Summer Student Project Report

Missing Energy Signals from b-flavored Hadrons at FCCee

Lucía Barreiro Rius

Supervised by Miguel Escudero and Maksym Ovchynnikov

Abstract

b-flavored hadrons decaying into a final state containing missing energy are of key importance to flavor physics, particularly in searches for new physics. In this context, electron-positron colliders running at the Z peak offer key conditions to explore this type of decays due to their clean environment and also because the initial energy is known. Given the high interest on a future TeraZ factory at FCCee, the aim of this project is to explore the type of signals that could be probed with 10^{12} Z bosons, specifically very suppressed SM decays and also BSM decays including dark sector particles (axion-like particles, dark photons) in the final state, as well as deeply understand the backgrounds for this measurements.

Contents

1 Introduction

1.1 Motivation

b-flavored hadrons decaying into a final state containing missing energy are of key importance to flavor physics. Specifically, this type of hadrons that contain at least one bottom quark have specific properties that make them a powerful tool for probing New Physics. For example, the first limit on the decay $b \to s\,\nu\,\bar{\nu}$, that is highly suppressed in the SM, comes from LEP [1]. More recently, the Belle-II experiment has reported an anomalously large $B^+ \to K^+\,\nu\,\tilde{\nu}$ decay rate [2]. This is very interesting as it could be a signal of physics BSM and may be hitting to the possibility of dark sectors, as the signal consists of missing energy, in other words, energy that escapes detection.

In this context, electron-positron colliders running at the Z peak have the ideal conditions for studying decays of b-flavored hadrons containing missing energy. There are two reasons behind this statement. On the one hand, these colliders provide a very clean environment compared to hadron colliders. As we are colliding fundamental particles, the number of particles produced in each event is much lower than when colliding protons, which contain quarks and gluons inside that lead to far more complex outcomes. On the other hand, the second reason has to do with the process itself and the fact that the initial energy of the b-flavored hadron decaying is known.

Given the high interest on a future TeraZ factory at FCCee, exploring this kind of processes is quite relevant as this collider can potentially create $6\,\cdot\,10^{12}$ Z bosons. Taking into account the branching ratio of the process $Z \to b\bar{b}$ [3] we would be talking about $2\,\cdot\,10^{12}$ b-flavored hadrons produced. Hence, FCCee could be very sensitive to the remarkably small branching ratios of b-flavored decays with large missing energy, which are very much of interest for BSM physics.

1.2 Objectives

The aims of this project are two:

1) explore the types of signals which could be probed with $\approx 10^{12}$ Z bosons. This will include $B \to \tau\,\nu$ decays, inclusive $b \to s\,\nu\,\nu$ decays as well as BSM modes such as $B \to$ invisibles decays containing dark sector states.

2) understand the backgrounds for such types of measurements, arising for instance, from semileptonic b decays, $b \to c\,l\,\nu$ or $b \to u\,l\,\nu$, which are very probable according to SM calculations and hence generate large quantities of missing energy [4].

2 Methodology

2.1 Process studied

The process we have focus on consists of an electron-positron collision at the Z peak with a subsequent decay of the Z boson in one bottom quark and one bottom anti-quark. Because of hadronization this process is followed by the formation of different hadrons in two back-to-back jets, from which we are interested in the b-flavored hadrons formed and its weakly decays. Events are partitioned into two hemispheres by the plane orthogonal to the thrust (as used in LEP/ALEPH event–shape analyses, e.g. [5]). The thrust quantifies the coherence of the group of particles resulting from one collision, in other words, indicates how much of a jet form a group of particles has. We defined it as

$$T = \max_{\vec{n}} \frac{\sum_i |\vec{p}_i \cdot \vec{n}|}{\sum_i |\vec{p}_i|} \quad , \tag{1}$$

where p_i is the momentum of particle i and $\vec{n}$ is the unit vector that maximizes T and defines the thrust axis [6]. Specifically, the signal we are looking for is large missing energy, more than 20 GeV at least, in one of the hemispheres of the event, which is a similar search as the one the ALEPH collaboration performed when obtaining the branching ratios of decays as $b \to \tau^- \bar{\nu}_\tau X$ and constraining SM highly suppressed decays like $b \to s\, \nu\, \bar{\nu}$ [7].

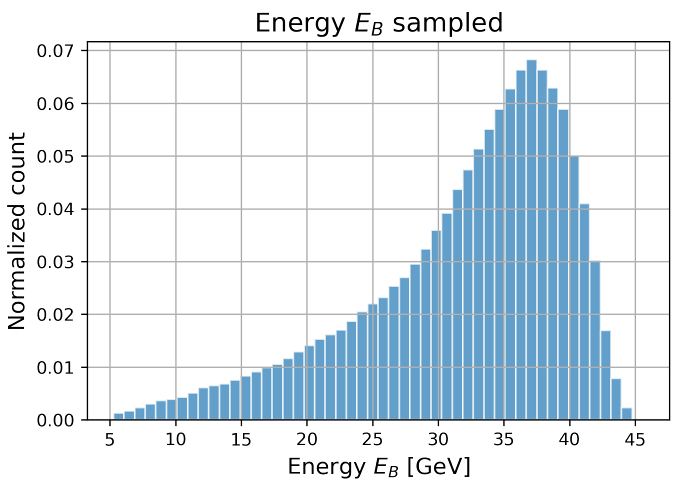

Figure 1: Weakly decaying meson energy distribution, obtained from ALEPH data of Figure 5 in [8].

As mentioned, the initial energy distribution of the weakly decaying hadron is known. In fact, the fragmentation of the b quarks into B mesons at the Z peak was studied by the ALEPH collaboration in [8]. From Figure 5 in [8] we have extracted the data for the initial energy, E_B, of one of those weakly decaying mesons. Specifically, we can obtain the fraction x_B, which is defined as the $E_B/E_{beam} \in [0,1]$. The E_{beam} at the Z pole is approximately 45 GeV.

