Università degli Studi di Padova
Dipartimento di Fisica ed Astronomia "G. Galilei"

Tesi di Laurea Magistrale in Fisica

Classification of muons produced in semileptonic B hadron decays in $pp \to t\bar{t}$ events at LHC

Candidato: **Alessio Boletti**

Relatore: **Dr. Martino Margoni**

Correlatori: **Dr. Paolo Ronchese**
Prof. Franco Simonetto

Anno Accademico 2013/2014

Contents

Abstract

This thesis work is part of the measurement of the time-integrated mixing probability, $\overline{\chi}$, of the b-hardons produced in pp collision at a centre of mass energy $\sqrt{s} = 8$ TeV, using the data acquired by the CMS experiment in the 2012 run. Only the b-hadrons generated in the decay of a top quark, in events where a top quark pair is produced, are used. This kind of events provides a self-tagged sample of b-hadrons. Indeed, if a semileptonic decay of at least one of the top quarks occurs, the charge of the produced lepton, ℓ_{hard}, provides information about their flavour and thus about the b-hadron flavour at the production. In events where also the b-hadron decays semileptonically, the charge of the lepton produced in this decay, ℓ_{soft}, tags the flavour of the hadron at the decay.

In this work, the discrimination of the muons that originates from the semileptonic decay of the b-hadron and muons that come from the subsequent charmed hadron decay is obtained by means of a Multi-Variate Analysis.

Also a set of selection criteria that reduces the background contamination is provided, as well as a second MVA method aiming at the discrimination of signal and background events.

A preliminary fit strategy is also tested, in order to extract the relative abundances of b-hadron, charmed hadron and background muons from the data sample.

Finally, the agreement between the data and Monte Carlo distributions for the variables used in the Multi-Variate Analyses is tested.

Chapter 1

The Standard Model

To describe the fundamental laws of Nature using the perturbative approach of a Quantum Field Theory (QFT) it is necessary to write down the Lagrangian of the model. To do that it is essential to set three axioms:

- the gauge symmetry;
- the representation of fermions and scalars under this symmetry;
- the pattern of spontaneous symmetry breaking.

Once defined these points the model is built and its Lagrangian is the most general renormalizable one that can be assembled with the fermion and scalar fields satisfying the gauge symmetry and providing the spontaneous symmetry breaking pattern.

The request of a renormalizable Lagrangian limits the number of its terms. This is a sort of approximation: all the terms that contain a power of $1/\Lambda$, where Λ is the UV energy scale, are neglected.

Furthermore, it is required the Lagrangian to be Poincaré invariant[1]. This is the same of assuming the space to be isotropic and homogeneous and the time to be homogeneous.

The theory contains a finite number of parameter which are not theoretically predicted. Those parameters need to be measured experimentally and a series of precision measurements of some of them is a way to test the consistency of the model with the observed phenomena.

The Standard Model is a theory defined as follows:

- the gauge symmetry is

$$G_{\mathrm{SM}} = SU(3)_{\mathrm{C}} \times SU(2)_{\mathrm{L}} \times U(1)_{\mathrm{Y}} \quad ;$$

[1] It is to transform like a scalar under Poincaré transformation.

- the five fermion representations are

  $$Q^I_{Li}(3,2)_{+1/6}\,, \quad U^I_{Ri}(3,1)_{+2/3}\,, \quad D^I_{Ri}(3,1)_{-1/3}\,, \quad L^I_{Li}(1,2)_{-1/2}\,, \quad E^I_{Ri}(1,1)_{-1}\,,$$

  where the index I indicates that these are the gauge interaction eigenstates, the L and R indices refer to the chirality eigenstates, the index $i = 1,2,3$ indicates the generation, the numbers in bracket indicate the multiplicity of the representation under $SU(3)_{\rm C}$ and $SU(2)_{\rm L}$ and the last number is the hypercharge. Furthermore there is one scalar representation:

  $$\phi(1,2)_{+1/2} \quad ;$$

- the scalar field assumes a vacuum expectation value:

  $$\langle\phi\rangle = \begin{pmatrix} 0 \\ v/\sqrt{2} \end{pmatrix},$$

  which implies the gauge group spontaneous breaking:

  $$SU(3)_{\rm C} \times SU(2)_{\rm L} \times U(1)_{\rm Y} \to SU(3)_{\rm C} \times U(1)_{\rm EM} \quad .$$

From these choices is possible to construct the Lagrangian of the Standard Model. It can be divided in three terms:

$$\mathcal{L}_{\rm SM} = \mathcal{L}_{\rm kin} + \mathcal{L}_{\rm Hig} + \mathcal{L}_{\rm Yuk} \quad .$$

The first part contains the kinetic term of the fermion and scalar fields where, in order to preserve the gauge invariance, the derivative is replaced by the covariant derivative:

$$D^\mu = \partial^\mu + ig_s G^\mu_a L_a + igW^\mu_b T_b + ig'B^\mu Y \quad ,$$

where G^μ_a are the eight gluon fields, W^μ_b the three weak interaction fields and B^μ the hypercharge field, L_a and T_b respectively the $SU(3)$ and $SU(2)$ generators (of the same multiplicity of the fermion representation they are applied to) and Y is the $U(1)_{\rm Y}$ charge. Finally g_s, g and g' are three parameters to be experimentally determined.

The second part is the Higgs potential, describing the scalar field self interaction:

$$\mathcal{L}_{\rm Hig} = \mu^2\phi^\dagger\phi - \lambda(\phi^\dagger\phi)^2 \quad ,$$

where $\lambda > 0$ and $\mu^2 < 0$, requesting the vacuum stability and a VEV different from zero.

Finally the third part contains the Yukawa couplings between the scalar and the fermion fields:

$$\mathcal{L}_{\rm Yuk} = -Y^d_{ij}\overline{Q^I_{Li}}\phi D^I_{Rj} - Y^u_{ij}\overline{Q^I_{Li}}\tilde{\phi}U^I_{Rj} - Y^e_{ij}\overline{L^I_{Li}}\phi E^I_{Rj} + \text{h.c.} \quad .$$

Because of the scalar field VEV at the available energy scales, these terms provide the fermion fields with the mass term (with the exception of the neutrinos, with which it is not possible to build a renormalizable gauge invariant Yukawa term). This part of the Standard Model Lagrangian contains 13 parameters, the 9 massive fermion masses and the 4 CKM matrix parameters, which will be introduced in section 1.2.

1.1 Symmetries

The gauge symmetry just introduced is a continuous local symmetry. In the Standard Model there are other types of symmetry, both continuous and discrete, exact, approximate or completely broken.

1.1.1 Discrete symmetries

There are three discrete symmetries: C, P, T. Any Lorentz invariant QFT must be invariant under CPT. This invariance guarantees that a particle and his antiparticle have the same mass and total decay rate. Then a theory violates T if and only if it violates CP.

The SM is not invariant under C and under P. These transformations change the chirality of the fermion fields but left handed and right handed fields have different representations, so the C and P symmetries are maximally violated. This fact is true independently of the value of the Lagrangian parameters.

