Search for long-lived exotic particles at LHCb

Veerle Anna Margeretha Heijne
"Zoektocht naar langlevende exotische deeltjes in LHCb"

omslagillustratie: "Maui sand grains arrangement"
door dr. G. Greenberg
http://sandgrains.com

druk: Gildeprint Drukkerijen - Enschede

Dit werk maakt deel uit van het onderzoekprogramma van de Stichting voor Fundamenteel Onderzoek der Materie (FOM), die deel uitmaakt van de Nederlandse Organisatie voor Wetenschappelijk Onderzoek (NWO). Het werk is gefinancierd door de NWO VIDI-subsidie getiteld "A search for long-lived heavy particles" van dr. W.D. Hulsbergen.
Search for long-lived exotic particles at LHCb
promotor: prof. dr. M.H.M. Merk
copromotor: dr. W.D. Hulsbergen
## CONTENTS

2.5 Software framework ................................................. 36

3 Data and simulation .................................................. 37
  3.1 Simulated signal ..................................................... 37
     3.1.1 Hidden valley ............................................... 37
     3.1.2 Detector acceptance ......................................... 39
     3.1.3 Generator level studies ................................. 39
     3.1.4 Expected number of events .............................. 40
  3.2 Simulated background ............................................ 43
     3.2.1 Beauty decays ................................................ 44
     3.2.2 Charm decays ................................................. 45
  3.3 Data ............................................................. 45

4 Trigger and vertex reconstruction .................................. 47
  4.1 Trigger L0 and HLT1 selection ................................ 47
  4.2 Trigger HLT2 selection ........................................... 49
  4.3 Displaced vertices HLT2 trigger lines ....................... 53
     4.3.1 Displaced vertex algorithms for VELO vertices ........ 53
     4.3.2 Displaced vertex algorithms for downstream vertices . 55
     4.3.3 Displaced vertex HLT2 lines .......................... 56
  4.4 Topological HLT2 trigger lines ................................ 61
  4.5 Stripping selection .............................................. 64
     4.5.1 Stripping vertex algorithm ............................... 64
     4.5.2 Stripping lines ............................................. 65
  4.6 Material veto ..................................................... 67
  4.7 Beam-related background veto ................................ 69
  4.8 Characteristics of the reconstructed displaced vertices ..... 71
  4.9 Minimum number of vertex tracks ............................. 71
  4.10 Outlook ......................................................... 74

5 Jet reconstruction .................................................... 77
  5.1 Jet reconstruction ................................................. 77
     5.1.1 Jet input .................................................... 78
     5.1.2 Jet reconstruction algorithm ............................ 79
     5.1.3 Jet reconstruction on signal Monte Carlo ............... 80
  5.2 Jet cone size ...................................................... 81
  5.3 Jet identification ................................................ 84
  5.4 Jet reconstruction efficiency ................................... 84
  5.5 Jet energy correction ............................................. 85
  5.6 Outlook ........................................................ 89
# Appendix

## A Tuning of parameters for vertex reconstruction

### A.1 VELO tuning

### A.2 Downstream tuning

## B 2011 data sample

# Bibliography

# Popular summary

# Populaire samenvatting

# Acknowledgements
Following the recent discovery of the Higgs boson, the standard model of particle physics (SM) can be regarded as an established theory. However, it leaves some mysteries unsolved, such as the nature of dark matter and dark energy in the universe, and the abundance of matter over antimatter. The experiments at the Large Hadron Collider (LHC) are currently in the best position to either discover or exclude theories beyond the standard model provided they affect physics at the electroweak scale. The analysis described in this thesis contributes to this effort by searching for evidence of alternative models in which long-lived particles occur that are massive and have a measurable non-zero lifetime. These exotic particles cannot be detected directly, but can be identified by their decay into SM particles at a decay vertex that is displaced from the primary interaction point. They can leave different types of signatures in the detector, such as events containing one or more displaced vertices that decay into individual particles, jets, or both jets and leptons.

Analysis overview

Theoretical models that propose long-lived particles are introduced in Chapter 1. The analysis described in this thesis is focussed on the search for a single long-lived 'hidden valley' \( \pi_0 \) particle decaying into two quark-jets. The search is performed using data from the LHCb experiment at the LHC proton-proton collider in Geneva, described in Chapter 2. Although most theoretical models, especially the hidden valley models, feature pair-produced long-lived particles, the limited acceptance of the LHCb detector reduces the efficiency with which both particles can be detected. It is therefore worthwhile to search for events with a single long-lived particle candidate.

The event selection is not fine-tuned to a specific quark flavour, lifetime or mass, such that a wide range of models can be included in the search. The results in this thesis are obtained by assuming a hidden valley signal in which a standard model Higgs boson decays to two \( \pi_0 \) metastable particles, each decaying into two \( b \)-jets.

The lifetime acceptance for the long-lived particles is mainly determined by the
size of the vertex locator (VELO detector). VELO tracks or track segments are re-
quired in both the trigger and the jet reconstruction. This implies that the particle has
to decay within the VELO volume, limiting the sensitivity to particles with a lifetime
less than approximately 100 ps. A lower limit on the lifetime acceptance of about 1 ps
is enforced by the requirement that candidate vertices are displaced from the beamline
by at least 0.4 mm, which is used to eliminate primary vertices.

The detector acceptance for different $\pi^0$ masses is limited due to the requirement
to detect two jets per candidate. For low-mass particles (approximately below 20 GeV)
the two $b$ quarks are reconstructed within a single jet. For masses above 50 GeV,
the transverse momentum of the $\pi^0$ particle is low due to the fact that it is produced in the
resonant decay of a Higgs boson. In this case, the jets will often be (partly) outside the
acceptance of the LHCb detector. Chapter 3 describes the data and simulated signal
and background samples.

The vertex reconstruction and trigger selection of signal-like events are described
in Chapter 4. The trigger selection at both the hardware level (L0) and the first soft-
ware high level trigger (HLT1) relies on the configurations that are used to efficiently
select $b$- and $c$-hadron decays in LHCb, since they are well-suited to trigger on dis-
placed vertices. At the hardware level, the signal is triggered by either hadron, lepton
or photon energy deposits. The HLT1 triggers on a reconstructed track with high
transverse momentum. The events are selected at the second software stage (HLT2)
using a dedicated displaced vertex trigger, which reconstructs vertices with a high
track multiplicity, a high mass and a minimal displacement from the primary interac-
tion point. Supplementary two-, three- and four-body topological triggers (ordinarily
used to select $b$-hadron decays) are added to increase the sensitivity to vertices with
less tracks, lower mass and lower lifetime. Offline, a vertex reconstruction algorithm
is run on the triggered events, which recreates the vertices found in the trigger stage.

Subsequently, a jet reconstruction is performed in order to make an estimate of
the long-lived particle mass (Chapter 5). This dedicated jet reconstruction is run on
all particles that point back to the displaced vertex candidate. The jets and vertices are
matched to create dijet objects, whose properties are described in Chapter 6.

The main backgrounds for heavy displaced vertices, as described in Chapter 7, are
material interactions, primary collision vertices and decays of standard model $b$- or
$c$-hadrons with additional tracks added to the vertex. The material interactions are
eliminated by vetoing a geometrically defined region around the detector elements.
The primary vertices are removed by excluding the primary interaction region, and
the standard model backgrounds are reduced by requiring the presence of a vertex
with a high mass and a high track multiplicity. However, due to combinations of real
detached tracks (e.g. from charm or beauty decays) with tracks from elsewhere in
the event, charm and beauty decays form the main source of background at the final
selection level. One of the main challenges in the analysis is to produce sufficient
Monte Carlo events to simulate the amount of $b\bar{b}$ and $c\bar{c}$ created in the data used for this analysis, namely $10^{11}$ and $10^{12}$ events, respectively. These large amounts cannot all be fully simulated, due to limited computing power. The solution is not to rely on simulated background, but to fit the mass of the remaining data events with a smooth background shape, from which a potential signal mass peak can be distinguished. This relies on a sufficiently good signal mass resolution, which is obtained from the jets.

The measurement is diluted by the systematic uncertainties described in Chapter 8. The main uncertainties arise firstly from the selection efficiency, which is estimated from simulation, and secondly from the uncertainty on the dijet mass of the candidates. The efficiencies of the trigger, offline vertex reconstruction and selection procedure are determined from simulation and the uncertainties are verified using control samples in data. The uncertainty on the invariant mass shape depends mainly on the jet energy scale, and is retrieved from control channels for the jet reconstruction. The validation of the jet algorithm constitutes a significant part of the analysis, since both the input particles, the reconstruction algorithm and the selection criteria on the jet properties differ from the standard jet reconstruction method in LHCB.

To determine whether a long-lived particle been observed in the current data set, a fit is performed to the mass shape in data, in bins of radial displacement from the beamline (as described in Chapter 9). The radial binning improves the sensitivity of the search, since the remaining background events are mostly at low mass and low lifetime. The fit is performed for a fixed set of generated $\pi^0$ mass assumptions between 25 and 50 GeV, and for different $\pi^0$ lifetimes between 1 and 200 ps.

The research presented in this thesis gives a detailed overview of all the steps that are involved in the search for exotic long-lived massive particles at the LHCb detector. It gives a good impression of both the possibilities and the limitations of new physics searches at high energy particle accelerators. The last chapter includes a critical review of the analysis and an outlook on future prospects.
Throughout the twentieth century, the development of new detectors and accelerators enabled the observation of the existence of a wide range of subatomic particles, some of which were predicted by theory. With the discovery of the Higgs particle, these observations completed the evidence for the standard model of particle physics (SM), the theoretical description of which was finished in 1974 [1]. This thesis investigates the existence of particles beyond the SM, focussing on theoretical models that predict long-lived massive particles.

1.1 The standard model

The standard model is a quantum field theory that describes the known particles and three of the four fundamental forces as gauge fields, based on the $SU(3)_C \times SU_L(2) \times U(1)_Y$ symmetry group. $SU(3)_C$ describes strong interactions, where $C$ represents the `colour charge' quantum number of quarks. Electroweak interactions are described by the combined $SU_L(2) \times U(1)_Y$ groups, where $Y$ denotes the hypercharge, and $L$ the weak isospin coupling to left-handed fermions only. The SM implements an invariance under local transformations in this symmetry group to describe interactions of quarks and leptons (both fermions with spin 1/2) via force carriers (bosons with integer spin). An illustration of the SM particle content is given in Fig. 1.1.

The quarks and leptons are divided into three generations. The first generation of quarks consists of the up- and down-quark, the second of the charm- and strange-quark, and the third generation of the top- and beauty-quark. The leptons and their associated lepton-neutrinos are divided into the electron, muon and tau. In addition,
each particle has an associated anti-particle with opposite internal quantum numbers. The gluon (g) is a quantum of the strong force, which couples to quarks through their colour charge. As a consequence of its self-coupling, the strong force has a limited range. However, due to its strong coupling constant, it is the dominant force inside nuclei, and it is responsible for keeping atomic nuclei together by binding the quarks in the protons and neutrons.

The photon (γ) corresponds to the electromagnetic force, which couples to the electrically charged particles. The electromagnetic force has a long range and binds the negatively charged electrons in atoms to the positively charged nucleus. The strong and electromagnetic interactions only couple to particles and antiparticles of one type, so both the photons and the gluons conserve flavour.

The weak field quanta couple to both quarks and leptons. Flavour is not conserved in charged current weak interactions that are mediated by the $W^\pm$ bosons, such that the quarks can change generation by the interchange of a $W$, as described by the Cabibbo-Kobayashi-Maskawa (CKM) matrix. Neutral weak interactions, mediated by the Z boson, are flavour conserving.

The electromagnetic and the weak force are closely connected through the $SU_L(2) \times U(1)_Y$ symmetry upon which the SM is based. The electroweak force carrying bosons ($W^+, W^-, Z^0$ and $\gamma$) should be massless as a consequence of the underlying gauge symmetry. However, the physical $W^\pm$ and $Z^0$ bosons do have mass. This contradiction is solved through the addition of a new scalar ‘Higgs’ field with a non-zero vacuum expectation value: so-called spontaneous symmetry breaking. Any particle
that interacts with this Higgs field acquires mass. This mechanism gives rise to the ex-
istence of a massive Higgs boson [3, 4], which was discovered in 2012 by the ATLAS
and CMS experiments [5, 6]. Even though it has proven to be a successful description
of particle physics, the SM still leaves several fundamental questions unanswered.

1.1.1 Limitations of the standard model

Unification of forces: The electromagnetic and the weak force both originate from
the $SU_L(2) \times U(1)_Y$ combined gauge symmetry, which at low energy appears
as two different forces. The strong force is not included in this symmetry. The
existence of a larger symmetry group embedding the SM is proposed in grand
unifying theories (GUT). However, the unification of the observed running of
the coupling constants requires the presence of supersymmetry.

Gravity: The most obvious limitation of the SM is that the fourth fundamental force,
gravity, is not included, and no gravitation quantum force carrier has been dis-
covered. The gravitational attraction is so weak at the subatomic particle level
that it does not influence any of the observations. String theory is one of the
attempts to incorporate gravity in an overall quantum theory [7].

Hierarchy problem: On the one hand, the requirements on the maximum allowed
value of the Higgs boson mass from precision measurements on weak inter-
actions are of the order of 100 GeV, which is in agreement with the observed
Higgs mass of $\sim 126$ GeV. On the other hand, if the SM were to be valid up to
the Planck scale of $10^{19}$ GeV, loop corrections in the self-coupling of the Higgs
would drive the mass towards very high energies [8, 9, 10]. The energy scales
in particle physics therefore seem to conflict, and an uncomfortable degree of
fine-tuning is needed to cancel the loop corrections to achieve the observed
Higgs mass using only the SM particles. It is therefore expected that physics
beyond the SM must appear at an energy scale of about 1 TeV, not too far above
the electroweak scale. Supersymmetric theories (SUSY) can solve the hierarchy
problem by introducing a supersymmetric partner for every SM particle, which
exactly cancels the loop contribution of the associated SM particle [11].

Dark matter and dark energy: Visible baryonic matter by itself cannot account for
observed orbital rotations of galaxies and velocities of stars throughout galax-
ies [12, 13, 14]. This supports the existence of ‘dark matter’, which only interacts
via gravity, and perhaps the weak force. Furthermore, the observed accelerated
expansion of the universe indicates that some ‘dark energy’ must act against
the gravitational attraction between the constituents of the universe. Dark en-
ergy and dark matter are estimated to make up respectively 68% and 27% of
the energy in the universe, whereas ordinary matter only accounts for 5% [15].
New particle physics models that attempt to solve the dark matter problem must feature at least one neutral stable particle, which serves as a dark matter candidate. Dark matter candidate searches are carried out through direct detection in large detector volumes such as the XENON experiment [16], through detection of high energy photons or positrons from dark matter annihilation in space, or through indirect detection of missing energy after producing the dark matter candidate in accelerators such as the LHC.

**Massive neutrinos:** For a long time, the neutrinos were assumed to be massless. The observation of neutrino flavour oscillations implies that neutrinos have a small mass [17], which is orders of magnitude smaller than the mass of the other fermions. Such small masses can be accommodated in the SM, but would be unnatural. Alternatively, the unusually small neutrino mass can be described by the addition of a Majorana mass term for neutrinos, resulting in both very small left-handed neutrino masses and heavy right-handed Majorana neutrinos, via the 'see-saw' mechanism, which will be discussed in more detail in Section 1.2.2 [18].

**Matter-antimatter asymmetry:** The electroweak sector of the SM allows charge-parity (CP) violation only at a very small level. This is not enough to account for the abundance of matter over antimatter in the universe [19]. Some process must have prevented the annihilation of all matter and antimatter into photons shortly after the Big Bang. There are models that introduce a large enough amount of CP-violation to explain the matter-antimatter asymmetry, such as sterile right-handed neutrinos that couple to the charged leptons [20].

### 1.2 Models featuring long-lived particles

The shortcomings of the standard model lead to searches for evidence of physics beyond the standard model (BSM) at particle accelerators. These searches for unknown processes are usually guided by theory. Many theoretical models are developed that can solve some of the questions that the SM leaves unanswered.

Most new physics searches use generalisations or simplifications of models, in order to be able to focus on a limited number of possible signatures. An example is the minimal supersymmetric model (MSSM). However, there is no reason why new physics would present itself in its most simplified form. For example, both the standard model and the BSM models could have additional gauge sectors, and extended Higgs sectors [21]. The fact that no evidence for physics beyond the SM has been found so far at the LHC supports these more complicated scenarios. They could reduce the experimental signatures of the most popular supersymmetrical signals such as missing energy, and increase the occurrence of other signatures such as displaced vertices,
which are created when new neutral massive long-lived particles decay into SM particles.

The next paragraphs introduce BSM models that motivate the existence of long-lived particles, and the displaced vertex signatures they produce in a hadron collider. The main focus will be on the hidden valley scenario, which is used as the signal model in the analysis described in this thesis.

1.2.1 SUSY with R-parity violation

Weak scale supersymmetry (SUSY) is the most popular model to solve the hierarchy problem of the standard model. It implies that for each SM boson, a supersymmetric fermion exists, and for each SM fermion a supersymmetric boson. Therefore the SUSY superpartners differ by spin $1/2$ from the SM particles. SUSY has to be a broken symmetry, for otherwise the SUSY particles would be identical to the SM particles except for their spin, and they would have been observed already. A detailed review of SUSY scenarios is given in e.g. Ref. [22, 11].

In SUSY, baryon number and lepton number are not conserved as they are in the SM at low energy, but a mechanism is needed to prevent proton decay. To this end, so-called 'R-parity' is introduced, defined as: $P_R = (-1)^{3(B-L)+2s}$, where $B$ and $L$ are the baryon and lepton number, respectively, and $s$ is the spin of the particle. By this definition, SM and SUSY particles receive opposite R-parity. Most SUSY models assume R-parity conservation. The multiplicative conservation of R-parity implies that SUSY interactions always require an even difference in the number of SUSY particles between the initial and the final state. A heavy SUSY particle will always decay into an odd number of lighter SUSY particles and an arbitrary number of SM particles, but never into only SM particles. Consequently, the lightest superpartner (LSP) cannot decay, and could therefore be a dark matter candidate.

However, R-parity violating (RPV) operators that allow the LSP to decay can be included in the description of SUSY within the current experimental bounds. R-parity violation can be achieved either through baryon number violation (BNV) or through lepton number violation (LNV) [23, 24, 25]. Both cannot be violated at the same time, because that would allow the proton to decay. In case of RPV, R-parity odd terms that are allowed by renormalisability and gauge invariance are included in the superpotential of the supersymmetric standard model, leading to the following expression:

$$W^{RPV} = \lambda_{ijk} L_i L_j E_k + \lambda'_{ijk} L_i Q_j D_k + \lambda''_{ijk} \bar{U}_i \bar{D}_j D_k$$

(1.1)

where $L$ and $\bar{E}$ are the lepton doublet and the antilepton singlet superfields, $Q$ and $\bar{U}, \bar{D}$ are the quark doublet and antiquark singlets superfields, and $\lambda_{ijk}$ are the couplings for the different terms. The $i, j, k$ are flavour indices. The first two terms violate l ep-
ton number, and the last term, which involves only quarks, violates baryon number. The Yukawa-like fermion–fermion-scalar interactions associated with the couplings of $W^\text{RPV}$ involve both SM and supersymmetrical particles. For example, the first term can be expressed as the following Lagrangian term, according to the derivation in Ref. [25]:

$$L_{L_i L_j E_k}^\text{RPV} \supset \lambda_{ijk} (\bar{\nu}_{iL} l_{jL} \tilde{L}_{kR} + \nu_{iL} l_{jL} \tilde{L}_{kR} + \bar{\nu}_{iR} l_{jL} \tilde{L}_{kR} - (i \leftrightarrow j))$$  \hspace{1cm} (1.2)$$

where the tildes represent supersymmetrical particles. As an example, the left diagram in Fig. 1.2 involving $\nu_{\mu}$, $\tau$ and $\bar{e}$ corresponds to the lepton number violating term $\lambda_{231} L_2 L_3 \tilde{E}_1$. The diagram on the right is a hadronic decay, involving baryon number violation, corresponding to the term $\lambda^\prime_{112} U_1 \bar{D}_1 \bar{D}_2$.

Supersymmetric theories with RPV through baryon number violation could contribute to baryon non-conservation that explains the matter-antimatter asymmetry. Lepton number violation can naturally generate the observed neutrino masses and mixing [25]. The couplings $\lambda_{ijk}$, $\lambda'_{ijk}$ and $\lambda''_{ijk}$ are constrained by measurements of processes such as lepton universality, top quark forward-backward asymmetry and neutrinoless double beta decay [24].

An additional motivation to search for RPV signals is that the current SUSY searches, relying mostly on `missing energy' due to the stable LSP escaping the detector, have not revealed any hints of signals yet. RPV has the consequence that the lightest superpartner is no longer stable. This would eliminate missing energy signatures in events with supersymmetric particles. The lightest superpartner can decay into standard model particles through an R-parity violating decay, which, depending on the amount of RPV, acquires a certain lifetime, and leaves a displaced vertex signature. The lifetime is inversely related to the strength of the RPV operator. An overview of possible hadron collider signatures in RPV SUSY models is given in Ref. [24], and
an update of the constraints using the latest LHC results for dijet resonances is summarised in Ref. [26].

SUSY with baryon number violation

The overview in Ref [27] proposes that the current experimental bound from 2012 LHC data leave room for SUSY signatures with displaced vertices. The rate of typical baryon decay processes scales as $\Gamma \sim |\lambda''|^2 m_{\tilde{q}}$, where $\lambda''$ is the BNV coupling, and $m_{\tilde{q}}$ is the typical mass of the LSP, in this case the squark (the SUSY partner of the quark) [25]. The long-lived particle decay length in the detector is inversely related to the strength of the BNV operator squared. As was mentioned before, BNV could have generated the matter-antimatter asymmetry in the universe at or after the electroweak phase transition. However, BNV interactions can also erase a baryon asymmetry that would be present before the electroweak phase transition in the early universe. The preservation of baryon number therefore poses an upper bound on the operator $\lambda''$, which translates into a lower bound on the decay length. The calculation in Ref. [27] gives a requirement on the rate of $|\lambda''|^2 m_{\tilde{q}} \lesssim O(10^9 \text{s}^{-1})$. At the LHC, this results in a minimal long-lived particle decay length of approximately 60 $\mu$m [27]. An upper bound of several meters is supported by the current experimental results that constrain missing energy signatures.

The minimal supersymmetric standard model (MSSM) is the simplest extension of the SM that includes SUSY. In case R-parity violation is caused by BNV, a Higgs boson could hadronically decay via two so-called ‘neutralinos’ ($\tilde{\chi}_1^0$) to a fermion and a sfermion (the SUSY partner of the fermion), which in turn decays to two fermions, as illustrated in Fig. 1.2b. Both decay processes from SUSY to SM particles violate R-parity, and lead to metastable lifetimes. If the neutralino lives sufficiently long, the Higgs decays to six displaced jets [28].

One of the MSSM models is minimal supergravity (mSUGRA), which features gravity mediated supersymmetry breaking. An R-parity violating mSUGRA model is proposed in Ref. [29]. It assumes baryon number violation, leading to purely hadronic neutralino decays. In this model the lightest neutralino has a mass in the range 20-60 GeV and it decays into three quarks through BNV. Such decays give rise to three soft jets with a total invariant mass equal to that of the original sparticle [30]. The final state is similar to that of the MSSM model mentioned above, and suitable signatures for LHCb with three jets (such as the one illustrated in Fig.1.2b) are studied in Ref. [31]. The neutralino decay length depends on its mass $m_{\tilde{\chi}_1^0}$, the mass of the squarks $m_{\tilde{q}}$ and the RPV couplings $\lambda''$. Neutralino lifetimes in the range from 3 to 25 ps, as proposed in Ref. [31], are compatible with the limits on the baryon number violating couplings $\lambda''$ from for example rare hadronic $B$-decays, in which any transition other than the $b \to c$ is suppressed in the SM and might be enhanced by RPV. The production of neutralinos mainly happens in pairs through the decay of a Higgs boson $h^0$. If the
parameter defining the Higgs couplings, $\tan(\beta)$, is small (typically $< 3$), the SUSY $h^0$ is essentially equivalent to the standard model Higgs, with an expected production cross-section of about 20 pb at 7 TeV proton-proton collisions [32].

**SUSY with lepton number violation**

In case R-parity violation is caused by lepton number violation, the allowed amount of LNV is constrained by the neutrino masses, and by limits on rare processes such as $\mu \rightarrow 3e$. An interesting scenario features a Higgs boson that, through two LSPs, ultimately decays to four jets and two leptons, creating two displaced vertex signatures [33].

mSUGRA scenarios can also incorporate RPV through LNV. In that case, the lightest neutralino decays either fully leptonically or semileptonically [34, 35]. The study in Ref. [36] proposes signatures with displaced vertices of dimuons and a neutrino, that would be detectable in LHCb. A preliminary study has been performed to search for such a model in LHCb, using a signature with a displaced vertex including one muon and two jets [37].

**GMSB**

In gauge mediated supersymmetry breaking models (GMSB), SUSY is broken in a hidden sector, and so-called messenger fields interact with the SM through gauge mediated interactions [38, 39]. The LSP is the gravitino, and the next-to lightest supersymmetric particle (NLSP) is long-lived because the coupling between the NLSP and the gravitino is small. The most promising signature is that of a NLSP decaying into a gravitino and an off-shell $Z^0$, with the $Z^0$ decaying into two leptons or two quarks. The potential of LHCb to search for signatures with a jet pair or an opposite-sign lepton pair is discussed in Ref. [40].

**Other SUSY models**

Various other extensions of SUSY can result in a wide scope of displaced vertex signatures. Some examples are: lepton number violating signals with multileptons and multijets, with multiple displaced vertices due to the presence of long-lived right-handed neutrinos [41], gluinos decaying to same-sign leptons and additional particles [42], and neutralinos decaying to any number of charged leptons [43].

**1.2.2 Right-handed neutrinos**

One of the striking observations in particle physics is the large difference between the neutrino masses and the SM fermion masses. This could be explained by adding Majorana mass terms that introduce new right-handed neutrinos at a high energy scale.
Such a Majorana neutrino, which is its own antiparticle, can be much heavier than the SM particles, since it is not related to the Higgs mass scale. However, as the sterile right-handed neutrino gets heavier, the SM left-handed neutrino gets lighter, hence this model is called the 'see-saw' mechanism [18]. The mass of the right-handed neutrino can be low enough for the particle to be produced at the LHC through couplings to SM particles at the TeV scale. If its mass is between 1-100 GeV, its lifetime is such that it can decay within the acceptance of LHC detectors. Particles with a smaller mass would decay outside the reach of the detector, and particles with a higher mass would decay too quickly to distinguish them from promptly decaying particles. An overview of heavy Majorana neutrino searches and the LHC potential for these searches can be found in Ref. [44].

The existence of Majorana neutrinos can be verified via lepton-number violating processes, but also via direct detection of their decay products. In LHCb, the LNV process $B^- \rightarrow \pi^+ \mu^- \mu^-$ is used to search for a light Majorana neutrino ($< 5$ GeV) [45], and a displaced vertex search could reveal a heavy neutrino candidate. There are various models that feature Majorana neutrinos, and depending on the exact implementation of the model, they have different displaced vertex signatures. A pair of displaced vertices could be the indication of a Higgs boson that decays into two right-handed neutrinos, each decaying into a quark pair and a lepton, thereby violating lepton number conservation [46]. A right-handed long-lived sneutrino in a next-to-minimal supersymmetric standard model (NMSSM) would predominantly decay into a lepton and two quarks or two leptons and neutrinos [47]. A more general proposal to search for Majorana neutrinos through displaced vertices with leptons and jets at the LHC is given in Ref. [48].

1.2.3 Hidden Valley

'Hidden valley' (HV) models feature a hidden sector at a low mass scale that is weakly coupled to the SM gauge sector through heavy mediators [49]. Hidden valley scenarios are often motivated by string theory [50]. The potential of LHCb in searches for hidden valley particles is pointed out in Ref. [51].

The term 'hidden valley' refers to the fact that the particles in the new sector have weak-scale masses, but they are hardly accessible because they are hidden by a high energy barrier, as illustrated in Fig. 1.3. The hidden sector is formed by a new non-abelian gauge group, which extends the SM gauge group. The SM particles are neutral under the hidden gauge group. The hidden sector contains 'v-particles', which are neutral under the SM but charged under the hidden gauge 'v'. An abelian group would result in radiation of v-photons that decay to SM fermion pairs, whereas a non-abelian group results in hadronisation into v-mesons that decay to SM particles such as quarks [52]. Here, only the non-abelian group is considered.

The interaction between the hidden particles and the SM group is mediated by
Figure 1.3: Schematic view of the production and decay of v-hadrons. While LEP was unable to penetrate the barrier separating the sectors, LHC may be able to produce v-particles. These form v-hadrons, some of which can decay to SM particles. From [51].

Figure 1.4: A hidden valley event in the two-light-flavours regime. The $\pi_\pm$ are v-charged, but neutral under the SM electrical charge. The $\pi_0$ are the long-lived particles that decay to SM quark pairs. Reproduced from [49].

higher dimension operators, for example loops of heavy particles (TeV scale) or a $Z'$. The high mass scale of the mediator would explain why such v-particles have never been observed at previous collider experiments, but could be produced at the LHC, through collisions at a higher center-of-mass energy. This model encompasses a large collection of v-particles, comparable to the SM particles, all with different decay widths, masses and decay modes.

The simplest hidden valley model is a QCD-like $Z'$ model [49], consisting of a $U(1) \times SU(3)$ group that is added to the SM. The scale of $SU(3)$ is $1 \text{ GeV} < \Lambda_\nu < 1$ TeV. The mediator, a vector particle $Z'$, has a mass of the order 1-6 TeV due to coupling to a scalar field $\phi$, with a non-zero expectation value, breaking the $U(1)$ symmetry. To this simple model two v-quark flavours ($U$ and $D$) can be added. The quarks acquire
mass by coupling to the scalar field $\phi$. In the case where there are two light flavours, and $m_U \sim m_D \ll \Lambda_\nu$, all v-hadrons decay into v-pions ($\pi_v^0, \pi_v^\pm$) and v-nucleons, as shown in Fig. 1.4. Particles with a v-charge cannot decay to SM particles, such that the lightest stable v-particle is a dark matter candidate. Of the v-particles, only the $\pi_v^0$, which is v-charge neutral, is unstable. The $\pi_v^0$ can decay via a $Z'$ into standard model fermion pairs. If the $\pi_v^0$ mass is in the range $2m_b < m_{\pi_v} < 2m_t$, it decays mainly to heavy flavour $b\bar{b}$, due to helicity conservation of a scalar $\pi_v^0$ into two spin $\frac{1}{2}$ quarks. The decay rate of the $\pi_v^0$ depends on $m_{\pi_v}$, $m_{Z'}$ and the coupling $g'$, which can all vary within a wide range, leaving the lifetime of the $\pi_v^0$ almost unconstrained.

The production of hidden valley particles in the two v-quark flavour regime can happen in various ways. First, just like the $\pi_v^0$ can decay via a $Z'$ into standard model fermion quark pairs, the process can also be reversed, such that the fusion of two quarks in a proton-proton collision produces a $Z'$ [49]. This $Z'$ can then decay into $U\bar{U}$, as shown in Fig. 1.4. It is expected that in the decay of $U\bar{U}$ to v-hadrons, there is a large spread of v-pion multiplicities, and events can have a more spherical or a more collimated structure depending on the center of mass energy of the collision. At the LHC, the dominant decay for the two-light-flavours model ($\pi_v^0 \rightarrow b\bar{b}$) would result in an unknown number of b-jet pairs and possibly missing energy. For an increasing $\Lambda_\nu$, the number of produced v-hadrons will decrease, but the jets resulting from their decay will be harder and easier to distinguish from background.

![Cross Section in fb](image)

**Figure 1.5:** v-quark production cross-section at the LHC in fb versus $m_{Z'}$, for different $m_{Z'}/g'$. Here SU(3) with two light v-quark flavours is assumed, and $\Lambda_\nu, m_U, m_D \ll m_{Z'}$. From [49].

Figure 1.5 shows the dependence of the v-quark production cross-section on the $Z'$ mass and on the coupling $m_{Z'}/g'$. This prediction is valid for the specific QCD-like model introduced above. There are some constraints on these cross-sections from previous experiments. For example, results from $Z$ decays at LEPI give a limit on $m_{Z'}/g' > 10$ TeV (for $m_{Z'} = 2.5$ GeV), such that the cross-section at the LHC is at most 20 fb [49]. More complicated models than the two light v-quark flavour model result in multi-object final states, which result in weaker bounds.
Figure 1.6: A hidden valley event with a resonant Higgs decay to $v$-hadrons, each of which decays to SM $b\bar{b}$ through a heavy $Z'$. Reproduced from [53].

A second way to produce $v$-particles is through the resonant decay of a Higgs boson [53, 49]. If the Higgs field mixes with the scalar field $\phi$ of the hidden sector, it can decay to pairs of $v$-particles through $gg \rightarrow h \times \phi \rightarrow Q\bar{Q}$, where $Q$ can be a $\pi_v^0$, if kinematically allowed. The $v$-particles decay through the heavy $Z'$, and are therefore long-lived. An illustration of such a process is given in Fig. 1.6. These decays are not affected by the LEP constraints. Even though the particle discovered in 2012 at the LHC is consistent with a SM Higgs boson, it could still have non-SM properties, such as a coupling to exotic particles. Given the results from 2012 LHC data, a SM Higgs with SM couplings can still accommodate a branching ratio up to 40-60% to new particles [54, 55].

A third production process of $v$-particles is through the decay of the lightest standard model superpartner (LSsP) in supersymmetric extensions of the hidden valley scenario [51]. In this case, both the SM and the hidden valley have additional superpartners. This model results in more complicated final states than the ones described before, including soft jets, leptons and displaced vertices. It also motivates the search for events with a single displaced vertex with a large track multiplicity, or events with cascades of displaced vertices.

The Monte Carlo simulation for the analysis described in this thesis implements the resonant Higgs decay to two $\pi_v^0$ particles. The current analysis searches for individual long-lived $\pi_v^0$ particles. Alternatively, a signature of two displaced vertices per event was used for a search for a Higgs boson in the 2010 LHCb data set [56].

1.2.4 Dark matter

Some dark matter candidates are already incorporated in previously discussed models, like the LSP in some SUSY models, and the lightest stable $v$-particle in hidden valley models. There are various other models with a dark matter candidate that feature displaced vertices in collider events. These will not be discussed in detail, but some examples of interesting signatures are: lepton jets that contain two to eight leptons in dark sector cascade decays [57], lepton pairs and missing energy in a GUT model [58], NLSP decay into pairs of dark matter particles and additional standard model particles.
in a model with dark matter particle-antiparticle asymmetry [59], and a vertex with leptons, high-\(p_T\) jets and missing energy in a pseudo-Dirac dark matter model [60].

### 1.3 Experimental constraints

Existing results from searches for displaced vertices are listed below. Most searches are inspired by the hidden valley model, as is the analysis described in this thesis.

**LEP** results have been studied in the context of hidden valley models with light \(v\)-hadrons. The results for \(Z\) decays at LEPI limit the branching fraction via \(Z \rightarrow Z'\) mixing into \(v\)-particles: \(Z \rightarrow QQ\). The \(Q\) are assumed to be either \(U\) or \(D\). Assuming that \(m_{Z'}/g' = 10\) TeV, and that the \(v\)-hadronic decays of the \(Z\) are easily distinguished from background, this branching fraction is less than \(10^{-7}\). Given this constraint, the cross-section for this process at LHC for \(m_{Z'} = 3.5\) TeV and \(g' = 0.25\) is about 20 fb. LEPII results give an additional cross section limit at 200 GeV on the process \(e^+e^- \rightarrow QQ\) of \(\approx 1\) fb, assuming that \(m_{Z'}/g' = 6\) TeV [49];

**D0** has published limits on a search for pair-produced \(\pi_\nu^0\) particles from a SM Higgs boson decaying to two \(b\) jets, in data with a center-of-mass energy of 1.96 TeV collected at the Tevatron [61]. The considered \(\pi_\nu^0\) mass is 15 or 40 GeV, with a proper decay length between 2.5 and 10 cm corresponding to a lifetime between 75 and 300 ps, and the Higgs boson mass is varied from 90 to 200 GeV. A high-\(p_T\) muon has to be present in order to trigger the event. Figure 1.7 shows the upper limits for one of the mass-points. Considering only 120 GeV Higgs mass points, the limits vary between 1.0 pb for a 15 GeV \(\pi_\nu^0\) with 2.5 cm flight distance, and 16 pb for a 40 GeV \(\pi_\nu^0\) with 5 cm flight distance. Figure 1.7 shows the limits for the region that is most compatible with the LHCb search region. Note that D0 has no measurements at a lifetime below 2.5 cm, and that the best limits are for low \(\pi_\nu^0\) mass;

**CDF** at the Tevatron has presented an analysis looking for two \(\pi_\nu^0\) particles from a SM Higgs boson decaying into two jets of any flavour [62], in 1.96 TeV data. The selection is mainly based on the jet kinematics, and the events are triggered using tracks that originate from displaced vertices. The Higgs masses considered are 130 and 170 GeV, the \(\pi_\nu^0\) mass lies between 20 and 65 GeV, and its lifetime is reweighted to values between 0.3 and 5 cm corresponding to a lifetime between 10 and 150 ps. Considering only the 130 GeV Higgs mass points, the limit is best for low \(\pi_\nu^0\) mass (6.2 pb for 20 GeV \(\pi_\nu^0\) mass and 1.0 cm lifetime) and low lifetime (17.8 pb for 40 GeV \(\pi_\nu^0\) mass and 0.3 cm lifetime). Figure 1.8 shows the limits for the region that is most compatible with the LHCb search region;
Figure 1.7: D0 upper limit on $\sigma(H) \times BR(H \rightarrow \pi_0^0 \pi_0^0) \times BR^2(\pi_0^0 \rightarrow b\bar{b})$ as a function of decay length for a 120 GeV Higgs mass, and a 15 GeV $\pi_0^0$ mass. From [61].

Figure 1.8: CDF upper limit on $\sigma(H) \times BR(H \rightarrow \pi_0^0 \pi_0^0) \times BR^2(\pi_0^0 \rightarrow b\bar{b})$ as a function of decay length for a 130 GeV Higgs mass, and a 40 GeV $\pi_0^0$ mass. From [62].
Figure 1.9: ATLAS upper limit on the process \( H \rightarrow \pi_0^0 \pi_0^0 \), expressed as a multiple of the SM cross section for Higgs production, as a function of \( \pi_0^0 \) proper decay length. The exclusion limits assume a 100% branching ratio for the Higgs decaying to \( \pi_0^0 \) particles. From [63].

Figure 1.10: CMS upper limit on \( \sigma(H) \times BR(H \rightarrow \pi_0^0 \pi_0^0) \times 2BR^2(\pi_0^0 \rightarrow q\bar{q}) \) as a function of decay length for the lightest Higgs (200 GeV) and \( \pi_0^0 \) (50 GeV) masses available. From [64].
ATLAS first presents a signature of a neutralino decaying to a muon and a multi-track vertex in 33 pb$^{-1}$ of 7 TeV data [65]. An update of this analysis with 4.4 fb$^{-1}$ of 7 TeV data is also presented, reporting an upper limit on between 100 and 2 fb, for lifetimes between $0.1 < c\tau < 100$ cm and neutralino masses between 108 and 494 GeV [66].