This spectrum is key to understand the

observed missing energy distribution pattern as it contains the fraction of the beam energy that b-hadrons carry at the time of their weakly decay. On average, at LEP, b-hadrons decay carrying a fraction of $\approx 70\%$ of the beam energy [9]. The ALEPH collaboration performed a dedicated study of the fragmentation function [8] and found that the energy spectrum of the B mesons was well described by the fragmentation function of Kartvelishvili et al. [10].

From the digitized points (x_i, y_i) of the scaled B-hadron energy x_B, we builded a non-negative probability distribution by linear interpolation. We integrated it on $[0,1]$ (trapezoidal rule) and normalize to obtain the CDF $F(x)$. Samples are then drawn by inverse–transform sampling and converted to energies via $E_B = x_B E_{\text{beam}}$ [11]. The final distribution we obtain is shown in Figure 1.

2.2 Simulation

Our analysis is based on a Pythia simulation of the full hadronization process with the Z boson as the starting point. On the one hand, we simulated the whole decay of the Z boson, meaning decays to a pair of quarks, leptons and neutrinos, taking into account the SM branching ratios. That is what ALEPH dealt with when searching for these signals of large missing energy in b-flavored decays. On the other hand, we specified the process to just $Z \to b\bar{b}$, obtaining only events with already our target B hadrons. In that case, we would leave in each event one B meson without decaying. This is important as for a later BSM implementation it is necessary to have that non-decayed particle that can be substituted, outside Pythia, with the chosen BSM products. More on this approach it would be discussed later on.

What leaving a B meson without decaying also allows is to compare the simulated energy distribution for that B meson with the ALEPH-based distribution. This is interesting as a double-check on the simulation, even though it has to be taken into consideration that the simulated distribution is at truth-level while ALEPH data is at detector level. This basically means there are effects coming from the detector resolution, the reconstruction acceptance bias etc. affecting the data that need to be implemented in the simulation in order to directly compare both. This has been done in the Appendix A and found agreement enough to continue with the Pythia simulation as a a validated baseline for BSM studies. Hence, we will now focus on the full hadronization picture of a Z decay in order to study the SM background of our BSM signals.

2.3 ALEPH-like analysis

The ALEPH experiment at LEP performed a search for rare B meson decays with significant missing energy aiming to select them out of their 4 million hadronic events measured between

1991 and 1995 at energies at and around the Z resonance [7]. We have implemented a selection algorithm based on the specific cuts ALEPH used in that search and we have applied them to our Pythia simulation of the whole Z decay. The cuts that ALEPH performed and how we have applied them are the main focus of this subsection.

Hadronic events Out of all the events, ALEPH selected those with two jets, consistent with a $Z \to q\bar{q}$ decay. The standard criteria was applied: at least five good tracks carrying at least 10 % of the centre-of-mass energy. That exact cut was applied in our case to 6 million general Z decay events and we correctly selected, approximately, 4 million hadronic ones, which corresponds to the 69% value of the branching ratio of the Z decay to quarks. Note that ALEPH defines a *good* track as a charged particle track reconstructed with at least four hits [7], which we implemented as that charged particle having enough energy to leave such hits, say more than 1 GeV.

In the same reconstruction algorithm, a Gaussian smearing was applied to the visible particles in order to mimic the ALEPH detector response [12]. The reported energy resolution of the electromagnetic calorimeter is given by

$$\frac{\sigma(E)}{E} = \frac{0.18}{\sqrt{E}} + 0.009, \quad (E \text{ in GeV}),$$

while the precision in the measurable total energy is parametrized as

$$\sigma(E) = 0.6\sqrt{E} + 0.6, \quad (E \text{ in GeV}).$$

Accordingly, the electromagnetic calorimeter resolution is applied to electrons and photons, whereas for hadrons and other charged particles (such as muons) the generic resolution in the energy flow is used, as this particles do not interact with the electromagnetic calorimeter but with the hadronic calorimeter and the muon chamber, in the muons case. As the specific resolution for this two elements is not reported by ALEPH we use the precision in the total energy that we have access to. This smearing allows the simulation to reflect the finite energy resolution of the detector, ensuring that the reconstructed quantities are comparable to those measured in the actual ALEPH data.

For the momentum, the ALEPH collaboration also reported a resolution, in this case of $\sigma(p)/p = 6\,10^{-4} p_T \oplus 0.005$ (p in GeV/c) [12]. Now, we have not included for simplicity momentum smearing in this analysis. We believe that the value reported by ALEPH wouldn't seem to have great significance in our results but we will discussed the implications of this approximation in the conclusions. Hence, every result reported is done with truth-level momenta for every particle. As invisible particles, no smearing is applied to the neutrino energies and all the values in their four-momenta are left untouched.

Thrust To keep only two-jet events well contained in the detector acceptance ALEPH performed two cuts to the thrust. First, the polar angle is required to satisfy $|\cos\theta_{thrust}| < 0.7$ to match the acceptance of the vertex detector. Also, the value of the thrust must exceed 0.85. This allows to correctly select events with two sufficiently back-to-back jets, as those are expected to have a value of the thrust axis near 1, and neglect those close to the direction of the incident beam, effectively avoiding signal looses. The same cuts were applied to the simulation after computing the thrust axis and the thrust value of the event. Several algorithms were used for maximizing the thrust value, all of them only using the momenta of visible particles to the ALEPH detector, in other words, stable-enough particles. This list includes pions, kaons, protons, neutrons, photons, electrons and muons and their antiparticles.