The CP invariance of the theory is more complicated to probe. In fact, a CP transformation applied on the Lagrangian transforms any term in its hermitian conjugate, but leaving untouched the coefficients. So the Lagrangian is invariant under CP transformation, unless any of its parameters contains a physical phase. The only term that could contain such a phase is the quark Yukawa coupling,

$$Y_{ij}\overline{\psi_{Li}}\phi\psi_{Rj} + Y_{ij}^{*}\overline{\psi_{Rj}}\phi^{\dagger}\psi_{Li} \quad .$$

A CP transformation on this term act as follows:

$$Y_{ij}\overline{\psi_{Li}}\phi\psi_{Rj} \leftrightarrow Y_{ij}\overline{\psi_{Rj}}\phi^{\dagger}\psi_{Li} \quad ,$$

so the Lagrangian is invariant under CP if $Y_{ij} = Y_{ij}^{*}$.

The kinetic and the scalar potential terms are CP invariant. An eventual CP violation is not intrinsic in the theory, but depends on the Lagrangian parameters and need to be measured experimentally.

1.1.2 Global symmetries

Furthermore there is a set of global symmetries, that can be divided in accidental and approximate symmetries. They can be very useful in amplitude calculation, although they are not fundamental like the gauge symmetry.

An accidental symmetry is not imposed to the Lagrangian, but thanks to the requests of renormalizability and gauge invariance all its terms result to be invariant under that symmetry. In the SM an accidental global symmetry is

$$U(1)_B \times U(1)_e \times U(1)_\mu \times U(1)_\tau \quad ,$$

where $U(1)_B$ is the baryon number and $U(1)_l$ are the lepton family numbers. Quarks carry baryon number $+1/3$ and no lepton number, leptons carry only lepton number $+1$ for the family they belong to.

On the other hand, the theory is not invariant under approximate symmetries. It would be only if some Lagrangian parameters were set to zero. If indeed these parameters are small (compared with an appropriate energy scale, if dimensional) it is anyway useful to exploit this kind of symmetries to simplify the calculations. An example of an approximate symmetry is the isospin, which transforms the up quark in down quark and vice versa. This symmetry is broken by the quark mass split

$$\frac{m_d - m_u}{\Lambda_{\mathrm{QCD}}}$$

and by the electric charge (the α_{em} parameter). Another similar approximate symmetry is $SU(3)_{\mathrm{flavour}}$ which equalizes the up, down and strange quark. This symmetry is less exact than isospin because of the greater quark mass split.

1.2 The CKM matrix

The fermion field representations presented in this chapter are the weak interaction eigenstates, so the covariant derivative is diagonal in this basis. However most of the measures use the mass eigenstates, which diagonalize the Yukawa coupling terms, so it is useful to perform a basis rotation. In the quark sector the flavour changing interactions guarantee that the weak interaction eigenstates and the mass ones are different.

The quark mass terms derive from the Yukawa coupling, substituting for the scalar field its VEV:

$$\mathcal{L}^q_{\mathrm{mass}} = -(M_d)_{ij}\overline{D^I_{Li}}D^I_{Rj} - (M_u)_{ij}\overline{U^I_{Li}}U^I_{Rj} + \mathrm{h.c.}\,, \qquad M_q = \frac{v}{\sqrt{2}}\,Y^q \quad .$$

To diagonalize the mass matrix M_q two unitary matrices are needed, V_{qL} and V_{qR}, such that

$$V_{qL}M_qV^\dagger_{qR} = M_q^{\mathrm{diag}} \quad ,$$

where M_q^{diag} is diagonal and real. Then the quark mass eigenstates are

$$q_{Li} = (V_{qL})_{ij}q^I_{Lj}\,, \qquad q_{Ri} = (V_{qR})_{ij}q^I_{Rj} \quad .$$

The only Lagrangian term that is changed by this basis rotation is the charge current coupling. Defining

$$W_\mu^\pm = \frac{W_{1\,\mu} \pm W_{2\,\mu}}{\sqrt{2}} \quad ,$$

after the basis rotation that term become

$$\mathcal{L}^q_{W^\pm} = -\frac{g}{\sqrt{2}} \overline{u_{Li}} \gamma^\mu (V_{uL} V_{dL}^\dagger)_{ij} d_{Lj} W_\mu^+ + \text{h.c.} \quad .$$

The unitary matrix

$$V = V_{uL} V_{dL}^\dagger = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix}$$

is the Cabibbo Kobayashi Maskawa (CKM) mixing matrix, that quantifies the charge current weak interaction mixing of quark from different generations.

An unitary 3×3 matrix has 9 free parameters, 3 real values and 6 phases. Only one of these phases is physical, the other 5 can be set to zero by some quark field phase redefinitions. There are different ways to parametrize the CKM matrix with the 4 physical parameters, the most used is the standard parametrization:

$$V = \begin{pmatrix} c_{12}c_{13} & s_{12}c_{13} & s_{13}\exp -i\delta \\ -s_{12}c_{23} - c_{12}s_{23}s_{13}\exp i\delta & c_{12}c_{23} - s_{12}s_{23}s_{13}\exp i\delta & s_{23}c_{13} \\ s_{12}s_{23} - c_{12}c_{23}s_{13}\exp i\delta & -c_{12}s_{23} - s_{12}c_{23}s_{13}\exp i\delta & c_{23}c_{13} \end{pmatrix} \quad ,$$

where $c_{ij} = \cos\theta_{ij}$ and $s_{ij} = \sin\theta_{ij}$. The three real parameters are the three θ_{ij} angles and δ is the physical Kobayashi-Maskawa phase. Another typical parametrization, sometimes more useful then the standard one, is the Wolfenstein parametrization. It makes use of the three parameters λ, A and ρ and the phase η. It is an expansion in power of λ and usually is written to the third order:

$$V = \begin{pmatrix} 1 - \frac{1}{2}\lambda^2 & \lambda & A\lambda^3(\rho - i\eta) \\ -\lambda & 1 - \frac{1}{2}\lambda^2 & A\lambda^2 \\ A\lambda^3(1 - \rho - i\eta) & -A\lambda^2 & 1 \end{pmatrix} + \mathcal{O}(\lambda^4) \quad .$$

A very useful tool that allows to summarize the measure of the CKM matrix elements taking into account its unitarity is the unitary triangle. The unitarity implies that

$$\sum_i V_{ij} V_{ik}^* = 0\,, \qquad j,k = d,s,b \quad j \neq k \quad .$$

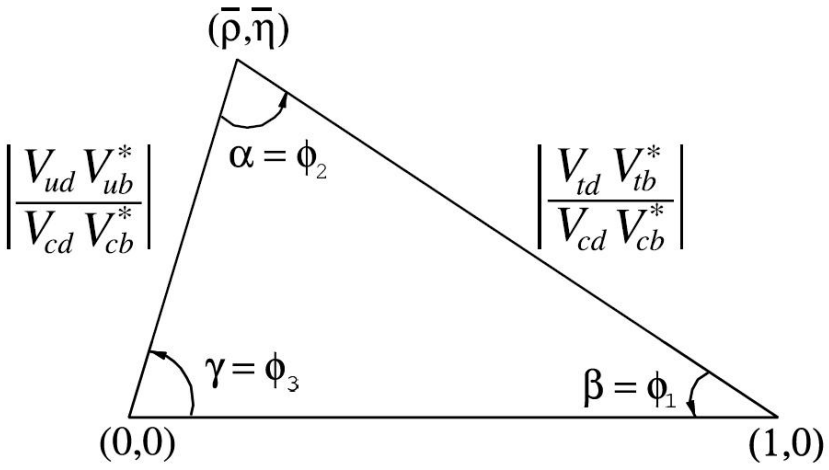


Figure 1.1: The unitary triangle for the (d, b) quark pair.