Second, ATLAS searches for a light Higgs (120 - 140 GeV) decaying to two long-lived particles resulting in an arbitrary final state with large multiplicity in the muon spectrometer system, in 1.94 fb$^{-1}$ of 7 TeV data [63]. For the latter, $\pi^0_\nu$ masses of 20 and 40 GeV are considered. Assuming a standard model cross-section for the Higgs, and a 100% branching ratio to dijets, ATLAS excludes lifetimes corresponding to a range between 0.5 and 20 meter (1.5 - 60 ns). This result is illustrated in Fig. 1.9;

A recent ATLAS conference report uses energy deposits in the hadronic calorimeter to reconstruct $\pi^0_\nu$ decay products in a single jet. This analysis extends the lifetime range down to about 10 cm, and uses a dataset of 20.3 fb$^{-1}$ at a center-of-mass energy of 8 TeV [67].

CMS also considers both leptonic and hadronic decay channels. First, a search is presented for heavy resonances decaying to two long-lived particles, each decaying to two leptons, on 4.1 (5.1) fb$^{-1}$ of 7 TeV data for the electron (muon) channel. Limits are set in the range 0.7-10 fb, for long-lived particles with lifetimes in the range $0.1 < c\tau < 200$ cm, masses between 20 - 350 GeV and a Higgs boson in the range 200 - 1000 GeV [68].

Second, CMS has presented a preliminary result on searches for long-lived dijets in 18.5 fb$^{-1}$ of 8 TeV data [64]. A scalar particle with a mass between 200 and 1000 GeV is considered, decaying into two long-lived neutral particles with mass between 50 and 350 GeV, each decaying into two jets. For lifetimes between 0.1 and 200 cm (3 ps - 6 ns) CMS obtains limits in the range 0.5 to 200 fb. The result for the lightest Higgs (200 GeV) and $\pi^0_\nu$ (50 GeV) masses that were investigated is illustrated in Fig. 1.10. Note that the masses are larger, while the decay lengths are smaller than those of the ATLAS result;

LHCb presented a preliminary analysis of 35.8 pb$^{-1}$ of 2010 data at 7 TeV, searching for a Higgs-like boson decaying into two neutralinos [56], following an R-parity violating SUSY model [28]. For a Higgs mass between 100–125 GeV and long-lived particles with masses between 30–55 GeV and a lifetime of 10 ps, production cross-sections upper limits are set between 29 and 179 pb;

Other constraints include measurements of the thermal relic energy density in the universe, which constrain the properties of a hidden valley dark matter candidate. Using these WMAP measurements [69], upper bounds can be set on
the mass of the $Z'$ mediator of the order of 10 TeV [70]. This upper bound ensures that the contribution from a hidden valley dark matter candidate fits the observed non-baryonic matter energy density. Dark matter searches for a relatively light ($\sim$GeV) dark sector gauge boson in beam-dump experiments and solar observations are summarised in Ref. [71].

### 1.4 Conclusion

With such a variety of long-lived particle models and signatures, it is needless to say that it is impossible to cover all of those in one analysis. However, an attempt is made to apply a selection that is sensitive to multiple scenarios. The analysis discussed in this thesis focusses on the decay of a long-lived particle to two jets, inspired by the hidden valley scenario. Contrary to ATLAS and CMS, LHCb will search in a relatively low mass and low lifetime regime. As an illustration, Fig. 1.11 schematically shows the regions of interest of the current experiments as a function of $\pi^{0}_v$ mass and lifetime. The LHCb result should mainly be compared to the CDF and D0 upper limits, that cover a similar region.

![Figure 1.11](image.png)

**Figure 1.11:** Schematic representation of the region of interest of different experiments at the Tevatron and at the LHC for hadronic signatures with a displaced vertex. The CMS and ATLAS regions extend to higher masses and lifetimes than displayed, respectively. All experiments use a different Higgs mass, which is indicated in the figure. Additional assumptions on the signatures are described in the main text.
Other ongoing or planned analyses in LHCb will specifically select displaced leptons, or a combination of leptons and jets, to cover an even wider range of proposed theoretical signatures. For some of these signatures, LHCb has an advantage over the general purpose detectors at the LHC. Non-isolated leptons are interesting for LHCb because of its displaced track trigger, whereas ATLAS and CMS depend on isolated leptons or high $p_T$ jets ($> 60$ GeV) to trigger the events. Due to its precise vertex reconstruction, LHCb also has an advantage over the general purpose detectors when it comes to distinguishing between multiple displaced vertices with tracks pointing in a common direction, in cascade decay topologies [72].
The world’s largest high-energy physics facility is situated near Geneva, at the European Organisation for Nuclear Research (CERN). It consists of a circular particle accelerator, the Large Hadron Collider (LHC), and four large experiments. ATLAS and CMS are both general-purpose detectors studying proton-proton interactions, and ALICE is designed to study lead-lead and lead-proton interactions. The detector that is used to collect the data for the measurements described in this thesis is the Large Hadron Collider beauty (LHCb) detector [73].

2.1 The Large Hadron Collider

A schematic overview of the CERN accelerator complex is shown in Fig. 2.1. The LHC is stationed in a 27 km long tunnel about 100 meter underground. Protons are accelerated in the Super Proton Synchrotron to 450 GeV, after which they are injected in opposite directions into the LHC for acceleration to the nominal multi-TeV energy. The two proton beams collide at four interaction points where the main detectors are located. During the running periods in 2011 and 2012, the proton bunches had a 50 ns bunch spacing, and the peak luminosity delivered by the LHC was $7.7 \times 10^{33} \text{cm}^{-2}\text{s}^{-1}$. In 2011 the accelerated protons collided at a centre-of-mass energy of 7 TeV, which was increased to 8 TeV in 2011, and which will eventually be raised to 14 TeV. Although the LHC delivers proton interactions for all four experiments, specific conditions apply to the LHCb interaction region.

The LHCb experiment received a luminosity of $2 \times 10^{32} \text{cm}^{-2}\text{s}^{-1}$ in 2011, which is lower than the luminosity delivered at the general purpose detectors ATLAS and
Figure 2.1: Schematic overview of the accelerator complex at CERN. Protons are accelerated consecutively in the linear accelerator LINAC 2, the BOOSTER, the Proton Synchrotron and the Super Proton Synchrotron, after which they are injected in opposite directions into the Large Hadron Collider. The four LHC experiments are indicated at the collision points.
CMS. This is achieved by defocussing the beams and by adjusting the overlap of the beams using magnets placed before and after the experiment. The luminosity is levelled throughout each proton fill period (lasting several hours), correcting for intensity loss of the beams, in order to keep the number of interactions per bunch-crossing constant [74]. Furthermore, the two beams are tilted by a 20° angle to ensure that the crossing angle is the same for different polarities of the large dipole magnet in the LHCb experiment.

A moderate luminosity is beneficial for the LHCb physics programme for several reasons. Firstly, a lower occupancy in the detectors shortens the reconstruction times and improves the reconstruction performance in both the online and the offline applications. At a luminosity of \(2 \times 10^{32} \text{ cm}^{-2}\text{s}^{-1}\), there are on average \(\mu = 1.4\) visible inelastic interactions per bunch crossing. Alternatively, the number of inelastic interactions for a given visible event, the so-called pile-up, is 1.9. Secondly, a moderate luminosity implies that the detector irradiation dose is reduced, which decreases the ageing of the detectors close to the beam. The integrated luminosity delivered to the LHCb experiment was 1.17 fb\(^{-1}\) at 7 TeV in 2011, and 2.19 fb\(^{-1}\) at 8 TeV in 2012.

![Figure 2.2: Pseudorapidity coverage of the CMS, ATLAS and LHCb detectors at the LHC. The zero-degree neutron and photon calorimeters at \(|\eta| > 8.5\) in CMS and ATLAS are mainly used for beam monitoring and tuning.](image)

2.2 The LHCb detector

The LHCb detector is built to study the physics in charm and beauty hadron decays; in particular studies of CP-violation and rare decays. The design of the LHCb detector differs from the general-purpose detectors at the LHC. Instead of being symmetric
around the interaction point, it covers only the forward direction, as shown in Figure 2.2. The reason for this specific layout is that at high center-of-mass beam energies, the $b$-hadron decays tend to be boosted in the direction of one of the beams, because the partons inside the colliding protons typically have unequal momenta. The angular distribution of the $b\bar{b}$ production peaks close to the polar angles $\theta = 0$ and $\theta = \pi$ with the beam axis, in other words the forward and backward direction. In terms of the pseudorapidity $\eta$ (defined as $\eta = -\ln|\tan(\theta/2)|$), the particles are spread more uniformly. The detector covers the pseudorapidity range $[2 < \eta < 5]$, i.e. the forward direction. The coordinate system of the detector is defined with the origin at the nominal interaction point of the protons, the $y$-axis pointing upwards, and the $z$-axis pointing along the beam towards the magnet. The $x$-axis completes a right-handed coordinate system.

![Figure 2.3: Schematic overview of the LHCb Detector. From [73].](image)

Figure 2.3 illustrates that the LHCb detector consists of different subdetectors. The proton-proton collisions take place around $z = 0$, where the Vertex Locator (VELO) is situated. The VELO is part of the tracking system, together with the Tracker Turicensis (TT) and the tracking stations T1-T3. This tracking system is used to reconstruct primary and secondary vertices in beauty and charm decays. There are two signatures that allow to distinguish these decays from other events. Firstly, the decay particles of a $b$-hadron have a higher transverse momentum than most other particles produced in proton-proton collisions. Secondly, at typical velocities produced in the LHC, both $b$-hadrons and $c$-hadrons have a decay length of a few mm. Therefore both the beauty and the charm events contain a vertex that is displaced from the proton collision point.
2.2.1 The VELO detector

The Vertex Locator is a silicon micro-strip detector, positioned closest to the proton-proton collision region [75]. A photograph of one half of the detector prior to installation is shown in Fig. 2.4a. It consists of 21 measurement planes complemented by two planes used exclusively in a trigger to veto events with high pile-up, perpendicular to the beam direction. The distance from the first to the last VELO station along the beam axis is 92.5 cm (106.5 cm including the pile-up detector). Each of these 23 stations consists of two halves or modules, which contain two sensors each. The active area of the concentric modules starts at a radius of 8.2 mm, and ends at 42 mm from the beam axis. Half of the sensors (one on each module) have a geometry of 2048 strips with varying pitch (between 38 and 92 μm) positioned radially, whereas the other half have strips oriented in the azimuthal direction, with a small stereo angle tilt. This allows to reconstruct both the r and \( \phi \) coordinates of the hits. The three-dimensional position is obtained by using the position of the detector plane in \( z \) as the third coordinate. The innermost strip is exposed to \( 5.5 \times 10^{13} \text{n}_{\text{eq}} \text{cm}^{-1} \) per \( \text{fb}^{-1} \) of integrated luminosity. To protect the detector from radiation damage during the injection and ramping of the beams, the detector modules are mechanically pulled away from the beam to a distance of 30 mm. The front-end electronics used for the VELO read-out, the Beetle chip, runs at a clock frequency of 40 MHz, and can accept trigger rates of 1.1 MHz, corresponding to a read-out time of 900 ns.

In order to shield the sensors against radio-frequency pickup from the traversing LHC beams, they are placed in a 300 μm thick aluminium box, the so-called RF-box. This box also serves as a secondary vacuum separating the VELO detector from the LHC vacuum. As is visible in Fig. 2.4b, the RF-box has a corrugated structure that surrounds the VELO sensors, which is designed to limit the amount of material which the particles traverse.

Figure 2.5 illustrates that the primary vertex resolution in data for a primary vertex with 25 tracks is about 85 μm in z and 13 μm in both x and y. A comparison of these resolutions with MC showed that the 2011 simulation has the same resolution as data in x and y, and is approximately 70 μm in z.

The resolution of the track impact parameter (IP), defined as the closest distance between the vertex and the trajectory of the particle, is below 40 μm for tracks with \( p_T > 1 \text{ GeV} \), as shown in Fig. 2.6a. The IP resolution is governed by three main factors: multiple scattering of particles by the detector material; the resolution on the position of hits in the detector from which tracks are reconstructed; and the distance required to extrapolate a track from its first hit in the detector to the primary vertex. Fig. 2.6a shows that there is no perfect agreement between the resolution in data and in simulation. This is due to an inaccurate description of the material distributions in the detector. The IP resolution as a function of azimuthal angle, shown in Fig. 2.6b, illustrates that the agreement between data and MC depends on \( \phi \). The amount of material...
Figure 2.4: (a) One half of the VELO detector prior to installation, showing all the detection planes, and (b) one half of the VELO detector and of the RF-shielding box.

Figure 2.5: Primary vertex resolution of events in 2011 data with exactly one PV as a function of track multiplicity. In (a) the x- and y-resolution, and in (b) the z-resolution. The data points are fitted with a function $A/N^B + C$. From [76].
is underestimated in most regions of $\phi$, resulting in an underestimated resolution in the simulation. There is an increase in IP resolution (both in data and in simulation) around $\phi = \pm \pi/2$. This increase reflects the increasing material density where not only the two halves of the VELO, but also the two sides of the RF foil overlap. In this region, the material description is better, and the IP resolution is not underestimated. The material distribution is more accurately described in the simulation of 2012 data, resulting in a better agreement between the resolutions in data and in MC [76].

2.2.2 The tracking stations

The VELO is part of the tracking system, which is supplemented by several other subdetectors. The second detector positioned upstream of the magnet is the Tracker Turicenisis (TT), which consists of silicon microstrips with 183 $\mu$m pitch. It contains four detection layers, of which the first and last have a vertical strip alignment, and the middle two layers are rotated over a small stereo angle, which results in a layer orientation of $0^\circ$, $+5^\circ$, $-5^\circ$ and $0^\circ$. The TT is complementary to the VELO because it enables the detection of tracks from particles produced in decays downstream from the primary interaction point. Furthermore, the additional position measurements improve the momentum resolution for tracks that traverse the magnet.

A warm dipole magnet with a bending power of 4 Tm bends particles in the $xz$-plane and enables the determination of the momenta of the particles by measuring their track curvature. Approximately half of the LHCb data is taken with magnet polarity ‘up’, and half with polarity ‘down’, in order to be able to study systematic detection efficiencies that depend on the charge of the particles.

The last component of the tracking system is located downstream of the magnet and consists of two parts: a silicon-strip inner tracker (IT) and a gaseous strawtube
outer tracker (OT). The inner tracker covers a small area close to the beam, and contains three stations (positioned at T1, T2 and T3 in Fig. 2.3) of four separate boxes of four-layer silicon-strip detectors, of which the middle two layers are again rotated by a five degree stereo angle. The inner region has a high particle flux, and therefore benefits from a silicon detector with a strip pitch of 198 \( \mu \text{m} \), which can cope with high occupancies.

The OT is built around the IT, and covers the full LHCb acceptance. The OT comprises three straw tube drift-time stations (positioned at T1, T2 and T3), each built out of four layers with the same stereo structure as the TT and IT. Since the outer region is less occupied with particles than the inner region, the OT uses gas-filled straw tubes with a diameter of 5 mm, in order to cover a large surface, be it with a slightly worse position resolution.

### 2.2.3 Particle identification

Besides a tracking system the LHCb detector also has a particle identification system separating pions, kaons, protons, muons and electrons, which is essential for the determination of exclusive final states in the decay of heavy flavour. The particle identification is performed by two Ring Imaging Cherenkov (RICH) detectors, an electromagnetic and a hadronic calorimeter and five muon stations. The muon detector and the calorimeters additionally provide the information needed for the hardware trigger decision. More specifically, the L0 trigger uses the transverse momentum of the muons and the energy deposits of electrons and hadrons. The trigger setup is discussed in more detail in Section 4.1.

The RICH system consists of two separate detectors. RICH1, positioned upstream of the magnet, is optimised for low-momentum particles between 2 and 40 GeV. RICH2, positioned behind the T-stations, can measure particles with momenta up to 100 GeV. The RICH detectors identify charged particles by measuring the Cherenkov radiation produced in a radiator that consists of gas in RICH2, and of both gas and silica aerogel in RICH1. The Cherenkov emission cone is projected onto a photon detection surface outside the active area of LHCb using spherical and flat mirrors. Angular resolutions of 1.618\( \pm 0.002 \) mrad and 0.68\( \pm 0.02 \) mrad have been achieved for Cherenkov angles in RICH1 and RICH2, respectively [77].

The calorimeter system is positioned downstream of the RICH detectors. First, the Scintillator Pad Detector (SPD) helps to determine the difference between charged particles that deposit energy in the scintillator material and neutral particles that do not interact, enabling the distinction between electrons on the one hand and photons and neutral pions on the other. Second, the Pre-Shower detector (PS), in combination with the rest of the electromagnetic calorimeter, provides longitudinal segmentation of the electromagnetic shower, such that the energy deposits of different particles can be measured. This can be used to separate between the charged pion background and
the more energetic electrons. A lead converter is placed in between the SPD and PS detectors.

The measurement of electromagnetic and hadronic energy deposits happens in two calorimeters positioned downstream of the SPD/PS system. Both calorimeters consist of alternating lead and scintillator layers. The electromagnetic calorimeter (ECAL) is used to measure the position and the energy of photons and electrons, whereas the hadronic calorimeter (HCAL) detects heavier particles like protons, pions and neutrons. The ECAL has a relative energy resolution $\sigma/E = 10%/\sqrt{E[GeV]} + 1\%$, and the HCAL $69%/\sqrt{E[GeV]} + 9\%$ [73].

Muons are identified by five stations of muon chambers. Four of those, made of multi-wire proportional chambers, are located behind the HCAL, alternated with 80 cm thick iron absorbers. The fifth station is positioned in between RICH2 and the calorimeters in order to provide a position measurement before the bulk of material of the calorimeters gives rise to multiple scattering. In order to cope with the high occupancy, this station is equipped with a triple-GEM (Gas Electron Multiplier) detector. The muon system can detect muons with a transverse momentum above 0.8 GeV and a momentum above 3.0 GeV with an efficiency of 98.13 ± 0.04 %. This number has been retrieved from a $J/\psi$ calibration sample. The misidentification probabilities are 1.033 ± 0.003 %, 1.025 ± 0.003 % and 1.111 ± 0.003 % for protons, pions and kaons, respectively [78].

2.3 Tracking

The track reconstruction algorithms in LHCb apply different strategies, depending on whether the reconstruction is performed in the online or in the offline environment. Both of these scenarios start with pattern recognition in the VELO detector, which is performed by the FastVelo algorithm [79]. It uses the hits in both the $\phi$ and the $R$ sensors in the VELO to fit a three-dimensional straight line through these clusters, exploiting the absence of a magnetic field in the VELO region.

In the online environment, the VELO seeds are then extrapolated to the T stations, where individual hits are added to the track (forward tracking). The forward tracking is implemented in the PatForward algorithm [80]. The expected position of a trajectory in the T stations is defined as a function of the VELO seed parameters and of the position of single hits in the first T station. The hits in the subsequent T stations are collected in a window around this expected position. Hits lying in between the VELO and the tracking stations, in the TT, are added if they are close enough. Some quality cuts are applied on the candidates to eventually store them as so-called 'long' tracks.

In the offline environment, there are two strategies to extend the VELO seeds. The first one is the forward tracking, as used in the online environment. The pattern recognition and the track fit that are used offline are similar to the online algorithms,
except that they are optimised for better precision instead of for fast execution. The second strategy is to start with the reconstruction of track segments in the individual subdetectors, after which those parts can be combined into trajectories traversing the complete detector (track matching). A stand-alone track finding is performed in the IT and OT tracking stations, where T station hits are collected to make seeds, on which several cuts are applied before they are accepted as tracks. The algorithm used for this seeding procedure is called PatSeeding [81]. The segments in the T stations are subsequently extrapolated through the magnet to the VELO to find the best matching VELO segment, after which TT hits are added.

![Figure 2.7: Track types in LHCb.](image)

By matching different segments together, several types of charged particle trajectories can be defined, as illustrated in Figure 2.7. The VELO segments can be matched to only TT hits (‘upstream’ track), or to both TT hits and T station segments (‘long’ track). The combination of a T segment with matching hits in the TT detector constitutes a ‘downstream’ track. Those downstream trajectories can originate from long-lived particles decaying outside the VELO acceptance. The downstream tracking is performed in the PatDownstream algorithm [82].

At the end of the reconstruction sequence, the best candidates are selected from the various track categories, at which stage duplicates (clones) are removed. After each track is fitted using a Kalman filter, all resulting tracks in the event are stored in a track list.

The track reconstruction efficiency must be known to calculate efficiencies to select signal events. Track reconstruction efficiencies in LHCb are estimated from simulated events. In addition, data-driven tag-and-probe methods are used to measure the single-track efficiency and can be used to correct the simulation. The reconstruction efficiency for long tracks has been determined from a tag-and-probe method using $J/\psi \rightarrow \mu^+\mu^-$ decays. A fully reconstructed muon is used as the tag, probing the ef-
ficiency of reconstructing the other muon. This is done for different track segments; the T station efficiency, for example, is calculated using tracks that have hits both in the VELO and in the muon chambers. The overall efficiency depends on the momentum and pseudorapidity of the tracks and on the particle multiplicity of the event. For example, for the momentum spectrum of the $J/\psi$ decay products, the overall efficiency to reconstruct a VELO track, given that there is a downstream track found in the detector, is 96.79 $\pm$ 0.07% [83].

The full tracking system provides a momentum resolution $\delta p/p$ between 0.4% for tracks with 5 GeV momentum and 0.6% for 100 GeV tracks. The momentum resolution is important for the invariant mass computation of decaying particles.

2.4 Trigger

The rate of bunch crossings with at least one visible proton-proton interaction was approximately 11 MHz in 2011. However, the nominal LHC bunch crossing rate at 25 ns bunch spacing, where almost each crossing provides a collision, results in an event rate for visible interactions of about 30 MHz. A rigorous trigger selection is needed in order to reduce the data rate. The flow chart of the trigger decision is shown in Fig. 2.8 [84]. First, a hardware trigger (L0) reduces the event rate from 30 MHz to 1.1 MHz, using information from the muon chambers and the calorimeters. Second, the software high level trigger (HLT), uses the full event data to further reduce the event rate in two stages, HLT1 and HLT2, from 1.1 MHz to 3 kHz. The latter is the event rate that is stored for offline analysis. The trigger stages are each divided into categories (depicted as boxes in Fig. 2.8) that select events from different physical processes. To reduce the output of certain trigger lines, a 'postscale' can be applied to the trigger output, which means that only a fraction of the selected sample is stored. One can also apply a 'prescale' to the trigger input, such that only a fraction of the input events is considered for the trigger selection. A general overview of the LHCb trigger framework is given in the next paragraphs.

2.4.1 L0

The level-0 (L0) hardware trigger reduces the rate of bunch crossings to the maximum read-out speed of the data acquisition system, namely 1.1 MHz. At the full bunch-crossing frequency, the L0 can reconstruct the two highest $p_T$ muons, and the highest $E_T$ hadron, electron and photon. The SPD detector distinguishes between electrons and photons, whereas the requirement of energy deposit in the PS decreases the contamination of hadrons.

The hardware trigger decision is based on the presence of either a hadron candidate, a photon candidate, an electron candidate, or one or two muon candidates that reach a certain $E_T$ threshold. Events are grouped accordingly into different trigger
lines, which are subsets of the categories listed in Fig. 2.8. A trigger line is a sequence of selections and reconstruction algorithms.

There is an additional so-called 'global event cut' on the maximum multiplicity in the SPD detector, which removes crossings in which the occupancy is too high to allow a successful online reconstruction of the event. Depending on the trigger line through which the event is selected, the SPD multiplicity cut is either 600 or 900 hits. The fraction of $b$-hadron events rejected by this cut is $8.8 \pm 0.6\%$ ($0.5 \pm 0.2\%$) in data with a luminosity corresponding to an average number of visible interactions per bunch crossing of $\mu = 1.4$, for a cut on 600 (900) SPD hits [84]. A more detailed overview of the selection criteria in the different trigger stages that are used for the long-lived particle analysis is given in Chapter 4.

2.4.2 HLT1

The first level software trigger is optimised to process the events at the maximum read-out speed of the hardware. The HLT1 can, due to CPU restrictions, only perform a partial event reconstruction. The track reconstruction uses the online tracking algorithm described in Section 2.3. VELO tracks that have a relatively small impact parameter to the primary interaction region are selected, in order to eliminate tracks that do not originate from the proton-proton collision. Subsequently, primary vertices with five
or more tracks that are within a transverse distance $R_{xy} < 0.3$ mm from the mean position of the proton-proton interaction point are reconstructed. Events that were selected in L0 because of the presence of a muon candidate are reconstructed in HLT1 by matching a VELO segment to hits in the muon chambers.

The HLT1 selects events with at least one VELO track with a large impact parameter ($IP > 0.1$ mm) to the primary interaction, or with one or two reconstructed muon candidates. These requirements are optimised for the selection of $b$-hadron decays, and are also well-suited for the selection of exotic long-lived particles. In the selected events, the VELO tracks are extended to long tracks using the forward tracking algorithm, as described in section 2.3, which enables the measurement of the track momenta. Eventually, events with at least one track with a transverse momentum larger than 1 GeV are selected. The HLT1 reduces the event rate to about 50 kHz.

### 2.4.3 HLT2

At the second software trigger level, the event rate is sufficiently low, such that a more extensive track and vertex reconstruction can be carried out. The forward tracking is applied to all VELO tracks, and the processing time is decreased by narrowing the search windows such that only tracks with a transverse momentum above 500 MeV are reconstructed [84]. The HLT2 comprises many different trigger lines with different event reconstruction algorithms and selections. A significant number of those lines base the selection decision on the presence of a displaced vertex containing two to four tracks, which is the signature of a charm or beauty decay. Other lines also use particle identification to select specific decay signatures. There are for example lines selecting one muon, di-muons, or displaced vertices with a large invariant mass. The total output of the various HLT2 trigger lines is around 3 kHz. The HLT2 lines that were designed for the long-lived particle search are discussed in detail in Section 4.3.

### 2.4.4 Definition of the trigger decision

In order to be able to calculate the efficiencies of the trigger lines, the online selected particle candidates should be matched to offline particles. This can be done using the so-called 'TIS-TOS' method, in which the events are divided into categories according to the way in which they were triggered. An event that was 'triggered-on-signal' (TOS) includes a trigger object in which all track candidates 'overlap' with the offline signal candidate. The candidates are considered to overlap if they share more than 70% of the hits. Alternatively, an event that is triggered 'independent-of-signal' (TIS) has no track overlap between the trigger object and the offline signal. In this case, two tracks are defined to overlap when more than 1% of their hits are shared. A looser definition of the TOS category is called TUS ('triggered-using-signal'), where at least one triggered track should overlap with the signal. The reason to define these trigger categories, is
to ease the definition of trigger efficiency, by knowing exactly on which offline object a certain line was triggered [85].

2.5 Software framework

LHCb uses a common software framework for the data acquisition, simulation and physics analysis. Whereas the full data processing is incorporated in the GAUDI project, several different projects group the software for specific tasks.

The simulation of Monte Carlo data happens within the GAUSS project[86]. The proton-proton collisions are generated using PyTHIA v6.4 [87], and the interactions of the final state particles in the detector are implemented with a GEANT4 simulation [88] using the detector geometry as described in GAUSS. The digitisation of the hits in the individual subdetectors is simulated in the BOOLE project. The MOORE project includes all high-level trigger algorithms, and the BRUNEL project accumulates the subsequent reconstruction algorithms for offline clustering of hits, tracking, vertex reconstruction and particle identification. The offline reconstruction is performed in a common reprocessing of all the data, the so-called 'stripping' process. Different stripping lines exist for various physics analyses, similar to the trigger lines described before. The stripping algorithms are grouped in the DAVINCI package, together with the physics analysis software tools. The BENDER package is the environment from which one can access those tools and algorithms using the PYTHON scripting language.

All the data that is processed through the stripping is stored on several GRID sites. The GRID unites the computing resources of particle physics institutes throughout the world, and offers the safe storage and permanent accessibility of data. The computing power of the grid is used for stripping, subsequent offline analysis, and the generation of simulated data.
Before starting a search for long-lived particles, it is essential to know how many of those particles are expected to be found in the LHCb detector acceptance. In order to calculate the expected number of events, and to study the properties of the signal particles, several MC signal samples are simulated. To get an estimate of the expected standard model background, various samples of inclusive beauty and charm events have been simulated and studied.

### 3.1 Simulated signal

#### 3.1.1 Hidden valley

The benchmark model used for this analysis is a hidden valley (HV) model in which two $\pi^0$ particles are produced in the decay of a SM-like scalar Higgs of 120 GeV. This model is discussed in Section 1.2.3, and illustrated in Fig. 1.6. The generated mass of the Higgs boson is set to 120 GeV, as the samples were generated just before the discovery of the Higgs-like boson of 126 GeV at the ATLAS and CMS experiments.

Different samples are generated with the $\pi^0$ lifetime set to 10 and 100 ps, which allows reweighting for intermediate lifetimes. The $\pi^0$ mass is set to either 15, 25, 35, 43 or 50 GeV. In the benchmark model, both long-lived particles are forced to decay to $b\bar{b}$. Two additional samples with the $\pi^0$ decaying into either $c$- or $s$-, $u$- and $d$-quarks have been generated (labelled 'HV10_CC' and 'HV10_SS'). The benchmark model (called 'HV10_M35') has a $\pi^0$ lifetime of 10 ps and a mass of 35 GeV. A full list of the simulated signal samples is given in Table 3.1. Some samples were generated with an incorrect primary vertex multiplicity. Therefore, a reweighting of events is applied at the final
<table>
<thead>
<tr>
<th>Eventtype</th>
<th>Number of events</th>
<th>Generator efficiency</th>
<th>ν</th>
<th>Parameters</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Signal</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HV10_M15</td>
<td>43900004</td>
<td>97497</td>
<td>0.278±0.002</td>
<td>2* m_H = 120, m_{π^+} = 15, τ_{π^+} = 10</td>
</tr>
<tr>
<td>HV10_M25</td>
<td>43900020</td>
<td>216500</td>
<td>0.294±0.003</td>
<td>2  m_H = 120, m_{π^+} = 25, τ_{π^+} = 10</td>
</tr>
<tr>
<td>HV10_M35</td>
<td>43900012</td>
<td>64200</td>
<td>0.309±0.002</td>
<td>2* m_H = 120, m_{π^+} = 35, τ_{π^+} = 10</td>
</tr>
<tr>
<td>HV10_M35</td>
<td>43900012</td>
<td>216998</td>
<td>0.309±0.002</td>
<td>2  m_H = 120, m_{π^+} = 35, τ_{π^+} = 10</td>
</tr>
<tr>
<td>HV10_M43</td>
<td>43900015</td>
<td>218499</td>
<td>0.309±0.004</td>
<td>2  m_H = 120, m_{π^+} = 43, τ_{π^+} = 10</td>
</tr>
<tr>
<td>HV10_M50</td>
<td>43900005</td>
<td>98000</td>
<td>0.313±0.003</td>
<td>2* m_H = 120, m_{π^+} = 50, τ_{π^+} = 10</td>
</tr>
<tr>
<td>HV10_CC</td>
<td>43900006</td>
<td>65500</td>
<td>0.324±0.003</td>
<td>2* m_H = 120, m_{π^+} = 35, τ_{π^+} = 10</td>
</tr>
<tr>
<td>HV10_SS</td>
<td>43900007</td>
<td>65500</td>
<td>0.289±0.002</td>
<td>2* m_H = 120, m_{π^+} = 35, τ_{π^+} = 10</td>
</tr>
<tr>
<td>HV100_M15</td>
<td>43900016</td>
<td>217999</td>
<td>0.278±0.003</td>
<td>2  m_H = 120, m_{π^+} = 15, τ_{π^+} = 100</td>
</tr>
<tr>
<td>HV100_M25</td>
<td>43900017</td>
<td>216998</td>
<td>0.294±0.003</td>
<td>2  m_H = 120, m_{π^+} = 25, τ_{π^+} = 100</td>
</tr>
<tr>
<td>HV100_M35</td>
<td>43900013</td>
<td>61500</td>
<td>0.309±0.002</td>
<td>2* m_H = 120, m_{π^+} = 35, τ_{π^+} = 100</td>
</tr>
<tr>
<td>HV100_M43</td>
<td>43900018</td>
<td>215498</td>
<td>0.309±0.004</td>
<td>2  m_H = 120, m_{π^+} = 43, τ_{π^+} = 100</td>
</tr>
<tr>
<td>HV100_M50</td>
<td>43900019</td>
<td>214999</td>
<td>0.315±0.004</td>
<td>2  m_H = 120, m_{π^+} = 50, τ_{π^+} = 100</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Background</th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>INCLB_5PS</td>
<td>10000022</td>
<td>7055964</td>
<td>0.311±0.003</td>
<td>2  τ_b = 5 ps, R_{xy} &gt; 0.4mm</td>
</tr>
<tr>
<td>INCLB_1DV</td>
<td>10000021</td>
<td>2533490</td>
<td>0.048±0.001</td>
<td>2  ≥ 1 b with R_{xy} &gt; 0.4mm</td>
</tr>
<tr>
<td>INCLB_2inacc</td>
<td>10000010</td>
<td>5029475</td>
<td>---</td>
<td>2  ≥ 2 b quarks in acceptance</td>
</tr>
<tr>
<td>INCLC</td>
<td>20000010</td>
<td>9148460</td>
<td>0.253±0.004</td>
<td>2.5 ≥ 2 c quarks in acceptance</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Systematics</th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Zmmumu_PHOTO</td>
<td>42112002</td>
<td>1033494</td>
<td>---</td>
<td>2.5 1 lepton with p_T &gt; 4GeV</td>
</tr>
<tr>
<td>B^0 \to J/ψ K^{*0}</td>
<td>11144001</td>
<td>10M</td>
<td>---</td>
<td>2 all daughters in acceptance</td>
</tr>
<tr>
<td>inclusive J/ψ</td>
<td>24142001</td>
<td>20M</td>
<td>---</td>
<td>2 both daughters in acceptance</td>
</tr>
</tbody>
</table>

**Table 3.1:** Main simulated signal samples used for this analysis, generated with the LHCb 2011 Monte Carlo simulation (MC11a) with the reconstruction version reco12a. The trigger simulation is performed with MOORE v12r8p1 and trigger configuration TCK 0x40760037. 'Eventtype' is the unique event generation number. 'Generator efficiency' is the efficiency of the generator level selection. 'ν' is the average number of proton-proton interactions per bunch crossing. 'Parameters' lists the signal parameters and the generator cuts made on the background samples. R_{xy} is the radial distance of a vertex to the pp interaction region. Masses m_H and m_{ψ^*} are in GeV; lifetimes τ in ps. The samples indicated with an asterisk (*) have an incorrect PV multiplicity distribution.
stage of the analysis in order to retrieve the correct multiplicity distributions. The HV10_M35 benchmark model, used for most of the studies described in the analysis, has the correct multiplicity.

### 3.1.2 Detector acceptance

Several cuts are applied at generator level to ensure that the generated events are in the acceptance of the LHCb detector. In order to be accepted, an event needs to include one $\pi^0_v$ with at least four stable 'reconstructible' daughters. A particle is considered 'reconstructible' if it is charged, has $p > 2$ GeV and an azimuthal angle $\theta < 400$ mrad. These are the only generator level cuts that are applied for the signal samples. The generator level efficiencies are listed in Table 3.1.

### 3.1.3 Generator level studies

The kinematic distributions of the Higgs boson illustrated in Fig. 3.1 are equal for all signal models. The Higgs events that pass the generator level selection lie mainly in the pseudorapidity range of LHCb, roughly between 1 and 6. A few events appear at negative $\eta$, which can be explained by the fact that, although the Higgs particles have negative $z$-momentum, one of the long-lived particle daughters can still decay within the LHCb acceptance and pass the generator cut. The fact that all the $\pi^0_v$ particles are pair-produced in the resonant decay from a Higgs limits the allowed kinematic region for the long-lived particles. The higher the $\pi^0_v$ mass, the lower its $p_T$, as shown in Figure 3.2. This $p_T$ limit also affects the radial distance from the beam axis at which the long-lived particle will decay. Low-mass particles with higher $p_T$ can decay at a larger radius, while remaining within the detector acceptance. For the higher masses, the acceptance of the detector starts to play a role. The plots in Figure 3.2 illustrate that although a high-mass $\pi^0_v$ produces more charged particles at generator level, those particles often tend to be outside the acceptance of LHCb. The mean number of tracks in the acceptance is approximately equal for the three masses shown.

Jets, collections of collimated neutral and charged particles, can be used to make an estimate of the invariant mass of the long-lived particles. The number of jets per $\pi^0_v$ in Fig. 3.3a is obtained by running a jet reconstruction algorithm only on particles origination from a true $\pi_v$ decay. Particles that point to the displaced vertex are used as input for the anti-$k_t$ algorithm [89], with a cone size radius $R = 0.7$. The jet algorithm for MC is explained in more detail in Section 5.1.3. Since the true jets in the MC are reconstructed from HepMC [90] particles (i.e. pions, kaons, protons, neutrons, hyperons, photons, electrons, muons and neutrinos), they represent the jets as they would be measured with a perfect detector. There are usually two jets, but for decreasing $\pi^0_v$ mass the jets are found to be merged more frequently, such that there are mostly one-jet candidates for the lowest $\pi^0_v$ mass. This is also illustrated by the jet
$p_T$ distribution in Figure 3.3b. The jet $p_T$ spectrum indicates that for the lowest mass sample one jet with large $p_T$ is reconstructed, while for the higher masses there will be multiple lower-$p_T$ jets. Figure 3.3c shows the reconstructible dijet mass for candidates with two jets. The number of remaining candidates with two jets is small for the low-mass sample. The median of these distributions is at the generated mass value, although the mass resolution is poor. The tails are mainly caused by jets that were incorrectly reconstructed by the jet algorithm, and the low-mass tail is enhanced by jets that fall partly outside the LHCb acceptance.

The distance between two jets can be defined using the pseudorapidity and the azimuthal angle between the jets as $\Delta R = \sqrt{\Delta \eta^2 + \Delta \phi^2}$. The $\Delta R$ distributions in Figure 3.3d indicate that the jets of heavy $\pi^0$s are more back-to-back. The cut-off at 0.7 corresponds to the jet cone size that determines the minimal distance between two jets.

In summary, a loss in detection efficiency is expected for low-mass candidates due to the dijet requirement. The mass resolution deteriorates in general due to the jet reconstruction algorithm, and for the high-mass candidates in particular due to the limited detector acceptance.

### 3.1.4 Expected number of events

To assess the sensitivity of LHCb to exotic long-lived particles, an estimate is made of the number of those particles that are expected to decay within the detector volume. At a center-of-mass energy $\sqrt{s} = 7$ TeV, the predicted SM Higgs production cross-section for the HV10_M35 model featuring a 120 GeV Higgs particle is [91]:

$$\sigma_H = 18.85 \pm 6.3\% \ (\text{QCD scale}) \pm 6.6\% \ (\text{PDF} \, \& \, \alpha_s) \ \text{pb}$$
Figure 3.2: Transverse momentum (a), radial distributions (b), number of tracks of the $\pi^0$ decay (c) and number of $\pi^0$ decay tracks in the LHCb acceptance (d) of $\pi^0$ particles at generator level for a simulated hidden valley signal model with $\tau_{\pi^0} = 10$ ps and various generated $\pi^0$ masses.
Figure 3.3: Number of jets per $\pi^0_\nu$ (a), the $p_T$ of jets associated to the $\pi^0_\nu$ (b), dijet mass of dijet $\pi^0_\nu$ candidates (c) and $\Delta R$ of dijet $\pi^0_\nu$ candidates (d) at generator level for a simulated hidden valley signal model with $\tau_{\pi^0_\nu} = 10$ ps and various generated $\pi^0_\nu$ masses. Note that the number of candidates with two jets is very small for the 15 GeV sample.
considering all production processes, namely gluon-gluon fusion, vector boson fusion, WH production, ZH production and ttH associate production. The production is dominated by gluon-gluon fusion.

