UNIVERSITE DE LAUSANNE
FACULTÉ DES SCIENCES

Direct search for Higgs boson in LHCb

Thèse de Doctorat

présentée à la Faculté des Sciences
de l'Université de Lausanne
pour l'obtention du grade de docteur ès sciences

par

Charles Currat
Ingénieur-physicien diplômé de l'EPFL

Jury

Professeur **Jacques Boéchat**, président du jury
Professeur **Aurelio Bay**, directeur de thèse
Professeur **Tatsuya Nakada**, expert interne
Docteur **Patrick Janot**, expert externe

Lausanne
2001

"Soyez réalistes, demandez l'impossible"

Remerciements

Première page de cette thèse, mais ultime étape d'un travail de longue haleine, la page des remerciements est l'occasion d'exprimer sa gratitude envers les mille et une personnes sans qui rien n'eût été possible.

Je tiens à remercier en premier lieu Monsieur le professeur Aurelio Bay, *il capo dei capi*, pour m'avoir guidé autant qu'accompagné fidèlement dans mon travail pendant ces années. Sa perception de la physique est surprenante de richesse, d'intuition et d'efficacité pour un jeune thésard jusque là habitué à apprendre la physique dans les livres ...

Je remercie Messieurs le professeur Claude Joseph et le Dr Minh-Tâm Tran qui m'ont permis de débuter dans le domaine à l'occasion de mon travail de diplôme, précédé d'un petit travail d'été qui m'a fait découvrir l'IPHE.

Mes remerciements vont aussi à toute l'équipe des ateliers de mécanique et d'électronique, en particulier Messieurs Raymond Frei, Rodolphe Gonzalez, Jean-Philippe Hertig et Alain Pinard pour leur serviabilité et leur efficacité dont j'ai pu constamment bénéficier.

Je remercie par la même occasion Annette Frenkel, Guiseppe Martellotti, Andrei Medvedkov et Gianni Penso avec qui j'ai pu participer à la "saga" des capillaires.

Je remercie également Monsieur le professeur Olivier Schneider, pour son regard attentif et pointu ainsi que pour ses conseils judicieux (et pointus) que j'ai sollicités à maintes reprises lors de mes errances.

Je remercie Messieurs le professeur Tatsuya Nakada et le Dr Patrick Janot pour l'intérêt qu'ils ont manifesté pour mon travail, ainsi que pour les nombreux conseils qu'ils ont pu me prodiguer depuis les "hot news" du LEP.

J'adresse aussi mes remerciements à tous les membres de la collaboration LHCb qui de près ou de loin ont participé à la réalisation de ce travail, en particulier, Paolo Bartalini, Themis Bowcock, Roger Forty, Oleg Iouchtchenko, Agnieszka Jacholkowska, Mike Koratzinos, Ivan Korolko, Martin McCubbin et Eric Van Herwijnen.

Mon appréhension des méandres de l'omniprésent "démon informatique" a fréquemment bénéficié des connaissances de Messieurs les "gourous" Hubert Degaudenzi, Augustin Kasser et Loris Renggli en particulier. Qu'ils en soient vivement remerciés.

Je tiens à remercier aussi les différents collègues dont j'ai partagé le bureau, successivement Giacomo, Claude et Benjamin, ainsi que Hubert pour l'esprit de camaraderie de tous les instants qui s'est établi entre nous.

Je remercie enfin tous les membres de l'Institut de Physique des Hautes Energies de l'Université de Lausanne pour l'atmosphère chaleureuse et stimulante que chacun(e) contribue à perpétuer.

Merci enfin à mes parents et à mon frère pour toujours avoir su me soutenir et m'encourager dans mon parcours.

Et finalement, je remercie tou(te)s les ami(e)s qui m'ont permis de remettre un peu les pieds sur terre et le nez en dehors de l'Institut en de nombreuses et revigorantes occasions, Arnaud, Claude, Claudie, Léon, Olivier, *Piccolo* Stéphane, Rénald et bien d'autres.

Lausanne, le 25 juillet 2001

Abstract

The LHCb detector is a forward one-arm spectrometer dedicated to precision measurements of CP violation in the B-meson systems. The motivation of the present work is to assess the potential of LHCb to observe a Standard Model (SM) Higgs signal.

The recent results obtained at LEP give a hint of a SM Higgs boson with a mass $m_H = 115.0^{+1.3}_{-0.9}$ GeV/c^2 with a statistical significance of 2.9 standard deviations.

Because of the high longitudinal boost encountered by the products in the pp collisions at LHC, a significant fraction ($\sim 30\%$) of light Higgs ($m_H = 115$ GeV/c^2) are produced in the LHCb acceptance $1.8 < \eta < 4.9$.

These facts potentially place LHCb in the race for the observation of the SM Higgs.

Given a relatively low running luminosity of 2×10^{32} cm^{-2}s^{-1}— compared to the nominal 10^{34} cm^{-2}s^{-1} at LHC — and a limited geometrical acceptance, we have shown that the channels accessible to LHCb are $H + W^{\pm}/Z^0 \to b\bar{b} + \ell^{\pm}X$, for Higgs masses in the range 100–130 GeV/c^2.

This work pioneered a setup for the production and the analysis of *hard jets* in the LHCb detector. We demonstrated in the full detector simulation that the LHCb baseline design allows to efficiently identify, reconstruct and trigger the b-jets coming from the Higgs.

Due to the impossibility to perform a detailed simulation of the huge amount of background, we have developed a "fast simulation" which includes the relevant detector effects.

The Higgs analysis requires a hard lepton isolated from the b-jets to reject the QCD background. At this stage, the cut-based study indicates that the dominant background comes from top pairs production. The signal significance is $S/\sqrt{B} \sim 0.7$ for one LHCb year (integrated luminosity of 2 fb^{-1}).

A realistic scheme for b-jets tagging and for the identification of the associated lepton is left for further studies. This work initiated an involved analysis that deserves to be continued. The signal significance is modest, however, potentially large improvements must not be neglected when performing a combined optimization on the discriminating variables. This thesis also suggests some new strategies to enhance the signal significance.

Part of the work done for this thesis was the participation in technical developments for LHCb. Appendix C summarizes the activity led in the context of the RD46 collaboration at CERN during the years 1996–1998 to develop the capillary layers technique for tracking purpose in high luminosity environment. The various publications to which the present work contributed are listed at the end of the appendix.

Résumé

Le détecteur LHCb est un spectromètre à un bras couvrant les petits angles polaires, conçu pour des mesures de précision de la violation CP dans les systèmes de mésons B. La motivation du présent travail consiste à évaluer le potentiel de l'expérience LHCb d'observer le boson de Higgs du Modèle Standard (MS).

Les récents résultats obtenus auprès de l'accélérateur LEP donnent une indication pour un Higgs d'une masse $m_H = 115.0^{+1.3}_{-0.9}$ GeV/c^2 avec une signification statistique de 2.9 écarts-types.

En raison du fort boost longitudinal des produits dans les collisions pp au LHC, une fraction importante ($\sim 30\%$) des Higgs légers (typiquement $m_H = 115$ GeV/c^2) vont être produits dans l'acceptance de LHCb, $1.8 < \eta < 4.9$ en pseudo-rapidité.

Ces faits placent potentiellement l'expérience LHCb dans la course pour observer le Higgs du MS.

Etant donné une luminosité de fonctionnement relativement basse de 2×10^{32} cm^{-2}s^{-1}— en comparaison avec la luminosité nominale de 10^{34} cm^{-2}s^{-1} au LHC — ainsi qu'une acceptance géométrique limitée, nous avons montré que les seuls canaux accessibles pour LHCb sont $H + W^{\pm}/Z^0 \to b\bar{b} + \ell^{\pm}X$, pour des Higgs de masse comprise entre 100–130 GeV/c^2.

Ce travail a mis en place *ab nihilo* un environnement pour la génération et l'analyse de *jets durs* dans le détecteur LHCb. Nous avons démontré, au niveau de la simulation complète du détecteur, que la conception de base du détecteur LHCb permet d'identifier, de reconstruire et de sélectionner les jets-b provenant du Higgs avec une efficacité élevée.

En raison du contingent de génération des bruits de fond au niveau de la simulation complète, une simulation rapide du détecteur incluant les effets dominants de l'appareillage a été développée.

L'analyse du signal demande un lepton dur et isolé des jets-b pour rejeter le bruit de fond QCD. A ce point, notre analyse en coupures indique que le bruit de fond dominant provient des paires de quarks $t\bar{t}$. La signification statistique du signal est $S/\sqrt{B} \sim 0.7$ pour une année d'exploitation à LHCb (2 fb^{-1}).

Une procédure réaliste dédiée à l'étiquetage des jets-b, ainsi qu'à l'évaluation de l'efficacité d'identification du lepton associé reste à être mise en oeuvre dans une étude ultérieure. Ce travail a permis d'initier une analyse difficile qui mérite

d'être approfondie. La signification statistique obtenue est modeste, toutefois, il ne faut pas négliger de potentiels grands gains par une optimisation combinée des variables discriminantes. Cette thèse suggère également diverses nouvelles stratégies pour augmenter la signification statistique du signal.

Une partie du travail accompli dans cette thèse fût la participation à des développements techniques pour LHCb. L'appendice C résume l'activité menée dans le contexte de la collaboration RD46 au CERN pendant les années 1996–1998 dans le but de développer la technique des capillaires en couches pour la reconstruction de traces à haute luminosité.

Les différentes publications auxquelles les présent travail a contribué sont listées à la fin de l'appendice.

Contents

Chapter 1

Introduction

1.1 Towards a Standard Model of particles

The current representation of particle physics contains two families of particles — the leptons and the quarks — interacting each other via mediators — the gauge bosons. The mathematical framework describing the theory has emerged around forty years ago with the first attempt by Murray Gell-Mann to classify the known-to-date subatomic particles: the *Eightfold Way*, in 1961 [10]. The Eightfold Way arranged the baryons and mesons into geometrical patterns[1], according to their charge and strangeness.

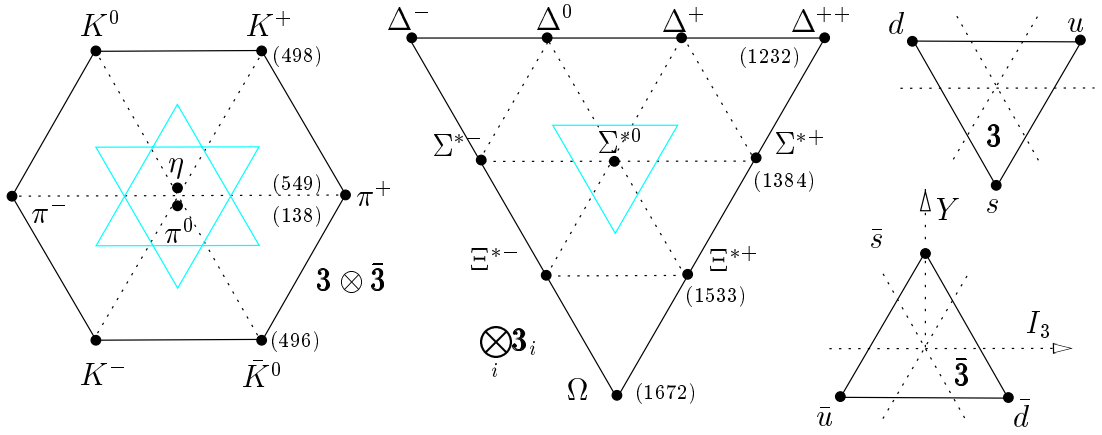


Figure 1.1: Hadron classification: $3 \otimes \bar{3} = 8 \oplus 1$ for mesons; $3 \otimes 3 \otimes 3 = 10 \oplus 8 \oplus 8 \oplus 1$ for baryons

The interpretation of these patterns soon led Gell-Mann and Zweig to independently propose the *quark model* in 1964, asserting that hadrons are in fact

[1] These patterns correspond actually to the diagram of the roots of the Lie algebra A_2 (Lie groups $SU(3)$ and $SU(3)/Z_3$).

composed of even more elementary constituents — the quarks — coming in three types (or "flavours"), forming a triangular "Eightfold Way" pattern: the u (for "up"), the d (for "down") and the s (originally "sideways", but now renamed "strange"). According to this model every baryon is composed of three quarks (antiquarks for the antibaryons) and every meson is composed of a quark and an antiquark. Hadron patterns can then be constructed by simple combination of these constituents, as suggested in Figure 1.1.

Since then, three other quark flavours have been observed as the explored energy range became wider: the c (for "charm") theoretically postulated in the GIM mechanism and then observed with the J/ψ resonance in 1974 independently by S. Ting and B. Richter [11], [12]; the b (for "beauty") observed in 1976 with the Υ resonance by L. Lederman [13] and felt as a member of a third quark generation as in the lepton picture (since the observation of the τ in 1975 by M. Perl [14]); and the t (for "top") "required" to complete the third family and effectively observed in 1995 by the CDF and D0 collaborations [15],[16].

Following the original Eightfold Way philosophy the flavour symmetry group had to be extended to SU(6). Nevertheless the symmetry is only approximate even in the SU(3) subgroup (u, d, s quarks strong isospin / hypercharge symmetry) with mass splittings within the baryon decuplet of $\sim 30\%$.

Spoiled symmetry and masses

In the minimal Standard Model (SM), the particles cited so far and their interactions obey the structure given by the non-abelian local gauge symmetry theory ruling the model (see next section)

$$SU(3)_c \otimes SU(2)_L \otimes U(1)_Y \tag{1.1}$$

which implies *massless* fermions and gauge fields, contrary to the experimentally observed mass spectrum shown in Table 1.1.

A mechanism which explicitely breaks the electroweak $SU(2)_L \otimes U(1)_Y$ part and generates massive leptonic and bosonic fields[2] is discussed in the next section: the Higgs mechanism. Nonetheless such a mechanism does not provide an explanation for the observed hierarchy of quark masses. These masses are then input parameters of the model.

On the other hand, the equality in the number of quark and lepton generations (families) is needed to obtain the triangular anomaly cancellation [17]. As illustrated above in the case of the b quark evidence, if a fourth generation of leptons were to be detected, one would expect to encounter a fourth generation of quarks as well. However, experimental studies of the Z decay at LEP have

[2]Massive quarks can naturally be accounted for in QCD itself as far as $SU(3)_c$ is concerned, but would then result from the dynamic symmetry breaking of chiral symmetry (leading to the chiral anomaly).

classification		doublet	mass	charge Q	weak-isospin hypercharge	colour
fermions	leptons	$\binom{e}{\nu_e}$	$\binom{511\text{ keV}}{0}$	$\binom{-1}{0}$	$\surd$	—
		$\binom{\mu}{\nu_\mu}$	$\binom{106\text{ MeV}}{0}$	$\binom{-1}{0}$	$\surd$	—
		$\binom{\tau}{\nu_\tau}$	$\binom{1.78}{0}$	$\binom{-1}{0}$	$\surd$	—
	quarks	$\binom{u}{d}$	$\binom{1-5\text{ MeV}}{3-9\text{ MeV}}$	$\binom{+\frac{2}{3}}{-\frac{1}{3}}$	$\surd$	$\binom{\text{RGB}}{\text{RGB}}$
		$\binom{c}{s}$	$\binom{1.15-1.35}{75-170\text{ MeV}}$	$\binom{+\frac{2}{3}}{-\frac{1}{3}}$	$\surd$	$\binom{\text{RGB}}{\text{RGB}}$
		$\binom{b}{t}$	$\binom{4.0-4.4}{174\pm5}$	$\binom{+\frac{2}{3}}{-\frac{1}{3}}$	$\surd$	$\binom{\text{RGB}}{\text{RGB}}$
mediators	e.m.	γ	0	0	$\surd$	—
	weak	$W^\pm$	80.4	±1		—
		Z^0	91.2	0		—
	strong	g	0	0	—	RGB$\otimes\overline{\text{RGB}}$
generations		2.9841 $\pm$ 0.0083 (LEP)				

Table 1.1: Elementary particles of the Standard Model. The last line shows the result of the measurement of the Z^0 lineshape at LEP, in particular its decay width into invisible channels, giving the number of neutrino species. Masses are in [GeV] with $c = 1$ unless otherwise specified.

ruled out a fourth light neutrino, as reported in Table 1.1 [18]. While not an absolute proof, this is a remarkable hint in favour of just three generations and of some "completeness" of the observed particle spectrum.

1.2 Mass generation through symmetry breaking

Peter W. Higgs (1929-) is the eponymous author of a scenario he proposed in the early sixties to generate masses, which was subsequently adopted to account for the mass spectrum of the gauge fields [19]. Essentially the same conclusions have been reached independently by F. Englert and R. Brout [20]. The Higgs mechanism exploits a spontaneous breaking of the gauge symmetry (SSB) and postulates the existence of *extra* (scalar) fields which are used to generate masses for the bosons. The fermion masses arise then through a coupling to these fields. We say the symmetry is *spontaneously* broken, as no external influence is necessary.

The gauge boson acquires mass at the expense of the massless quantum of excitation generated by a continuous symmetry transformation[3]. For each vector

[3]This last statement is known as the Goldstone theorem: if a theory has a continuous

gauge field that gets massive one complex scalar field is needed, one piece of which becomes the longitudinal mode of the vector field leaving one real scalar physical field, the Higgs boson.

1.2.1 SSB and the Higgs mechanism

Let's illustrate the mechanism by putting forward, in the simplest case, a Lagrangian term typical of what would be necessary to be added to the theory to achieve our goal. To do so we consider a self-interacting complex scalar field ϕ, as in the $\lambda\phi^4$ model (with units where $\hbar = c = 1$) [17].

The $\lambda\phi^4$ Lagrangian density can be written in the classical $\mathcal{L} = \mathcal{T} - \mathcal{U}$ form

$$\mathcal{L} = \underbrace{\frac{1}{2}(\partial_\mu \phi)^*(\partial^\mu \phi)}_{\mathcal{T}} - \underbrace{\left(-\frac{1}{2}\mu^2(\phi^*\phi) + \frac{1}{4}\lambda(\phi^*\phi)^2\right)}_{\mathcal{U}} \quad \begin{array}{l}\text{with } \phi \equiv \phi_1 + \imath\phi_2 \\ \text{and } \mu^2 > 0, \lambda > 0\end{array} \tag{1.2}$$

with the first term being the kinetic energy of the field, the second term looking like a mass term (but with the wrong sign, i.e. an imaginary mass) and the third one being a (self-)interaction with a coupling intensity given by λ.

The proposed Lagrangian exhibits invariance under the $U(1)$ global phase transformation. Actually, the Lagrangian can be required to even be invariant under local[4] gauge transformations

$$\phi \to e^{\imath q\theta(x)} \cdot \phi \tag{1.3}$$

by introducing a massless gauge field A^μ with minimal coupling[5],

$$\mathcal{L} = \frac{1}{2}(\mathcal{D}_\mu \phi)^*(\mathcal{D}^\mu \phi) + \frac{1}{2}\mu^2(\phi^*\phi) - \frac{1}{4}\lambda(\phi^*\phi)^2 - \frac{1}{16}F^{\mu\nu}F_{\mu\nu} \tag{1.4}$$

where $\mathcal{D}_\mu = \partial_\mu + \imath q A_\mu$ is the covariant derivative, $F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$ the field-strength tensor and q the "charge" representing the coupling constant between interacting fields (e.g. the *electric* charge in the U(1) case).

To find an adequate interpretation of the mass term, one has to keep in mind that a particle is an excitation of the ground state (the "vacuum") in quantum field theory.

---

symmetry of the Lagrangian which is not a symmetry of the vacuum, there must exist one or more massless bosons (Goldstone bosons).

[4]There is no compelling physical argument for why a global invariance should necessarily hold locally. In some sense, conversely, one should be able to carry out gauge transformations independently at space-like separated points (which are after all, out of communication with one another).

[5]In QED, A^μ is the field of the photon. The field has indeed to be massless without what U(1) symmetry will be spoiled and consequently the charge conservation (the U(1) Noether current).

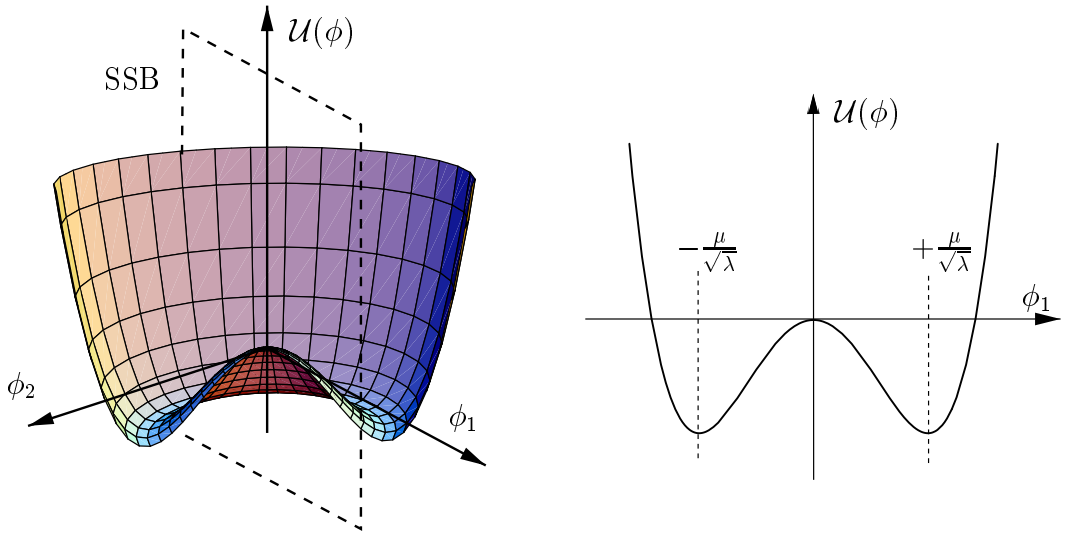


Figure 1.2: The $\lambda\phi^4$ potential (and so the ground state) is symmetric under $U(1)$. A particular choice of vacuum state breaks the symmetry; if this choice is arbitrary, we have SSB.

As the locally invariant Lagrangian keeps the global symmetry (it is always possible to rewrite $\theta(x) \to \rho + \theta(x)$ with $\rho =$ constant), it turns out that the trivial field $|\phi| = 0$ is not suited to describe the "vacuum", for the minima $\phi_{\min}$ of $\mathcal{U}$ are infinitely degenerate

$$\frac{\partial \mathcal{U}}{\partial \phi}(\phi_{\min}) = 0 \Rightarrow |\phi|^2 = \phi_{1_{\min}}^2 + \phi_{2_{\min}}^2 = \frac{\mu^2}{\lambda}, \tag{1.5}$$

lying on a circle of radius $\mu/\sqrt{\lambda}$ as depicted in Figure 1.2.

To express deviations from one of these ground states ϕ_{vac} (e.g. $\rho = 0$ in the previous brackets), we introduce a new field variables $\phi = \phi_1 + \imath\phi_2 = (\phi_1; \phi_2) \to \phi_{\text{vac}} + (\eta; \xi)$ so that:

$$\phi_{\text{vac}} = (\phi_{1_{\min}}; \phi_{2_{\min}}) = (v; 0) \qquad (\eta; \xi) = (\phi_1 - v; \phi_2)$$

with

$$v \equiv \langle 0|\phi|0\rangle = \frac{\mu}{\sqrt{\lambda}} \quad \text{, real}$$

as the resulting vacuum expectation value of the field ϕ. It has to be noticed that our explicit choice for ϕ_{vac} is affected by U(1) transformations: although arbitrary, it broke the symmetry.

Now the Lagrangian 1.2 becomes:

$$\begin{aligned}\mathcal{L} = &\left[\frac{1}{2}(\partial_\mu\eta)(\partial^\mu\eta) - \lambda v^2\eta^2\right] + \left[\frac{1}{2}(\partial_\mu\xi)(\partial^\mu\xi)\right] \\ &+ \left[-\frac{1}{16\pi}F^{\mu\nu}F_{\mu\nu} + \frac{1}{2}(qv)^2 A_\mu A^\mu\right] - 2\imath qv(\partial_\mu\xi)A^\mu \\ &+ \Big\{q[\eta(\partial_\mu\xi) - \xi(\partial_\mu\eta)]A^\mu + q^2 v\eta(A_\mu A^\mu) + \frac{1}{2}q^2(\xi^2+\eta^2)(A_\mu A^\mu) \\ &-\lambda v(\eta^3 + \eta\xi^2) - \frac{1}{4}\lambda(\eta^4 + 2\eta^2\xi^2 + \xi^4)\Big\} + \frac{1}{4}\lambda v^4 \quad (1.6)\end{aligned}$$

where the first line describes a scalar particle η of mass $\sqrt{2\lambda}v$ and a massless Goldstone boson ξ; the second line the free gauge field A^μ, which has *acquired* a mass

$$m_A = qv \quad (1.7)$$

and the third and fourth lines specify the various couplings of ξ, η and A^μ. The last term, a constant energy density, plays no rôle from the viewpoint[6] of the field dynamics as it will disappear from the equations of motion.

However an unnatural term remains, bilinear in two different fields, which apparently allows the gauge field to transform into a ξ when it propagates

$$-2\imath qv(\partial_\mu\xi)A^\mu$$

indicating badly chosen physical states. More importantly, the bookkeeping of the number of degrees of freedom between lagrangians 1.4 (four) and 1.6 (five) reveals that our particular choice for the vacuum introduced an extra degree of freedom. One can remedy to this problem by observing that that bilinear term indicates redundancy in the counting of the degrees of freedom. The seemingly unphysical Goldstone boson can indeed be "gauged away" by an appropriate choice of θ in 1.3:

$$\theta = -\frac{1}{q}\arctan\left(\frac{\phi_2}{\phi_1}\right)$$

and Lagrangian 1.6 can be rewritten (as it is invariant) by setting $\xi = 0$:

$$\begin{aligned}\mathcal{L} = &\left[\frac{1}{2}(\partial_\mu\eta)(\partial^\mu\eta) - \lambda v^2\eta^2\right] + \left[-\frac{1}{16\pi}F^{\mu\nu}F_{\mu\nu} + \frac{1}{2}(qv)^2 A_\mu A^\mu\right] \\ &\left\{q^2 v\eta(A_\mu A^\mu) + \frac{1}{2}q^2\eta^2(A_\mu A^\mu) - \lambda v\eta^3 - \frac{1}{4}\lambda\eta^4\right\} + \frac{1}{4}\lambda v^4 \quad (1.8)\end{aligned}$$

[6]The term $|\mathcal{U}^0_{\text{Higgs}}| \equiv \frac{1}{4}\lambda v^4$ can be interpreted as a contribution to the vacuum energy and would as such contribute to the cosmological constant of general relativity by a dramatic amount. See further discussion in Section 1.4.

We are left with a single real (and hence neutral) massive scalar η — the "Higgs" particle — and a massive gauge field A^μ having gained a longitudinal degree of freedom by absorption of the Goldstone boson.

All we have done is sacrify the manifest symmetry of 1.2 in favour of a notation which makes the physical content more explicit. To put it the other way around, the true symmetry of the system is "hidden" by the arbitrary selection of a particular (asymmetrical) ground state.

The whole of the above analysis can be adapted to any non-Abelian gauge theory.

1.2.2 The Higgs mechanism in the SM

Going from $\lambda\phi^4$ to full SM, an analogous Higgs mechanism would be responsible for the masses of the weak interaction gauge bosons $W^\pm$ and Z^0 and for the quarks and leptons (via Yukawa couplings) in the GWS[7] electroweak model, though the Higgs potential $\mathcal{U}(\phi)$ is yet unknown (although v can be obtained from a measurement of the mass 1.7 of the gauge field A_μ, the parameter λ is not constrained).

The abelian Higgs mechanism described above has then to be extended to the appropriate symmetry, local $SU(2)_L \otimes U(1)_Y$, where $SU(2)_L$ refers to weak isospin (giving the triplet of gauge fields $\boldsymbol{W_\mu}$, and $\boldsymbol{I_W}$ as generators), involving only left-handed particles, while $U(1)_Y$ refers to weak hypercharge (gauge field B_μ, generator Y_W) in which particles of both chiralities are involved. Neutrinos with zero mass[8] are assumed in the following and are therefore uniquely assigned to left-handed doublets.

The above cited quantities satisfy the Gell-Mann–Nishijima relation

$$Q = I_{3W} + \frac{1}{2}Y_W$$

under which operator the vacuum state has to remain invariant to ensure electric charge conservation. In accordance with the symmetry the Higgs field is now a doublet:

$$\phi = \begin{pmatrix} \phi^+ \\ \phi^0 \end{pmatrix} \qquad \phi_{\text{vac}} = \begin{pmatrix} 0 \\ v + H \end{pmatrix}$$

The corresponding Lagrangian density is given by 1.11 where the first line describes the kinetic energies of the gauge fields and the second line the fermions kinetic energies and their interactions with the gauge fields.

[7] For Glashow, Weinberg and Salam.

[8] There is now rather convincing evidence that neutrinos have non-zero mass from the apparent observation of neutrino oscillations [21]. One will stay here with the original GWS description of the Standard Model, with massless neutrinos.

The "Higgs" Lagrangian 1.12 including the kinetic terms in 1.11 is of the same form as in 1.4 and is responsible for the mass generation of the gauge and Higgs bosons ($W^{\pm}, Z^0, \gamma, H$) and their mutual couplings.

To complete the SM, the fermions have to acquire mass. It is done by coupling the fermions to the Higgs boson through the $SU(2)_L \otimes U(1)_Y$ invariant Lagrangian 1.13 (called Yukawa couplings).

The mechanism however does not help to calculate these masses since the Yukawa coupling constants $g_{1,f}, g_{2,f}$ are themselves unknown, but satisfy the relation:

$$m_f = \frac{g_{i,f}\, v}{\sqrt{2}} \qquad \begin{array}{l} i = 1 \text{ for } f \text{ up} \\ i = 2 \text{ for } f \text{ down} \end{array} \tag{1.9}$$

The conjugate Higgs doublet is needed, as for the quarks the second member of the doublet is massive.

The last Lagrangian density 1.14, invariant under $SU(3)_c$, describes the strong interaction, namely the gluons kinetic energies and the quark-gluon couplings.

The SM lagrangian density shown here is not in its explicitely broken form.

$$\mathcal{L}_{\text{SM}} = \mathcal{L}_{\text{GWS}} + \mathcal{L}_{\text{Higgs}} + \mathcal{L}_{f\,\text{mass}} + \mathcal{L}_{\text{QCD}} \tag{1.10}$$

$$\begin{aligned} \mathcal{L}_{\text{GWS}} = & -\frac{1}{4}\boldsymbol{W_{\mu\nu}} \cdot \boldsymbol{W^{\mu\nu}} - \frac{1}{4} B_{\mu\nu} \cdot B^{\mu\nu} \\ & + \bar{L}\gamma^{\mu}(\imath\partial_{\mu} - \frac{g}{2}\boldsymbol{\tau} \cdot \boldsymbol{W_{\mu}} - \frac{g'}{2} Y B_{\mu}) L + \bar{R}\gamma^{\mu}(\imath\partial_{\mu} - \frac{g'}{2} Y B_{\mu}) R \end{aligned} \tag{1.11}$$

$$\mathcal{L}_{\text{Higgs}} = \left| (\imath\partial_{\mu} - \frac{g}{2}\boldsymbol{\tau} \cdot \boldsymbol{W_{\mu}} - \frac{g'}{2} Y B_{\mu})\, \phi \right|^2 - \mathcal{U}(\phi) \tag{1.12}$$

$$\mathcal{L}_{f\,\text{mass}} = -\left(g_{1,f} \bar{L} \phi R + g_{2,f} \bar{L} \tilde{\phi} R + \text{H.C.}\right) \tag{1.13}$$

$$\mathcal{L}_{\text{QCD}} = -\, G^a_{\mu\nu} \cdot G^{\mu\nu}_a + \frac{g_s}{2} (\bar{\psi}^j_q \gamma^{\mu} \lambda^a_{jk} \psi^k_q) G^a_{\mu} \tag{1.14}$$

where L represents a left-handed fermion doublet, R a right-handed singlet, ϕ ($\tilde{\phi}$) the Higgs doublet and its conjugate, and $\boldsymbol{\tau}$ the Pauli matrices. The quarks colour fields are represented by ψ^j_q with λ^a_{jk} the Gell-Man matrices, and G^a_{μ} the gluon gauge fields. The coupling constants of the various interactions are g, g' (electroweak) and g_s (strong). The term 1.13 prevents the appearance of a mass term and interaction term for the (left-handed only) neutrinos. Manifestly, the overall Lagrangian $\mathcal{L}_{\text{SM}}$ 1.10 is invariant under the symmetry group mentioned in 1.1.

Out of this model can be extracted predictions on several physical quantities:

$$m_W = \frac{1}{2} v g \quad \text{with} \quad v = (\sqrt{2} G_F)^{-1/2} \sim 246 \text{ GeV}$$

the vacuum expectation value depending only on G_F, giving

$$m_W = \left(\frac{\pi\alpha_{\rm em}}{\sqrt{2}G_F}\right)^{1/2} \frac{1}{\sin\theta_W} \qquad \text{with} \quad \alpha_{\rm em} = \frac{e^2}{4\pi} \tag{1.15}$$

The Weinberg angle θ_W must be measured experimentally. An estimate is

$$\sin^2\theta_W \sim 0.23$$

which gives at tree-level, in agreement with the direct mass measurements shown in Table 1.1

$$m_W \approx 80 \text{ GeV/c}^2 \qquad \text{and} \qquad m_Z = \frac{m_W}{\cos\theta_W} \approx 91 \text{ GeV/c}^2$$

and the coupling constants

$$g = \frac{e}{\sin\theta_W} = (\sqrt{2}G_F)^{1/2} \cdot m_W \qquad g' = \frac{e}{\cos\theta_W} = \frac{\sqrt{4\pi\alpha_{\rm em}}}{\cos\theta_W} \qquad g_s = \sqrt{4\pi\alpha_s} \tag{1.16}$$

Thus finally, the model in the version depicted here requires as input nineteen independent parameters:

- 3 charged-lepton masses, 6 quark masses,
- 3 gauge coupling constants (or e, θ_W and α_s),
- 3 quark-mixing angles and 1 complex phase,
- the Higgs mass and the vacuum expectation value v,
- the QCD vacuum angle (a consequence of which is the non-zero electric dipole of the neutron).

The precise measurement of each of these parameters is the object of permanent efforts along with the test of scenarios enlarging the perspectives (one of which is discussed in Section 1.4) of this minimal and somehow unsatisfactory SM given the number of questions it leaves without answer. In particular, the measurement of the three quark-mixing angles and complex phase is giving rise for a decade to a vast experimental program in the community, as is discussed in the next chapter with the LHCb experiment at CERN.

It should be stressed that despite its great success, the SM will not be proved until the symmetry breaking mechanism is tested and a possible Higgs boson (or equivalent) is discovered. Whether or not the Higgs mechanism is indeed the correct symmetry-breaking mechanism remains to be assessed. All we have shown up to here is that in its simplest form the theory requires the existence of an electrically neutral scalar Higgs boson. Many extensions of the SM retain the Higgs mechanism as the primary method for mass generation for gauge bosons, but with more complicated Higgs sectors and more Higgs bosons (e.g. supersymmetry).

SM	All fermions	WW	ZZ
H	$\frac{e\, m_f}{2 m_W \sin\theta_W}$	$\frac{e\, m_W}{\sin\theta_W}$	$\frac{e\, m_Z}{\sin\theta_W \cos\theta_W}$

Table 1.2: SM Higgs couplings to fermions and massive gauge bosons. The expressions can be rendered dimensionless with appropriate definitions [22].

1.3 Higgs production and decays

The SM does not predict the mass of the Higgs boson, rendering its search a real challenge for experimentalists. Nonetheless, regions of mass can be excluded as is discussed in Section 1.3.3. For instance, a lower bound can be obtained by comparing the experimental data with the theoretical expectations for the production (and probable decays) of the Higgs particle.

1.3.1 Decays

Let's recall that the Higgs coupling constant to fermions in 1.9 is fixed:

$$g_{f\bar{f}H}\bar{\psi}_f H \psi_f \qquad \text{with} \qquad g_{f\bar{f}H} = \frac{g_{i,f}}{\sqrt{2}} = (\sqrt{2}G_F)^{1/2} \cdot m_f$$

with the property of being proportionnal to the fermion mass. By comparison, 1.12 and 1.16 indicate a coupling to the gauge bosons proportional to the square of their mass, as summarized in Table 1.2.