The final algorithm used takes as a unique seed the direction of the largest momentum norm $|p|$ and performs a 20 iterations maximization with a certain tolerance. As we are working with hadronic events already, the thrust distribution of a sample of events has its peak very near one. This is consistent with two quarks being generated and hence the thrust cut selects almost 90% of events, which is a bit above what ALEPH reported, which is more in the 60% to 70% region [13]. Combined with the more selective cut on $\cos\theta_{thrust}$ 55% of events survive these two cuts. Each event is then divided in two hemispheres, which are defined by the plane perpendicular to the thrust axis of the event.

Hemispheres In each hemisphere, ALEPH measures the missing energy as $E_{\text{miss}} = E_{beam} - E_{\text{vis}}$. The visible energy we have calculated it summing the energies in each hemisphere of the visible particles mentioned earlier. The beam energy is performed calculating the invariant mass in each hemisphere, also using only information of the visible particles, which is what ALEPH had access to, and applying four-momentum conservation. The missing energy in each hemisphere is then calculated exactly like ALEPH.

Large missing energy The signal region of ALEPH was defined as $E_{miss} > 35$ GeV. In our case, our study is focused in the high energy tail, defined as $E_{miss} > 20$ GeV. Hence, that cut defines our signal hemisphere, the one with large missing energy.

b-tagging To enhance sensitivity to $Z \to b\bar{b}$ events, ALEPH applied b-tagging in the opposite hemisphere, the one without large missing energy, to ensure that the missing energy signal observed was indeed coming from a b-flavored hadron decay. This can be done because the b quark has a relatively long lifetime and hence hadrons containing them travel a longer distance (order of mm), compared to hadrons containing light quarks, before decaying. This gives to the event quite a distinctive signature for identifying it: what is normally called a secondary vertex. That distance traveled is known as the "flying" distance and can be used, as explained, to distinguish jets coming from b-flavored hadrons from jets coming from other type of hadrons.

The b-tagging algorithm has been based as every other cut on [7]. In order to tag the opposite hemisphere as coming from a b quark, each good track in it has assigned a probability of originating from the primary interaction point. That probability is calculated from the impact parameter significance of the track, where the impact parameter is defined as the distance of closest approach of a particle to the primary vertex. Its sign is often defined using the jet/flight direction so that tracks from a displaced secondary vertex tend to have positive IP. The secondary-vertex displacement can be quantified by the decay length, which in our case was given by Pythia alongside the four-momentum and PDG of every particle. The confidence level α^{hemi} that all good tracks come from the primary interaction point is then required to be smaller than 1% in ALEPH case. Also, ALEPH reported a certain resolution, of $\sigma = 25\mu\text{m} + 95\mu\text{m}\,/p$ (p in GeV/c) [12] on the impact parameter measurement that has been implemented in the simulation.

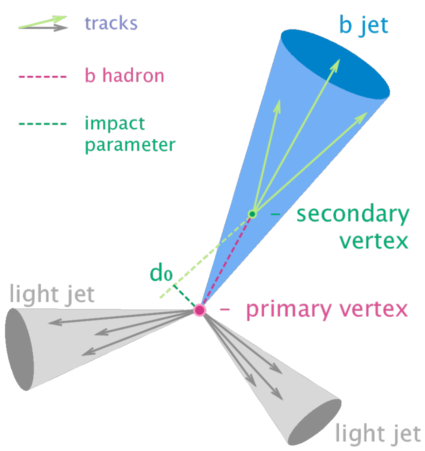

Figure 2: Three jet event with a secondary vertex, coming from a b-flavored hadron.

Now, the algorithm has been tested on a sample of hadronic events as well as on a sample of events coming from a b quark. The goal was to find a balance between the number of events that survive in the inclusive sample, that should be ideally around a 15% (the branching ratio of $Z \to b\,\bar{b}$ [3]) and the ones that survive in the only b quarks sampled, which ideally should be around a 100%. Of course, there are several aspects to consider. First, that a c quark does also have a "flying" distance of the order of mm, and hence events coming from this type of hadron can be tagged as coming from a b incorrectly. Secondly, being very strict with the tag can cause that truly b quark events are lost. Finally, the method we follow is the ALEPH one explained with a slight modification of the specific parameters for the selection that maximizes the efficiency of it. For example, we are more strict with the value of α_{hemi} than ALEPH was and also with the allowed missing energy in the opposite hemisphere, reducing it from 25 GeV TO 20 GeV, but we maintain the number of good tracks in the tag-hemisphere in 7 [7]. Our best result is approximately a 65% selection in $b,\bar{b}$ events and a 20% selection in inclusive hadronic ones.

Semi-leptonic veto Events containing moderately energetic electrons or muons are rejected to suppress backgrounds from semi-leptonic decays such as $b \to \ell \nu X$, which have a large branching fraction and produce significant amounts of missing energy. In our analysis, events with leptons in the signal hemisphere with momentum above 2 GeV are discarded. This threshold is slightly stricter than the 1 GeV cut typically used in ALEPH analyses [12]. We would discuss later the implications of this difference.

2.4 BSM implementation

The next step in the analysis is to implement the BSM decay chosen to study in the sample of b, $\bar{b}$ events. As mentioned, one B meson in each event is left undecayed so its four-momentum can be replace by the four-momenta of the products of the corresponding BSM decay. Hence, a study of two-body and three-body decay kinematics is needed. We have considered a b-flavored hadron decaying to two or three generic particles, one or more of which can be then treated as being invisible. Examples of this are the two-body decay $B_d \to \rho\ a$, where an axion-like particle is involved, or the FCNC (flavour changing nuclear current) decay $B \to X_s\ \nu\ \nu$, where the two neutrinos would be the source of missing energy.

2.4.1 Two-body decay

We first consider a generic two-body decay of a B hadron to two particles, $B \to 1+2$.