The main uncertainties on the production cross-section are the theoretical uncertainties due to the choice of the quantum chromo dynamics scale (QCD scale), the uncertainties on the parton distribution function (PDF) sets and uncertainties in the computation of the strong coupling ($\alpha_s$). At $\sqrt{s} = 8$ TeV the cross section increases to:

$$\sigma_H = 23.93 \pm 6.4\% \text{ (QCD scale)} + 6.5\% \text{ (PDF & } \alpha_s \text{)} \text{ pb.}$$

The expected number of signal events $s$ can be calculated as follows:

$$s = \mathcal{L} \times \epsilon \times \sigma(H \rightarrow \pi^0_v \pi^0_v) \times 2BR(\pi^0_v \rightarrow b\bar{b}) \quad (3.1)$$

For the 0.62 fb$^{-1}$ of 2011 data at 7 TeV used in this analysis, the cross section $\sigma_H$ corresponds to approximately 11.7k produced Higgs events. Forcing the Higgs to decay only via the process $H \rightarrow \pi^0_v \pi^0_v \rightarrow b\bar{b}b\bar{b}$, about 4000 of those decays are in the detector acceptance. This can be deduced from the generator level selection efficiency for this sample. Therefore, in order to set a meaningful limit, i.e. to have a few expected signal events in the data set, the selection efficiency $\epsilon$ on the signal events in the acceptance must be larger than approximately 0.1%.

### 3.2 Simulated background

One of the main challenges of the exotic long-lived particle search is the prediction of the standard model background. The high-mass and high-lifetime range that is considered for this analysis should exclude SM particles. However, this background can be enhanced by mis-reconstructions, or the combination of several particles into one long-lived candidate. The events that can constitute the background are weakly decaying hadrons. Since strange hadrons are both too light to pass the mass requirement, and they have a very long lifetime, the only backgrounds considered here are beauty and charm events. Note that these decays cannot by themselves constitute a 10 GeV vertex mass. Other particles have to be included in the vertex to add up to the mass threshold.

The largest problem is that it is not possible to simulate a full detector response of a representative inclusive $b\bar{b}$ or $c\bar{c}$ sample of the LHCb data, namely $10^{11}$ and $10^{12}$ events, respectively, with the current computing resources. An attempt is made to generate a selected sample that still represents the events that pass the final selection. However, because the background sources are not purely caused by physics, but also by mis-reconstructions and combinatoric backgrounds, it is difficult to define a set
of criteria that selects the representative events at generator level. Some attempts have been made, and a more thorough report of the results of these studies will be presented in Chapter 7. Since the resulting statistics are still far from what is desired, these Monte Carlo background estimates will not be used for the final results of the analysis.

### 3.2.1 Beauty decays

The production cross-section of $b\bar{b}$ is $288 \pm 40$ μb, which is a weighted average of results reported in [92, 93]. This results in $1.8 \times 10^{11}$ events in 0.62 fb$^{-1}$, or $8.6 \times 10^9$ events in the detector acceptance. It is not possible to generate sufficient fully simulated Monte Carlo events to match the statistics of inclusive $b$-hadron decays in data; at most a few million events can be produced.

The first MC sample, which is expected to be characteristic for the background in the long-lived particle search, is an inclusive $b\bar{b}$ sample ('INCLB_1DV') requiring at least one of the $b$-hadrons to be produced in the detector acceptance ($1.5 < \eta < 5.0$), to have a transverse momentum $p_T > 2$ GeV, to have its decay vertex with a distance $R_{xy} > 0.4$ mm from the beam line and to have at least four charged daughters. The requirement on the radial distance serves to select particles with long lifetimes, and the requirement of at least four daughters favours events that are likely to contain high-mass candidates. It turns out that the radial cut alone does not enhance the lifetime distribution enough to get a reasonable background yield after the full selection (the selected yields are listed in Section 7.1).

In an attempt to enhance the lifetime distribution, a second sample ('INCLB_5PS') is produced, in which the average lifetime of the $b$-hadrons is increased to 5 ps. The number of $b$-hadrons with a lifetime over 10 ps produced in this simulation should be larger than the number produced in the 2011 data set. At least one of the $b$-hadrons must be in the detector acceptance ($1.5 < \eta < 5.0$), have a transverse momentum $p_T > 2$ GeV and have its decay vertex with a distance $R_{xy} > 0.4$ mm from the beam line. A difference with the INCLB_1DV sample is that only one charged daughter of the $B$-vertex is required to be in the acceptance. In this way, the selection includes events with $B$-vertices that pick up tracks in the reconstruction that do not necessarily come from the $b$-hadron decay itself. Therefore the sample will on the one hand be more representative for the data surviving the final selection, but on the other hand it will have a reduced overall mass distribution (compared to the INCLB_1DV sample), which lowers the probability that events survive the selection criteria for high mass candidates.

Unfortunately, the adjusted lifetime does not have the desired effect, due to the requirement that both $b$-hadrons get a 5 ps average lifetime. To correct for the increased lifetime, events at large decay length get a weight assigned to match their frequency in the data. Events with a large lifetime get a small weight, and for events where both hadrons live long, the weight is even smaller. Since it is likely that the background in
data consists of events where multiple decays are combined to form one candidate, it would be incorrect to consider the weight of only one of the two $b$-decays in the event. The requirement to have two $b$-hadrons with increased lifetime therefore produces very improbable events with low weights, thereby reducing the statistical power of the sample. An alternative approach would be to generate a sample where only one of the $b$-hadrons has an enhanced lifetime.

In the end, the INCLB_1DV and the INCLB_5PS sample result in comparable selection yields, so both samples will be used to make a rough estimate of the background at intermediate selection stages. However, the statistical error on this estimate is too large to use these samples for a background estimation to set an upper limit or claim a discovery. Instead, the data will be modelled by a smooth background mass shape.

3.2.2 Charm decays

In addition to beauty decays, also charm decays can contribute to the background for long-lived particles. The background from charm is expected to be lower than that from beauty, mainly because the $c$ quark is lighter. The problem of limited statistics enters here as well. After the unsuccessful attempts to enrich the inclusive $b$ events with signal-like candidates, no effort is made to make a more representative sample for charm. The existing simulation will be used to study the yields at intermediate selection levels, but the final background shape will be extracted from data.

The inclusive $c\bar{c}$ sample (INCLC) that is generated only requires that both of the $c$ quarks decay in the detector acceptance ($1.5 < \eta < 5.0$). This sample also includes charmed mesons from $b$-hadron decays, which constitute about 10% of the events.

The production cross-section of $c\bar{c}$ is retrieved from the measurement of inclusive states in LHCb. It amounts to $\sigma(c\bar{c}) = 68.2 \pm 7.6$ mb for the phase space $2.5 < y < 4.0$ and $p_T > 3$ GeV. Using PYTHIA, this number can be extrapolated to full space, which results in $\sigma(c\bar{c}) = 5.180 \pm 0.577$ mb [94]. This corresponds to $(3.3 \pm 1.1) \times 10^{12}$ events in $0.62$ fb$^{-1}$, or $8.2 \times 10^{11}$ events in the LHCb acceptance. The selection yields of the INCLC sample are listed in Section 7.1.

3.3 Data

The long-lived particle analysis is performed using the 2012 processing of the 2011 LHCb data set. The real data and the simulated samples are processed with identical software chains, as discussed in Chapter 2. The trigger and stripping used to select the data are described in Chapter 4. The stripping lines used for this analysis are grouped in the ‘electroweak’ (EW) stream. Table 6 in Appendix B lists the magnet-up and magnet-down event yields in the 2011 EW stream, divided into periods with
different trigger configurations (TCKs). The total data set recorded at LHCb in 2011 corresponds to an integrated luminosity of $1.013 \pm 0.017$ fb$^{-1}$.

Due to an optimisation of the displaced vertex triggers half-way through the year, the first and second part of the data set have different selection criteria, and different efficiencies. Here, only the second part of the data is used, namely the part with the highest trigger efficiency, processed with trigger software version MOORE v12r8 or later. This implies that only $0.624 \pm 0.011$ fb$^{-1}$ is used for this analysis, constituting approximately two-thirds of the 2011 data sample.
Trigger and vertex reconstruction

The selection of events with long-lived particle candidates in the HLT2 trigger and offline selection (stripping) stages relies on the presence of a displaced vertex. The invariant mass of the vertex, its number of tracks, the total $p_T$ of those tracks and the flight distance of the particle distinguish the signal from background. Dedicated algorithms and selection procedures are applied on the data to select signal events. This chapter starts with describing the hardware (L0) and first level software trigger (HLT1), followed by the algorithms used in the second level software trigger (HLT2). The last part of the chapter discusses the offline reconstruction and event selection in the stripping procedure. Figure 4.1 gives a schematic view of the different steps involved.

4.1 Trigger L0 and HLT1 selection

To achieve a trigger selection with a high efficiency, all L0 and HLT1 physics trigger lines described in Section 2.4 are considered, and eventually the ones with the highest efficiency on the long-lived particle simulated signal are selected. In order to make this decision, the simulated signal events are first reconstructed and selected up to the final selection stage, without applying any trigger requirements. The full selection is described in chapters 5 and 6. Only at the end of the procedure, the trigger selections are applied, and the loss in efficiency due to the trigger is measured per for each individual trigger line. The result of this procedure is shown in Table 4.1 for L0, and in Table 4.2 for HLT1. Note that the trigger decision 'TOS' is required (see Section 2.4.4 for the definition of 'TOS' or 'triggered-on-signal'). Requiring a link between
Figure 4.1: Dataflow of the displaced vertex analysis. The ellipses represent data types, and the boxes are selection and reconstruction stages. The first three rows of boxes correspond to the 'L0', 'HLT1' and 'HLT2' trigger stages in Fig. 2.8. 'DST' stands for Data Summary Tape, a data format which contains information about the reconstructed physics quantities and some additional information of the raw event.
the triggered object and the offline candidate facilitates the evaluation of the trigger efficiencies and the systematic uncertainties on those efficiencies. The inputs for the 'TOS' decision are the offline reconstructed particles in the displaced vertex, as well as the particles in the jets associated to that vertex. Both the L0 and HLT1 selections trigger on objects such as single tracks or calorimeter clusters. It is likely that such an object is contained in the offline vertex candidate, such that the 'TOS'-requirement is met. An overview of the selection criteria in the L0, HLT1 trigger lines used for this analysis is given in Table 4.5.

All L0 lines except the L0DiMuon line will be taken into account for the final selection, since L0DiMuon adds a negligible amount to the total efficiency, as shown in Table 4.1. The combined efficiency of these lines on offline selected candidates in HV10_M35 is $88.9 \pm 0.3\%$.

Contrary to the L0, many of the HLT1 lines in Table 4.2 do not give a significant contribution to the efficiency on the sample of offline selected signal events. The lines with the highest HLT1 efficiencies are Hlt1TrackAllL0, Hlt1TrackPhoton and Hlt1TrackMuon; the other lines are discarded. The combined efficiency of these three lines on offline selected candidates (after applying L0) in HV10_M35 is $94.0 \pm 0.3\%$.

### 4.2 Trigger HLT2 selection

In the HLT2, a selection is made with a combination of 'displaced vertex' lines (designed specifically for long-lived heavy particles) and 'topological' lines (designed to trigger on exclusively reconstructed $B$ decays through 2-, 3- or 4-track vertices). The latter lines are included to increase the efficiency on signal candidates with a short flight distance and a low mass. This section describes the efficiencies of the HLT2 selection. The 'displaced vertex' and the 'topological' trigger algorithms and selection criteria are discussed in more detail in Sections 4.3 and 4.4.

The two displaced vertex HLT2 lines considered for this analysis are lines: 'Hlt2DVSingleHighMassPS' (for candidates with high mass) and

---

<table>
<thead>
<tr>
<th>Line</th>
<th>Efficiency</th>
<th>Cum. efficiency</th>
</tr>
</thead>
<tbody>
<tr>
<td>L0Hadron</td>
<td>0.716 ± 0.005</td>
<td>0.716 ± 0.005</td>
</tr>
<tr>
<td>L0Electron</td>
<td>0.373 ± 0.005</td>
<td>0.807 ± 0.004</td>
</tr>
<tr>
<td>L0Muon</td>
<td>0.216 ± 0.004</td>
<td>0.859 ± 0.004</td>
</tr>
<tr>
<td>L0Photon</td>
<td>0.215 ± 0.004</td>
<td>0.889 ± 0.003</td>
</tr>
<tr>
<td>L0DiMuon</td>
<td>0.039 ± 0.002</td>
<td>0.889 ± 0.003</td>
</tr>
</tbody>
</table>

*Table 4.1: Efficiency of L0 'TOS' on simulated hidden valley (HV10_M35) candidates that pass the full selection without the trigger requirements. The third column shows the cumulative efficiency after adding the line ('OR' with the preceding lines). The lines are sorted by their contribution to the cumulative efficiency. The lines printed in bold are used for this analysis.*
### Table 4.2: Efficiency of HLT1 'TOS' on simulated hidden valley (HV10_M35) candidates that pass the offline selection and the L0 selection. The third column shows the cumulative efficiency after adding the line ('OR' with the preceding lines). The lines are sorted by their contribution to the cumulative efficiency. The lines printed in bold are used for this analysis. *) The Hlt1SingleMuonNoIP was prescaled in the data.

<table>
<thead>
<tr>
<th>Line</th>
<th>Efficiency</th>
<th>Cum. efficiency</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hlt1TrackAllL0</td>
<td>0.909 ± 0.003</td>
<td>0.909 ± 0.003</td>
</tr>
<tr>
<td>Hlt1TrackPhoton</td>
<td>0.449 ± 0.006</td>
<td>0.934 ± 0.003</td>
</tr>
<tr>
<td>Hlt1TrackMuon</td>
<td>0.178 ± 0.004</td>
<td>0.940 ± 0.003</td>
</tr>
<tr>
<td>Hlt1SingleElectronNoIP</td>
<td>0.021 ± 0.002</td>
<td>0.940 ± 0.003</td>
</tr>
<tr>
<td>Hlt1SingleMuonHighPT</td>
<td>0.075 ± 0.003</td>
<td>0.941 ± 0.003</td>
</tr>
<tr>
<td>Hlt1DiProton</td>
<td>0.082 ± 0.003</td>
<td>0.941 ± 0.003</td>
</tr>
<tr>
<td>Hlt1DiMuonLowMass</td>
<td>0.014 ± 0.001</td>
<td>0.941 ± 0.003</td>
</tr>
<tr>
<td>Hlt1DiMuonHighMass</td>
<td>0.006 ± 0.001</td>
<td>0.941 ± 0.003</td>
</tr>
<tr>
<td>Hlt1SingleMuonNoIP</td>
<td>0.001 ± 0.000</td>
<td>0.941 ± 0.003</td>
</tr>
<tr>
<td>Hlt1DiProtonLowMult</td>
<td>0.000 ± 0.000</td>
<td>0.941 ± 0.003</td>
</tr>
</tbody>
</table>

'Hlt2DVSingleHighFDPS' (for candidates with high flight distance). In spite of the 'PS' (postscale) appendix in the names, they were not postscaled for the 2011 data taking, since the retention was well within the required boundaries. The selection efficiencies of these two displaced vertex lines on offline selected candidates are compared with those of the topological trigger lines in Table 4.3. Note that the trigger requirement on the offline candidate is 'TOS' ('triggered-on-signal') for the topological lines, and 'TUS' ('triggered-using-signal') for the displaced vertex lines. The 'TOS' selection requires a better overlap between the triggered object and the offline candidate than the 'TUS' selection.

For the topological lines, the reasoning behind the choice for 'TOS' is similar as for the L0 and HLT1 lines. Since the triggered object consists of only 2, 3 or 4 tracks, it is likely that those tracks are contained in the offline vertex candidate. The displaced vertex trigger objects on the other hand are vertices with a large track multiplicity. Slight differences in input tracks and vertex reconstruction in the online and offline environments cause a reduced overlap between the trigger object and the offline vertex. The 'TUS' requirement takes this into account by selecting the event even if only part of the on- and offline vertex tracks overlap. It turns out that the topological lines with a multivariate selection ('BBDT') are more efficient than the 'simple' lines. The final HLT2 selection requires events to pass either one of the lines printed in bold in Table 4.3. The combined efficiency of these lines on offline selected candidates (after applying the L0 and HLT1) in HV10_M35 is 92.4 ± 0.3%.

The topological lines and the displaced vertex lines are complementary. The selection efficiency of the different HLT2 triggers for various signal regions is illustrated in Fig. 4.2. Figures 4.2a and 4.2b show the efficiency as a function the reconstructed...
Figure 4.2: Selection efficiency of different HLT2 trigger lines as a function of offline reconstructed vertex mass (a), (b) and as a function of $R_{xy}$ (c), (d) on simulated hidden valley signal with a variety of lifetimes and masses: (left) on all the simulated candidates without L0 or HLT1 applied, and (right) on the offline selected dijet candidates with L0 and HLT1 applied.
Table 4.3: Efficiency of HLT2 Topo 'TOS' or DisplacedVertex 'TUS' lines on simulated hidden valley (HV10_M35) candidates that pass the offline, L0 and HLT1 selection. The third column shows the cumulative efficiency after adding the line ('OR' with the preceding lines). The lines are sorted by their contribution to the cumulative efficiency. 'PS' stands for postscale, which has not been applied in 2011 data. The lines printed in bold are used for this analysis.

<table>
<thead>
<tr>
<th>Line</th>
<th>Efficiency</th>
<th>Cum. efficiency</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hlt2Topo3BodyBBDT</td>
<td>0.687 ± 0.005</td>
<td>0.687 ± 0.005</td>
</tr>
<tr>
<td>Hlt2DVSingleHighFDPS</td>
<td>0.520 ± 0.006</td>
<td>0.853 ± 0.004</td>
</tr>
<tr>
<td>Hlt2Topo2BodyBBDT</td>
<td>0.648 ± 0.006</td>
<td>0.893 ± 0.004</td>
</tr>
<tr>
<td>Hlt2DVSingleHighMassPS</td>
<td>0.395 ± 0.006</td>
<td>0.915 ± 0.003</td>
</tr>
<tr>
<td>Hlt2Topo4BodyBBDT</td>
<td>0.604 ± 0.006</td>
<td>0.924 ± 0.003</td>
</tr>
<tr>
<td>Hlt2Topo2BodySimple</td>
<td>0.573 ± 0.006</td>
<td>0.931 ± 0.003</td>
</tr>
<tr>
<td>Hlt2Topo4BodySimple</td>
<td>0.416 ± 0.006</td>
<td>0.932 ± 0.003</td>
</tr>
<tr>
<td>Hlt2Topo3BodySimple</td>
<td>0.427 ± 0.006</td>
<td>0.932 ± 0.003</td>
</tr>
</tbody>
</table>

vertex mass. The left figure shows the HLT2 selection efficiency on all the generated events, whereas the right figure shows the efficiency on events that passed the full offline selection (including L0 and HLT1). Note that the mass shown here is not the dijet mass, but the mass of the vertex itself, computed from the charged tracks in the vertex (the vertex mass and the dijet mass are compared in Fig. 4.6 in Section 4.3.3). The HighMass line is more efficient for higher masses, whereas both the HighFD and the topological lines lose efficiency in that region. All three categories contribute significantly to the total efficiency.

The selection efficiency of the HLT2 triggers as a function of the radial distance parameter $R_{xy}$ is shown in Figs. 4.2c and 4.2d. The drop in efficiency between 5-12 mm is caused by a material veto, which is only applied in the HighFD and HighMass lines. Both at low and at high radius, the HighFD line selects candidates that were disregarded by the topological lines. The efficiency of the various HLT2 lines depends on the generated signal mass and lifetime. For example, the topological lines are more efficient on low-mass $\pi^0$ samples, whereas the HighMass displaced vertex line is more efficient on samples with a high $\pi^0$ mass.

To summarise the number of events that survive the trigger selection in the 2011 data set, Table 4.4 gives the event yields of the HLT2 trigger lines, counting only events that pass any of the offline displaced vertex stripping selection.

The offline stripping selection is described in Section 4.5. An overview of the selection criteria in the trigger (L0, HLT1 and HLT2) is given in Table 4.5. The HLT2 selection is discussed in detail in the following sections.
### Table 4.4: Number of candidates passing the trigger on 2011 data, consisting of 0.62 fb⁻¹. Only candidates selected by the displaced vertex offline stripping are considered. The ‘L0’ and ‘HLT1’ stages only consider the trigger lines used for this analysis. The lines printed in bold are used for this analysis.

<table>
<thead>
<tr>
<th>Trigger Line</th>
<th>N candidates (x Stripping)</th>
</tr>
</thead>
<tbody>
<tr>
<td>EW stream</td>
<td>14,656,953</td>
</tr>
<tr>
<td>L0</td>
<td>13,668,199</td>
</tr>
<tr>
<td>HLT1 (xL0)</td>
<td>13,036,318</td>
</tr>
<tr>
<td>Hlt2DVSingleHighFDPS TUS (xL0xHLT1)</td>
<td>2,065,703</td>
</tr>
<tr>
<td>Hlt2DVSingleHighMassPS TUS (xL0xHLT1)</td>
<td>666,268</td>
</tr>
<tr>
<td>Hlt2DVDouble TUS (xL0xHLT1)</td>
<td>963,957</td>
</tr>
<tr>
<td>Hlt2DVSingleFD TUS (xL0xHLT1)</td>
<td>1,428,612</td>
</tr>
<tr>
<td>Hlt2DVSingleHighMass TUS (xL0xHLT1)</td>
<td>592,222</td>
</tr>
<tr>
<td>Hlt2DVSingleDown TUS (xL0xHLT1)</td>
<td>14,185</td>
</tr>
<tr>
<td>Hlt2Topo2BodyBBDT TOS (xL0xHLT1)</td>
<td>2,022,104</td>
</tr>
<tr>
<td>Hlt2Topo3BodyBBDT TOS (xL0xHLT1)</td>
<td>1,832,392</td>
</tr>
<tr>
<td>Hlt2Topo4BodyBBDT TOS (xL0xHLT1)</td>
<td>1,487,560</td>
</tr>
<tr>
<td><strong>Total for DV L0xHLT1xHLT2</strong></td>
<td><strong>4,872,722</strong></td>
</tr>
</tbody>
</table>

#### 4.3 Displaced vertices HLT2 trigger lines

The reconstruction in the HLT2 displaced vertex trigger is described in Section 4.3.1. It only uses tracks with a VELO segment (VELO, upstream and long tracks). Because this VELO algorithm severely limits the acceptance for long-lived particles, another algorithm to reconstruct vertices from downstream tracks has been developed for online usage, which is described in Section 4.3.2. Ideally, an algorithm using a combination of both downstream and VELO tracks is to be developed. Due to the fact that the HLT1 requires a VELO segment in the tracks, the efficiency gain of these downstream lines is small and their output has not been used for the analysis of 2011 data. Section 4.3.3 describes the selection criteria applied in the various displaced vertex lines that were developed for this analysis.

#### 4.3.1 Displaced vertex algorithms for VELO vertices

The algorithm sequence to reconstruct displaced vertices in the HLT2 is the same as the default one for primary vertices, be it with an adjusted tuning. A detailed description of the algorithms can be found in [95, pp. 116-123]. A summary is given here, starting with the primary vertex (PV) algorithm. The vertex reconstruction is split up into a seeding and a fitting step, which are repeated until no new vertices are found in the event.

A first algorithm called PVSeed3DToo1[96] creates vertex seeds by combining
Trigger and Vertex Reconstruction

<table>
<thead>
<tr>
<th>Trigger</th>
<th>Selection</th>
</tr>
</thead>
<tbody>
<tr>
<td>L0Hadron</td>
<td></td>
</tr>
<tr>
<td>L0Electron</td>
<td></td>
</tr>
<tr>
<td>L0Photon</td>
<td></td>
</tr>
<tr>
<td>L0Muon</td>
<td>muon $p_T &gt; 1.48$ GeV, $0 &lt;$ SPD hits $&lt; 600$</td>
</tr>
<tr>
<td>Hlt1TrackAllL0</td>
<td></td>
</tr>
<tr>
<td></td>
<td>$IP &gt; 0.1$ mm, $IP\chi^2 &gt; 16$, $p_T &gt; 1.7$ GeV, $p &gt; 10$ GeV, $\chi^2/ndf &lt; 2.5$,</td>
</tr>
<tr>
<td></td>
<td>VELO hits $&gt; 9$, missed VELO hits $&lt; 3$, OT+IT×2 hits $&gt; 16$</td>
</tr>
<tr>
<td>Hlt1TrackPhoton</td>
<td></td>
</tr>
<tr>
<td></td>
<td>$IP &gt; 0.1$ mm, $IP\chi^2 &gt; 16$, $p_T &gt; 1.2$ GeV, $p &gt; 6$ GeV, $\chi^2/ndf &lt; 2.5$,</td>
</tr>
<tr>
<td></td>
<td>VELO hits $&gt; 6$, missed VELO hits $&lt; 3$, OT+IT×2 hits $&gt; 15$</td>
</tr>
<tr>
<td>Hlt1TrackMuon</td>
<td>Passing L0Muon, muon track with:</td>
</tr>
<tr>
<td></td>
<td>$IP &gt; 0.1$ mm, $IP\chi^2 &gt; 16$, $p_T &gt; 1$ GeV, $p &gt; 8$ GeV, $\chi^2/ndf &lt; 2$</td>
</tr>
<tr>
<td>Hlt2DVSingleHighMassPS</td>
<td></td>
</tr>
<tr>
<td>Hlt2DVSingleHighFDPS</td>
<td></td>
</tr>
<tr>
<td>Hlt2Topo2BodyBBDT</td>
<td></td>
</tr>
<tr>
<td>Hlt2Topo3BodyBBDT</td>
<td></td>
</tr>
<tr>
<td>Hlt2Topo4BodyBBDT</td>
<td>BBDT $&gt; 0.3$</td>
</tr>
</tbody>
</table>

Table 4.5: Summary of the cuts made in the three different trigger stages for the displaced vertex analysis. In each stage, the 'OR' of the listed lines determines the trigger decision. The input variables to the BBDT are: $\sum p_T$, $p_T^{\text{min}}$, $m$, $m_{\text{corrected}}$, distance of closest approach, candidate IP$\chi^2$ and flight distance $\chi^2$ [84].
pairs of VELO tracks with a distance of closest approach smaller than $\text{TrackPairMaxDistance} = 0.3$ mm. If there are at least four tracks ($\text{MinCloseTracks} = 4$) within a distance of $z\text{MaxSpread} = 3$ mm to the seed, the seed is accepted for further processing.

The seed algorithm is followed by a vertex algorithm, called $\text{LSAdaptPV3DFitter}$. In the online environment the PV algorithm runs on the tracks found by the VELO pattern recognition. Because these tracks have unknown momenta, scattering corrections cannot be reliably estimated. For this reason, as well as to save CPU time, the tracks are not fitted with the standard track fit, leading to a somewhat worse track parameter resolution and relatively poorly known parameter uncertainties. Consequently, the online $\text{LSAdaptPV3DFitter}$ makes no use of estimated track parameter uncertainties and uses a simpler, linear track extrapolation to calculate distances between the tracks. The gain in CPU timing in the online environment is about 20%. The linear track extrapolation in the online algorithm results in larger errors on the extrapolated tracks, such that more tracks are assigned to a vertex than in the offline environment.

Starting from the vertex seed position, the $\text{LSAdaptPV3DFitter}$ combines the seed with all nearby tracks in an adaptive least squares fit. The weight of each track is $\left(1 - \frac{\chi^2}{c^2}\right)$, where $c$ (Tukey's constant) is set to 3. If the track is of bad quality and $\chi^2 > c^2$, the weight is set to zero, thereby excluding the track from the fit. Tracks with a weight larger than $\text{minTrackWeight} = 1 \times 10^{-5}$ are selected. The fit requires at least $\text{minTr} = 5$ tracks to pass this requirement on the maximum impact parameter $\chi^2$ to the seed. In addition the online algorithm selects input tracks by their impact parameter to the seed vertex; if the impact parameter is larger than 2 mm, the track is removed from the selection before the fit. The minimisation is iterated, and it is checked for convergence by requiring that the maximum difference between the $z$ coordinates of the current and previous iteration is smaller than $\text{maxDeltaZ} = 0.1$ μm. After a successful fit, tracks with a $\chi^2$ to the seed below $\text{trackMaxChi2Remove} = 25$ are removed from further vertex searches.

To reconstruct displaced vertices, the same $\text{PVSeed3DTool}$ and $\text{LSAdaptPV3DFitter}$ algorithms are used. The parameters of those algorithms are adjusted, as described in Appendix A.1. The different configurations for the online primary and displaced vertex reconstruction are summarised in Table 5 in the appendix.

### 4.3.2 Displaced vertex algorithms for downstream vertices

Complementary to the search for displaced vertices in the VELO, a sequence has been developed to look for vertices beyond the VELO volume, using a different vertex algorithm. The input for this vertex reconstruction has to include downstream tracks. The reason that $\text{LSAdaptPV3DFitter}$ is not used here is that it relies on a 'straight-
extrapolation for the tracks, which is a poor approximation in between the VELO and TT detectors. Therefore, the downstream vertex reconstruction uses the primary vertex algorithm developed for the offline environment: LSAAdaptPVFitter.

The first difference between LSAAdaptPVFitter and LSAAdaptPV3DFitter is that the first algorithm can use downstream tracks that are fitted with the standard track fit as input. Secondly, instead of applying a selection on the impact parameter of the tracks to the seed vertex, a selection is made on the maximum \( \chi^2 \) of the track, to remove tracks of bad quality. The minimisation is similar to the one in LSAAdaptPV3DFitter, with the additional requirement that the difference in the \( \chi^2 \) per number of degrees of freedom between iterations is smaller than \( \text{maxDeltaChi2NDoF} = 0.002 \), in case the \( z \) coordinates of the current and previous iteration do not converge.

The LSAAdaptPVFitter algorithm is deployed in the downstream displaced vertex reconstruction of both the trigger and stripping. In order to apply the downstream vertexing in the trigger algorithms, the parameters have been tuned following the procedure for the VELO vertices. The tuning for the downstream seed and vertex algorithms (described in Appendix A.2) is slightly different than for VELO vertices.

The 'downstream strategy' improves the efficiencies for larger decay times in both the online and offline environments. For the offline environment, this is illustrated by the difference between the black and the red points in Fig. 4.3. However, due to the VELO segment requirement in the HLT1, the gain in efficiency is currently negligible. The requirement of a HLT1'TOS' decision (not shown in the figure), discards all the candidates from the downstream line. The blue points in Fig. 4.3 show the efficiency when requiring HLT1 without 'TOS'. As explained in Section 4.1, the 'TOS' requirement is needed to allow for the calculation of the trigger efficiency and its uncertainties. The VELO requirement has been removed from the HLT1 selection for 2012 data, such that the downstream line will be of more use for future analyses.

### 4.3.3 Displaced vertex HLT2 lines

After the vertex reconstruction algorithms are completed, a collection of vertices is available for each event. In the next step, a selection is applied using the properties of those vertices to save only the events containing long-lived particle candidates. To optimise the signal efficiency within a limited output rate, various trigger lines are defined that focus on a different phase space of the signal; either candidates with high mass and intermediate flight distance, or candidates with high flight distance and intermediate mass.

**Minimal requirements on displaced vertices**

The vertex algorithm reconstructs displaced vertices as well as primary interaction vertices. In order to be tagged as a displaced vertex (DV), the trigger object should
Figure 4.3: Fraction of displaced vertex candidates found at the stripping level with the 'VELO' and 'downstream' strategies, in a simulated hidden valley signal sample with $m_{\pi^0} = 35$ GeV and $\tau_{\pi^0} = 100$ ps. Contrary to the analysis selection, the HLT 1 here requires only a positive decision, without 'TOS'.

have:

- no backward tracks;
- a radial displacement $R_{xy} > 0.4$ mm from the beamline;
- at least 4 tracks with VELO hits;
- an invariant mass of at least 3 GeV and scalar sum of transverse momentum of the tracks of at least 3 GeV;

Furthermore, there has to be at least one primary vertex in the event, upstream of the DV, with at least ten tracks, including one backward and one forward track, $R_{xy} < 0.4$ mm and $|z| < 400$ mm. These selection criteria are discussed in detail below.

The first selection criteria assure that the trigger object is not a primary vertex. It has to have a radial distance to the beam axis $R_{xy} > 0.4$ mm, to select only vertices that are not reconstructed in the primary interaction region, which extends up to approximately 0.3 mm, as shown in Fig. 4.4. The selection criterion on the radius is varied between 0.4 mm for high mass candidates and 4.0 mm for candidates with a lower mass.

The position of the beamline is different for each LHC fill. The VELO detector halves move towards the interaction region when stable collisions take place, and are positioned such that the interactions always take place exactly in the centre, by reconstructing the beam collisions. The VELO resolvers or stepper motors hold the information of the beamline position in every single fill [97]. A tool has been developed
The event should contain at least one primary vertex upstream of the displaced vertex, to ensure that the displaced vertex is not created by an interaction between the beam and the gas in the detector. This background is discussed in more detail in Section 7.3.

After the confirmation that the vertex is displaced, its track multiplicity is the next selection criterion. The track multiplicity of the candidates in simulated signal depends on the type of signal model. For example, for a $\pi^0$ decaying to light quarks, the track multiplicity of the displaced vertex is high, whereas a $\pi^0$ decaying into $b\bar{b}$ has a lower displaced vertex multiplicity and more tracks in tertiary vertices. Therefore the selection on the track multiplicity is made as inclusive as possible within the allowed output rate. Depending on the trigger line, a minimum of 4 to 6 tracks is required.

The subsequent selection concerns the minimal mass of the candidates. At the trigger level, jet reconstruction cannot be performed, because it is too CPU-intensive. The mass estimate that is used instead is the invariant mass calculated from the reconstructed charged tracks in the vertex. Since this calculation does not include tertiary decay vertices, and no neutral particles, the mean of this variable will be significantly lower than the dijet mass. A comparison of the invariant mass from the vertex tracks
Figure 4.5: Position of the beamspot as determined from the VELO resolvers, for the 2011 data set used for this analysis. Subsequent fills are connected by a line.

Figure 4.6: Invariant mass of candidates that passed the trigger and stripping selections, calculated using the tracks in the vertex (black), using the jets for candidates with one jet (red), and using the jets for candidates for candidates with two jets (blue). (a) shows simulated hidden valley signal, and (b) shows 2011 data.
and the dijet mass is given in Fig. 4.6. At this selection stage, one can assume that the data mainly consists of background events. Comparing the vertex mass at trigger level (black histograms) between simulated signal (Fig. 4.6a) and data (Fig. 4.6b), shows that it is difficult to distinguish between signal and background using the vertex mass. The reconstructed dijet mass however (blue histograms) allows easier separation between signal and background. As long as the jet reconstruction is not fast enough to be used online, the dijet mass can currently only be used in the offline analysis. Therefore the online HLT2 selection only uses the invariant mass of the vertex, requiring at least 3 to 10 GeV, depending on the trigger line.

One of the main backgrounds outside the primary interaction region is the interaction of particles with detector material. Therefore, a material veto (MV) is applied for the trigger lines that have a good reconstructed vertex position resolution; namely all the lines with vertices constructed from tracks with a VELO segment. The implementation of the material veto is described in detail in Section 4.6.

**Displaced vertex lines**

Various HLT2 lines have been developed, each of which focuses on a different region in phase space: a high-mass particle with low flight distance or vice versa, a particle decaying outside the VELO volume, events with at least two displaced vertices, etc. A detailed overview of the selections applied in the displaced vertex HLT2 lines is shown in Table 4.6. The top row in the table ('Hlt2RV2P') shows the minimal requirements on all the candidates, described in the previous paragraph. The two displaced vertex lines used for the analysis described in this thesis are 'Hlt2DVSingleHighMassPS' and 'Hlt2DVSingleHighFDPS', printed in bold in the table. These lines can optionally be postscaled, to reduce the data output rate if necessary. In 2011, these two lines were run without a postscale, which makes them the loosest selections. The other lines ('Hlt2DVSingleHighMass', 'Hlt2DVSingleHighFD' and 'Hlt2DVSingleMedium') each cover a different region in mass and lifetime. They will most likely be used for future analyses, e.g. when running at increased luminosity, when the output rate of the 'PS' lines becomes so high that the postscales have to be applied.

The so-called 'efficiency lines' are postscaled, which means that only a small random portion of the selected sample is stored. These lines apply only a minimal selection, and are used to study reconstruction and selection efficiencies. The line that does not have the material veto applied (Hlt2DVSinglePS) is used to study the efficiency of this veto.

The remaining two lines in Table 4.6 are not used for this analysis. First, the 'Hlt2DVDouble' line selects events with at least two \( \pi^0 \) candidates per event, which is necessary for the search of Higgs-like bosons decaying into two long-lived exotic particles [56]. Second, the 'Hlt2DVSingleDown' line uses the vertices reconstructed from downstream tracks, as described in Section 4.3.2. Since the timing is one of
the most important limitations in the trigger, it is not possible to execute the downstream algorithms for each event. In order to reduce the number of events considered for downstream vertexing, only those events that already contain a displaced vertex constructed from VELO tracks are taken into account. Although this reduces the acceptance for particles with high lifetimes, it is currently the only way to reduce the input rate while still retaining signal-like events. The downstream line is not used yet because the efficiency gain is almost completely cancelled by the HLT1 selection. This issue has been solved for the 2012 data sample.

The total output rate of the HLT2 displaced vertex lines on 2011 data is 15-20 Hz. To sum up, the two displaced vertex lines used to select signal-like events are 'Hlt2DVSingleHighMassPS', selecting candidates with a high mass and an intermediate flight distance, and 'Hlt2DVSingleHighFDPS', selecting candidates with a high flight distance and an intermediate mass.

### 4.4 Topological HLT2 trigger lines

For the purpose of selecting $B$ decays, a set of inclusive trigger lines exists in the HLT2 that are designed to reconstruct (part of) any type of $B$ decay topology; hence the name 'topological' triggers [98]. Although they are tuned for relatively short lifetimes, these

<table>
<thead>
<tr>
<th>Line</th>
<th>$R_{xy}$ [mm]</th>
<th>$m &gt; [GeV]$</th>
<th>$m &gt; 2nd$</th>
<th>$\sum p_T &gt; [GeV]$</th>
<th>$N_{tr} \geq$</th>
<th>material veto</th>
<th>PS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Preselection</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hlt2RV2P</td>
<td>0.4</td>
<td>3</td>
<td>-</td>
<td>3</td>
<td>4</td>
<td>no</td>
<td>-</td>
</tr>
</tbody>
</table>

**Table 4.6:** Definition of DisplacedVertex HLT2 lines implemented in June 2011. 'PS' stands for postscale. The postscale between brackets were not applied on 2011 data. Values in the line Hlt2DVSingleDown marked with * apply only to the VELO candidate. The lines printed in bold are used for this analysis.
triggers are complementary to the displaced vertex HLT2 lines. The topological lines are efficient for the long-lived \( \pi^0 \) signal, firstly because they are inclusive by reconstructing a few tracks of the decay, and secondly because their allowed output rate is not as limited as the rate of the dedicated displaced vertex lines. There are six topological lines, three of which use a multivariate selection. They are divided into 2-, 3- and 4-track vertex lines. The efficiencies of these lines on signal were listed in Table 4.3.