Because of its simple coupling the calculation of the Higgs decay widths can be carried out without difficulties and gives at leading order:

$$, (H \to f\bar{f}) = N_{\text{cf}} \frac{G_F m_f^2 m_H}{4\pi\sqrt{2}} \left(1 - \frac{4m_f^2}{m_H^2}\right)^{3/2} \tag{1.17}$$

$$, (H \to W^+W^-) = \frac{G_F m_H^3}{8\pi\sqrt{2}} \left(1 - \frac{4m_W^2}{m_H^2}\right)^{3/2} \cdot \left(1 - \frac{4m_W^2}{m_H^2} + 12\frac{m_W^4}{m_H^4}\right) \tag{1.18}$$

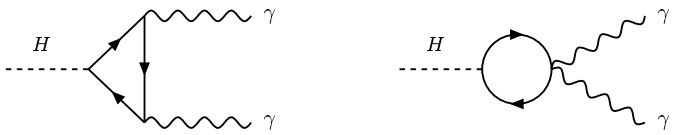
$$, (H \to Z^0Z^0) = \frac{1}{2}, (H \to W^+W^-)\Big|_{m_W \to m_Z} \tag{1.19}$$

where N_{cf} is the number of colours (3 for quarks and 1 for leptons). It can be noticed that due to its privileged coupling to heavy objects, diagrams involving coupling to light fermions pairs are suppressed in a ratio:

$$\frac{g_{WWH}}{g_{f\bar{f}H}} \simeq \frac{1}{2}\frac{m_W^2}{m_f} >> 1 \tag{1.20}$$

for heavy enough Higgs particles, and top quark excluded.

The decay mode $H \to \gamma\gamma$ is especially relevant for Higgs bosons in the intermediate mass range $m_W < m_H < 2m_Z$ [23]. It occurs via one-loop diagrams involving W bosons, quarks and leptons



Spins $0, \frac{1}{2}, 1$ Spins $0, 1$

allowed by the scalar structure of the Higgs. The result of a lengthy calculation [24] gives:

$$\begin{aligned} , (H \to \gamma\gamma)[\text{GeV}] &\simeq \frac{G_F}{\sqrt{2}} \frac{m_H^3}{8\pi} \left(\frac{\alpha_{\text{em}}}{4\pi}\right)^2 \cdot \mathcal{F}_{\text{spin}}(m_H) \\ &\simeq 1 \times 10^{-13} (m_H[\text{GeV}])^3 \cdot \mathcal{F}_{\text{spin}}(m_H) \end{aligned} \tag{1.21}$$

where $\mathcal{F}_{\text{spin}}(m_H)$ represents the spin terms. The corresponding branching is expected to remain small ($\mathcal{O}(10^{-3})$) over the intermediate mass range but has the advantage to provide a clear signature and benefits from the suppression of such a decay channel for the Z^0 (Landau-Yang theorem violating, and provided a good e/γ separation could be achieved).

Finally, the Higgs decay into gluons is analogous and can be obtained from 1.21 keeping just the quark loop contributions and replacing the coupling constant. Keeping the main contribution, a reasonable estimate for $, (H \to gg)$ is

$$, (H \to gg)[\text{GeV}] \simeq 1.5 \times 10^{-10} (m_H[\text{GeV}])^3 (\alpha_s/0.15)^2 n_h^2$$

where n_h is the number of quarks heavier than the Higgs.

All of the above results are summarized in Figure 1.4 (a) and (b), showing the predicted branching ratios and total width of the Higgs as a function of its mass.

1.3.2 Production

According to what has been described for the Higgs couplings and decays, it is possible to determine the various channels through which Higgs particles would be produced.

The dominant Higgs production mechanisms in hadron collisions are indicated in Figure 1.3.

The cross section for Higgs production involving light hadrons (last diagram in Figure 1.3) is extremely suppressed as shown by expression 1.20. On the same basis, the gluon fusion reaction $gg \to H$ is much more favorable when involving

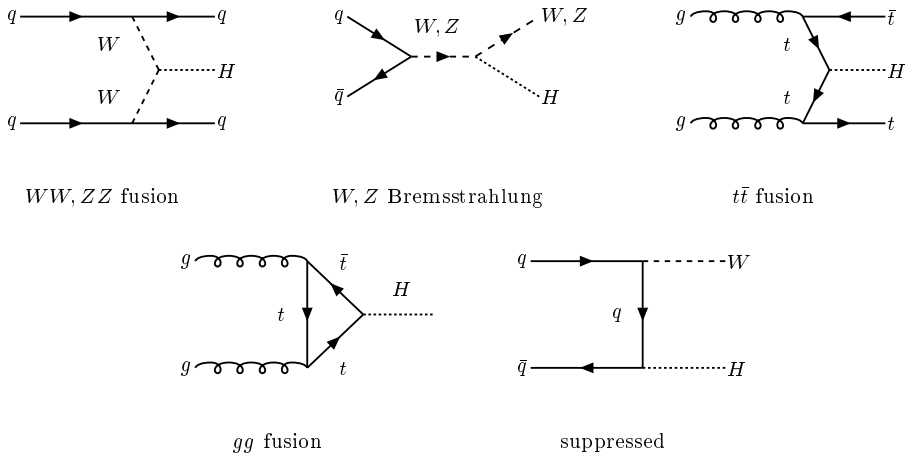


Figure 1.3: Dominant diagrams for SM Higgs production in hadron collisions.

heavy quarks in the loop. Gluon fusion is even the dominant production mechanism for a top quark loop and favoured by the simple phase space configuration of the final state.

Along with the knowledge of the parton distribution functions (pdf), one can compute the Higgs production cross sections. Figure 1.4 (c) shows the different contributions as a function of the Higgs mass for pp collisions at $\sqrt{s} = 14$ TeV, with a top mass $m_t = 175$ GeV [25]. Calculations include the next-to-leading order QCD corrections ($t\bar{t}H$ excepted). The uncertainty associated to the pdf's has been evaluated to be $\pm 20\%$.

The experimental approaches for Higgs hunting can already be stated. The factor m_f^2 in 1.17 implies that decays into heavier particles would be favoured and H would appear as a (small) peak in the continuous mass distribution of the pair. But the tiny production cross section implies that very high statistics experiments are needed.

1.3.3 Precision measurements and bounds on m_H

It has been said that the Higgs mass is a free parameter in the SM. Nevertheless, bounds on m_H can be extracted from precision measurements of the electroweak parameters through its effect in radiative corrections. Moreover, the direct search for Higgs particles and the measurement of the top mass by CDF an D0 allowed

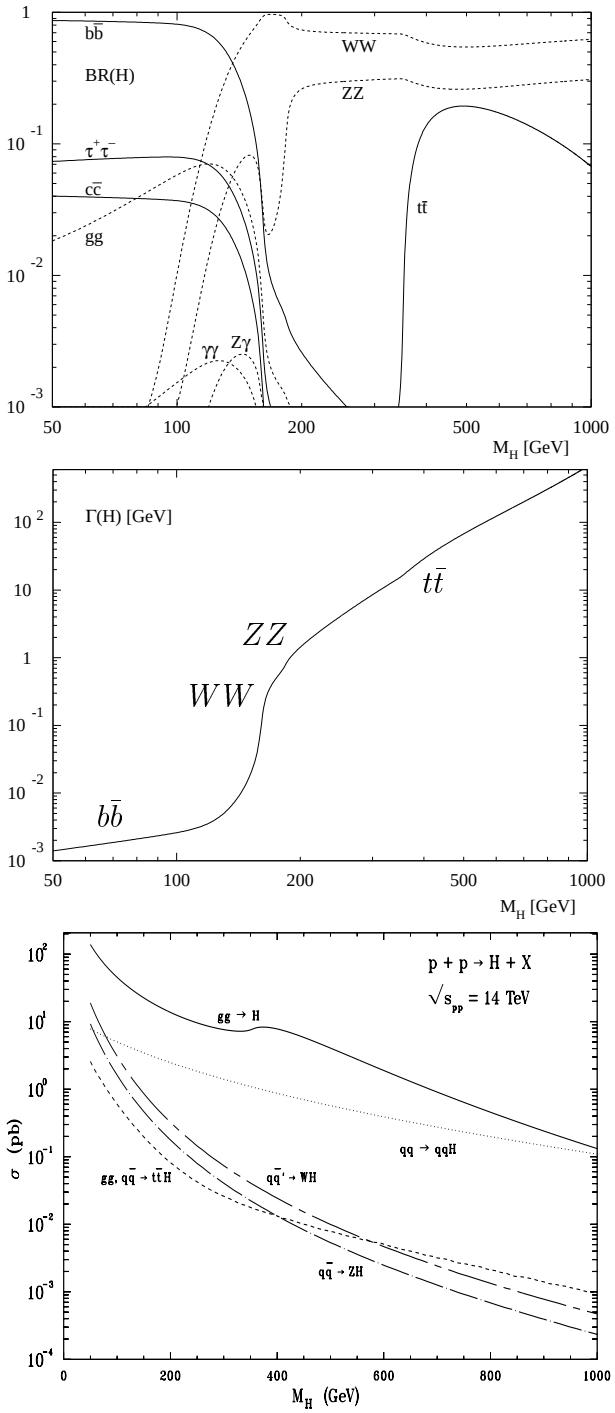


Figure 1.4: (a) Branching fractions of the main decay channels of the SM H^0 versus m_H. (b) Width of the Higgs particle versus m_H. The various kinks in the curve correspond to the opening of new decay channels as shown in (a). (c) Higgs production reactions in pp collisions at $\sqrt{s} = 14$ TeV versus m_H. From [25].

to significantly shrink the likely mass range over the last decade:

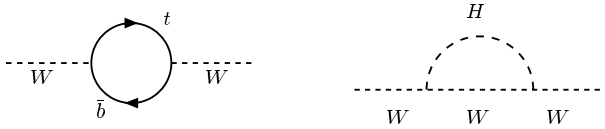
$$\begin{aligned} 1989: &\quad 0 \text{ GeV} \lesssim m_H \lesssim 1 \text{ TeV} \\ 1990: &\quad 40 \text{ GeV} \lesssim m_H \lesssim 1 \text{ TeV} \\ 2000: &\quad 113 \text{ GeV} < m_H < 170 \text{ GeV} \quad \text{CL=95\%}, \end{aligned}$$

the current lower limit coming from the direct search at LEP II (see Section 1.5 for a description of the latest results).

The lowest order relation 1.16 with 1.15 between the Fermi coupling constant G_F, the electromagnetic coupling constant α and m_W is modified by adding all the one-loop corrections to the muon decay diagram, resulting in a redefinition of G_F:

$$\frac{G_F}{\sqrt{2}} = \frac{e^2}{8 s_W^2 c_W^2 m_Z^2} \cdot (1 + \Delta r) \qquad \text{where} \quad \begin{aligned} s_W^2 &= \sin\theta_W^2 \\ c_W^2 &= \cos\theta_W^2 \end{aligned}$$

Δr involves vacuum polarization effects and other weak corrections which depend on all the parameters of the model, including the Higgs and top masses m_H and m_t. The most important diagrams are expected to be:



Boson self-energy Boson loop

The complete expression for Δr can be written as

$$\Delta r = \Delta\alpha - \frac{c_W^2}{s_W^2}\Delta\rho + \Delta r_{\text{rem}}$$

where, to one loop and for heavy Higgs particles ($m_H \gg m_W$), the largest contributions have the following form:

$$\Delta\alpha = 1 - \frac{\alpha}{\hat{\alpha}(m_Z)} \tag{1.22}$$

$$\Delta\rho \sim \frac{3e^2}{64\pi^2 s_W^2 c_W^2 m_Z^2} \cdot m_t^2 \tag{1.23}$$

$$\Delta r_{\text{rem}} = \Delta r_{\text{rem}}^{\text{Higgs}} + \Delta r_{\text{rem}}^{\text{top}} \tag{1.24}$$

$$\Delta r_{\text{rem}}^{\text{Higgs}} = \frac{e^2}{64\pi^2 s_W^2}\frac{11}{3}\left[\ln\left(\frac{m_H^2}{c_W^2 m_Z^2}\right) - \frac{5}{6}\right] \tag{1.25}$$

$$\Delta r_{\text{rem}}^{\text{top}} = \frac{e^2}{64\pi^2 s_W^2} 2\left(\frac{c_W^2}{s_W^2} - \frac{1}{3}\right)\ln\left(\frac{m_t^2}{c_W^2 m_Z^2}\right) \tag{1.26}$$

The $\Delta\alpha$ term results from the need to have a running coupling constant to account for the effect of $f\bar{f}$ self-energy loops in the photon propagator. Thus

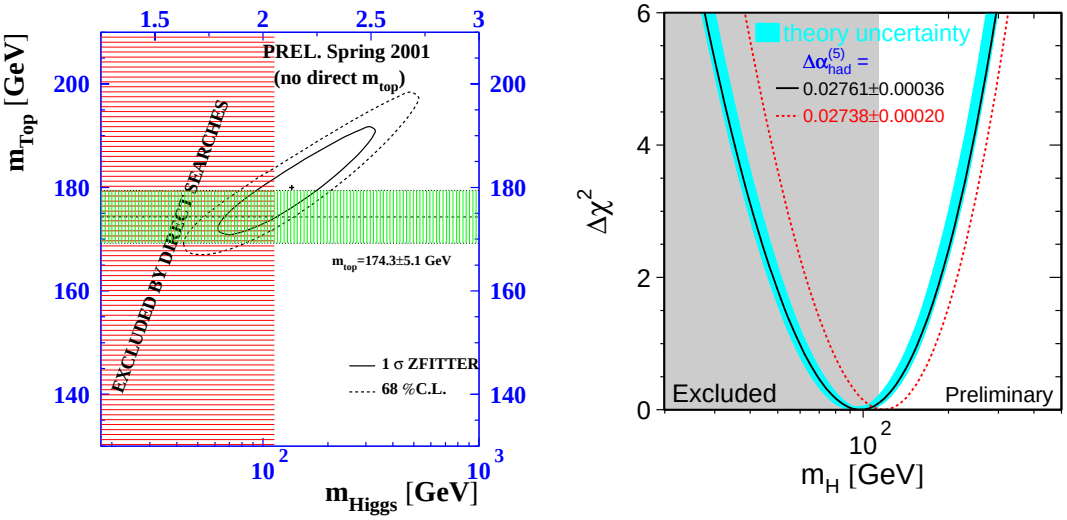


Figure 1.5: Fit of the precision measurements of the SM electroweak parameters [18].

radiative corrections to electroweak parameters give a leading quadratic sensitivity on the top mass $\mathcal{O}(m_t^2/m_Z^2)$ but only a logarithmic sensitivity on the Higgs mass $\mathcal{O}(\ln(m_H/M_Z))$.

SM predictions are calculated using various semi-analytical programs. To test the agreement between the data and the SM, one performs a fit to the LEP data including the experimental determination of m_W but leaving m_H and m_t as free parameters [18]. This yields confidence intervals on the left of Figure 1.5 in very good agreement with the result of the direct measurement of m_t by CDF and D0 (horizontal band). Also included is the result of direct search on m_H at LEP II (vertical band). The fit shows that LEP data prefer a light Higgs boson, albeit with very large (and asymmetric) uncertainties due to the $\ln(m_H)$ sensitivity.

The observed $\Delta\chi^2 \equiv \chi^2 - \chi^2_{\min}$ of the fit to all data as a function of m_H, including direct measurements of m_W (LEP II) and m_t (CDF,D0) is shown on the right of Figure 1.5. That fit results in $\log(m_H/\mathrm{GeV}) = 1.99^{+0.20}_{-0.21}$ corresponding to $m_H = 98^{+58}_{-38}$ GeV. The uncertainty in $\alpha(m_Z^2)$ arising from the contribution of the light quarks to the photon vacuum polarization ($\Delta\alpha^{(5)}_{\mathrm{had}}(m_Z^2)$) causes a (non-dominant) error of 0.1 on the fitted value of $\log(m_H)$, included in the result. Preference of LEP data for a light Higgs is confirmed with a 95% CL upper limit given on the hypothesis $m_H < 212$ GeV/c^2. The shaded region shows the lower exclusion limit from the direct searches $m_H > 113.5$ GeV/c^2. Taking into account the estimate of the theoretical error due to missing higher order corrections (dark band), the central value of the (black) fit and the lower bound given by the direct

search are in good agreement.

1.4 Alternative models

Being a self-interacting scalar, the radiative corrections to the SM Higgs mass are quadratically divergent[9]. Absorbing them into renormalized parameters for cutoff values (say of the order of the Planck mass $M_{\mathrm{Pl}} \sim 10^{19}$ GeV) while keeping $m_H \sim \mathcal{O}(100\ \mathrm{GeV})$ at all perturbations orders is difficult and unnatural, though technically possible (fine-tuning techniques). This is the *gauge hierarchy problem* which typically led to search for alternatives to a single scalar Higgs and to suggest that new physics should exist at higher energy scale.

Another issue of the SM has been mentioned in Section 1.2.1 when discussing the vacuum energy term $|\mathcal{U}^0_{\mathrm{Higgs}}| \equiv \frac{1}{4}\lambda v^4$ in Equation 1.6. The Higgs vacuum energy would contribute to the cosmological constant by an amount $|\Lambda_{\mathrm{Higgs}}| = 8\pi\, G_{\mathrm{N}}\, \mathcal{U}^0_{\mathrm{Higgs}}$. Taking the present lower limit on the Higgs mass $m_H > 113\ \mathrm{GeV/c^2}$ (see Section 1.3.3), $|\Lambda_{\mathrm{Higgs}}| \sim 7 \times 10^{-7}\ \mathrm{cm^{-2}}$ would yield a contribution more than forty orders of magnitude larger than the present determination of the cosmological constant $\Lambda \sim \mathcal{O}(10^{-48})\ \mathrm{cm^{-2}}$ [21]. There must then be some important and non-trivial physics which forces the suppression or cancellation of the vacuum energy.

One of the proposed model is that of *supersymmetry* (SUSY) which establishes a symmetry between bosons and fermions. Each fermion and each vector boson of a gauge theory lagrangian will have a corresponding superpartner and the quadratic divergence in the Higgs self-energy is removed by cancellations between supersymmetric partners.

Despite of a few elegant features (solution to hierarchy problem, origin of CP-violation due to spontaneous symmetry breaking in the Higgs sector, cancellation of cosmological constant contribution, extension to gravity) SUSY models do not benefit yet from any direct observation of the superpartners (in particular the "lightest supersymmetric particle"). One may also wonder if a theory which doubles the particle spectrum really looks more fundamental.

1.5 Status of Higgs search

As discussed in Section 1.3.3, the Higgs has been searched for during several years. Before the 1990's, several experiments excluded various regions for m_H but all had weak points [26]. The first rigorous results have been produced by LEP I running at the Z^0 peak.

[9] The analogy with the $\lambda\phi^4$ model still holds, and as such the renormalization proceeds the same way.

In particular, the lower bound $m_H > 40$ GeV reached in 1990 is the result of the search for the "rare" $Z^0 \to Hf\bar{f}$ decay mode at LEP I.

From 1996, the phase 2 of LEP operation (LEP II, running at $\sqrt{s} > 2m_W$) allowed to exclude higher masses for a Higgs produced in association with a vector boson (see second diagram in Figure 1.3 and relation 1.20).

After an extra running period of two months, LEP II operation has been stopped on the 2$^{\text{nd}}$ of November 2000 with a hint of a signal.

Hint of the Higgs boson at LEP !

The combined result of the four LEP experiments (ALEPH, DELPHI, L3 and OPAL) for the 2000 run indicates a hint of a Higgs boson with mass 115 GeV/c^2 with a significance of 2.9σ [27]. The distribution of the events is in agreement with expectations in the four experiments and in the different Z^0 decay channels.

The result of each of the experiments and of the combination is obtained with a standard log-likelihood ratio [28]. The number of candidate events which then contribute a non-negligible amount to the total significance of the effect is found to be fourteen, divided into:

- nine four-jet events ($Hq\bar{q}$ channel)
- three missing energy candidate events ($H\nu\bar{\nu}$ channel)
- one leptonic candidate event ($H\ell^+\ell^-$ channel)
- one candidate with tau's ($H\tau^+\tau^-$ channel)

corresponding to an integrated signal-to-noise ratio slightly smaller than unity for an *expected* signal of seven events. Figure 1.6 shows a four-jet candidate event passing tight selection[10] recorded in ALEPH at a center-of-mass energy $\sqrt{s} = 206.7$ GeV [29]. The Higgs boson jets are both very well b-tagged with well measured displaced vertices, and give a reconstructed mass of 114.3 GeV/c^2. The measured invariant mass of the two *non* b-tagged jets is 92.1 GeV/c^2, consistent with a Z^0 boson. The missing momentum in the event points into the b-jet candidate containing an identified muon coming from the secondary vertex, as shown in the lower right closeup. This is a strong indication that, except for the unmeasured neutrino from the semileptonic b quark decay, the rest of the event is well measured. Finally, the measured invariant mass of the b-tagged jets and the missing momentum is 114.4 GeV/c^2, which renders unlikely the ZZ hypothesis.

With the data collected in the 2000 run, around 490 pb^{-1} combined for the four experiments at energies $\sqrt{s} > 206$ GeV, the likely mass range of the SM Higgs boson currently is (see [27] and Section 1.3.3)

[10] Meaning a large "weight", i.e. a large contribution to the (logarithm of the) likelihood ratio defined in the analysis.

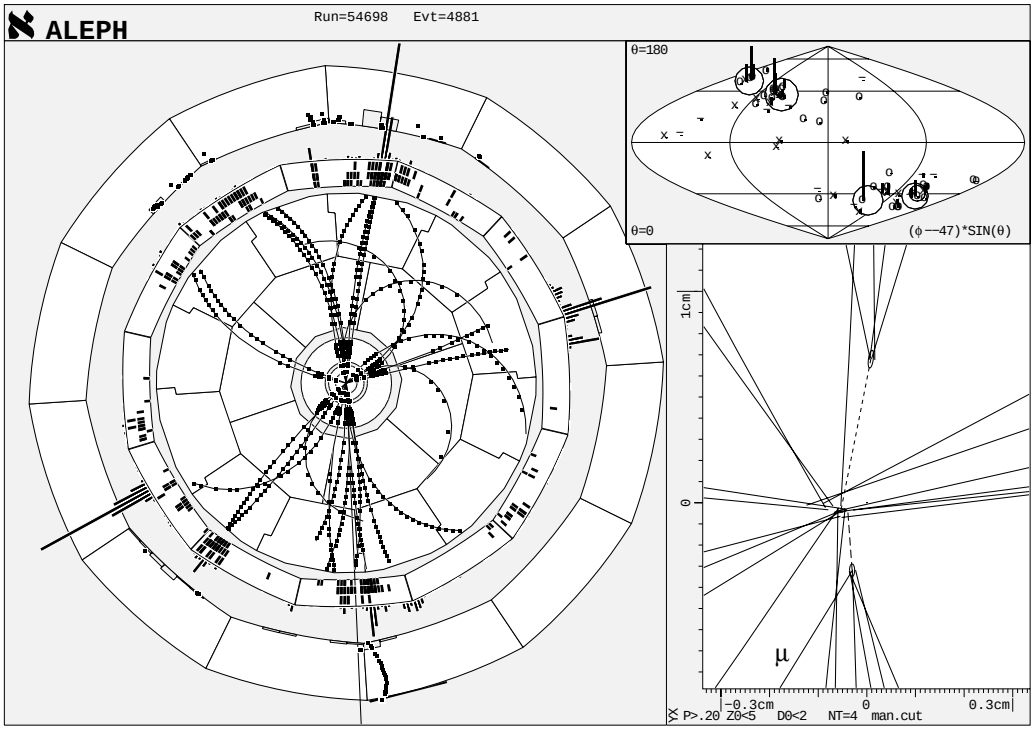


Figure 1.6: A four-jet Higgs boson candidate event seen in ALEPH with a reconstructed Higgs boson mass of 114.3 GeV/c^2 [29]. The two Higgs boson jets are well b-tagged. The event is shown in the view transverse to the beam direction, the θ-$\phi \sin\theta$ view, and in a closeup of the charged particles in the vertex region.

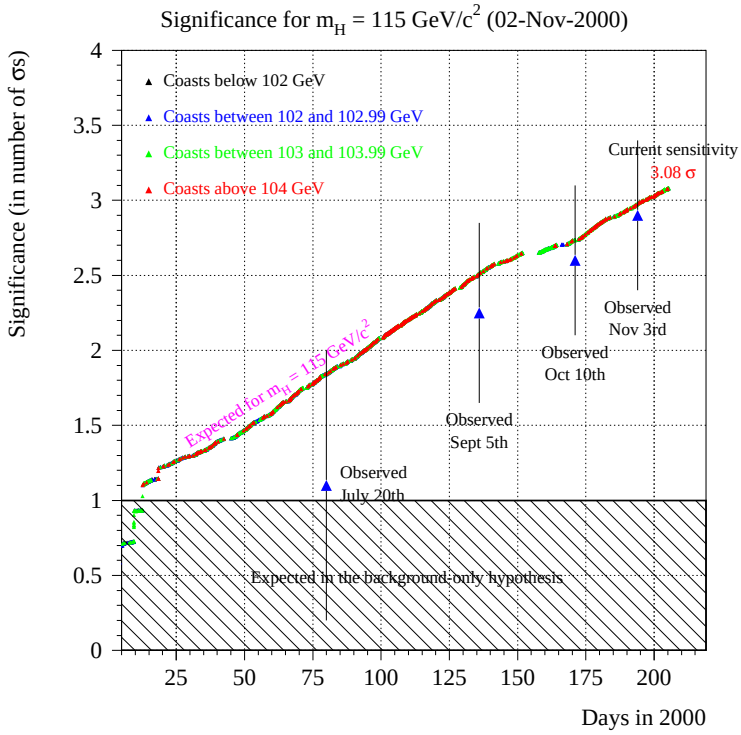


Figure 1.7: Increase of the observed combined LEP II data significance for $m_H = 115$ GeV/c^2 in 2000, compared with an online estimate of the significance expected in the signal-plus-background hypothesis [30]. The error bars are statistical only, with large point-to-point correlation. Also indicated (hatched) is the expectation in the background-only hypothesis.

$$\underset{\text{LEP II limit}}{113.5\ \text{GeV}} < m_H < \underset{\text{EW measurements}}{212\ \text{GeV}} \qquad \text{CL=95\%}$$

for an expectation value $m_H = 115.3$ GeV/c^2, and a simultaneous compatibility with the SM Higgs cross section for

$$m_H = 115.0^{+1.3}_{-0.9}\ \text{GeV/c}^2$$

Figure 1.7 shows the time evolution of the data significance for the $m_H = 115$ GeV/c^2 hypothesis during the 2000 run [30].

These results may give, depending on the reader's belief, a more or less strong conviction of the Higgs presence.

The high energy physics community has been focused on these results, and it is with much more motivation that the next generation of projects are being prepared — LHC at CERN and Tevatron II at FNAL.

Chapter 2

LHC and the LHCb experiment

It has been shown that LEP data are so accurate that they are sensitive to phenomena which occur at energies beyond those of the machine itself, giving a "preview" of exciting discoveries which may be made at higher energies. LHC is a machine designed to make these discoveries.

2.1 The LHC project

To explore physics up to the TeV scale, the next generation of experiments at CERN is under development within the LHC project. Given that most of the interesting physics requires high interaction rates, the construction of a proton-proton collider at a center-of-mass energy[1] $\sqrt{s} = 14$ TeV with a design luminosity of 10^{34} cm^{-2}s^{-1} is under way and should be commissioned in 2006. The project will benefit from the existing infrastructure, namely the 27 km long circular underground tunnel used for LEP under the Jura, and its versatile and well-known accelerator injection complex. So, it profits in terms of know-how and cost-saving.

In addition to pp operation, the LHC will be able to collide heavy nuclei (Pb-Pb) produced in the existing CERN accelerator complex, giving an energy of 1150 TeV in the center of mass (2.76 TeV/u and 7 TeV per unit charge). The LHC injectors, including the heavy ion route, are shown in Figure 2.1 together with the main design parameters of the LHC machine [31].

The LHC will be installed on the tunnel floor after removing the LEP equipments.

2.1.1 The layout

The basic layout of the machine mirrors that of LEP, with eight straight sections each approximately 538 m long, available for experimental insertions or utilities.

[1] So giving on average a $\sim$2 TeV scale for partonic interactions.

Machine circumference	[m]	26659
Beam energy	[TeV]	7
Dipole field	[T]	8.4
Luminosity	$[\mathrm{cm}^{-2}\mathrm{s}^{-1}]$	10^{34}
Number of bunches		2835
Particles per bunch		10^{11}
Bunch spacing	[ns]	25
Energy loss per turn	[keV]	6.7
crossing angle	[μrad]	200
r.m.s. IP beam radius	[μm]	16
r.m.s. IP length	[cm]	5.3
luminosity lifetime	[h]	10

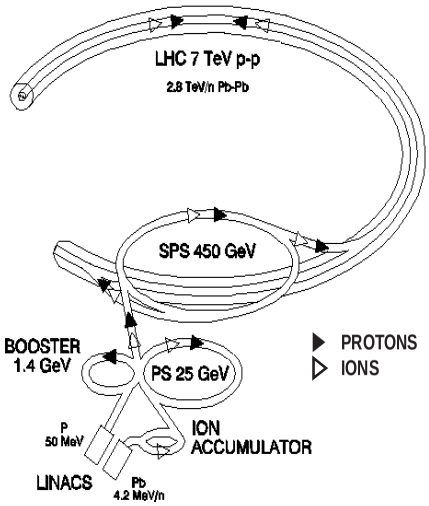


Figure 2.1: *Left:* Main parameters of the LHC machine. *Right:* LHC injectors

Four of these sections will have the beams crossing from one ring to the other and are therefore dedicated to experimental sites, two of which will also host the injection systems. Two insertions will contain collimation systems using only classical robust magnets (betatron and momentum cleaning). One insertion will contain the RF systems and the last straight section will contain the beam dump insertion to remove the beam safely from the collider at the end of a physics run, when the luminosity has degraded.

After the commissioning period, the LHC will deliver in 2006 beams for physics with a starting luminosity of 5×10^{32} $\mathrm{cm}^{-2}\mathrm{s}^{-1}$ to be steadily increased to its nominal value of 10^{34} $\mathrm{cm}^{-2}\mathrm{s}^{-1}$ over the first three years of operation.

The choice for a *proton* accelerator was driven by the fact that losses by synchrotron radiation for electrons of the same energy are prohibitive, as illustrated by LEP II beams whose energy has been pushed to the limit $\gtrsim 104$ GeV (intermittent) despite of a full use of superconducting technology. The huge RF power was then consumed just to compensate for the losses.

Two identical *proton* beams have been chosen to satisfy the high luminosity requirement. An antiproton beam would have simplified the technical conception but, given the low efficiency to produce them, the resulting luminosity would have been too small. As a consequence, and in order to manage with the room in the existing tunnel, the magnet configuration is unusual as shown in Figure 2.2. Two coil assemblies surrounding the two beam pipes are enclosed in the same iron yoke and cryostat. Given the radius of curvature of the orbit (imposed), the required huge operating field of 8.4 T can only be obtained at an acceptable cost by cooling the magnets to 1.9 K. There will be 1232 of such 14 m long dipole

1 beam pipe
2 superconducting coils
3 steel collars
4 iron yoke
5 iron yoke insert
6 shrinking cylinder / He II vessel
7 heat exchanger tube
8 dipole bus-bars
9 arc quadrupole
10 wires for protection/instrumentation

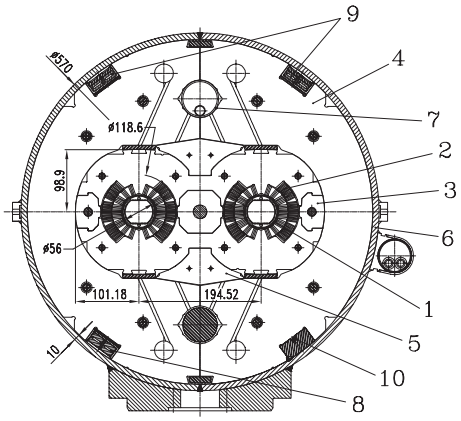


Figure 2.2: LHC main arc dipole description. Quoted lengths are in [mm].

magnets in the main arcs.

The synchrotron energy loss per turn amounts to 6.7 keV and is unsignificant for 7 TeV protons in terms of RF power load. However the emitted power of 3.7 kW cannot be neglected as it has to be absorbed by the beam pipe working at cryogenic temperature, and will then affect the power of the refrigeration system. An additionnal issue is the release of absorbed gas molecules when synchrotron light impinges on the beam pipe (hard UV photons), which increases the residual gas pressure.

Let's mention that 10^{-7} of the stored beam intensity (or 300 ppm of a single bunch) is enough to quench a magnet and consequently abort the run. This shows how demanding the design of this new machine is.

2.1.2 The experiments

Four experiments have been approved to be installed at the intersecting points of the LHC:

- The ATLAS (A Toroidal Lhc ApparatuS, see Figure 2.3(right)) collaboration proposes to build a general-purpose *pp* detector designed to exploit the full discovery potential of the LHC [32]. In particular, the quest for the origin of the SSB mechanism in the electroweak sector of the SM and beyond (e.g. SUSY signals), as discussed in Chapter 1. The investigation of CP violation in B-decays will be part of the addressed topics as well. ATLAS is planned to run also in heavy ion mode, searching for some rare processes.

- The CMS (Compact Muon Solenoid, see Figure 2.3(left)) general-purpose detector has been designed to detect cleanly the various signatures for new

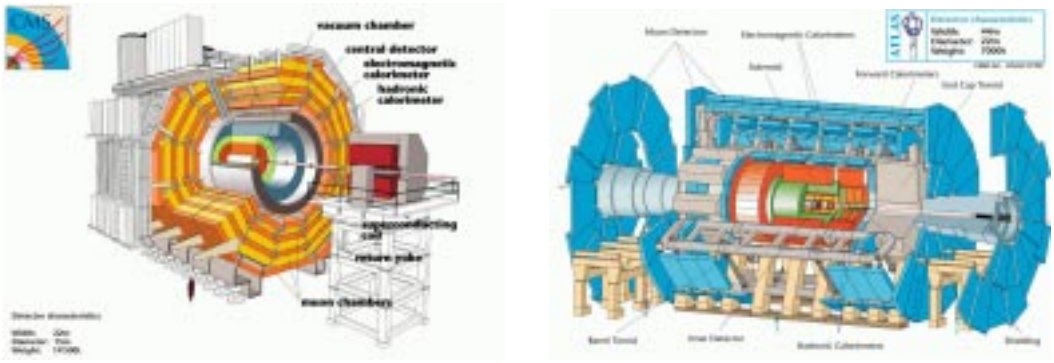


Figure 2.3: *Left:* Layout of the CMS detector with a cutaway view showing the outer four layers for detecting muons (interleaved with three layers of iron), the central calorimeters and the inner tracking system. *Right:* Equivalent view of the layout of the ATLAS detector. Magnetic fields will be both longitudinal and solenoidal.

physics by identifying and precisely measuring muons, electrons and photons over a large energy range and at high luminosity [33]. CMS could be competitive as well for B-physics in some particular channels to measure two of the angles of the unitarity triangle (see next section). The detector will also be used to detect low momentum muons produced in heavy ion collisions.

- The LHCb detector is designed to exploit the large number of b-hadrons produced at the LHC in order to make precision studies of CP asymmetries and of rare decays in the B-meson systems [3]. Its main goal is to measure all angles of the unitarity triangles in many different ways, over-constraining the CKM matrix including higher order terms, and thereby searching for an inconsistency in the CKM picture which would reveal new physics.

- ALICE (A Large Ion Collider Experiment) is the only detector fully dedicated to the physics of nuclear collisions [34]. It is designed to cover the full richness of hadronic and leptonic signals expected at the LHC allowing to establish and to study the phase transition from hadronic matter to deconfined partonic matter, the quark gluon plasma (QGP).