Rest-frame of the B hadron When considering the rest-frame of the B hadron and applying simple relativistic kinematics we can obtain both particles energy in this frame (Equation 2 and Equation 3) as well as the magnitude of their momenta (Equation 4). Because of being in the rest-frame of the decaying particle both particles are created with opposite direction momenta.

$$E_1 = \frac{m_B^2 + m_2^2 - m_1^2}{2m_B} \tag{2}$$

$$E_2 = \frac{m_B^2 - m_2^2 + m_1^2}{2m_B} \tag{3}$$

$$p_1 = p_2 = \frac{\sqrt{(m_B^2 + m_2^2 - m_1^2)^2 - 4m_1^2 m_B^2}}{2m_B} \tag{4}$$

LAB-frame of the B hadron The results in the rest-frame of the B hadron can be boosted to the LAB frame considering that the particle is moving with a velocity v_B in the x direction and performing the corresponding Lorentz transformation, given by Equation 5, in which $\gamma = \frac{E_B}{m_B}$ and $v_B = \beta = \frac{P_B}{E_B}$.

$$E^{LAB} = \gamma(E - v_B p_x) \tag{5}$$

Hence, the resulting energies in the LAB frame for particle 1 and 2 are the following:

$$E_1^{LAB} = \gamma \left[\frac{m_B^2 + m_2^2 - m_1^2}{2m_B} - v_B p_1 \cos\vartheta \right] \tag{6}$$

$$E_2^{LAB} = \gamma \left[\frac{m_B^2 - m_2^2 + m_1^2}{2m_B} + v_B p_1 \cos\vartheta \right] \tag{7}$$

where θ is the angle between the particle 1 in the rest frame and the direction of the boost, which in this case and without loose of generality has been chosen to be the x direction. It is easy to notice that E_1^{LAB} + E_2^{LAB} = E_B = $m_B\gamma$ as it should because of energy conservation.

2.4.2 Three-body decay

We now consider the decay of a b-flavored hadron to three particles, namely 1, 2 and 3. In order to simplify the problem we can define $P_{ij} = P_i + P_j$ using the four-momenta P_i where $i = 1,2,3$ and consequently the Lorentz invariant masses of two particles $s_{ij} = m_{ij}^2 = P_{ij}^2$ $(i < j)$ [14]. It is forward to prove that $m_{12}^2 + m_{23}^2 + m_{13}^2 = m_B^2 + m_1^2 + m_2^2 + m_3^2$, where m_i is the mass of the particle i. Hence, the defined invariant masses are not independent of each other.

Rest-frame of the B hadron Our approach in order to solve the three-body decay problem, which has a total of 5 degrees of freedom, is to consider an effective two-body decay in the rest-frame of the decaying B meson. Hence, we will first solve the decay $B \to 1 + (2+3)$ and later solve the decay $23 \to 2+3$ in its own rest-frame and boost our results back to the rest-frame of the b-flavored hadron. Then, by fixing s_{23}, the energy and momentum of particle 1 come from the results from a two-body decay and can be written as:

$$E_1 = \frac{m_B^2 - m_1^2 + s_{23}}{2m_B} \tag{8}$$

$$p_1 = \frac{\sqrt{\lambda(m_B^2, m_1^2, s_{23})}}{2m_B} \tag{9}$$

where we have used the Källén function, defined as $\lambda(a,b,c) = a^2 + b^2 + c^2 - 2ab - 2ac - 2bc$.

Considering now the decay $23 \to 2+3$ in its own frame we can similarly calculate the energies and momenta of particles 2 and 3 as in a two-body decay:

$$E_2^* = \frac{s_{23} - m_2^2 + m_3^2}{2\sqrt{s23}} \tag{10}$$

$$E_3^* = \frac{s_{23} - m_3^2 + m_2^2}{2\sqrt{s23}} \tag{11}$$

$$p_2^* = p_3^* = \frac{\sqrt{\lambda(s_{23}, m_2^2, m_3^2)}}{2\sqrt{s_{23}}} \tag{12}$$

We finally boost these last results back to the rest-frame of B with the following parameters $\gamma_{23} = \frac{E_{23}}{\sqrt{s_{23}}}$ and $\vec{\beta}_{23} = \frac{\vec{p}_{23}}{E_{23}} = -\frac{\vec{p}_1}{E_{23}}$, where $\vec{p}_1 = p_1(\cos\alpha, 0, \sin\alpha)$. The angle α is the angle between the direction of the boost to the LAB-frame of the hadron decaying, which we choose to be the x axis in its rest-frame, and the momentum $\vec{p}_1$. By choosing the component y of $\vec{p}_1$ to be null we are simplifying the problem to the xz plane. The equations for a boost in a general direction are the following:

$$E' = \gamma(E + \vec{\beta} \cdot \vec{p}) \tag{13}$$

$$\vec{p}' = \vec{p} + \left(\frac{\gamma - 1}{\beta^2} (\vec{\beta} \cdot \vec{p}) + \gamma E \right) \vec{\beta} \tag{14}$$

By substituting E_2^*, E_3^*, $\vec{p}_3^* = p_3^*(\cos\theta, 0, \sin\theta)$, $\vec{p}_2^* = -\vec{p}_3^*$ and the corresponding parameters for this specific boost we can solve the three-body decay problem in the rest-frame of the b-flavored hadron. Note that the angle θ is the angle in the 23 rest-frame between the 23 particle boost direction and the particle 3 momentum.

LAB-frame of the B hadron Solving the three-body decay in the LAB-frame is conceptually easy as it simply consists in boosting the results for the rest-frame of B, using Equation 13 and Equation 14. The parameters of this boost are $\gamma_B = \frac{E_B}{m_B}$ and $\vec{\beta}_B = (\frac{p_B}{E_B}, 0, 0)$, where E_B and p_B are the known energy and momentum of the decaying hadron.

