Jet reconstruction in LHCb searching for Higgs-like particles

Alessandro Camboni
Estructura i Constituents de la Matèria Dept., Av. Diagonal 647, 08028 Barcelona, Spain
E-mail: acamboni@ecm.ub.es

Abstract. One of the greatest challenges in High Energy Physics is the discovery of the Higgs boson which is responsible for the Standard Model particles mass generation. Below $\sim 150 \text{ GeV}/c^2$ the Higgs decay into two $b$-quarks, $H \rightarrow b\bar{b}$ dominates. The two quarks form a string which fragments, giving rise to hadronization in jets containing $b$-hadrons. The study is focused on the channel where the Higgs boson is produced in association with a gauge boson decaying leptonically $H + W \rightarrow b\bar{b} + l\nu$ and $H + Z \rightarrow b\bar{b} + ll$ and Higgs masses are in the range $115 - 140 \text{ GeV}/c^2$. The gauge bosons decay produces hard leptons quite often isolated from the $b$-jets. Hence an isolated lepton with high transverse momentum is required in order to reject the large QCD background.

The aim of this work is to explore the feasibility to observe Higgs-like particles at the LHCb experiment at CERN by exploiting the detector capabilities to identify $b$-jets.

1. Motivations

The LHCb experiment is one of the four experiments at the LHC, a circular accelerator in which protons-protons collide at a center-of-mass energy of $\sqrt{s} = 14 \text{ TeV}$. This will generate a large number of high energy $b\bar{b}$ pairs which are predominantly produced in the same forward direction. The LHCb detector is a forward single arm spectrometer designed to exploit the large $b\bar{b}$ production cross section ($\sigma_{b\bar{b}} \sim 500 \mu\text{b}$) and to perform precise measurements of CP violation and rare $b$-hadrons decays.

One of the actual greatest challenges in High Energy Physics is the discovery of the Higgs boson which is responsible for the Standard Model particles mass generation. The Higgs mass is not known but electroweak fits suggest a light mass Higgs [1]. Below $\sim 150 \text{ GeV}/c^2$ the Higgs decay into two $b$-quarks $H \rightarrow b\bar{b}$ dominates, giving rise to hadronization in jets containing $b$-hadrons.

The aim of this study is to investigate the feasibility to detect a Higgs boson with intermediate mass at LHCb by using the detector sensitivity to $b$-hadrons in order to reconstruct $b$-jets.

LHCb relies on a precision vertexing, since it has a primary vertex resolution of 10 $\mu\text{m}$ in the transverse plane and 60 $\mu\text{m}$ in the longitudinal direction. In addition, LHCb, which will operate at a luminosity of $2 \cdot 10^{32} \text{ cm}^{-2}\text{s}^{-1}$, has a trigger designed to maximise sensitivity to B decays, and will also be well calibrated thanks to the huge amount of B meson produced.

1.1. SM Higgs

Jet reconstruction algorithms are needed for the search of the Higgs boson in LHCb, as well as for particles predicted by Hidden Valley phenomenological models and as a probe for SUSY...
models with long lived neutralinos. Even though this is a field of search outside the main LHCb scope, the potentiality of the detector are worth to be investigated.

Precision electroweak fits suggest that Higgs boson is most likely to have a low or intermediate mass with an upper limit at 95% C.L. of \( m_H < 154 \) GeV. It is worth emphasizing that \( b\bar{b} \) is the dominant decay mode of a Higgs boson with mass up to \( \sim 130 \) GeV, as can be seen in figure 1.

On the other hand, due to the very restricted LHCb acceptance of 350 mrad in the forward direction as much as 30% of SM Higgs events can be observed. Besides, the standard algorithms already available on the market are conceived for \( 4\pi \) detectors.

Among the possible Higgs production mechanisms, one in particular is a candidate for the Higgs search in LHCb: the associated production of a Higgs boson with a gauge boson decaying leptonically \( H + W \to b\bar{b} + l\nu \) and \( H + Z \to b\bar{b} + ll \) with Higgs masses in the range \( 115 - 140 \) GeV/c, see figure 2. The gauge bosons decay produces high \( P_T \) leptons isolated from the \( b\bar{b} \)-jets, see figure 3, which shows the transverse momentum for muons coming from associated \( W \) or \( Z \) and for muons coming from inclusive \( b\bar{b} \). Hence an isolated lepton with high transverse momentum must be required in order to reject the large QCD background.

![Figure 1. Branching ratio of Higgs boson decay modes as a function of the boson mass.](image1)

![Figure 2. Cross section of various Higgs production mechanisms as a function of Higgs mass.](image2)
1.2. Hidden Valley and SUSY particles

Interesting applications of jet algorithms are those involving the search for Hidden Valley [2] and supersymmetric particles [3].