2.2 Physics of pp collisions, a summary

Essentially all physics aspects of the LHC, from SUSY particle searches to electroweak precision measurements and studies of heavy quarks, are connected to the

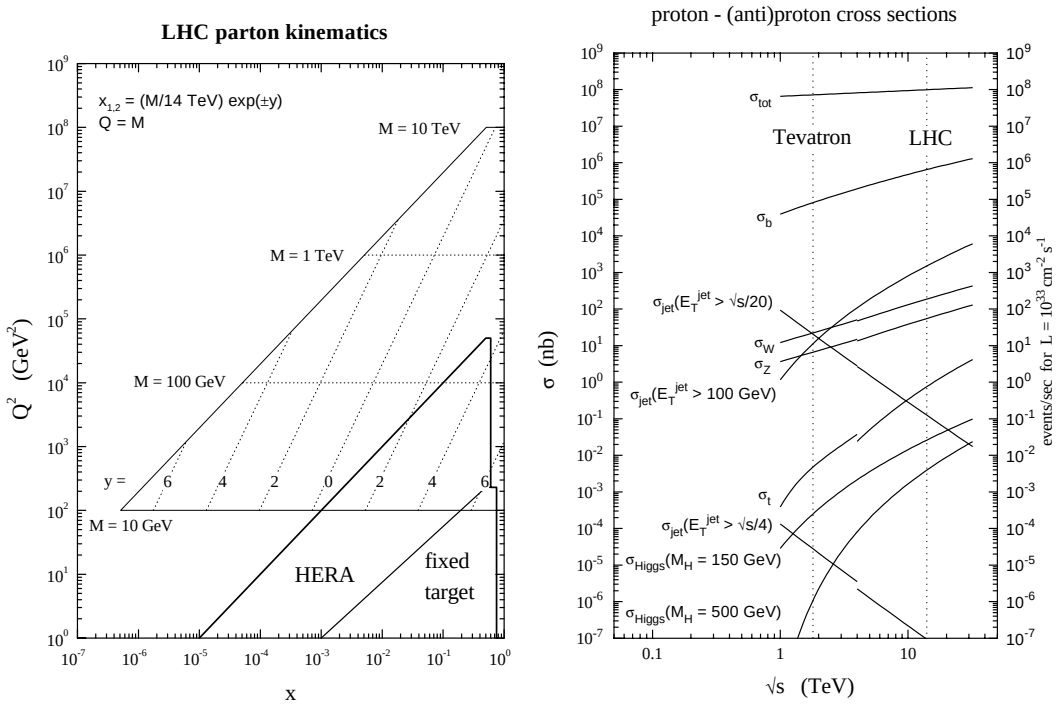


Figure 2.4: *Left:* Values of x and Q^2 probed in the production of an object of mass M and rapidity y at the LHC, $\sqrt{s} = 14$ TeV. *Right:* Cross sections for hard scattering versus $\sqrt{s}$, calculated using the latest MRST parton density functions [35] except $\sigma_{\rm tot}$. From [36].

interactions of quarks and gluons at large transferred momentum. A solid understanding of QCD and its modelling is therefore necessary. Figure 2.4 (left) shows the kinematical regions accessible to the partons in the inelastic pp collisions at LHC (together with a comparison with the accessible range at HERA)[36]. In the case of the production of a heavy particle of mass $M^2 = Q^2$ and rapidity y (in the laboratory frame), the dominant values of the momentum fractions carried by the partons are $x_{1,2} \sim \frac{M}{\sqrt{s}} e^{\pm y}$. Accordingly, the heavier the produced particle, the more central it is and the harder the collision ($x_{1,2} \sim 1$). As a consequence, detectors designed to scan for (new) physics up to the TeV scale, like ATLAS and CMS, have to cover the full solid angle[2]. One sees also that, as $\sqrt{s}$ increases, the parton density functions are probed in a kinematic range that extends towards larger values of Q and smaller values of $x_{1,2}$.

An important chapter of LHC physics will be the study of heavy quark sys-

[2]Although the equivalence between the rapidity y and the pseudorapidity $\eta \to y, E \gg M$ would not hold for heavy (but slow) objects, the central coverage is necessary to detect their decay products.

tems. In fact, already at initial lower luminosities the LHC will be a high-rate beauty and top quarks factory as shown in Figure 2.4 (right) with the energy dependence of cross sections (NLO) for a selection of hard processes and the event rates expected at the LHC. The curves for the lower values of $\sqrt{s}$ are for $p\bar{p}$ collisions, as at the Tevatron, while the curves for the higher values of $\sqrt{s}$ are for pp collisions, as at the LHC. The ratio $\sigma_{p\bar{p}}/\sigma_{pp}$ is expected to scale to 1 with increasing $\sqrt{s}$ (Pomeranchuk theorem, see Appendix A).

Cross sections for production of objects with a fixed mass or jets with a fixed transverse energy E_T rise with $\sqrt{s}$ because the required $x_{1,2}$ values decrease as discussed above, and there are more partons at smaller x (sea quarks and gluons) [37]. Conversely, cross sections for jets with transverse momentum which is a fixed fraction of $\sqrt{s}$ fall with $\sqrt{s}$, mostly because the partonic cross sections $\hat{\sigma}$ (of Drell-Yan type) fall with E_T like E_T^{-2}.

Large-p_T W, Z production will doubtlessly provide an interesting test of QCD, but it is also a major source of background for the potential Higgs signal because of their yield of hard and isolated leptons and dileptons, as will be shown in Chapter 4.

An estimate of the total cross section (based on an extrapolation of a standard Regge parameterization, see Appendix A) is displayed, out of which a reference value for the soft hadronic inelastic processes cross section can be extracted $\sigma_{\text{inelastic}} = \sigma_{\text{tot}} - \sigma_{\text{elastic}}$.

As discussed in Appendix A the properties of such minimum-bias[3] events are not calculable in standard perturbative QCD, thus one has to rely on various models. The understanding of the minimum-bias event together with the "underlying event" in a hard scattering process, is important as they can limit or spoil the lepton (or γ) isolation which is a crucial selection criterion in Higgs searches. The underlying event is part of the hard collision event itself and is thus unavoidable, whilst the minimum-bias event enters through event superpositions during the same bunch crossing when running at high luminosity.

Former abundant pp (ISR at $\sqrt{s} \sim 63$ GeV) and $p\bar{p}$ (Sp$\bar{\text{p}}$S at $\sqrt{s} = 630, 900$ GeV) data are not shown on the plot but are part of the same fitting procedure described in Appendix A.

2.3 The LHCb experiment

The LHCb detector is a forward one-arm spectrometer dedicated to the study of CP violation in the B-meson systems and indirect search for new physics [3].

The choice of the geometry is motivated by the fact that at high energies both the b and $\bar{b}$ quarks are predominantly produced in the same forward cone with a

[3]Referring to the terminology proposed in Section A.4.

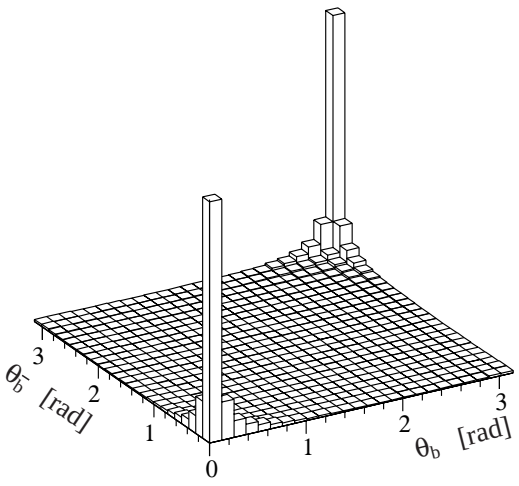


Figure 2.5: Polar angles of the b and $\bar{b}$ quarks at generator level according to equation 2.1.

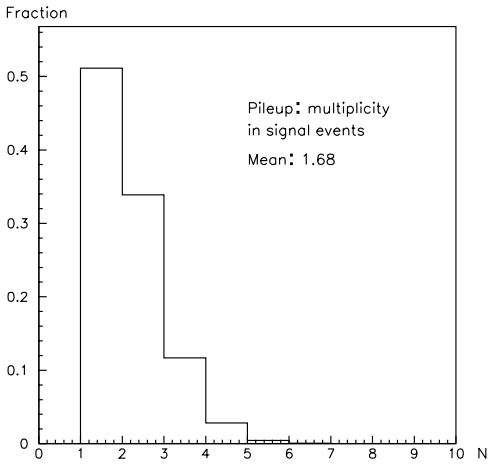


Figure 2.6: Pileup multiplicity distribution for Higgs events at LHCb with a luminosity 2×10^{32} cm^{-2}s^{-1}.

typical correlation of one unit in rapidity. At tree level [37],

$$q\bar{q} \to Q\bar{Q}\,(1\%)\ ,\ gg \to Q\bar{Q}\,(99\%)$$
$$\frac{d^4\sigma}{dy_Q dy_{\bar{Q}} d^2p_T} \propto \frac{1}{64\pi^2 m_T^4 (1+\cosh(\Delta y))^2} \tag{2.1}$$

The cross section is suppressed for values of p_T much greater than the quark mass (due to the term $m_T^2 = m^2 + p_T^2$), as well as for large rapidity differences Δy between the two heavy quarks. The corresponding correlation between the polar angles of the quark and antiquark is shown in Figure 2.5. The LHCb rapidity coverage is slightly larger than 3 units, ranging from $\eta = 1.8$ to $\eta = 4.9$. From Figure 2.4(left), one can notice that light objects up to $M \sim \mathcal{O}(100$ GeV) are observable with significant efficiencies.
Appendix B gives an overview of the main physics program of LHCb.

2.3.1 Apparatus

The overall length of the LHCb apparatus, whose components are shortly described hereafter, is determined by the size of the experimental hall on the LHC ring assigned to LHCb (used previously at LEP for the DELPHI detector). In this hall, the interaction point will have to be displaced by ~ 11 m with respect to the center, and extra room needed for the vertex detector will be excavated around the pipe tunnel. The layout of the detector is shown in Figure 2.7.

At the LHCb interaction point, the beam will be defocused to produce an average luminosity of 2×10^{32} cm^{-2}s^{-1}. This luminosity is chosen to optimize the

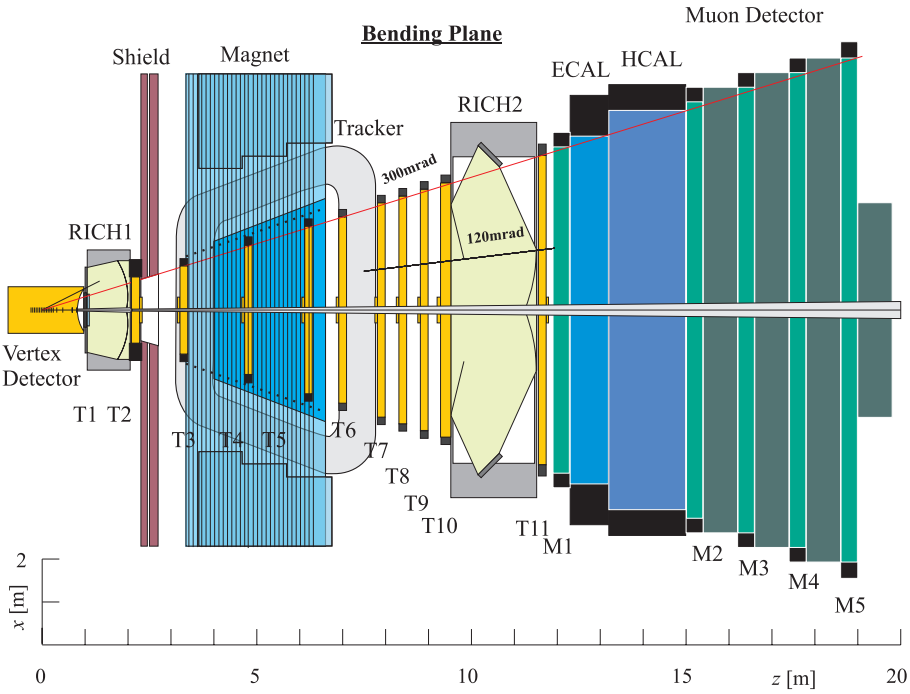


Figure 2.7: Layout of the LHCb detector in the bending (horizontal) plane.

fraction of single pp interactions per bunch crossing which are cleaner and easier to reconstruct; it also minimizes both the radiation damage and the detector occupancy. Predictions for the pileup multiplicity for (Higgs) signal events with a running luminosity 2×10^{32} cm^{-2}s^{-1}are shown in Figure 2.6, according to [38],[39].

2.3.2 The VErtex LOcator (VELO) [1]

The VELO has been designed for a precision measurement of the decay length of the b-hadrons. An excellent proper time resolution is indeed necessary to resolve the fast $B_s^0 - \bar{B}_s^0$ oscillations, as needed for the observation of time-dependent CP asymmetries in B_s^0 decays. The VELO provides a lifetime resolution of up to $\sim$40 fs depending on the decay channel.

The VELO consists in 25 Si stations (disc shaped, see Figure 2.8) distributed over 1 m along the beam axis, around the interaction point. They will be mounted perpendicular to that axis on Roman pots inside a vacuum tank and will be retracted from the beams during injection.

The 300 μm thick single-sided silicon detectors will have r or ϕ strips with an active radius from 8 to 42 mm, and analogue readout. This device will provide a resolution $\sigma_L = 42$ μm on the interaction point along the beam axis ($\sigma_T = 10$ μm

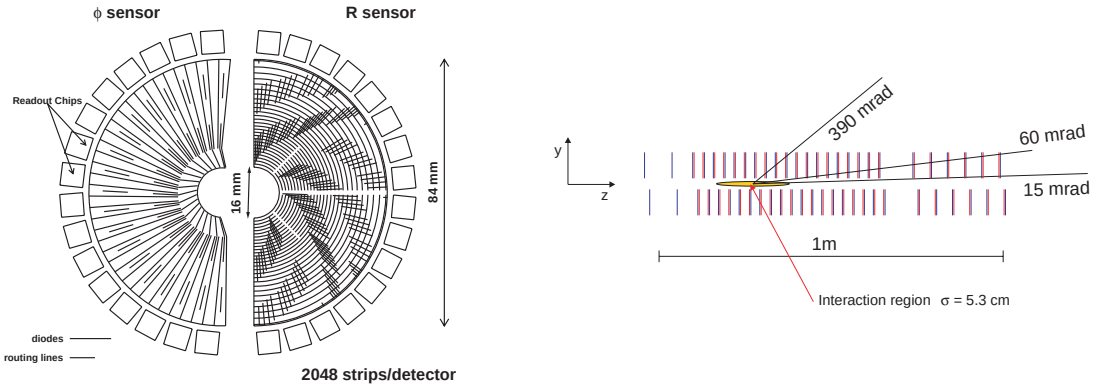


Figure 2.8: Layout of the vertex detector given in the technical design report [1]. The detector has 25 stations, each consisting of 2 silicon discs (with r and ϕ strips), divided into 2 sectors.

transverse) and a typical hit resolution between $\sim 3.6\,\mu$m for 100 mrad tracks and 40 μm pitch.

2.3.3 The magnet [2]

A warm dipole magnet has been designed with a bending power of 4 Tm. The field is oriented vertically and has a maximum value of 1.1 T. The warm option will ease the regular reversal of the magnetic field, needed to reduce a possible systematic bias in the CP asymmetries due to non-uniformities in the detection efficiency. An iron shield upstream of the magnet reduces the stray field in the vicinity of the vertex detector and of RICH1 (see section 2.3.5). Charged particles passing through the magnet will receive a p_T kick of ~ 1 GeV/c.

2.3.4 Tracking [3],[4]

The principal task of the tracking system is to provide efficient reconstruction of charged-particle tracks and precise measurements of their momenta. It has also to provide measurements of track directions in both x and y projections for the reconstruction of Cherenkov rings in the RICH detectors. Finally, tracking hits are used in the Level-2 trigger. The main tracking system comprises ten tracking stations, located between the vertex detector vessel and the calorimeters. These are shown as T1–T10 in Figure 2.7, with T3–T5 mounted inside the magnet to allow for a continuous tracking. Tracking station T11 has been removed since the original design. The geometry of the detectors and the choice of technologies are driven by occupancy considerations. Because of the high particle density close to the beam axis, the tracking system is split into outer and inner subsystems at a

radius of approximately 0.5 m. Most of the detector acceptance is subject to a particle density which is low enough to use drift chambers with small cells: the *outer* tracking system consists of straw-tube-like chambers.

On the other hand, the *inner* tracking system covers in most stations an area of 60×40 cm^2 around the beam pipe, where a much finer granularity is required. For the inner region several options have been tested (MSGC, Micromegas, GEM+Micromegas) and discarded for being particularly sensitive to discharges in the presence of highly ionizing particles [40]. The last competitive technology is triple-GEM.

Despite of a higher cost per channel, silicon detectors are envisaged as a fall-back solution. With radiation hardness treatment (by oxygenation) silicon is expected to survive around 5 years. In the present design, four planes of single sided 300 μm thick strips will be used in each station.

The tracking system provides precise particle trajectories in the bending plane of the spectrometer using near-vertical wires and strips. Three dimensional track reconstruction is achieved using small stereo angles of $\pm5^\circ$ with respect to the vertical.

The track reconstruction program is based on the technique of Kalman filter. The reconstruction is done in the upstream direction.

The momentum resolution is limited by multiple scattering, with a relative precision $\sigma_p/p = 0.3\%$ approximately constant for tracks momenta from 5 GeV to 200 GeV.

2.3.5 The RICH system [5]

Two Ring Imaging Cherenkov detectors (RICH) have been designed to identify charged hadrons over a large momentum range. The first one, located just in front of the magnet, will cover polar angles up to 330 mrad and contains two radiators: C_4F_{10} gas and aerogel. The latter has a high refractive index ($n = 1.03$), ensuring a good K-π separation down to 1 GeV/c, and thereby efficient flavour tagging with charged kaons from $b \to c \to s$ decays. High momentum tracks, which tend to be emitted at low angles, will be identified in front of the calorimeter by a second RICH detector with CF_4 gas. This will extend the K-π separation capability up to 100 GeV/c, fairly close to the high end of the momentum spectrum of tracks from two-body decays like $B^0 \to \pi^+\pi^-$.

2.3.6 Calorimetry [6]

The LHCb calorimeter system consists of three components: a preshower detector, electromagnetic and hadronic calorimeters. The transverse segmentation is shown on Figure 2.9 and parameters of the whole assembly are summarized in Table 2.1.

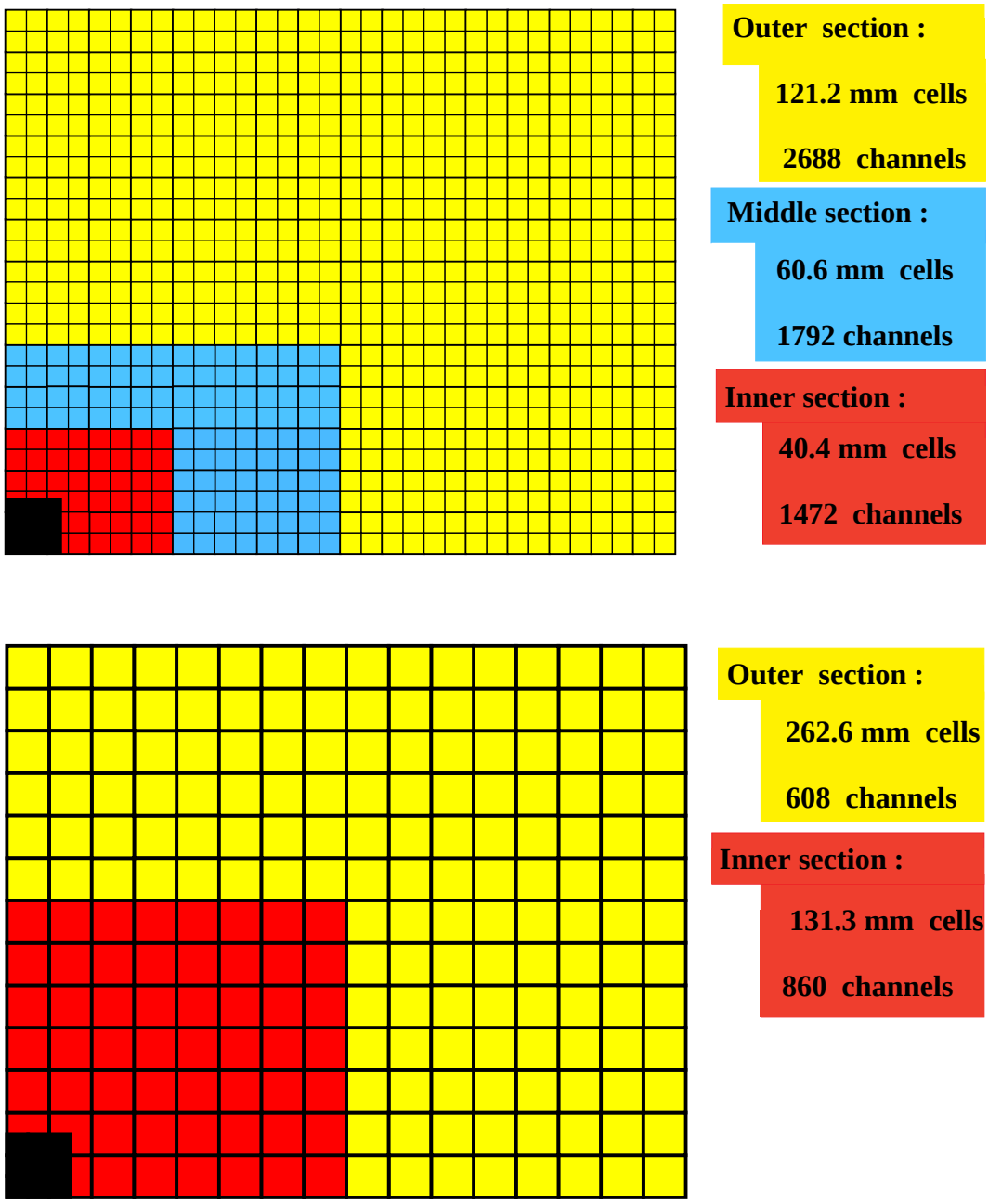


Figure 2.9: *Top:* Lateral segmentation of the SPD/PS and ECAL. One quarter of the detector front face is shown. The cell dimensions are given for ECAL. The size of the SPD/PS is reduced by $\sim 1.5\%$ for SPD/PS. *Bottom:* Lateral segmentation of HCAL. One quarter of the detector front face is shown. The black squares are the central hole crossed by the beam pipe: 64×64 cm^2 for ECAL, 78×78 cm^2 for HCAL (i.e. $\sim$30 mrad in both direction). The grid pitch corresponds to the cell size in the outer section. See text and Table 2.1 for further description. From [6].

Preshower

The preshower (PS) detector plays an important rôle in electron and photon identification, and in pion rejection. Immediately upstream of the PS is the scintillator pad detector (SPD; at present the first muon chamber), providing position measurement for electron/photon discrimination, with a pad size of 2×4 cm^2, about half of the preshower detector cell size. The preshower detector matches the granularity of the electromagnetic calorimeter; its readout is achieved with fast photomultipliers (PMs).

Electromagnetic calorimeter

The electromagnetic calorimeter (ECAL) plays a major rôle in electron and photon identification at the first LHCb trigger level, and is designed in a Shashlik-type sampling structure. Lead and scintillators are coupled to wavelength shifting fibers (WLS) which are read out by fast PMs. There are 70 longitudinal layers giving a total depth of $25X_0$ (1.1 λ_I), with inner, middle and outer transverse regions with granularities chosen to keep particle occupancy everywhere below 5–10%. The inner section is composed of 2688 4×4 cm^2 cells, the middle section has 1792 6×6 cm^2 cells, and the outer section has 1472 12×12 cm^2 cells, with a Moliere radius equal to 3.6 cm. The geometrical acceptance of the ECAL is 300×250 mrad^2 with an inner beam hole of 30×30 mrad^2. The dynamic range is adapted to the different kinematical zones to cover a constant E_T scale (0–10 GeV). The calorimeter signals are digitized by 12 bit ADCs with a sampling time of 25 ns. The expected resolution is $\frac{\sigma_E}{E} = \frac{10\%}{\sqrt{E}} \oplus 1.5\%$ (E in GeV).

Hadronic calorimeter

The main aim of the hadron calorimeter (HCAL) is to provide cluster energy and position measurement of hadrons, for use in the Level-0 hadron E_T trigger (see section 2.3.8). It will also be used offline in muon identification and pile-up rejection. The HCAL has a scintillator/iron (4 mm/16 mm) sampling structure, with scintillating tiles parallel to the beam axis giving a total depth of 5.6 interaction lengths. The number of cells in the inner and outer sections are 860 and 608, with cell sizes of 13×13 and 26×26 cm^2, respectively. The scintillators are coupled to WLS fibers read out with fast PMs. The dynamic range of the readout allows an energy measurement between 100 MeV and 300 GeV. The expected resolution is $\frac{\sigma_E}{E} = \frac{80\%}{\sqrt{E}} \oplus 5\%$ (E in GeV).

The main characteristics of the calorimeter system are summarized in Table 2.1 and detailed information concerning the readout and the calibration scheme can be found in the technical design report [6]. Energy resolutions are rather modest as these calorimeters are basically designed to be used for triggering purposes. Part of the work done for this thesis shows that the LHCb calorimetry

sub-detector	SPD/PS	ECAL	HCAL
number of channels	2×5952	5952	1468
overall lateral dimensions in x, y	6.2 m×7.6 m	6.3 m×7.8 m	6.8 m×8.4 m
depth in z	180 mm, 2 X_0, 0.1 λ_I	835 mm, 25 X_0, 1.1 λ_I	1655 mm, 5.6 λ_I
basic performance (E in GeV)	20–30 photoelectrons per MIP	$\sigma(E)/E =$ $10\%/\sqrt{E} \oplus 1.5\%$	$\sigma(E)/E =$ $80\%/\sqrt{E} \oplus 10\%$
dynamic range	0–100 MIPs 10/1 bits (PS/SPD)	0–10 GeV E_T 12 bits	0–10 GeV E_T 12 bits

Table 2.1: Requirements to the calorimeter sub-detectors.

is performant enough to handle with jet physics. As an example, Figure 2.10 shows a QCD generic $b\bar{b}$ event with an invariant mass $M(J_b, J_{\bar{b}}) = 52$ GeV/c^2 with clear two jets clearly separated in the innermost part of the electromagnetic calorimeter.

2.3.7 The muon system [7]

The muon detector provides muon identification and Level-0 trigger information. It consists of four stations M2–M5 on Figure 2.7 embedded in an iron filter and a special station M1 in front of the electromagnetic calorimeter. All stations have pad readout to achieve fast trigger response. Multigap Resistive Plate Chambers (MRPCs) are proposed for most of the coverage of M2–M5, where particle fluxes are below 5×10^3 cm^{-2}s^{-1}. Station M1 and the inner regions of stations M2–M5 experience the highest fluxes and are therefore constructed from Cathode Pad Chambers (CPCs).

2.3.8 The trigger system [3],[8]

The 25 ns bunch spacing at LHC implies that the pipeline of the front-end electronics will have to be clocked at 40 MHz. However, some of the bunch crossings in LHCb will involve empty bunches, reducing the bunch crossing rate to an effective average value of 30 MHz.

As mentioned earlier, LHCb is planning to use mostly single proton-proton interactions. The detector will therefore be operated at a modest 2×10^{32} cm^{-2}s^{-1}luminosity, and the beams will be defocused at the LHCb interaction point as the LHC machine gradually delivers its design luminosity of 10^{34} cm^{-2}s^{-1} to ATLAS and CMS. Under these conditions and assuming an inelastic cross section of 80 mb, 30% (10%) of the bunch crossings will have one (more than one) pp

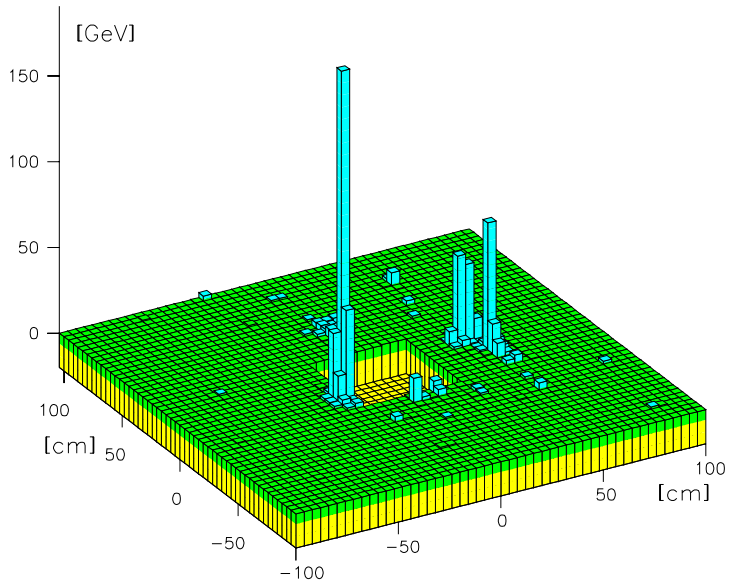


Figure 2.10: Generic $b\bar{b}$ event as seen in the innermost section of the ECAL in LHCb. The invariant mass of the pair $M(J_b, J_{\bar{b}}) = 52$ GeV/c^2 generates a clear di-jet signature.

interaction. In addition, a pile-up veto system is foreseen at the earliest trigger level to reject events with more than one interaction vertex, from the information provided by two dedicated silicon stations located upstream of the VELO. After this cut, the event rate is around 10 MHz with more than 90% single interactions. For comparison, a typical bunch crossing at ATLAS and CMS will contain on average 20 pp interactions.

The trigger scheme will be implemented in four levels to selectively extract the B decays of interest while rejecting the non-b events. The main parameters and the architecture of the trigger system are summarised in Table 2.2 and in Figure 2.11.

Level-0

This first level of trigger, based on calorimeter and muon chamber information only, will reduce the event rate to 1 MHz by requiring a muon, an electron or a hadron with a transverse momentum p_T or energy E_T above some threshold, typically, 1, 2.1 and 2.4 GeV, respectively: tracks from b-hadron decays have indeed a harder p_T spectrum than tracks from non-b events. The fractions of the bandwidth attributed to the muon, electron and hadron triggers have been chosen to maximize the overall CP reach and will approximately be 20%, 10% and 60%, respectively, the rest being allocated to other triggers such as high E_T

photon, dimuon, etc. As illustrated in Figure 2.11, the Level-0 decision will take 4 μs during which the data will be kept in the pipeline of the front-end electronics.

Level-1

The Level-1 trigger achieves a further reduction in rate by a factor 25. It uses vertex detector information to identify secondary vertices produced by the $b-$hadron decays. Tracks are found from the hits in the VELO and a primary vertex is reconstructed; secondary vertices are then formed with large impact parameter tracks. The Level-1 buffer will reside on off-detector electronics; its depth will determine the maximum latency, which will be as large as 1024 or 2048 μs.

Levels 2-3

Some secondary vertices found at Level-1 are fake, due to low momentum tracks undergoing multiple scattering. The aim of the Level-2 trigger is to reconstruct large impact parameter tracks in the VELO *and* the tracking chambers and use the momentum information to refine the secondary vertex requirement. At Level-3, specific b-hadron decay modes will be reconstructed and selected with loose cuts, using *all* available information.

2.3.9 Front-end electronics [3]

Figure 2.11 shows the general architecture of the trigger and data acquisition system, including the main data and control flows involved in transporting the subdetector data from the front-end electronics to the event storage. The exact operating point of the trigger scheme can be adjusted to running conditions without significant loss in physics. Approximately 200 events/s will be written to tape at a rate of 20 Mbytes/s.

The subdetectors will use a similar architecture for the front-end electronics, which has to accommodate the specific trigger requirements of LHCb, making maximum use of existing components. All analogue and digital signals arriving at 40 MHz will be stored in Level-0 pipelined buffers, 128 cell deep, to await the Level-0 trigger decision taken after a fixed delay of 4 μs. Events accepted at an average rate of 1 MHz are transmitted to short derandomizing buffers to avoid overflow due to limited output speed. The data are then multiplexed and digitized, if they were still analogue, and sent to Level-1 buffers, at least 1024 events deep, to allow $\gtrsim 1$ ms for the next trigger selection. The average rate of events accepted by Level-1 is 40 kHz. Accepted events pass zero suppression and data formatting, are multiplexed and sent via the "front-end links" to the data acquisition system, located approximately 60 m from the detector.

The front-end electronics located inside the detector must be radiation hard or tolerant, the dose integrated over 10 years amounting to 0.2 MRad at 30 cm.

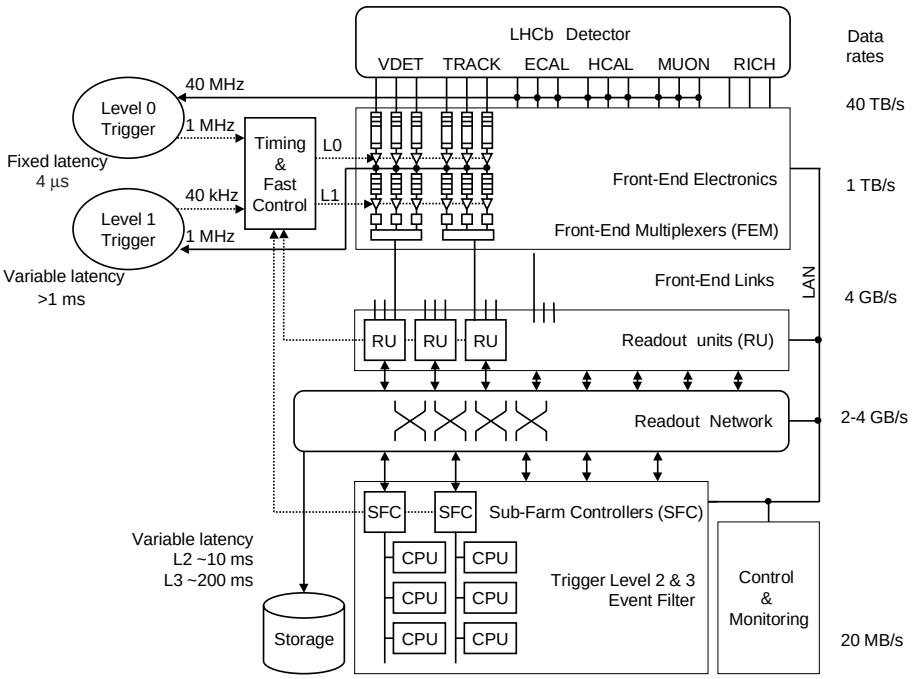


Figure 2.11: General architecture of the trigger and DAQ system in LHCb [3].

Parameter	Value
Readout channels	950,000
Average event size	100 kB
Average Level-0 accept rate	1 MHz
Level-0 latency (fixed)	4 μs
Average Level-1 accept rate	40 kHz
Level-1 latency (variable)	>1 ms
Front-end Links	165
Readout Units (RUs)	$\sim$120
Event Building bandwidth	2-4 GB/s
Sub-farm Controllers (SFCs)	$\sim$120
Average Level-2 processing time (on a 1000 MIPS CPU)	10 ms
Average Level-2 accept rate	5 kHz
Average Level-3 processing time (on a 1000 MIPS CPU)	200 ms
Average Level-3 accept rate	200 Hz
Data storage rate	20 MB/s

Table 2.2: Trigger/DAQ parameters in LHCb [3],[8].

Part of the electronics, probably from the Level-1 buffer onwards, will be installed at least a few metres away from the beam to permit standard components to be used. The possibility is envisaged to transport the VELO and inner tracker L1 electronics to the counting barracks.

2.4 Light Higgs in LHCb

The issue of the next chapters is the investigation of the LHCb potential to observe a Higgs particle signal in the context of the most recent knowledge about its likely mass range, given the particular geometry and design of the experiment. The first two chapters outline that the dominant $b\bar{b}$ decay channel for light mass Higgs should be accessible to LHCb. It will be shown that this specific channel — though not "gold-plated" — turns out to be the LHCb's "warhorse" for Higgs hunting in the very confuse hadronic environment reigning over the 100 GeV mass scale.

Key features which have to be addressed are the suitability of the (modest) calorimetry to reconstruct jets with an acceptable resolution, the flexibility of the trigger scheme to accept Higgs events with very high efficiency and the ability of the tracking/vertexing to tag b-jets.

Chapter 3

Jet physics in LHCb

3.1 What is a jet ?

We have seen in Appendix A that high energy hadronic interactions are dominated by events in which secondaries have small p_T and are mostly confined to the nearly forward direction as can be inferred from cross section 2.1. This led to the picture that strong interactions at high energy are generally rather "soft". To this extend LHC can be seen as a *gluon* collider.

On the other hand, large p_T events proceed via "hard scattering" involving the collision of just one constituent from each initial particle. The partons involved are scattered at large p_T and are supposed to materialize as a set of fairly well collimated hadrons called a "jet" — a shower of hadrons.

An idealized description of such a situation is given in Figure 3.1.

Jets for a given initial parton can vary widely in shape, particle content, and energy spectrum. There is, of course, also substantial smearing due to instrumental effects: the finite resolution and granularity of detectors (calorimeter cells and muon measurements), and escaping neutrinos.

3.1.1 Historical considerations

The first evidence for jets was in e^+e^- collisions, producing secondary hadrons (at SLAC and DESY, see next section); subsequently, they were also observed in hadron collisions (e.g. UA experiments and ISR at CERN). Frequently, two or three main jets are observed which dominate the energy balance of the collision.

In hadronic collider events, however, the balance is observed only in the transverse plane due to the difficulty of observing at large (absolute) rapidity, and due to the structure function, which leaves the hard quark undergo a longitudinal boost. The subsequent undetermination of the center-of-mass (CM) of the hard process imposes a definition of the reference energy or scale event per event.

Jet definition is not without ambiguity, hadrons in the final state not being rigorously associated to the partons (which in turn also undergo some final-state

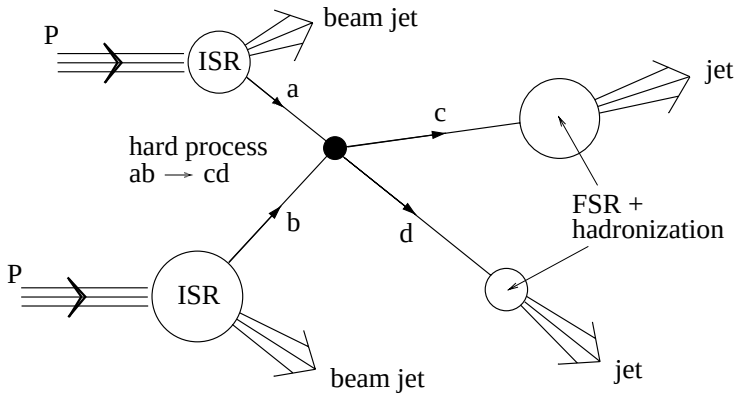


Figure 3.1: Schematic description of a large p_T hadronic interaction $ab \to cd$ producing four jets. Initial and final state radiations (respectively ISR, FSR) represent the parton showering process.

interactions before they hadronize). Many methods can be used to define what is meant by a jet. There is no best definition, but one must be sure that both theoretical and experimental analyses use the same definition, typically in the determination of the jet cross sections $d^2\sigma_{nJ}/dE_T d\eta, n = 1, 2, ...$ This implies for instance the jets to be covariant objects.