These are three independent relations between the CKM elements. Using one of these constrains it is possible to construct a triangle in the complex plane so that any term of the summation represents one side of the triangle. Now choosing the phase convention such that the "charm addend" ($V_{cj}V_{ck}^*$) is real and rescaling the triangle such that the "charm side" is equal to 1, one obtains the unitary triangle for the (j, k) quark pair.

An example, for the (d, b) pair, is shown in figure 1.1. In this triangle the complex coordinates of the vertex are the Wolfenstein parameters ρ and η and the amplitude of the angles is related to the phases of the CKM matrix elements:

$$\alpha = \arg\left[-\frac{V_{td}V_{tb}^*}{V_{ud}V_{ub}^*}\right], \quad \beta = \arg\left[-\frac{V_{cd}V_{cb}^*}{V_{td}V_{tb}^*}\right], \quad \gamma = \arg\left[-\frac{V_{ud}V_{ub}^*}{V_{cd}V_{cb}^*}\right] \quad .$$

The γ angle is also equal to the standard parametrization phase δ_{KM}. This triangle is the easiest to probe and the most precisely measured, thanks to the fact that its side lengths have the same order of magnitude. A plot that summarizes all the measurements of the quantities related to this triangle is show in figure 1.2.

Also the unitary triangle for the (s, b) quark pair has been probed. It has a side two order of magnitude smaller than the other two, so it is more difficult to constrain the position of its vertex. The plot that summarizes the measures related to the elements of this triangle is shown in figure 1.3.

The third triangle is useless, being one side three order of magnitude smaller than the other sides and compatible with their error.

Another important feature of the unitary triangles is that before the rescaling they have the same area which is equal to $|J_{\mathrm{CKM}}|/2$, where J_{CKM} is the Jarlskog invariant, that is a parametrization independent estimator

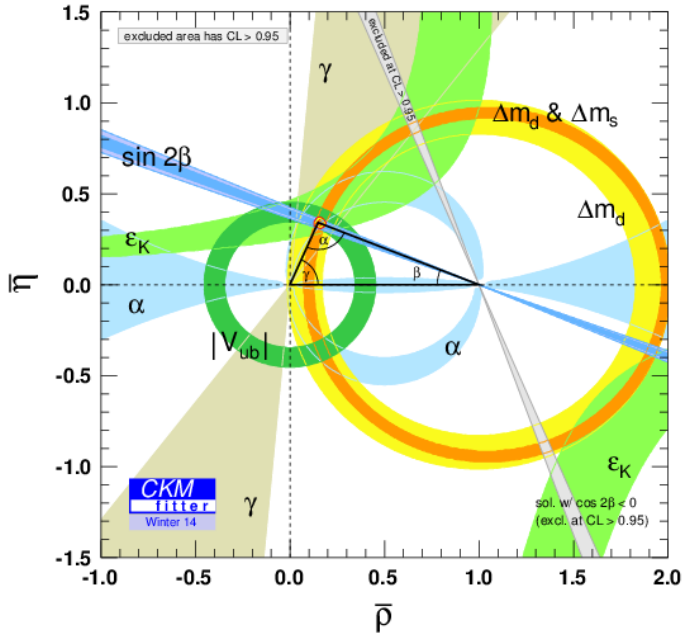


Figure 1.2: Plot of the measurement status for the (d, b) unitary triangle.

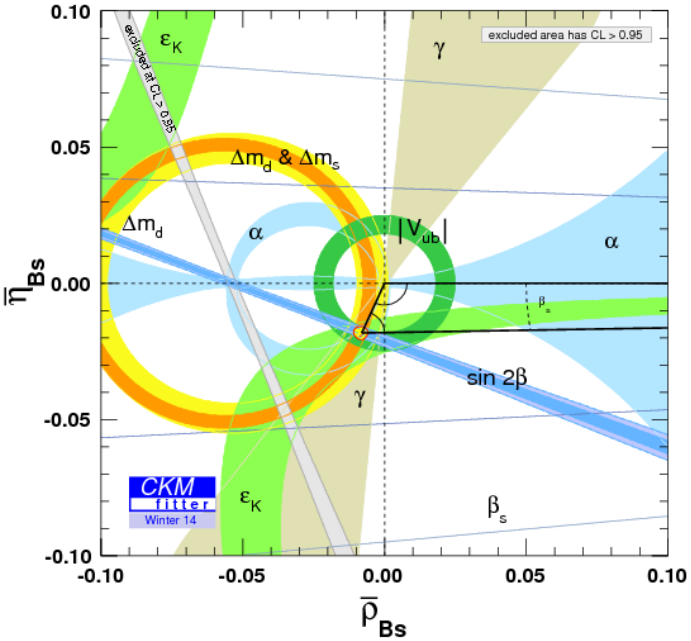


Figure 1.3: Plot of the measurement status for the (s, b) unitary triangle.

of the CP violation provided by the CKM matrix. This parameter will be introduced in section 1.4.

1.3 Meson mixing

In the SM there are no flavour changing neutral currents (FCNC) at tree level. This because all the neutral interactions are diagonal and universal, such that any basis rotation leaves them diagonal and then flavour conserving. Thanks to this fact, the leading order FCNC processes are generated by one loop charge current mediated diagrams, producing some interesting observable interference effects. One of these effects is the GIM mechanism which produces the cancellation of the quark mass independent terms in the FCNC loop amplitudes.

Another interesting FCNC process is the meson mixing. When a particle and its antiparticle can decay in a common final state (also a virtual state) the phenomenon of oscillation is present. This request and the conservation of the baryon and lepton numbers and of the electric charge impose that only the neutral mesons that do not coincide with their antiparticle can oscillate, in particular the neutral kaons ($K^0, \overline{K^0}$), the neutral charmed mesons ($D^0, \overline{D^0}$) and the two neutral B meson pairs ($B^0, \overline{B^0}$) and ($B_s^0, \overline{B_s^0}$).

Due to the meson peculiarities, these systems behave in different ways. We are interested in the B meson system, but the oscillation formalism here presented is the same for all the systems.

1.3.1 Oscillation formalism

In the oscillation phenomena there is an open system with two meson orthogonal states $|P^0\rangle$ and $|\overline{P^0}\rangle$ and the continuum of states they can decay to. Before the decay the system can be described by a coherent superposition of the two meson states and can be represented by a two component vector.