3 Results

We have accomplished two main results. Firstly, the comparison between the ALEPH data and our Pythia simulation after the cuts described in subsection 2.3 in the high missing energy tail, defined as mentioned as $E_{miss} > 20$ GeV. Secondly, a constraint on the branching ratio of the decay $B_d \to \rho\, a$, which involves an axion-like particle, based on the events observed by ALEPH, the events of the SM predicted by the simulation and the events with BSM implementation, also in that high missing energy tail.

3.1 Pythia simulation and ALEPH data on the high-E_{miss} tail

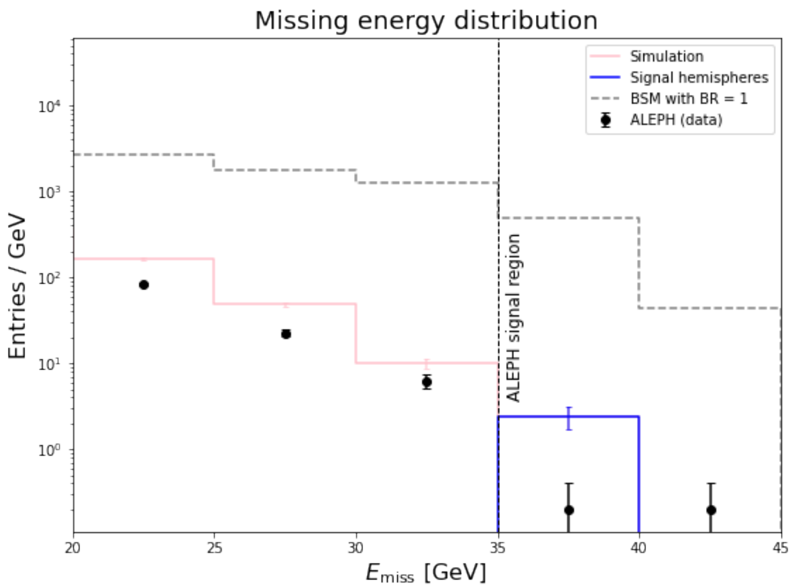

Figure 3: ALEPH data and simulated events that survive on our signal region of $E_{miss} > 20$ Gev. We have also indicated the ALEPH signal region, $E_{miss} > 35$ GeV, and the selected bins, in blue, that in our simulation survive in that region. Finally, in gray, we observe the BSM events for a branching ratio for $B_d \to \rho\, a$ of 1%.

ALEPH reported on [7] their main results for the specific search we have reproduced, which we have summarize in Table 1. Following ALEPH, we use five exclusive bins in the missing-energy tail, $[20,25]$, $[25,30]$, $[30,35]$, $[35,40]$, $[40,45]$ GeV and we use our selecting algorithm to obtain the number of counts in each event, with a starting sample of approximately 4 million hadronic events, a number that mimics the sample ALEPH worked with. Our result can be found also in Table 1.

Table 1: Observed ALEPH counts and the Standard Model (SM) expectation from the simulation, after identical selections.

bin [GeV]	20–25	25–30	30–35	35–40	40–45
N_i^{obs}	420	113	31	1	1
N_i^{SM}	839	245	51	12	0

We find an appreciable difference in the number of events selected in each bin compared to what ALEPH reported [7]. This can be properly observed in Figure 3, where we are showing the high missing energy tail we are interested in. The discrepancy is specially notable in the region of 35 GeV onward, where ALEPH reported 2 events and we find 12. This difference cannot be justified by Poisson fluctuations and hence has to do with the selection process.

By studying the number of events discarded and selected by each cut we have detected this discrepancy with ALEPH results is likely due to the fact that the cut in the value of the thrust is looser than what ALEPH reported and that the algorithm used for b-tagging is also looser than it should be in this high missing energy tail. This two discrepancies are probably leading to the observed behavior. Also, we applied a slightly higher lepton veto than ALEPH in order to have a closer result, which wouldn't be necessary if both the thrust and the b-tag cuts had an efficiency more similar to the one ALEPH reported. Hence, a combination of this three aspects is what more likely is causing the discrepancy in the missing energy distribution that we observed.

However, acknowledging this discrepancy and considering its effect allows for continuing the study, as the SM simulated background needed has been built. As mentioned, we have focus on the BSM decay $B_d \to \rho\, a$ due to being a two-body decay that is still unconstrained experimentally.

3.2 Limit on $BR(B_d \to \rho\, a)$ from the high-E_{miss} tail

The BSM template is built by forcing $B_d \to \rho a$ decays for every B_d in the sample of $Z \to b\,\bar{b}$ events (the other b hadron in the event has been left SM). This basically means that we are implementing the BSM decay with a branching ratio of 1%. The implementation is done using the two-body decay equations from subsection 2.4 and considering the ρ later decays to $\pi^+ \pi^-$, which the SM model predicts to be the only decay of this neutral meson [4]. Also, in this whole analysis the ALP a is considered to be massless, so we fix $m_a = 0$ GeV.

Because in the simulation other type of B hadrons are also left undecayed, events containing $B^\pm$ and B_s are discarded. As a consequence, out of a simulation of 10^5 events, we are left with a initial number of BSM B_d decays of $N^{\rm in}_{B_d} = 490{,}249$. This number agrees with $f_{B_d} = 0.407 \pm 0.007$ [15], which is the probability of a b-quark hadronizing and leading to a weakly decaying B_d in Z decays. Applying the same selection explained in subsection 2.3 to this BSM template we can form *absolute per-B_d* selection probabilities:

$$p_i \equiv \frac{N_i^{\rm sel}}{N^{\rm in}_{B_d}} = \{0.022105,\ 0.018305,\ 0.012985,\ 0.004987,\ 0.000447\},$$

with N_i^{sel} the number of counts per bin i that survive the cuts. Because $\sum_i p_i < 1$, these numbers already include the total selection efficiency.