In hadron-hadron collisions, jets are particularly difficult to separate: there is an underlying event due to the remnants from the spectator quarks not participating in the hard interaction[1], and at high luminosity, there may be pileup of multiple collisions which cannot be separated in time.

In recent jet analyses, former topology-based jet variables used in symmetric electron colliders have been somewhat abandoned (see next section); jets are analysed in cones defined by a cutoff in angular radius $R(\Delta\eta, \Delta\phi)$ around an axis defined in various ways. Jets found in this way are often subsequently combined into fewer jets, using some clustering algorithm.

The jet structure is studied as a function of the jet radius; particle multiplicities, rapidity distributions, fragmentation functions, and others (and again their gradients, when varying the cone radius) are compared to the values obtained from phenomenological Monte Carlo programs.

A lot of effort has been devoted to setting up a standard jet definition and an agreement on a cone definition was reached at the 1990 Snowmass Workshop, known as the "Snowmass Accord" [41].

[1] We are back to the considerations of Appendix A, about soft interactions which cannot be treated in pQCD.

3.1.2 Jet characterization

Central detectors at symmetric e^+e^- colliders are ideal to characterize high p_T manifestations that jets are meant to be, as taught by history (see data from SPEAR [42], PETRA [43], [44],[45] and, more recently, from LEP). In the framework of QCD studies, the description of mostly two- or three-jet situations was done with scalar variables[2] such as planarity, sphericity or thrust [46] reflecting the topology of the event.

On the contrary, collisions involving hadrons let's presume of a much less easier situation to define jets, and their associated variables will be different.

For e^+e^- annihilation, one wants to emphasize rotational invariance. Thus the natural variables are energies E and spherical angles θ, ϕ and associated functions of these variables. For hadron-hadron collisions, one wants to emphasize invariance under boosts along the beam axis since, the CM frame of the hard scattering is typically moving in the hadron-hadron CM frame. Thus the natural variables are transverse momenta p_T or the corresponding "transverse energy" $E_T \equiv E \sin\theta$, azimuthal angle ϕ, and pseudorapidity η.

In e^+e^- collisions one normally uses a jet definition which associates every final-state hadron uniquely with one of the jets. In hadron-hadron collisions producing high p_T jets, only a small fraction of the final state hadrons are associated with the high p_T jets. The other particles present in the event are associated with the beam jet (the soft underlying event plus the initial state radiation). One wants to keep the high p_T jets distinct from the hadronic debris in the beam jets.

For this reason, one has typically used a cone definition [47], which was actually inspired by the original theoretical definition for jets in e^+e^- collisions [48]. A jet in this definition is a set (or cluster) of particles whose momentum vectors lie within a certain angular cone of longitudinal invariant size R (typically between 0.5 and 1.0)

$$\Delta R = \sqrt{(\Delta y)^2 + (\Delta\phi)^2} < R \tag{3.1}$$

In practice, the rapidity $y = \frac{1}{2}\log\frac{E+p_L}{E-p_L}$ is normally replaced by the pseudorapidity η

$$\eta = -\ln\tan(\frac{\theta}{2})$$

which coincides with the rapidity in the $m \to 0$ limit and is a more convenient variable experimentally, since the angle θ from the beam direction is measured directly in the detector.

Such a definition suppresses the effect of the beam jets, since only a small fraction of the low p_T particles in the beam jets will fall into the cone of a high p_T jet.

[2] Mostly based on a principal component analysis of the activity in the event.

There will also be an essentially kinematical correlation induced by the fact that the jet selection or trigger process will generally be biased to choose events where the beam jets have higher than average global E_T and multiplicity (i.e. the underlying event is noisier than average). This emphasizes the necessity for an efficient jet characterization: according to its substructure and specific configuration in the event, a jet is being associated with a probability to come from a gluon, a light-, or heavy-flavour quark and will be "tagged" as such.

3.2 Jets in the forward region

Remark about conventions: In what follows, Higgs events used to illustrate the discussion are generated, unless mentioned, with $m_H = 115$ GeV/c^2. Furthermore, the term "LHCb acceptance" refers to a pseudorapidity coverage $1.8 < \eta < 4.9$ (or, 15 mrad $< \theta <$ 330 mrad).

As mentioned in 3.1, the concern in hadronic collisions is the separation of jets originating from the hard scattering subsystem and the underlying event, left by the beam jets. The question is ever more relevant in the forward region where the beam jets are expected to be more "invasive".

Figure 3.2 shows the average event multiplicity (a,c) and transverse energy (b,d) profiles for a typical hard scattering process such as $gg \to H \to b\bar{b}$. The multiplicity and energy plotted refer to final state particles with an energy threshold $E > 1$ GeV. The profiles are plotted for both b and $\bar{b}$ quarks emitted in the central region $|\eta| < 2.5$ (top plots) and in the forward region $1.8 < \eta < 4.9$ (bottom plots). No acceptance cut on final state particles has been performed and profiles are integrated over the azimuthal angle.

The profiles of the underlying event debris look similar in both kinematical regions. However, the integrated multiplicity of Higgs products (hatched) is significantly more important for b's in the forward configuration ($\sim$ 50 particles/event with $E > 1$ GeV) than in the central one ($\sim$ 33 particles/event with $E > 1$ GeV) due to the forward boost. On the other hand, integrated E_T profiles are left untouched by the longitudinal boost and are naturally identical ($\langle E_T \rangle \sim 100$ GeV per event).

The relative importance of the Higgs products (average E_T/particle) over the underlying debris is similar in both configurations, leaving hope to observe and separate jets at LHC even in the forward region[3]. Illustrations of the former statement are Figure 2.10 showing a QCD generic $b\bar{b}$ event forming an invariant mass of $M_{jj} \sim 50$ GeV/c^2, and Figure 3.12 showing a $HW \to b\bar{b}\ell\nu_\ell$ event, both

[3]The meaning of jet separation becomes clearer taking into account that quantities shown here are integrated in azimuth. Assuming that the beam jets are outside the acceptance and that only soft gluon radiation contribute to the underlying event, the debris tend to be uniformly distributed in ϕ whereas decay products of the Higgs are clearly not.

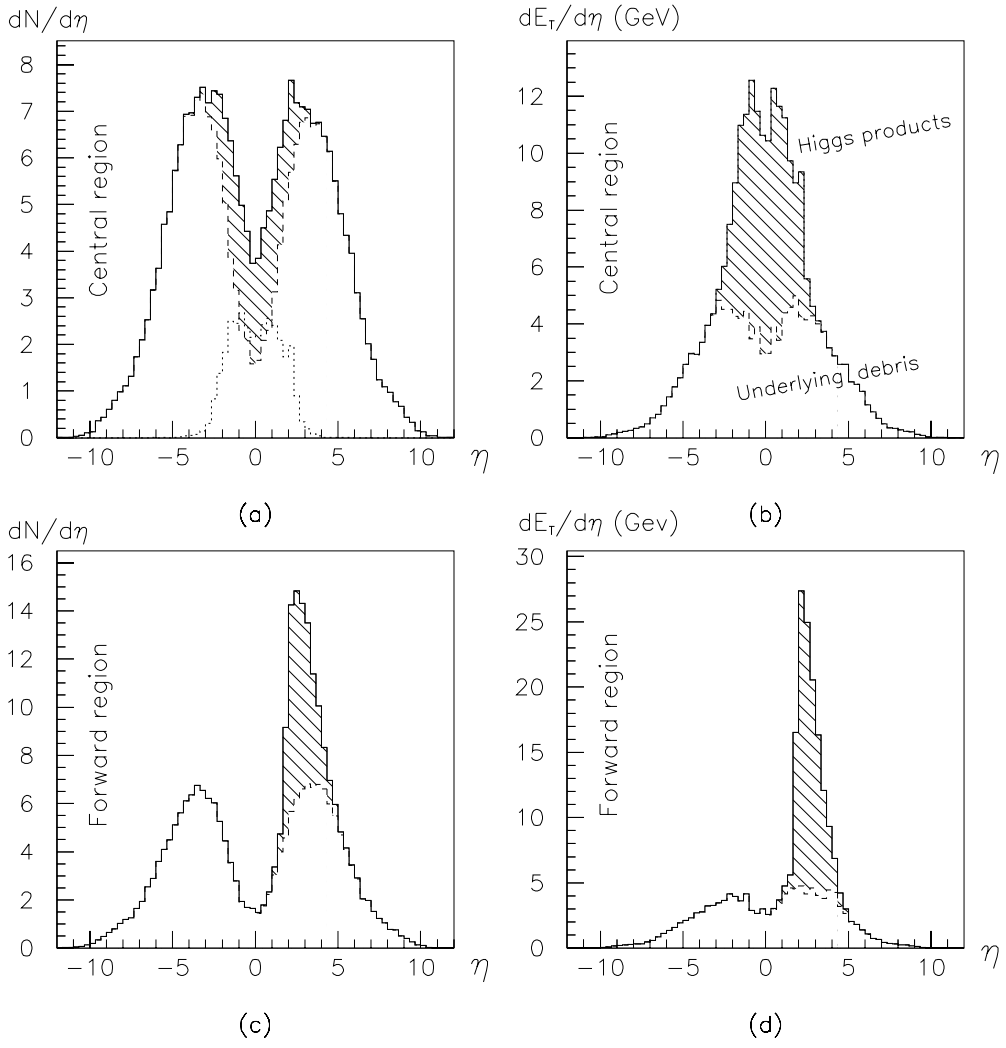


Figure 3.2: Average $H \to b\bar{b}$ event multiplicity (a,c) and transverse energy (b,d) profiles in pseudorapidity of Higgs fragments (hatched) and underlying event debris (plain white) for $b\bar{b}$ pair in central (a,b: $|\eta_{b,\bar{b}}| < 2.5$) and forward (c,d: $1.8 < \eta_{b,\bar{b}} < 4.9$) regions. In (a), the profile for Higgs fragments alone is superimposed.

with clear collimated jet structures.

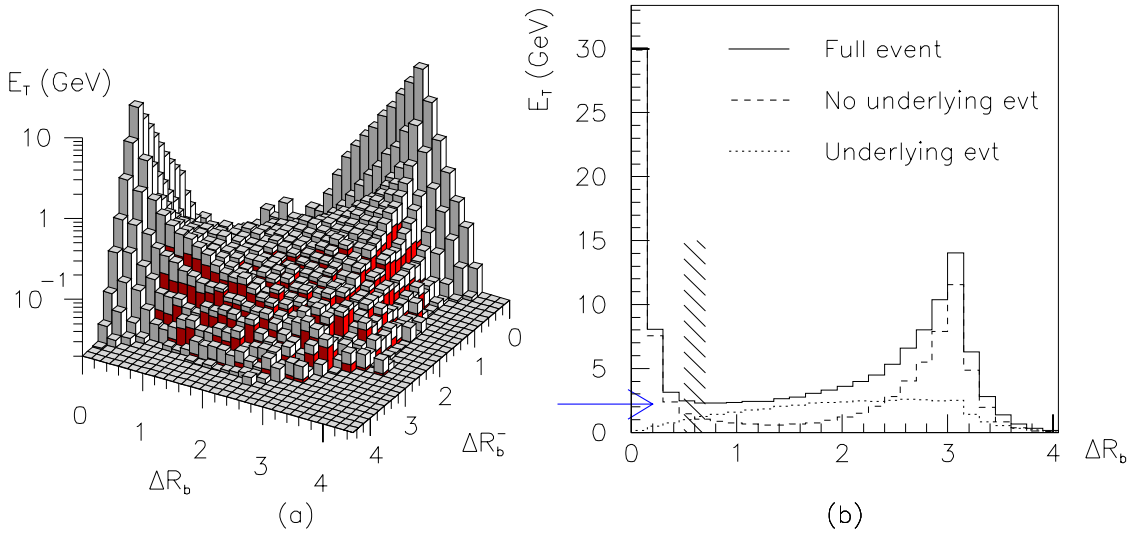


Figure 3.3: *Left:* Transverse energy profile as a function of the distances from b and $\bar{b}$ of final state particles originating from the Higgs (light) and the underlying event (dark). *Right:* Projection of the left plot. A typical radius for a jet reconstruction algorithm (shaded) would be around 0.6. The arrow shows the transverse energy pedestal formed by the underlying debris.

More specifically, Figure 3.3 shows the average transverse energy profile of final state particles in $gg \to H \to b\bar{b}$ events as a function of the distances ΔR_b and $\Delta R_{\bar{b}}$ from the two quarks of the Higgs decay, with ΔR defined in Equation 3.1. In Figure 3.3(a), the Higgs decay product distribution is plotted on top of the underlying event debris (dark).

One sees a clear indication for two jets which are well balanced in the transverse plane ($\Delta R \sim \Delta\phi \sim \pi$ in Figure 3.3(b)) on top of a roughly uniform distribution of the underlying debris. The projection 3.3(b) of the symmetric bidimensional distribution in (a) shows the relevant distance within which collection of Higgs products should proceed. We deduce from this plot that an efficient b-jet reconstruction should be limited to the $R < 0.5-0.7$ region. Beyond 0.7, the "transverse energy pedestal" [49] (indicated by the arrow) begins to deteriorate the resolution.

The main issue addressed in this section resides is the separation of the different contributions to the jet (or more precisely *di*-jet) mass resolution. After a brief introduction to the simulation framework that has been setup to study the jet reconstruction in LHCb and technical aspects of jet reconstruction algorithms,

Section 3.4.1 will more quantitatively discuss the jet reconstruction optimization and performance of such algorithms as a function of ΔR.

3.3 The LHCb simulation package

The LHCb detector simulation SICB[4] [50] is a GEANT3-based [51] implementation of the detector as described in Chapter 2. The event generation is handled by Pythia [52],[53] relayed by the QQ package. QQ was developped by the CLEO collaboration, for a deeper description of the particle decays, especially for b-hadrons (true matrix element approach with polarizations).

For the needs of the present work a Fast SIMulation (FSIM) has been developed by the author for the study of background channels to Higgs signal as will be discussed in the next chapter. Most of these channels have production cross sections orders of magnitude larger than the Higgs signal. The generation of the required amount of background events is prohibitive at the level of the full simulation. The FSIM skips the simulation step in GEANT (propagation of particles through matter). The relevant detector effects are then implemented by smearing in the four-vector level Pythia output the key variables with the ad hoc distribution.

The processing time for hard scattering processes such as Higgs production is typically $\mathcal{O}(10^{-1})$ s/event in FSIM compared to $\mathcal{O}(4)$ min/event using the full detector simulation (SICB), hence an overall gain of ~ 2500.

Remark about conventions: In what follows, the full detector simulation will simply be referred to as "SICB" whereas four-vector level studies will be referred to as "FSIM".

Fast simulation (FSIM)

Table 3.1 lists the various effects which have been implemented in the FSIM and are described hereafter.

By default, the fast simulation reproduces the LHCb geometrical acceptance by selecting only particles which are inside a forward conical ring whose inner and outer radii are given by the angular coverage of the detector in the non-bending plane: $\theta_{\min} = 15$ mrad and $\theta_{\max} = 300$ mrad. For studies involving the magnet such as the ones performed in [54], the volume is defined by a prism given by a virtual calorimeter front face at $z \sim 13$ m with the conventional lateral dimensions given in Section 2.3.6.

The Level-0 (L0) trigger conditions described in [3] are applied at four-vector level. This implementation certainly has to be seen as crude because of the absence of detector effects. However, hard-p_T events such as $H(W,Z) \to b\bar{b}(\ell\nu_\ell)$ and their backgrounds will fire the LHCb L0 with a probability close to unity

[4] SImulation CoBex, an ancestor of LHCb.

FSIM accounts for	Yes	No	Comment
geometrical acceptance	✘		cylindrical
L0 trigger	✘		at single particle level
pileup		✘	not presently, but feasible
primary vertex resolution	✘		$f(N_{\text{tracks}})$
first hit coordinates	✘		$f(p)$ along z
magnet p_T-kick	✘		no curling tracks
momentum resolution	✘		charged tracks
shower shape		✘	no clusters consequently
energy resolution	✘		calorimeter resolution
segmentation	✘		not explicit
misidentification		✘	no, mostly difficult
neutrinos	✘		

Table 3.1: Summary of the implementation of the FSIM. The symbol $f(x)$ means "as a function of x".

as is discussed in the next chapter. More detailed studies are needed, however, when describing the rejection of minimum bias events achieved by L0 *lepton* and L0 *hadron* triggers. The pileup (multiple interactions per bunch crossing) is presently ignored in FSIM given the explicit tuning of the luminosity optimized for single interactions per bunch crossing as mentioned in Section 2.3.1. However, scenarios are being discussed to run at higher luminosity and an implementation of the corresponding pileup would then be needed to re-assess jet separation.

As illustrated in Section 3.4.3 the resolution on the impact parameter has to be reproduced in order to perform meaningful b-tagging at jet level. The distribution of the measured impact parameter of tracks originating from the primary vertex in minimum bias events is compared in Figure 3.4 between SICB and FSIM for independent samples corresponding to the same integrated luminosity. On the same basis, and with the method discussed in Section 3.4.3, Figure 3.5 gives the transverse momentum distribution of the two jets with the highest p_T found in minimum bias events in full simulation (solid) and with the FSIM (dashed). Both samples correspond to the same integrated luminosity.

Several factors affect the precision on the impact parameter measurement given in Figure 3.4: mainly the resolution on the primary vertex and the resolution on the track slope measurement induced by multiple scattering. The effect of the latter component is amplified by the propagation of the first hit back to the production vertex of the particle [1]. Finally, but to a lesser extend, the first hit resolution will affect as well the determination of the impact parameter.

The resolution on the primary vertex is reproduced by a Gaussian smearing which agrees with the values predicted by SICB. The dependence of that reso-

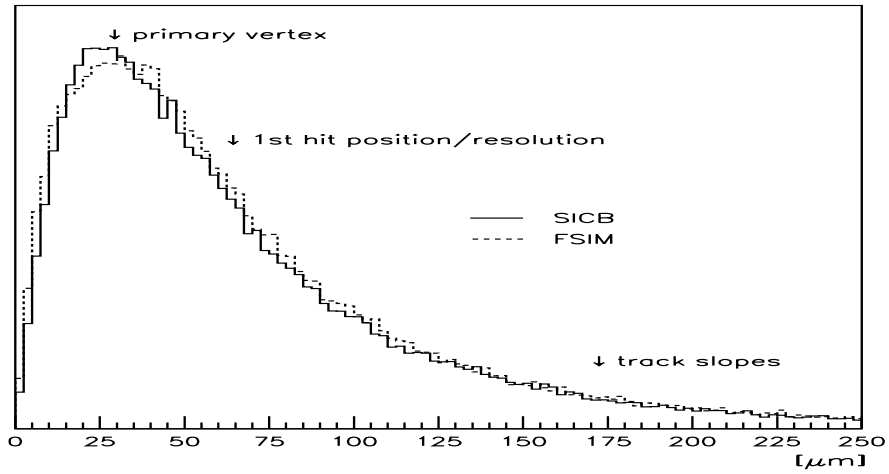


Figure 3.4: Distribution of the impact parameter of tracks in minimum bias events originating from the primary vertex, measured in SICB (solid) and reproduced in FSIM (dashed). Both samples correspond to the same integrated luminosity.

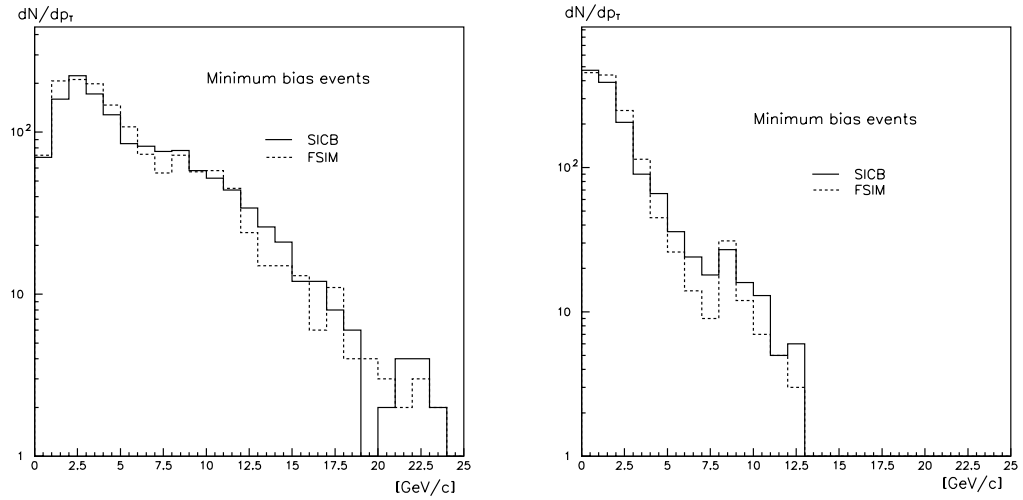


Figure 3.5: Transverse momentum distributions dN/dp_T of the more and less energetic jets in dijets found in minimum bias events. In both cases, the FSIM gives results in fair agreement with the simulated measurement in SICB, for samples corresponding to the same integrated luminosity.

lution with the number of (primary) charged tracks in the event — upon which the primary vertex reconstruction sensibly depends — is included.

The precision in the measurement of a vertex is also a function of its distance from the first measurement station. This effect has been included as a function of the track momentum p and is responsible for the long tail in Figure 3.4.

In reference to the dependence induced by multiple Coulomb scattering, the resolution on track slopes measurement has been also parameterized as a function of the track momentum.

An approach consisting in the use of the track transverse momentum p_T as parameter (instead of p) has been assessed to give equivalent results.

A switch in FSIM allows to provide a p_T kick to the charged particles which reproduces the average effect of the 4 Tm magnetic field. Curling of low energy particles is not implemented, particles with $p < 1$ GeV/c are discarded in FSIM. As will be discussed in Section 3.4.1, the use of the tracking information for charged tracks allows to correct for the magnet effect in the jet reconstruction. Generally, the magnet has been kept off in FSIM, unless otherwise specified.

Energy smearing can finally be applied to particles to account for calorimeter finite resolution. The amount of smearing is inferred from the resolution predicted by the full simulation [54] in agreement with the parameterizations shown in Section 2.3.6. The calorimeter segmentation and the shower shape are not taken into account explicitly, but are effectively included by the parameterized resolution.

There is unfortunately no simple way to introduce particle misidentification in calorimetry (e.g. electron/hadron) without a faithful modelling of shower shapes[5] [55]. This feature is strongly dependent on the configuration of the event (e.g. mini-jets of π^0s and $\pi^{\pm}$) and randomly assigning a lepton identity to a hadron at four-vector level with some probability would most likely just bias the study rather than rendering it more realistic. This issue is presently ignored in FSIM, which assumes perfect particle identification.

Finally, a switch allows to remove neutrinos from the list of particles. They are normally excluded, except for special tests.

3.4 Jet finding in LHCb

This section deals with the jet finding and reconstruction performance in LHCb. The identification of jets originated from b quark hadronization will be presented.

Higgs events decaying into a $b\bar{b}$ pair will be used as benchmark to drive the tuning of the algorithms in terms of mass resolution and tagging efficiency.

In Section 3.4.1 will be discussed at FSIM level the two methods mentioned in Section 3.1: an exclusive 2-jets reconstruction which makes use of a cone-based

[5]The next release of GEANT, v4, will provide standalone routines to simulate shower shapes.

algorithm, and an inclusive jet reconstruction based on clusterization. The two methods show equivalent performances.

To complete our analysis, a full simulation study has been done to define a realistic energy calibration procedure, which will be presented in Section 3.4.4.

3.4.1 Jet algorithms [9]

Jet finding algorithm have to find collimated structures in a set of four-vectors, regardless of the source (partons, hadrons or detector signal).

For the jet algorithm, the final state of the collision is represented as a starting set of defined *"particles"* with four-momenta p_i^μ. The starting p_i^μ may be the four-vector of reconstructed particles track or hit calorimeter cells at detector level, or simply identified to the final state particles at four-vector level.

Iterative cone algorithm

The cone algorithm is typically defined [56] in terms of the particles n whose momenta $\vec{p}_n$ lie within a cone of radius R centered on the jet axis (η_J, ϕ_J), in pseudorapidity η and azimuthal angle ϕ, as shown on Figure 3.6:

$$\sqrt{(\eta_n - \eta_J)^2 + (\phi_n - \phi_J)^2} < R \tag{3.2}$$

The jet angles (η_J, ϕ_J) defining the jet axis are the transverse-energy-weighted means of the angles of the particles:

$$\begin{cases} \eta_J = \sum\limits_{n \in \text{cone}} \frac{p_{T,n}\,\eta_n}{E_{T,J}} \\ \\ \phi_J = \sum\limits_{n \in \text{cone}} \frac{p_{T,n}\,\phi_n}{E_{T,J}} \end{cases} \quad \text{with} \quad E_{T,J} = \sum_{n \in \text{cone}} p_{T,n} \tag{3.3}$$

This process is iterated until the cone center matches the jet axis computed with 3.3. In principle, the algorithm requires to find *all* the possible solutions to 3.2 and 3.3, so giving an inclusive output list of jets. In practice, however, the search for valid jet cones begins by using "initiator" calorimeter cells, typically with E_T above some threshold value.

In conclusion, the cone algorithm consists in the search of the direction which maximizes the energy (or transverse energy) flowing inside the cone as illustrated in Figure 3.6(left). The complications start when two of these cones overlap. The properties of the resulting jets and hence the jet cross section predictions depend[6] strongly on the exact treatment of the overlap region. Various strategies exist to overcome the problem such as drawing a dividing line half way between the centers or giving all the energy to the more energetic of the two jets. However none of these is without ambiguities.

[6] A fortiori, according to whether the algorithm is run at parton, hadron, or detector level.

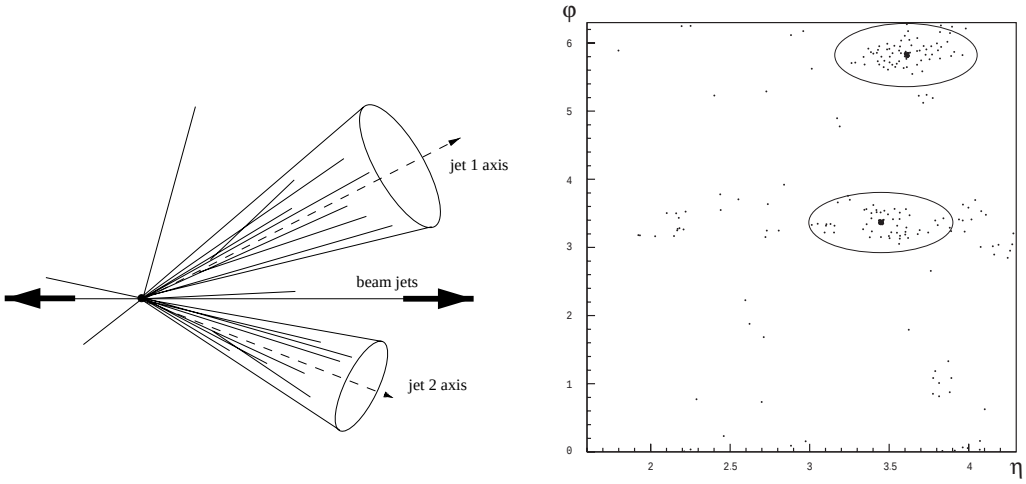


Figure 3.6: *Left:* Schematic representation of the cone algorithm. *Right:* illustration at calorimeter level in full simulation. The dark dots correspond to the direction of the jet axis. In the present study, we use a fixed cone algorithm with the direction of b-hadrons as jet axes.

Fixed cone algorithm

This type of algorithm does not calculate iteratively all possible jet axes but starts from a list of *predefined primary seeds* around which the energy is summed up (in a cone according to 3.2). The primary seeds are taken to be (calorimeter) cells or clusters of cells above a chosen cutoff $E_{(T)}^{\text{seed}}$. Beginning with the most energetic jet, all entries contributing to it are removed for further processing.

This operation — with arbitrary thresholds — is delicate because seeds of less energetic jets may be discarded as well and may therefore cause a bias on the final results. However, given the "dedicated" aspect of our search — $HX \to b\bar{b}Y$, giving exclusive final states — we have chosen to define the jet axis as the direction of the b-hadron. This choice is also based on the specific capability of LHCb to (fully) reconstruct b-hadrons.

Section 3.4.3 shows how b-hadron directions are reconstructed to get the cone axis.

Figure 3.6(right) shows the application of our fixed cone algorithm definition in full simulation on a $H \to b\bar{b}$ event mapped into (η, ϕ) space. Displayed are the (electromagnetic) calorimeter cells with energy deposition above 100 MeV. The two fat dots are the directions of the two b and $\bar{b}$ quarks four-vectors taken as seed for the jet axis (see discussion in Section 3.4.3). The circles ($R = 0.45$) show the extension of the cones into which merging proceeds.

The K_T cluster algorithm [57]

Clustering algorithms work in a very different way from cone algorithms: instead of globally finding the jet direction, they start by finding pairs of particles which are close in space and merge them together to form new pseudoparticles. This continues iteratively until the event consists of a few well-separated pseudoparticles, which are the output jets [58].

Numerous "flavours" of clustering algorithms exist, basically differing in the way the "nearness" (or distance) of two particles is defined. Historically, we can cite the original one developped by the JADE collaboration (distance $\sim E_i E_j(1-\cos\theta_{ij})$, i.e. a mass) and the Durham algorithm (distance $\sim \min(E_i^2 E_j^2)(1-\cos\theta_{ij})$) used in e^+e^- collisions. The K_T algorithm is a generalization of the Durham one intended to allow for the definition of inclusive jet cross sections in collisions producing beam jets ($p\bar{p}, ep, pp$).

Briefly, the algorithm proceeds as follows:

1. For every pair of particles, define a distance

$$d_{ij} = \min(E_{T_i}, E_{T_j})^2 \Delta R_{ij}^2, \qquad \Delta R_{ij}^2 \equiv \Delta\eta_{ij}^2 + \Delta\phi_{ij}^2$$

   which makes the algorithm invariant under boosts along the beam axis as in the cone algorithm. This feature is referred to as "longitudinal invariance".

2. For every particle, define a closeness to the beam direction,

$$d_{ib} = E_{T_i}^2 R^2$$

   where R is an adjustable parameter of the algorithm

3. Find the minimum in both sets of inter-particles distance $\{d_{ij}\}$ and particles-to-beam distance $\{d_{ib}\}$. If $\min\{d_{ij}\} < \min\{d_{ib}\}$ *merge* particles i and j

4. If $\min\{d_{ib}\} < \min\{d_{ij}\}$ jet i is *complete*

These steps are iterated until a given *stopping condition* is satisfied. Depending on what kind of studies one is interested in, different stopping conditions are useful. For *inclusive* jet studies, one iterates the above steps until all jets are complete. In this case, all opening angles within each jet are smaller than R and all opening angles between jets are larger than R. This means that the resulting jets are very similar to those produced by the cone algorithm. If one instead wants to reconstruct exclusive final states, one iterates the above steps until all jet pairs have $d_{ij} > d_{\text{cut}}$, an adjustable parameter in the algorithm which acts as a resolution parameter. All complete jets with $d_{ib} < d_{\text{cut}}$ are discarded (merged with the beam remnants).

Although the cluster jets are very similar to the cone jets in terms of the results of inclusive cross section measurements, their main practical advantage is that the overlap problem has completely disappeared. The algorithm unambiguously assigns every particle to a single jet in a dynamic way, adjusting to the shapes of the individual jets.

Recombination schemes

Different ways of merging two four-vectors into one define different recombination schemes. The two most common are the (covariant) "E scheme" — simple four-vector addition $p^\mu_{ij} = p^\mu_i + p^\mu_j$ — which seems the most natural for reconstructing particle decays; and the "p_T-weighted scheme"

$$\begin{aligned} p_{T,ij} &= p_{T,i} + p_{T,j} \\ \eta_{ij} &= (p_{T,i}\eta_i + p_{T,j}\eta_j)/p_{T,ij} \\ \phi_{ij} &= (p_{T,i}\phi_i + p_{T,j}\phi_j)/p_{T,ij} \end{aligned}$$

which produces massless jets.

Both schemes become equivalent for small opening angles. Although there are some practical differences, there is no strong preference for either scheme over the other. In the following, the E-scheme will be used in order to exploit the mass of single jets as preliminary discriminant variable against badly reconstructed jets.

3.4.2 Contributions to the jet energy resolution

The performance of both the fixed cone and K_T clustering algorithms has been studied using the dijet mass spectrum $M(j_1 j_2) = |j_1 + j_2|$ from $H \to b\bar{b}$ events, where j_1, j_2 are the four-vectors of the two jets reconstructed by the algorithm.

The study has been performed in FSIM with the default tunings described in Section 3.3. In particular, we assume a perfect energy calibration of the calorimeters and a perfect detection efficiency, to focus on the intrinsic acceptance and hadronization contributions.

One recalls that the intrinsic width of the SM Higgs boson is 3 MeV/c^2, for $m_H = 115$ GeV/c^2 (see Figure 1.4) and will be ignored in the following discussion.

Typical results can be seen in Figure 3.7 for the K_T clustering algorithm for different values of the radius R.

The resulting dijet mass distributions have non-negligible tails which affect the standard deviations. In order to meaningfully parameterize the width of the peak, a smooth [59] fit is performed, out of which the full width at half maximum FWHM is determined. The central value M found is taken as peak value and the ratio FWHM/M is used as the relative resolution[7].

[7]From a practical point of view, the mass relative resolution may be estimated with a Gaussian best fit around the peak and would become σ_M/M with $\sigma_M = \mathrm{FWHM}/2\sqrt{2\ln 2}$.

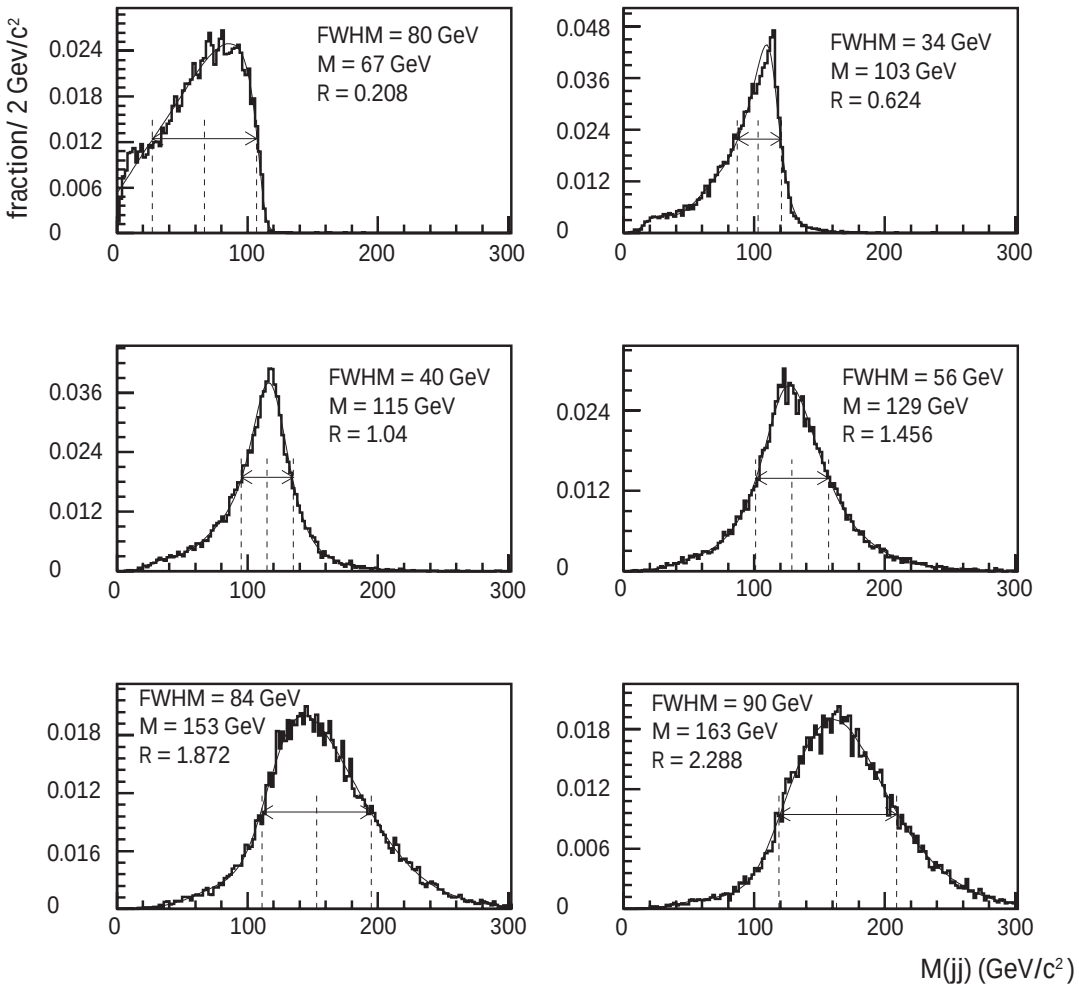


Figure 3.7: FSIM: output of the best fit procedure to estimate the dijet mass resolution for $H \to b\bar{b}$ events with increasing radius values obtained with the K_T jet clustering algorithm ($m_H = 115$ GeV/c^2).