The topological lines can reconstruct and trigger any \( B \) decay with at least two charged daughter particles. The input consists of all tracks with momentum \( p > 5 \) GeV and transverse momentum \( p_T > 500 \) MeV that pass a selection on the track quality to eliminate ghost tracks: track \( \chi^2/dof < 5 \), and an impact parameter \( IP\chi^2 > 16 \) to any primary vertex, to eliminate prompt tracks.

The candidates are made by forming a track pair and adding particles to it one by one, thereby building two- three- or four-body decay candidates. The prerequisite to add a track is that its distance of closest approach to the track pair must be smaller than 0.15 mm. Subsequently, the \( n \)-body candidates have to pass the selection criteria described below.

Typically, only part of the \( B \) decay is reconstructed, and it is therefore not desirable to cut on the invariant mass of the objects. Instead, a selection is made using a variable called ‘corrected mass’, which is illustrated in Fig. 4.7. The corrected mass is calculated by adding a correction to the mass of the candidate due to missing tracks or neutral particles.

![Figure 4.7: Schematic representation of a long-lived particle produced at a primary vertex (PV), with the momentum vector of the candidate \( \vec{p} \), and the vector \( \hat{d} \) pointing from the PV to the displaced vertex (DV).](image)

The unit vector pointing from the primary vertex to the displaced vertex is indicated with \( \hat{d} \). Suppose that the pointing to the PV is wrong because there is missing momentum. The missing three-momentum is denoted by \( \vec{q} \). The observed momentum vector \( \vec{p} \) and the vector \( \hat{d} \) can be made parallel by adding a transverse momentum \( \vec{p}_\perp \). The observed momentum \( \vec{p} \) can be decomposed in a component along and perpendicular to \( \hat{d} \) as follows:

\[
\vec{p}_\parallel = (\vec{p} \cdot \hat{d})\hat{d} \quad \quad \vec{p}_\perp = \vec{p} - \vec{p}_\parallel
\]
such that the missing transverse momentum:

\[ \vec{p}_\perp = \vec{p} - (\vec{p} \cdot \hat{d})\hat{d} \]  

(4.2)

The missing momentum is \( \vec{q}_\perp = -\vec{p}_\perp \), such that the transverse momentum in the sum \( q + p \) will cancel. The observed three-momentum \( \vec{p} \) and its mass \( m \) constitute the observed energy \( \sqrt{m^2 + p^2} \). The minimal missing mass can be determined as follows. First by writing the four-vector as:

\[
\begin{pmatrix}
\sqrt{p^2 + m^2} \\
p_\parallel \\
\vec{p}_\perp
\end{pmatrix}
\]  

(4.3)

where the first coordinate is the energy, \( p_\parallel \) is the momentum along \( \hat{d} \), and \( \vec{p}_\perp \) represents the two coordinates perpendicular to it. Next, the system can be boosted along \( \hat{d} \) with a velocity such that \( p_\parallel \) vanishes, while the perpendicular momenta do not change. Assuming massless missing particles, in the boosted system the total four-vector can be written as:

\[
p' + q' = \begin{pmatrix}
\sqrt{p'^2 + m^2} \\
0 \\
\vec{p}_\perp
\end{pmatrix} + \begin{pmatrix}
\sqrt{q'^2 + q'^2} \\
q'_\parallel \\
q'_\perp
\end{pmatrix} = \begin{pmatrix}
\sqrt{p^2 + m^2 + q'^2 + p'^2} \\
qu'\parallel \\
qu'\perp
\end{pmatrix}
\]  

(4.4)

The minimum mass correction is obtained when \( q'_\parallel = 0 \), such that the corrected mass becomes [98]:

\[ m_{\text{corr}} = \sqrt{m^2 + p^2 + \vec{p}_\perp} \]  

(4.5)

The selection criterion on the corrected mass in the topological lines is: \( 4 < m_{\text{corrected}} < 7 \) GeV. For the \( \pi^0 \) candidates, which typically have a higher mass and higher track multiplicity than \( B \) vertices, this criterion is not ideal, since the missing momentum is likely to be larger than for \( B \) candidates. Exploiting the high momentum of the \( B \) daughters, further cuts are made on the highest \( p_T \) track (\( > 1.5 \) GeV) and on the \( \sum p_T \) of the tracks: 4 GeV, 4.25 GeV and 4.5 GeV for 2-, 3- and 4-body decays, respectively. Furthermore, at least one track of good quality (track \( \chi^2/\text{dof} < 3 \)) is required.

Further selections are made on the flight distance (flight distance \( \chi^2 > 64 \)) and the total impact parameter \( \chi^2 \) of all tracks to the primary vertex (\( IP \chi^2 > 100, 150, 200 \) for 2-, 3- and 4-body lines, respectively), in order to exploit the displacement of the \( B \)
decay vertex. Finally, pollution from $D$ decays is avoided by requiring that the $n$-body vertex has either a high invariant mass ($m > 2.5\,\text{GeV}$) or a large impact parameter to the primary vertex ($IP\chi^2 > 16$).

The requirements mentioned previously apply to the so-called 'simple' topological lines. However, additional lines use multivariate selection techniques [85]. A boosted decision tree (BDT) is run on the $n$-body candidates surviving a preselection. Such a decision tree is a multivariate classifier performing repetitive one-dimensional splits of the data. The boosted decision tree is 'trained' to learn the difference between signal and background on simulated samples as well as on data. The figure of merit used for determining the splitting point is the signal significance, namely the ratio of signal over background in various $B$-decays [99]. The output of the boosted decision tree is a single variable that separates signal from background. The input variables for the BDT are the ones also used for the simple selection: the minimum transverse momentum $p_T^{\text{min}}$ of each track, the distance of closest approach between the tracks, the $\sum p_T$ of the tracks, the mass of the candidate, its corrected mass $m_{\text{corrected}}$, the $IP\chi^2$ of the candidate with the PV, and its flight distance $\chi^2$. The BDT is trained on different $B$ signal samples, to make the signal profile as inclusive as possible. The final selection criterion on the BDT output variable depends on the multiplicity of the topological candidate.

4.5 Stripping selection

The first stage of the offline reconstruction and selection is called 'stripping'. It is similar to the online trigger selection procedure, but with a less constrained CPU requirement, allowing for more precise reconstruction algorithms. The algorithms and selection lines are presented in the following paragraphs. Although both the trigger and the stripping selections rely on the presence of a displaced vertex candidate, the selection criteria vary between the two.

4.5.1 Stripping vertex algorithm

The offline primary vertex algorithm is different than the one used in the online environment. The primary vertices in the stripping are reconstructed with the LSAdaptPVFitter algorithm, which is described in the context of downstream vertex reconstruction in Section 4.3.2. The only difference with that description is that the input to the primary vertex reconstruction consists only of tracks with a VELO segment.

The algorithm to reconstruct $displaced$ vertices in the stripping, however, is identical to the displaced vertex algorithm used in the trigger (LSAdaptPV3DFitter), although the input tracks and the parameter tuning are different. More specifically, the parameters for the seeding algorithm are loosened in order to improve the seed
Table 4.7: Definition of the displaced vertex stripping physics lines in stripping 17. ‘MV’ stands for material veto. The lines printed in bold are used for this analysis.

Table 4.8: Definition of the efficiency-related displaced vertex stripping lines in stripping 17. ‘MV’ stands for material veto, ‘PS’ for postscale.

finding efficiency. The downstream displaced vertex reconstruction in the stripping is identical to the one performed in the trigger, which is described in Section 4.3.2. The only exception are the input tracks, which in the offline environment include all track types. The parameter tuning of the offline stripping algorithms can be found in Table 5 in Appendix A.

4.5.2 Stripping lines

The long-lived particle candidates are selected with the stripping lines ‘StrippingDVSingleHighMass’, ‘StrippingDVSingleMedium’ and ‘StrippingDVSingleHighFD’. They are printed in bold in Table 4.7, which lists the selection criteria for the main displaced vertex stripping lines. The offline vertex reconstruction includes primary interaction vertices, as it did in the trigger. To eliminate these, a selection similar to the one described in Section 4.3.3 is applied to all lines. The grouping of the different physics analysis lines in the stripping is similar to the trigger set-up, as shown in Table 4.7.

Furthermore, there are stripping lines to obtain control samples, labelled ‘efficiency lines’, as listed in Table 4.8.

The ‘StrippingDVSinglePS’ stripping line has the same settings as the Hlt2DVSinglePS trigger line, with an additional postscale factor, thereby further limiting the output rate. This line provides a sample with a loose selection, which can be used for background Monte Carlo studies.

The ‘StrippingDVSingleHLTPS’ and ‘StrippingDVDoubleHLTPS’ lines do not re-
construct the event, which saves CPU time. Instead, they store every event passing the corresponding postscaled trigger lines indicated in Table 4.8. An offline reconstruction is applied afterwards on these events. The 'StrippingDVSingleHLP' line for example contains events that pass trigger lines with and without the material veto applied, which be used to compute the number of events rejected by the material veto.

'StrippingDVJPSiHLT' is designed to determine trigger efficiencies and systematics. It selects events that were triggered in HLT2 independently of the displaced vertex selections, namely by the $J/\psi$ di-muon line. It then applies the stripping displaced vertex reconstruction on this independent sample.

Events passing any of the displaced vertex lines are collected in the electroweak data stream (see Section 3.3). Table 4.9 shows the number of events in the subset of the 2011 electroweak stream that pass each of the stripping lines. A comparison of the third and the rightmost column shows that there is limited overlap between the displaced vertex stripping lines and the displaced vertex trigger lines. This is mainly because the tracks that are used as input to the on- and offline vertexing algorithms are not identical, leading to differences in the reconstructed mass of the vertices in the online and offline environment. The HLT2 only reconstructs tracks using an algorithm

<table>
<thead>
<tr>
<th>Offline selection line</th>
<th>N candidates</th>
<th>N candidates</th>
<th>N candidates</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>(xL0xHT1 (DV</td>
<td></td>
</tr>
<tr>
<td>Stripping all (non-PS)</td>
<td>14,422,113</td>
<td>4,731,724</td>
<td></td>
</tr>
<tr>
<td>StrippingDVSingleHighMass x Hlt2DVHighMassPS</td>
<td>3,228,663</td>
<td>1,599,207</td>
<td>419,765</td>
</tr>
<tr>
<td>StrippingDVSingleMedium</td>
<td>3,054,695</td>
<td>1,844,031</td>
<td></td>
</tr>
<tr>
<td>StrippingDVSingleHighFD x Hlt2DVHighFDPS</td>
<td>3,624,340</td>
<td>2,046,191</td>
<td>1,234,803</td>
</tr>
<tr>
<td>StrippingDVSingleDown x Hlt2DVSingleDown</td>
<td>3,127,778</td>
<td>3,186</td>
<td>377</td>
</tr>
<tr>
<td>StrippingDVDouble x Hlt2DVDouble</td>
<td>5,560,695</td>
<td>1,850,547</td>
<td>371,921</td>
</tr>
<tr>
<td>Stripping lines for DV</td>
<td>6,333,548</td>
<td>3,250,006</td>
<td></td>
</tr>
</tbody>
</table>

*Table 4.9: Offline vertex reconstruction event yields on 0.62 fb$^{-1}$ of 2011 data. The third column lists the output of the specified stripping line, after requiring the L0, HLT1, and the HLT2 lines used for this analysis (DV and Topo). The right-most column gives the overlap between the Stripping line and the corresponding HLT2 line. The lines printed in bold are used for this analysis.*
based on VELO track seeds (the so-called 'forward tracking', explained in Section 2.3), whereas the offline tracking implements two different tracking techniques, leading to a higher offline track-finding efficiency. Furthermore, to limit the time used by the online event reconstruction, only tracks with $p > 5$ GeV and $p_T > 0.5$ GeV are selected. The events that do not pass the displaced vertex trigger lines because of the online-offline differences, can partly be recovered by the topological trigger lines, as explained in Section 4.2. One can also see in this table that the 'StrippingDVSingle-Down' stripping line has almost no overlap with the HLT1 trigger, since no downstream tracks are available in HLT1. The events that do pass, are triggered by another object in the event within the VELO acceptance.

To conclude, the long-lived particle analysis uses candidates from the stripping lines 'StrippingDVSingleHighMass', 'StrippingDVSingleMedium' and 'StrippingDVSingleHighFD'. The number of events selected in data by the combination of these three lines is given in the last line of Tab. 4.9, with and without the additional trigger requirements.

### 4.6 Material veto

Most displaced vertices at a radius above approximately 5 mm from the beamline originate from interactions with the detector material, as shown in Fig. 4.8a. The figure illustrates that the interaction vertices can be identified by their position distribution. For example, the VELO sensors and the RF foil can be clearly distinguished in the zoom-in of the $R_{xy}$ versus $z$ distribution.

To eliminate vertices from material interactions, a veto is applied on vertex positions close to material, using a geometrical description of the detector volume. The material veto defines approximately 1-2 mm wide envelopes around the VELO modules and the RF foil. It is described in detail in [56]. The material positions are determined from the detector geometry in the running conditions database, and the veto is implemented separately for the left and right detector modules. Figure 4.8 shows the vertices rejected by such a veto in black, and the ones that pass in red. Since the figure depicts real data, the red points are mostly decaying beauty and charm hadrons, which typically have a flight distance below a few millimeter. The material veto is applied at trigger level, since the material interactions could easily saturate the trigger bandwidth reserved for the displaced vertex analyses.

The material veto applied in the HLT2 and in the stripping is identical. However, since the position of vertices in the trigger and stripping is not the same, the final envelop is still slightly larger than the stripping veto itself. When analysing a large set of data, it was found that the MV still missed a small region of material interactions at large $z$-position, close to the beam axis. Furthermore, the resolution of the vertices at large $z$ is significantly worse than at low $z$. Therefore an extension of the material veto
Figure 4.8: Distribution of displaced vertices in (a) the $R_{xy}z$-plane and (b) the $xz$-plane in LHCb data. The box in (a) shows a zoom where only vertices with $x > 0$ are selected, which shows the separate sensors in the VELO plane, enveloped by the RF-foil. The vertices in red are accepted by the material veto while those in black are rejected. The data has been obtained with a postscaled displaced vertex trigger without the material veto applied.
Another significant contribution to the background comes from events induced by interactions of particles in the beam halo, so-called 'beam-splash' events. The particles possibly originate from an interaction between the beam halo and an object upstream or downstream of the VELO detector, for example a collimator. These events are characterised by a large number of charged particles that traverse one side of the VELO in almost parallel trajectories. The proximity of the hits of these particles leads to an even larger number of fake tracks with very small slopes. These tracks can easily be combined into a vertex at multiple points, leading to fake vertices at a large distance to the beam. Figure 4.10 shows the position of decay vertices in the offline analysis, after the material veto has been applied. Two large clouds of vertices are visible at the negative x-side, where most parallel tracks pass through. The tracks are most likely to create vertices when they cross real tracks at a sizeable angle, for example from the primary interaction region or from material interactions. This explains the diagonal shape of the clouds, which point back to the collision region. Two selection criteria are applied
Figure 4.10: Position of decay vertices in the xz plane after application of the extended material veto. Events in red have been identified as beam-splash events, and will be rejected.

that reject beam-splash events.

First, a selection is applied to the ratio of the number of reconstructed VELO track segments relative to the number of VELO clusters. If there is more than one track reconstructed for every ten clusters, i.e. $N_{\text{tracks}}/N_{\text{clusters}} > 0.1$, the event is rejected. The distribution of this variable is shown in Fig. 4.11a. Although it is efficient for signal, this selection still leaves a considerable amount of background.

Therefore, in addition, a second veto is developed that uses the distribution of VELO hits. The hits associated to beam-splash particles are all close in $\phi$ in the VELO. To exploit this feature, a vector sum can be defined in the transverse plane:

$$D_{\phi} = \sum_{i} \left( \frac{\cos \phi_i}{\sin \phi_i} \right)$$  \hspace{1cm} (4.6)$$

where the sum runs over all hits in the VELO $\phi$-strips, and $\phi_i$ is the direction of the strip. If all hits are in $\phi$-strips close together (in the same $\phi$ direction), they add up to a large $D_{\phi}$. If they are more random, they decrease the value of $D_{\phi}$. The distribution of $D_{\phi}$, both for simulated signal events (green) and data (red and blue) is shown in Figure 4.11b. To reject beam-splash events, the threshold is set at $D_{\phi} < 250$.

Neither the cut on the track-to-hit ratio nor the cut on $D_{\phi}$ has been deployed in the 2011 trigger or stripping. However, they are applied in the offline analysis and will be used for future trigger and stripping selections.
4.8 Characteristics of the reconstructed displaced vertices

The selections made in the trigger and the stripping can be studied on simulated signal samples, but they are not easily visualised in data since the trigger selection is already applied online, before the data is stored. Fig. 4.12 shows the relevant distributions of the MC signal candidates passing the loosest stripping line SinglePS (in red), before any subsequent cuts are made, and before the triggers are applied. The cut-offs due to the SinglePS stripping selection are clearly visible. Fig. 4.12d, depicting the radial distance from the beamline, shows that the data contain many material interactions, for example around $R_{xy} = 6$ mm.

Figure 4.13 shows the same data, after the trigger selection has been applied (passing L0, HLT1 and either topological or displaced vertex HLT2). Since the histograms show events passing a mixture of stripping lines, and since the online and offline reconstructed quantities are not necessarily equal, the various trigger cuts are not all clearly visible in the distributions. However, one can still see the effect of the cut at $R_{xy} > 2$ in the data in Fig. 4.13d. The decrease at $R_{xy} = 5$ is due to the material veto.

4.9 Minimum number of vertex tracks

The displaced vertex candidates created in the stripping consist of at least five tracks. The distribution of the number of tracks in the displaced vertices, shown in Fig. 4.13c,
Figure 4.12: Vertex (a) mass, (b) $\sum p_T$, (c) number of tracks and (d) distance from the beamline of vertices accepted by the StrippingDVSinglePS stripping line in simulated hidden valley signal events and in data. No further trigger requirements are made on the MC, and the material veto is not applied. These histograms are normalised to unit area, and do not represent the full data set used for the analysis.
Figure 4.13: Vertex (a) mass, (b) $\sum p_T$, (c) number of tracks and (d) distance from the beamline of vertices accepted by the StrippingDVSinglePS stripping line in simulated hidden valley signal events and in data. Both are triggered by L0, HLT1, and either the topological or the displaced vertex HLT2 lines, and the material veto has been applied. These histograms are normalised to unit area, and do not represent the full data set used for the analysis.
indicates that at least six tracks can be required in the vertex without much signal loss. This requirement rejects a large fraction of the background and increases the signal over background ratio. A tighter selection on the number of tracks could be motivated with the same argument, as shown in Fig. 4.14. However, in order to keep the analysis sensitive to decays with a relatively small displaced vertex track multiplicity, the selection criterion is to have at least six tracks per vertex.

![Figure 4.14](image)

**Figure 4.14**: Optimisation of the selection on the displaced vertex track multiplicity. (a) selection efficiency versus the value of the selection criterion in simulated hidden valley signal (HV10\_M35) and in data. (b) figure of merit $N_{\text{signal}}/\sqrt{N_{\text{background}}} + 1$ versus the value of the selection criterion.

### 4.10 Outlook

**Trigger selection**

The main limitation of the HLT1 trigger is its inefficiency to select vertices outside the VELO detector acceptance. Firstly, a track must contain at least ten VELO clusters to be accepted for the HLT1. Secondly, the number of VELO planes in which a hit is expected but not observed is required to be smaller than three. This results in a large decrease of the vertex finding efficiency above $R_{xy} > 10$ mm. A dedicated HLT1 displaced vertex line is added in 2012 to circumvent part of this problem. This line selects only those tracks that have a large impact parameter ($IP_{PV} > 1.5$ mm) to the primary vertex. It then creates two-track combinations, of which at least one track has a distance of closest approach to the beamline larger than 2 mm and at least three VELO planes with clusters, of which at least two are consecutive. Only events with
two-track vertices with a distance of closest approach to the beamline above 0.3 mm, and a radial displacement larger than $R_{xy} > 12$ mm, are selected. This development increases the HLT1 efficiency on signal by about 60% for high radii.

The additional HLT1 line still does not solve the efficiency loss due to the requirement of a VELO segment in each HLT1 track. Both in the HLT2 and in the stripping an algorithm is implemented that makes use of downstream tracks, in addition to the standard search for displaced vertices in the VELO. However, this output has not been used since the HLT1 rejects vertices outside the VELO acceptance. A possible solution is to implement a faster and more precise track reconstruction in the detectors downstream of the VELO detector.

In the HLT2, the vertex algorithm reconstructs displaced vertices as well as primary interaction vertices. This is inefficient in terms of timing, and therefore the algorithms written for future data collection only receive input tracks that do not point to any primary vertex ($IP_{PV} > 0.1$ mm), and that do not point in the backward direction. These improvements are propagated to the stripping procedure.

**Stripping selection**

To obtain a high efficiency, the vertex reconstruction algorithms apply loose criteria, allowing a large fraction of the tracks in the $\pi^0$ decay chain to be absorbed into a single vertex. To decrease the systematic uncertainties related to the position of the vertex and to get a more precise estimate of the flight distance of the long-lived particle, some effort has been made to improve the vertex position resolution. Additional tuning of the vertex algorithm to improve this resolution did not produce the desired effect without a notable loss in efficiency.

Several attempts to improve the vertex reconstruction efficiency, such as the use of a different vertex algorithm and a looser parameter tuning, have not yielded any significant increase in efficiency. Both in the stripping and in the trigger selection, the room for improvement is constrained by the maximum output bandwidth.
In the analysis of the 2010 dataset [56], the mass of the long-lived particle candidates was calculated as the invariant mass of the charged particles originating from the displaced vertex. This leads to a mass estimate of signal candidates below the generated mass. Since then, a common jet algorithm has been developed in the LHCb collaboration, and a customised version for jets from displaced vertices was developed in parallel. Although the emphasis of this analysis lies on retrieving a correct mass estimate using the jets, the flavour of the underlying quarks can also be used for selection and identification of the long-lived particles. The candidate selection and mass reconstruction based on the jets are discussed in Chapter 6. This chapter describes the jet input and reconstruction for long-lived particles in LHCb, the optimisation of the jet reconstruction parameters and the selection of good jets.

5.1 Jet reconstruction

A jet is defined as a collimated cluster of energetic particles, produced by the fragmentation of partons (either quarks or gluons). Since partons are not directly observable, the properties of hadron jets are the only available means to relate the observed particles to the initial quarks and gluons. The relations between the particles observed in the detector and the associated parton-level production processes are not easily identified, but jet-finding algorithms have been developed which come close to reconstructing the kinematic properties of the initial partons [89]. To successfully perform the jet reconstruction, one needs to start with feeding the jet algorithm an accurate list of input particles.
5.1.1 Jet input

The selection of particles that serve as input to the jet reconstruction algorithm can be done in several ways. One way is to start from energy clusters in the calorimeters, and to match charged particle tracks to those clusters. This method is used for example in the ATLAS experiment, which benefits from a good calorimeter resolution [100]. A second way, called ‘particle flow’, is to start from the tracks reconstructed in the tracking and vertexing detectors, and to extrapolate these to the calorimeters to match the energy clusters. The latter method is used in CMS [101] and also in LHCb, since it relies primarily on good tracking performance to reconstruct the jets, thereby reducing the contribution of the calorimeters to the jet energy resolution.

The standard jet finding algorithm used in LHCb [102] starts with a partitioning of the charged particles in the event. By default this partitioning is performed by associating tracks to primary vertices using their distance to the vertex $\chi^2_{IP}$. In contrast, to reconstruct jets from long-lived particles, the event is partitioned by associating tracks to the displaced vertices, only selecting tracks that explicitly point away from any primary vertex. The jet finding is performed independently for each candidate in the event, although in practice there is usually only a single displaced vertex candidate. The input to the jet search is selected as follows:

- tracks with a VELO segment (long, upstream and VELO tracks) that have an $\chi^2_{IP} > 20$ to any PV, a distance to the displaced vertex $IP_{DV} < 2$ mm and $IP_{DV} < IP_{PV}$ for any PV;
- downstream tracks with an $\chi^2_{IP} < 30$ to the displaced vertex;
- neutral calorimeter clusters which cannot be associated to a charged track.

The cut on the track $IP_{DV} < 2$ mm was shown in the simulation to be large enough to include over 95% of tracks from the $b$-hadron decay chain of $\pi^0$s. The selection of the downstream tracks takes into account the limited resolution of the pointing information of those tracks due to the long extrapolation distance from the TT detector hits. For the neutral energy deposits in the calorimeter, care is taken to prevent double-counting of the energy. If the neutral energy deposits in the calorimeters are matched to charged particles, the expected calorimeter energies associated to the momentum measurements of the tracks are subtracted from the neutral energy. The clusters that are far enough from any track are considered as neutral particles. These neutral particles are used as input to each of the displaced vertices in the jet finding, since they contain no pointing information.

Additionally, the effect of quality cuts to reduce the fraction of reconstructed ghost tracks (fake tracks resulting from random combinations of hits in the tracking detectors) has been studied, such as a selection on the maximum $\chi^2$/dof in the VELO, on the maximum pseudorapidity of the track, and on the number of missing hits in the
JET RECONSTRUCTION

VELO (compared to the expectation from the tracking). Although a tighter selection reduces the ghost rate, it is observed not to improve the jet energy resolution. Furthermore, more selection criteria on the tracks lead to additional systematic uncertainties on the efficiency or jet energy scale. Consequently, these selection criteria are not applied in the analysis presented in this thesis.

5.1.2 Jet reconstruction algorithm

The charged particles and the neutral energy clusters are combined into jets using the anti-$k_t$ jet algorithm [89] within the FASTJET package [103]. The jet finding procedure searches for jets within a certain cone size around the most energetic objects in the event. There are two categories of jet reconstruction algorithms. The first category are cone jet algorithms that assume the conical energy flow in the event is unchanged by the hadronisation and QCD branching. The jets are reconstructed based on the conical energy deposit in the detector. The second category are sequential recombination algorithms that start from the largest energy deposits in the detector, and repeatedly add surrounding particles according to a distance measure that depends on the energy and the spatial distance of the particles. Especially for LHCb, with its limited acceptance, jets with a smooth conical shape are preferred, because they enable the jet to be reconstructed even though part of its energy falls outside the acceptance. The anti-$k_t$ algorithm combines the advantages of sequential recombination and cone jet algorithms in such a way that soft radiation does not cause the jet boundaries to be irregular.

The main characteristic of the anti-$k_t$ algorithm is that it is infrared safe; particles with a high momentum can modify the shape of the jet, while softer particles cannot. This means that the jet boundary in the algorithm is robust with respect to soft radiation, but flexible for hard radiation. The anti-$k_t$ algorithm is also collinear safe; the output is not dependent on whether a certain amount of energy is carried within one single particle, or within two collinear particles.

The jet finding sequence starts from the hardest (highest transverse momentum) object in the event, which is labelled ‘1’. Soft particles (labelled ‘i’) within a radius $R$ of the hardest object are collected in one jet, that is, if no other hard objects are within that radius. The clustering is performed by first merging the elements with the smallest distance factor $d_{ij}^2$. Defining the transverse momenta of the hardest particle $p_{T1}$ and any other particle $p_{Ti}$, the distance factor between them is:

$$d_{ij}^2 = \min\left(\frac{1}{p_{T1}^2}, \frac{1}{p_{Ti}^2}\right) \frac{\Delta R_{ij}^2}{R^2}$$  \hspace{1cm} (5.1)$$

where $\Delta R_{ij}^2$ is the spatial distance in pseudorapidity and azimuthal angle between the particles ($\Delta R_{ij}^2 = \Delta \phi_{ij}^2 + \Delta \eta_{ij}^2$), and $R$ is the maximum cone radius. This definition,
using the inverse transverse momenta, ensures that the distance factor between two softer particles is larger than the factor between a hard and a softer particle, so softer particles will cluster more easily with a harder particle than with other soft particles. If any other energetic particles are within a distance $R < \Delta R_{ij} < 2R$ of the initial hard object, they will form a jet of their own, and depending on their energy and overlap, at least one of the resulting jets will be conical.

Jets in LHCb are by default reconstructed with a minimal $p_T$ threshold of 10 GeV. However, MC studies showed that lowering the $p_T$ requirement to 5 GeV improves the signal efficiency of the low-mass $\pi^0_v$ samples. Therefore the jets associated with displaced vertices are reconstructed with $p_T > 5$ GeV. The reconstructed jet $p_T$ distribution of the jets associated to displaced vertices is shown in Fig. 5.1. The jets in data typically have a lower transverse momentum than the signal jets.

**Figure 5.1:** Transverse momentum of the individual jets of dijet candidates in data and in simulated hidden valley signal. The distributions are normalised to unit area.

### 5.1.3 Jet reconstruction on signal Monte Carlo

The effect of varying input parameters of the algorithm and of applying cuts on the output jets is studied using simulated signal events. 'True' Monte Carlo jets are reconstructed by running the jet reconstruction algorithm only on particles originating from a true $\pi^0_v$ decay. Since the true jets are reconstructed on top of HepMC [90] particles (i.e. pions, kaons, protons, neutrons, hyperons, photons, electrons, muons and neutrinos), they represent the jets that could be measured with a perfect detector, without any pile-up or underlying event. To retrieve the 'reconstructed' jets, the jet algorithm as used in the displaced vertex analysis is run on the whole event (on all the reconstructed particles instead of on 'truth' MC particles from the $\pi^0_v$ decay only).

From the true jets, the ones with sufficient transverse momentum ($p_T > 5$ GeV),
and with a pseudorapidity $2.0 < \eta < 4.5$ are selected, mimicking the acceptance of the LHCb experiment. An extra selection is made on the true jets to make sure they are reconstructible: they need to originate from a $\pi^0\nu$ that has an associated reconstructed displaced vertex. This has the effect that only jets from $\pi^0\nu$ decays within the VELO acceptance are called reconstructible. Reconstructed jets are matched to MC jets by requiring that they are spatially close, i.e. that $\Delta R < 0.5$. The distribution of $\Delta R$ is shown in Fig. 5.2a. To evaluate the quality of the reconstructed jets, the ratio $E_{\text{reco}}/E_{\text{MC}}$ is studied in Fig. 5.2b. To call a jet 'reconstructed', $0.6 < E_{\text{reco}}/E_{\text{MC}} < 2.0$ is required. This removes about 10% of the jets.

5.2 Jet cone size

An important parameter in the jet clustering is the maximum cone size $R$. The smaller the value of $R$, the larger the fraction of jet energy that is lost, and the poorer the energy resolution. Although larger values of $R$ lead to smaller energy loss, they also increase the noise from the underlying event and the probability that otherwise well-separated parton jets are merged.

By default, the LHCb jet reconstruction is performed with a distance parameter $R = 0.5$, for example in the 'Z + jet' analysis [102]. This cone size is also used for the displaced vertex analyses in the CMS experiment, whereas the ATLAS experiment uses a lower value of $R = 0.4$. However, the analysis presented here uses a larger cone size, of $R = 0.7$, because LHCb has a lower pile-up than the multipurpose experiments, and the sought-after $b$-jets have a relatively large size.
To motivate the choice for the cone size, the algorithms to retrieve true and reconstructed jets (as explained in section 5.1.3) are applied with varying cone size, between $R = 0.4$ and $R = 0.9$. Fig. 5.3 shows the resulting dijet invariant mass of reconstructed jets, and Table 5.1 lists the reconstructed mean and resolution values of the various cone sizes. The aim is to get a good resolution at a mean mass close to the original $\pi^0_v$ mass, in this case 35 GeV. The best combination of reconstruction efficiency and mass resolution, determined by the ratio of the RMS and the mean of the dijet mass, is achieved for $R = 0.7$.

**Figure 5.3:** Reconstructed dijet mass using different cone sizes, for simulated hidden valley signal (with $m_{\pi^0_v} = 35$ GeV and $\tau_{\pi^0_v} = 10$ ps).

<table>
<thead>
<tr>
<th>Cone size</th>
<th>Dijet mass mean [GeV]</th>
<th>RMS [GeV]</th>
</tr>
</thead>
<tbody>
<tr>
<td>$R = 0.4$</td>
<td>18.1</td>
<td>7.9</td>
</tr>
<tr>
<td>$R = 0.5$</td>
<td>21.0</td>
<td>8.1</td>
</tr>
<tr>
<td>$R = 0.6$</td>
<td>23.3</td>
<td>8.5</td>
</tr>
<tr>
<td>$R = 0.7$</td>
<td>25.3</td>
<td>9.0</td>
</tr>
<tr>
<td>$R = 0.8$</td>
<td>27.2</td>
<td>9.7</td>
</tr>
<tr>
<td>$R = 0.9$</td>
<td>29.2</td>
<td>10.6</td>
</tr>
</tbody>
</table>

**Table 5.1:** Reconstructed dijet mass mean and resolution for various jet cone sizes, on the HV10_M35 signal sample (with $m_{\pi^0_v} = 35$ GeV and $\tau_{\pi^0_v} = 10$ ps).
Figure 5.4: Distribution of jet ID variables for jets associated to offline selected displaced vertices in data and simulated hidden valley signal before the jet ID cuts are applied. The distributions are normalised to unity. The following jet ID criteria define a good jet: $\text{MTF} < 0.7$, $\text{MPT} > 0.9$ GeV and $\text{CPF} > 0.1$. 
5.3 Jet identification

Among the objects reconstructed by the jet finding algorithm, there is still a certain fraction of poorly reconstructed and fake jets. Jet identification (jet ID) variables can be used to select good jets:

- **MTF** is the fraction of transverse momentum carried by the highest \( p_T \) track. \( \text{MTF} < 0.7 \) rejects low multiplicity jets that consist essentially of one track, such as leptons;

- **MPT** is the \( p_T \) value of the track with most \( p_T \) in the jet. \( \text{MPT} > 0.9 \text{ GeV} \) selects jets with at least one charged particle with sufficient transverse momentum. It rejects jets that are composed only of neutrals and soft tracks;

- **CPF** is the fraction of charged transverse momentum in the jet, with respect to the total \( p_T \) of neutral and charged particles. \( \text{CPF} > 0.1 \) rejects jets with a small charged \( p_T \) fraction. A minimum amount of charged \( p_T \) is required since vertexing will be performed within the jet in a later stage of the reconstruction;

- **N90** is the minimum number of charged and neutral particles required to contain 90% of the jet \( p_T \). No selection is made on this variable.

Fig. 5.4 shows the distributions of the jet ID variables listed above, for candidates that pass the offline selection in data and signal MC. The jet ID selection values were determined from these distributions. There is a contribution in data from fake jets that have either many low-momentum tracks (visible in the MTF and MPT distributions), one or two high-momentum tracks (N90), or either a large fraction of neutral energy or only charged energy (CPF). These are removed from the sample by the jet identification quality criteria.

5.4 Jet reconstruction efficiency

The efficiency of reconstructing jets is defined using the reconstructed and true jets as described in Section 5.1.3. To call a true jet 'reconstructible', it must have: \( p_T > 5 \text{ GeV}, 2.0 < \eta < 4.5 \), and there has to be a reconstructed displaced vertex associated to the true \( \pi^0 \).

The jet reconstruction efficiency before and after the jet ID selection is shown as a function of \( p_T, \eta, \phi \), the decay time \( c\tau \) and the radial distance \( R_{xy} \) in Fig. 5.5. The efficiency drops below \( \eta = 2 \) and above \( \eta = 4.5 \) due to the detector acceptance. The decrease at low \( p_T \) can be related to the drop in efficiency at high \( \eta \). Furthermore, low-\( p_T \) particles might not reach the calorimeters because they are bent outside the detector acceptance due to the magnetic field. The efficiency is reasonably independent of the
true decay time and $R_{xy}$, as long as the decay vertex is well within the fiducial volume of the VELO and outside the primary interaction region. Studies of the jet ID efficiencies as a function of $p_T$ showed that the MPT cuts mostly affect low $p_T$ jets, while the CPF efficiency is approximately independent of $p_T$. A summary of the efficiency of the jet ID cuts on signal MC, integrated over all $p_T$ values, is given in Table 5.2.

<table>
<thead>
<tr>
<th>cut</th>
<th>$\text{eff}_{MC}$ only this cut</th>
<th>$\text{eff}_{MC}$ this cut last</th>
</tr>
</thead>
<tbody>
<tr>
<td>MTF</td>
<td>99.55 ± 0.02</td>
<td>99.54 ± 0.02</td>
</tr>
<tr>
<td>CPF</td>
<td>98.49 ± 0.04</td>
<td>99.55 ± 0.02</td>
</tr>
<tr>
<td>MPT</td>
<td>97.68 ± 0.05</td>
<td>98.72 ± 0.04</td>
</tr>
<tr>
<td>all</td>
<td>96.79 ± 0.06</td>
<td>-</td>
</tr>
</tbody>
</table>

*Table 5.2: Efficiencies of jet ID cuts on reconstructed jets associated to the $\pi^0$ candidates in the simulated hidden valley signal (with $m_{\pi^0} = 35$ GeV and $\tau_{\pi^0} = 10$ ps). The requirement $0.6 < E_{\text{reco}}/E_{\text{MC}} < 2.0$ has been applied on the reconstructed jets.*

### 5.5 Jet energy correction

A jet energy correction is applied to recover the non-reconstructible neutral particles and charged particles that fall outside the detector, to correct for the finite resolution of the calorimeter, and to correct for noise and pile-up in the event. This correction will be applied as a function of the number of primary interactions per event.

Jets with low $p_T$ have a deteriorated $p_T$ resolution, as is shown in Fig. 5.6a. After a selection on the 'reconstructed' jets, which requires $0.6 < E_{\text{reco}}/E_{\text{MC}} < 2.0$, the differences in resolution between the $p_T$ bins decrease. Figure 5.6b shows the ratio of the transverse momenta of the reconstructed and associated true jets, for various pile-up multiplicities. For events without pile-up ('1 PV'), the distribution is slightly narrower, than for higher multiplicities.

Figure 5.7a compares the energy of the reconstructed jets to that of the 'true' jets in signal MC, for events with one primary interaction. A jet energy correction is applied for the pile-up; the number of primary interactions in the event. Fig. 5.8 shows the ratio $E_{\text{reco}}/E_{\text{MC}}$ for different primary vertex multiplicities. The ratio features on the one hand a shift due to pile-up contributions (a positive bias) and on the other hand it features losses from reconstruction (a negative bias). A jet energy correction factor is retrieved by taking the mean of a fit to this ratio, using a landau function convoluted with a gaussian. The fit is performed only on the jets with an energy in the range 60-600 GeV. The correction factor retrieved in this way for the events with one primary vertex is depicted by the slope of the grey line in Fig. 5.7b. The correction factors as a function of the number of primary vertices are listed in Table 5.8.
Figure 5.5: Jet reconstruction efficiency excluding jet ID selection (black) and including jet ID (green) on the simulated hidden valley signal. For MC jets with $p_T > 5$ GeV from $\pi^0$ decays inside the VELO acceptance.
Figure 5.6: Ratio of the transverse momenta of the reconstructed and associated true jets in simulated hidden valley signal, (a) for various bins of jet $p_T$ (before the selection on the energy ratio), and (b) for various pile-up multiplicities (after the selection on the energy ratio). All distributions are normalised to unity.