Being an open system, the Hamiltonian is non-Hermitian and can be decomposed as sum of two 2×2 Hermitian matrices:

$$\mathcal{H} = M - \frac{i}{2}\Gamma \quad .$$

Here M is a typical mass matrix, but with also off-diagonal components, $M_{12} = M_{21}^*$, which are proportional to the $P^0 \leftrightarrow \overline{P^0}$ oscillation amplitude. Moreover, because of the CPT invariance, $M_{11} = M_{22}$. This matrix is associated to the system evolution via off-shell (dispersive) intermediate states. Also in the Γ matrix there are off-diagonal components, $\Gamma_{12} = \Gamma_{21}^*$, related to the common channel decay amplitude and also here, due to CPT, $\Gamma_{11} = \Gamma_{22}$. This matrix, on the other hand, is related to the system evolution via the on-shell (absorptive) intermediate states.

If the Hamiltonian is not diagonal, the two flavour eigenstates (P^0, $\overline{P^0}$) are not mass eigenstate and thus do not have mass and decay rate well defined. The mass eigenstates are

$$|P_{L,H}\rangle = p|P^0\rangle \pm q|\overline{P^0}\rangle \quad ,$$

where p and q depends on the Hamiltonian matrix elements. P_L is defined as the eigenstate with the smaller mass[2] ($m_L < m_H$).

The two mass eigenstate are normalized, so the parameters must satisfy

$$|p|^2 + |q|^2 = 1 \quad ,$$

$$\left(\frac{q}{p}\right)^2 = \frac{M_{12}^* - (i/2)\Gamma_{12}^*}{M_{12} - (i/2)\Gamma_{12}} \quad .$$

If CP is not violated, M_{12} and Γ_{12} have the same phase (in section 1.4 is discussed the case with CPV). Thus,

$$\left|\frac{q}{p}\right| = 1$$

and the mass eigenstates are orthogonal

$$\langle P_L|P_H\rangle = |p|^2 - |q|^2 = 0 \quad ,$$

which was not guaranteed, because the Hamiltonian is not Hermitian.

The mass difference Δm and the width difference $\Delta\Gamma$ are defined as

$$\Delta m = m_H - m_L \quad ,$$

$$\Delta\Gamma = \Gamma_H - \Gamma_L \quad .$$

While Δm is positive by definition, $\Delta\Gamma$ can be positive or negative. The average mass and the average width are

$$m = \frac{m_H + m_L}{2} \quad ,$$

$$\Gamma = \frac{\Gamma_H + \Gamma_L}{2} \quad .$$

It is useful also to define the dimensionless parameter x and y:

$$x \equiv \frac{\Delta m}{\Gamma}\,, \qquad y \equiv \frac{\Delta\Gamma}{2\Gamma}$$

and the phase

$$\theta \equiv \arg(M_{12}\Gamma_{12}^*) \quad .$$

[2]The kaon system mass eigenstate (K_S, K_L) are defined on the base of their width, being their large width difference more peculiar than their small mass difference.

The solution of the eigenvalue equation gives

$$(\Delta m)^2 - \frac{1}{4}(\Delta\Gamma)^2 = 4|M_{12}|^2 - |\Gamma_{12}|^2 \quad .$$

$$\Delta m \Delta\Gamma = 4\mathcal{R}e(M_{12}\Gamma_{12}^*) \quad ,$$

which, in the CP conservation hypothesis, give

$$\Delta m = 2|M_{12}|\,, \qquad |\Delta\Gamma| = 2|\Gamma_{12}| \quad .$$

1.3.2 System evolution

In this section the time evolution of the meson system is presented, in the approximation of CP conservation. The corrections to take CPV into account are presented in section 1.4.

Choosing an adequate phase convention, we can obtain $|q| = |p| = 1/\sqrt{2}$, so that

$$|P_{L,H}\rangle = \frac{1}{\sqrt{2}}(|P^0\rangle \pm |\overline{P^0}\rangle) \quad .$$

Since in experiments these systems are often generated as interaction eigenstates with definite flavour, it is more useful to take a pure $|P^0\rangle$ initial state. Applying the time evolution operator

$$|P^0(t)\rangle = g_+(t)|P^0\rangle + g_-(t)|\overline{P^0}\rangle \quad ,$$

where

$$g_+(t) = e^{-it(m-i\Gamma/2)}\left[\cosh\left(\frac{\Delta\Gamma t}{4}\right)\cos\left(\frac{\Delta m t}{2}\right) - i\sinh\left(\frac{\Delta\Gamma t}{4}\right)\sin\left(\frac{\Delta m t}{2}\right)\right] \tag{1.1}$$

$$g_-(t) = e^{-it(m-i\Gamma/2)}\left[-\sinh\left(\frac{\Delta\Gamma t}{4}\right)\cos\left(\frac{\Delta m t}{2}\right) + i\cosh\left(\frac{\Delta\Gamma t}{4}\right)\sin\left(\frac{\Delta m t}{2}\right)\right] \tag{1.2}$$

From these formulas and with the approximation $y \ll 1$ (valid for D, B and B_s systems) it is possible to derive the time-dependent mixing probability

$$\mathcal{P}(P^0 \to \overline{P^0})[t] = |\langle P^0(t)|\overline{P^0}\rangle|^2 = e^{-\Gamma t}\,\frac{1-\cos(\Delta m t)}{2} \quad ,$$

$$\mathcal{P}(P^0 \to P^0)[t] = |\langle P^0(t)|P^0\rangle|^2 = e^{-\Gamma t}\,\frac{1+\cos(\Delta m t)}{2} \quad .$$

Then, measuring the oscillation frequency, it is possible to estimate the mass split.

1.3.3 Mixing measurements

In order to measure the time dependent oscillation probability it is necessary to determine the flavour of the system at two different times.

One of these measurements is at the decay time: the meson can decay in channels common to both flavour eigenstates, but also in channels that produce different final states for each of them, like the semileptonic decay in which the lepton charge tag the charge of the heavy quark in the meson.

The other flavour measurement depends on the process used to produce the meson. For brevity only the B hadron systems are considered here. There are three ways to produce a B meson.

- From coherent b quark pair production, typical of the b-factories, when both the quarks hadronize producing two B_0 mesons. The flavour can be tagged from the measurement of the opposite side meson decay. Being in a coherent state, the two mesons have at any time opposite flavours, so tagging the opposite one also the analysed meson flavour is known at that time.

- From incoherent b quark pair production, typical of high-energy lepton and hadron colliders, when the opposite side b quark hadronizes in a non oscillating hadron. If its decay final state gives information about its flavour, also the flavour of the analysed meson at the production is known.

- From the semileptonic decay of a top quark. In this scenario the produced lepton and the top has the same charge sign, thus it gives also information about the B meson flavour at the production. This type of flavour tagging is used for the current analysis and further details are presented in chapter 3.