Assessing the mass resolution

Figure 3.8(top) shows the evolution of the dijet mass resolution for increasing values of the radius parameter R. A comparison has been done between the fixed cone algorithm (circles) defined above and the K_T algorithm (triangles). Jet algorithms have been run on the same sample of events in fast simulation using all final state particles visible in the LHCb acceptance, then after reintroduction of the neutrinos, and finally after removal of the particles from the underlying event.

The cone axis is taken to be the direction of the $b(\bar{b})$ quark. For the K_T algorithm the two jets in the inclusive list lying the closest to the $b(\bar{b})$ quarks are used. In case the same jet is found to be the closest to simultaneously the b and $\bar{b}$ quarks no second jet is picked and the mass of the jet is taken to be the mass of the mother particle (typically for large values of the R parameter).

Curves 1–4 show three regimes: up to radii values $R \lesssim 0.5$ the mass resolution improves by a highly pure collection of the b-jets fragments. For radii $0.5 \lesssim R \lesssim 1.2$ the mass resolution is about constant but one starts to collect underlying event particles and the mass value increases linearly for both algorithms as shown on Figure 3.8(bottom) as a manifestion of the pedestal effect seen in Section 3.2, Figure 3.3b). A third regime, is visible for radii $R \gtrsim 1.2$ where the mass resolution degrades due to a dominant collection of uncorrelated particles from the underlying event.

Curve 5 gives the mass resolution in the case where only the Higgs decay products in the acceptance – including the neutrinos – are considered (i.e. no underlying event). The mass resolution improves monotonically as one may expect and tends towards an irreducible resolution term of around 5–7% reflecting the loss of information out of the acceptance. One can also see on Figure 3.8(bottom, dotted curve) that the peak value in the dijet invariant mass spectrum quickly saturates to the nominal Higgs mass.

Curves 3 and 4 come from the full event in the acceptance, with neutrinos included. The gain in resolution, with respect to curves 1 and 2, monotonically diminishes with increasing radius due to a growing contribution of the underlying event.

The cone and K_T clustering algorithms show equivalent performances over a large range of radius parameter values. However, a difference is noticeable for large radii, typically when the cone algorithm faces overlap problems. The mass resolution steadily worsens after $R \gtrsim 1$ whereas K_T saturates, giving the crossover for $R \simeq 2$ in Figure 3.8(top). The constant rise comes from the fact that one does not select non-overlapping cones. By construction K_T precludes overlaps and for large enough R values the event is simply merged into a single jet[8].

[8]Regarding the issue of cross sections determination in QCD, the cone algorithm does not allow to separate overlapping jets, but the definition of a single jet may not be possible either.

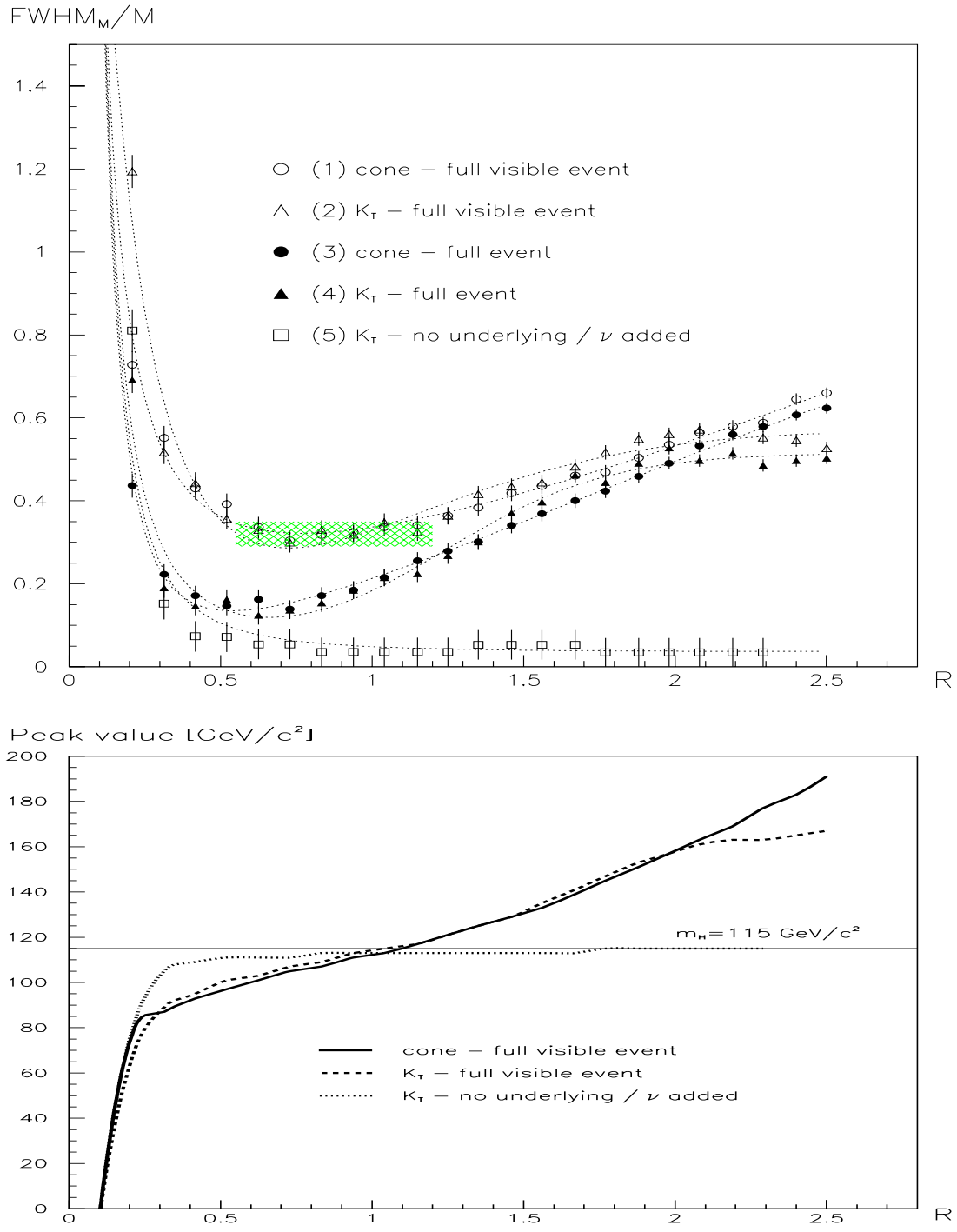


Figure 3.8: FSIM – *Top:* Dijet mass resolution for $H \to b\bar{b}$ events as a function of the parameter R used in the cone ($\circ, \bullet$) and K_T ($\triangle, \blacktriangle, \square$) jet algorithms. The errors are correlated. The curves are used for eye guidance. *Bottom:* Displacement of the peak value in the dijet mass spectrum as a function of R.

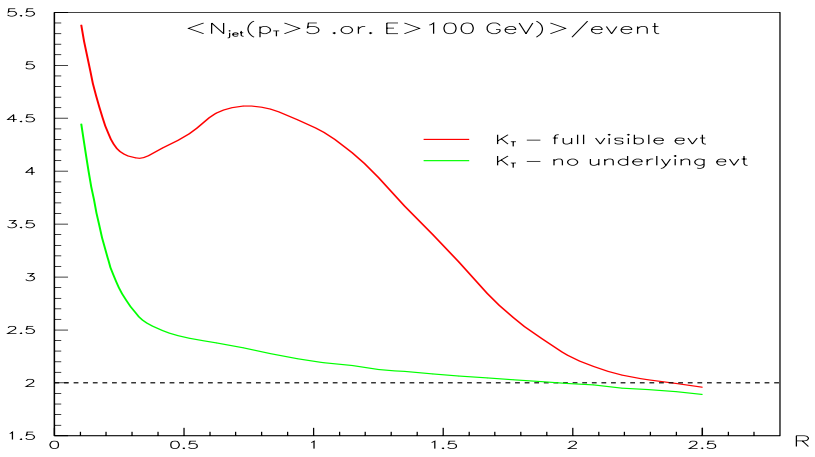


Figure 3.9: FSIM: Average number of jets per event with $p_T > 5$ GeV/c or energy > 100 GeV found with the K_T algorithm, as a function of R. Statistical errors are negligible.

For jet reconstruction using the full visible information, the hatched region emphasizes the plateau giving a best mass resolution of around 30% (corresponding to $\sigma_M/M \simeq 13\%$, see Section 3.4.4).

It has to be noticed that for optimal values of the radius parameter, $R \sim 0.5$, the main contribution to the dijet mass resolution comes from the escape of the neutrinos ($\Delta\frac{\text{FWHM}}{M} \sim 20\%$ between curves 1 and 3). Most of the remaining contribution to resolution comes from inefficient energy collection, originating from either a too small jet radius or energy flowing out of the detector acceptance (no cuts have been used to define a fiducial volume). It will be shown in Section 3.4.4 that the optimal dijet mass resolution found above is reproduced at *detector level* after proper calibration of the jet energies.

Towards improvement

As a consequence of our search limited to two (b-)jets, a contribution to the mass resolution will come from events with hard gluon radiation in the final state. In principle, the use of the K_T algorithm allows to reconstrut jets inclusively, including thus the jets other than the two leading ones.

Figure 3.9 shows the evolution of the number of "hard" jets found by K_T as a

This is a controversial feature of the cone algorithms: a single jet consisting in two widely separated subjets may not be correctly detected as such by lack of initiator cells in between. This is a well documented issue, see for instance [9],[56].

function of the R parameter, for Higgs events with and without underlying event. For both cases the average event is seen as of two-jets type at large radii.

The jet finding with $R \sim 0.6$ indicates on average one extra jet every three events in the absence of underlying event, while four jets per event are found in the full event. The rising between $0.3 \lesssim R \lesssim 1$ is another manifestation of the underlying event. Extra hard jets would have to be found with a small[9] radius $R \sim 0.25$. In this regime, there would be on average one extra good jet to be caught out of the four in the event: the two b-jets, one radiation jet (gluon) and one from the underlying event.

An additional strategy is required to identify the extra jet as part of the Higgs products. Figure 3.3, at particle level, shows that this kind of jet will be hard to isolate on the basis of relative distance to the leading b-jets for instance. One may think of a neural network analysis to try to identify characteristic patterns in the angular-ordered parton cascade picture.

The expected improvement, however, should not exceed $\Delta\frac{\text{FWHM}}{M} \sim 7\%$, referring to the difference between the valley on curve 4 and the asymptotical mass resolution on curve 5 in Figure 3.8(top). This kind of studies has not been performed yet.

3.4.3 The seed finder

We now switch to the *fixed cone* algorithm presented in Section 3.4.1 and discuss the determination of the jet axis.

The motivation for the choice of the b-hadron direction as jet axis is twofold. On the one hand, we expect the b quarks to be leading in the fragmentation. This can be seen by considering the quark-antiquark pair production in the Lund string fragmentation model [53]. The production is given by a quantum mechanical tunnelling picture, with a probability $\propto \exp(-\frac{\pi m_T^2}{\kappa})$, where m_T is the transverse mass of the quarks and κ the string energy (typically $\mathcal{O}(1\ \text{GeV/fm})$). The expression leads to a suppression of heavy quark production in a ratio:

$$\mathcal{P}_u : \mathcal{P}_d : \mathcal{P}_s : \mathcal{P}_c \approx 1 : 1 : 0.3 : 10^{-11}$$

Charm and heavier quarks are not likely to be produced in the soft fragmentation, as experimentally verified, and the probability for a light quark jet to be contaminated by a c or b quark is therefore negligible. There is actually in Pythia no production at all of flavours heavier than u, d, s in the fragmentation process, but only as part of the parton shower evolution ($g \to b\bar{b}, g \to c\bar{c}$).

On the other hand, in the hard fragmentation region, the attachment of a light antiquark $\bar{q}$ to the leading b quark only produces a slight deceleration of the latter. The corresponding probability function $f(z)$ for the hadron to take a

[9] Or, equivalently, by studying the substructure of larger jets with a comparable distance resolution.

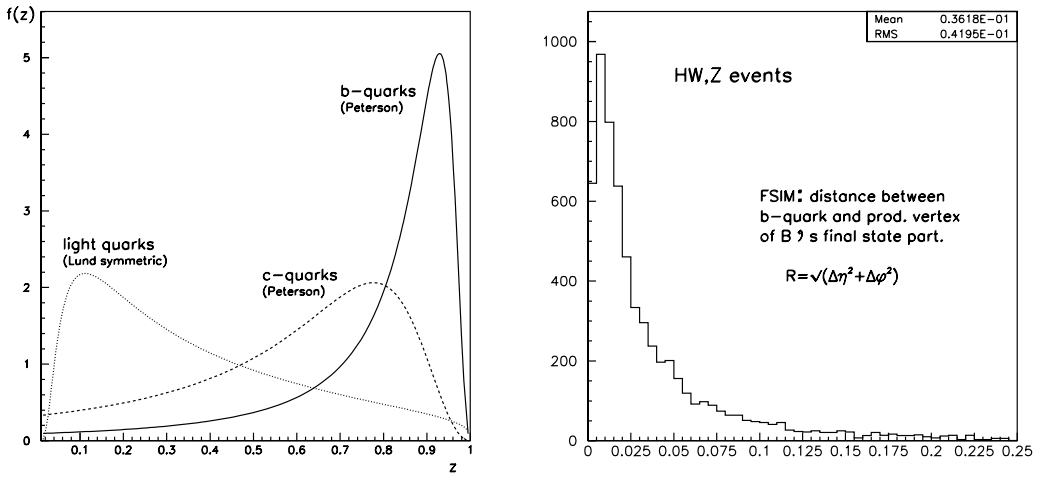


Figure 3.10: *Left:* Fragmentation functions as function of the parton momentum fraction taken by the hadron. *Right:* R distance between the b quark momentum and the direction of the production vertex of the charged particles in the b-hadron decay cascade.

fraction z of the momentum of the leading quark is a function of the quark mass. Figure 3.10(left) shows the functions used commonly for light and heavy quarks, compatible with hadro-production data. The Lund symmetric fragmentation function, with a parameterization almost independent of the flavour (u, d, s) and the Peterson function for charm and bottom are displayed.

The strong peak towards $z \approx 1$ in the Peterson model for the bottom quark thus reflects at hadron level a b-hadron strongly collinear to the b quark.

The transverse distance in Equation 3.1 between the b quark and the b-hadron, is small with respect to the typical extension of the shower $\Delta R_{\text{shower}} \sim 0.5$ (Figure 3.3(b)). This can be inferred from Figure 3.10(right) reporting the R distance between the b quark and the direction of the production vertex of all the particles in the b-hadron decay *cascade*[10]. The distribution vanishes for small ΔR values (first bin) because the b quark is displaced by the hadronization.

The bulk of the distribution is peaked at $\Delta R < 0.02$ with an average displacement $\langle \Delta R \rangle \sim 0.04$ small compared to the optimal cone aperture of $R \sim 0.5$, and also comparable to the calorimeters spatial resolution (see Figure 2.9).

The LHCb dedicated capabilities to identify exclusive b-hadron channels is another motivation to define the jet axis as the direction of the b-hadron.

[10]Referring to a typical decay chain of the type $B \to D + X \to K + X + Y$.

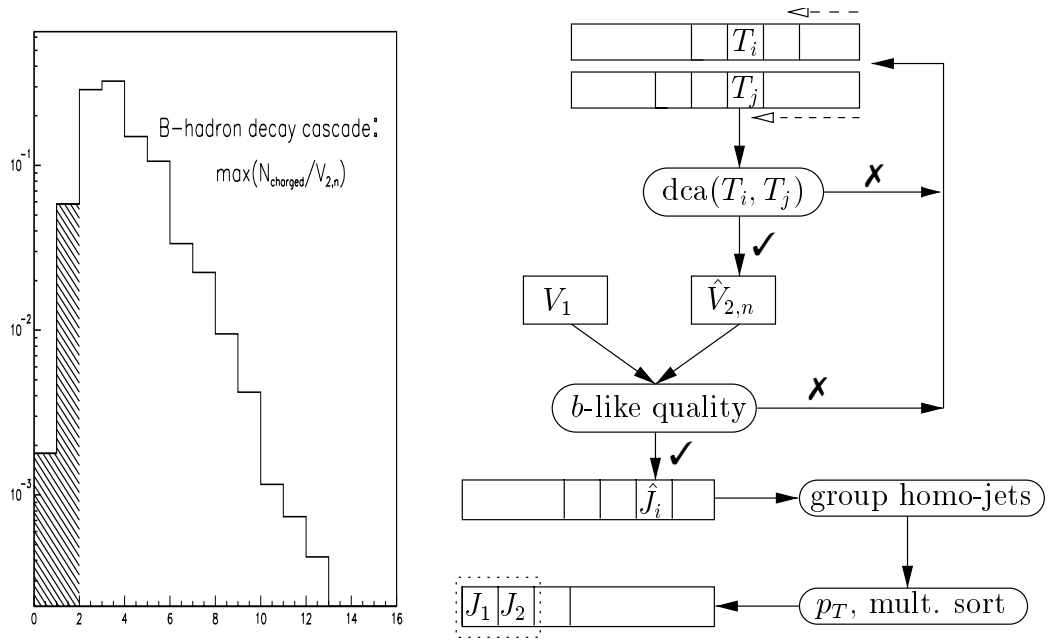


Figure 3.11: *Left:* Maximum multiplicity in charged particles originating from the same vertex in a b-hadron decay cascade. *Right:* Schematical description of the seed-finder algorithm.

Determination of the b-hadron direction

We define the b-hadron direction as the line joining the primary vertex to the b-hadron decay vertex. The procedure consists thus in identifying all significantly displaced vertices in the event. The cuts will be weaker than in CP studies to keep an efficiency as high as possible on Higgs events.

The algorithm has been implemented exactly in the same way in FSIM and SICB and is called "seed-finder" in the following to make it distinct from the jet proper reconstruction. The light quark jet retention is discussed in Section 4.4.

Figure 3.11(right) shows the organization of the algorithm motivated by the high multiplicity of charged particles originating from a same vertex in a b-hadron decay cascade. This can be seen in Figure 3.11(left) showing the distribution of the maximum multiplicity for the vertices in the cascade. The hatched portion represents the fraction ($\sim 6\%$) of B-hadron decays giving less than two tracks as maximum vertex multiplicity. The direction of such B-hadrons will not be determined by our algorithm.

All possible 2-track vertices ($\hat{V}_{2,n}$) are formed in a preselection of reconstructed tracks with a distance of closest approach (dca) below $\mathsf{dca}_{\max} = 500\,\mu$m. The

Scenario	Di-jet (single) efficiency	Di-jet (single) purity
standard	85% (92%)	97% (98.5%)
$N_{\min}^{\mathrm{IP}} = 3$	72% (85%)	97.5% (99%)

Table 3.2: Efficiency and purity of the seed finder for Higgs events with a dijet mass in the mass window $95 < M(JJ) < 130$ GeV/c^2. The efficiency and purity for a *single* jet can be estimated by the square root of the latter numbers (given in parenthesis).

value of $\mathsf{dca}_{\max}$ has been determined from the distance distribution between pairs of tracks originating from the same vertex in the full simulation. To pass the preselection, tracks are asked to have a momentum $p > 1$ GeV with a significance $p/\sigma_p > 3$, or a significance $s_{x,y}/\sigma_{s_{x,y}} > 10$ on the slopes in both the bending and non-bending planes. These criteria prevent the rejection of tracks with correctly reconstructed slopes, but with a biased momentum estimate due to a conversion before the magnet. Further preselection keeps the tracks (T_i) with high impact parameter (IP) in the vertex detector, with $30\,\mu\mathrm{m} < \mathrm{IP} < 10$ cm and a significance $\mathrm{IP}/\sigma_{\mathrm{IP}} > 3$.

The 2-track vertices $\hat{V}_{2,n}$ are required to be downstream with respect to the primary vertex V_1: $z(\hat{V}_{2,n}) > z(V_1)$. The displacement $\delta(V_1, \hat{V}_{2,n})$ between the primary and each of the 2-track vertices has to satisfy $2\,\mathrm{mm} < \delta(V_1, \hat{V}_{2,n}) < 40$ cm with a significance of at least 3σ to isolate b-hadron-like times of flight.

At this point, each proto-vertex is taken as axis to form a proto-jet $\hat{J}_i$. The high multiplicity in the b-hadron decay cascade will typically give a set of homothetical proto-jets with proportionnal four-vectors

$$(E_J, \vec{p}_J)_i \simeq \lambda \cdot (E_J, \vec{p}_J)_k$$

which are clusterized together within a distance $\Delta R_{ik} < 0.2$. Each set of proto-jets is then used to form the seed of the jet J_i. A b-jet seed candidate is calculated from the set as the average direction of the proto-jets $\hat{p} = \langle \vec{p}_i \rangle$. Finally, the two jets with the highest proto-jets multiplicity are picked as final seed candidates. In case two or more jets with the same multiplicity are found, the one (or two ones) with the highest p_T is (are) taken. The possibility exists to set a lower cutoff in p_T for the jets.

Figure 3.12 shows a Higgs event $HW \to b\bar{b}e\nu_e$, where the two b-jets have been reconstructed around the jet axes found by the seed finder (dark cells). The extra activity at high pseudorapidity is produced by remnants of the beam jet and has been discarded by the seed finder. The associated electron is isolated from both b-jets.

The efficiency and purity of the algorithm are assessed in Table 3.2 by comparing the number of entries in a fixed mass window with respect to the case where the MC b quarks are taken as jet seed, as shown in Figure 3.13. The default

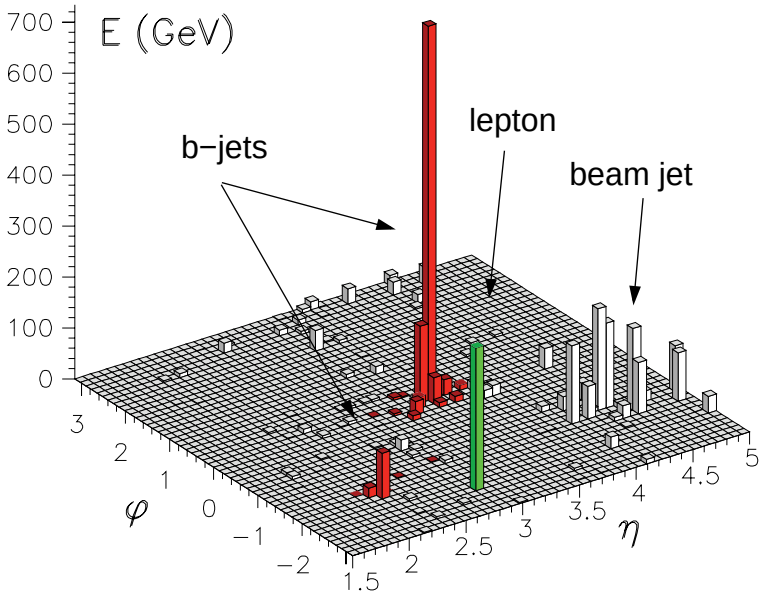


Figure 3.12: $HW \to b\bar{b}e\nu_e$ event at calorimeter level (full simulation). The two $b-$jets have been reconstructed by the jet algorithm. The associated electron is isolated.

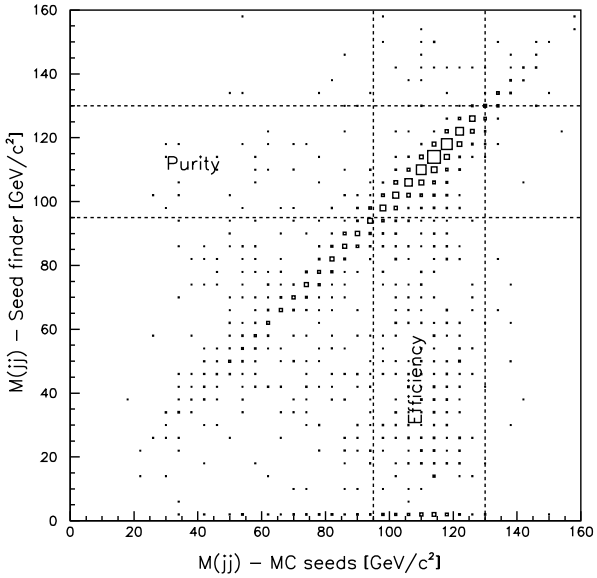


Figure 3.13: Definition of the mass regions used to determine the efficiency and purity of the seed finder algorithm. On the horizontal axis is the dijet mass obtained when taking the MC b quarks as jet seeds, and on the vertical axis the dijet mass obtained when using the seed finder on the same event.

cone radius used to form the proto-jets is $R = 0.6$. The second line corresponds to a selection of jets with a minimal number $N_{\rm min}^{\rm IP}$ of tracks with large impact parameter inside the jet, as a crude b-tagging method.

To conclude this section, let's mention that by construction the present algorithm could inspire the trigger algorithms for efficient selection of the Higgs events (e.g. in the steps leading to L1 decision).

3.4.4 Jet energy calibration

A first study of a realistic jet energy calibration was done in SICB and presented in [54]. The jet reconstruction was performed at calorimeter level except for the muon information from the tracking. The jet calibration coefficients were compared to those obtained for the calibration of single electron or pion at different energies and were found to be close. The energy resolution for the combined calorimetry obtained in particle-gun mode[11] were then used as input of FSIM. Finally, the dijet mass resolution in full simulation of $H \to b\bar{b}$ was found to be in agreement with the resolution obtained in FSIM. This reinforces our confidence in FSIM.

In the present work, the jet reconstruction considers the scenario where the momentum information is given by the tracking system, leaving the calorimeter information for neutrals. The main goal is to assess the impact of the calorimeters segmentation and energy smearing on the dijet mass resolution.

At this stage, we define the neutrals as the calorimetric cells hit by tracks sharing no common MC reference with any reconstructed track. No study has been done yet to set up a procedure for associating tracks to clusters.

At detector level, particles are considered as massless and the recombination scheme is the natural covariant E scheme $p_{ij} = p_i + p_j$.

The reconstructed jet energy E_j is given by the linear parameterization

$$E_j = a_0 \sum_i^{\rm tracks} \|p_{j,i}\| + a_1 \sum_i^{\rm cells} E_{j,i}^{\rm ECAL} + a_2 \sum_i^{\rm cells} E_{j,i}^{\rm HCAL}$$

where the i is an index running over all the tracks ($\|p_{j,i}\|$) and cells ($E_{j,i}^k$, $k =$ ECAL,HCAL) inside a cone with defined radius R, and a_k ($k = 0, 1, 2$) are calibration constants. A least squares minimization of the difference with the jet energy using Monte-Carlo information gives the calibration constants quoted in Table 3.3. Performing an independent calibration on the two jets of the event sorted in p_T gives a variation in the calibration constants covered by the last digit.

Figure 3.14 shows the dijet invariant mass for a cone radius $R = 0.6$ reconstructed once at four-vector level (solid) in FSIM and after calibration in SICB

[11] One particle at a time with defined momentum is simulated inside the detector.

a_0 (tracks)	a_1 (ECAL)	a_2 (HCAL)
1.0	1.0	0.7

Table 3.3: Calibration constants to get the reconstructed jet energy in full simulation (including pileup events).

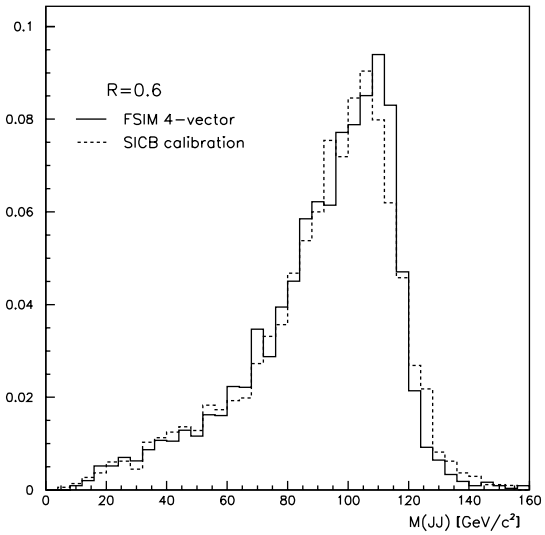


Figure 3.14: Dijet invariant mass obtained in FSIM (solid) and in SICB after calibration (dashed). The jet radius is $R = 0.6$.

(dashed). In the fast simulation, the magnet has been kept off and the energy resolution assumed to be perfect. The full simulation was run with the pileup conditions described in Section 2.3 (Figure 2.6) and with finite dynamic range of the calorimeters (see Table 2.1). The saturation has been modelled with a step function. It has been shown, however, that the saturation effects are expected to be minor [60]. The upper tail next to the right flank of the peak is somewhat more important in full simulation, mostly due to the presence of pileup events, while the peak value is slightly lower due to the saturation. In both cases however, the resolution is estimated to be $\sigma_M/M \sim 13\%$ according to the procedure described previously (see Figure 3.7).

The contributions to the mass resolution coming from the segmentation of the calorimeters and their own resolution in energy may therefore be considered as marginal, of the order of the error inherent to the fit procedure. Conversely, the effect of the magnet is major as was demonstrated in [54], pushing the mass

resolution to $\sigma_M/M = 18\%$. The key issue to achieve the predicted resolution presently resides in the track-to-cluster association procedure which remains to be worked on.

Chapter 4

Higgs search

4.1 Higgs kinematics

As seen in Section 2.2, light objects (with respect to $\sqrt{s}$) are emitted at high values of rapidity (see Figure 2.4).

Figure 4.1 summarizes common kinematical variables for SM light Higgs of $m_H = 115$ GeV/c^2 at LHC. The pseudorapidity distribution $dN_H/d\eta$ (a, solid histogram) shows that a significant fraction ($\sim 30\%$) of the Higgs bosons are emitted forward, within the LHCb acceptance.

We also represent the distribution of the b quarks from the Higgs decay. The correlation in pseudorapidity of the decay products is given in Figure 4.1(b), where the two boxes show the LHCb acceptance and that of a "central" detector. The transverse momentum distributions of the b-products within these boxes are similar in both the central and forward regions (see Figure 4.1(c)). According to the Lund string fragmentation model, the rapidity plateau (a, dashed histogram) in the distribution of the Higgs final state products may be related to the colour neutral state of the Higgs: the string is stretched between the two $b\bar{b}$ quarks of the decay. The distribution in azimuthal angle of the b quarks is flat (d).

4.2 Insight of the accessible channels

The point is to explore now the Higgs observation potential in the forward region with the LHCb detector.

The values for the Higgs production cross sections and for the Higgs branching ratios as a function of m_H are obtained with a set of standalone computer programs developed by various authors (HIGLU[61], HDECAY[62], PPHV[63]) including up to next-to-leading order QCD corrections. The parton density functions used for the calculations are the CTEQ4L according to the present tuning proposed by the LHCb generator group [64].

Figure 4.2 summarizes the SM Higgs decay channels rates for one LHCb year

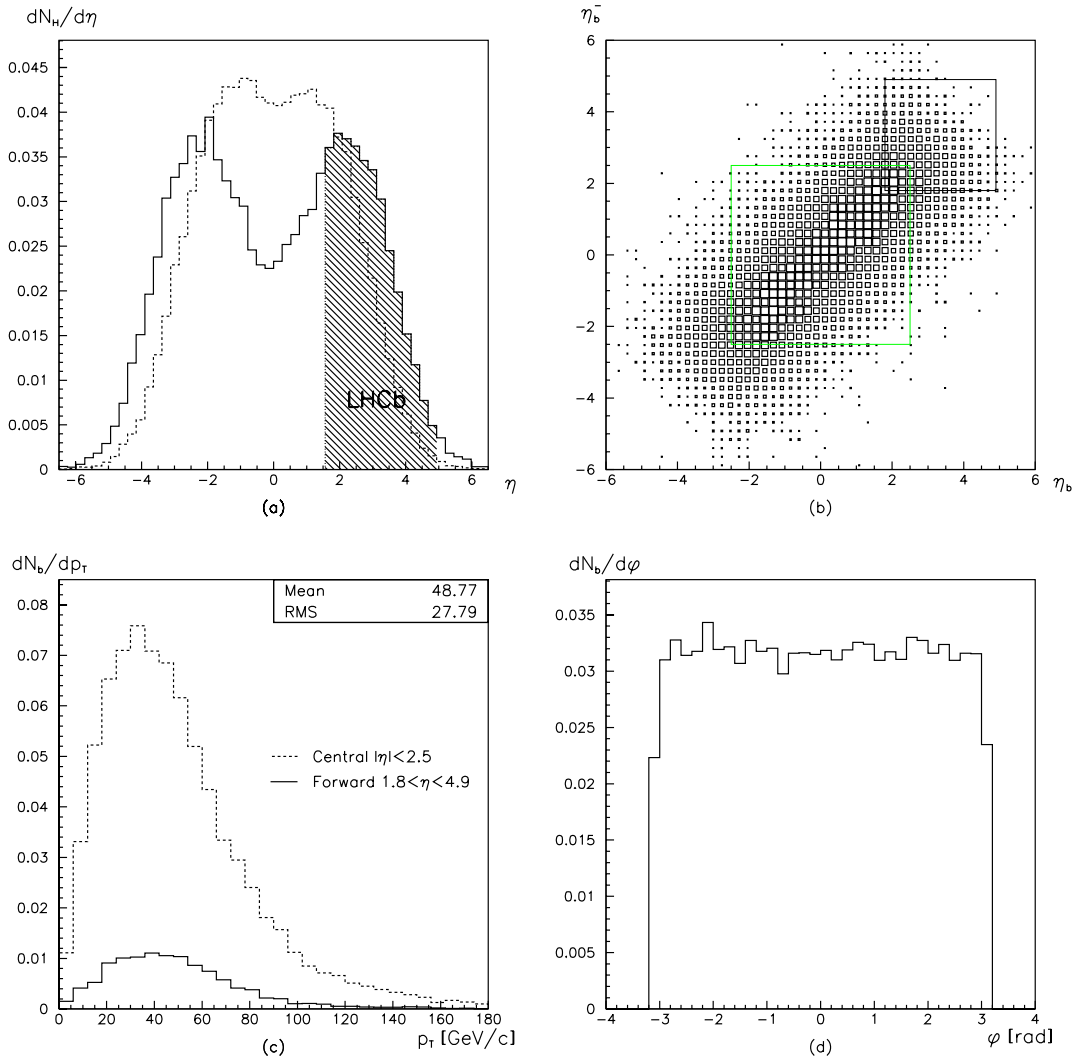


Figure 4.1: Kinematics of the $H(b\bar{b})W$ process. (a) Pseudorapidity distributions of Higgs particles (solid) and final state particles from the Higgs decay (dashed). The fraction covered by the LHCb detector is hatched. (b) Pseudorapidity of the b versus $\bar{b}$ quarks from Higgs decay. The forward and central region are indicated by the two boxes. (c) Transverse momentum distribution of the b quarks, in the forward (solid) and central (dashed) regions corresponding to the boxes in (b). (d) Azimuthal angle distribution of the b quark.