Both ALEPH data and our SM sample correspond to approximately 4 million hadronic Z events. In the simulation case the exact number is $N_{\rm had} = 4\,176\,778$, but ALEPH number is also of that order. For a test value $\mu \equiv BR(B_d \to \rho a)$ the expected signal in bin i is

$$\nu_i^{\rm sig}(\mu) \ = \ N_b\, \mu\, p_i\,, \qquad N_b \equiv N_{\rm had}\, BR(Z \to b\bar{b}) \times 2, \tag{15}$$

with $BR(Z \to b\bar{b}) = 0.151$ [3]. We omit an explicit B_d fragmentation factor here because it has been already indirectly applied in the B_d-only injection used in the analysis.

Assuming independent Poisson fluctuations per bin, the likelihood for μ is

$$\mathscr{L}(\mu) = \prod_i \mathrm{Pois}\big(N_i^{\rm obs} \,\big|\, \mu_i(\mu)\big), \qquad \mu_i(\mu) = N_i^{\rm SM} + \nu_i^{\rm sig}(\mu). \tag{16}$$

We scan μ on a logarithmic grid (including $\mu = 0$) and use the *bounded* profile–likelihood ratio

$$q_\mu \ = \ -2\ln\lambda(\mu) \ = \ 2\,[\ln\mathscr{L}(\hat{\mu}) - \ln\mathscr{L}(\mu)] \ \geq 0,$$

which vanishes at the best fit $\hat{\mu}$. Under Wilks' theorem, q_μ is asymptotically distributed as a χ^2 variable with one degree of freedom [16]. The one-sided 90% CL upper limit is then defined as the first μ for which $q_\mu = 2.71$ [17].

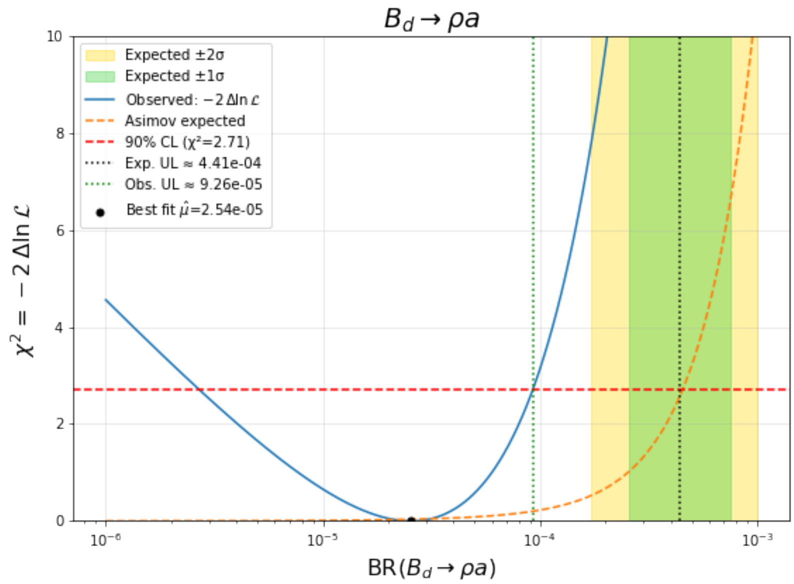

Figure 4: Observed and Asimov expected curves for the decay $B_d \to \rho\, a$. We have marked the best fit for the branching ratio as well as the 90% confidence level upper limit, for both observed and expected. The $\pm 1\sigma$ and $\pm 2\sigma$ expected bands are also plotted.

The expected sensitivity is first estimated using the *Asimov* dataset, defined by setting $N_i^{\rm obs} = N_i^{\rm SM}$, which gives the median expectation in the asymptotic approximation [17]. To quantify the statistical spread around this median, we generate $N_{\rm toys} = 3000$ pseudo–experiments in which each bin is fluctuated according to a Poisson distribution with mean $N_i^{\rm SM}$. For each toy dataset the upper limit is determined, and the central 68% and 95% intervals of the resulting distribution are quoted as the $\pm 1\sigma$ and $\pm 2\sigma$ expected bands.

From the likelihood fit we obtain a best–fit signal strength $\hat{\mu} \simeq 2.54 \times 10^{-5}$. The observed 90% CL upper limit is

$$\mu_{\rm up}^{\rm obs} = 9.26 \times 10^{-5}.$$

For comparison, the expected limit from the Asimov dataset is $\mu_{\rm up}^{\rm A} \approx 4.53 \times 10^{-4}$, while the median of the toy experiments is $\tilde{\mu}_{\rm up} \approx 4.41 \times 10^{-4}$; both numbers agreeing quite well. The toy distribution also yields the expected one– and two–sigma bands:

$$\pm 1\sigma: \ [2.55, 7.56] \times 10^{-4}, \qquad \pm 2\sigma: \ [1.71 \times 10^{-4}, 1.00 \times 10^{-3}].$$

The observed upper limit and the expected median one as well as the $\pm 1\sigma$ and $\pm 2\sigma$ expected bands are plotted in Figure 4.

The observed upper limit is noticeably stronger than the median expected one. This is because the ALEPH data contains fewer events than predicted by the Standard Model simulation in the highest missing–energy bins, which are the most sensitive to a potential signal. Such a deficit disfavors the presence of any additional positive contribution, and the likelihood fit therefore prefers a smaller signal strength. Now, we should recognize that the Standard Model prediction from the simulation does not perfectly reproduce the ALEPH distributions, mainly due to the limitations in our selection algorithm discussed earlier. Hence, this mismatch contributes to the discrepancy between the observed and expected limits, and underlines that the present result should be regarded as preliminary.