Figure 5.7: (a) Energy of the reconstructed jet versus energy of the associated true jet for events with one primary vertex (before the selection on $0.6 < E_{\text{reco}}/E_{\text{MC}} < 2.0$). For jets originating from $\pi^0$ samples with 10 ps lifetime and masses of 25, 35 and 43 GeV. (b) Profile of reconstructed versus true jet energy for events with one primary vertex. The slope of the grey line (0.843) is the JEC factor, which is the mean of the fit to events with no pile-up in Fig. 5.8.
Figure 5.8: Ratio of the energy of the reconstructed jet and the associated true jet, for different PV multiplicities. For jets originating from a $\pi^0$ with 10 ps lifetime and a mass of either 25, 35 or 43 GeV. The jet energy correction factors in the table on the right represent the mean value of a fit (landau function convoluted with a gaussian) of each curve in the left figure.

As illustrated in Fig. 5.9, applying the pile-up corrections to the jet four-momenta leads to a shift of the mass to higher values, but does not improve the mass resolution. The ratio of data over signal events under the signal peak (within two standard deviations) slightly decreases after the correction, which suggests that the correction will not worsen the signal from background separation. The jet energy correction will be applied to the dijet mass of the candidates, after all the selections have been performed.

In other LHCb analyses that use jets, a jet energy correction is applied as a function of the number of primary interactions in the event, of the pseudorapidity and the transverse momentum of the jets, and of the number of charged particles in the jet. The correction factors are retrieved from the energy or $p_T$ ratio of the true and reconstructed Monte Carlo jets [104]. The LHCb ‘$Z + \text{jet}$’ analysis showed that the jet energy response is well described by the simulation [102].

Since this analysis uses a larger cone size than the default jet reconstruction, all the correction factors would have to be retrieved anew. In order to judge the importance of the jet energy correction for this analysis, the effect of a correction on the dijet mass mean and resolution is studied first. The result is that a correction does not improve the mass resolution: it only shifts the reconstructed mass to higher values. Since the corrections are larger for low $p_T$ jets, which appear more often in background than in signal, it is expected that the low mass background will be moved into the signal region. Therefore, no jet energy correction depending on jet kinematics is applied.

Not only does this analysis use a different cone size than the default LHCb jet reconstruction, it also uses displaced rather than prompt jets. Therefore any dependence on the flight distance to the PV of the jet energy response (for example caused by the
Figure 5.9: Dijet invariant mass distribution (a) before and (b) after the jet energy correction is applied. Data is shown in black, simulated hidden valley signal in red. For visibility, the simulated signal is scaled to 0.62 fb$^{-1}$ assuming a Higgs cross-section of 10 nb and branching fractions of 100% for $B(H \to \pi^0\pi^0)$ and $B(\pi^0 \to b\bar{b})$.

VELO tracking efficiency might influence the quality of the MC description of the jets. Whereas the 'Z + jet' analysis showed that the energy response is described in the MC within a few percent, a decline of the resolution as a function of $R_{xy}$ could possibly worsen this result. However, no significant dependence is observed of the MC jet energy response on the radial displacement of the candidate, or on the $z$ position of the primary vertex. The jet reconstruction efficiency shows no decline as a function of radius, as can be seen in Fig. 5.5e and 5.5f. Therefore, no additional systematic uncertainty is expected on the energy scale of displaced jets within the VELO acceptance.

5.6 Outlook

The selection of input particles plays a major role in the jet reconstruction. Various studies on the exclusion of ghost-like and bad quality tracks have not yielded any notable improvement in the jet energy resolution. However, studies to exclude jets with large numbers of ghost tracks in the outer tracker could lead to a solution for the peaking distribution of low-quality jets in the $\phi$ plane, that will be discussed in more detail in Chapter 7. An alternative way to solve this is to apply a multi-variate selection using input variables such as the charged particle density, photon density, and energy distribution in the jet.

To reconstruct a cleaner sample of jets, one could increase the $p_T$ threshold for jets
from 5 GeV to 10 GeV. The lower-mass signals, for which the efficiency would decrease, could be recovered using the selection of a single large jet instead of two small jets. The reconstruction of single-jet candidates is an interesting approach to enhance the sensitivity for lower signal masses. However, the background for single-jet candidates is much larger than for dijets. Therefore one could reconstruct the substructure within the jet to distinguish signal from background. A separation can be made between the 'direct' decay vertex, and 'subvertices' from the $b$-decays. This also enables a separation between different quark flavours (mainly $b$ versus $c$ and lighter quarks), and could additionally be used as a means of jet flavour tagging for the dijet analysis.

One of the issues that influences the selection efficiency is the inconsistency between the online and offline reconstructed tracks and vertices. Currently the displaced vertex is reconstructed twice, in the HLT2 and in the stripping. The largest efficiency loss in the offline selection is caused by the requirement of two jets. This is solved by including the jet reconstruction in the stripping procedure, thereby omitting the stripping vertex reconstruction. This decreases the systematic uncertainties related to the online-offline difference and it increases the selection efficiency.
Candidate reconstruction

After the jets are reconstructed and selected, further selections are made to ensure that the jets are well-matched to the displaced vertex. Once the candidates are defined, the jet multiplicity and the jet or multi-jet properties can be used to distinguish the signal candidates from background.

6.1 Jet-to-vertex matching

To ensure that the jets and the vertex belong to the same $\pi^0$ candidate, a jet-to-vertex matching is performed in two ways. Firstly, as was described in Section 5.1.1, each jet is constructed from individual particles that point back to the displaced vertex, using the impact parameter of the tracks to the vertex. Secondly, the resulting jet should also point back to the vertex, which is achieved by a selection using the maximal impact parameter of the jet to the displaced vertex.

In order to calculate the impact parameter (IP) of a jet to a vertex, the ‘trajectory’ of the jet needs to be defined. If the jet has only one track segment with VELO information, the jet IP is defined as the IP of this segment to the vertex. Otherwise, if the jet has more than one track segment, an adaptive vertex fit to all the tracks is performed. The position of the resulting vertex is combined with the jet momentum direction to compute an impact parameter to the displaced vertex.

To assure that the jet contains sufficient information to compute a reliable impact parameter, jets are only accepted in case they contain at least one charged particle with a VELO segment. Fig. 6.1 shows the track content of the jets associated to each vertex after this requirement, illustrating that the tracks without a VELO segment (down-
Figure 6.1: Track content of the jets associated to displaced vertices. Number of tracks of a certain type divided by the total number of tracks in the jet, normalised to unity in data and simulated hidden valley signal.

Figure 6.2: (a) Impact parameter between the jet trajectory and the displaced vertex candidate and (b) the smallest impact parameter of the jet to any primary vertex, in data and simulated HV10_M35 events. The distributions are normalised to unity.
stream tracks) only constitute a small portion of the tracks in the jets. The difference in track content between data and simulation is due to the fact that the simulation contains only signal events, whereas the data contains mainly background events.

Figure 6.2a shows the IP distribution of the jets to the displaced vertex and 6.2b shows the smallest IP of the jets to any primary vertex, in both data and in simulated signal. As expected, the IP of the jets to the displaced vertex is peaked at zero. The tail in data is larger than in signal due to combinatorial background. The distance to the PV is larger in signal events than in the data because the majority of candidates in data, consisting typically of $B$-decays, have a shorter decay length than the signal candidates. Jets are associated to a candidate if their impact parameter to the displaced vertex is smaller than 2 mm, and if that IP is also smaller than the IP to any PV. This selection is over 98% efficient for simulated signal events with two jets.

### 6.2 Jet multiplicity

In Chapter 3 the jet multiplicity of simulated signal was studied (Fig. 3.3), by performing the jet reconstruction on true $\pi^0$ decay particles at generator level. Most candidates have two associated jets, which is what is expected for a $\pi^0$ decaying to $b\bar{b}$. Only for the lowest mass sample (15 GeV) the two jets often merge into one.

![Jet multiplicity plot](image)

**Figure 6.3:** Distribution of the number of jets per vertex candidate in data and in simulated hidden valley signal events with various masses, after the global event cut and material veto. The distributions are normalised to unity.

The same illustration can be made for candidates made with reconstructed particles instead of ‘true’ MC particles. Figure 6.3 shows the jet multiplicity (with all the quality cuts on the jets applied) of offline selected candidates in signal samples with different masses. Most of the candidates in the benchmark model HV10_M35 have two associated jets, although the second jet is less often found for the lower mass sam-
Figure 6.4: Mass distribution of the candidates with exactly two jets, after the global event cut and material veto. (a) comparing various simulated hidden valley signal samples, and (b) comparing data and simulated signal. All distributions are normalised to unity.

More than two jets are expected to be found for a small fraction of the candidates due to gluon radiation. The tight jet ID requirements, in particular on the minimum $p_T$, reduce the fraction of such soft jets. Both in signal and in data, only few candidates have three or more jets. Figure 6.5 illustrates that the mass distribution of those signal candidates has a prominent tail towards higher mass. In signal, the mean of the mass peak only change moderately, whereas the three-jet candidates in data have a much higher mass than the dijet candidates. This makes the distinction between signal and background (assuming most of the data consists of background) more difficult, as shown in Fig. 6.5b. Because of these reasons, and because the contribution of the three-jet candidates to the signal efficiency is small, they are ignored for the current analysis.

The candidates with one jet cannot easily be distinguished from background, since most of the background also consists of single $b$-jets. Additionally, the separation of signal from background is harder, because the kinematic properties of combinations of jets (like $\Delta R$ and $\Delta \phi$) cannot be used. These candidates are therefore also ignored, but could be investigated in future searches for low-mass long-lived particles.

Hence, only dijet candidates are considered in this analysis. Although the dijet requirement removes a large fraction of the background (as will be discussed in Sec. 6.7), further selections are needed to eliminate events that do not have a signal-like topol-
Selection using the jet pointing

In this section, the variable $m/m_{\text{corr}}$ will be derived, to select dijet candidates that 'point back' to the primary vertex. Ideally, by reconstructing all the particles of the decay, the original momentum of the long-lived particle can be retrieved. The momentum of a good candidate points back to the primary interaction, whereas background candidates, for example created by combinatorics, will have a more random momentum direction. The pointing angle $\theta$ can be defined as the angle between the vector $\vec{d}$ pointing from the PV to the DV and the momentum vector $\vec{p}$ of the candidate, as illustrated in Fig. 6.6. The angle $\theta$ gives an indication of the quality of the reconstructed momentum of the candidate.

The resolution of the pointing angle is determined by the reconstruction of both $\vec{p}$ and $\vec{d}$ vectors. Figure 6.7 illustrates that the contribution to the uncertainty on $\theta$ from the displacement vector $\vec{d}$ is small compared to the contribution of the momentum vector $\vec{p}$. Reasons for this are the relatively large decay length of the particles and the good displaced vertex position resolution. The uncertainty on the dijet momentum is determined by the jet energy and the jet direction resolutions, which are discussed in more detail in Chapter 8.

The pointing angle could be used to correct the mass of decaying particles, by

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{fig6_5.jpg}
\caption{Mass distribution of the candidates with three jets, after the global event cut and material veto. (a) comparing various simulated hidden valley signal samples, and (b) comparing data and simulated signal. All distributions are normalised to unity.}
\end{figure}
Figure 6.6: Schematic representation of a long-lived particle produced at a primary vertex (PV) decaying into two jets, with the momentum vector of the candidate $\vec{p}$, and the vector $\vec{d}$ pointing from the PV to the displaced vertex (DV). The angle between the two is defined as $\theta$.

Figure 6.7: Contributions of the dijet momentum $\vec{p}$ (grey) and the vertex displacement $\vec{d}$ (black) to the resolution of the pointing angle. The angle between the true momentum and $\vec{p}$ respectively $\vec{d}$ is shown, for simulated hidden valley signal (with $m_{\nu^0} = 35$ GeV and $\tau_{\nu^0} = 10$ ps), after the global event cut and material veto.
forcing $\vec{p}$ to be parallel to $\vec{d}$ through the addition of missing transverse momentum. As explained in Section 4.4, the corrected mass is the minimal mass obtained from the resulting four-momentum, and is given by equation 4.5. Knowing that the missing transverse momentum is related to the pointing angle as $|\vec{p}_\perp| = \sin \theta |\vec{p}|$, the corrected mass is also given by:

$$m_{\text{corr}} = \sqrt{m^2 + (|\vec{p}| \sin \theta)^2} + |\vec{p}| \sin \theta.$$  \hspace{1cm} (6.1)

For the long-lived particle mass estimate, a correction of the mass is not applied. The main reason is that the corrected mass variable can only compensate for missing energy, and therefore increase the mass. It does not enable both downwards and upwards corrections, which would be required to correct for uncertainties in the momentum measurement. These uncertainties include the added momentum from pile-up and the energy resolution in the calorimeter. As illustrated in Fig. 6.8, the corrected mass distribution for dijet candidates has a higher mean than the uncorrected mass distribution, but worse resolution, and a tail at high mass. The only correction that will be applied to the candidates is the jet energy correction, described previously in Section 5.5.

![Figure 6.8: Distributions of (a) the mass and (b) the corrected mass in data and simulated hidden valley signal (with $m_{\pi^0} = 35$ GeV and $\tau_{\pi^0} = 10$ ps), before the selection on the pointing of the jets. The distributions are normalised to unity.](image)

Since the invariant dijet mass spectrum will be used for the fit to calculate the signal and background yields, a selection that implicitly applies an invariant mass threshold, like a cut on $p_T$ or $\sum p_T$, should be avoided. The quantity $m/m_{\text{corr}}$ is an example of a more mass-independent variable. The quantity $m/m_{\text{corr}}$ can be retrieved from the
following formula:

\[
\frac{m_{corr}}{m} = \sqrt{1 + \left(\frac{p}{m} \sin \theta\right)^2 + \frac{p}{m} \sin \theta}.
\] (6.2)

The analysis should be as inclusive as possible for different long-lived particle models, so independence of the boost \( p/m = \beta \gamma \) is also required. Studies showed that \( m/m_{corr} \) is almost independent of the boost, unlike the pointing angle \( \theta \) itself. For this reason a selection using \( m/m_{corr} \) is preferred.

A threshold selection value is optimised using the figure of merit \( \epsilon/\sqrt{\epsilon\sigma} \), where \( \epsilon \) is the efficiency of the selection, and \( \sigma \) is the invariant mass resolution after the selection. The optimum selection threshold is found at \( m/m_{corr} > 0.7 \). The distribution of \( m/m_{corr} \), before this selection on the pointing is applied, is shown in Fig. 6.9.

**6.4 Selection using the jet opening angle**

A second variable that can be used for selection is the opening angle between the two jets, which is represented by their distance \( \Delta R = \sqrt{\Delta \phi^2 + \Delta \eta^2} \). The signal jets are differently distributed in \( \eta \) than the jets in data (compare Fig. 6.10a and 6.10c). Figure 6.10e, showing only the high-mass candidates in data, illustrates that this additional contribution of 'back-to-back' jets in data results in candidates with a high dijet mass. This is confirmed by the distribution of the azimuthal angles of the jets. Figure Fig. 6.10b gives the distribution for HV10_M35, showing correlations in \( \phi \). These bands are caused by the \( p_T \) distribution of the jets; for signal samples with a higher or
Figure 6.10: Distributions of the pseudorapidity (left) and the azimuthal angle (right) of the candidate jets in (a), (b) simulated signal (HV10_M35), in (c), (d) data, and in (e), (f) data for candidates with a dijet mass over 20 GeV. All distributions are made before the selection on the pointing and opening angle of the jets is made.
lower $\pi^0$ mass, the bands are in a different place. Figure 6.10d illustrates the $\phi$ distributions of the jets in data. Again, when selecting the high-mass candidates (Fig. 6.10f), the 'back-to-back' jet contribution is clearly visible. This feature is not observed in simulated signal, as shown in Fig. 6.10b.

It is likely that the events at high $\Delta R$ correspond to $b\bar{b}$ events. The production of $b\bar{b}$ happens predominantly through the processes of flavour creation in hard scattering (gluon-gluon fusion and $q\bar{q}$ annihilation), flavour excitation and gluon splitting. The dominant process is the hard scattering flavour creation process \[105\]. The production mechanism of the $b\bar{b}$ pair determines the topology of the quark pair and its decay products. For example, $b$-quarks produced in gluon-gluon fusion are produced 'back-to-back' in the transverse plane and result in azimuthally 'back-to-back' jets. The gluon splitting process on the other hand typically results in jets that are close together in the azimuthal plane.

Samples of $b\bar{b}$ events are available in data and in MC through the so-called 'double-topo' selection. The sample consists of events with two 2-, 3-, or 4-prong vertices selected by the topological lines, that add up to the $B$ meson mass. There is no low-mass double-topo sample available; a minimal dijet mass of 19 GeV is required for the candidates in the stripping selection. Note that in these events, the $b$ jets originate from the primary vertex, rather than from a displaced vertex. The $\phi$ and $\eta$ distributions of the jets in these candidates are shown in Fig. 6.11, and the 'back-to-back' contribution is clearly visible in the $\phi$ distribution.

There is an additional contribution in data in the $\phi$ plane of the two jets in Fig 6.10d around $(0, 0)$ and $(\pi, \pi)$, which is likely due to fake jets that are reconstructed just to

---

Figure 6.11: Distributions of (a) the pseudorapidity and (b) the azimuthal angle of dijet candidates in a sample of double $B$-decays in data, selected with the double topo selection.
Figure 6.12: Distribution of the dijet mass versus $\Delta R$ of the candidates in simulated hidden valley signal with (a) $m_{\pi^0} = 35$ GeV, (b) $m_{\pi^0} = 50$ GeV and (c) data. Before the selection on the pointing and opening angle of the jets is made. (d) shows the same distribution for dijet candidates in a sample of double $B$ events in data, selected with the double topo selection. The selection on $\Delta R$ is set at $\Delta R < 2.2$. 
the left and right of the beam pipe because of the large occupancy in the tracker and calorimeter in that region.

A selection using a maximal value of $\Delta R$ eliminates jets which are far apart in either $\eta$, $\phi$, or both. Figure 6.12 shows the distribution of the dijet mass versus $\Delta R$ for several signal and data samples. Data, shown in Fig. 6.12c, has two contributions, one at low $\Delta R$, and one at high $\Delta R$. The contribution at high $\Delta R$ is likely to come from events with two $B$-decays, since those events, selected through the double topo selection, also peak at high $\Delta R$, as shown in Fig. 6.12d.

![Delta R distribution](image)

**Figure 6.13:** The $\Delta R$ distributions of the candidate jets in data and simulated hidden valley signal, before the selection on the pointing and opening angle of the jets is made. All distributions are normalised to unity. The selection on $\Delta R$ is set at $\Delta R < 2.2$.

The high $\Delta R$ contribution can be removed by the requirement: $\Delta R < 2.2$, which is in between the two contributions in data shown in Fig. 6.12c. This unfortunately implies that the high-mass signal samples, for which the jets are evenly distributed in $\phi$, suffer efficiency loss. The loss in efficiency can be deduced from the $\Delta R$ distributions in Fig. 6.13. From the same figure one can conclude that the low-mass signal sample has a similar shape as the data. As stated before, the dijet requirement is less suitable for the lower mass, where it is found that if two jets are reconstructed, they have low momentum. In this case the signal sample resembles the background where a $b$-jet is split into two jets.

An improvement for future analyses to prevent efficiency loss for high-mass signals could be to apply a different selection for every mass hypothesis.
6.5 Candidate multiplicity

Since the purpose of this analysis is to search for a single long-lived particle, all the efficiencies and distributions are given for individual candidates. It is however possible that more than one candidate appears per event. This is for example the case in the generated MC samples, where a Higgs boson always decays to two $\pi^0$ particles, which can both decay in the detector acceptance.

<table>
<thead>
<tr>
<th>Selection</th>
<th>HV10_M35 # candidates</th>
<th>Data 2011 # candidates</th>
<th># events</th>
<th># events</th>
</tr>
</thead>
<tbody>
<tr>
<td>Generated</td>
<td>$2 \times 193666$</td>
<td>193666</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>InTuple</td>
<td>114359</td>
<td>71012</td>
<td>6637942</td>
<td>6475590</td>
</tr>
<tr>
<td>L0</td>
<td>74153</td>
<td>51339</td>
<td>3231457</td>
<td>3226553</td>
</tr>
<tr>
<td>HLT1</td>
<td>58712</td>
<td>42996</td>
<td>2774557</td>
<td>2772522</td>
</tr>
<tr>
<td>HLT2</td>
<td>40009</td>
<td>31244</td>
<td>1931924</td>
<td>1930988</td>
</tr>
<tr>
<td>Stripping</td>
<td>15118</td>
<td>14020</td>
<td>1862390</td>
<td>1861519</td>
</tr>
<tr>
<td>GEC</td>
<td>15090</td>
<td>13996</td>
<td>1856336</td>
<td>1855478</td>
</tr>
<tr>
<td>MV</td>
<td>15046</td>
<td>24178</td>
<td>1833816</td>
<td>1839580</td>
</tr>
<tr>
<td>Ntracks</td>
<td>14481</td>
<td>13479</td>
<td>1423144</td>
<td>1422651</td>
</tr>
<tr>
<td>Two Jets</td>
<td>8689</td>
<td>8333</td>
<td>56218</td>
<td>56216</td>
</tr>
<tr>
<td>$M/M_{\text{corr}} &gt; 0.7$</td>
<td>7423</td>
<td>7143</td>
<td>31739</td>
<td>31738</td>
</tr>
<tr>
<td>$\Delta R &lt; 2.2$</td>
<td>6549</td>
<td>6306</td>
<td>29921</td>
<td>29920</td>
</tr>
</tbody>
</table>

Table 6.1: Number of candidates and number of events surviving the selection in simulated hidden valley signal (HV10_M35) and in data. On data, the triggers L0, HLT1 and HLT2 have been applied online, and are re-applied offline with the TOS or TUS requirements.

Table 6.1 lists the difference between the number of candidates and the number of events, for various selection stages. The collection of candidates that are stored in a tuple for further analysis (denoted as 'InTuple') consists of all cases where a displaced vertex is reconstructed in the stripping. In MC, the trigger requirements can be applied afterwards, while in data non-triggered events cannot be retrieved. After the final selection, the average number of candidates per event in the simulated sample is 1.04, i.e. close to unity. Almost all events with more than one candidate have exactly two candidates, which correspond to different true $\pi^0$'s. In the data only one event with two candidates survives the full selection.

If there are two candidates in an event, it is possible to combine them to reconstruct the mass of the mother particle, in this case the Higgs. The invariant mass distribution of events with two candidates is shown in Fig. 6.14, illustrating that the reconstructed mass is consistent with the generated Higgs mass of 120 GeV.
Figure 6.14: Mass distribution of the mother of the $\pi^0$ candidates for events with two candidates. For simulated hidden valley signal, after the jet energy correction has been applied.
6.6 Characteristics of the selected candidates

The offline selection for the long-lived particle analysis can be summarised as follows:

**stripping**
Select events from the 'StrippingDVSingleHighMass', 'StrippingDVSingleMedium' or 'StrippingDVSingleHighFD' stripping lines, as described in Chapter 4;

**material interaction veto**
To reject candidates from material interactions, the 'extended MV' is used, as explained in Section 4.6. A looser version of the veto is also used in displaced vertex trigger and stripping lines.

**VELO GEC**
To reject events from interactions of beam halo tracks with material in front of the VELO, a global event cut (GEC) is applied on the \( \phi \) distribution of VELO hits and on the ratio of VELO hits to tracks;

**N tracks**
At least six tracks are required in the displaced vertex;

**two jets**
Only dijet candidates are considered;

**pointing or corrected mass**
To improve the mass resolution and to remove more background \( m/m_{\text{corr}} > 0.7 \) is required;

**jet opening or \( \Delta R \)**
To remove a contribution from back-to-back dijet events and combinatorics, \( \Delta R < 2.2 \) between the two jets is required.

The mass and the radius from the beamline are the two variables that will be used to fit the distribution of the offline selected candidates. The distribution of the number of events for the selection stages as listed in Table 6.1 is shown as a function of either mass or \( R_{xy} \) in Fig. 6.15 (for data) and 6.16 (for simulated signal).

For the 15 GeV sample (Fig. 6.16a and 6.16b), the requirement of two jets is far from ideal, for it removes single jet candidates with a well-reconstructed mass. However, for the other two mass samples, this criterion is necessary to eliminate the single jet candidates peaking at low mass. The selection on \( \Delta R \) is inefficient for the 50 GeV sample, rejecting events at high mass (Fig. 6.16e). This cut is nevertheless well-motivated to eliminate fake dijet candidates, as was explained in Section 6.4. Since the MC includes no background events, this effect does not appear in the MC distributions. The distributions of signal events in \( R_{xy} \) show the effect of the material veto at the stripping level, which removes events with \( R_{xy} \) in the range 5 – 14 mm. The final selection leaves only a small fraction of events at a radius above 5 mm.
Figure 6.15: Number of candidates versus (a) mass and (b) radial distance from the beamline after different selection levels, for 2011 data.

6.7 Selection efficiencies

Table 6.1 lists the numbers of events in data passing various selection stages. They can be compared to the number of selected events in signal MC, shown in Tables 6.2, 6.3 and 6.4. Since non-triggered events cannot be retrieved in data, the trigger stages quoted in Table 6.1 are obtained by applying the TUS or TOS requirement on offline candidates. A comparison between the selection efficiency on data and on simulated events can therefore only be made from the line 'Stripping' onwards.

Surprisingly, although the stripping is required for all events that enter this table, some candidates are lost when re-applying the 'Stripping' selection. The reason for this is that the definition of the tracks associated to the vertex is slightly different in the analysis environment than in the stripping. When storing the vertices retrieved in the stripping as particles to be analysed, a selection is applied on the quality of tracks that are not VELO-only: $\chi^2/NDF < 5.0$, which possibly reduces the number of tracks. Although these events are selected by the stripping, they can feature a candidate with less tracks than the threshold value. The cut on the minimal number of tracks is re-applied in the analysis. This procedure rejects about 4% of the stripped candidates in data, and 1% in simulated signal.

The global event cut ('GEC') rejects only a small portion of events, since most of the beam-splash events (as described in Section 4.7) are already vetoed by the triggers and the stripping. However, the GEC is crucial to remove the few remaining beam-splash events, since they occur mainly at high mass. The selection that removes most
Figure 6.16: Number of candidates from generated hidden valley signal events (with 10 ps lifetime and various masses) after different selection levels, versus mass (left) and radial distance from the beamline (right). The grey histograms show the distribution of generated candidates.
events in data is the requirement of the presence of two jets.

The requirement of two jets is the most rigorous cut in the offline selection. Still, whereas the rejection on data is a factor 20, the efficiency on signal is around 50%. An exception is the 15 GeV mass signal, for which a different reconstruction sequence might be more beneficial, as explained in Section 6.2.

The inefficiency on signal events of both the trigger and the stripping procedure is mainly due to the limited output rate and reconstruction time that is allowed at these stages. The line 'HLT2' indicates how many interesting events could be reconstructed with a loose (SinglePS) stripping selection without these restrictions. The efficiency loss caused by the combination of the stripping selection and the jet selection cannot be recovered for the 2011 analysis, but the inclusion of the jet reconstruction in the stripping procedure should solve this issue for future analyses.

Tables 6.2 and 6.3 list the signal selection efficiencies, which are calculated as a fraction of all generated events. The generated events include events with a displaced vertex or jets outside the acceptance, so the estimated efficiencies calculated here are lower than the efficiencies on 'reconstructible' long-lived particles.

The sensitivity of this analysis for several masses and lifetimes can be deduced from the final efficiencies shown here. The selections are not fine-tuned to one specific model, which is supported by the fact that for hidden valley samples where the $\pi^0_V$ decays into light quarks, the selection efficiency is higher than for the benchmark model, as shown in Table 6.4. Long-lived particles decaying into light quarks usually have lighter jets, and more tracks in the secondary vertex, which makes the stripping selection more efficient.

The selection efficiencies reported in this section will be used to obtain the upper limits on signal models with several masses and lifetimes.

<table>
<thead>
<tr>
<th>Selection</th>
<th>HV10</th>
<th>HV10</th>
<th>HV10</th>
<th>HV10</th>
<th>HV10</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>M15</td>
<td>M25</td>
<td>M35</td>
<td>M43</td>
<td>M50</td>
</tr>
<tr>
<td># cand %</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Generated events</td>
<td>94440</td>
<td>196684</td>
<td>193666</td>
<td>194306</td>
<td>86852</td>
</tr>
<tr>
<td>InTuple</td>
<td>36995</td>
<td>39.17</td>
<td>100276</td>
<td>50.98</td>
<td>114359</td>
</tr>
<tr>
<td>L0</td>
<td>27899</td>
<td>29.54</td>
<td>69866</td>
<td>35.52</td>
<td>74153</td>
</tr>
<tr>
<td>HLT1</td>
<td>19531</td>
<td>20.68</td>
<td>53658</td>
<td>27.28</td>
<td>58712</td>
</tr>
<tr>
<td>HLT2</td>
<td>12903</td>
<td>13.66</td>
<td>36820</td>
<td>18.72</td>
<td>40009</td>
</tr>
<tr>
<td>Stripping</td>
<td>1533</td>
<td>1.62</td>
<td>9028</td>
<td>4.59</td>
<td>15118</td>
</tr>
<tr>
<td>GEC</td>
<td>1510</td>
<td>1.60</td>
<td>8990</td>
<td>4.57</td>
<td>15090</td>
</tr>
<tr>
<td>MV extended</td>
<td>1456</td>
<td>1.54</td>
<td>8883</td>
<td>4.52</td>
<td>15046</td>
</tr>
<tr>
<td>Ntracks ≥ 6</td>
<td>1336</td>
<td>1.41</td>
<td>8395</td>
<td>4.27</td>
<td>14481</td>
</tr>
<tr>
<td>Two Jets</td>
<td>89</td>
<td>0.99</td>
<td>3799</td>
<td>1.93</td>
<td>8689</td>
</tr>
<tr>
<td>$M/M_{\text{corr}} &gt; 0.7$</td>
<td>68</td>
<td>0.07</td>
<td>3362</td>
<td>1.71</td>
<td>7423</td>
</tr>
<tr>
<td>$\Delta R &lt; 2.2$</td>
<td>67</td>
<td>0.07</td>
<td>3259</td>
<td>1.66</td>
<td>6549</td>
</tr>
</tbody>
</table>

Table 6.2: Number of candidates and selection efficiencies on simulated hidden valley signal with a lifetime of 10 ps and various masses.
### Table 6.3: Number of candidates and selection efficiencies on simulated hidden valley signal with a lifetime of 100 ps and various masses.

<table>
<thead>
<tr>
<th>Selection</th>
<th>HV10_BB</th>
<th>HV10_CC</th>
<th>HV10_SS</th>
</tr>
</thead>
<tbody>
<tr>
<td># cand</td>
<td>%</td>
<td># cand</td>
<td>%</td>
</tr>
<tr>
<td>Generated events</td>
<td>193666</td>
<td>54184</td>
<td>59806</td>
</tr>
<tr>
<td>InTuple</td>
<td>114359</td>
<td>59.05</td>
<td>37348</td>
</tr>
<tr>
<td>L0</td>
<td>74153</td>
<td>38.29</td>
<td>25819</td>
</tr>
<tr>
<td>HLT1</td>
<td>58712</td>
<td>30.32</td>
<td>20311</td>
</tr>
<tr>
<td>HLT2</td>
<td>40009</td>
<td>20.66</td>
<td>15504</td>
</tr>
<tr>
<td>Stripping</td>
<td>15118</td>
<td>7.81</td>
<td>9464</td>
</tr>
<tr>
<td>GEC</td>
<td>15090</td>
<td>7.79</td>
<td>9455</td>
</tr>
<tr>
<td>MV extended</td>
<td>15046</td>
<td>7.77</td>
<td>9428</td>
</tr>
<tr>
<td>Ntracks ≥ 6</td>
<td>14481</td>
<td>7.48</td>
<td>9223</td>
</tr>
<tr>
<td>Two Jets</td>
<td>8689</td>
<td>4.49</td>
<td>6017</td>
</tr>
<tr>
<td>M/M_{corr} &gt; 0.7</td>
<td>7423</td>
<td>3.83</td>
<td>5186</td>
</tr>
<tr>
<td>∆R &lt; 2.2</td>
<td>6549</td>
<td>3.38</td>
<td>4371</td>
</tr>
</tbody>
</table>

### Table 6.4: Number of candidates and selection efficiency on simulated hidden valley signal with a π^0_ν lifetime of 10 ps and a π^0_ν mass of 35 GeV, decaying to either b, c or u/d/s.

<table>
<thead>
<tr>
<th>Selection</th>
<th>HV10_BB</th>
<th>HV10_CC</th>
<th>HV10_SS</th>
</tr>
</thead>
<tbody>
<tr>
<td># cand</td>
<td>%</td>
<td># cand</td>
<td>%</td>
</tr>
<tr>
<td>Generated events</td>
<td>210873</td>
<td>196844</td>
<td>54961</td>
</tr>
<tr>
<td>InTuple</td>
<td>58327</td>
<td>27.66</td>
<td>75108</td>
</tr>
<tr>
<td>L0</td>
<td>44575</td>
<td>21.14</td>
<td>53621</td>
</tr>
<tr>
<td>HLT1</td>
<td>8768</td>
<td>4.16</td>
<td>14478</td>
</tr>
<tr>
<td>HLT2</td>
<td>5221</td>
<td>2.48</td>
<td>9066</td>
</tr>
<tr>
<td>Stripping</td>
<td>634</td>
<td>0.30</td>
<td>2150</td>
</tr>
<tr>
<td>GEC</td>
<td>627</td>
<td>0.30</td>
<td>2136</td>
</tr>
<tr>
<td>MV extended</td>
<td>569</td>
<td>0.27</td>
<td>2046</td>
</tr>
<tr>
<td>Ntracks ≥ 6</td>
<td>514</td>
<td>0.24</td>
<td>1938</td>
</tr>
<tr>
<td>Two Jets</td>
<td>35</td>
<td>0.02</td>
<td>848</td>
</tr>
<tr>
<td>M/M_{corr} &gt; 0.7</td>
<td>27</td>
<td>0.01</td>
<td>760</td>
</tr>
<tr>
<td>∆R &lt; 2.2</td>
<td>26</td>
<td>0.01</td>
<td>713</td>
</tr>
</tbody>
</table>
6.8 Outlook

The event selection is not fine-tuned on a specific model, which is motivated by the idea that a looser selection is sensitive to a range of exotic long lived heavy particles. For example, the long-lived particles are not required to be pair-produced, and the $b$-flavour of the jets is not exploited. Therefore a similar selection can be used to set limits on other models such as the SUSY models discussed in Chapter 1. For a future analysis, a more stringent selection could be applied that might lead to a better sensitivity for particular models. For example, the maximum dijet $\Delta R$ selection used to reject combinatorial background in the horizontal plane could be released to obtain a better efficiency for higher masses. Additionally, the use of a multi-variate selection has been investigated. As long as the selection is ensured to be independent of the dijet mass distribution, it could be possible to improve the current selection efficiency and background rejection by using a multi-variate selection.
CHAPTER 7

Background

There are several sources of background for a displaced vertex signal: long-lived standard model particles, particle interactions with the detector material, material interactions, mistakes in the reconstruction, and combinations thereof. Concerning the long-lived standard model particles, the main contribution originates from weakly decaying charm and beauty hadrons. Strange hadrons are too light to pass the selection criteria. Charm and beauty backgrounds are considered in more detail in the following paragraphs. Errors in the reconstruction enhance the standard model backgrounds, as will be discussed alongside. Detector material interactions and beam-gas interactions are estimated individually.

7.1 Beauty decays

Recalling the discussion in Section 3.2, it is not possible to make a reliable estimate of standard model backgrounds from simulated background events due to the large statistical uncertainty on the simulated samples. Instead, the data will be modelled by a smooth background mass shape. Nevertheless, an effort is made to obtain a rough estimate of the predicted background yields and distributions, in order to gather more information about the content of the observed backgrounds.

The event yields on the simulated inclusive $b\bar{b}$ samples ('INCLB_1DV' and 'INCLB_5PS' with increased lifetimes, as described in Section 3.2.1) are given in Table 7.1, along with an inclusive charm sample ('INCLC'). The number of MC background events surviving the final selection is insufficient to use it for a mass fit to retrieve the number of background and signal events in the limit extraction. Looser
selections, as listed in Table 7.2, are used to get a qualitative understanding of the background. The 'SinglePS' selection is the same as the selection applied in the StrippingDVSinglePS line (see Table 4.8), with additionally the MV and the GEC cuts applied. The 'Stripping' requirement is equal to the one used for the main analysis, again with additional MV and GEC cuts. Requiring two jets in both cases allows for the evaluation of the dijet mass distributions of the candidates. The number of expected background events in the 2011 data set, deduced from the two $b\bar{b}$ samples, is listed in Table 7.3. The mass, number of tracks and radial distributions of candidates with two jets at the SinglePS level are shown in Figure 7.2.

There are two factors that explain the variations in distributions and selected yields between the two MC samples and the data.

Firstly, the INCLB_1DV sample contains events with at least four charged particles in the acceptance, whereas the INCLB_5PS sample requires only one. The first sample excludes events in which up to three $b$-tracks and a number of other particles in the event constitute a candidate, and is therefore not completely representative for the $b\bar{b}$-events seen in the data. This is most likely the reason why the number of selected events from INCLB_1DV is short of the number in data (and in INCLB_5PS). It also explains why the mass distribution of the INCLB_1DV sample is high compared to data, as shown in Fig. 7.2a.

Secondly, the average lifetime of the $b$-hadrons in INCLB_5PS is increased to 5 ps, which implies that each event gets a weight factor assigned, enabling the retrieval of the correct lifetime distribution. Events with a long lifetime get a low weight, since they are unlikely to occur in real data. As the selection favours the long-lived candidates (see Fig. 7.1), the selected sample consists mainly of events with a low weight, on which the

<table>
<thead>
<tr>
<th>Selection</th>
<th>INCLB_5PS</th>
<th>INCLB_1DV</th>
<th>INCLC</th>
</tr>
</thead>
<tbody>
<tr>
<td>Generated events</td>
<td>7,055,964</td>
<td>2,533,490</td>
<td>9,148,460</td>
</tr>
<tr>
<td>InTuple</td>
<td>723,287</td>
<td>262,994</td>
<td>49,746</td>
</tr>
<tr>
<td>L0</td>
<td>105,072</td>
<td>33,844</td>
<td>6,242</td>
</tr>
<tr>
<td>HLT1</td>
<td>75,037</td>
<td>23,534</td>
<td>3,903</td>
</tr>
<tr>
<td>HLT2</td>
<td>19,138</td>
<td>7,015</td>
<td>1,098</td>
</tr>
<tr>
<td>Stripping</td>
<td>1,017</td>
<td>351</td>
<td>39</td>
</tr>
<tr>
<td>GEC</td>
<td>1,017</td>
<td>351</td>
<td>39</td>
</tr>
<tr>
<td>MV extended</td>
<td>997</td>
<td>351</td>
<td>39</td>
</tr>
<tr>
<td>Ntracks $\geq$ 6</td>
<td>752</td>
<td>309</td>
<td>36</td>
</tr>
<tr>
<td>Two Jets</td>
<td>50</td>
<td>8</td>
<td>2</td>
</tr>
<tr>
<td>$M/M_{\text{corr}} &gt; 0.7$</td>
<td>28</td>
<td>4</td>
<td>1</td>
</tr>
<tr>
<td>$\Delta R &lt; 2.2$</td>
<td>27</td>
<td>4</td>
<td>1</td>
</tr>
</tbody>
</table>

Table 7.1: Number of selected candidates in background $b\bar{b}$ and $c\bar{c}$ MC models.
Table 7.2: Number of selected events in background MC before reweighting. The INCLC sample contains both $c\bar{c}$ and $b\bar{b}$ events, of which the pure $c\bar{c}$ events are selected for the last column. The last two rows give the generator level efficiency (detector acceptance) and the resulting scale factor required to derive the equivalent yields for $0.62 \text{ fb}^{-1}$ of data. (*) For the INCLB_5PS sample, a per-event weighting is needed to correct for the lifetime.