The so-called Hidden Valley is a sector of non-abelian gauge group which couples weakly to the Standard Model via higher dimension operators. Several, possibly long-lived v-hadrons are predicted, with masses typically of the order of the v-confinement scale $\Lambda_v$. Some v-hadrons may be stable, providing dark matter candidates and missing energy signals, while others decay to neutral combinations of SM particles. Decay lifetimes can vary over many orders of magnitude producing displaced vertices in the detector.

Within the supersymmetric framework, MSSM with R-parity violation models are particularly interesting, as they feature weakly interacting neutralinos $\chi$, whose decay would produce displaced secondary vertices, as for example the decay of a supersymmetric Higgs into neutralinos $h_0 \rightarrow \chi \chi$.

2. Jet algorithms

Jet reconstruction can be summarized into two different approaches, based on two different criteria: nearness of tracks in angle (Cone algorithm) or in transverse momentum (Kt algorithm).

The performance of the LHCb detector to identify secondary vertices can be used to develop a displaced-vertex cone-type algorithm. Information in the form of 4-momenta can be used to build 2 initial seeds, namely 2 secondary vertices formed by two tracks. Other tracks are then added within a cone of a given aperture $\Delta R = \sqrt{\Delta \eta^2 + \Delta \phi^2}$, around the seeds direction. The process ends up with 2 b-jet per event.

Otherwise, the same 4-momenta information can be used by Kt algorithm [4] to build a variable number of jets. A B-tagging of jets is then performed at the end looking at the kinematical properties of the tracks inside the jets.

2.1. Displaced-vertex Cone algorithm

A track selection is performed in order to build a 2-track secondary vertex using kinematic cuts on $\chi^2$, momentum and transverse momentum of tracks, impact parameter significance with respect to primary vertex. Possible $K_s$ candidates are excluded in the mass window $M = [0.490; 0.505]$ GeV. Once a set of seeds is selected for each event, we make use of a multivariate classifier in order to B-tag them. The following step is to choose the best pair of seeds that will represent
the candidate pair of secondary vertices for the event. The choice is made again by using a multivariate classifier. It performs a classification of all seeds pairs, provided that two seeds do not have any tracks in common, selecting the highest probability pair.

The last step is the jet construction: a track is added to form one of the two jets if it beholds in a \((\eta, \phi)\) cone of given aperture around the seeds direction, namely the direction identified by the seed and the primary vertex. In this approach, the direction of the cone does not change after any track addition.

Results shown are obtained by making use of fully reconstructed signal events in which a Higgs boson of \(M_H = 120\) GeV is produced in association with a W or Z boson. The lepton coming from them is required to be within the LHCb acceptance. The lepton is also required to have a transverse momentum \(P_T > 10\) GeV. The efficiency of this cut at generator level is \(\sim 22\%\).

2.1.1. Seed selection

Two well reconstructed tracks are used to form a seed. A seed is defined as true if the tracks are both coming from the same true \(b\) meson.

Seed selection process is performed by a multivariate classifier trained to separate true seeds from the remaining (note that no background has been considered in the tuning phase) by means of kinematic variables of tracks. Figure 4 shows the correlation in \(\phi\) and \(\eta\) of selected seeds versus the true \(b\)-meson direction. The resolution is 114 mrad in both \(\eta\) and \(\phi\). No bias is seen in the reconstructed jet direction.

![Figure 4. \(\phi\) (left plot) and \(\eta\) correlation between reconstructed (x axes) and true secondary vertices.](image)

2.1.2. Seed pair selection

A seeds pair is defined as true if both seeds contain only tracks from \(b\) mesons. This is quite a stringent definition: we are requiring that the 2 tracks of the first seed comes from a true \(B\) and the 2 tracks of the second seed come from the other true \(B\). In principle, less stringent definition of truth can be adopted. The fraction of true visible events, for which there are at least 4 true reconstructed tracks, 2 coming from a \(B\) and 2 from the other \(B\), represents about the 25% of the data sample. Pair classifier is then trained to distinguish true pairs defined in this way. All possible pairs of an event are then ranked according to the classifier output, the best pair among all is finally selected as the candidate secondary vertices pair.
The selected sample after the cut on likelihood has a purity of $\sim 72\%$, with a reconstruction efficiency of $\epsilon_{\text{rec}} = 8.4\%$. The fraction of reconstructed true events over the true visible events corresponds to $24.6\%$.