It is evident the presence of two time scale. One is the oscillation time scales, given by the oscillation period or Δm, the other is the decay time scale, given by the average width Γ. The best situation to perform mixing measurements is when these two time scales have the same order of magnitude, that is $x \sim 1$. This situation occurs for the kaon and B^0 meson systems. If the oscillation time scale is smaller than the particle mean life ($x \gg 1$) the system oscillate many times before decaying and without a very good time resolution and tagging power the oscillation probabilities mediate to 1/2. This is the case of B_s^0 meson system. Finally, if the oscillation time scale is bigger than the decay time scale ($x \ll 1$) the system has no time to exhibits oscillation phenomena before decaying. This is the case of D^0 system.

Time integrated probability

In addition to the time dependent probabilities, there is another parameter interesting to measure: the time integrated mixing probability, χ. It is simply the probability that the two flavour measurements tag the meson in two different eigenstates:

$$\chi = \int \mathcal{P}(P^0 \to \overline{P^0})\,\mathrm{d}t = \frac{x^2+y^2}{2(x^2+1)} \quad .$$

This quantity is more simple to measure, because it does not require a optimal time resolution but only a good event reconstruction.

For the B^0 system, χ_d measurements was published by ARGUS and CLEO collaborations [1–3], obtaining an average of $\chi_d = 0.182 \pm 0.015$ [4]. In the approximation of no CP violation and null $\Delta\Gamma_d$, also the time-dependent measurements of Δm_d performed by the ALEPH [5], DELPHI [6, 7], L3 [8], OPAL [9, 10], BaBar [11], Belle [12], CDF [13], DØ [14] and LHCb [15] collaborations can be used to compute χ_d. The resulting world averages [4] are then

$$\Delta m_d = 0.510 \pm 0.003\,\mathrm{ps}^{-1} \quad ,$$

$$\chi_d = 0.1874 \pm 0.0018 \quad .$$

For the B_s^0 system, a χ_s direct measurement was never performed. Time-dependent measurements of Δm_s were published by the CDF [13] and LHCb [16] collaborations, providing a combined measurement of $\Delta m_s = 17.761 \pm 0.022\,\mathrm{ps}^{-1}$ [4]. Assuming no CP violation and a width difference $\Delta\Gamma_s = 0.091 \pm 0.008\,\mathrm{ps}^{-1}$, it results

$$\chi_s = 0.499311 \pm 0.000007 \quad .$$

Furthermore, mixing measurements can significantly improve our knowledge on the fractions f_u, f_d, f_s and f_{baryon}, defined as the fractions of B_u, B_d^0, B_s^0 and b-baryons in an unbiased sample of weakly decaying b-hadrons produced in high-energy collisions. Indeed, a time-integrated mixing analysis, performed at high energy without the identification of the b-hadron but only observing a semileptonic decay, measures the quantity

$$\overline{\chi} = f'_d\chi_d + f'_s\chi_s \quad ,$$

where f'_d and f'_s are the fractions of B_d^0 and B_s^0 hadrons in a sample of semileptonic b-hadron decays. Assuming that all b-hadrons have the same semileptonic decay width implies $f'_q = f_q/(\Gamma_q\tau_b)$ ($q = s,\, d$), where τ_b is the average b-hadron lifetime.

The average of the $\overline{\chi}$ measurements performed at the LEP collider [17], with e^+e^- collision at $\sqrt{s} = m_Z$, is

$$\overline{\chi} = 0.1259 \pm 0.0042 \quad ,$$

	LEP	Tevatron	LHC
$\overline{\chi}$	0.1259 ± 0.0042	0.147 ± 0.011	
$f_u = f_d$	0.404 ± 0.009	0.339 ± 0.031	
f_s	0.103 ± 0.009	0.111 ± 0.014	
f_{baryon}	0.089 ± 0.015	0.212 ± 0.069	
f_s/f_d	0.254 ± 0.025	0.328 ± 0.039	$0.267^{+0.021}_{-0.020}$

Table 1.1: Averages of $\overline{\chi}$ and b-hadron fractions measurements at different colliders [4].

while at the Tevatron collider [18], with $p\overline{p}$ collision at $\sqrt{s} = 1.8 - 2$ TeV, the average result is

$$\overline{\chi} = 0.147 \pm 0.011 \quad .$$

In LHC collider, with pp collision at $\sqrt{s} = 7 - 8$ TeV, this measurement is not performed yet. However it could be useful, in order to probe an eventual energy dependence of the b-hadron fractions. A complete summary on the status of the b-hadron fraction measurements is reported in table 1.1.

CKM parameter measurements

The measurement of the oscillating meson system parameters can lead to an estimation of the related CKM matrix elements. In order to do that, it is necessary to provide a relation between the Hamiltonian matrix element and the SM Lagrangian parameters. Here for simplicity only the formalism for B^0 system is described. The extension to the other meson systems is trivial.

The most interesting term to calculate is M_{12}. It is basically the transition amplitude between a B and a $\overline{B}$ at zero momentum transfer. In terms of states with the conventional normalization we have

$$M_{12} = \frac{1}{2m_B}\langle B|\mathcal{O}|\overline{B}\rangle \quad .$$

The operator $\mathcal{O}$ is the one that can create a B and annihilate a $\overline{B}$. Remembering that a B^0 meson is made of a $\bar{b}$ and a d quark, while $\overline{B}$ from the conjugate combination, in terms of quark field annihilation and creation operators $\mathcal{O}$ is in the form

$$\mathcal{O} \sim (\bar{b}d)(\bar{b}d) \quad ,$$

where the explicit Dirac structure is not presented.

Since $\mathcal{O}$ is a FCNC operator, in the SM it cannot be generated at tree level and must represent an one-loop diagram, which is called "box diagram" for its shape, shown in figure 1.4. The calculation of the amplitude

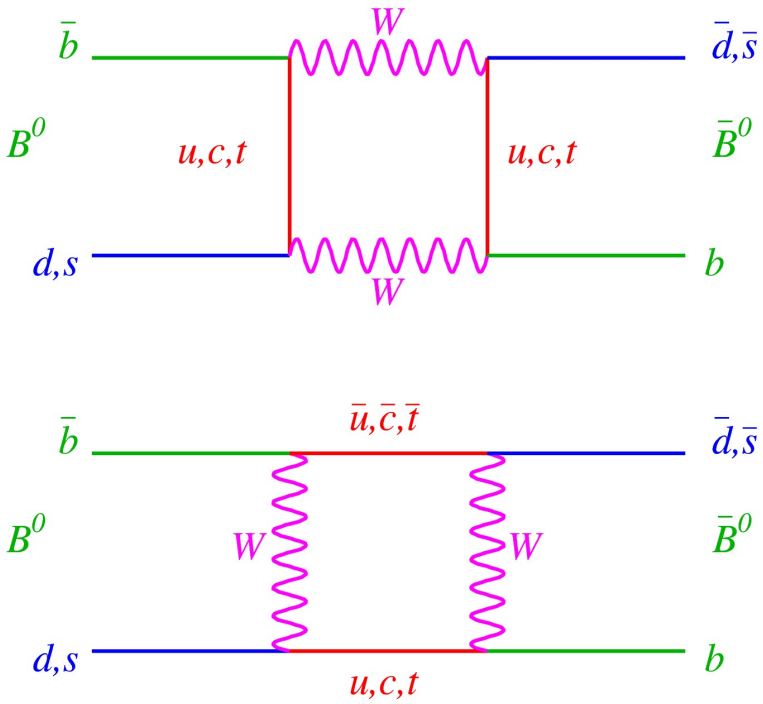


Figure 1.4: Box Feynman diagrams mediating the $B \leftrightarrow \overline{B}$ oscillation.