<table>
<thead>
<tr>
<th>Selection</th>
<th>INCLB_5PS</th>
<th>INCLB_1DV</th>
<th>INCLC</th>
<th>INCLC</th>
</tr>
</thead>
<tbody>
<tr>
<td>Generated events</td>
<td>7.1M</td>
<td>2.5M</td>
<td>9.1M</td>
<td>8.1M</td>
</tr>
<tr>
<td>SinglePS</td>
<td>14,798</td>
<td>6,493</td>
<td>1,006</td>
<td>3</td>
</tr>
<tr>
<td>SinglePS+2jets</td>
<td>1219</td>
<td>307</td>
<td>67</td>
<td>0</td>
</tr>
<tr>
<td>Stripping</td>
<td>997</td>
<td>351</td>
<td>39</td>
<td>0</td>
</tr>
<tr>
<td>Stripping+2jets</td>
<td>76</td>
<td>12</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>Final selection</td>
<td>27</td>
<td>4</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>$\varepsilon_{\text{gen}}$</td>
<td>0.311</td>
<td>0.0482</td>
<td>0.253</td>
<td>0.253</td>
</tr>
<tr>
<td>scale-factor (0.62 $\text{fb}^{-1}$)</td>
<td>(*)</td>
<td>$(3.5 \pm 0.5) \cdot 10^{3}$</td>
<td>$(8.9 \pm 0.3) \cdot 10^{4}$</td>
<td>$(10.1 \pm 0.3) \cdot 10^{4}$</td>
</tr>
</tbody>
</table>

Table 7.3: Number of expected background events in $0.62 \text{ fb}^{-1}$ and the number of events found in data. The SinglePS events in data are corrected for the 0.005 postscale in the SinglePS stripping line. The INCLB_5PS event yields are corrected for the lifetime reweighting.

<table>
<thead>
<tr>
<th>Selection</th>
<th>2011 data</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Simulation</td>
</tr>
<tr>
<td></td>
<td>INCLB_5PS</td>
</tr>
<tr>
<td>SinglePS</td>
<td>69M ± 1.0M</td>
</tr>
<tr>
<td>SinglePS+2jets</td>
<td>5.0M ± 278k</td>
</tr>
<tr>
<td>Stripping</td>
<td>2.3M ± 153k</td>
</tr>
<tr>
<td>Stripping+2jets</td>
<td>133k ± 33k</td>
</tr>
<tr>
<td>Final selection</td>
<td>28k ± 12k</td>
</tr>
</tbody>
</table>
uncertainty is relatively large. This explains why, even though the absolute number of events is larger, the uncertainty on the final selected INCLB_5PS events is higher than on INCLB_1DV events (see Tables 7.2 and 7.3). It is not clear, however, why the total number of predicted events from INCLB_5PS is significantly higher than the number of events observed in data, mainly at the looser selection levels. This difference is still under investigation, but does not affect the results of the analysis.

Figure 7.1: Number of reconstructed candidates as a function of $R_{xy}$ at various selection stages, on 1 million events from the background samples (a) 'INCLB_1DV' and (b) 'INCLB_5PS'. The grey histograms show the distributions of generated events.

A study of the track content of $b\bar{b}$ candidates passing the SinglePS selection indicates that most tracks originate from beauty and charm decay vertices. The pollution from primary vertex tracks is about one percent. For this loose selection about 4% of the candidates in INCLB_1DV background sample have a displaced vertex with tracks from two different $b$-hadrons in the event. There is an additional contribution of approximately 10% from tracks that cannot be associated to any MC particle, which might be ghost tracks (misreconstructed tracks). They can lead to the high mass required to pass the candidate selection. The conclusion is that there is no single source that accounts for the background events; they result from a combination of $b$-hadron tracks with tracks from the other $b$-hadron, ghost tracks and other tracks in the event.
Figure 7.2: Distributions of (a)(b) mass, (c)(d) number of tracks and (e)(f) $R_{xy}$ of reconstructed displaced vertices in data, compared to (left) INCLB_1DV background, and to (right) INCLB_5PS background. The MV and GEC are applied, as well as the triggers, SinglePS stripping and the dijet requirement. The number of events in MC is scaled to the number of expected events in $0.62 \, fb^{-1}$. The number of data events is corrected for the prescale factor 0.005.
7.2 Charm decays

The 'INCLC' simulated sample contains events that have at least two $c$-hadrons in the acceptance. This includes the case in which the $c$-hadron originates from a decaying $b$-hadron (non-prompt charm), which constitute about 11% of the INCLC sample. In 4% of the events both a $c\bar{c}$ and a $b\bar{b}$ pair are produced in the primary interaction, and in 7% there is only a $b\bar{b}$ pair produced. The remaining 89% of the generated events contain only $c\bar{c}$. It turns out that the long-lived particle selection favours the $b$-hadron decays. The fraction of events containing only a $c\bar{c}$ at the primary interaction decreases from 89% at generator level, to 0.3% at the SinglePS selection level, and to zero at the final selection, as shown in Table 7.2. The remaining candidates come either from events with both $c\bar{c}$ and $b\bar{b}$ produced (25 – 50%, depending on the selection level), or from events containing only $b\bar{b}$ (75 – 50%). Apparently, the prompt $c\bar{c}$ events do not feature particles that are heavy enough and live long enough to pass the selection.

Even though the statistics are poor, an estimate can be made for the number of charm events surviving the full selection. Starting from about 8 million pure $c\bar{c}$ events (after excluding the 11% $b$-events from the 9 million generated events), the SinglePS selection has three events left. This corresponds to $307k \pm 177k$ expected events in 0.62 fb$^{-1}$ of data, which is calculated using the scale factor in Table 7.2. Comparing this number to the yields from INCLB_1DV and INCLB_5PS in Table 7.3, it can be concluded that the number of charm events surviving the SinglePS selection is $\sim 0.5\%$ of the number of beauty events. When assuming that $c\bar{c}$ has the same probability to survive the final selection as $b\bar{b}$, only $281 \pm 163$ charm candidates would remain.

To conclude, since a sample of 8 million prompt $c\bar{c}$ events does not produce a single candidate that survives the final selection, only a rough estimate can be made for the final yield. This estimate indicates that the prompt $c\bar{c}$ background in data is negligible compared to the background from beauty decays.

Considering the large uncertainties, the agreement between the data and the prediction from the INCLB_5PS sample at the final selection level is reasonable. Some inconsistencies remain because the MC samples are not completely representative for the selected data. The MC background estimates will not be used to obtain the result of this analysis. Events involving beauty decays are found to be the only significantly contributing standard model background to the long-lived particle analysis.

7.3 Beam-gas and beam-beam interactions

Proton-proton collisions occur inside the region $R_{xy} < 0.4$ mm, which is excluded by the displaced vertex candidate selection. This is supported by Fig. 4.4 in Section 4.3.3. However, a displaced vertex signature can occur due to beam-beam, beam-gas or beam halo interactions.
Firstly, although the proton-proton collision region is kept in vacuum, there is still some residual gas, which might cause interactions. Such interactions can take place between particles originating in the proton-proton interaction and the gas (so-called 'particle-gas' interactions). The particle-gas interaction rate has been estimated on MC for the analysis on 2010 data, and was found to be negligible \[56\].

Secondly, similar interactions can occur between one of the beams and the gas ('beam-gas' interactions). This type of interaction is not well represented by the MC. The beam-profile in Fig. 7.3 shows the 'beam-empty' crossings in blue. In these events, only one beam crosses the interaction point, while the bunch of the other beam is empty, such that the only possible interactions are between the beam and gas in the detector. The distribution of reconstructed vertices in those events drops fast with \(R_{xy}\), defining the profile of the beam. The contribution of beam-gas interactions to displaced vertex events is expected to be negligible, because the number of these vertices above \(R_{xy} > 0.4\) mm is zero. In principle, the beam-gas interaction yield could be extracted by looking at the selected displaced vertex event yield in 'beam-empty' bunch-crossings. However, in order to select long-lived particle candidate events, a primary vertex is required to be upstream of the signal candidate, in both the trigger and the stripping selections. Such a primary vertex is absent in beam-gas events, since there is only one beam crossing. Due to these trigger and stripping selections, no beam-gas interactions can be retrieved from the selected events.

Thirdly, despite the continuous collimation of the beams, there is a 'beam halo' around the beams due to beam losses. Interactions of particles in the beam halos of the two beams might interact ('beam-beam' interactions), leading to a displaced vertex signature. These interactions are taken into account in the simulation of the

\[\text{Figure 7.3: Distributions of the } x \text{ and } y \text{ coordinates of reconstructed primary vertices in beam-empty (blue) and empty-empty (green) beam-crossing events. Empty-empty vertices are created by residual gas in the empty bunches, and spill-over from the neighbouring bunches. From [106].}\]
data. Beam-beam interaction vertices are efficiently removed by the requirement to have no backward tracks in the vertex.

All in all, the beam interactions are expected to be negligible compared to the background from beauty decays.

### 7.4 Material interactions

The veto for material interactions is described in Section 4.6. Since the reconstructed vertex position resolution is not perfect, some material interactions may escape the material veto. The yield of these remaining interactions is estimated using a sample of postscaled events for which the MV was not applied in trigger or stripping. The lines used are Hlt2SinglePSLonglived and SingleHLTPSDV, described in Sections 4.3.3 and 4.5.2. The total postscale factor applied on these lines is \( f_{PS} = 0.0001 \). The material veto is only defined outside a radius \( R_{xy} > 5.0 \text{ mm} \). Table 7.4 shows the event yields in this sample for events with \( R_{xy} > 4.8 \text{ mm} \) inside and outside the material (as determined by the MV) per selection stage.

By conservatively assuming that all events with \( R_{xy} > 4.8 \text{ mm} \) before the selection are in fact material interactions, the probability that the MV by mistake misses a material interaction is at most \( \eta_{MV} = 121 / (20233 + 121) = (5.9 \pm 0.5) \times 10^{-3} \). The probability for a material interaction to pass the final selection criteria is estimated to be: \( \varepsilon_{sel}^{\text{material}} = 22 / 20233 = (1.1 \pm 0.2) \times 10^{-3} \). Therefore, the total yield of material interactions in the final sample is estimated to be smaller than:

\[
N_{\text{material}} = \frac{1}{f_{PS}} \times \varepsilon_{sel}^{\text{material}} \times \eta_{MV} = 0.064 \pm 0.015 \quad (7.1)
\]

This estimate is made from the loose 'SinglePS' selection. In the usual stripping selection for long-lived particle candidates, more stringent selections are applied on the invariant mass, sum-\( p_T \) and the number of tracks. Therefore the actual yield of material candidates can be expected to be even smaller than the one quoted in Eq. 7.1.

<table>
<thead>
<tr>
<th>Inside material</th>
<th>Outside material</th>
</tr>
</thead>
<tbody>
<tr>
<td>SinglePS</td>
<td>20411</td>
</tr>
<tr>
<td>GEC</td>
<td>20233</td>
</tr>
<tr>
<td>( R_{xy} &gt; 4.8 \text{ mm} )</td>
<td>20233</td>
</tr>
<tr>
<td>Two Jets</td>
<td>47</td>
</tr>
<tr>
<td>( M/M_{corr} &gt; 0.7 )</td>
<td>24</td>
</tr>
<tr>
<td>( \Delta R &lt; 2.2 )</td>
<td>22</td>
</tr>
</tbody>
</table>

*Table 7.4: Number of events in data surviving various selections, inside the MV region and outside the MV region. The events are retrieved from the Hlt2SinglePSLonglived trigger and SingleHLTPSDV stripping lines.*
The fraction of material interactions in the final sample must therefore be negligible.

7.5 Outlook

The current trigger and stripping selections both require a primary vertex in the interaction region that is upstream of the signal candidate. This poses some difficulties when estimating the background from beam-gas interactions, as discussed in Section 7.3. Furthermore, it is possible that the primary vertex algorithm fails to reconstruct a primary vertex. This can happen for example if a primary vertex does not pass the minimal requirement of four tracks. In that case, an additional contribution from primary vertex tracks can enter the jets and the displaced vertex. In order to be able to estimate this background, future trigger and stripping selections do not require the presence of a primary vertex upstream of the displaced vertex. Furthermore, to exclude the background of these primary vertex tracks, all tracks pointing towards the interaction region as opposed to the primary vertices are excluded from both the displaced vertex and the jet search.
Systematic uncertainties

The systematic uncertainties of the long-lived particle search can be divided into three categories: the uncertainty on the selection efficiency, the uncertainty on the estimated signal yield from the background and the signal mass shapes in the fit for the limit extraction, and the uncertainty on the luminosity of the data sample.

The selection efficiency is influenced by the reconstruction and selection procedures in trigger, stripping, jet reconstruction, etc. The efficiency is solely determined from simulation, and the main task is therefore to demonstrate that the simulation correctly describes the detector performance. Differences between data and simulation arise because the MC can be inaccurate in the description of detector geometry (e.g. alignment and material position), of the detector response (e.g. single-hit efficiency and hit resolution) and of physical and machine related backgrounds (e.g. pile-up). The uncertainties on the selection efficiency, retrieved in the following sections, are summarised in Table 8.1.

To compare data and MC, one would need to select events with similar characteristics as the signal candidates (a high dijet mass, high vertex track multiplicity and a large displacement), independently of the signal sample. Unfortunately, there is no such sample available in both data and MC. Therefore several samples with either one of those characteristics are used to determine the differences between data and MC, and the results are extrapolated to the signal region.
### 8.1 Jet reconstruction and selection

There are several reasons why the systematic uncertainties on the jet reconstruction efficiency and resolution as used in this analysis cannot be extracted from existing analyses with jets in LHCb (e.g. [102]). The long-lived particle jet reconstruction algorithm is not the one used for the default LHCb jet reconstruction. Different input particles are used, the jets are more displaced, and the clustering uses a different cone size. Furthermore, the jet identification selection is different.

The systematic uncertainties related to the jets are obtained using two control channels, namely Z + jet events (Section 8.1.1) and inclusive $b\bar{b}$ events (Section 8.1.5). The jet selection efficiency can be divided into two parts: the jet reconstruction from the tracks and neutral constituents in the event (Section 8.1.2) and the selection on the jet ID variables that follows afterwards (Section 8.1.3). The uncertainty on the jet direction resolution contributes to the systematic uncertainty of the candidate selection, which enters when requirements are made on the pointing or the opening angle of jet pairs (Section 8.1.4).

#### 8.1.1 Introduction: Monte Carlo jet validation with Z + jet events

The Z + jet samples in data and in MC are used for most of the studies involving jets. This section describes how to retrieve a suitable sample to evaluate different systematic uncertainties for the jets. The jets in Z + jet events originate from the primary interaction point, rather than from a displaced vertex. Therefore, a dedicated 'prompt' version of the displaced jet algorithm is developed, which uses primary vertices in-
instead of displaced vertices to group the input particles for the jet finding. The tracks pointing to other primary vertices in the event are vetoed from the jet input tracks list, using the same requirements that are used to veto tracks from primary vertices in the DV jet reconstruction. The extrapolation from prompt to displaced jets relies on Monte Carlo.

The signal MC sample consists of 1 million generated $Z \rightarrow \mu\mu$ events, for which one lepton with $p_T > 4$ GeV is required at generator level. The data consists of events from the electroweak stream. Both samples have to pass the 'Z02MuMuLine' stripping line, after which the following selection criteria are applied to the $Z \rightarrow \mu\mu$ candidates, selecting approximately 25000 events in data and 80000 events in MC:

- both muons have $p_T(\mu) > 20$ GeV, track $\chi^2$ probability $> 0.001$, and are in the acceptance $2.0 < \eta(\mu) < 4.5$
- the $Z$ candidate has $p_T(Z) > 5$ GeV and $60 < m(Z) < 120$ GeV.

The kinematic distributions in $\eta$, $\phi$, and $p_T$ for the selected $Z$ candidates in data and simulation in Fig. 8.1 illustrate that the distributions are similar in data and MC. The $Z$ mass distribution in data is slightly worse, which is probably due to an overestimation of the resolution of the muon tracks in the simulation.

The jets are reconstructed with a minimum $p_T$ of 5 GeV, and they are required to originate from the same primary vertex as the $Z$ candidate. Subsequently, the jets must pass the following selection criteria:

- $\Delta R(\mu, jet) > 0.7$ for both muons, to isolate the jets from the muons;
- $2.0 < \eta(jet) < 4.5$, such that the jets are in the acceptance.

The jet multiplicity of in the $Z$ events is shown in Fig. 8.2a. The hadronic processes that produce jets are difficult to model in MC, which is why it is not surprising that the distributions in data and MC before selection do not have the same jet multiplicities. Detailed information about the jet multiplicity in $Z + \text{jet}$ events in LHCb can be found in [102].
Figure 8.1: The comparison of the Z distributions for (a) $\eta$, (b) $\phi$, (c) $p_T$ and (d) mass in data and MC. For all Z-candidates with any number of jets. All distributions are normalised to unit area.
Figure 8.2: (a) Jet multiplicity of all Z candidates, in data and MC. (b) Multiplicity of associated jets that have passed the $|\Delta \phi(Z, \text{jet})| > \frac{3}{4} \pi$ selection. All distributions are normalised to unit area.

Figure 8.3: (a) $\Delta \phi$ distribution for Z-jet combinations, in data and MC. (b) $\phi$ of all Z-candidates that have at least one jet passing the $\Delta \phi$ selection. All distributions are normalised to unit area.
All previously mentioned selections are applied for each study of jet systematics. In addition, a selection is applied depending on the purpose of individual studies. The selection criteria are explained below, and Table 8.2 summarises which selection criteria are used for which study.

<table>
<thead>
<tr>
<th>Selection:</th>
<th>Study:</th>
</tr>
</thead>
<tbody>
<tr>
<td>$</td>
<td>\Delta \phi(Z,\text{jet})</td>
</tr>
<tr>
<td>$\frac{1}{2} \pi &lt;</td>
<td>\phi(\text{jet})</td>
</tr>
<tr>
<td>$p_T \text{ balance: } 0.5 &lt; p_T(\text{jet})/p_T(Z) &lt; 1.5$</td>
<td>X</td>
</tr>
<tr>
<td>$p_T(\text{second jet})/p_T(\text{leading jet}) &lt; 0.25$</td>
<td>X</td>
</tr>
<tr>
<td>leading jet passes jet ID</td>
<td>X</td>
</tr>
</tbody>
</table>

*Table 8.2: Selection criteria applied to the $Z + \text{jet}$ sample, as they are applied in the individual systematic studies presented below.*

**$\Delta \phi(Z,\text{jet})$**

Because of momentum conservation, the $Z$ and the jet are expected to have a 'back-to-back' topology. This is supported by Fig. 8.3a, which shows that the azimuthal angle between the $Z$ and the jet peaks at $\pi$. The same feature is visible in Fig. 8.4a, which shows the distribution of the azimuthal angle of the $Z$ versus that of the jet. The diagonal correlation correspond to 'back-to-back' $Z + \text{jet}$ candidates. The vertical correlation corresponds to (mainly soft) jets reconstructed in the horizontal plane. The back-to-back topology is better visible when selecting only high $p_T$ jet-candidates, in Fig. 8.4b. To select a sample in which all the candidates are fully 'back-to-back', the selection $|\Delta \phi(Z,\text{jet})| > \frac{3}{4} \pi$ is used. This selection results in the jet multiplicity distribution shown in Fig. 8.2b. Most $Z$ candidate events contain no jet, and the fraction of candidates with more than one jet is small.

**$\phi(\text{jet})$**

Figure 8.4a does not only show diagonal bands, but also vertical ones. There is a pollution at low $p_T$ around $\phi(\text{jet}) = [0, \pi]$, mainly due to ghost tracks in the outer tracker detector. It was shown in Fig. 8.1b that the $\phi$ distribution for the $Z$ is flat, so the $\phi$ dependency must be introduced by the jet requirement. This feature appears mainly at low jet $p_T$, and the reconstructed jets are of low quality. The low-quality jets in the horizontal plane can be eliminated by the selection $\frac{1}{4} \pi < |\phi(\text{jet})| < \frac{3}{4} \pi$. Since this selection criterion removes a large number of events, it is only used in the L0 hadron trigger efficiency study.
Figure 8.4: Distribution of the azimuthal angles $\phi(Z)$ versus $\phi(jet)$, for Z-candidates with at least one jet in MC. (a) lowest $p_T(Z)$ bin, with $5 < p_T(Z) < 10$ GeV, (b) highest $p_T(Z)$ bin, with $p_T(Z) > 30$ GeV. The distributions in data have similar characteristics.

Figure 8.5: The ‘$p_T$ balance’ $p_T(jet)/p_T(Z)$ of Z + jet candidates in data and in MC. The distributions are normalised to unit area.
$p_T$ balance

Momentum conservation also causes the transverse momenta of the $Z$ and of the jets to be correlated. The ratio $p_T(\text{jet})/p_T(Z)$ is illustrated in Fig. 8.5. A selection can be made on the '\(p_T\) balance' of the $Z$ and the jet as follows: $0.5 < p_T(\text{jet})/p_T(Z) < 1.5$.

$p_T$ ratio jets

Only those candidates with little other activity than the $Z + \text{jet}$ are considered for jet studies. Besides the single-jet candidates, multi-jet candidates with a transverse momentum ratio between the leading jets: $p_T(\text{second jet})/p_T(\text{leading jet}) < 0.25$ are selected. Figure 8.6 illustrates the distribution of the $p_T$ ratio of multi-jet events. In practice, the effect of this criterion is that only events a leading jet with $p_T > 20 \text{ GeV}$ are considered, since the minimum $p_T$ of the second jet is 5 GeV. From those events, one selects only the events in which the $p_T$ of the second jet is small. A typical distribution of the ratio in events with a leading jet with $p_T > 30 \text{ GeV}$ is shown in Fig. 8.6b. Furthermore, since the $Z$ candidates with lower $p_T$ usually have only a single jet, this selection mostly influences the highest $p_T$ sample.

Figure 8.6: The $p_T$ ratio $p_T(\text{second jet})/p_T(\text{leading jet})$ of $Z + \text{jet}$ candidates with at least two jets in data and in MC. (a) all events, and (b) events with $p_T(\text{leading jet}) > 30 \text{ GeV}$. The distributions are normalised to unit area. Events with $p_T(\text{second jet})/p_T(\text{leading jet}) < 0.25$ are selected.
Jet ID

Good quality jets are defined by the jet identification selection that is also applied to the $\pi^0$ signal candidates, as discussed in Section 5.3.

8.1.2 Jet reconstruction efficiency using Z + jet events

The main source of jet reconstruction inefficiencies is the minimum $p_T$ threshold of 5 GeV. Any uncertainties on the efficiency of this selection are caused by the jet energy resolution difference in data and MC.

![Figure 8.7: Comparison of the mean of the $p_T(jet)/p_T(Z)$ distributions in Z + jet data and MC, as a function of $p_T(jet)$. The table shows the ratio of the two means, in bins of transverse momentum.](image)

<table>
<thead>
<tr>
<th>$p_T(jet)$</th>
<th>$\mu_{data}/\mu_{MC}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>5-10</td>
<td>1.0176 ± 0.0122</td>
</tr>
<tr>
<td>10-15</td>
<td>0.9884 ± 0.0169</td>
</tr>
<tr>
<td>15-20</td>
<td>1.0245 ± 0.0249</td>
</tr>
<tr>
<td>20-25</td>
<td>1.0272 ± 0.0283</td>
</tr>
<tr>
<td>25-30</td>
<td>1.0186 ± 0.0290</td>
</tr>
<tr>
<td>&gt;30</td>
<td>1.0476 ± 0.0161</td>
</tr>
<tr>
<td>all bins</td>
<td>0.9996 ± 0.0082</td>
</tr>
</tbody>
</table>

The energy resolution is determined by the detector response, which needs to be correctly simulated in MC. In order to verify that this is the case, 'back-to-back' Z + jet candidates can be used to determine the accuracy of the energy scale of the jets, since the Z and the jet should balance in $p_T$. This sample of 'back-to-back' Z + jet candidates is selected as described in the category 'Jet energy' in Table 8.2.

The mean of $p_T(jet)/p_T(Z)$ of selected candidates is given as a function of jet $p_T$ in Figure 8.7. Due to the definition of the mean, this distribution increases with jet $p_T$. The overall agreement between data and MC is satisfactory, and a systematic uncertainty is assigned using the ratio of the mean of the distributions in data and MC. The ratio of the means of the distributions in data and in MC: $r = \mu_{data}/\mu_{MC}$ is listed as a function of $p_T$ in the accompanying table.

As a cross-check to see if the data and MC samples consist of clean Z + jet candidates, the additional selection: $\frac{1}{4}\pi < |\phi(jet)| < \frac{3}{4}\pi$ is applied, which selects only jets outside the horizontal plane. This supplementary selection does not significantly
change the ratio of data and MC.

In order to retrieve the uncertainty on the jet reconstruction efficiency, the jet $p_T$ threshold is increased by a conservative 2\% (the difference from unity of the ratio of the means $p_T(\text{jet})/p_T(Z)$, in the lowest $p_T$ bins in the table in Fig. 8.7). The efficiency loss of candidates passing the final selection on the hidden valley samples is given in Table 8.3. The resulting systematic uncertainties, obtained from the difference with unity summed in quadrature with the statistical error, are summarised in Table 8.1. The uncertainty on the jet energy scale also affects the mass resolution, which is discussed in Section 8.5.

![Table 8.3: Relative change in efficiency of the jet $p_T$ selection on offline selected candidates as a result of increasing the minimal jet $p_T$ requirement by 2\%. For various simulated hidden valley signal masses and lifetimes.](/images/document/130.png)

<table>
<thead>
<tr>
<th>$\pi^0$ mass [GeV]</th>
<th>10 ps</th>
<th>100 ps</th>
</tr>
</thead>
<tbody>
<tr>
<td>15</td>
<td>0.940 ± 0.032</td>
<td>1.000 ± 0.039</td>
</tr>
<tr>
<td>25</td>
<td>0.991 ± 0.002</td>
<td>0.990 ± 0.004</td>
</tr>
<tr>
<td>35</td>
<td>0.995 ± 0.001</td>
<td>0.998 ± 0.003</td>
</tr>
<tr>
<td>43</td>
<td>0.998 ± 0.001</td>
<td>0.998 ± 0.001</td>
</tr>
<tr>
<td>50</td>
<td>0.998 ± 0.001</td>
<td>0.998 ± 0.001</td>
</tr>
</tbody>
</table>

8.1.3 Jet ID using Z + jet events

Figure 8.8 shows the distributions of several jet ID variables as defined in Section 5.3 of the jets of selected 'back-to-back' Z candidates, indicating a good agreement between data and MC.

<table>
<thead>
<tr>
<th>cut</th>
<th>Only this cut</th>
<th>This cut last</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$\text{eff}_{MC}$</td>
<td>$\text{eff}_{Data}$</td>
<td>$\text{eff}_{MC}$</td>
<td>$\text{eff}_{Data}$</td>
</tr>
<tr>
<td>MTF</td>
<td>99.10 ± 0.06</td>
<td>98.91 ± 0.15</td>
<td>98.95 ± 0.07</td>
<td>98.70 ± 0.17</td>
</tr>
<tr>
<td>CPF</td>
<td>93.42 ± 0.16</td>
<td>91.99 ± 0.38</td>
<td>99.51 ± 0.05</td>
<td>99.37 ± 0.12</td>
</tr>
<tr>
<td>MPT</td>
<td>85.53 ± 0.23</td>
<td>84.62 ± 0.50</td>
<td>91.02 ± 0.20</td>
<td>91.32 ± 0.41</td>
</tr>
<tr>
<td>all</td>
<td>84.22 ± 0.24</td>
<td>83.01 ± 0.52</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

Table 8.4: Efficiencies of jet ID selection criteria on jets associated to the Z candidates in the Z + jet samples in data and simulation.

The efficiencies of the jet ID selections on the Z + jet events, given in Table 8.4, show that the agreement between data and MC is within 1\%. The MPT and CPF variables are correlated: the CPF selection is almost fully efficient if applied after the MPT selection.

Figure 8.9 shows the efficiencies of the three jet ID selections as a function of transverse momentum, illustrating that the efficiencies are a strong function of $p_T$. To ac-
Figure 8.8: Jet ID properties of jets associated to the Z candidates in the Z + jet samples in data and MC simulation before the jet ID selection is applied. The variables MTF, CPF, MPT and N90 are defined in Section 5.3.
Figure 8.9: Efficiency in data and MC of the jet ID selection (MTF, CPF and MPT) on the Z + jet sample.

Table 8.5: Estimated efficiency correction factor $\epsilon_{data}/\epsilon_{MC}$ for the jet ID selection for $\pi_{\nu}$ candidates in different HV models estimated by extrapolating from the Z + jet sample.
count for this dependence, the efficiency ratio is evaluated in bins of jet $p_T$.

The efficiency correction factor for signal is estimated by integrating the $p_T$-dependent efficiency factors over the jet $p_T$ distribution in signal MC. Since each candidate contains two jets, the efficiency factors of the jets must be multiplied. When labelling the $p_T$ bins with an index $i$ for the first jet and an index $j$ for the second jet, the efficiency correction factor for hidden valley (HV) signal can be evaluated as:

$$
\langle r^\text{HV} \rangle = \sum_{i,j \leq i} f_{ij} \cdot r_{i}^{Z+\text{jet}} \cdot r_{j}^{Z+\text{jet}}
$$

(8.1)

where $f_{ij}$ is the fraction of signal events in bin $i,j$. This method will also be used to evaluate the other jet systematics.

The resulting estimated efficiency correction factors for signal are shown in Table 8.5. The reported uncertainties reflect the statistical uncertainties in the control channel in data and MC and in the signal hidden valley samples. All the scale factors are consistent with unity. The difference with unity summed in quadrature with the statistical error is used as an estimate of the jet ID efficiency uncertainty, and listed in Table 8.1.

### 8.1.4 Jet direction resolution using Z + jet events

The direction resolution of the jets influences the efficiency of selections made on the pointing and the opening angle of the dijet candidates. To estimate the uncertainty from the direction resolution, the 'back-to-back' Z + jet candidates are used, selected as described by 'Jet direction' in Table 8.2.

The direction resolution of the remaining jets is retrieved from a Gaussian fit to the $\Delta \phi$ between the $Z$ and the jet, which is shown in Fig. 8.10. The fit is parameterised as: $a + b \cdot \exp\left[-\frac{(x-\bar{x})^2}{2\sigma^2}\right]$. The resolutions estimated from the width of the Gaussian are shown in Table 8.6 as a function of $p_T(\text{jet})$. By assuming that the difference in the $\phi$ resolution in data and MC is a result of a difference in the resolutions of the slopes $t_x \equiv p_x/p_z$ and $t_y \equiv p_y/p_z$, the effects of an inconsistency in the slope resolution can be simultaneously assessed on both $\eta$ and $\phi$.

The $\phi$ and $\eta$ resolutions in the signal MC are similar, and never exceed 0.5, as shown in Fig. 8.11. For $p_T > 20$ GeV they are around 0.1, which is better than the width of the $\Delta \phi$ distribution in the Z + jet sample. It is unlikely that the difference in the $\Delta \phi$ width in data and MC in the Z + jet sample is only the result of a difference in jet resolution. The LHCbZ + jet analysis gives an explanation why the MC overestimates the number of events in which the jet and the $Z$ are produced back-to-back: The MC relies on a parton shower to produce jets around the hardest parton in the event. The parton shower is only accurate in the collinear approximation, and it underestimates jets which are produced at large separations from the hardest jet.
Figure 8.10: $\Delta \phi$ distributions in data (black) and MC (grey). The fit is parameterised as a Gaussian plus a constant.

Figure 8.11: Jet $\phi$ and $\eta$ resolution versus the reconstructed jet $p_T$ in the HV10_M35 sample before (black) and after (red) smearing the jet slopes by 30%. The final selection was applied apart from the cuts on $m/m_{\text{corr}}$ and $\Delta R$.

<table>
<thead>
<tr>
<th>$p_T$(jet)</th>
<th>$\sigma_{\text{data}}$</th>
<th>$\sigma_{\text{MC}}$</th>
<th>$\sigma_{\text{data}}/\sigma_{\text{MC}}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>5-10</td>
<td>0.742 ± 0.037</td>
<td>0.593 ± 0.012</td>
<td>1.25 ± 0.07</td>
</tr>
<tr>
<td>10-15</td>
<td>0.465 ± 0.021</td>
<td>0.465 ± 0.010</td>
<td>1.00 ± 0.05</td>
</tr>
<tr>
<td>15-20</td>
<td>0.344 ± 0.022</td>
<td>0.372 ± 0.010</td>
<td>0.92 ± 0.06</td>
</tr>
<tr>
<td>20-25</td>
<td>0.335 ± 0.027</td>
<td>0.323 ± 0.010</td>
<td>1.039 ± 0.09</td>
</tr>
<tr>
<td>25-30</td>
<td>0.268 ± 0.020</td>
<td>0.285 ± 0.011</td>
<td>0.94 ± 0.08</td>
</tr>
<tr>
<td>&gt;30</td>
<td>0.252 ± 0.011</td>
<td>0.237 ± 0.005</td>
<td>1.06 ± 0.05</td>
</tr>
<tr>
<td>all bins</td>
<td>0.429 ± 0.012</td>
<td>0.396 ± 0.004</td>
<td>1.0822 ± 0.0321</td>
</tr>
</tbody>
</table>

Table 8.6: $\Delta \phi$ distributions in data (black) and MC (grey). The fit is parameterised as a Gaussian plus a constant. The resolutions extracted from the fit in Fig. 8.10, in bins of jet $p_T$. 
The hardest jet is usually produced back to back with the $Z$ boson, but may not be in the LHCb acceptance. Jets that are not back to back with the $Z$ boson will usually be produced in the parton shower, and the number of these is likely to be underestimated \[107\]. Nonetheless, the ratio of the widths in data and MC on that sample is used to set the scale for the jet direction resolution. Since the largest ratio found in Table 8.6 is an approximate 30\%, the $\phi$ resolution in signal MC is smeared by 30\%. The resulting change in efficiency of the selection $m/m_{\text{corr}} > 0.7$ is given in Table 8.7, and of the selection $\Delta R < 2.2$ in Table 8.8. Apart from the 15 GeV sample, which suffers from small statistics, the results are all within 5\% from unity and independent of the lifetime. The efficiency ratios for the two selection criteria are multiplied to get the systematic uncertainty on the jet direction resolution for the signal samples, as reported in Table 8.1.

<table>
<thead>
<tr>
<th>$\pi^0$ mass [GeV]</th>
<th>10 ps</th>
<th>100 ps</th>
</tr>
</thead>
<tbody>
<tr>
<td>15</td>
<td>$1.0015 \pm 0.0143$</td>
<td>$1.0000 \pm 0.0392$</td>
</tr>
<tr>
<td>25</td>
<td>$0.9970 \pm 0.0010$</td>
<td>$0.9986 \pm 0.0020$</td>
</tr>
<tr>
<td>35</td>
<td>$0.9948 \pm 0.0009$</td>
<td>$0.9882 \pm 0.0054$</td>
</tr>
<tr>
<td>43</td>
<td>$0.9933 \pm 0.0010$</td>
<td>$0.9919 \pm 0.0022$</td>
</tr>
<tr>
<td>50</td>
<td>$0.9964 \pm 0.0016$</td>
<td>$0.9940 \pm 0.0023$</td>
</tr>
</tbody>
</table>

**Table 8.7:** Relative change in efficiency of the $m/m_{\text{corr}} > 0.7$ selection as a result of the smearing of the jet direction as described in the text.

<table>
<thead>
<tr>
<th>$\pi^0$ mass [GeV]</th>
<th>10 ps</th>
<th>100 ps</th>
</tr>
</thead>
<tbody>
<tr>
<td>15</td>
<td>$0.9015 \pm 0.0394$</td>
<td>$0.9385 \pm 0.0613$</td>
</tr>
<tr>
<td>25</td>
<td>$0.9543 \pm 0.0037$</td>
<td>$0.9522 \pm 0.0081$</td>
</tr>
<tr>
<td>35</td>
<td>$0.9712 \pm 0.0021$</td>
<td>$0.9532 \pm 0.0100$</td>
</tr>
<tr>
<td>43</td>
<td>$0.9747 \pm 0.0020$</td>
<td>$0.9744 \pm 0.0039$</td>
</tr>
<tr>
<td>50</td>
<td>$0.9807 \pm 0.0034$</td>
<td>$0.9849 \pm 0.0036$</td>
</tr>
</tbody>
</table>

**Table 8.8:** Relative change in efficiency of the $\Delta R < 2.2$ selection as a result of the smearing of the jet direction as described in the text.
8.1.5 Cross check Monte Carlo jet validation with $b\bar{b}$ events

Samples of inclusive $b\bar{b}$ events in data and in MC are used to validate the reconstruction of $b$-quark jets. For this purpose, the so-called 'double-topo' selection is used. The sample consists of events with two 2-, 3-, or 4-prong vertices selected by the topological lines, which add up to the $B$ meson mass. In these events the $b$ jets originate from the primary vertex, rather than from a displaced vertex. Therefore, the 'prompt' version of the displaced jet algorithm is used to reconstruct the jets. The jets are subsequently matched to the double-topo candidates. In this way, a sample of $b$-jets is obtained in data and in MC.

The events in data are taken from the `BHadron' stream, and the events in MC from the inclusive $b$ sample with two $b$ decays in the acceptance ('INCLB_2inacc' as listed in Table 3.1). The events in both samples have to pass the double-topo stripping line. The minimum required invariant mass of the double-topo candidate in the stripping is 19 GeV. The minimum opening angle between the two $b$ candidates is 0.035 rad, which rejects badly reconstructed or overlapping combinations.

The number of selected events in data is approximately 8M, whereas 620 candidates are selected in MC. These event numbers agree well: after correcting the MC for generator efficiency and the luminosity seen in data, about 8.5M events are expected. Figure 8.12 shows the jet ID properties and the reconstructed transverse momentum of these candidates. The agreement between data and simulation is good. Due to the strict requirement on the double-topo invariant mass made in the 2011 stripping, the sample lacks jets with low $p_T$, which makes it less suited as a control sample for the long-lived particle analysis. Consequently, no quantitative results are extracted from this sample.

8.2 Trigger efficiency

The trigger efficiencies on the hidden valley models shown in Tables 6.2 and 6.3 are evaluated from MC. In order to assess the level of uncertainty associated to the MC simulation of the trigger, the trigger efficiencies of events with displaced vertices in data and MC are compared using generic $B \rightarrow J/\psi X$ events. The comparison is performed as a function of relevant variables such as the vertex mass, number of tracks and radial distance to the beamline. The integrated efficiency difference is used as systematic uncertainty, while the differential difference allows to evaluate whether or not the MC description is valid over the full parameter space. The trigger efficiency is evaluated relative to the stripping efficiency.

Events containing an offline reconstructed displaced vertex are selected through the 'StrippingDVJPsiHLT' stripping line (described in Section 4.5.2), to ensure they are triggered by the 'Hlt2DiMuonJPsi' $J/\psi$ trigger line. The offline $J/\psi$ candidate is triggered by muon or dimuon lines. Furthermore, the $J/\psi$ is required to have a radial
Figure 8.12: Jet ID properties and $p_T$ of jets associated to either of the $b$ candidates in the double-topo samples in data and $b\bar{b}$ MC simulation, before the jet ID cuts are applied.
displacement of $R_{xy} > 0.5$ mm. On this selected sample, one counts how often a trigger line used for the long-lived particle analysis triggers on the offline displaced vertex candidate. The trigger efficiency difference between data and MC is retrieved for loosely selected vertices (SinglePS stripping).