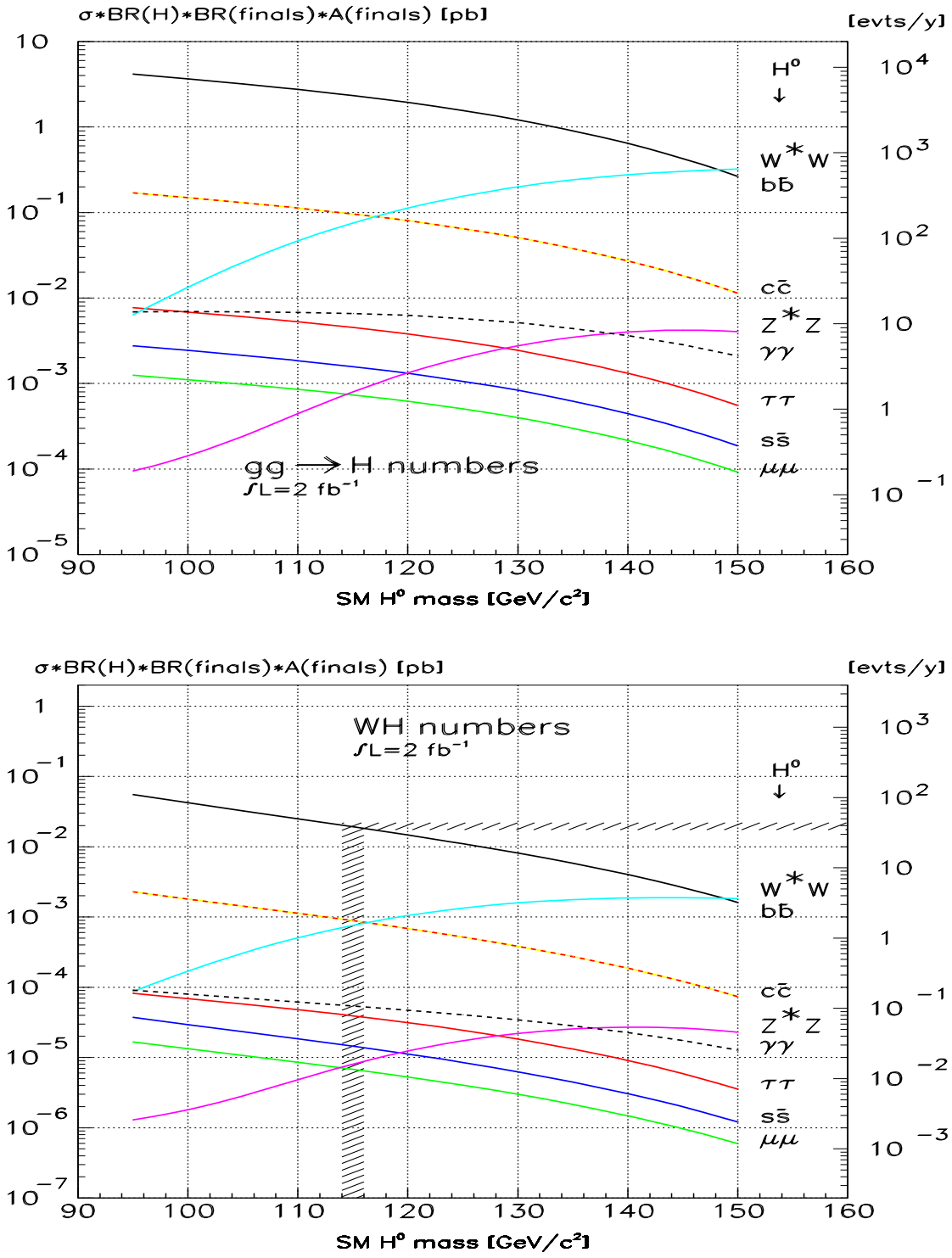


Figure 4.2: Cross sections at LO (left scales) and event rates (right scales) in the LHCb acceptance for the SM Higgs. *Top:* direct production via gluon fusion $gg \to H$. *Bottom:* production in association with a vector boson $f\bar{f}' \to HW$. The decay products of the W, Z and τ are in the acceptance as well. The hatches indicate the yield for Higgs with a mass in the range $m_H = 114 - 116$ GeV/c^2.

run (integrated luminosity of 2 fb^{-1}). The left scale shows the signal cross section, at leading order (LO), for the direct production $gg \to H$ (top plot) and for the production in association with a vector boson $f\bar{f}' \to HW$ (bottom plot) — the W, Z Bremsstrahlung process in Figure 1.3. The latter channel would provide a lepton to tag the presence of the vector boson. The potential for observing each channel is briefly discussed below:

The "visible" channels which have been considered are the following

$$\begin{array}{c} \mathrm{BR}(W \to q\bar{q}') \sim 68\% \\ \left.\begin{array}{ll} \mathrm{BR}(Z \to \ell^+\ell^-) & \sim 6.7\% \\ \mathrm{BR}(Z \to b\bar{b}) & \sim 15\% \end{array}\right\} \sim 22\% \, (\ell = e, \mu) \\ \mathrm{BR}(\tau \to 3\pi^{\pm}(\pi^0)\nu_\tau) \sim 15\% \end{array}$$

The effect of the LHCb acceptance has been approximated by requiring the relevant partons and leptons to be in the 15 mrad $< \theta <$ 330 mrad polar angle window. The number of events produced in one year at nominal luminosity (2 fb^{-1}) is indicated on the right scale. On the bottom plot the branching ratio $W \to \ell\nu_\ell \, , \ell = e, \mu$ of the associated W is taken into account and the lepton must be in the LHCb acceptance.

For the τ lepton, only the "3-prong" modes have been considered to allow for a reasonable di-lepton mass reconstruction.

The combined effects of the acceptance, production cross sections and selected decay channel of the associated W in the associated production mode give a yield about 100 times smaller than the single Higgs.

The potential in direct Higgs production may be summarized along the following criteria:

- Up to ∼130 GeV, the dominant decay channel is $b\bar{b}$. This signal would provide a tag, but is overhelmed by the QCD generic $b\bar{b}$ pairs forming an irreducible continuum. According to the QCD generator predictions (see next section):

  $$\sigma(M(b\bar{b}) > 80 \text{ GeV/c}^2) \simeq 250 \cdot 10^3 \text{pb}$$

  for generic $b\bar{b}$ pairs in the LHCb acceptance, to be compared to $\sigma(H \to b\bar{b}) \sim 3$ pb;

- The design inherent lack of hermeticity in LHCb substantially reduces the potential for channels with missing energy. For the sake of completeness, the favourite decay channel $H \to WW^*$ in the light to intermediate mass region has been however represented for W's decaying into hadrons, though mostly difficult to isolate from the generic QCD light jets irreducible background. The channel is furthermore strongly penalized due to the requirement to accept four jets in LHCb;

- The $H^0 \to \gamma\gamma$ channel ($\sim$ 12 events/year in acceptance for $m_H = 115$ GeV/c^2) is overwhelmed[1] by the generic Born ($q\bar{q} \to \gamma\gamma$), box ($gg \to \gamma\gamma$), and Bremsstrahlung continua ($qg \to q\gamma \to q\gamma\gamma$) and is precluded at low luminosity regime. Moreover, the calorimetry in LHCb has not been designed for this purpose and would not provide the required resolution and dynamic range;

- The $H^0 \to \tau\tau$ channel with τ decaying into ($\pi^{\pm}$-prongs)ν_τ may be of interest given the ability of LHCb to identify pions in the RICH detectors. Resolution on Higgs mass would be around 10% but the signal would be dominated by the tail (Breit-Wigner) of the irreducible $Z^0 \to \tau\tau$ decays. The same conclusion applies for the ~ 2 (!) events from $H^0 \to \mu\mu$.

Given then the difficulty to isolate the signal in direct Higgs production mode with respect to the generic QCD (and EW) backgrounds continua (see next section), the search has to be performed in the production mode in association with a W, Z boson.

The point is the search for final states with a $b\bar{b}$ pair together with a charged lepton (e, μ) from the associate W, yielding $\mathcal{O}(10)$ events per year up to masses $m_H \sim 130$ GeV/c^2 (see Figure 4.2(bottom)). The production channel in association with either a partially or fully reconstructed Z^0, though contributing modestly, is considered as well.

The Higgs production channel in association with a top pair, $t\bar{t}H^0$, would provide an interesting alternative channel for a lepton tag, but is however strongly suppressed in the forward region — as could be inferred from Figure 2.4(left) for an object weighting ~ 500 GeV/c^2.

4.3 SM Higgs search strategy in LHCb

At this point, the set of tools for jet reconstruction introduced in Chapter 3 is used. The analysis shown here has been performed at four-vector level with the fast simulation, unless otherwise specified. The coverage $1.8 < \eta < 4.9$ has been considered for the geometrical acceptance of the detector.

4.3.1 Event generation

As stated above, the signal to be investigated would be of the type $HV \to b\bar{b}\ell X$.

Table 4.1 gives the cross sections at leading order for the associate Higgs production with a vector boson, and the branching ratio into $b\bar{b}$ for increasing Higgs masses which will be taken as reference values for this study. The next-to-leading order (NLO) estimates are also indicated in parentheses, and are about 15% larger than LO.

[1] Not to mention the background from π^0 decays.

m_H [GeV/c^2]	$H^0W^\pm$ [pb]	H^0Z^0 [pb]	BR($H \to b\bar{b}$)
105	2.13 (2.44)	1.13 (1.29)	0.79
115	1.61 (1.84)	0.86 (0.98)	0.73
125	1.24 (1.42)	0.66 (0.76)	0.61

Table 4.1: Total cross sections at LO (NLO) for associate SM Higgs production at LHC and branching ratio into $b\bar{b}$ for three values of m_H in the mass range accessible to LHCb. The parton densities used are the CTEQ4L. Computed with PPHV[63] and HDECAY[62]. The LO values are taken as reference values for the present study.

In order to keep a coherent picture of the factorization scheme in pQCD, the signal and backgrounds have been generated using tree-level matrix element calculations together with partonic showering to reproduce the higher orders in the pertubation theory, as provided by the event generator Pythia [52]. The processes are simulated to all orders but the normalization for the Monte Carlo (MC) remains that of a leading order calculation[2].

Table 4.2 summarizes the different sources of signal and background. We give the total production cross section for the different channels and their relative ratio to the $H + W, Z$ cross section for a Higgs with mass $m_H = 115$ GeV/c^2. The number of events with all products of the tagging signature within LHCb acceptance are given for an integrated luminosity of 2 fb^{-1}. The leptons are required to have $p_T > 5$ GeV/c. The production channel in association with a top pair, $t\bar{t}H^0$, is also given though without entering the total due to its marginal contribution.

The $W^\pm + b\bar{b}$ background has been determined from the inclusive $W^\pm$ production cross section. The $b\bar{b}$ jets are found[3] in the initial state parton shower. The $\gamma^*/Z^0 + b\bar{b}$ background is estimated in the same way, but the channel is not exclusively resonant as the Drell-Yan process is also included. A lower cutoff has been set at $\hat{m} = 12$ GeV/c^2 to take into account the fraction of events yielding one or more lepton with a high p_T. Typically, the fraction of events with a lepton of $p_T > 20$ GeV coming from a Drell-Yan (off Z^0 resonance) is 10%.

[2]The parton showering process redistributes the event particles in phase space, but does not change the total cross section. Regarding Higgs production, the branching $q \to q + H$ is however not included in the shower, because giving only a negligible contribution to the higher orders [65].

[3]In the low-p_T region, this choice has to be preferred over a calulation using explicit matrix elements for higher order diagrams, such as $q\bar{q}' \to gW$ or $qg \to q'W$, in order to avoid double-counting [66].

$m_H = 115$ GeV/c^2	σ [pb]	Ratio	Events/year	$bb\ell, \eta_{\rm LHCb}$	$bb\ell\ell, \eta_{\rm LHCb}$
Signal					
$H^0W^\pm$	1.61	0.65	3220	37	—
H^0Z^0	0.86	0.35	1720	8	4
$t\bar{t}H^0$	(0.43)		(860)	(2.1)	($< 10^{-2}$)
Total	2.47	1	4940	45	4
Background					
$b\bar{b}$	$500 \cdot 10^6$	$2 \cdot 10^8$	10^{12}	$3 \cdot 10^9$	$\mathcal{O}(10^7)$
$t\bar{t}$ ($WbW\bar{b}$)	610	247	$1.2 \cdot 10^6$	$1.9 \cdot 10^4$	5300
$Z^0W^\pm, Z^0$	38	15.4	$7.6 \cdot 10^4$	138	57
$W^\pm$+jets	$160 \cdot 10^3$	$6.5 \cdot 10^4$	$3.2 \cdot 10^8$	$2.2 \cdot 10^4$	1750
γ^*/Z^0+jets	$73 \cdot 10^3$	$3 \cdot 10^4$	$1.46 \cdot 10^8$	$1.7 \cdot 10^4$	8200
Total	$500 \cdot 10^6$	$2 \cdot 10^8$	10^{12}	$3 \cdot 10^9$	$\mathcal{O}(10^7)$

Table 4.2: Summary of SM Higgs signals and backgrounds channels into $X \to b\bar{b}\ell X$ for a luminosity of $\int \mathcal{L} = 2$ fb^{-1} (1 LHCb year) and $m_H = 115$ GeV/c^2. The quoted cross sections are at LO, using the CTEQ4L parton density functions. The symbol $\eta_{\rm LHCb}$ designates particles entering the LHCb acceptance. The lepton must have $p_T > 5$ GeV/c. The $W^\pm$+jets and γ^*/Z^0+jets backgrounds are estimated by the inclusive production cross sections of respectively $W^\pm$ and γ^*/Z^0. The Drell-Yan events are generated above a cutoff $\hat{m} = 12$ GeV/c^2 (see text).

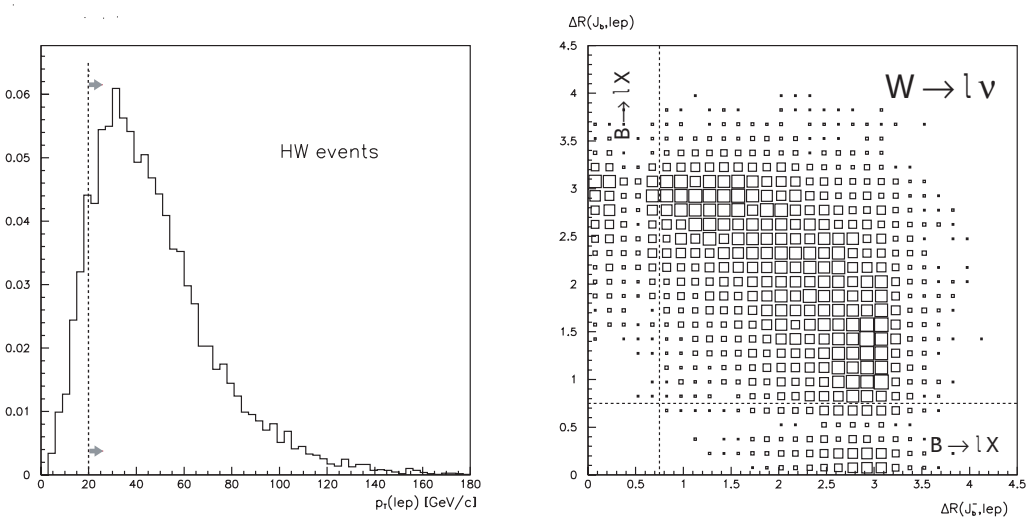


Figure 4.3: *Left:* Transverse momentum distribution of the lepton in $W \to \ell\nu_\ell$. *Right:* Isolation of the lepton from the associated W. On the vertical (horizontal) axis is the ΔR distance (3.1) between the $b(\bar{b})$-jets and the lepton (with $p_T > 5$ GeV/c) found in the event.

4.3.2 Analysis setup

As a first step, we limit ourself to the background consisting of (explicit) $b\bar{b}$ pairs. A jet cone aperture $R = 0.6$ is used, opened around the direction of the b quark. A study of the analysis with a realistic jet finding, tagging, and including trigger efficiency will be discussed in Section 4.4.

We set up a cut-based analysis in terms of the following discriminating variables:

- *Hard lepton*: a prompt lepton with $p_T > 20$ GeV/c is required. The lepton is requested to be compatible with the primary vertex, with an impact parameter below 30 μm (see Figure 3.4);
- *Isolation*: the lepton is required to be distant from both b-jets. As illustrated in Figure 4.3(left) for $HW \to b\bar{b}\ell$ events, the typical value required for isolation is $\Delta R(j, \ell) > 0.75$. This value discards the leptons coming from b-hadrons decays (see also Figure 3.10(right));
- *Jets*: the two jets are required to be separated by at least $\Delta R(j, j) > 1.2 = 2 \times R_{\text{cone}}$ to preclude overlaps. The invariant mass of each jet has to be compatible with a typical b-jet mass[4]: jets with $M(J) < 3$ GeV/c^2 are

[4]This cut is implemented as a quick and raw selection at an early stage to get rid of badly reconstructed jet candidates for further processing. This prevents as well τ contamination.

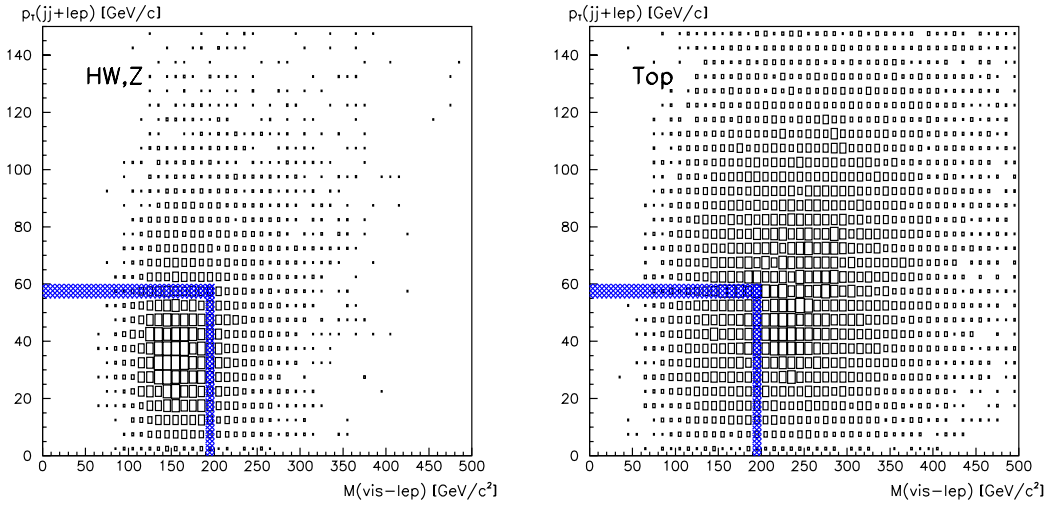


Figure 4.4: Distribution of the variables assessing extra activity for Higgs events and top pairs (see text). The bands delimit the region where the Higgs signal significance is the highest.

discarded;

- *Extra activity*: the events with top pairs have more activity (multiplicity, energy) in the detector compared to the Higgs signal. Cuts are performed on the visible mass ($M_{v\ell} = M(\text{vis}-\ell)$) and on the transverse momentum of the tagging products ($p_T(\text{jet}+\text{jet}+\ell)$). Figure 4.4 shows the distribution of these two discriminant variables for HW events and top pairs. The squares are the regions selected by the typical cuts $M(\text{vis}-\ell) < 200$ GeV/c^2 and $p_T(j+j+\ell) < 60$ GeV/c;

- A veto on a second isolated lepton can be applied to reduce the Z^0 background. Its effect is shown in Table 4.3.

According to Table 4.2, the main concern in the present case is to reject the QCD $q\bar{q}$ pairs, especially the 10^{12} $b\bar{b}$ pairs per year (paradoxically, the source for the signal for the CP studies).

The demand for an associated lepton is sufficient to completely kill the QCD background. With the latest tuning of the event generator proposed by the collaboration [67], the cross section of QCD $b\bar{b}$ pairs[5] forming an invariant mass

[5]In that scheme, a selection of the $b\bar{b}$ pairs is performed among the minimum bias events. To date, the model includes three mechanisms which contribute to the b total cross section: flavour creation, flavour excitation and gluon splitting. The 10^9 sample used in the present study only contains flavour creation, which gives the hardest spectrum.

Mass window $80 > M_{jj} > 120$	$bb\ell$/ year	$\Delta R(j,j)$ > 1.2	$\Delta R(j_b,\ell)$ > 0.75	$p_{T,\ell}$ > 20	$M_{v\ell} < 200$ $p_{T,jj\ell} < 60$	$2^{\text{nd}}\ell$ veto
Signal, $m_H = 115$ $H^0W^\pm, Z^0$	29	28	26	23	14	13.5
Background $b\bar{b}$	$\mathcal{O}(10^6)$			$\mathcal{O}(1)$	$\mathcal{O}(1)$	$\mathcal{O}(1)$
$t\bar{t}$ $(WbW\bar{b})$	2170	2050	1560	1260	210	200
$Z^0W^\pm, Z^0$	57	55	47	39	25	23
W+jets	123	118	66	55	36	36
γ^*/Z+jets	256	251	205	188	125	100
Total *background*	$\mathcal{O}(10^6)$			~ 1543	~ 397	~ 360
$S/\sqrt{B}$	$\mathcal{O}(10^{-2})$			0.59	0.70	0.71

Table 4.3: Number of signal and background events in the mass window $80 < M(jj) < 120$ for 1 LHCb year (2 fb^{-1}) after analysis. Mass and momentum are in GeV/c$^{(2)}$. Each column (new cut) cumulates the effect of the previous ones. Both jets must have an invariant mass $M_j > 3$ GeV/c^2 and, unless otherwise specified, the lepton must have $p_T > 5$ GeV/c and an impact parameter $IP < 30\,\mu$m. The analysis for the $b\bar{b}$ background has been performed in one step.

$M(b\bar{b}) > 80$ GeV/c^2 is estimated to be $\sigma \sim 250 \cdot 10^3$ pb giving $5 \cdot 10^8$ events per year. A sample of 10^9 such $b\bar{b}$ pairs (matrix element calculation with massive quarks) has been generated (on the MAP[68] facility, Liverpool) and is rejected in applying the above mentioned cuts by a factor $R \sim 10^8$ in the mass window $80 < M(jj) < 120$ GeV/c^2.

Table 4.3 gives the expected signal and background contributions in the mass window $80 < M(jj) < 120$ GeV/c^2 after application of the above cuts.

The dijet invariant mass distributions for signal and backgrounds are shown in Figure 4.5 with all the cuts applied (last column of Table 4.3). The contribution of the generic QCD $b\bar{b}$ pairs is neglected. The status is the following:

- Given the high power of the high-p_T and isolation cuts on the associated lepton, the $b\bar{b}$ pairs are a negligible background;
- The dominant contribution comes from the top pairs despite of a strong reduction by the veto on extra activity;
- The production of vector boson pairs is an irreducible background. Given the achievable dijet mass resolution, the signal would appear as a shoulder on the high side of the Z^0 peak;
- According to Table 4.2 the integrated contribution of $b\bar{b}W(\ell\nu_\ell)$ is ~ 1.5 times higher than $b\bar{b}Z^0/\gamma^*(\ell^+\ell^-)$, but significantly different mass distri-

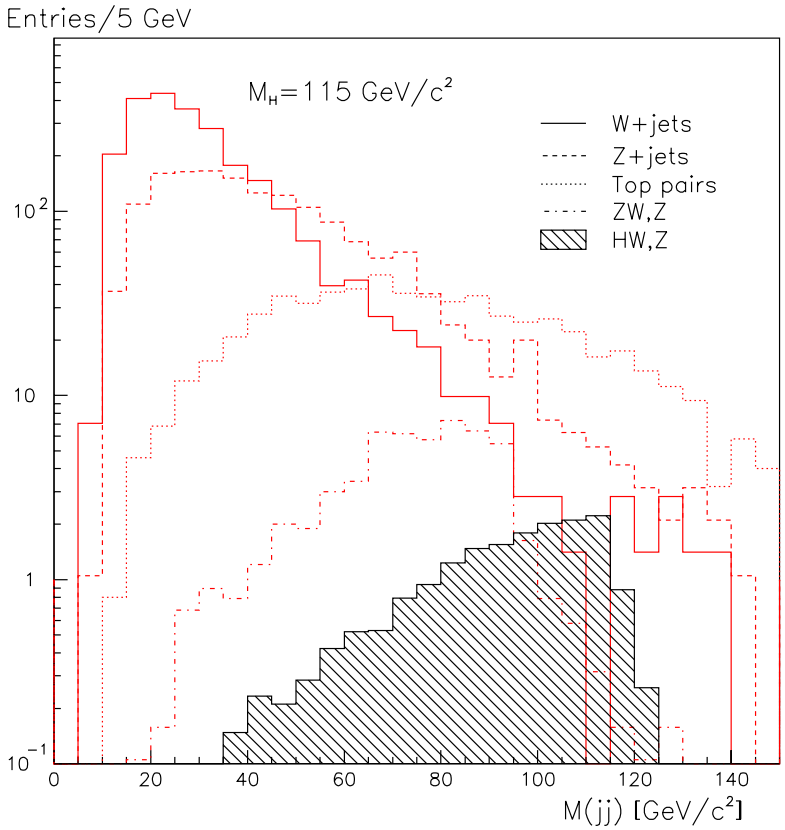


Figure 4.5: Signal and background distributions after analysis (see text) with a normalization to one LHCb year (2 fb^{-1}). The background contribution coming from the generic QCD $b\bar{b}$ pairs has been neglected. Appendix D gives the size of the original generated samples.

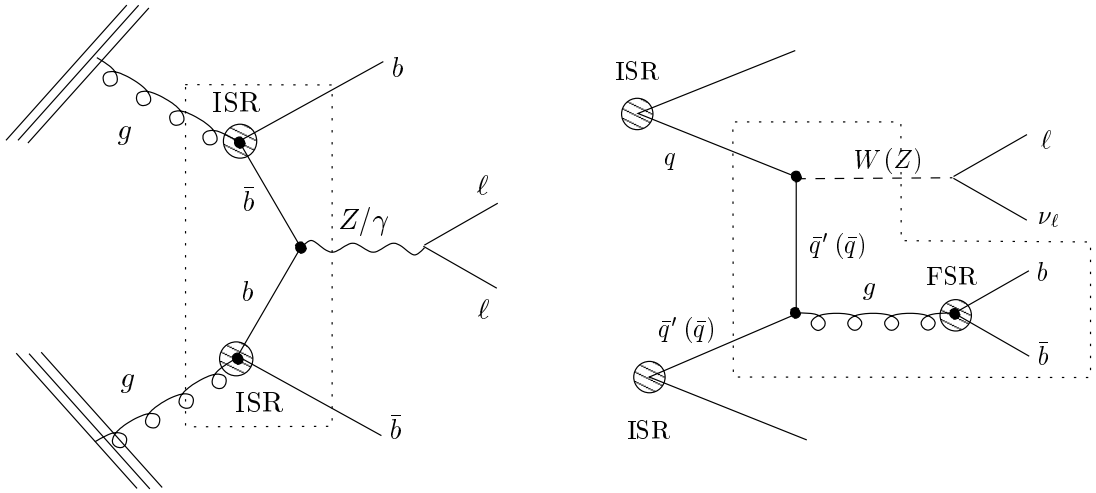


Figure 4.6: Production of a $b\bar{b}$ pair in the parton showers (hatched blobs) as higher order diagrams for Z (left,right) and W (right) production as depicted in Pythia. Distinction is made between initial and final state radiation (respectively ISR,FSR).

butions make the latter more active in the selected mass window. The explanation resides in the production mechanisms of the $b\bar{b}$ pairs in the parton shower, as illustrated in Figure 4.6. For the $W^{\pm}$ case, the pairs are produced as a pure QCD correction diagram to single W production (right, boxed). On the other hand, the bulk of the Z^0 production will proceed mainly via $b\bar{b}$-fusion (left, boxed), leaving an extra $b\bar{b}$ pair with a p_T spectrum significantly harder than in gluon splitting.

Note that Pythia predictions have been compared with the analytical calculation performed with CompHEP [69]. In both cases, a ratio ($b\bar{b}$-fusion : gluon splitting) $\sim (9:1)$ in the Z^0 production has been observed.

In principle, further discrimination could be obtained by rejecting events with a second (prompt) isolated lepton, which is not very effective in LHCb given the limited angular coverage. Only $\sim 20\%$ are rejected (see Table 4.3).

- The channel $HZ \to b\bar{b}\ell^+\ell^-$ has been analysed separately. In addition to the above cuts, a second isolated prompt lepton with opposite charge *is* required now to form the Z^0 mass (i.e. anti-veto). Only events with a dilepton mass $82 < M(\ell^+\ell^-) < 100$ GeV/c^2 are selected. The dominant background will come from the generic $Z/\gamma^* \to \ell^+\ell^-$. For one year, the average number of events in the mass window $80 < M(jj) < 120$ GeV/c^2 is ~ 0.5 for the signal. The channel Z/γ^* gives ~ 20 background events, which is in agreement with the reduction produced by the veto on a second

Description/Source [pb]	W+jets	Top pairs
① $\sigma(jj+\ell), \eta_{\rm LHCb}$	4000 (5.9)	23 (9.7)
② jet finder (JF)	107	9.15
③ standard cuts	45 (1.1)	0.84 (0.32)
④ $b-$tag: $N_{\rm min}=3$	3.4 (0.55)	0.36 (0.24)

Table 4.4: Overview of the light flavour reduction in W inclusive production and top pairs. The first line gives the cross section corrected in acceptance for two partons and a lepton with each $p_T > 5$ GeV/c. In parentheses is the value when the two partons are b quarks. Line 2 is the cross section for two reconstructed jets with $p_T > 5$ GeV/c. The analysis and a crude $b-$tagging are applied in lines 3 and 4. The typical statistical uncertainty is $\pm 5\%$.

lepton in Table 4.3. The resulting signal significance is marginal.

4.4 Towards b-tagging and trigger

We show now some results when the jet finder is used for the determination of the jet axis, as described in Section 3.4.3. A possible scenario for a trigger scheme in LHCb, to select the Higgs events with high efficiency will also be presented in the second part.

4.4.1 Jet b-tagging

The results shown in Tables 4.2 and 4.3 were obtained with an explicit identification of the b-jets, i.e. with an efficiency and purity equal to one.

We want to assess the rejection of the light flavour jets using the jet algorithm described in Section 3.4.3. The algorithm was designed for the maximum efficiency to reconstruct $H \to b\bar{b}$ in LHCb (Table 3.2), but without considering the jet tagging as a goal.

Table 4.4 shows an indication on the level of rejection of the light flavour jets with the present tuning of the seed finder algorithm. Two background sources have been studied, the top pairs and the inclusive $W^{\pm}$ production, which is potentially the most dangerous through the parton showering process. The number of real $b\bar{b}$ in the sample is given in parenthesis. An attempt for a stronger selection of the b-jets consists in requiring a minimal number $N_{\rm min}$ of tracks inside the jet, with impact parameter larger than a value $\mathrm{IP}_{\rm min}$. The quoted values are obtained with $N_{\rm min}=3$ and $\mathrm{IP}_{\rm min}=30\mu$m.

The two last rows of Table 4.4 correspond to the distributions illustrated in Figure 4.7 (respectively solid and dashed histograms) for inclusive W events and

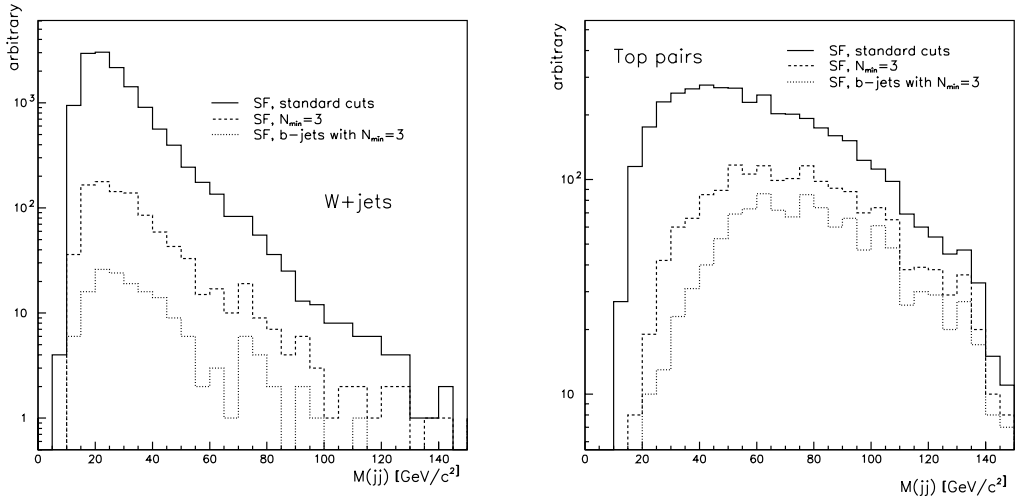


Figure 4.7: Reduction of light flavour jets background in W inclusive (left) and top pairs (right) production. A crude b-tagging consists in rejecting events containing less than 3 tracks per jet with impact parameter IP $> 30\,\mu$m .

top pairs. We have also represented the distribution (dotted histogram) for the real $b-$jets with $N_{\min} = 3$ (values in parentheses on line 4).

We can see a reduction of the light flavour jets by a factor ~ 1400 in the W inclusive sample after analysis and b-tagging. However, a contamination of around 3 pb is still present because of the large production cross section. That contamination typically comes from D mesons and neutral kaons originating in the partonic shower. For the top pairs, the two jets total cross section is smaller, but the fraction giving two b-jets is large: $\sim 40\%$, to be compared to $\sim 0.15\%$ in the W sample. Furthermore, the $b\bar{b}$ pairs in the top events are part of the hard process, and the specifity of the seed finder to pick up the hardest jets in the event makes their identification efficient. Note that the reconstruction efficiency for two "tagged" b-jets in top pairs ($\frac{0.24}{0.32}$) is similar to the values we obtain for Higgs events.

A more evolved analysis should make use of the full LHCb potential (particle identification, in particular, has never been considered in the present work). However, the capability of LHCb to reconstruct and identify exclusive B-hadrons final states let presume ultimately of a high power for b-jet tagging.

4.4.2 Higgs trigger

Section 3.4.3 showed a potentially high efficiency to reconstruct Higgs events in LHCb. The $b\bar{b}\ell$ signature satisfies the standard trigger setup of LHCb for CP

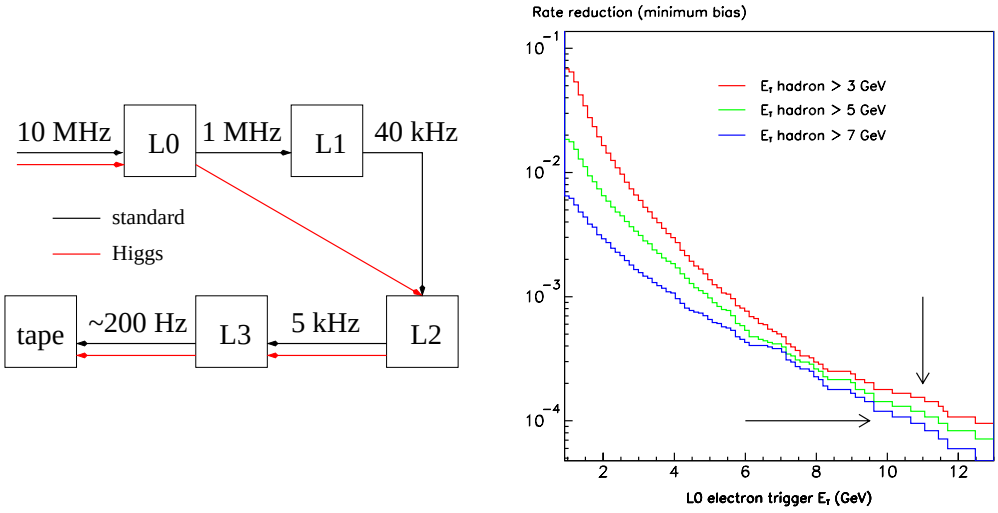


Figure 4.8: Level-0 electron trigger efficiency on minimum bias events versus the required transverse energy of the electron. The curves correspond to different E_T threshold values for a simultaneous hadron. The arrows suggest a working point (rate reduction by a factor 10^4) which would allow to skip L1 without saturating the L2 bandwidth.

physics described in Section 2.3.8. In full simulation, HW events fire the L0 logic with a probability ~ 1. However, the same events pass the L1 with a probability of 40%, similar to the performance on the QCD generic $b\bar{b}$ for the CP physics [3],[8]. Such a reduction is not affordable for the Higgs study, with low statistics. We show here an alternative trigger scheme to accomodate the Higgs events while remaining compatible with the bandwidth downstream.

Given the hardness of the signal, the L0 conditions can be strengthened to reach enough reduction on the minimum bias events which would allow to skip the L1 trigger with no bandwidth deterioration (see Section 2.3.8 and Figure 2.11). This scheme is illustrated in Figure 4.8(left).

A possible scenario for the extra requirements on the electron trigger L0 (calorimetry) is shown in Figure 4.8(right). Combining thresholds on the transverse energy for both an electron and a hadron ($E_{T,\mathrm{electron}} \gtrsim 10$ GeV and $E_{T,\mathrm{hadron}} \gtrsim 5$ GeV) in the event allows a rate reduction $\mathcal{O}(10^4)$ on minimum bias events to be reached, which corresponds to a change in the L2 bandwidth by only 2%.

In comparison to the standard setup the proposed settings increase the efficiency on Higgs events up to 85%.

Isolation of the electron cluster with respect to the hadron cluster has not been asked at this level and might contribute to enhance the Higgs efficiency.

A corresponding scenario for the level-0 muon trigger has not been assessed yet, but is expected to give at least equivalent performances.

Though not definitive, the above mentioned scenario exploits the flexible design of the LHCb apparatus to make it adequate to measure hard events.

4.5 Outlook

This section reviews some observations brought by the investigation of the characteristics of the top pairs background. These may be taken as ideas for further studies. We also mention a few general remarks related to a potential enhancement of the signal significance.