is straightforward and gives

$$M_{12} \propto \frac{g^4}{m_W^2} \langle B|(\overline{b_L}\gamma_\mu d_L)(\overline{b_L}\gamma^\mu d_L)|\overline{B}\rangle \sum_{i,j} V_{id}^* V_{ib} V_{jd}^* V_{jb} F(x_i, x_j) \quad ,$$

where

$$x_i \equiv \frac{m_i^2}{m_W^2} , \qquad i = u, c, t \quad ,$$

and the function F is known. The biggest theoretical uncertainties rise from the hadronic matrix element, $\langle B|(\overline{b_L}\gamma_\mu d_L)(\overline{b_L}\gamma^\mu d_L)|\overline{B}\rangle$. The difficulties are due to the fact that this operator creates and annihilate free b and d quarks, which is different from creating and annihilating B mesons. The lattice QCD calculations give an estimation of this quantity.

Thanks to GIM mechanism the quark mass independent components in the function F vanish, and the only term of the summation that is not suppressed is the one that contains two top quarks. It is possible to conclude that the Hamiltonian matrix element M_{12}, and thus the mass split, is proportional to the CKM matrix elements:

$$\Delta m_B \propto |V_{tb} V_{td}|^2 \quad ,$$

where the proportionality constant is known with an uncertainty at the 10% level. In figures 1.2 and 1.3 the constrains to the unitary triangle due to Δm_d and Δm_s measurements are shown.

Also the matrix element Γ_{12} has been calculated, considering the on-shell components of the box diagrams. But with on-shell mediator, QCD becomes important and so the theoretical uncertainties of Γ_{12} are larger than that of M_{12}.

1.4 CP violation

As obtained in section 1.1.1, the only Lagrangian term that in SM allows CP violation is the charge current interaction. This term, in facts, is the only one that could contain a physical phase (δ in the standard parametrization, η in the Wolfenstein parametrization). It is useful to define a phase parameter that does not depend on the parametrization. This parameter is the Jarlskog invariant, J_{CKM}, defined through

$$\mathcal{I}m(V_{ij} V_{kl} V_{il}^* V_{kj}^*) = J_{CKM} \sum_{m,n} \epsilon_{ikm} \epsilon_{jln} \quad ,$$

where ϵ_{ijk} is the Levi-Civita symbol. In terms of explicit parametrizations it is

$$J_{CKM} = c_{12} c_{23} c_{13}^2 s_{12} s_{23} s_{13} \sin\delta \approx \lambda^6 A^2 \eta$$

and, as mentioned in section 1.2, its modulus equals the double of the unitary triangle areas. Now, writing the request of a physical phase in terms of CKM parameters and quark masses, a necessary and sufficient condition for CP violation in the quark sector of the SM is

$$\Delta m_{tc}^2 \Delta m_{tu}^2 \Delta m_{cu}^2 \Delta m_{bs}^2 \Delta m_{bd}^2 \Delta m_{sd}^2 J_{CKM} \neq 0 \quad ,$$

where $\Delta m_{ij}^2 \equiv m_i^2 - m_j^2$.

1.4.1 Observing CPV

The CP symmetry relates particles with their anti-particles. Thus, probing CP violation is the same as observing different behavior between them. One example is a decay process with different decay rate:

$$\Gamma(A \to B) \neq \Gamma(\overline{A} \to \overline{B}) \quad ,$$

or different oscillating behavior:

$$\mathcal{P}(P^0 \to \overline{P^0}) \neq \mathcal{P}(\overline{P^0} \to P^0) \quad .$$

In order to measure the CP violation effects on the physical observables, it is necessary to study the processes with interference between more than one amplitude. In these cases, any phase difference between two or more amplitudes propagates in the observable quantities.

Two type of phase are required in the amplitudes.

- The first type ones are called "weak phases". They originate from the phases of the Lagrangian parameters so they are CP-odd, that is in the CP-conjugate process amplitude they appear with opposite sign. These phases are related to the charge current weak interactions.
- The other type ones are the "strong phases". They can appear even if the Lagrangian is real, since they come from possible contributions of intermediate on-shell states in the decay or oscillation process. These phases are the same both in an amplitude and in its CP-conjugate, so they are CP-even. They can originate from the trivial time evolution, $\exp(iEt)$, or from strong interaction rescattering.

Thus, for a process $B \to f$ with two interfering amplitudes, it is

$$A_f = |a_1| e^{i(\delta_1 + \phi_1)} + |a_2| e^{i(\delta_2 + \phi_2)} \quad , \tag{1.3}$$

$$\overline{A_f} = |a_1| e^{i(\delta_1 - \phi_1)} + |a_2| e^{i(\delta_2 - \phi_2)} \quad , \tag{1.4}$$

where $\overline{A_f}$ is the amplitude of the CP-conjugate process $\overline{B} \to \overline{f}$, δ_i and ϕ_i are the strong and weak phases, respectively. Furthermore for an oscillating neutral meson system it is

$$M_{12} = |M_{12}| e^{i\phi_M} \qquad \Gamma_{12} = |\Gamma_{12}| e^{i\phi_\Gamma} \quad .$$

Each phase appearing in these equation are convention dependent, but the phase differences $\delta_1 - \delta_2$, $\phi_1 - \phi_2$ and $\phi_M - \phi_\Gamma$ are physical.

Since there are two types of amplitudes in these kind of processes, decay and mixing, there are three way to observe CP violation, depending on how they interfere:

1. CP violation in decay, also called direct CP violation. In this case the interference is between two decay amplitude. The strong phase is due to rescattering.

2. CP violation in mixing, also called indirect CP violation. Here the absorptive (on-shell) and dispersive (off-shell) mixing amplitudes interfere. The strong phase is due to the time evolution of the oscillation.

3. CP violation in interference between mixing and decay. Here both oscillation and decay amplitudes interfere. The dominant contribution is provided by the dispersive mixing amplitude and the leading decay amplitude. The strong phase is due to the time evolution of the oscillation.

Direct CPV

This is the case when

$$|A(P \to f)| \neq |A(\overline{P} \to \overline{f})| \quad .$$

And it is useful to define a CPV estimator, the asymmetry, defined as

$$a_{CP} \equiv \frac{\Gamma(\overline{B} \to \overline{f}) - \Gamma(B \to f)}{\Gamma(\overline{B} \to \overline{f}) + \Gamma(B \to f)} = \frac{|\overline{A}/A|^2 - 1}{|\overline{A}/A|^2 + 1} \quad .$$

Equations (1.3) and (1.4) can be written as

$$A(P \to f) = A(1 + r\exp[i(\delta + \phi)]) \quad ,$$

$$A(\overline{P} \to \overline{f}) = A(1 + r\exp[i(\delta - \phi)]) \quad ,$$

where δ and ϕ are, respectively, the strong and weak phase difference and $r \leq 1$. Then the asymmetry is

$$a_{CP} = r\sin\phi\sin\delta \quad ,$$

which shows that in order to observe direct CP violation it is necessary a decay with at least two interfering amplitudes ($r \geq 0$) with different strong phases ($\delta \neq 0, \pi$) and different weak phases ($\phi \neq 0, \pi$).