The trigger efficiency is defined for the sample of events with offline reconstructed DV candidates as the ratio of the number of triggered DV events over the number of triggered $J/\psi$ events:

$$
\epsilon_{\text{Trigger}(DV)} = \frac{\bigcup_{i \in \text{DVLines}}(\text{Trigger}(DV)_i) \& \bigcup_{i \in J/\psi Lines}(\text{TOS}(J/\psi)_i)}{\bigcup_{i \in J/\psi Lines}(\text{TOS}(J/\psi)_i)}
$$

where $\text{Trigger}(DV)_i$ means that a trigger line $i$ is fired (TUS or TOS depending on the lines) by the offline displaced vertex candidate, and $\text{TOS}(J/\psi)_i$ means that the Hlt2DiMuonJPsi line was triggered TOS on the $J/\psi$ candidate.

The distributions of the mass, radial distance and number of tracks of the offline displaced vertex candidate are consistent in data and MC over several order of magnitude for all trigger levels. For the displaced vertex HLT2 triggers, this is shown in Fig. 8.13. Note that the mass of the displaced vertex peaks roughly at the $B$-mass. Since the Hlt2DVSingleHighMassPS trigger requires a mass over 10 GeV, most of the candidates pass the Hlt2DVSingleHighFDPS line instead. This explains the cut-off at 2 mm in the radial distance (as defined in Table 4.5).

The integrated efficiencies for data and MC, as well as the corresponding ratio, are shown in Fig. 8.14. The errors on the efficiencies are binomial while the error on the ratio $\sigma_r$ consists of the efficiency errors added in quadrature. Table 8.9 summarises the ratio between the integrated efficiencies in data and MC in this $J/\psi$ triggered DV sample. The integrated efficiencies for this sample of displaced vertices with low mass, low displacement and low track multiplicity are smaller than the integrated efficiencies for the signal that is made of heavier, more displaced objects with a higher number of tracks. The systematic uncertainties retrieved from this method are therefore conservative. The associated systematic uncertainty is conservatively taken as

$$
\sqrt{(1 - r)^2 + \sigma_r^2}, \quad \text{where} \quad r = \frac{\text{eff}_{\text{data}}}{\text{eff}_{\text{MC}}}.
$$

The decrease in signal efficiency is obtained by removing $x\%$ of the triggered

<table>
<thead>
<tr>
<th>Trigger stage</th>
<th>$r$</th>
<th>Syst. uncertainty (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>L0</td>
<td>0.955$^{+0.009}_{-0.009}$</td>
<td>4.5</td>
</tr>
<tr>
<td>Hlt1</td>
<td>0.959$^{+0.008}_{-0.008}$</td>
<td>4.1</td>
</tr>
<tr>
<td>Hlt2Topo</td>
<td>0.946$^{+0.022}_{-0.022}$</td>
<td>5.8</td>
</tr>
<tr>
<td>Hlt2Phys</td>
<td>0.962$^{+0.047}_{-0.047}$</td>
<td>6.0</td>
</tr>
</tbody>
</table>

Table 8.9: Systematic uncertainties for each trigger stage. The systematic uncertainty corresponds to $\sqrt{(1 - r)^2 + \sigma_r^2}$, where $r = \text{eff}_{\text{data}}/\text{eff}_{\text{MC}}$. 

The decrease in signal efficiency is obtained by removing $x\%$ of the triggered
Figure 8.13: Distributions of HLT2 displaced vertex triggered candidates (left) and associated $J/\psi$ candidates (right) in MC and data events. All distributions are normalised to unity.
events for each trigger stage, with \( x \) being the uncertainty per trigger stage (reported in Table 8.9). A weighted average over the events in each sample results in the overall trigger systematic uncertainties quoted in Table 8.1. For the HLT2, the events get an uncertainty assigned depending on whether they passed the topological or the displaced vertex lines. The relative fraction of topological and displaced vertex triggered candidates varies as a function of the mass and lifetime of the signal candidate, as was shown in Figures 4.2a and 4.2c.

8.2.1 Cross-check L0 efficiency using Z + jet events

The method described before focusses on the trigger efficiency on the tracks in the displaced vertex. However, the offline candidate also contains two jets that can trigger the most efficient L0 line: L0Hadron (see Table 4.1). As a cross-check for the method described in Section 8.2, the systematic uncertainty on the L0Hadron trigger efficiency is estimated from Z + jet events in data and in MC. A clean 'back-to-back' Z + jet sample is retrieved using the selection described by 'L0Hadron' in Table 8.2.

In order to obtain the trigger efficiency, one counts the fraction of jets of the offline selected Z + jet candidates containing particles that fired the L0Hadron line, using the TOS requirement on the jet particles. Figure 8.15 gives the L0Hadron efficiencies in data and in MC, and the ratio of the two: \( r_{HV} = \frac{\epsilon_{\text{data}}}{\epsilon_{\text{MC}}} \), as a function of \( p_T \).

By integrating the efficiency for MC over the transverse momentum distribution
Figure 8.15: (a) L0Hadron efficiency on the Z + jet events in data and in MC as a function of jet $p_T$, and (b) the ratio $\langle r^{HV} \rangle$ as a function of jet $p_T$.

<table>
<thead>
<tr>
<th>$\pi^0$ mass [GeV]</th>
<th>10 ps</th>
<th>100 ps</th>
</tr>
</thead>
<tbody>
<tr>
<td>15</td>
<td>0.5365 ± 0.0318</td>
<td>0.5474 ± 0.0435</td>
</tr>
<tr>
<td>25</td>
<td>0.7860 ± 0.0204</td>
<td>0.7525 ± 0.0198</td>
</tr>
<tr>
<td>35</td>
<td>0.8150 ± 0.0241</td>
<td>0.7975 ± 0.0208</td>
</tr>
<tr>
<td>43</td>
<td>0.8251 ± 0.0286</td>
<td>0.8137 ± 0.0242</td>
</tr>
<tr>
<td>50</td>
<td>0.8312 ± 0.0261</td>
<td>0.8113 ± 0.0247</td>
</tr>
</tbody>
</table>

Table 8.10: Estimated efficiency of the L0Hadron trigger on the simulated hidden valley signal sample, using the trigger efficiency estimated from the Z + jet sample, after integration over the jet $p_T$ distribution in the signal sample.
of the jets in the hidden valley signal sample, the L0Hadron efficiencies listed in Table 8.10 are obtained. Only one of the two jets must fire the trigger for the candidate to be selected. When labelling the $p_T$ bins with an index $i$ for the first jet and an index $j$ for the second jet, the efficiency correction factor can be evaluated as

$$ \langle r_{HV} \rangle = \frac{\sum_{i,j} f_{ij}^{HV} (1 - (1 - \epsilon_i^{data}) \cdot (1 - \epsilon_j^{data}))}{\sum_{i,j} f_{ij}^{HV} (1 - (1 - \epsilon_i^{MC}) \cdot (1 - \epsilon_j^{MC}))} \quad (8.2) $$

where $f_{ij}$ is the fraction of signal events in bin $i, j$. The absolute efficiencies on signal in Table 8.10 can be compared to the efficiencies for L0Hadron in Table 4.1, which shows that they are in reasonable agreement. Finally, Table 8.11 lists the estimated efficiency correction factors on signal MC, which can be used to obtain the systematic uncertainty on the L0Hadron trigger efficiency. The efficiency ratios are consistent, although systematically lower, with the $0.955^{+0.009}_{-0.009}$ L0 efficiency ratio obtained in Table 8.9. However, since it has smaller statistical uncertainty, the latter one will be used to estimate the uncertainties in this analysis.

8.3 Vertex finding efficiency

The displaced vertex finding efficiency may be affected by differences in the detector response and the primary vertex multiplicity. This systematic uncertainty can be split into two contributions: the vertex finding efficiency and the track finding efficiency.

8.3.1 Vertex finding efficiency using $B^0 \rightarrow J/\psi K^{*0}$

The performance of the displaced vertex algorithm is studied on well-reconstructed $B^0 \rightarrow J/\psi K^{*0}$ decays in real and simulated data. This decay channel is selected because it contains candidates with a single high-multiplicity vertex (consisting of four tracks) and because sufficiently large and pure samples can be obtained in data and MC.
The candidates are selected to have four well-reconstructed tracks, and to constitute a good-quality candidate in the $B^0 \rightarrow J/\psi K^{*0}$ stripping, such that a vertex finding algorithm should be able to find all of them. Two further selections are applied on the candidates to match the long-lived particle signal: they should be outside of the material veto region, and the decay vertex should have a radial displacement $R_{xy} > 0.2$ mm from the collision point. Since the kinematic distributions in data and MC do not completely match, a reweighting as a function of pseudorapidity is applied to the MC in order to match the data.

![Figure 8.16: Vertex finding efficiency for $B^0 \rightarrow J/\psi K^{*0}$ data (red), unweighted MC (green) and MC after reweighting in $\eta$ (blue), as a function of (a) lifetime and (b) radial displacement from the beamline.](image)

The inclusively reconstructed displaced vertices are matched to the exclusive $B^0 \rightarrow J/\psi K^{*0}$ candidates by requiring that the vertex contains all four tracks used to reconstruct the exclusive decay. The vertex reconstruction efficiency as a function of the displacement is shown in Fig. 8.16. Losses in the vertex finding occur mostly because tracks that belong to the signal vertex have been erroneously assigned to another vertex. Furthermore, the vertex algorithm in the stripping has been tuned to reconstruct vertices with high track multiplicity, with typically seven to eight tracks, rather than the four-track vertices available in this sample. The overall behaviour of the efficiency is well described by the simulation, although the MC efficiency is about 5% higher.

To understand if the difference between data and MC is related to a difference in track parameter resolution, a standard track parameter smearing recipe [108] is applied to the tracks in the simulation. This parameter smearing makes a negligible difference.

Another possible source for the data-MC difference is the ‘confusion’ caused by other tracks in the event. The multiplicity of tracks from the PV is larger in data than in simulation, as can be seen in Fig. 8.17. This can affect the vertex finding efficiency.
since vertices 'compete' for tracks, favouring vertex seeds with higher multiplicity.

By integrating the efficiency difference over the radial distribution in the signal HV100_M35 sample, a total uncertainty of 7.5 ± 1.5% is found. As the long-lived particle signal vertices have higher multiplicity than the \( B^0 \rightarrow J/\psi K^{*0} \) vertices, the sensitivity of the vertex finding efficiency to data-MC differences is expected to be smaller than 7.5%.

### 8.3.2 Tracking efficiency

Since the efficiency above is measured relative to a reconstructed \( B^0 \rightarrow J/\psi K^{*0} \) decay, it does not account for the effects of the per-track inefficiency. For tracks with VELO hits at a small displacement to the beam axis, this efficiency has been estimated with control channels to be well described by the simulation with an uncertainty of approximately 1% [110].

To obtain an estimate of the tracking efficiency at higher radius, it is evaluated on the \( \pi^0_v \) signal MC, as shown in Fig. 8.18a. The efficiency to reconstruct 'reconstructible' tracks (tracks that are within \( 2 < \eta < 4.5 \), have \( p_T > 0.5 \) GeV, and that leave sufficient hits in the VELO to reconstruct the VELO track segment) for this sample drops slightly with \( R_{xy} \), as shown in Fig. 8.18a.

Efficiencies at some distance to the beam axis have also been studied extensively in the context of the \( B \rightarrow J/\psi X \) lifetime analysis [111]. In that case, the VELO track
Figure 8.18: (a) efficiency to reconstruct tracks with a VELO segment versus the radius \( R_{xy} \) of the true origin vertex of the tracks with respect to the beamline, on a hidden valley signal sample with \( m_{\pi^0} = 35 \) GeV and \( \tau_{\pi^0} = 100 \) ps (HV100_M35).

(b) VELO reconstruction efficiency for kaon tracks reconstructed using the offline algorithms as a function of the kaon DOCAz, in simulation and (c) the same in 2011 data. The red solid lines show the result of an unbinned maximum log-likelihood fit to the data. From [111].
Table 8.12: Change in efficiency of reconstructing a vertex candidate with at least six tracks, after randomly removing 2% of the reconstructed tracks, for simulated hidden valley signal (HV10_M35).

<table>
<thead>
<tr>
<th>$R_{xy}$ [mm]</th>
<th>Efficiency ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>0. - 1.</td>
<td>0.9922 ± 0.0200</td>
</tr>
<tr>
<td>1. - 2.</td>
<td>0.9900 ± 0.0187</td>
</tr>
<tr>
<td>2. - 4.</td>
<td>0.9829 ± 0.0145</td>
</tr>
<tr>
<td>4. - 30.</td>
<td>0.9878 ± 0.0173</td>
</tr>
<tr>
<td>0. - 30.</td>
<td>0.9898 ± 0.0086</td>
</tr>
</tbody>
</table>

finding algorithm features a strong dependence of the efficiency on $R_{xy}$, as shown in Fig. 8.18b for simulation and 8.18c for data. The efficiency relative to prompt tracks decreases by about 10% for $R_{xy} = 5$ mm, and by 40% for 10 mm. The statistical uncertainty on this efficiency increases at large radii, because vertices from $B$ decays rarely make it out of the beam pipe. The radial dependence of the hidden valley signal efficiency is not as strong as the $B \rightarrow J/\psi X$ analysis suggests. However, as a conservative estimate, the ratio of data and MC efficiencies from that analysis is integrated over the radial distribution of the HV100_M35 sample, which gives an efficiency correction factor of $0.991 \pm 0.005$ for the radial dependence.

Adding up the 1% uncertainty on the tracking efficiency at small radius, and the 1% uncertainty due to the radial dependence, conservatively, a total uncertainty of 2% is assigned to the offline vertex reconstruction. Table 8.12 gives the change in the efficiency of the vertex selection if 2% of the tracks are removed in the HV10_M35 sample.

If the number of reconstructed VELO segments falls below the threshold of six tracks applied in the selection, the event is lost. The change in tracking efficiency leads to less than 2% loss in vertex reconstruction efficiency, practically independent of the radius, as shown in Table 8.12. This additional uncertainty is assigned to the offline vertex reconstruction to account for the effects of the per-track inefficiency.

To summarise, a 2% uncertainty is conservatively assigned to the track finding efficiency due to differences in the accuracy on the estimated track parameters. Adding this in quadrature to an uncertainty of 7.5% for the performance of the displaced vertex finding algorithm, a total uncertainty of 8% is obtained on the efficiency to find the displaced vertices.

## 8.4 Vertex selection

The stripping applies a selection on the mass and sum-$p_T$ of the long-lived particle, calculated from the charged tracks, without using any jet information. The thresh-
Table 8.13: Relative change in selection efficiency as a result of increasing the stripping vertex mass and sum-$p_T$ thresholds by 2%.

<table>
<thead>
<tr>
<th>$\pi_0^0$ mass [GeV]</th>
<th>10 ps</th>
<th>100 ps</th>
</tr>
</thead>
<tbody>
<tr>
<td>15</td>
<td>0.985 ± 0.021</td>
<td>0.962 ± 0.054</td>
</tr>
<tr>
<td>25</td>
<td>0.971 ± 0.003</td>
<td>0.973 ± 0.006</td>
</tr>
<tr>
<td>35</td>
<td>0.977 ± 0.002</td>
<td>0.979 ± 0.007</td>
</tr>
<tr>
<td>43</td>
<td>0.980 ± 0.002</td>
<td>0.984 ± 0.003</td>
</tr>
<tr>
<td>50</td>
<td>0.984 ± 0.003</td>
<td>0.984 ± 0.004</td>
</tr>
</tbody>
</table>

The tracking efficiency is the main source of uncertainty for this selection. By assuming that the reconstructed mass and sum-$p_T$ are proportional to the number of reconstructed tracks in the vertex, the thresholds on the mass and sum-$p_T$ should be increased by 2% to obtain the efficiency loss. A tracking efficiency loss of 2% translates in a decrease of the efficiency as shown in Table 8.13. The resulting uncertainties per sample are summarised in Table 8.1.

8.5 Dijet mass shape

The dijet mass distribution of the long-lived particle candidates is affected by the jet-energy scale uncertainty. It was shown Section 8.1.2 that the jet energy scale is understood within a few percent. The extrapolation from the Z + jet sample to the different signal samples is reported in Table 8.14, and it is used to scale or 'smear' the mass distribution in the maximum likelihood fit used to extract the background and signal yields. Table 8.15 summarises the systematic uncertainties used to smear the mass shape for different signal MC samples.

<table>
<thead>
<tr>
<th>$\pi_0^0$ mass [GeV]</th>
<th>10 ps</th>
<th>100 ps</th>
</tr>
</thead>
<tbody>
<tr>
<td>15</td>
<td>1.024 ± 0.041</td>
<td>1.021 ± 0.035</td>
</tr>
<tr>
<td>25</td>
<td>1.026 ± 0.033</td>
<td>1.026 ± 0.032</td>
</tr>
<tr>
<td>35</td>
<td>1.026 ± 0.031</td>
<td>1.026 ± 0.032</td>
</tr>
<tr>
<td>43</td>
<td>1.026 ± 0.032</td>
<td>1.026 ± 0.033</td>
</tr>
<tr>
<td>50</td>
<td>1.026 ± 0.032</td>
<td>1.028 ± 0.033</td>
</tr>
</tbody>
</table>

Table 8.14: Correction factor for the jet energy scale for $\pi_0^0$ candidates in different HV models estimated by extrapolating from the Z + jet sample.
<table>
<thead>
<tr>
<th>$\pi^0$ mass [GeV]</th>
<th>10 ps</th>
<th>100 ps</th>
</tr>
</thead>
<tbody>
<tr>
<td>15</td>
<td>4.7%</td>
<td>4.1%</td>
</tr>
<tr>
<td>25</td>
<td>4.2%</td>
<td>4.1%</td>
</tr>
<tr>
<td>35</td>
<td>4.1%</td>
<td>4.1%</td>
</tr>
<tr>
<td>43</td>
<td>4.2%</td>
<td>4.2%</td>
</tr>
<tr>
<td>50</td>
<td>4.2%</td>
<td>4.3%</td>
</tr>
</tbody>
</table>

*Table 8.15:* Systematic uncertainty related to the jet energy scale for different hidden valley signal samples, calculated as: $\sqrt{(1 - r)^2 + \sigma_r^2}$, where $r$ is the correction factor in Table 8.14.

### 8.6 Luminosity

The integrated luminosity used for this analysis is $0.624 \pm 0.011$ fb$^{-1}$. The uncertainty on the luminosity measurement at the LHCb intersection point is determined with two methods. Firstly by Van der Meer scans which measure the beam profile by transversely moving the colliding beams across each other, and secondly by a beam-gas imaging method [112]. The uncertainty on the luminosity measurement for the corresponding data set amounts to 1.7% [113].

### 8.7 Other systematic uncertainties

Several other systematic uncertainties have been considered and found to be negligible.

The estimated number of remaining material interactions after the material veto is much smaller than one event, as shown in Eq. 7.1 in Section 7.4. The uncertainty of the MV on the selection efficiency, can be neglected for the following reason. As shown in Section 6.6, the number of candidates at a radius higher than 5 mm is small. The fit to the data that is used to set an upper limit or to claim a discovery, is performed in bins of $R_{xy}$, and since the bin with the highest radius (4.8-40 mm) does not improve the significance of the result, this bin is discarded. Since the MV only rejects events above $R_{xy} > 5$ mm, any uncertainty on the MV does not influence the results of the analysis.

The position and the size of the primary interaction region could be different in data and MC. The $z$ position of the primary vertex is the primary source for possible systematic uncertainties. Therefore, the $z$ position is weighted in signal MC to match the data, after which the change in signal efficiency is studied. The change in efficiency is $0.5\% \pm 0.07\%$, and is not expected to vary as a function of mass. Since this effect is small, the contribution of uncertainties in the position of the primary interaction region to the selection efficiency will be neglected.

The Global Event Cut (GEC) selection applied on the VELO hits removes less than $0.5\%$ of the candidates in both signal and data (as shown in Table 6.2 for example).
Furthermore, these candidates are typically at high radius, outside the region $R_{xy} < 4.8$ mm that is used to obtain the final results. The contribution of uncertainties in the GEC to the selection efficiency will therefore be neglected.

In the L0, a selection is made on the maximum number of SPD hits: $N_{\text{SPD}} < 600$. The distribution of the SPD hit multiplicity in data and signal MC is different. An uncertainty is retrieved from the efficiency ratio of this selection in data and MC in the $Z + \text{jet}$ samples. This is done in bins of primary vertex multiplicity, since the number of SPD hits is correlated with collision pile-up of the event. The resulting correction factor, is independent of the mass of the signal sample. The efficiency loss of the SPD cut is always $3.2 \pm 0.2\%$ more in data than in MC. This translates into a decrease in signal efficiency, which is taken into account when determining the signal yield in the selected data sample.
SYSTEMATIC UNCERTAINTIES
The mass distribution of the candidates that are left after the final offline selection is used to estimate the number of observed $\pi^0_\nu$ signal candidates, assuming several values of the $\pi^0_\nu$ mass and lifetime. In the following, the procedure will be explained using only the benchmark model with $m_{\pi^0_\nu} = 35$ GeV and $\tau_{\pi^0_\nu} = 10$ ps, whereas the results will be quoted for all signal masses.

Since the statistical errors on the background estimates from simulated events are large, the Monte Carlo statistics would effectively dominate the results if they would be used as an estimate of the actual background in data. Fortunately, the signal mass resolution is high compared to the features in the background mass distribution, as is shown in Fig. 9.1. Therefore the background is parameterised from the data itself, extrapolating with a smooth function, and a signal peak is searched for on top of the smooth background.

The parameter of interest that is retrieved from the fit to the data is the number of signal candidates. The expected number of signal candidates, $\theta$, can be calculated as follows:

$$\theta = L \times \epsilon \times \sigma(H) \times \mathcal{B}(H \rightarrow \pi^0_\nu \pi^0_\nu) \times \mathcal{B}(\pi^0_\nu \rightarrow b\bar{b})(2 - \mathcal{B}(\pi^0_\nu \rightarrow b\bar{b}))$$

(9.1)

where the luminosity $L = 0.62$ fb$^{-1}$, and $\epsilon$ is the total selection efficiency on signal. In the simulation it is assumed that both $\pi^0_\nu$ particles decay to the same final state. The efficiency $\epsilon$ represents the number of selected candidates divided by the number of generated events. However, as the selection efficiencies for the two $\pi^0_\nu$ particles in an event are practically independent, the fraction of selected events with more than one candidate is less than 1%. In data no events with more than one $\pi^0_\nu$ candidate are
selected. Using this Eq. 9.1, $\sigma(H) \times B(H \to \pi^0_\nu \pi^0_\nu)$ can be deduced from the number of signal-like candidates observed in the data.

### 9.1 Statistical method

The RooStats [114] tools are used to obtain the results. The value of the parameter of interest $\theta$, the number of observed signal events, can be determined by using a set of measurements $x$ described by a probability density function (p.d.f.) such as $f(x | \theta)$.

However, the fit model not only depends on $\theta$, but also on the nuisance parameters $\nu$. These are parameters, other than the number of signal candidates $\theta$, that are not known a priori, in this case the signal efficiency, the width and mean of the signal mass distribution, the slope of the exponential background fit and the number of background events. By including these, the p.d.f. is extended to $f(x | \theta, \nu)$. For the long-lived particle search, this p.d.f. consists of a signal and a background distribution:

$$f(x | \theta, \nu) = f_{signal}(x | \theta, \nu) + f_{background}(x | \nu). \quad (9.2)$$

The addition of uncertainties to the nuisance parameters results in an increased statistical uncertainty on the parameter of interest, and a loss in sensitivity [115]. When replacing the variable $x$ by the observed data sample $x_i$ one obtains the likelihood,
which is defined as the product of the probability density functions:

\[
L(\theta, \nu | \{x_i\}) = \prod_{i}^{\text{Nevents}} f(x_i | \theta, \nu).
\] (9.3)

If the fit model describes the data accurately, regardless of the amount of signal truly present, the values for \(\theta\) and \(\nu\) that maximise \(L\) are unbiased estimators of \(\theta\) and \(\nu\).

It is most common, in order to set an upper limit, to define a null hypothesis \(H_0\), which states that the data are consistent with a signal of a given strength plus background. The aim is to test if \(H_0\) can be rejected with the given data set. An alternative hypothesis is defined as the background-only hypothesis \(H_1\). The level of agreement between the data and \(H_0\) is reflected by a 'test statistic' that is a function of the measured variables, denoted by \(t(\theta)\). Furthermore, the \(p\)-value \(p_0\) is the probability, assuming the hypothesis \(H_0\), to find data less or equally compatible with \(H_0\) than the observed data. The hypothesis \(H_0\) can be excluded if the observed \(p\)-value is smaller than a certain threshold. Likewise, \(H_1\) can be rejected if the observed \(p_1\) is small.

The sensitivity of an analysis can be studied by computing the 'expected significance'. This is the \(p\)-value one would expect to find given a certain hypothesis, e.g. background plus a predefined signal contribution. The expected significance can ultimately be compared to the observed significance, which gives an indication of the presence of a signal.

Each hypothesis gives a p.d.f. for the statistic \(t(\theta)\), such as \(g(t|H_0)\). The value measured in data is indicated by \(t_{\text{obs}}\). The \(p\)-value for the case where large values of \(t\) correspond to a bad agreement with \(H_0\) can be written as follows, using the probability density function \(g\):

\[
p_0 = \int_{t_{\text{obs}}}^{\infty} g(t | H_0) dt
\] (9.4)

When the \(p\)-value is low, \(H_0\) can be rejected. An illustration of this procedure is given in Fig. 9.2. The \(p\)-value can also be expressed in terms of a one-sided Gaussian 'significance', such that a \(p\)-value of \(2.87 \times 10^{-7}\) corresponds to a significance of \(5\sigma\).

When including nuisance parameters \(\nu\), the \(p\)-value amounts to:

\[
p_0(\nu) = \int_{t_{\text{obs}}}^{\infty} g(t | H_0, \nu) dt
\] (9.5)

This \(p\)-value now depends on the nuisance parameters, and in principle has to be evaluated for all possible values of \(\nu\). However, the test statistic \(t\) can be defined such that its distribution is independent of the nuisance parameters. This can be achieved by using a profile likelihood ratio, which is defined using the likelihood in Eq. 9.3.
Figure 9.2: Schematic representation of the probability density functions for hypothesised background: $g(t|H_1)$, and signal plus background: $g(t|H_0)$. The test statistic $t$ is shown on the horizontal axis. The observed value of $t$ in data is indicated by $t_{obs}$. The p-value $p_0$ to reject the hypothesis $H_0$ is indicated by the grey area.

The profile likelihood is divided by the value of the likelihood at its maximum to obtain the profile likelihood ratio. The following profile likelihood ratio tests a particular value of $\theta$ (a measure of the signal strength for $H_0$):

$$\lambda(\theta) = \frac{L(\theta, \hat{\nu}(\theta))}{L(\hat{\theta}, \hat{\nu})}$$

(9.6)

where $\hat{\theta}$ is the best estimate for $\theta$ from the fit to the observed data (the so-called estimator) and $\hat{\nu}$ is the estimator of $\nu$. Both $\hat{\theta}$ and $\hat{\nu}$ are evaluated to maximise the unconditional likelihood function $L$. The conditional likelihood estimator $\hat{\nu}(\theta)$ maximises $L$ for a specific $\theta$, given the observed data. Note that $0 \leq \lambda \leq 1$, where a smaller value of $\lambda$ implies worse agreement between the data and the hypothesised $\theta$. The test statistic is now chosen as the log likelihood value:

$$t(\theta) = \begin{cases} 
-2 \ln \lambda(\theta) & \text{if } \theta \geq \hat{\theta} \\
0 & \text{if } \theta < \hat{\theta}
\end{cases}$$

(9.7)

which is independent of $\nu$ [115, 116]. The case $t = 0$ ensures that the hypothesis $H_0$ is not rejected when more signal events are observed than predicted by the signal plus background hypothesis, and that the hypothesis $H_1$ is not rejected when the number of observed events is below the background estimation. The test statistic resembles a $\chi^2$ distribution (following Wilks’ theorem [117]), and the higher the value of $t$, the larger the incompatibility between the data and the hypothesised value of $\theta$. 
9.2 Modelling the data

To achieve maximum sensitivity, the data is divided into bins of $R_{xy}$ (the radial distance of the vertex to the beamline) that are aligned with the selections applied in the trigger and the stripping, as shown in Fig. 9.3a. Above $R_{xy} = 4.8$ mm, the signal efficiency is low due to the material veto and the HLT1 selection. Therefore, only radii between 0.4 and 4.8 mm are used to obtain the final result. Figure 9.3b shows the selection efficiency versus $R_{xy}$ within this range.

The yields in data and the efficiencies on simulated signal (taking into account the lifetime reweighting and the correct primary vertex multiplicity distribution), are given in Table 9.1. The selection efficiency $\epsilon$ is determined by the generator level efficiency (roughly defining the LHCb acceptance) and the total selection efficiency of the analysis, such that $\epsilon = \epsilon_{\text{generator}} \cdot \epsilon_{\text{selection}}$. The signal selection efficiency $\epsilon_{\text{selection}}$ is retrieved from the last line of Table 6.2, and $\epsilon_{\text{generator}}$ is retrieved from Table 3.1. Since the signal selection efficiency of the sample with $m_{\pi^0} = 15$ GeV is too low to reliably
<table>
<thead>
<tr>
<th>(R_{xy} \text{ [mm]})</th>
<th>Data11 yield</th>
<th>Signal MC yield</th>
<th>(\epsilon/\text{bin} \times 10^{-4})</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.4-1.0</td>
<td>1897</td>
<td>576</td>
<td>6.84 ± 0.38</td>
</tr>
<tr>
<td>1.0-2.0</td>
<td>5654</td>
<td>1533</td>
<td>18.76 ± 0.625</td>
</tr>
<tr>
<td>2.0-2.5</td>
<td>2836</td>
<td>777</td>
<td>9.82 ± 0.46</td>
</tr>
<tr>
<td>2.5-4.0</td>
<td>11087</td>
<td>2129</td>
<td>25.13 ± 0.71</td>
</tr>
<tr>
<td>4.0-4.8</td>
<td>6126</td>
<td>883</td>
<td>10.88 ± 0.48</td>
</tr>
<tr>
<td>total 0.4-4.8</td>
<td>27600</td>
<td>5887</td>
<td>73.2 ± 2.1</td>
</tr>
</tbody>
</table>

Table 9.1: Data yield, number of signal candidates surviving in the generated hidden valley signal sample (HV10_M35) and simulated signal efficiency \(\epsilon\) (including the generator efficiency, and taking into account the lifetime reweighting and the correct primary vertex multiplicity distribution) per bin of \(R_{xy}\).

Figure 9.4: Fit of the bifurcated Gaussian model to simulated hidden valley signal (HV10_M35), in radial bin 2.5 < \(R_{xy}\) < 4.0 mm.

extract a signal shape from the events remaining after all selections (see Table 6.2), no results are obtained for this model.

The fit models are implemented using RooFit [118]. The background is modelled as a single sided exponential function with slope \(\tau\):

\[
D_{bkg}(m) = \begin{cases} 
0 & \text{if } m < 0 \\
\frac{1}{\tau} \exp\left[-\frac{m}{\tau}\right] & \text{if } m > 0
\end{cases}
\]  

(9.8)

convoluted with a bifurcated Gaussian, to account for the offset at low mass due to the minimal \(p_t\) threshold of the jets. It is obtained by requiring that the shape follows a Gaussian with different widths on both sides of \(m = \mu\), that is continuous in \(m = \mu\),
MODELLING THE DATA

and normalised to unity:

\[ R_{bkg}(m) = \begin{cases} R_L(m) = \sqrt{\frac{2}{\pi} \frac{1}{\sigma_L + \sigma_R}} \exp\left[-\frac{(m-\mu)^2}{2\sigma_L^2}\right] & \text{if } m \leq \mu \\ R_R(m) = \sqrt{\frac{2}{\pi} \frac{1}{\sigma_L + \sigma_R}} \exp\left[-\frac{(m-\mu)^2}{2\sigma_R^2}\right] & \text{if } m \geq \mu \end{cases} \]  \tag{9.9}

such that the fit model becomes:

\[ (D * R)(m) = \int_{-\infty}^{+\infty} D(m') R(m - m') dm', \]  \tag{9.10}

with \( D \) the exponential distribution and \( R \) the bifurcated gaussian distribution. This expression accounts for the term \( f_{\text{background}}(x \mid \nu) \) in Eq. 9.2.

The signal is modelled by a bifurcated Gaussian function, as illustrated in Fig. 9.4. The parameters of the signal fit are fixed using the Monte Carlo signal. The results of the binned fits of the background-plus-signal model to data are shown in Figure 9.5.

The observable in the fit is the mass of the candidate (after applying the jet energy correction), and the parameter of interest is the number of signal events. Different \( \pi^0 \) masses are assumed from the various generated samples (25, 35, 43 and 50 GeV). Lifetimes of 2, 5, 10, 20, 50 and 100 ps are obtained by reweighting the generated 10 and 100 ps samples.

<table>
<thead>
<tr>
<th>( R_{xy} [\text{mm}] )</th>
<th>Fitted yield</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.4-1.0</td>
<td>1.48 ± 0.93</td>
</tr>
<tr>
<td>1.0-2.0</td>
<td>4.0 ± 2.5</td>
</tr>
<tr>
<td>2.0-2.5</td>
<td>2.1 ± 1.4</td>
</tr>
<tr>
<td>2.5-4.0</td>
<td>5.4 ± 3.4</td>
</tr>
<tr>
<td>4.0-4.8</td>
<td>2.3 ± 1.5</td>
</tr>
</tbody>
</table>

Table 9.2: Fitted signal yields in data, using the HV10_M35 sample for the signal model, from a fit for the cross section, as in the limit setting procedure.

Figure 9.5 shows the best fit of the background-plus-signal model to the data for the HV10_M35 sample. The size of the fitted gaussian signal is translated into a signal yield per \( R_{xy} \) bin. A single yield can be fitted that is distributed across the \( R_{xy} \) bins following the lifetime distribution in MC, as listed in Table 9.2. The combination of the yields in all \( R_{xy} \) bins is used in the procedure to obtain an estimate of the signal production cross section.

The systematic uncertainties are incorporated as constraints on the nuisance parameters in the fit model. The three sources of systematic uncertainties are the total uncertainty on the selection efficiency, the uncertainty on the dijet invariant mass scale and the uncertainty on the shape parameters and the relative normalisation of the signal model due to finite statistics in the simulated samples. Gaussian constraint
Figure 9.5: Mass fit to the data sample for a hidden valley model with $m_{\nu^0} = 35$ GeV and $\tau_{\mu^0} = 10$ ps, for the five $R_{xy}$ bins. The dotted line is the background-only model, and the red dashed line is the fitted signal. The blue line indicates the total fit to the data.
terms are added to the relevant parameters: the width and mean of the signal Gaussian and the signal efficiency per bin.

9.3 Result on 2011 data

Since no significant signal is observed in the data, an upper limit is set on the signal production cross section. The $\text{CL}_s$ method [119] is used, with the test statistic as in Eq. 9.7. The $\text{CL}_s$ method has a frequentist approach to defining a confidence interval, but prevents the exclusion of parameters when there is low sensitivity. This is favourable for analyses in which the background-only hypothesis is hard to distinguish from the background-plus-signal hypothesis.

The upper limit procedure relies on the generation of MC pseudo-experiments that approximate $g(t | \theta, \nu)$ for a range of hypotheses (different values of $\theta$). Test statistic distributions for those values of $\theta$ are obtained by generating pseudo-data sets from the fit model with the $H_0$ hypothesis (background-plus-signal, using the current value of $\theta$) and with the alternative $H_1$ (background-only, where $\theta = 0$) hypothesis. The nuisance parameters are profiled for the null model, and those fixed values of the nuisance parameters are used for generating the pseudo-experiments. This results in a broadening of the test statistic distribution, which makes the separation of the hypotheses more difficult. The level of agreement $p_0$ between data and the hypothesis $H_0$ is calculated using Eq. 9.5. The $p$-values can be extracted from the pseudo-experiments, both for the background-only hypothesis: $\text{CL}_b \equiv p_1$ and for the background-plus-signal hypothesis: $\text{CL}_{s+b} \equiv p_0$. The signal strength is defined as $\text{CL}_s = \frac{\text{CL}_{s+b}}{\text{CL}_b} \equiv \frac{p_0}{p_1}$.

Figure 9.6 shows an example of the result of the hypothesis testing. The observed $\text{CL}_{s+b}$ and $\text{CL}_s$, determined from the observed data set, are shown as blue and red points. In this analysis, results are quoted for the confidence level CL = 95\%, depicted by the horizontal line at $1 - \text{CL} = 0.05$. For the example in Fig. 9.6, this amounts to a $\text{CL}_s$ upper limit of 7 pb. The dashed black line in the figure is the expected $\text{CL}_s$ limit as a function of $\theta$, with $1\sigma$ and $2\sigma$ errors in green and yellow bands, determined from the pseudo-experiments. In this example, although the observed limit is statistically compatible with the expected limit, it is slightly higher. This could indicate that there are more background events than expected, or that there is a signal that is not significant.

The 95 % CL upper limits on $\sigma(H) \times B(H \to \pi^0_\nu \pi^0_\nu)$ for hidden valley models with different $\pi^0_\nu$ masses are shown as a function of lifetime in Figure 9.7. The different mass and lifetime points are not statistically independent. Figure 9.8 and Table 9.3 summarise the final result of this analysis: the observed upper limits for several masses, as a function of lifetime.
Figure 9.6: Example of the result of generating pseudo-data sets at various values of $\theta$ (the assumed cross section in pb on the horizontal axis). The extracted $p$-value is shown on the vertical axis. The red, black and blue points show the results on data, assuming a given signal model. The 95\% CL$_s$ is indicated by the crossing of the horizontal line with the red points, and results in an upper limit on the cross section of about 7 pb. The black dashed line and the green and yellow bands show the expected $p$-value.

<table>
<thead>
<tr>
<th>$\pi^0_\nu$ mass / lifetime</th>
<th>1 ps</th>
<th>2 ps</th>
<th>5 ps</th>
<th>10 ps</th>
<th>20 ps</th>
<th>50 ps</th>
<th>100 ps</th>
<th>200 ps</th>
</tr>
</thead>
<tbody>
<tr>
<td>25 GeV</td>
<td>106.3</td>
<td>54.6</td>
<td>43.8</td>
<td>54.2</td>
<td>80.0</td>
<td>164.1</td>
<td>285.7</td>
<td>588.5</td>
</tr>
<tr>
<td>35 GeV</td>
<td>19.0</td>
<td>10.4</td>
<td>8.0</td>
<td>8.9</td>
<td>13.3</td>
<td>25.4</td>
<td>46.5</td>
<td>89.8</td>
</tr>
<tr>
<td>43 GeV</td>
<td>10.5</td>
<td>5.6</td>
<td>4.4</td>
<td>4.7</td>
<td>6.7</td>
<td>12.4</td>
<td>22.7</td>
<td>42.8</td>
</tr>
<tr>
<td>50 GeV</td>
<td>10.6</td>
<td>5.1</td>
<td>3.7</td>
<td>3.8</td>
<td>4.8</td>
<td>9.3</td>
<td>16.2</td>
<td>29.3</td>
</tr>
<tr>
<td>$35$ GeV, $\pi^0_\nu \rightarrow c\bar{c}$</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>2.5</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td>$35$ GeV, $\pi^0_\nu \rightarrow s\bar{s}$</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>2.3</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
</tr>
</tbody>
</table>

Table 9.3: Observed 95\% CL cross section upper limits (in pb) on a hidden valley model with different masses and reweighted lifetimes.
Figure 9.7: Expected (dashed line) and observed (solid line) cross section upper limits at the 95\% confidence level, as a function of lifetime, for four values of the hidden valley $\pi^0_v$ mass.
Figure 9.8: Observed 95\% CL cross section upper limits for different hidden valley models as a function of lifetime.