Nevertheless, it is instructive to compare with previous bounds in the literature. In particular, [18] derived limits from a recast of the ALEPH analysis targeting partially invisible B decays into an axion. For a massless axion they report

$$\mathrm{Br}(B \to \rho a) < 3.9 \times 10^{-4} \quad (90\%\ \mathrm{CL}).$$

However, their analysis combines both B^+ and B_d contributions to the ρa final state, so the effective bound relevant for comparison is closer to

$$\mathrm{Br}(B \to \rho a) \lesssim 9.75 \times 10^{-4},$$

where we are considering the probability of fragmentation to B_d mentioned earlier. Our present study therefore provides complementary sensitivity, though the systematic uncertainties and background mismodelling must be further refined before a robust comparison can be established.

4 Conclusion

We can conclude from the results discussed several key ideas. Firstly, we have implemented a machinery that reproduces to a certain extent ALEPH results on their search of b-flavored hadron decays with large missing energy. There are still improvements to do to it, mainly to the algorithm that finds the thrust axis of each event as well as to the b-tagging algorithm, which, combined, is what is probably causing the main disagreements. Also, it would be very interesting to consider the momentum resolution of the ALEPH detector, as for example the thrust is quite dependent on the momenta values. Secondly, that machinery has allowed us to obtain a constraint on the BSM decay $B_d \to \rho\ a$, knowing is only a preliminary result until the SM prediction can be improve to be closer to what ALEPH observed.

Hence, what is next to do is pretty straightforward. The SM prediction needs to be refine so an improved result can be obtained for the specific decay considered in this analysis, but also for others. A full $B \to invisible$ decay could be considered, as in [18], but also a FCNC decay as $B \to X_s\, \nu\, \bar{\nu}$ or a decay involving a dark matter particle like $B^+ \to p\, \psi$. Moreover, there has been going on an effort at CERN to recast the ALEPH data to a modern format. This means this whole analysis, with the pertinent improvements, could be applied to real data, which is quite a promising possibility.

A final note is that every code discussed in this report is available on GitHub.

References

[1] Yuval Grossman, Zoltan Ligeti, and Enrico Nardi. "First limit on inclusive decay and constraints on new physics". In: *Nuclear Physics B* 465.3 (Apr. 1996), pp. 369–398. ISSN: 0550-3213. DOI: 10.1016/0550-3213(96)00051-x. URL: http://dx.doi.org/10.1016/0550-3213(96)00051-X.

[2] I. Adachi et al. "Evidence for $B^+ \to K^+\nu\overline{\nu}$ decays". In: *Phys. Rev. D* 109 (11 June 2024), p. 112006. DOI: 10.1103/PhysRevD.109.112006. URL: https://link.aps.org/doi/10.1103/PhysRevD.109.112006.

[3] R. L. Workman et al. "Review of Particle Physics". In: *Prog. Theor. Exp. Phys.* 2022.8 (2022), p. 083C01. DOI: 10.1093/ptep/ptac097.

[4] Particle Data Group. "Review of Particle Physics". In: *Prog. Theor. Exp. Phys.* 2024.8 (2024), p. 083C01. DOI: 10.1093/ptep/ptae083. URL: https://pdg.lbl.gov.

[5] ALEPH Collaboration. "Properties of hadronic Z decays and test of QCD generators". In: *Z. Phys. C* 55 (1992), pp. 209–234. DOI: 10.1007/BF01482583.

[6] Edward Farhi. "A QCD Test for Jets". In: *Phys. Rev. Lett.* 39 (1977), pp. 1587–1588. DOI: 10.1103/PhysRevLett.39.1587.

[7] R. Barate et al. "Measurements of BR (b —> tau- anti-nu(tau) X) and BR (b —> tau- anti-nu(tau) D*+- X) and upper limits on BR (B- —> tau- anti-nu(tau)) and BR (b—> s nu anti-nu)". In: *Eur. Phys. J. C* 19 (2001), pp. 213–227. DOI: 10.1007/s100520100612. arXiv: hep-ex/0010022.

[8] A. Heister et al. "Study of the fragmentation of b quarks into B mesons at the Z peak". In: *Phys. Lett. B* 512 (2001), pp. 30–48. DOI: 10.1016/S0370-2693(01)00690-6. arXiv: hep-ex/0106051.

[9] J. Abdallah et al. "A study of the b-quark fragmentation function with the DELPHI detector at LEP I and an averaged distribution obtained at the Z Pole". In: *The European Physical Journal C* 71.2 (Feb. 2011). ISSN: 1434-6052. DOI: 10.1140/epjc/s10052-011-1557-x. URL: http://dx.doi.org/10.1140/epjc/s10052-011-1557-x.

[10] D. Buskulic et al. "Measurement of the Bs0 lifetime". In: *Physics Letters B* 322.3 (1994), pp. 275–286. ISSN: 0370-2693. DOI: https://doi.org/10.1016/0370-2693(94)91120-7. URL: https://www.sciencedirect.com/science/article/pii/0370269394911207.

[11] Luc Devroye. *Non-Uniform Random Variate Generation*. New York: Springer, 1986. ISBN: 978-0-387-96305-7.

[12] D. Buskulic et al. "Measurement of the B_s 0 lifetime and production rate with $D_s - \ell^+$ combinations in Z decays". In: *Phys. Lett. B* 361 (1995), pp. 221–233. DOI: 10.1016/0370-2693(95)01173-N.

[13] Anthony Badea et al. *Analysis note: measurement of thrust in e^+e^- collisions at $\sqrt{s}$ = 91 GeV with archived ALEPH data*. 2025. arXiv: 2507.14349 [hep-ex]. URL: https://arxiv.org/abs/2507.14349.