Mixing related CPV

Considering the decay of an oscillating B meson system in a possible final state f, the decay amplitudes are

$$A_f \equiv A(B \to f)\,, \qquad \overline{A}_f \equiv A(\overline{B} \to f) \quad ,$$

and it is defined

$$\lambda_f \equiv \frac{q}{p}\frac{\overline{A}_f}{A_f} \quad .$$

The time evolution of the flavour eigenstates, without the restriction of CP conservation, is

$$|P^0(t)\rangle = g_+(t)|P^0\rangle + \frac{q}{p}g_-(t)|\overline{P^0}\rangle \quad ,$$

$$|\overline{P^0}(t)\rangle = g_+(t)|\overline{P^0}\rangle + \frac{p}{q}g_-(t)|P^0\rangle \quad ,$$

where g_+ and g_- are the same defined in equations (1.1) and (1.2). Also in this case a good CPV estimator is the asymmetry:

$$\mathcal{A}_f(t) \equiv \frac{\Gamma[\overline{B}(t) \to \overline{f}] - \Gamma[B(t) \to f]}{\Gamma[\overline{B}(t) \to \overline{f}] + \Gamma[B(t) \to f]} \quad .$$

Two f states are interesting. When f is a CP eigenstate ($f = \overline{f}$). In this case it is considered the situation without direct CP violation ($|\overline{A}_f| = |A_f|$) and with the approximations $\Delta\Gamma = 0$ and $|q/p| = 1$ (valid for B system). The asymmetry is then

$$\mathcal{A}_f(t) = \mathcal{I}m(\lambda_f)\sin(x\Gamma t) = \sin\left[\arg(\lambda_f)\right]\sin(\Delta m t) \quad .$$

This provides a good measure for the λ_f parameter. In the B systems it is related to the β angle of the unitary triangles by the relations

$$\mathcal{I}m(\lambda_{\psi K_S}) = \sin 2\beta\,, \qquad \mathcal{I}m(\lambda_{\psi\varphi}) = \sin 2\beta_s \quad .$$

The other interesting f state is a pure flavour state. In this case $\lambda = 0$ and it is possible to isolate the effects of CP violation in mixing, defined by

$$|q/p| \neq 1 \quad .$$

An important example is the charge-current semileptonic neutral meson decay $(P, \overline{P}) \to l^{\pm}X$. In this case the decay amplitudes obey $|A_{l^+X}| = |\overline{A}_{l^-X}|$ and $A_{l^-X} = \overline{A}_{l^+X} = 0$, at the leading order (one loop diagram), and then $\lambda = 0$. Thus, the semileptonic asymmetry is

$$\mathcal{A}_{\mathrm{SL}}(t) \equiv \frac{\Gamma[\overline{B}(t) \to l^+X] - \Gamma[B(t) \to l^-X]}{\Gamma[\overline{B}(t) \to l^+X] + \Gamma[B(t) \to l^-X]} = \frac{1 - |q/p|^4}{1 + |q/p|^4} \quad ,$$

which is time independent.

Chapter 2

The Large Hadron Collider and the Compact Muon Solenoid Experiment

The Large Hadron Collider (LHC) [19] is an accelerator located at the European Laboratory for Particle Physics Research (CERN) in Geneva. It has been conceived to collide proton beams at a centre-of-mass energy of $\sqrt{s} = 14$ TeV and a nominal instantaneous luminosity of $\mathcal{L} = 10^{34}$ cm^{-2} s^{-1}, representing a seven-fold increase in energy and a hundred-fold increase in integrated luminosity over the previous hadron collider experiments. Its main purpose is to search for rare processes like the production of Higgs or new particles with mass of 1 TeV and beyond. Two experiments have been installed around the LHC to pursue these results: ATLAS [20] and CMS [21]. Furthermore, the LHCb [22] experiment studies the properties of charm and beauty hadrons produced with large cross sections in the forward region in collisions at the LHC, and the ALICE [23] experiment analyses the data from relativistic heavy ion collisions to study the hadronic matter in extreme temperature and density conditions (i.e. high quark-gluon density).

2.1 The Large Hadron Collider

The LHC has been installed in the same tunnel which hosted the e^+e^- collider LEP (Large Electron-Positron collider). Accelerated electrons and positrons suffer large energy loss due to the synchrotron radiation, which is proportional to $E^4/(Rm^4)$, where E is the electron energy, m its mass and R the accelerator radius. To obtain energies of the order of TeV, at the fixed accelerator radius, only massive charged particles could have been used: protons and heavy nuclei. The energy loss is reduced by a factor $(2000)^4$ for a given fixed energy E for protons, respect to electrons. Another important aspect of the LHC is the collision rate. To produce a sufficient number of

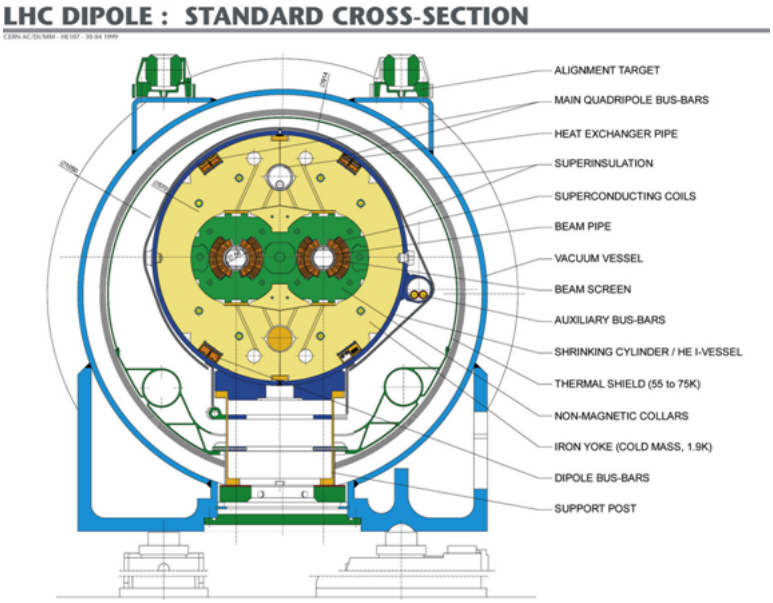


Figure 2.1: LHC dipole magnet section scheme.

rare processes, the collision rate needs to be very high. Beam protons are collected in packets called bunches. The collision rate is proportional to the instantaneous luminosity of the accelerator, defined as:

$$\mathcal{L} = \frac{f k n_p^2}{4\pi\sigma_x\sigma_y} \quad ,$$

where f is the bunch revolution frequency, k the number of bunches, n_p the number of protons per bunch, σ_x and σ_y their transverse dispersion along the x and y axis. At the nominal 14 TeV LHC conditions ($\mathcal{L} = 10^{34}$ cm^{-2} s^{-1}) the parameter values are: $k = 2808$, $n_p = 1.5{\cdot}10^{11}$ and $\sigma_x\sigma_y = 16.6\,\mu$m^2 (with $\sigma_z = 7.6$ cm along the beam). The integrated luminosity is defined as $L = \int \mathcal{L} dt$. For comparison we can consider the Tevatron accelerator at Fermilab, which produced proton-antiproton collisions since 1992. Its centre of mass energy was 1.8 TeV until 1998 and 1.96 TeV since 2001. To increase $\mathcal{L}$ by two orders of magnitude, protons are injected in both LHC beams. The antiprotons, in fact, are obtained by steering proton beams onto a nickel target and represent only a small fraction of the wide range of secondary particles produced in this interactions, thus have a production rate lower than the proton one.