9.4  Comparison to existing limits

The upper limits are set in a region of low $\pi_\nu^0$ mass and low lifetime, which so far has been inaccessible for other LHC experiments. For comparison, the closest regions with limits by ATLAS and CMS are shown in Fig. 1.9 and 1.10. The preliminary result from CMS was discussed in Chapter 1. For the lightest Higgs (200 GeV) and $\pi_\nu^0$ (50 GeV) masses available [64], the results are summarised in Fig. 1.10. LHCb probes a lower Higgs mass (120 GeV) and lower $\pi_\nu^0$ masses (25-50 GeV). Note that the data set used by CMS is larger than the one used in this analysis. The ATLAS upper limits extend to lifetimes up to several meters, but ATLAS has a decreasing sensitivity in the region below 100 ps (30 mm) [63, 67] that is covered by the LHCb analysis.

The limits from CDF and D0 (Fig. 1.7 and 1.8) are in the same region as LHCb, but weaker than the LHCb limits. The LHCb analysis has a high sensitivity that is primarily due to the high trigger efficiency of the LHCb experiment and the high center-of-mass energy of the LHC. This is reflected in the fact that the current LHCb result is obtained with only 0.62 fb$^{-1}$, whereas both Tevatron experiments used a data set of 3.6 fb$^{-1}$. 
9.5 Outlook

The results of this analysis are complementary to the results of the ATLAS and CMS experiments and comparable to Tevatron results. The good sensitivity of the LHCb analysis opens up possibilities for improving the results with the 2 \( fb^{-1} \) of data that is currently recorded at a center-of-mass energy of 8 TeV, and with the additional 3 \( fb^{-1} \) of data to be taken in the coming years at \( \sqrt{s} = 14 \) TeV. The increase in center-of-mass energy has a positive effect on the production cross section of the Higgs boson, which would increase the number of expected long-lived particles for the Hidden Valley model. Nonetheless, additional improvements can be made to increase the signal selection efficiency and to improve the accuracy of the results. Most of those were mentioned at the end of the corresponding chapters in this thesis, and a few of them are already implemented for the 2012 data analysis.

Around 2020, an upgrade of most of the LHC detectors is planned. At LHCb, the largest interventions concern the VELO detector, which will be replaced by a silicon pixel detector, and the inner and outer tracker, which will be replaced by a scintillating fiber detector. These developments are expected to improve the precision with which tracks can be measured, and to suppress the ghost rate of reconstructed tracks. About 50 \( fb^{-1} \) of data is expected to be collected with the upgraded detector.

The analysis presented in this thesis introduces many novel techniques and concepts that are not used in any other LHCb analysis, and it can be used as a blueprint for a variety of displaced vertex searches. Ongoing searches cover a signature of a displaced vertex with two jets and a muon, a displaced vertex with two muons, and a vertex with a high track multiplicity without jet requirements, all motivated by the models introduced in Chapter 1.

This thesis gives an impression of both the possibilities and the limitations of physics analyses at high energy particle accelerators. The accelerators provide an overwhelming amount of collision data, which the current generation of detectors can hardly process. Due to limited readout speed and capacity, events need to be rejected at a very early stage by a trigger selection. Furthermore, the high number of interactions and particles present in the collisions puts a burden on the track- and vertex reconstruction algorithms, and makes it challenging to distinguish background from signal events. Because of such reasons, the current experiments are for example not sensitive to low-mass particles that are produced at a high center-of-mass energy.

Therefore, the experiments exploit the properties of distinctive signatures such as missing energy, displacement and high transverse momentum. Since numerous LHC results rapidly confine the possibility for new physics to appear in its most simplified form, it becomes ever more important to initiate less conventional searches, such as the one presented in this thesis.
A Tuning of parameters for vertex reconstruction

A.1 VELO tuning

The tuning of the input parameters of the LSAdaptPV3DFitter algorithm, which reconstructs vertices from VELO tracks, is performed and described in detail in [95]. The following paragraph gives a summary of this procedure, and lists the resulting parameters in Table 4. A similar procedure is applied to optimise the downstream vertex algorithm, which is described in Section A.2.

<table>
<thead>
<tr>
<th>PVSeed3DTool</th>
</tr>
</thead>
<tbody>
<tr>
<td>TrackPairMaxDistance</td>
</tr>
<tr>
<td>MinCloseTracks</td>
</tr>
<tr>
<td>zMaxSpread</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>LSAdaptPVFitter</th>
</tr>
</thead>
<tbody>
<tr>
<td>minTr</td>
</tr>
<tr>
<td>maxIterations</td>
</tr>
<tr>
<td>minIter</td>
</tr>
<tr>
<td>maxChi2</td>
</tr>
<tr>
<td>maxDeltaZ</td>
</tr>
<tr>
<td>maxDeltaChi2/ndof</td>
</tr>
<tr>
<td>minTrackWeight (acceptTrack)</td>
</tr>
<tr>
<td>trackMaxChi2</td>
</tr>
<tr>
<td>trackMaxChi2Remove</td>
</tr>
<tr>
<td>trackChi</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>LSAdaptPV3DFitter</th>
</tr>
</thead>
<tbody>
<tr>
<td>minTr</td>
</tr>
<tr>
<td>Iterations</td>
</tr>
<tr>
<td>minIter</td>
</tr>
<tr>
<td>maxIP2PV</td>
</tr>
<tr>
<td>maxDeltaZ</td>
</tr>
<tr>
<td>minTrackWeight(acceptTrack)</td>
</tr>
<tr>
<td>trackMaxChi2</td>
</tr>
<tr>
<td>trackMaxChi2Remove</td>
</tr>
<tr>
<td>trackChi</td>
</tr>
</tbody>
</table>

Table 4: Definitions of the fit parameters for the vertex reconstruction algorithms PVSeed3DTool, LSAdaptPVFitter and LSAdaptPV3DFitter.
### Table 5: Vertex fit parameters for primary vertex (PV), VELO displaced vertex (DV) and downstream DV algorithms, for the online as well as the offline environment. Where the VELO DV and default PV parameters differ, the numbers are printed in bold. The 'Best' track input includes all track types.

The parameters that have most influence on the performance of the vertex algorithm for the displaced vertex (DV) reconstruction are: \( \text{trackMaxChi2} \), \( \text{trackMaxChi2Remove} \), \( \text{trackPairMaxDistance} \), \( \text{zMaxSpread} \), \( \text{minTr} \) and \( \text{maxIP2PV} \). The optimisation considers the following properties:

- Resolution of the reconstructed vertex;
- Track efficiency, defined by the \( \sum p_T \) of the reconstructed MC tracks in the DV, divided by the \( \sum p_T \) of all the MC tracks;
- Track purity, defined by the \( \sum p_T \) of the MC tracks in the DV, divided by the \( \sum p_T \) of all the tracks in the DV;
- Reconstruction efficiency, defined by the number of reconstructed DVs divided by the number of MC DV candidates;

- Split displaced vertices, defined by the percentage of events where one MC metastable particle is associated to more than one reconstructed vertex;

- Timing of the algorithm. The CPU timing is important since the HLT2 requires a fast online reconstruction;

- Retention rate on sample selected with minimum bias. The event output rate should be kept as low as possible, in order to stay within the allowed trigger output rate. The minimum bias sample used for testing consists of collision data of 2010, with approximately 180,000 events.

The main difference between a $\pi^0$ vertex and a primary vertex is that the first contains less tracks. The parameters $\text{minCloseTracks}$ and $\text{minTr}$ are therefore smaller than for the default PV reconstruction. Though this improves the efficiency, it leads to insufficient retention at trigger level. Therefore the values of $\text{trackPairMaxDistance}$ and $\text{zMaxSpread}$ in the seeding are tightened in the HLT compared to the offline settings. They are furthermore optimised for the highest reconstruction efficiency, with the lowest probability to split the vertex.

A second difference between displaced and primary vertices is that the tracks in the long-lived particle decay can originate from secondary beauty or charm decays. The value of $\text{trackMaxChi2}$ ($\text{trackMaxChi2Remove}$) can be decreased (increased) to avoid splitting the vertex, but it also lowers the efficiency and is therefore not changed. $\text{maxIP2PV}$ could also be increased to include tracks from secondary decays, but studies have shown that 2 mm is the best value considering the track purity and the timing of the algorithm.

### A.2 Downstream tuning

The optimisation of the downstream vertex algorithm $\text{LSAdaptPVFitter}$ is performed on a hidden valley signal sample with a 100 ps $\pi^0$ lifetime, which features more decays outside the VELO than the 10 ps sample. Table 5 shows that both the algorithm and its tuning differ from the VELO displaced vertex reconstruction.

Firstly, starting with the vertex seeding, the value of $\text{MinCloseTracks}$ is tuned. Requiring four tracks improves the track efficiency, purity, number of split vertices and the resolution. There is a small loss in reconstruction efficiency, but the seeds that are found have a higher number of tracks and the timing becomes much faster compared to a requirement of three tracks. In the next optimisation, $\text{zMaxSpread}$ and $\text{TrackPairMaxDistance}$ are varied, because it is expected that the distance between tracks and their spread in $z$ are one order of magnitude larger for downstream
tracks than for tracks with a VELO segment. For an optimal resolution, track efficiency and track purity, small values of $z_{\text{MaxSpread}}$ and $\text{TrackPairMaxDistance}$ should be chosen, whereas the reconstruction efficiency increases with higher values. The number of split vertices is small in all cases, and an optimum of all those observables can be retrieved by setting $z_{\text{MaxSpread}}$ to 3.0 and $\text{TrackPairMaxDistance}$ to 20. However, the timing and retention show a large improvement when slightly adjusting $\text{TrackPairMaxDistance}$ to 2.0, which is chosen as the optimal value.

Secondly, the vertex fit parameters are optimised, namely $\text{trackMaxChi2}$ and $\text{trackMaxChi2Remove}$. A low $\text{trackMaxChi2}$ results in a better resolution and track purity, such that the value $\text{trackMaxChi2}=9$ is preferred. $\text{trackMaxChi2Remove}$ does not influence the vertex reconstruction efficiency and the resolution. In terms of timing, $\text{trackMaxChi2Remove}=64$ is optimal, since the timing depends most on the number of reconstructed vertices. Choosing a high value for $\text{trackMaxChi2Remove}$ implies that usually only one vertex is found. Since the selection efficiency for downstream line requires only one downstream vertex, it is expected that this improvement in the timing will not cause an efficiency loss.

Changing the parameters $\text{maxChi2}$, $\text{maxDeltaZ}$, $\text{maxDeltaChi2NDoF}$ and $\text{minTrackWeight}$ does not influence the track efficiency, track purity, resolution and vertex reconstruction efficiency. The default parameters result in the fastest timing, and are therefore not changed.
## 2011 data sample

<table>
<thead>
<tr>
<th>Version</th>
<th>TCK</th>
<th>Runs</th>
<th>Lumi [ nb$^{-1}$ ]</th>
<th>nEvents</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Magnet up</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MOORE_v12r5</td>
<td>0x5d0033</td>
<td>89333-89351</td>
<td>2.25</td>
<td>329421</td>
</tr>
<tr>
<td></td>
<td>0x5b0032</td>
<td>89388-90181</td>
<td>1.98</td>
<td>336623</td>
</tr>
<tr>
<td></td>
<td>0x5a0032</td>
<td>89489-90207</td>
<td>33.5</td>
<td>5181971</td>
</tr>
<tr>
<td>MOORE_v12r6p1</td>
<td>0x730035</td>
<td>93398-94385</td>
<td>134</td>
<td>28240185</td>
</tr>
<tr>
<td></td>
<td>0x740036</td>
<td>93558-93723</td>
<td>5.18</td>
<td>611940</td>
</tr>
<tr>
<td>MOORE_v12r8</td>
<td>0x760037</td>
<td>96214-100256</td>
<td>106</td>
<td>19833817</td>
</tr>
<tr>
<td></td>
<td>0x790037</td>
<td>101375-101761</td>
<td>39.0</td>
<td>6436069</td>
</tr>
<tr>
<td></td>
<td>0x790038</td>
<td>101762-104263</td>
<td>125</td>
<td>22487804</td>
</tr>
<tr>
<td><strong>All</strong></td>
<td></td>
<td>89333-104263</td>
<td>447</td>
<td>83457830</td>
</tr>
<tr>
<td><strong>Magnet down</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MOORE_v12r3</td>
<td>0x360032</td>
<td>87665-87861</td>
<td>3.31</td>
<td>739124</td>
</tr>
<tr>
<td>MOORE_v12r4</td>
<td>0x480032</td>
<td>87880-87894</td>
<td>1.94</td>
<td>337878</td>
</tr>
<tr>
<td></td>
<td>0x4a0033</td>
<td>87962-87977</td>
<td>2.09</td>
<td>423417</td>
</tr>
<tr>
<td>MOORE_v12r5</td>
<td>0x5b0032</td>
<td>90256-90762</td>
<td>0.03</td>
<td>6940</td>
</tr>
<tr>
<td></td>
<td>0x5a0032</td>
<td>90257-90763</td>
<td>28.4</td>
<td>442934</td>
</tr>
<tr>
<td>MOORE_v12r6p1</td>
<td>0x6d0032</td>
<td>91657-92906</td>
<td>99.1</td>
<td>21487810</td>
</tr>
<tr>
<td></td>
<td>0x700034</td>
<td>91732-91733</td>
<td>1.13</td>
<td>281848</td>
</tr>
<tr>
<td></td>
<td>0x710035</td>
<td>92821-92826</td>
<td>0.90</td>
<td>189726</td>
</tr>
<tr>
<td></td>
<td>0x730035</td>
<td>92929-93282</td>
<td>60.7</td>
<td>13026535</td>
</tr>
<tr>
<td>MOORE_v12r8</td>
<td>0x760037</td>
<td>97114-98882</td>
<td>189</td>
<td>35278259</td>
</tr>
<tr>
<td></td>
<td>0x790038</td>
<td>101891-103686</td>
<td>208</td>
<td>38474244</td>
</tr>
<tr>
<td><strong>All</strong></td>
<td></td>
<td>87665-103686</td>
<td>595</td>
<td>114675123</td>
</tr>
</tbody>
</table>

*Table 6: Trigger software versions, trigger configurations keys, run numbers, online measured luminosity and event yields in the EW stream for magnet up (top) and magnet down (bottom) 2011 data. The trigger configurations (TCK) used for this analysis are printed in bold. The luminosity after offline reconstruction is about 10% lower than the one listed here.*
Bibliography


[35] F. de Campos et al., Probing bilinear R-parity violating supergravity at the LHC. JHEP, 05, 048, 2008. DOI: 10.1088/1126-6708/2008/05/048. (Cit. on p. 12)


[54] ATLAS collaboration, *Updated coupling measurements of the Higgs boson with the ATLAS detector using up to 25 fb⁻¹ of proton-proton collision data*. 2014. (Cit. on p. 16)


[72] M. Strassler, *Some thoughts on long-lived particles (etc.) at LHCbeyond*. Presented as the LHCb Implications Workshop, CERN, 2013. (Cit. on p. 22)


[78] F. Archilli et al., *Performance of the Muon Identification at LHCb*. JINST, 8, P10020, 2013. doi: 10.1088/1748-0221/8/10/P10020. (Cit. on p. 31)


[94] A. Camboni et al., *Measurement of \( \sigma(\bar{b}b) \) and \( \sigma(c\bar{c}) \) with inclusive final states in LHCb at 7 TeV*. LHCb-ANA-2014-013, 2014. (Cit. on p. 45)


The cover of this thesis shows a collection of exotic massive long-lived particles: grains of sand from the Hawaiian island Maui. On the beach this sand has a gray colour, but when you look at it more closely, colourful shapes of fossils, coral and volcanic rock become visible, weathered over the course of time. The picture is made using three-dimensional light microscope technology, which combines many microscopic photographs with different focal lengths into one sharp image.

These sand grains are many times larger than the elementary particles that are studied in high-energy physics. Such elementary particles might seem mysterious, but are nowadays quite easily visualised. Traces of charged particles that traverse a detector such as a cloud chamber or a silicon detector give an impression of the underlying invisible world that there is to discover.

Studying particles

The microscope that was used to visualise the sand grains can reveal objects with a size larger than 1 micrometer. The photon, which carries an energy of about 1 electron volt (eV), transmits the visible light that allows us to see the sand grains. The smaller the object, the smaller the wavelength of the probing particles has to be in order to visualise it. Smaller wavelengths can be achieved by giving the particles more energy, for example by accelerating them. Table S.1 gives an overview of the energies needed to visualise objects of different sizes, and the tools needed to reach those energies.

The most natural acceleration 'tool' is the universe itself. Particles that are produced at high energies, for example in a supernova, can be detected on earth by using the atmosphere, an ocean or polar ice as a detection medium. However, such interesting interactions occur only a few times per year. The investigation of rare physical processes requires more data. Particle accelerators are built to generate a high number of particle interactions in a controlled environment.

The experiment described in this thesis is performed at the LHCb detector, positioned at the Large Hadron Collider (LHC) in Geneva. At the LHC, two proton beams...
A certain amount of energy is needed to give particles a wavelength small enough to probe small objects. The table lists objects of decreasing size and the energy needed to visualise them. The tools with which these energies can be reached are listed in the last column.

<table>
<thead>
<tr>
<th>object</th>
<th>size</th>
<th>energy</th>
<th>tool</th>
</tr>
</thead>
<tbody>
<tr>
<td>sand grain</td>
<td>$10^{-4}$ m</td>
<td>1 eV</td>
<td>optical microscope</td>
</tr>
<tr>
<td>atom</td>
<td>$10^{-10}$ m</td>
<td>10 keV</td>
<td>electron microscope</td>
</tr>
<tr>
<td>nucleus</td>
<td>$10^{-14}$ m</td>
<td>100 MeV</td>
<td>alpha rays</td>
</tr>
<tr>
<td>proton</td>
<td>$10^{-15}$ m</td>
<td>1 GeV</td>
<td>electron beam</td>
</tr>
<tr>
<td>quark</td>
<td>$&lt;10^{-20}$ m</td>
<td>$\gg100$ GeV</td>
<td>proton beam</td>
</tr>
</tbody>
</table>

Collide with an energy of 7 TeV. The protons in the beams are accelerated in bunches, and each crossing of bunches (also called 'event') can contain multiple proton-proton collisions ('primary' interactions). When the protons collide, many different particles are produced. The higher the energy of the protons, the heavier those particles can be. The LHCb experiment detects the particles that are produced by measuring for example their position, energy and charge. This allows the reconstruction of the particles in the event. When the LHC is running at full capacity, it produces 40 million collisions per second, of which the LHCb experiment stores about 3000 events per second. Events have to be rejected because limitations on the reconstruction speed and the storage capacity make it impossible to study each event in detail. The investigation of reconstructed events allows us to test the theory that describes the subatomic world: the standard model.

### The standard model of elementary particles

The known particles and the forces through which they interact are described by the standard model of particle physics. The elementary particles can be grouped into quarks and leptons. Heavy particles can decay into lighter particles, which is why the stable world around us is built of only the lightest quarks: the up- and down-quarks. The two lightest quarks in combination with the lightest lepton (the electron) form protons, neutrons, nuclei and atoms. Additionally, there are heavier quarks such as the charm-, strange-, beauty- and top-quark, and also heavier leptons together with their associated lepton-neutrinos. To each particle corresponds an anti-particle with opposite charge, which is produced simultaneously with the particle.

Three fundamental forces are also included in the standard model. The strong force is the dominant force inside nuclei. It is responsible for keeping atomic nuclei together by binding the quarks in the protons and neutrons. The electromagnetic force acts between electrically charged particles. The weak force couples to both quarks and leptons. The inclusion of a Higgs field in the standard model ensures that particles acquire mass, and also describes the existence of a Higgs particle, recently discovered at CERN.
Beyond the standard model

Even though it has proven to be a successful description of subatomic particles, the standard model still leaves several fundamental questions unanswered. Gravity is not included in the standard model; it is so weak at subatomic scales that it does not influence particle interactions. The standard model also does not explain the existence of dark energy and dark matter, while those together make up 95% of the energy in the universe. Some of the anomalies can be explained by theories that go beyond the standard model, and the present-day goal of experimental high-energy physics is to observe evidence to support or to reject such hypotheses.

Figure S.1: Schematic view of the production and decay of v-particles. While the LEP accelerator was unable to penetrate the barrier separating the hidden valley from the standard model, LHC may be able to produce v-particles due to its higher collision energy. The $\pi^0_v$ particles can subsequently decay to standard model particles via a heavy mediator that passes back through the barrier.

A particular interesting theory is the so-called 'hidden valley' model, which is used as a theoretical guideline for the analysis described in this thesis. The hidden valley model features particles just like the standard model particles, that are 'hidden' because they have a very small interaction probability with the standard model particles. Figure S.1 visualises this with a barrier between the standard model and the hidden valley. Only a massive mediator particle can traverse the barrier. The LHC might provide sufficient energy to cross the barrier and produce hidden valley particles. In order to observe those particles, they somehow need to come back to the standard model region. The lightest hidden valley particle, the $\pi^0_v$ (pronounced vee-pion), can only decay into standard model particles by interacting with the heavy mediator, which is able to pass back through the energy barrier. This process is not instantaneous, because of the high mass of the mediator involved. Therefore, the $\pi^0_v$ acquires a certain lifetime before it decays into standard model particles.

A long-lived exotic particle cannot be detected directly since it does not interact
like the standard model particles. However, when it decays to standard model particles within the detector volume, the presence of a displaced decay vertex (the point from which the standard model particles emerge) reveals its existence. The longer the particle lives, the further it can fly away from the primary interaction point.

Research goal

To make new discoveries in science, one can take different approaches: either to put current knowledge to the test, or to specifically search for theoretically predicted new phenomena. At the LHCb experiment, many physicists use the first approach by looking for inconsistencies with the established standard model theory through precision measurements in beauty and charm decays. However, the analysis presented in this thesis follows the second method. By identifying and eliminating a background of known physical processes, an 'exotic' signal process can be singled out. The aim of this thesis is to conclude whether the exotic long-lived $\pi^0_v$ particle is observed in the LHCb data.

![Figure S.2: Signature of a hidden valley $\pi^0_v$ particle created at a primary vertex (PV), decaying at a displaced vertex into two standard model $b$-quarks, which fragment into jets.](image)

The results are obtained by assuming a hidden valley model in which a standard model Higgs particle decays to two relatively massive (about 35 GeV) long-lived $\pi^0_v$ particles, each decaying into two $b$-quarks. The quarks fragment into other particles and are detected as jets: collections of charged traces (tracks) and energy deposits. The signature that distinguishes the exotic long-lived particle from other interactions consists of a displaced vertex with many tracks, a high reconstructed $\pi^0_v$ mass, and two jets that originate from the displaced vertex, as illustrated in Fig. S.2.

Experimental sensitivity

Until now, no hidden valley $\pi^0_v$ particles have been observed, and different experiments have set upper limits on their production rate. Contrary to the ATLAS and CMS experiments, LHCb will search for $\pi^0_v$ particles with a low mass and low lifetime. As an illustration, Fig. S.3 schematically shows the regions of interest of the current experiments as a function of $\pi^0_v$ mass ($m$) and lifetime ($c\tau$). The LHCb result should
Figure S.3: Schematic representation of the region of interest of different experiments at the Tevatron and LHC accelerators for hidden valley signatures with a displaced vertex. The CMS and ATLAS regions extend to higher masses and lifetimes than displayed, respectively. All experiments assume a different Higgs mass for the production of the $\pi^0_v$ particles, which is indicated in the figure.

be compared to the results from the CDF and D0 experiments (at the Tevatron accelerator in Chicago), which cover a similar region.

Since the particle has to decay within the volume of the detector, the region in which the long-lived particles can be detected in LHCb is determined by the size of its vertex locator, which is about 1 meter long. This puts a limit on the sensitivity to long $\pi^0_v$ lifetimes. The sensitivity to different $\pi^0_v$ masses is limited due to requirements on the minimal momentum of the jets and the particles that originate from the $\pi^0_v$ decay.

Background and signal event selection

Standard model processes are abundant in comparison to the expected occurrence of the exotic process. The events need to be reconstructed with high precision in order to distinguish the $\pi^0_v$ decay (signal) from other processes (background).

The main backgrounds for heavy displaced vertices are interactions of particles with detector material, primary proton-proton collision vertices and decays of standard model composite particles that contain a beauty or charm quark. The material interactions are eliminated by rejecting vertices that are positioned within or nearby
the detector elements. Primary vertices are excluded by considering only particles that
decay at a radial distance of at least 0.4 mm from the proton-proton interaction region. The $b$- and $c$-particles are lighter than the $\pi^0_v$ particle, and can be rejected by requir-
ing the presence of a vertex with a high track multiplicity and a high reconstructed
mass. However, due to combinations of these decays with tracks from elsewhere in
the event, their mass and track multiplicity is sometimes higher than expected, such
that they form the main source of background.

![Dijet mass distributions for real data and simulated hidden valley ('HV') signal with
mass $m_{\pi^0_v} = 35$ GeV and a lifetime $\tau_{\pi^0_v} = 10$ ps. For visibility, the
simulated signal is scaled to an arbitrary production rate.](image)

After a selection has been applied to single out events containing a displaced ver-
tex, jets are reconstructed and associated to this vertex. The jets provide enough in-
formation about the decaying particle to reconstruct its mass. Figure S.4 shows the
reconstructed mass distributions of potential $\pi^0_v$ candidates in simulated signal and in
real data events. There is a clear difference between the two distributions, which will
be exploited to determine whether the events in real data contain any signal.

**Results**

The mass distribution in Fig. S.4 suggests that the real data events, which are assumed
to consist mainly of background, can be modelled by an exponential function with
a certain offset at low mass. A combination of two functions is fitted to the mass
distribution: an exponential function (background) and a gaussian peak (signal) with
a shape that is retrieved from the simulated signal sample. The number of events that
is observed in the fitted signal mass peak can be related to the production rate at which
the $\pi^0_v$ particle is created.

The result of this analysis is that there is no evidence for long-lived exotic particles
Figure S.5: Observed upper limits for different hidden valley models as a function of lifetime. The data is collected at LHCb in 2011 and corresponds to an integrated luminosity of 624 pb$^{-1}$. 
in the 2011 LHCb data set. Since no signal is observed, upper limits can be set on the production rate of $\pi^0_\nu$ particles, as shown in Fig. S.5. Since the theoretical model cannot predict the exact mass and lifetime of the $\pi^0_\nu$, different combinations of mass and lifetime are simulated. The figure shows the result as a function of the $\pi^0_\nu$ lifetime. Different $\pi^0_\nu$ masses are indicated with different colours. The result is expressed as a limit on $\sigma(H) \times B(H \rightarrow \pi^0_\nu \pi^0_\nu)$, the production rate of the process in which a Higgs particle decays into two $\pi^0_\nu$ particles. The lower the upper limit, the less likely it is that the process occurs.

The background events that remain after the selection mostly have a low reconstructed $\pi^0_\nu$ mass and low $\pi^0_\nu$ lifetime. For this reason, the limits get higher when the $\pi^0_\nu$ is assumed to have a lower mass or a lower lifetime. It is hard to distinguish the signal mass peak from the background in these regions. The best performance is obtained for a mass around 40 GeV. The limit gets worse at high $\pi^0_\nu$ lifetimes (above 10 ps) due to the limited acceptance of the vertex detector.

The results are obtained in a region of low $\pi^0_\nu$ mass and low lifetime that so far has been inaccessible for other LHC experiments. The upper limits retrieved in this thesis are in a similar $\pi^0_\nu$ mass-lifetime region as results from the CDF and D0 experiments at the Tevatron accelerator, but stronger.

Even though the $\pi^0_\nu$ particle has not been found, the results do not exclude its existence yet. It might have a mass or a lifetime outside any of the regions that were studied by the different experiments, or its production rate might be lower than the upper limits that were set. The analysis of additional LHCb data that is already recorded and that is to be collected in the coming years can shine more light on long-lived exotic particles. Even though the LHCb detector was not designed to perform analyses like the one presented in this thesis, it has proven to be a worthy competitor to other experiments in unconventional searches for physics beyond the standard model.
Op de omslag van dit proefschrift staat een foto van exotische langlevende deeltjes: zandkorrels van het eiland Maui in Hawaï. Hoewel een zandstrand er van veraf grijs uitziet, verschijnt er onder de microscoop een bonte verzameling van koraal, fossielen en vulkanisch gesteente, verweerd door een lange wereldreis. De foto is gemaakt met een lichtmicroscoop, en de zandkorrels zijn vele malen groter dan de elementaire deeltjes die worden bestudeerd in de hoge-energiefysica. Ook deze elementaire deeltjes laten een andere wereld laten zien als ze van dichtbij worden bekeken.

Deeltjes observeren

Hoe kleiner het object, des te kleiner moet de golflengte van het deeltje zijn waarmee het object zichtbaar gemaakt wordt. De fotonen die het licht van de zandkorrels weerkaatsen geven enkel objecten groter dan 1 micrometer weer. Kleinere golflengtes worden verkregen door de deeltjes meer energie mee te geven, bijvoorbeeld door ze te versnellen. Dit kan met behulp van deeltjesversnellers zoals de Large Hadron Collider (LHC) op het CERN instituut in Genève. In de LHC worden twee bundels met protonen in tegengestelde richting versneld. Op de punten waar ze elkaar kruisen staan detectoren zoals de LHCb detector, waarin de sporen van de proton-proton botsingen waargenomen kunnen worden. In de botsing worden vele verschillende deeltjes geproduceerd, en het LHCb experiment detecteert deze deeltjes door hun energie, lading en positie te meten.

Langlevende exotische deeltjes

Het zogenaamde standaard model geeft een theoretische beschrijving van de bekende elementaire deeltjes en hun interacties. Hoewel dit model vele proeven heeft doorstaan, laat het nog verschillende vragen onbeantwoord. Het geeft bijvoorbeeld geen verklaring voor het bestaan van donkere energie en donkere materie, die samen 95% van de energie in het heelal beslaan. De overgebleven vraagstukken kunnen worden opgelost door nieuwe theorieën die aansluiten bij het standaard model. Het doel
van de hedendaagse experimentele hoge-energie fysica is om bewijs te vinden voor zulke hypotheses.

Enkele van die aanvullende theorieën voorspellen het bestaan van een nieuw exotisch deeltje: het zogenaamde $\pi^0_v$ (uitgesproken als vee-pion). Het $\pi^0_v$ deeltje gedraagt zich niet zoals de standaard model deeltjes, en vliegt dan ook door de detector heen zonder een spoor achter te laten. De enige mogelijkheid om het deeltje waar te nemen, is wanneer het verval naar standaard model deeltjes die wel detecteerbaar zijn. Dit verval is niet eenvoudig en het kan dan ook enige tijd duren voordat het plaatsvindt. Hoe langer het exotische deeltje blijft leven, hoe verder het weg vliegt van het proton-proton botsingspunt. Dit gegeven maakt het mogelijk om het exotische proces te onderscheiden van normale standaard model interacties: eerst is er niets te zien, maar vervolgens verschijnt er een aantal sporen uit het niets.

**Doel van het onderzoek**

Er zijn verschillende methodes om tot nieuwe ontdekkingen te komen binnen de wetenschap: de huidige kennis kan getoetst worden, of er kan specifiek gezocht worden naar nieuwe hypotheses die door theoretici voorspeld worden. De meeste fysici in het LHCb experiment passen de eerste methode toe. Ze zoeken naar tegenstrijdigheden met het standaard model in precisiemetingen van het verval van welbekende deeltjes. De analyse die in dit proefschrift beschreven wordt, volgt echter de tweede aanpak. Door het elimineren van een achtergrond van bekende processen kunnen exotische signaalprocessen geselecteerd worden. Het doel van dit proefschrift is om te concluderen of een exotisch langlevend deeltje geobserveerd is in de LHCb dataset. Tot nu toe hebben andere experimenten nog geen $\pi^0_v$ deeltje geobserveerd.

**Selectie van het signaal**

De verwachting is dat exotische fysica veel minder vaak voorkomt dan de reguliere standaard model processen. Daarom wordt er een strenge selectie toegepast om het $\pi^0_v$ verval (signaal proces) te onderscheiden van reguliere processen (achtergrond), zodat er uiteindelijk enkel interacties overblijven die kandidaten bevatten voor het exotische proces. Wanneer de LHC op volle toeren draait vinden er 40 miljoen botsingen plaats per seconde, waarvan LHCb er ongeveer 3000 per seconde opslaat. De meeste botsingen kunnen niet in detail worden bestudeerd, door de beperkte snelheid waarmee de reconstructie plaats kan vinden en de beperkte opslagcapaciteit.

Om het onderscheid te maken tussen signaal en achtergrond is het nodig om te voorspellen hoe het signaal eruit ziet in de detector. Daarom worden er modellen gemaakt, aan de hand waarvan een groot aantal botsingen gesimuleerd kunnen worden waarin het verval van het $\pi^0_v$ deeltje nagebootst wordt. Een voorbeeld daarvan is te zien in Figuur S.1. De rode lijn geeft de gereconstrueerde massa-afhankelijkheid weer van de $\pi^0_v$ deeltjes in het gesimuleerde signaalmodel. De zwarte punten laten de echte data
zien. Het is duidelijk dat de signaaldeeltjes een hogere massa hebben dan de kandidaten in de echte data. Dat betekent dat de data vooral uit achtergrond bestaan.

Figure S.1: Gereconstrueerde massaverdelingen voor echte data (zwart) en gesimuleerd $\pi^0$ signaal (rood). De x-as geeft de massa weer, en de y-as telt het aantal kandidaten dat is geobserveerd. Het aantal kandidaten in het gesimuleerde signaal is willekeurig opgeschaald, zodat het goed zichtbaar is.

Resultaten

Aangezien er geen exotische deeltjes geobserveerd zijn in de data, zet dit proefschrift een limiet op de waarschijnlijkheid dat het exotische $\pi^0_\nu$ deeltje bestaat. De bovenlimieten die gezet worden op de mate waarin het $\pi^0_\nu$ deeltje geproduceerd wordt zijn beter dan die van de CDF en D0 experimenten aan de Tevatron versneller in Chicago. Daarbij zijn ze aanvullend aan de limieten die gezet zijn door de ATLAS en CMS experimenten aan de LHC versneller, omdat ze een $\pi^0_\nu$ deeltje met een relatief lage massa en lage levensduur uitsluiten. De reden dat verschillende experimenten verschillende gebieden beslaan, is dat het zoekgebied gelimiteerd is door de vorm, grootte en functie van de detectoren.

Het resultaat van dit onderzoek betekent niet dat het deeltje absoluut niet bestaat. Het is mogelijk dat het bijvoorbeeld een lagere massa heeft, of langer leeft, of dat het slechts in zeer kleine mate geproduceerd wordt. In deze gevallen zou LHCb het deeltje (nog) niet kunnen observeren. Om andere massas en levensduren te bestuderen kan de bestaande analyse uitgebreid worden, of kunnen de andere experimenten bijspringen. Om het bestaan van het $\pi^0_\nu$ deeltje volledig uit te sluiten (of te bevestigen!) kan eenzelfde analyse worden uitgevoerd op de data die vanaf 2015 verzameld gaan worden in LHCb.
Acknowledgements

I would like to take this opportunity to thank all the people that have contributed in any way to this thesis. My first words of gratitude go to Wouter for his enthusiasm, guidance and understanding. As the most critical reviewer of our own results, his dedication proved to be vital for delivering work of high quality. Even though it was not always too subtle, I very much appreciated his straightforward honesty, and it was a pleasure to work together. I am also grateful to Marcel for enthusiastically teaching me my first course in particle physics at Nikhef, and guiding me all the way to finishing my thesis. He was indispensable for creating the group atmosphere that enabled the discussion of work-related topics in a positive environment.

I am greatly indebted to Pieter for his help with the analysis. Besides the fact that my thesis would not have been finished without his involvement, I also appreciated the moral support and the semi-Flemish ambience in our office. Of course I should not forget Victor, whose ever-optimistic point of view helped me a lot in difficult times. And he turned out to be right: we were indeed almost there... merci beaucoup! Furthermore I would like to thank our colleagues from Lausanne for the instructive discussions about exotic particles while enjoying the view of Lake Geneva.

During my work for the VELO, Eddy and Kazu always patiently helped out. I learned many things, and I particularly enjoyed helping out with the activities in the control room.

A good working environment needs skilled colleagues and a diverse social programme, which were always present, both at CERN and Nikhef. B-fys outings, jamborees, emergency work meetings at Polder, ice skating lunch, summer schools, PhD paper discussions and Friday afternoon drinks were always good stimuli to continue working with renewed energy. In particular thanks to Jeroen, my office mate at the start of my PhD, and a good friend towards the end. I appreciated the interesting discussions with Suvayu, Roels ability to liven up any conversational topic, Jacco’s eagerness to help people, Siim’s sense of style, Maarten’s guitar skills and Vasilis’ amazing dinners. Rose, it was good to have some support and a faithful roommate in a world dominated by men. Suzanne, thanks for always believing in the existence of exotic par-
articles. Furthermore I should not forget to mention Antonio, Gerhard, Hella, Jeroen, Lennaert, Laurent, Maurizio, Manuel, Niels, Tjeerd and all the other members of the LHCb community who have supported me during the past four years.

One of the most positive sides of Nikhef is the interaction between people from different backgrounds and experiments. I would like to thank Rosemarie, Priscilla, Rogier, Jörn, Stefan, Ingrid, Hartger, Geert-Jan, Erwin, Tino, Panos and Rolf in particular for the good times, and the opportunity to share the worries and joys of working at Nikhef. And Robin, it was good to be able to escape the office every now and then for a tennis match.

Verder wil ik Ellie, Liv and Reinout bedanken voor het aanhoren van al mijn frustraties; mijn humeur was altijd een stuk beter na een avondje uit. Hetzelfde geldt natuurlijk voor Ankie. Alle EMMers, bedankt voor het begrip voor mijn regelmatige afwezigheid de afgelopen jaren. Hopelijk blijft ’t muziek nog lang klinken op de Koewacht! Maarten, Wilbert en Saskia zijn al zolang ik me kan herinneren de beste voorbeelden die ik me kan wensen. Doorzettingsvermogen, academische interesse en oog voor de rest van de wereld heb ik van jullie meegekregen. Natuurlijk wil ik bij deze ook mijn ouders bedanken. Tijdens weekenden op ’t Zand wist mijn moeder me altijd te overtuigen wat ’echt werk’ te verzetten, een goede afleiding voor muis en toetsenbord. Mijn vaders enthousiasme, originele ideeën en rondleidingen bij de testbundel op CERN hebben mijn interesse voor de natuurkunde gewekt, en tot een mooie reis door de wereld van de deeltjesfysica geleid. Martin was marvellous at understanding all the troubles a PhD student can have with both physics and physicists. Luckily he was always able to divert my thoughts, making the many train rides to Bonn definitely worthwhile.

Thanks!

Veerle