- The present analysis is cut-based. A multi-dimensional optimization with a neural network or a likelihood-based approach to combine the discriminating variables remain to be assessed. However, a preliminary attempt for the optimization of the present set of cuts with a neural network indicated a factor ~ 1.3 enhancement on $S/\sqrt{B(\text{top})}$, treating the top pairs background individually. This has to be compared to $S/\sqrt{B(\text{top})} \sim 1$ in Table 4.3. On the same basis, a real strategy remains to be developed for tagging the (b-)jets.

- The study has been done with LO estimates of the cross sections. The use of $K-$factors will raise the significance of the signal. For instance, the NNLO cross sections at LHC for $W^{\pm}, Z^0$ in [36] and the NLO cross section for $HW^{\pm}$ in Table 4.1 both indicate a typical $K \sim 1.2$ to be introduced.

- The key strategy against the top pairs has been found in rejecting events with high visible activity inside the detector, apart from the two jets and the isolated lepton. As a consequence, most of the remaining background events are expected to have their energy flowing immediately outside of the detector acceptance.

  A toy study has been done at four-vector level to compare the activity *outside* the LHCb acceptance for Higgs events and top pairs. The results are shown in Figure 4.9. The standard cuts have been applied. The plots (a) and (b) show the mean $dE_T/d\eta$ for charged pions and photons (with $p_T > 1$ GeV/c), for Higgs and top events as a function of the pseudorapidity. We defined the transverse energy $E_T^2 = (\sum p_x)^2 + (\sum p_y)^2$ for the particles within the selected slice in pseudorapidity.

  The uppermost value on the abcissa corresponds to the limit of the LHCb detector. For both $\pi^{\pm}$ and γ, the top pairs yield an average transverse energy flow significantly more important than the Higgs events in the next unit of pseudo-rapidity outside the detector (dashed line). The signal shows to be largely contained in the acceptance, as requested by the standard analysis.

Plots (c) and (d) show the ΔE_T sum of respectively $\pi^{\pm}$ and γ in the next unit of pseudo-rapidity for Higgs and top events. As an illustration, rejecting the events with pions giving a $\Delta E_T > 10$ GeV would keep virtually all the signal but reject half of the top pairs. Plots (e) and (f), in the same configuration, show the distribution in multiplicity of the considered particles.

This extra unit of pseudo-rapidity is covered by the shield in front of the magnet (see Figure 2.7, to cut the stray field upstream in RICH1) and one may imagine to instrument it (~ 25 cm thick Fe $\simeq 1.5\lambda_I, 14X_0$, and an area of $\sim 4 \times 9$ m^2) with a gross padding.

Ideally, an extra coverage would be around the VELO. However, the tight occupancy around it (cabling route for the off-detector electronics and cooling) makes this scenario very unprobable.

- It has been shown that Higgs events in association with a lepton are fairly well contained in the transverse plane. In the case of W Bremsstrahlung, the transverse momentum of the neutrino may indeed be estimated by

$$\vec{p}_{T,\text{vis}} = \sum_{i \in \eta_{\text{LHCb}}} \vec{p}_{T,i} \approx -\vec{p}_{T,\nu}$$

as shown at four-vector level on Figure 4.10(a). The sum is taken over all the visible particles in the LHCb geometrical acceptance, and $\vec{p}_{T,\nu}$ is the transverse momentum of all the neutrinos found in the event.

From the relation above, the W transverse mass can be approximated by

$$m^2_{T,W} = (|\vec{p}_{T,\ell}| + |\vec{p}_{T,\nu}|)^2 - (\vec{p}_{T,\ell} + \vec{p}_{T,\nu})^2 \quad \text{with } \vec{p}_{T,\nu} \equiv -\vec{p}_{T,\text{vis}}$$

using the prompt isolated lepton ℓ selected in the analysis. The resulting m_T mass spectrum shown on Figure 4.10(b) provides a strong indication on the presence of a W boson in the event.

The situation is different for top pairs for which it has been shown in the previous point that a significant part of the total momentum flows outside the acceptance. The correlation between the total visible transverse momemtum $\vec{p}_{T,\text{vis}}$ and the transverse momentum of the neutrinos $\vec{p}_{T,\nu}$ is shown on Figure 4.10(c) and the resulting W transverse mass on Figure 4.10(d).

The transverse mass $m_{T,W}$ would clearly provide additional discriminating power between Higgs events and top pairs when combined in the multi-parametric optimizations cited above.

Furthermore, the $m_{T,W}$ spectrum would provide a check that the lepton blindly selected in the event is indeed the correct one.

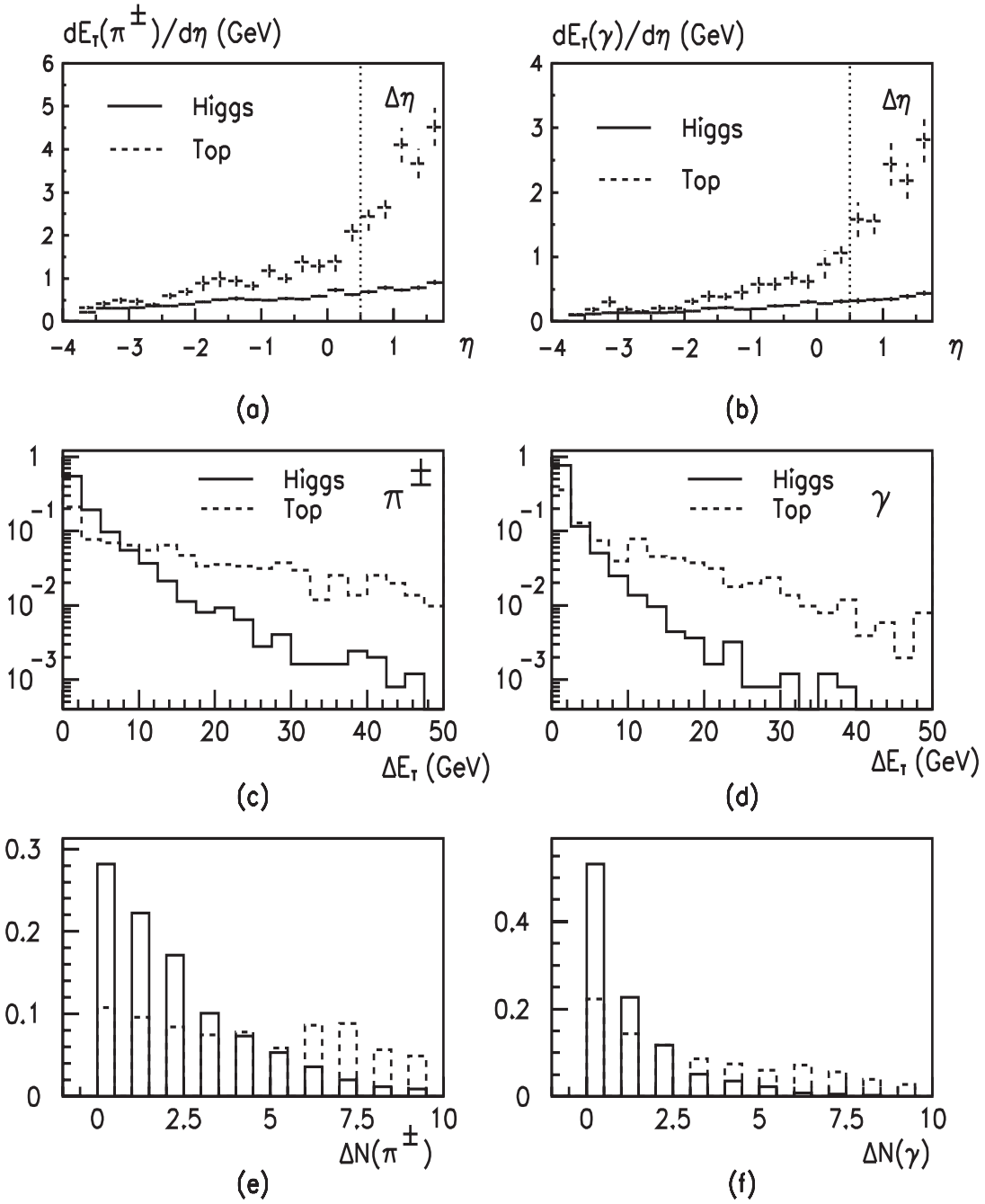


Figure 4.9: Comparison of the multiplicity and transverse energy yields for an extra coverage $\Delta\eta \sim -1$ for Higgs signal and top pairs background. Four-vector level study. The first column shows, for charged pions with $p_T > 1$ GeV/c, (a) the mean transverse energy profile in pseudorapidity, (c) the transverse energy distribution in the slice $\Delta\eta$ and (e) the multiplicity in the slice $\Delta\eta$. The second column gives the same plots for photons with $p_T > 1$ GeV/c (respectively b,d,f).

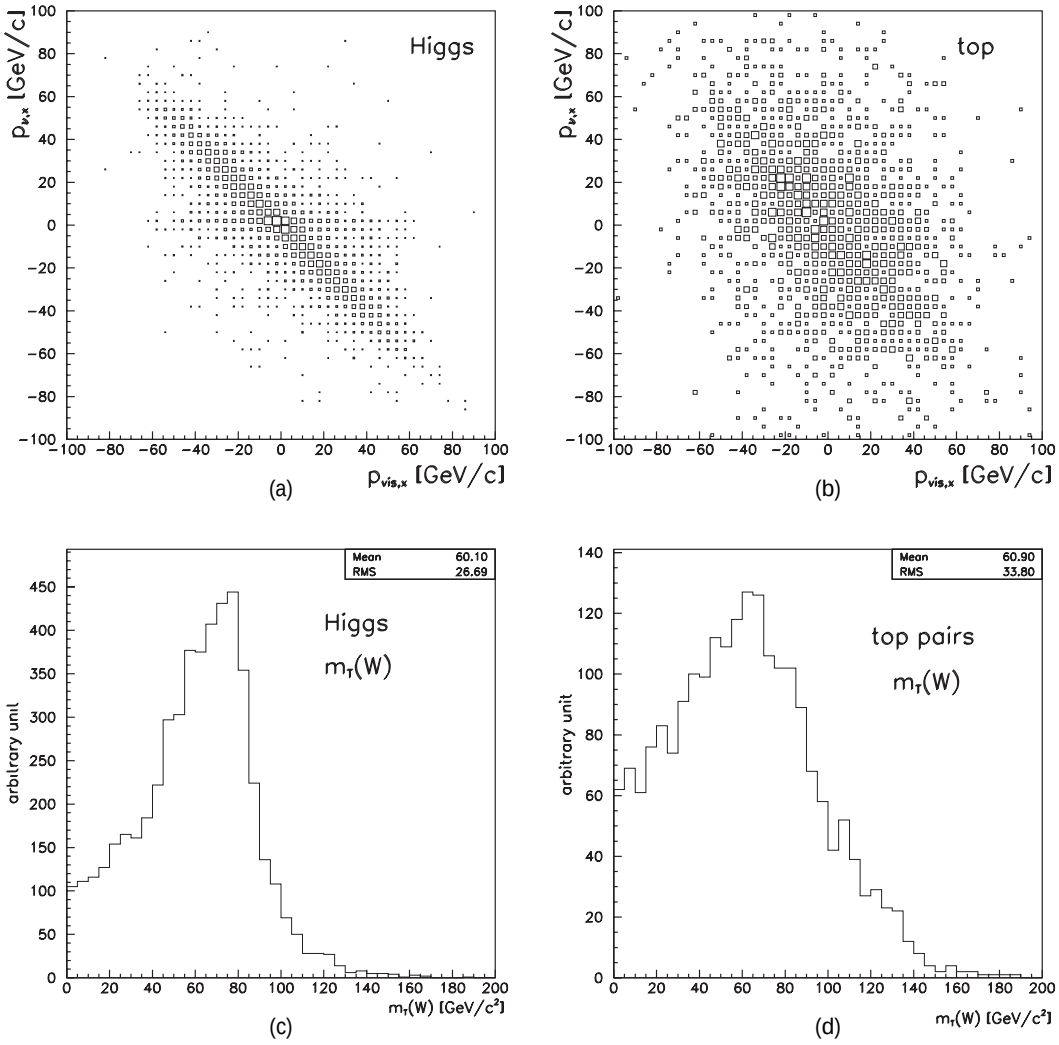


Figure 4.10: Correlation between the total visible transverse momemtum $\vec{p}_{T,\mathrm{vis}}$ and the transverse momentum of the neutrinos $\vec{p}_{T,\nu}$ for (a) Higgs events and (b) top pairs. Transverse mass of the presumed $W \to \ell\nu_\ell$ in the event for (c) Higgs events and (d) top pairs (see text).

- The colour flow in hadronization is different for the Higgs decay products and the QCD generic $q\bar{q}$ pairs. The Higgs being a colour singlet, the string is exclusively formed between the two b quarks. Conversely, the $q\bar{q}$ pairs will reconnect to the beam remnants. Some difference in the jet structure may be expected between the two configurations. The K_T jet algorithm would be the choice to analyse it.

- Scenarios are being discussed in the collaboration for running at a higher luminosity while remaining within acceptable range for ageing and occupancy of the detectors. This would certainly bring a significant improvement of the signal significance, but with new pileup conditions potentially affecting both the jet energy resolution and b-tagging performance.

- The identification efficiency of the tagging lepton is a crucial point. For instance, we never requested an *absolute* isolation. This could discard mini-jets with large eletromagnetic energy, faking an electron. Together with the tagging of the b-jets, this would be the next steps in the continuation of this analysis. Meaningful studies would imply at this stage a general use of the full detector simulation. However, dedicated computing resources would be probably required in order to generate large samples of background.

Conclusion

The motivation was to assess the potential of LHCb to observe a SM Higgs signal. Because of the high longitudinal boosts encountered in the pp collisions at LHC, a significant fraction $\sim 30\%$ of light Higgs ($m_H = 115\,\mathrm{GeV/c^2}$) are produced in the LHCb acceptance $1.8 < \eta < 4.9$.

Given a low running luminosity of 2×10^{32} $\mathrm{cm^{-2}s^{-1}}$— to be compared to the nominal 10^{32} $\mathrm{cm^{-2}s^{-1}}$ at LHC — and a limited acceptance, we showed that the only channel worth to be studied is $HW, Z \to b\bar{b}\ell$ for Higgs masses in the range 100–130 $\mathrm{GeV/c^2}$.

The recent results obtained at LEP give a hint of a SM Higgs boson with a mass $m_H = 115.0^{+1.3}_{-0.9}$ with a significance $\sigma = 2.9$.

The physics runs at LHC should start in 2006, with a low luminosity. The design luminosity should be steadily reached after three years. At that time only, the central detectors ATLAS and CMS will have the opportunity to fully expoit their potential for Higgs in the low mass region. However, LHCb will be operational at its design luminosity from day one.

These facts place LHCb on the line for the observation of the SM Higgs.

This work pioneered a setup for the production and the analysis of hard jets in the LHCb detector, in the forward region.

We demonstrated in the full detector simulation that the LHCb baseline design allows to efficiently identify, reconstruct and trigger the b-jets coming from the Higgs.

Due to the heavy task of background generation in full simulation, a *fast* simulation, including the relevant detector effects, has been developped to perform the analysis.

The analysis requires a hard lepton isolated from the b-jets to reject the QCD background. We assessed it on a sample of 10^9 $b\bar{b}$ generated on a computer farm, and observed a rejection of the events by a factor $\sim 10^8$ in the mass window $80 < M(jj) < 120$ $\mathrm{GeV/c^2}$, leaving hope for further studies.

At this stage, the cut-based study indicates a significance $S/\sqrt{B} \sim 0.7$ for an integrated luminosity of 2 $\mathrm{fb^{-1}}$(one LHCb-year) assuming a perfect identification of the signal and backgrounds.

The next steps in the analysis to be performed would be to setup a real analysis for b-jets tagging, and evaluate the efficiency of identification of the associated

lepton.

The significance is modest, however, potentially large improvements must not be neglected when performing a combined optimization on the discriminant variables.

Due to its particular requirements for the background generation, this study triggered the creation of a "generator working group" in the LHCb collaboration. The group is presently in charge of all the issues related to the physics at generator level.

As mentioned, this study has been catalyzed by the recent hint of a light mass Higgs at LEP and echoed with an involvement of other groups (Bologna, Milano, Montpellier) to investigate the field more deeply in the frame of the LHCb collaboration.

The general feeling is that this work boostrapped a delicate analysis that deserves to be continued.

Appendix A

Total cross sections at LHC

This appendix gathers information scattered in various (more or less historical) papers. We found useful to put them together in a coherent manner.

With some assumptions and adequate parameterizations we will briefly show the current way to estimate inclusive hadronic cross sections, and in particular their asymptotic behaviour at high energies.

Much details will be skipped to focus on the basic and factual issues which lead to the reference values taken for LHC, which are commonly used to cross-evaluate the Monte Carlo predictions.

We propose the following decomposition of the total hadronic cross section:

$$\sigma_{\text{tot}} = \sigma_{\text{inelastic}} + \sigma_{\text{elastic}} \tag{A.1}$$

$$\sigma_{\text{inelastic}} = \sigma_{\text{min.bias}} + \sigma_{\text{diffractive}} \tag{A.2}$$

each term of which will be discussed in the subsequent sections.

A.1 Overview of the Regge theory

In the late fifties Regge showed that, for technical aspects of the description of particle scattering by a potential, it is useful to regard the angular momentum ℓ as a complex variable [70]. For example, the singularities which occur in the plane given by analytic continuation of the angular momentum to complex values are related to the asymptotic behaviour of the scattering amplitude. Roughly speaking, the amplitudes corresponding to the exchange of particles in a standard perturbation theory, A_{pert}, will look like A_{Regge} in the Regge approach,

$$A_{\text{pert}} \sim \frac{g(q^2)}{q^2 - m^2 + \imath\epsilon}, \qquad A_{\text{Regge}} \sim \frac{g'(t)}{\ell - \alpha(t)},$$

where $t = (p_1 - p_3)^2$ is the momentum transfer in a $1 + 2 \to 3 + 4$ scattering. These (moving) poles in the angular momentum plane give rise to composite

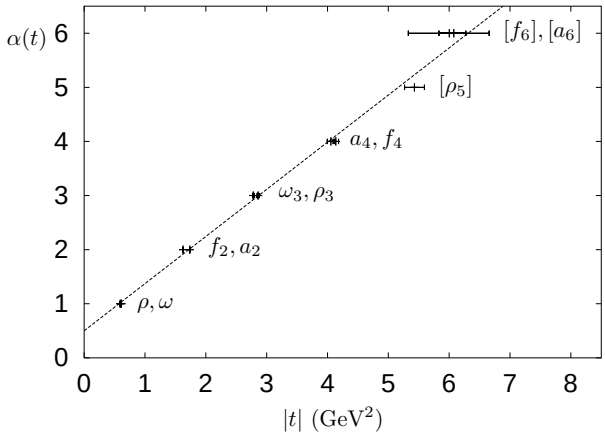


Figure A.1: The meson spectroscopy shows a clear linear relation between the spin J of the particles and their mass M_J^2 giving a Regge trajectory. See relation A.3.

particles (or resonances), so that the asymptotic behaviour of a scattering amplitude is determined by the particles which can be exchanged. The so-called Regge-trajectories $\alpha(t)$ consist in the whole set of particles — the Reggeon — giving the poles $\ell \to \alpha(t)$. Rather than considering the exchange of individual particles[1], Regge considers the exchange of a complete trajectory of particles with the benefit that divergences, found in the treatment of individual particles with spin, mutually cancel. The hadron spectroscopy reveals a simple parameterization of the trajectories:

$$\alpha(t) = \alpha_0 + \alpha' t + \imath|\text{small}|$$

and the set of particles whose spin $J = Re[\alpha(0)]$ and mass $M_J^2 = t$ obey the relation

$$J = \alpha_0 + \alpha' M_J^2 \tag{A.3}$$

is said to lie on a Regge trajectory with a Regge intercept α_0 and a Regge slope α'. Figure A.1 shows the trajectory of the ρ, ω, f, a mesons, illustrating the relation A.3. The particles between brackets are not yet clearly identified. The theory then predicts that the high energy behaviour of the scattering amplitude $A(s,t)$ will be

$$|A(s,t)|_{\text{Regge}} \xrightarrow{s\to\infty} s^{\alpha(t)}$$

where $s = (p_1 + p_2)^2$ (centre-of-mass energy) and $t = (p_1 - p_3)^2 = Q^2$ (momentum transfer) are the standard Mandelstam variables.

[1]Likewise to the Yukawa's hypothesis that the long-range part of the inter-nucleon force is due to the exchange of π mesons.

A.2 Parameterization of σ_{tot}

The Regge theory leads to a good description of high-energy, low-$|t|$ experimental data[2], where QCD is not perturbative any more since $\alpha_s \sim \mathcal{O}(1)$. Moreover, pQCD cannot apply with large parton densities — even if partons are at short distances where the coupling is small — typically at high energies.

Within this framework, the behaviour of the experimental data shown on Figure A.2 suggests to use the following *phenomenological* Regge description of the total cross section:

$$\sigma_{\mathrm{tot}}^{pp} = X^{pp} s^{\epsilon} + Y^{pp} s^{-\eta} \tag{A.4}$$

The *assumption* made is that the rise of cross sections with s in the first term is the result of the presence of a particular non-resonant Reggeon trajectory called the (soft) *Pomeron*. It is widely agreed that if there are particles on the Pomeron trajectory they are glueballs (bound state of gluons), for which strong candidates exist [72],[73].

The second term arises from meson trajectory exchange (ρ, ω, f, a ; leading trajectory) and is important only at low energies.

The success of this approach resides in the fact that the overall data of hadron scattering ($p\pi, pp, p\bar{p}, Kp, \gamma p$) are simultaneously well fitted with the same power parameters [74]

$$\epsilon \approx 0.0808 \qquad \eta \approx 0.4525$$

indicating some kind of universality of the trajectories, with the X, Y amplitudes being the only process dependent parameters. For example, the fit on pp data shown in Figure A.2 gives [71]

$$X^{pp} \sim 21.7 \qquad \text{and} \qquad Y^{pp} \sim 56.1 \tag{A.5}$$

From rather low energies onwards, total cross sections are therefore predicted to rise as a single power of the c.m. energy $\sigma_{\mathrm{tot}} \propto s^{\epsilon}$, which in the asymptotic limit would violate the so-called Froissart bound which imposes an increase at most as $(\ln s)^2$ to preserve unitarity. But with the smallness of ϵ the bound would only be reached at energies of 10^{23} GeV, and one might therefore expect reasonable estimates of total cross sections at LHC energies.

It is clear that smooth functions cannot describe the resonance region; hence the fits can be trusted only above a certain energy $\sqrt{s_{\mathrm{min}}}$, indicated in Figure A.2 with a vertical arrow.

The point at $\sqrt{s} = 30$ TeV on the pp data is the result of the measurement of atmospheric air showers giving the total cross section $\sigma_{p\text{-air}}$ [75]. Although the extraction of σ_{pp} from $\sigma_{p\text{-air}}$ is subject to large uncertainties (because of model

[2] One speaks also of soft hadronic processes, or low-p_T processes, i.e. processes which occur at long distances.

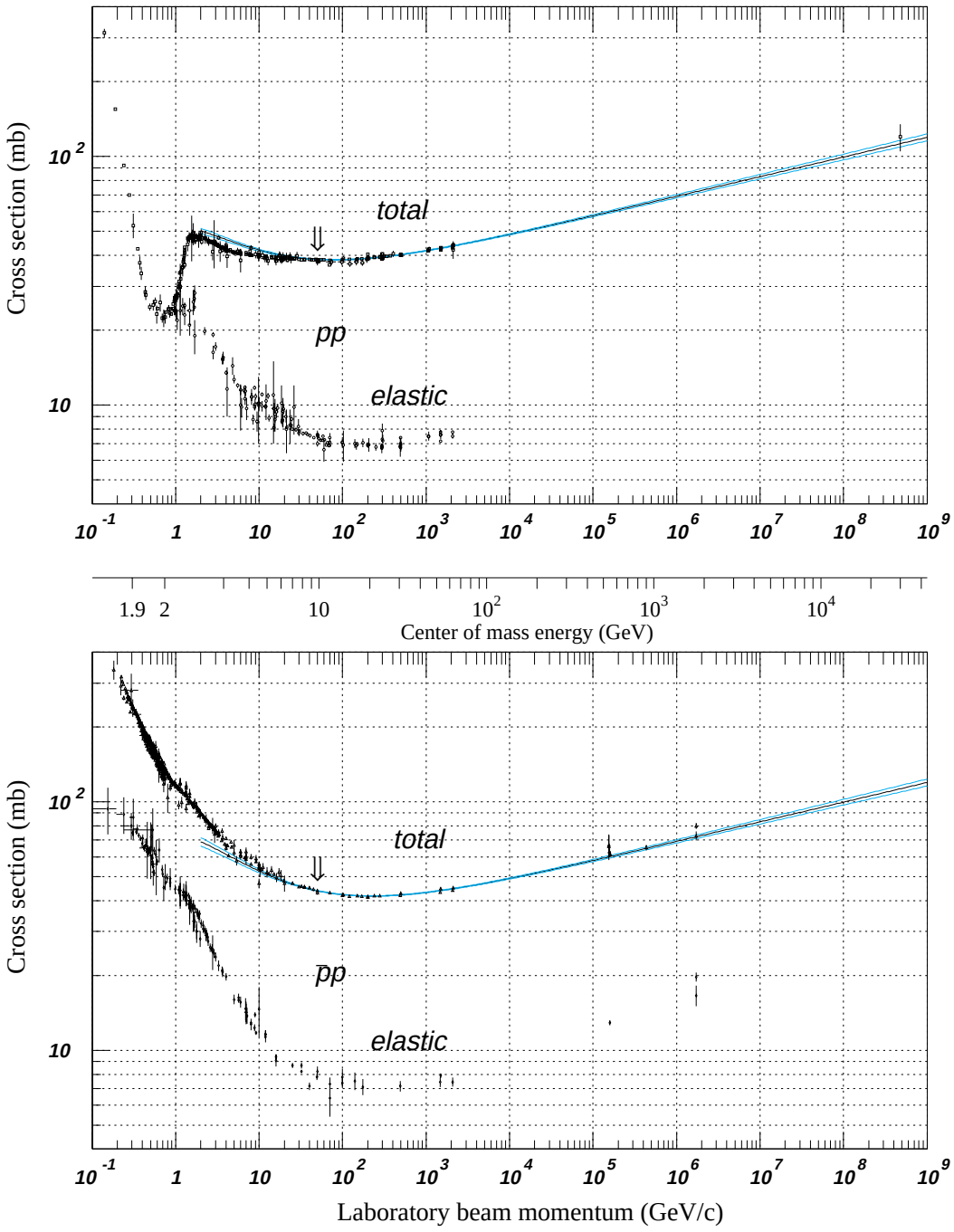


Figure A.2: Total and elastic cross sections for pp and $p\bar{p}$ collisions as a function of laboratory beam momentum and total centre-of-mass energy [71] (Courtesy of the COMPAS Group, IHEP, Protvino, Russia, 1996).

dependencies), the point is compatible with the fit despite of the important lever-arm.

It has to be noticed that the Pomeron has to carry the (internal) quantum numbers of the vacuum, since the scattering amplitudes, and hence the singularities (poles), for a diagram $1 + 2 \to 3 + 4$ remain the same under CPT and/or leg-crossing. As consequence, total cross sections should scale with energy to identical values if a particle is replaced by its antiparticle, $\sigma_{\rm tot}(12) = \sigma_{\rm tot}(\bar{1}2)$ for large s, since the Pomeron would experience the same couplings. This result is known as the *Pomeranchuk theorem* [76], and is verified so far by the data of Figure A.2.

A.3 Determination of $\sigma_{\rm elastic}$

In full generality the optical theorem relates the elastic and total cross sections according to

$$\left.\frac{d\sigma_{\rm elastic}}{dt}\right|_{t=0} = \frac{1+\rho^2}{16\pi}\sigma_{\rm tot}^2 \tag{A.6}$$

At not too large momentum transfers t, where the bulk of the (Rutherford-like) cross section is, the elastic cross section can be approximated by a simple exponential fall-off (form factor of the particle which couples to the Pomeron). If one assumes accordingly to the data a vanishing ratio of the real to imaginary part of the forward scattering amplitude, $\rho \sim 0$, one gets:

$$\frac{d\sigma_{\rm el}}{dt} = \frac{\sigma_{\rm tot}^2}{16\pi}\exp(B_{\rm el}t) \tag{A.7}$$

In the Pomeron exchange picture, the slope parameter $B_{\rm el}$ is predicted to be

$$B_{\rm el} = 2b_A + 2b_B + (4s^\epsilon - 4.2)$$

where b_X is the form factor for the particle $X = A, B$ and the last term is an effective way of including single and *multi*-pomeron exchange contributions to remain consistent with the Pumplin bound $\sigma_{\rm elastic}/\sigma_{\rm tot} \leq \frac{1}{2}$. The form factor is typically of the order of unity, e.g. for the proton $b_p = 2.3$.

The total elastic cross section is so:

$$\sigma_{\rm el} = \frac{\sigma_{\rm tot}^2}{16\pi B_{\rm el}}$$

and the ratio $\sigma_{\rm el}/\sigma_{\rm tot}$ becomes constant in the asymptotic limit in agreement with the $p\bar{p}$ data at high energies (see Figure A.2)

$$\frac{\sigma_{\rm el}}{\sigma_{\rm tot}} \xrightarrow{s\to\infty} \frac{\sigma_{\rm tot}}{64\pi s^\epsilon}\cdot(2.58\ {\rm GeV}^{-2}{\rm mb}^{-1}) = \frac{2.58\,X^{pp}}{64\pi} \simeq 28\%$$

where natural units have been reintroduced in the intermediate term.

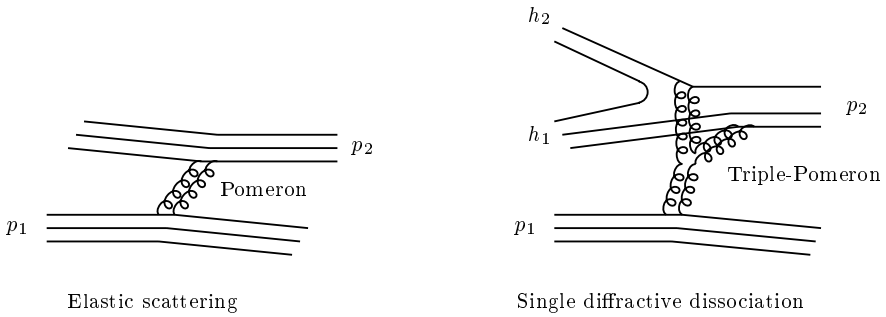


Figure A.3: Illustration of contributions to σ_{elastic} and $\sigma_{\text{diffractive}}$ in terms of Pomeron exchange viewed as gg glueballs.

A.4 Determination of $\sigma_{\text{inelastic}}$

The inelastic cross section can then be determined by subtraction $\sigma_{\text{inelastic}} = \sigma_{\text{tot}} - \sigma_{\text{elastic}}$ and according to the previous parameterizations should not produce an additionnal rise of the total cross section asymptotically.

Following the classification suggested in relation A.2, the basic distinction between inelastic diffractive events and inelastic non-diffractive events[3] is that there is no net colour flow for the latter and the events show a reduced multiplicity and large gaps in the rapidity distribution of the products. The fraction of inelastic cross section which can be attributed to diffractive dissociation is not certain because the separation of diffractive and non-diffractive components is not without ambiguities. As a fall-out, the term "minimum-bias" is often used as a generic term to refer to the full inelastic "low-p_T" production.

Nonetheless, the diffractive cross section can be modelled from the viewpoint of the Regge-theory by a triple-Pomeron exchange mechanism as shown in Figure A.3. It amounts to about 1/4 of the inelastic cross section.

The remaining non-diffractive inelastic interactions can be described with standard perturbative QCD under the assumption, as simplest model, of *multiple* parton-parton interactions within the incoming hadrons, with a proper regularization for $p_T \to 0$. The model reveals a similar asymptotical behaviour cross section like in the Reggeon exchange picture, though this would involve a more complex Pomeron structure (the "hard" Pomeron), basically a formal construct, about which very little is known presently.

Assuming the picture of the glueball, Figure A.3 illustrates, at the lower order, interactions in the Pomeron exchange picture.

[3]This type of events is commonly called *minimum-bias*, because that is what we get with an unselective trigger.

Quantitatively, the estimate of the contributions in relation A.2 we will consider as reference values for pp collisions at $\sqrt{s} = 14$ TeV are

	σ_{tot} [mb]	σ_{elastic} [mb]	$\sigma_{\text{inelastic}}$ [mb]
$pp, \sqrt{s} = 14$ TeV	102	22	80

in agreement with the most recent parameterizations [21],[77].

A.5 Epilogue

As of a *warning* regarding what has been mentioned in this appendix, it has to be stressed that there is no deep theoretical reason for the Pomeron hypothesis and it only appears to be the simplest attempt to comply with the experimental data on total cross sections. The common belief is that the "soft" Pomeron exchange gives the correct high energy asymptotic description of simultaneously the total cross section, the $d\sigma/dt$ behaviour of elastic scattering, and the phenomenology of diffractive dissociation. So, in short, the Pomeron is still unknown but needed to describe experimental data.

It has to be mentioned, however, that the Regge picture allows to explain — alternatively to the quark model of Gell-Mann — the particle spectrum of the hadrons discussed in Section 1.1. One can illustrate indeed the self-consistency of strong interactions under the hypothesis that Regge exchanges provide the binding forces between particles which in turn generate Regge trajectories: the so-called "boostrap" mechanism. All strongly interacting particles arise then as a consequence of just analyticity and unitarity requirements.

The bootstrapping character of the exchanges is however not incompatible with the quark picture: when the parameterization A.4 is used to fit $\sigma_{\text{tot}}^{\pi p}$ data, the amplitudes $X^{\pi p}$ and X^{pp} are found to be in a ratio $\sim 2/3$. Equivalently, $X^{Kp}/X^{\pi p} \sim 1$ suggests that Pomeron couples to single (constituent) quarks, showing then a "quark-counting" property.

Appendix B

Overview of LHCb physics

To account for the flavour changing charged currents between quark generations, the quark sector of 1.13 has to be more precisely written, recalling the SM Lagrangian 1.10

$$\mathcal{L}_{q\,\mathrm{mass}} = -(g_{1,f}\bar{Q}_L\phi V^{\mathrm{CKM}}D_R + g_{2,f}\bar{Q}_L\tilde{\phi}U_R + \mathrm{H.C.})$$

where V^{CKM} is the matrix relating the weak interaction eigenstates to the mass eigenstates $D_i' = \sum_j V_{ij}^{\mathrm{CKM}} D_j$. Such a mixing does not occur in the lepton sector of the SM for the neutrinos are degenerated to zero masses. The so-called Cabibbo-Kobayashi-Maskawa (CKM) matrix V^{CKM} depends on four independent parameters, including a complex phase which is responsible for CP violation.

CP violation was discovered in 1964 in the neutral kaon decays [78]. It has been shown that the effect could be accounted for if a third quark generation existed in addition to the two generations known at that time: for two generations, the Cabibbo 2×2 matrix is necessarily real, whereas for three generations the general 3×3 matrix is complex: the presence of an irreducible phase is a way to introduce CP violation in the SM[1].

Using the Wolfenstein parameterization, the CKM matrix can be expanded in powers of the sine of the Cabibbo angle, $\lambda = \sin\theta_C$, and approximated as

$$V^{\mathrm{CKM}} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} \simeq V_{\mathrm{CKM}}^{(3)} + \delta V_{\mathrm{CKM}} \tag{B.1}$$

$$V_{\mathrm{CKM}}^{(3)} = \begin{pmatrix} 1-\lambda^2/2 & \lambda & A\lambda^3(1-\imath\eta) \\ -\lambda & 1-\lambda^2/2 & A\lambda^2 \\ A\lambda^3(1-\rho-\imath\eta) & -A\lambda^2 & 1 \end{pmatrix} \tag{B.2}$$

It has to be noticed that the condition $\eta \neq 0$ is required to generate CP violation. The Cabibbo angle is measured to be $\theta_C = 0.221 \pm 0.002$ from decays involving

[1] CP violation is actually a requirement for baryogenesis (in fact, on request of the Sakharov conditions). On the other hand, the model of baryogenesis seems to indicate that non-SM sources of CP violation must exist.