2.1.3. Jet construction

An optimized value of $\Delta R = 0.7$ has been chosen for cone aperture in $(\phi, \eta)$. Only well reconstructed charged tracks and neutral particles with $\Delta R < 0.7$ are added to the jet. High $P_t$ leptons coming from the W or Z are not included in jets. Figure 5 shows the transverse momentum distribution of the two B-tagged jets, while the distribution peaked at low $P_t$ corresponds to the total transverse momentum of particles not associated to neither of the two jets.

![Figure 5. Transverse momentum of the two b-jets, along with the total $P_t$ of remaining particles, peaking at zero.](image)

The correlation in $\theta$ angle between the reconstructed and the true Higgs polar angle is shown in figure 6.a, with a resolution of $\sim 10$ mrad. Figure 6.b shows the invariant mass of the 2 jets: it peaks at 80 GeV with a resolution of $\sim 27$ GeV. The shift in the mass peak is induced by losses of particles (including neutrinos), showing that energy of jets needs to be calibrated. The energy calibration will be evaluated initially by means of a Monte Carlo analysis and then using $Z \rightarrow b\bar{b}$.

![Figure 6. Correlation in $\theta$ angle between true and reconstructed Higgs (a). Di-jet invariant mass is also shown (b).](image)
2.2. Kt algorithm
The so-called Kt algorithm is the standard jet-algorithm used by many experiments, with several implementations available on the market.

It builds a number of jets that can be different for each event according to nearness of particles in transverse momentum space. Various conditions are required for a jet to be tagged as a $b$-jet, considering the number of particles in the jets, the impact parameter significance of particles, the energy of charged particles over the total energy and the energy in a cone around jet axis over the total energy.

A neural net trained to discriminate against $c$-jets and light jets is also applied using kinematic variables. Selection efficiency for $b$-jets is about 80% with a very good non $b$-jet rejection of 99.7%.

In order to correct the jet energy response, a correction function is applied using the dependence of the ratio $E_{jet}/E_B$ on $\eta$ and $P_t$ of the reconstructed jets. Figure 7 shows the di-jet mass before and after this correction.

Improvements in mass resolution are crucial in order to suppress the background.

A first estimation of $S/\sqrt{B}$ has been done, using associated production $HZ$ and $HW$ as signal and taking into account backgrounds such as $t\bar{t}$, $ZW$, $Wb\bar{b}$ etc. The expectation is about 15-20 reconstructed signal events in one year of data taking ($2 \text{ fb}^{-1}$) and a $S/\sqrt{B} = 0.34$ is obtained after this background rejection. The dominant background contribution comes from $t\bar{t}$ and more efforts are needed to suppress it.

Table 1 shows the possible sources of background and their cross sections. The cross section of the signal is 0.22 pb.

![Figure 7. Di-jet mass before (empty histogram) and after calibration using the Kt algorithm.](image_url)

| Source              | $\sigma$ (pb) | Details                      |
|---------------------|--------------|------------------------------|
| $b\bar{b}$          | $7 \cdot 10^8$ | associated production        |
| $t\bar{t}$          | 570          | extra jet activity           |
| $WZ$                | 0.86         | di-jet mass resolution       |
| $ZZ$                | 0.77         | di-jet mass resolution       |
| $\gamma^*/Z +$ jets | $10^4$       | $b$-jet identification       |
| $W +$ jets          | $10^3$       |                              |
| Signal              | 0.24         |                              |
3. Conclusions

A dedicated approach based on displaced vertices for jet reconstruction has been developed. Jet calibration is mandatory to discriminate various sources of backgrounds. It is very challenging reconstructing a SM Higgs at LHCb in the mass region up to $\sim 130$ GeV. At the present the largely dominant background is $t\bar{t}$, so work is mainly focused in trying to optimize background rejection and di-jet mass resolution. The high performance of LHCb in detecting secondary vertices to identify $b$-jets can be applied to any exotic particle decaying into 2 $b$-quarks giving rise to high momentum jets in the forward region.

References

[1] 2008 Precision Electroweak Measurements and Constraints on the Standard Model *CERN Report CERN-PH-EP/2008-020* (*Preprint* arXiv:0811.4682 [hep-ex])

[2] Strassler M J and Zurek K M 2007 *Phys. Lett. B* **651** 374 (*Preprint* hep-ph/0604261)

[3] Hesselbach S, Franke F and Fraas H 2000 *Phys. Lett. B* **492** 140 (*Preprint* hep-ph/0007310)

[4] Butterworth J M, Couchman J P, Cox B E and Waugh B M 2007 KtJet: A C++ implementation of the K(T) clustering algorithm *Comput. Phys. Commun.* **153** 85 (*Preprint* hep-ph/0210022v1)

[5] Coco V 2008 Ph.D. thesis (*CERN Report* CERN-THESIS-2008-101)