[14] Eero Byckling and K. Kajantie. *Particle Kinematics: (Chapters I-VI, X)*. Jyvaskyla, Finland: University of Jyvaskyla, 1971.

[15] Y. Amhis et al. "Averages of b-hadron, c-hadron, and τ-lepton properties as of 2021". In: *Phys. Rev. D* 107 (5 Mar. 2023), p. 052008. DOI: 10.1103/PhysRevD.107.052008. URL: https://link.aps.org/doi/10.1103/PhysRevD.107.052008.

[16] S. S. Wilks. "The Large-Sample Distribution of the Likelihood Ratio for Testing Composite Hypotheses". In: *Ann. Math. Statist.* 9.1 (1938), pp. 60–62. DOI: 10.1214/aoms/1177732360.

[17] Glen Cowan et al. "Asymptotic formulae for likelihood-based tests of new physics". In: *Eur. Phys. J. C* 71 (2011), p. 1554. DOI: 10.1140/epjc/s10052-011-1554-0. arXiv: 1007.1727 [physics.data-an].

[18] Gonzalo Alonso-Álvarez and Miguel Escudero Abenza. "The first limit on invisible decays of B_s mesons comes from LEP". In: *Eur. Phys. J. C* 84.5 (2024), p. 553. DOI: 10.1140/epjc/s10052-024-12936-x. arXiv: 2310.13043 [hep-ph].

Appendices

Appendix A: Pythia and ALEPH comparison on the energy distribution of B mesons

Since the ALEPH collaboration only considers non-strange B mesons, the data in the Pythia distribution was selected to include only the mesons consider by ALEPH, B^o and $B^{\pm}$. In order to make a physical comparison a simplified detector reconstruction model was implemented and tuned directly to the ALEPH measured spectrum.

The tuning procedure was designed to account for two main effects: (i) the smearing of the true hadron energy due to the finite detector resolution, and (ii) the acceptance bias introduced by the event reconstruction, which disfavors high-energy B hadrons.

The smearing was modeled by applying an event–by–event Gaussian fluctuation to the visible energy fraction, parametrized by a mean visible fraction μ_{vis} and a constant fractional resolution term σ_{vis}. The detector resolution was further modeled as

$$\sigma(E_{\text{vis}}) = \sqrt{A^2 + \left(B\sqrt{E_{\text{vis}}}\right)^2}, \tag{17}$$

where A and B are energy–independent and energy–dependent resolution parameters, respectively. The reconstruction acceptance was described by a logistic function

$$\varepsilon(E_{\text{reco}}) = e_{\min} + \frac{1 - e_{\min}}{1 + \exp\left(\frac{E_{\text{reco}} - E_0}{w}\right)}, \tag{18}$$

where $e_{\min}$, E_0 and w control the low–energy plateau, the inflection point and the slope of the turn–on.

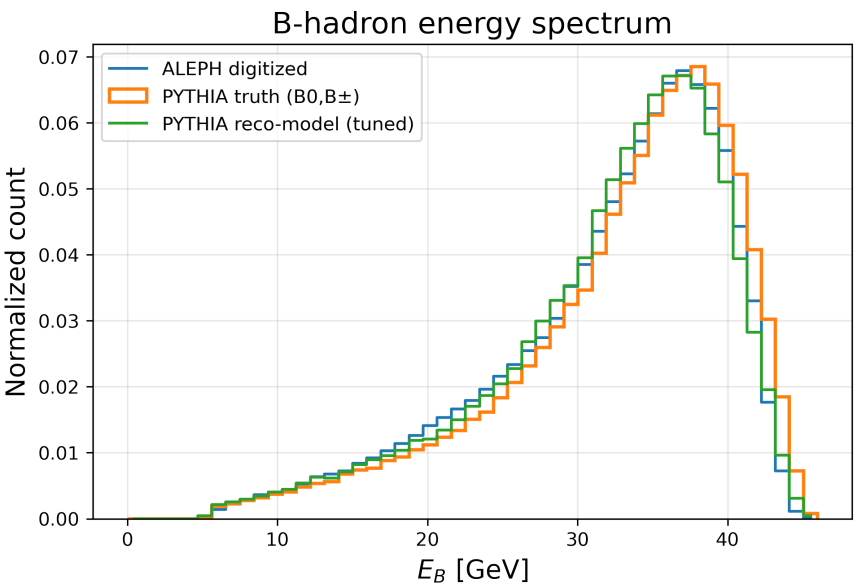

Figure 5: b-flavored hadron energy distribution comparison between the ALEPH digitized data, the Pythia data only including the mesons considered by ALEPH and the model for that Pythia data when including the effects coming from the finite detector resolution and the event reconstruction

The parameters μ_{vis}, A and $e_{\min}$ have been tuned iteratively so that the reconstructed PYTHIA 8 spectrum for non-strange B mesons (B^0 and $B^{\pm}$) matched three key observables of the ALEPH distribution: the peak position (E_{peak}), the RMS width, and the fraction of events in the high-energy tail ($E_B > 38$ GeV). The final tuned values, $\mu_{\text{vis}} = 0.984$, $A = 0.205$, and $e_{\min} = 0.83$, reproduce the ALEPH peak exactly and match the width and tail fraction within $\Delta E \simeq 0.1$ GeV and $|\Delta f| \lesssim 0.01$, respectively.

From this analysis we can conclude that the tuned, minimal reconstruction model reproduces the ALEPH weakly decaying B hadron energy spectrum very well: the peak position matches exactly, the width differs by ~ 0.1 GeV, and the high–energy tail fraction agrees within ~ 0.01. The largest residual discrepancy appears at low E_B. Given these uncertainties, we consider the agreement sufficient for the use of our simulation as a validated baseline for BSM studies.