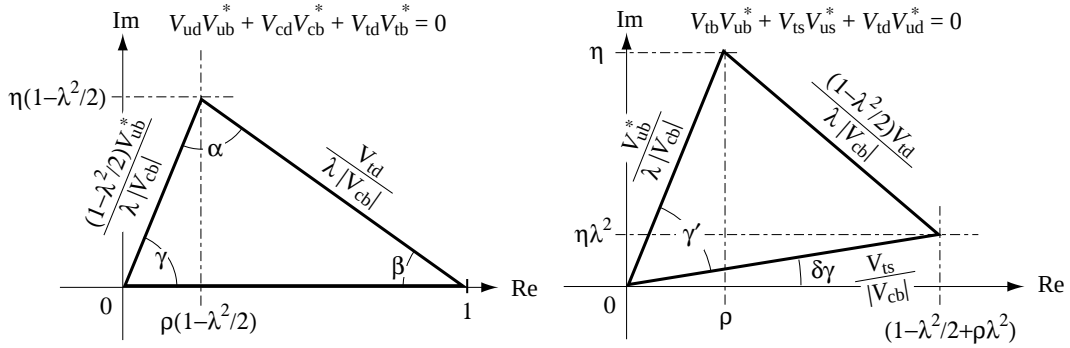


Figure B.1: Two unitarity triangles in the Wolfenstein parameterization with an approximation valid up to $\sim \mathcal{O}(\lambda^5)$.

s-quarks. For a qualitative discussion of CP violation in the B-meson systems $V_{\rm CKM}^{(3)}$ is sufficient and the second term $\delta V_{\rm CKM}$ is usually ignored. Six of the nine unitarity conditions $V_{\rm CKM}V_{\rm CKM}^\dagger = \mathbb{I}$ can be drawn as triangles in the complex plane. The two triangles relevant for the B-meson systems are shown in Figure B.1 with $\gamma = -{\rm Arg}V_{ub}$, $\beta = -{\rm Arg}V_{td}$, $\alpha = \pi - \beta - \gamma$ and $\delta\gamma = -{\rm Arg}V_{ts} - \pi$.

With just the first year of data, LHCb will determine all the angles of the CKM triangles with unprecedented precision. The following briefly summarizes the measurements in the B-meson systems which can be related to the CKM matrix elements, more precisely to the angles and sides of the unitarity triangles [79]:

Sides

- The length of the side opposite to the angle β, proportionnal to $|V_{ub}|$, will be determined from $b \to u + W$ measurements
- The length of the side opposite to the angle γ, proportionnal to $|V_{td}|$, will be determined from B-mixing measurements

Although with some non-negligible hadronic uncertainty, the knowledge of these two sides determines the triangle (the third side having a unit length), allowing for an *indirect* measurement of the angles.

Angles

- The angle β will be directly measured from the CP asymmetry in $B^0 \to J/\psi K_s^0$ decays, with an uncertainty on $\sin(2\beta)$ expected as low as 0.01, thanks to high statistics and to very small theoretical uncertainties.
- The angle α (actually $\pi - (\beta + \gamma)$) will be determined from $B^0 \to \pi^+\pi^-$ or $B^0 \to \pi^+\pi^-\pi^0$ with a precision limited by the importance of the second

order "penguin" diagrams. Interpreting this CP asymmetry in terms of CKM parameters is difficult as long as the ratio of the penguin to tree-level contributions is not known or predicted accurately. Nonetheless, the $B_d^0 \to \rho\pi \to \pi^+\pi^-\pi^0$ channel would provide an alternative measurement with the opportunity to extract tree and penguins terms separately [80].

- Several methods are investigated to measure the angle γ. The most promising ones involve the decays $B^0 \to D^{*\mp}\pi^{\pm}$ to extract $\gamma+2\beta$ and $B_s^0 \to D_s^{\mp}K^{\pm}$ to extract $\gamma - 2\delta\gamma$. In each case, the absence of penguin diagrams allows a clean determination of γ after the mixing phase (as measured from $B^0 \to J/\psi K_s^0$ or $B_s^0 \to J/\psi\phi$) has been subtracted out.

- In the SM, $\delta\gamma \simeq \lambda^2\eta$ is expected to be of the order 10^{-2} and the CP asymmetry in $B_s^0 \to J/\psi\phi$ decays should be very small. Nonetheless, the angle $\delta\gamma$ could be extracted with a precision of 0.01 and as such is adequate to look for new physics in B_s^0 mixing.

Appendix C

Liquid scintillator filled capillaries

The principle of operation of the liquid-filled capillaries is the same as for glass and plastic scintillating fibres, but the main advantage of capillaries over fibres and silicon is their outstanding resistance to radiation damages (no opacification), which is of special concern at LHC [1].

An important effort has been concentrated to produce high quality layers, ~ 2 mm thick by ~ 4 cm wide, out of manufactury-delivered raw hexagonal or square rods of packed 20 μm diameter capillaries as depicted on Figure C.1.

Up to three such layers have been mounted in an aluminium frame, as a tracking device prototype. Several tests with cosmics have been run to assess the performance of both the readout and the detector.

C.1 Technical achievements

Neat front- and back-edges of the mounted layers are required for optical contact with the readout. This has been obtained by setting up an optimized cutting procedure. A fine granularity circular diamond saw is used with water sprays and back end Ar blow to evacuate the glass dust. The procedure has been found not to break the sticks and very efficient to avoid single capillary clogging.

A semi-automated filling procedure has been developed in Lausanne to inject the scintillating liquid into the capillaries under controlled environment. The scintillating liquid — an organic solvent (1-methylnaphthalene, $C_{11}H_{10}$) to match the refractive indices and a scintillating dye (R39, a pyrazoline derivative) — is extremely sensitive to oxygen pollution and a closed circuit design had to be realized where the liquid is pushed through the capillaries by Ar overpressure. The mounting scheme is illustrated on Figure C.2. A set of flowgates G1–G6 are enslaved by the gauge control system to make the liquid constantly flow through the layers along the same direction (indicated by the arrows). When reservoir 1

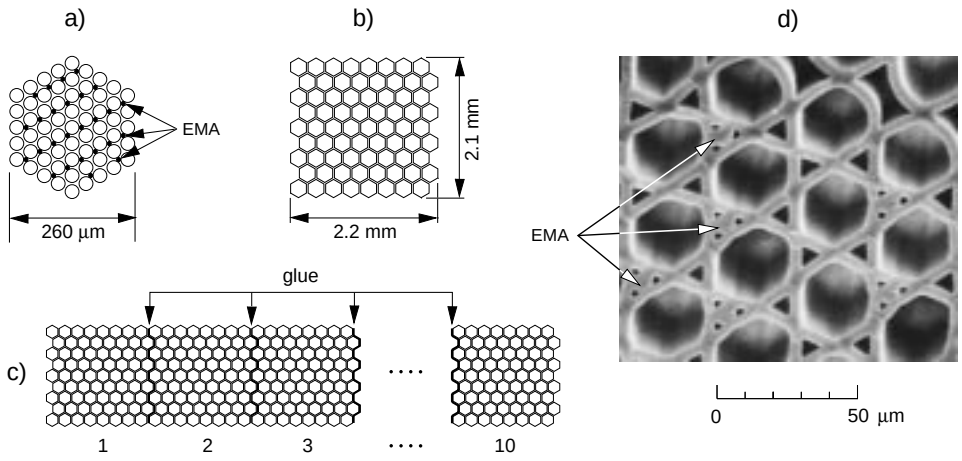


Figure C.1: Cross-sectional views illustrating layer construction. The optical noise produced by crosstalk is strongly reduced by filling the interstices between capillaries with a light-absorbing material like black glass (Extra-Mural Absorber).

is empty, the liquid is automatically sent back from reservoir 2 through a bypass. Compared to the previous design based on a peristaltic pump [2], the filling cycle time has been lowered from 24 hours to 30 minutes with identical quality or better.

The importance of the purity of the solvent on light attenuation in the capillaries can be seen from Figure C.3. Systematic light attenuation measurements have been performed with a $\sim$ 2 m long, $\sim$ 500 μm inner diameter single fibre, filled with the same dye dilution in different solvents. The products delivered by two manufacturers have been compared at various levels of purification. The attenuation length is enhanced by a factor > 3 after two successive purifications (reaching $\gtrsim 99\%$ purity) of the original solvent ($\sim 95\%$).

C.2 Tests with cosmic runs

Prototypes have been tested with cosmic rays with an opto-electronic chain comprising an Electron Bombarded CCD as readout. Figure C.4 is a side view of the setup used for the cosmic runs; on the left part are the three layers of capillaries (L1,L2,L3) in contact with the photocathode of the EBCCD on the right part. The chain showed single-photoelectron sensitvity and a readout cycle time of 100 ms (mainly due to the phospor screens response of the image intensifiers).

Cosmic events are shown on Figure C.5, raw recorded by the CCD. One can notice that the pattern formed by the hits in a same layer indicates the incident

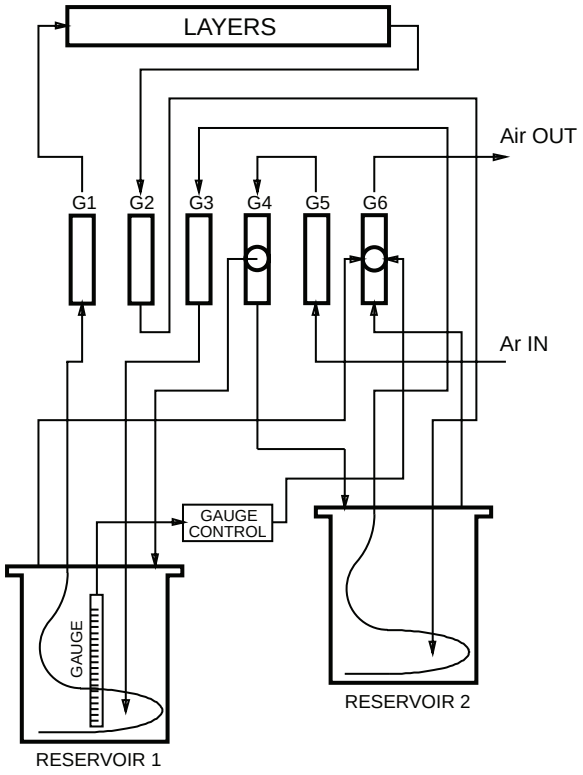


Figure C.2: The capillary layers semi-automatic filling system. The liquid is pushed by Ar overpressure from one reservoir to the other. The set of flowgates G1–G6 is driven by the gauge control system to keep a constant flow direction through the layers (see text).

track direction ("minivector"), foreshadowing the potentiality for an efficient and unambiguous pattern recognition and tracking.

The main points of interest to characterize the device performance are the track residuals which can be obtained from the tracking and the hit density attenuation with the distance from the readout surface. Figure C.6 shows the distribution of the track residuals for a linear fit performed on the hits in the three layers. The RMS $\sigma_{\mathrm{tr},3} = 48\,\mu$m is compatible with the RMS obtained for a fit on individual layers $\sigma_{\mathrm{tr},3} \sim \sqrt{3}\,\sigma_{\mathrm{tr},1}$. Data shown are for tracks crossing the layers at a distance d =15 cm from the readout surface.

By performing cluster reconstruction in events taken at 6 different values of d (along the 21 cm long layers) an exponential fit of the measured hit density reveals a light-attenuation length in the capillary layers of 48 $\pm$ 9 cm. The observed hit density was 6.6 hits/mm for particles crossing the layers perpendiculary at d=1

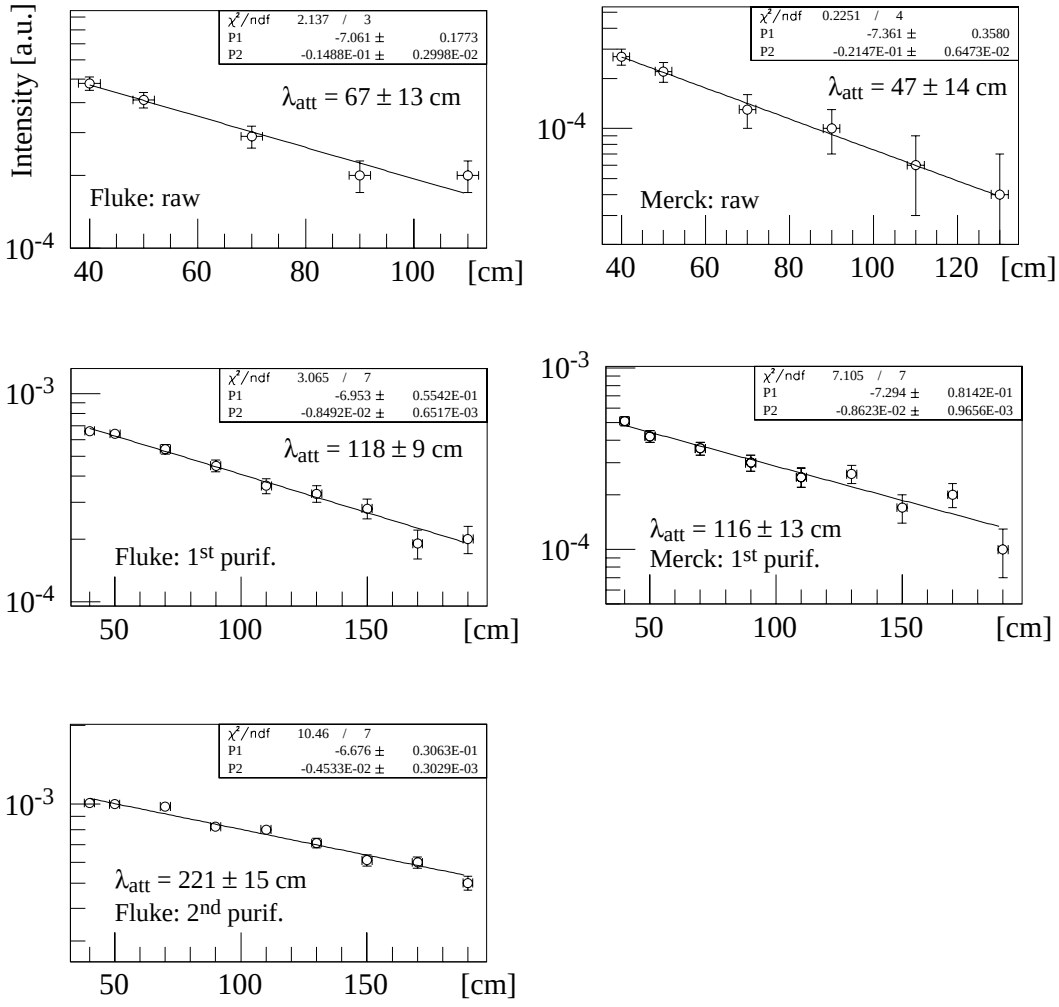


Figure C.3: Determination of the light attenuation length λ_{att} in a single macroscopic fibre for different levels of purification of the same solvent. The raw solvents delivered by two manufacturers are compared.

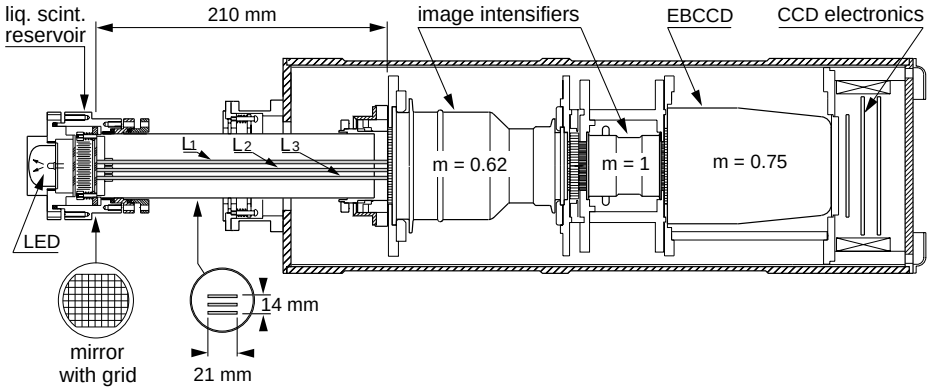


Figure C.4: The three capillary layers (L1,L2,L3) and their optoelectronic readout system, comprising two image intensifiers and a megapixel EBCCD. The magnification of each component is indicated. A reference grid, placed at the non-readout end of the capillaries, could be illuminated from behind by a LED to measure image distortion.

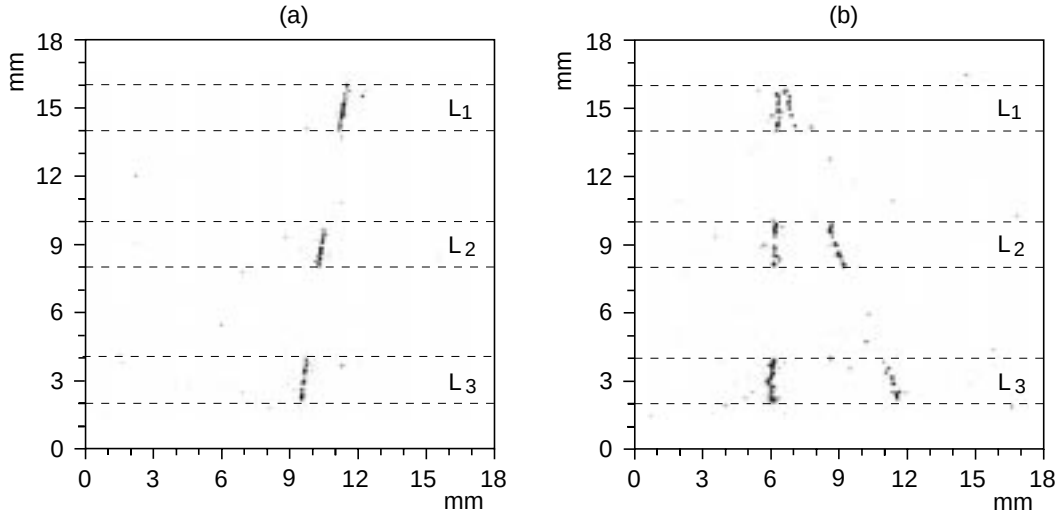


Figure C.5: Examples of cosmic-ray events recorded in the capillary layers. Interaction events (b) are not used in the event analysis. The dashed lines delimit the regions corresponding to the three layers L1, L2 and L3.

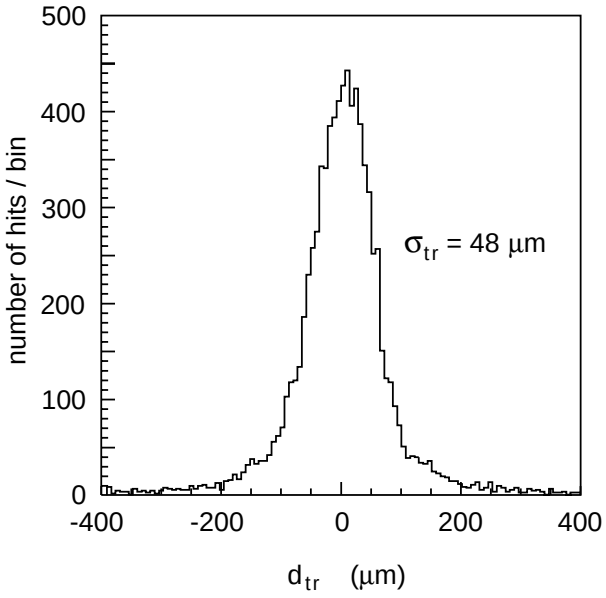


Figure C.6: Distribution of track residual $d_{\rm tr}$ for tracks fitted in all three layers.

cm and dropped to 4.2 hits/mm for particles at d=20 cm.

C.3 Further work

It has been shown in [3],[4] that a minimal hit density of 8 hits/mm is required to be able to perform efficient tracking in a collider environment with a realistic setup close to LHCb needs (efficiency becoming ~ 1 for 10 hits/mm). The value of 6.6 hits/mm obtained with our setup is still a bit low. The reason is associated to the quantum efficiency of the photocathodes $\sim 10\%$ for the wavelength of scintillation light of the R39 dye (green $\sim$ 500 nm), and the scintillation yield itself is poor ($\sim 2\%$).

Nevertheless the situation is felt to be significantly improved when considering recent progress done in the field of new scintillating media [5] showing light yields ranging up to 10 times higher for wavelengthes down to the blue ($\sim$ 400 nm), attainable for specific temperature conditions.

Initially thought as an option for the vertex detector system in LHCb [6], capillaries layers technique is currently missing an efficient and fast readout system to be competitive in a high luminosity environment. As mentioned above, the EBCCD chain used with cosmics has a readout cycle of 100 ms, excluding any perspective at LHC.

Approaches based on a high-speed gateable image pipeline [7] or more particularly on fast pixel detectors [3] have been studied and remain of special concern

given the increasing progress in the field (APDs, pad-HPDs, multi-anode PMTs) and the concomittant price per channel reduction.

In the context of LHCb, the project has been terminated in 1999 as the deadline for the choice of a vertex detector candidate forced the collaboration to preferrably opt for a more mature technology as is silicon.
The latest activity report has been submitted to the LHC committee in 2000 [8].

Bibliography

[1] A. Bay *et al. A high-resolution tracking hodoscope based on capillary layers filled with liquid scintillator.* Nucl. Instr. and Meth. **A 457** (2001) 107.

[2] P. Annis *et al. High resolution tracking using large capillary bundles filled with liquid scintillator.* Nucl. Instr. and Meth. **A 449** (2000) 60.

[3] C. Currat *et al. Capillary detectors.* Nucl. Instr. and Meth. **A 418** (1998) 186.

[4] P. Koppenburg. *Travail de diplôme.* IPHE - Université de Lausanne (1997).

[5] S. Golovkin *et al. New scintillating media based on liquid crystals for particles detectors.* Instrum. Exp. Tech. **42** (1999) 771.

[6] The LHC-B Collaboration. *LHC-B Letter Of Intent.* CERN/LHCC 95-5 (1995).

[7] A.G. Berkovsky. *A high speed image gateable image pipeline.* Nucl. Instr. and Meth. **A 380** (1996) 537.

[8] The RD46 Collaboration. *RD46 status report.* LHCC 2000-027 (2000).

Other publications related to this work:

- The RD46 Collaboration. *RD46 status report.* LHCC 97-038 (1997).
- T. Wolff *et al. High resolution tracking devices based on capillaries filled with liquid scintillator.* Proceedings of SCIFI'97, American Inst. Phys. Vol. 450 (1998) 200.
- I.E. Kreslo *et al. A single-photon multichannel detector: the megapixel EBCCD.* Proceedings of SCIFI'97, American Inst. Phys. Vol. 450 (1998) 509.

Appendix D

Simulated data production

This appendix summarizes various technical informations about the production of the event samples used in the present analysis in both FSIM and SICB.

D.1 SICB

A custom version of SICB has been built to deal with Higgs production in full simulation.

The studies performed in full simulation were dedicated to assess jet reconstruction and identification. In order to optimize the generation, the allowed phase space has been biased in Pythia to efficiently produce events with the relevant decay products inside LHCb acceptance.

The following routines have been modified to produce SICBMC version 233r2higgs.

- egpyinit.F: force the range of allowed scattering subsystem rapidities y^* in the CM frame of the event (CKIN(7)=2.2, CKIN(8)=4.5), force decays to custom channels (MDME array: $H \to b\bar{b}$, and $W, Z \to$ leptons), set SM Higgs mass (PMAS array) to 115 GeV/c^2.

- edpyinit.F: force Higgs (PID 25) and vector bosons W, Z (PID 24,23) to decay at an early stage, in Pythia.

Under these conditions, the fraction of events with the two b quarks and a charged lepton (respective decay products from the Higgs and the W, Z) within LHCb acceptance is $\sim 40\%$.

The Higgs production channel by W, Z Bremsstrahlung is selected by setting the datacard MSEL 17. The version of Pythia used was v.6134.

A total of 16000 $H + W, Z \to b\bar{b}\ell X$ events have been generated and processed into DST2 data tapes including pileup events.

The corresponding tape list can be found in the LHCb bookkeeping database (see [81]) for the event type 170000.

D.2 FSIM

The analysis on the Higgs observation potential in LHCb, described in chapter 4, has been performed with the fast simulation.

The samples of signal and backgrounds listed in Table D.1 have been generated with Pythia v.6134 with the tuning of multiple interactions described in [64].

Whenever possible without biasing the event weights, the allowed decay channels have been restrained to produce the desired signature in the final state as mentioned in the third column. The corresponding cross section is quoted in the second column.

Columns 4 and 5 give respectively the number of events generated for each sample and the corresponding number of years running with an integrated luminosity of 2 fb^{-1}. This latter number latter number gives the normalization factor used in Tables 4.2 and 4.3. The statistical error on the numbers shown in these tables has therefore to be scaled by a corresponding factor $1/\sqrt{\text{Years}}$.

Sample	σ [pb]	Decays ($\ell = \mu, e$)	Size [events]	Years/2 fb^{-1}
Signal				
$H^0W^\pm, Z^0$	1.80	$H \to b\bar{b}$	10^6	277
Background				
$t\bar{t}$ $(WbW\bar{b})$	610	free	$6 \cdot 10^6$	5
$Z^0W^\pm, Z^0$	38	free	$1.5 \cdot 10^6$	19
$W^\pm$+jets	$34 \cdot 10^3$	$W \to \ell\nu_\ell$	$49 \cdot 10^6$	0.7
γ^*/Z^0+jets	$10 \cdot 10^3$	$Z \to \ell^+\ell^-$	$20 \cdot 10^6$	0.95

Table D.1: Description of the samples generated in FSIM for the signal/background study performed in Chapter 4. The Higgs sample has been generated with $m_H = 115$ GeV/c^2. Some decays have been forced to specific channels to enhance generation efficency (see text). The quoted cross sections for pp collisions at $\sqrt{s} = 14$ TeV are at LO, using the CTEQ4L parton density functions. The $W^\pm$+jets and γ^*/Z^0+jets backgrounds are estimated by the inclusive production cross sections of respectively $W^\pm$ and γ^*/Z^0. The Drell-Yan events are generated above a cutoff CKIN(1) $= \hat{m} = 12$ GeV/c^2 (default value in Pythia).

Bibliography

[1] LHCb Collaboration. *LHCb Vertex Locator, Technical Design Report.* CERN/LHCC/2001-011.

[2] LHCb Collaboration. *LHCb Magnet, Technical Design Report.* CERN/LHCC/2000-007.

[3] LHCb Collaboration. *A Large Hadron Collider Beauty Experiment for Precision Measurements of CP-Violation and Rare Decays, Technical Proposal.* CERN/LHCC 98-4, LHCC/P4.

[4] LHCb Tracker Working Group. *Home Page.* (URL) http://lhcb-track.web.cern.ch/lhcb-track.

[5] LHCb Collaboration. *LHCb RICH, Technical Design Report.* CERN/LHCC/2000-0037.

[6] LHCb Collaboration. *LHCb Calorimeters, Technical Design Report.* CERN/LHCC/2000-0036.

[7] LHCb Collaboration. *LHCb Muon System, Technical Design Report.* CERN/LHCC/2001-010.

[8] LHCb Trigger Working Group. *Home Page.* (URL) http://lhcb-trig.web.cern.ch/lhcb-trig.

[9] M.H. Seymour. *Jets in QCD.* CERN-TH/95-176, hep-ph/9506421.

[10] M. Gell-Mann and Y. Ne'eman. *The Eightfold Way.* Benjamin, New York, 1964.

[11] J.J. Aubert *et al. Experimental Observation of a Heavy Particle J.* Phys. Rev. Lett. **33** (1974) 1404.

[12] J.-E. Augustin *et al. Discovery of a Narrow Resonance in e^+e^- Annihilation.* Phys. Rev. Lett. **33** (1974) 1406.

[13] S.W. Herb *et al. Observation of a Dimuon Resonance at 9.5 GeV in 400-GeV Proton-Nucleus Collisions.* Phys. Rev. Lett. **39** (1977) 252.

[14] M.L. Perl *et al. Evidence for Anomalous Lepton Production in e^+e^- Annihilation.* Phys. Rev. Lett. **35** (1975) 1489.

[15] The CDF Collaboration. *Observation of Top Quark Production in $\bar{p}p$ Collisions with the Collider Detector at Fermilab.* Phys. Rev. Lett. **74** (1995) 2626.

[16] The D0 Collaboration. *Observation of the Top Quark.* Phys. Rev. Lett. **74** (1995) 2632.

[17] L.H. Ryder. *Quantum field theory.* Cambridge University Press, Cambridge, 1985.

[18] LEP Collaborations / LEP Electroweak Working Group / SLD Heavy Flavour and Electroweak Groups. *A Combination of Preliminary Eletroweak Measurements and Constraints on the Standard Model.* LEPEWWG/2001-01.

[19] P.W. Higgs. *Broken symmetries and the masses of gauge bosons.* Phys. Rev. Lett. **13** (1964) 508.

[20] F. Englert and R. Brout. *Broken symmetry and the mass of gauge vector mesons.* Phys. Rev. Lett. **13** (1964) 321.

[21] The Particle Data Group. *Particle and Fields.* Eur. Phys. J. **C 15** (2000).

[22] J.F. Donoghue, E. Golowich, and B.R. Holstein. *Dynamics of the Standard Model.* Cambridge University Press, Cambridge, 1992.

[23] A. Djouadi *et al. QCD corrections to $\gamma\gamma$ decays of Higgs particles in the intermediate mass range.* Phys. Lett. **B 257** (1991) 187.

[24] P.J. Franzini *et al.* In: *Z Physics at LEP1,* eds. G. Altarelli, R. Kleiss and C. Verzegnassi, Vol. 2, p. 59. CERN 89-08.

[25] Z. Kunszt, S. Moretti, and W.J. Stirling. *Higgs production at the LHC: an Update on Cross Sections and Branching Ratios.* hep-ph/9611397v2.

[26] S. Rabi, G.B. West, and C.M. Hoffman. *Are there really any experimental limits on a light Higgs boson ?* Phys. Rev. **D 39** (1989) 828.

[27] P. Igo-Kemenes. *Status of Higgs Boson Searches.* Talk given at LEP Seminar, Nov. 3rd.

[28] The LEP Higgs Working Group. *A Hint of the Standard Model Higgs Boson at LEP: Frequently Asked Questions, and Their Answers.* Submitted to the CERN Research Board and to the CERN Directorate.

[29] The ALEPH Collaboration. *Observation of an Excess in the Search for the Standard Model Higgs Boson at ALPEH.* CERN-EP/2000-138 (Submitted to Physics Letters B).

[30] The LEP Higgs Working Group. *Standard Model Higgs Boson at LEP: Results with the 2000 Data, Request for Running in 2001.* Submitted to the LEP Comittee and to the CERN Research Board.

[31] The LHC Study Group. *The Large Hadron Collider, Conceptual Design.* CERN/AC/95-05 (LHC).

[32] ATLAS Collaboration. *ATLAS Technical Proposal.* CERN/LHCC/94-43, LHCC/P2.

[33] CMS Collaboration. *The Compact Muon Solenoid, Technical Proposal.* CERN/LHCC/94-38, LHCC/P1.

[34] ALICE Collaboration. *A Large Ion Collider Experiment, Technical Proposal.* CERN/LHCC 95-71, LHCC/P3.

[35] A.D. Martin, R.G. Roberts, W.J. Stirling, and R.S. Thorne. hep-ph/9907231.

[36] Various contributing authors. *Proceedings of the Workshop on Standard Model Physics (and more) at the LHC,* eds. G. Altarelli and M.L. Mangano. CERN 2000-004.

[37] R.K. Ellis, W.J. Stirling, and B.R. Webber. *QCD and Collider Physics.* Cambridge University Press, Cambridge, 1996.

[38] P. Bartalini. *Pile-up in signal events.* Minutes of the LHCb physics meeting, October/24/2000.

[39] O. Schneider. *Physics Generator Issues.* Talk given at the 6^{th} LHCb Software Week, November/1/2000.

[40] J.-P. Perroud on behalf of the LHCb collaboration. *Test beam results of the precision gaseous detectors for the LHCb inner tracker.* Nucl. Inst. and Methods in Physics Research **A 453** (2000) 159-164.

[41] F. Aversa *et al. Towards a Standardization of Jet Definitions.* Proceedings of the Summer Study, Snowmass, Colorado, 1990.

[42] G. Hanson *et al. Evidence for Jet Structure in Hadron Production by* e^+e^- *Annihilation.* Phys. Rev. Lett. **35** (1975) 1609.

[43] PLUTO Collaboration. *A study of jets in electron positron annihilation into hadrons in the energy range 3.1 to 9.5 GeV.* Phys. Lett. **B 78** (1978) 176.

[44] TASSO Collaboration. *Properties of hadron final states in e^+e^- annihilation at 13 GeV and 17 GeV center of mass energies.* Phys. Lett. **B 83** (1979) 261.

[45] The JADE Collaboration. *Experimental Studies on Multijet Production in e^+e^- Annihilation at PETRA Energies.* Z. Phys. **C 33** (1986) 23.

[46] R.K. Bock and A. Vasilescu. *The Particle Detector BriefBook.* Springer Verlag, 1998.

[47] S.D. Ellis *et al. Research Directions For the Decade.* Proceedings of the Summer Study, Snowmass, Colorado, 1990.

[48] G. Sterman and S. Weinberg. Phys. Rev. Lett. **39** (1977) 1436.

[49] G. Marchesini and B.R. Webber. *Associated transverse energy in hadronic jet production.* Phys. Rev. **D 38** (1988) 3419.

[50] LHCb Computing Working Group. *SICB – the LHCb Geant3 based simulation program.* (URL) http://lhcb-comp.web.cern.ch/lhcb-comp/SICB/default.htm.

[51] CN division Application Software Group. *GEANT, Detector Description and Simulation Tool.* CERN Program Library Long Writeup W5013.

[52] T. Sjöstrand. Computer Physics Commun. **82** (1994) 74.

[53] T. Sjöstrand. *Pythia and Jetset Home Page.* (URL) http://www.thep.lu.se/ torbjorn/Pythia.html.

[54] Ch. Currat *et al. Jet studies in LHCb.* LHCb 99-016 PHYS.

[55] A. Bay. *Private communications.*

[56] S.D. Ellis and D.E. Soper. *Successive combination jet algorithm for hadron collisions.* Phys. Rev. **D 48** (1993) 3160.

[57] M.H. Seymour. *KTCLUS information (*with code source*).* (URL) http://hepwww.rl.ac.uk/theory/seymour/ktclus.

[58] S. Catani, Yu. L. Dokshitzer, and B.R. Webber. *The $k_\perp$-clustering algorithm for jets in deep inelastic scattering and hadron collisions.* Physics Letters **285** (1992) 291.

[59] J. Allison. *Multiquadratic Radial Basis Functions for Representing Multi-dimensional High Energy Physics Data.* Computer Physics Commun. **77** (1993).

[60] Ch. Currat. *SM Higgs study – update.* Minutes of the LHCb physics meeting, March/8/2001.

[61] M. Spira. Report DESY T-95-05, hep-ph/9510347.

[62] A. Djouadi, J. Kalinowski, and M. Spira. Comp. Phys. Commun. **108** (1998), hep-ph/9704448.

[63] Various contributing authors. *M. Spira Home Page.* (URL) http://m.home.cern.ch/m/mspira/www/proglist.html.

[64] P. Bartalini *et al. Tuning of Multiple Interactions Generated by* Pythia. LHCb 99-028 PHYS.

[65] J. Huston C. Balázs and I. Puljak. *Higgs production: A Comparison of Parton Showers and Resummation.* hep-ph/0002032, FERMILAB-Pub-00/032-T.

[66] T. Sjöstrand. *Private communications.*

[67] P. Bartalini. *Status of the event generator studies in LHCb.* Talk given at the 2$^{\text{nd}}$ LHCb generator miniworkshop, March/2/01, IPHE Lausanne.

[68] LHCb public web pages. (URL) http://lhcb-public.web.cern.ch/lhcb-public/html/MAP.htm.

[69] A. Pukhov *et al. CompHEP – a package for evaluation of Feynman diagrams and integration over multi-particle phase space.* Preprint INP-MSU 88-41/542, hep-ph/9908288.

[70] P.D.B. Collins. *An Introduction to Regge Theory & High Energy Physics.* Cambridge University Press, Cambridge, 1977.

[71] The Particle Data Group. *Particle and Fields.* Phys. Rev. **D 54** (1996).

[72] D. Barberis *et al.* (WA102 Collaboration). Phys. Lett. **B 432** (1998) 436.

[73] S. Abatzis *et al.* (WA91 Collaboration). Phys. Lett. **B 324** (1994) 509.

[74] A. Donnachie and P.V. Landshoff. *Total cross sections.* Phys. Lett. **B 296** (1992) 227.

[75] R.M. Baltrusaitis *et al. Total Proton-Proton Cross Section at* $s^{1/2} = 30$ *TeV.* Phys. Rev. Lett. **52** (1984) 1380.

[76] I. Ya. Pomeranchuck. Sov. Phys. JETP **7** (1958) 499.

[77] J.R. Cudell *et al. High-energy forward scattering and the Pomeron: Simple pole versus unitarized models.* Phys. Rev. **D 61** (2000).

[78] J.W. Cronin *et al. Evidence for the 2π decay of the K_2^0 meson.* Phys. Rev. Lett. **13** (1964) 138.

[79] O. Schneider for the LHCb collaboration. *Overview of the LHCb experiment.* Nucl. Inst. and Methods in Physics Research **A 446** (2000) 213-221.

[80] A. Jacholkowska for the LHCb collaboration. *The LHCb EM calorimeter and π^0 detection.* Nucl. Inst. and Methods in Physics Research **A 446** (2000) 259-263.

[81] LHCb Computing Working Group. *LHCb bookkeeping.* (URL) http://oraweb01.cern.ch:9000/lhcb_prod/owa/book_procedures.general?procname=init.

List of Figures

List of Tables