The LHC is composed by 1232 super-conducting dipole magnets each 15 m long, providing a 8.3 T magnetic field to let the beams circulate inside their trajectories along the 27 km circumference. Two vacuum pipes are used to let beams circulate in opposite directions. A scheme representing

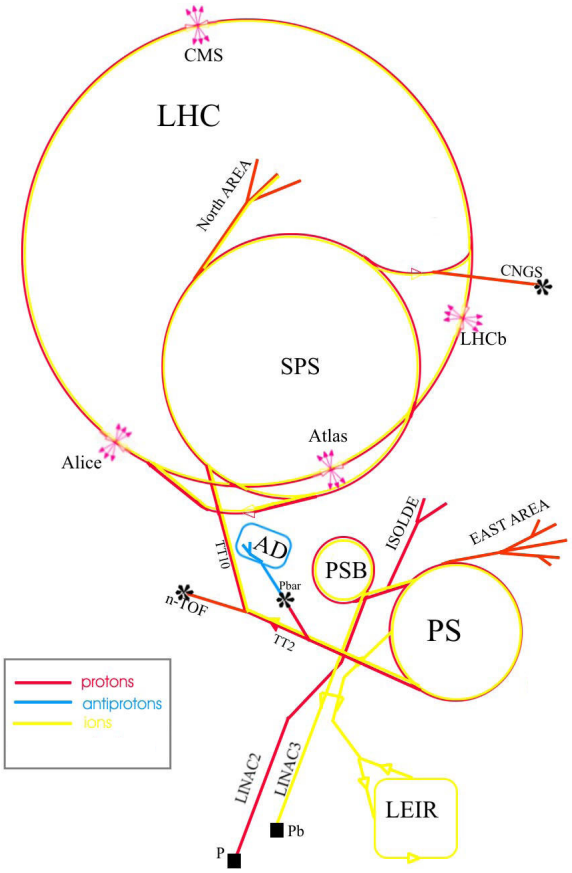


Figure 2.2: Scheme representing the CERN accelerator complex.

the transverse dipole magnet section is represented in figure 2.1. More than 8000 other magnets are utilized for the beam injection, their collimation, trajectory correction, crossing. All the magnets are kept cool by superfluid helium at 1.9 K temperature. The beams are accelerated from 450 GeV (the injection energy from the SPS) to 7 TeV with 16 Radio Frequency cavities (8 per beam) which raise the beam energy by 16 MeV each round with an electric field of 5 MV/m oscillating at 400 MHz frequency.
Before the injection into the LHC, the beams are produced and accelerated by different components of the CERN accelerator complex. Being produced from ionized hydrogen atoms, protons are accelerated by the linear accelerator LINAC, Booster and the Proton Synchrotron (PS) up to 26 GeV energy, the bunches being separated by 25 ns each. The beams are then injected into the Super Proton Synchrotron (SPS) where they are accelerated up to 450 GeV. They are then finally transferred to the LHC and accelerated up to 7 TeV energy per beam. The CERN accelerator complex is illustrated in figure 2.2.

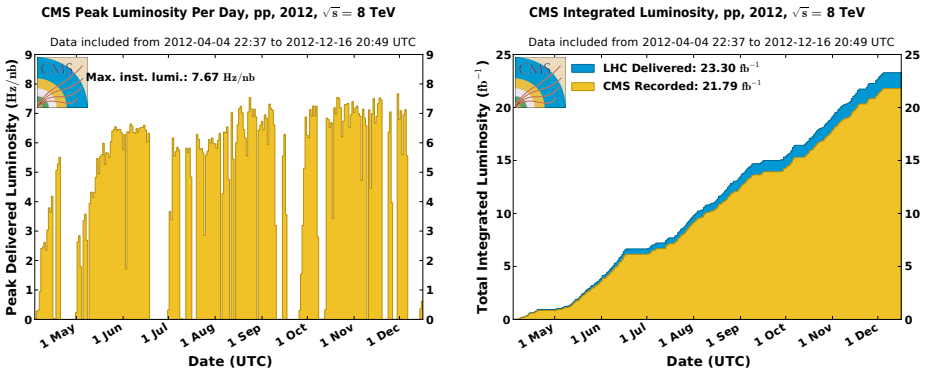


Figure 2.3: LHC performance in 2012. Left: CMS detected peak luminosity; right: CMS detected integrated luminosity.

The LHC started its operations in December 2009 with centre of mass energy for the proton-proton collision $\sqrt{s} = 0.9$ TeV. The centre of mass energy was set to $\sqrt{s} = 7$ TeV in the 2010 and 2011 runs and raised to $\sqrt{s} = 8$ TeV in the 2012 runs. Here are reported the CMS detected peak and integrated luminosities for proton-proton runs. In 2010 the peak luminosity reached $\mathcal{L} = 203.80\,\mathrm{Hz}/\mu\mathrm{b}$ and the integrated luminosity has been $L = 40.76\,\mathrm{pb}^{-1}$, while during 2011 the peak luminosity increased to $\mathcal{L} = 4.02\,\mathrm{Hz/nb}$ and the integrated luminosity has been $L = 5.55\,\mathrm{fb}^{-1}$. In the 2012 runs the peak luminosity reached $\mathcal{L} = 7.67\,\mathrm{Hz/nb}$ and the integrated luminosity has been $L = 21.79\,\mathrm{fb}^{-1}$, as graphically summarized in figure 2.3.

2.2 CMS Experiment

The Compact Muon Solenoid [21] is a general purpose detector situated at interaction point 5 of the CERN Large Hadron Collider. It is designed around a 4 T solenoidal magnetic field provided by the largest superconducting solenoid ever built. The structure of CMS is shown in figure 2.4, where particular emphasis is put on the volumes of the different subsystems: the Silicon Pixel Detector, the Silicon Strip Tracker, the Electromagnetic and Hadronic Calorimeters, and Muon Detectors.

We can briefly summarize the aims of the CMS detector [24]. They are mainly:

- search for SM and MSSM Higgs boson decaying into photons, b quarks, τ leptons, W and Z bosons,
- search for additional heavy neutral gauge bosons predicted in many superstring-inspired theories or Great Unification Theories and decaying to muon pairs,

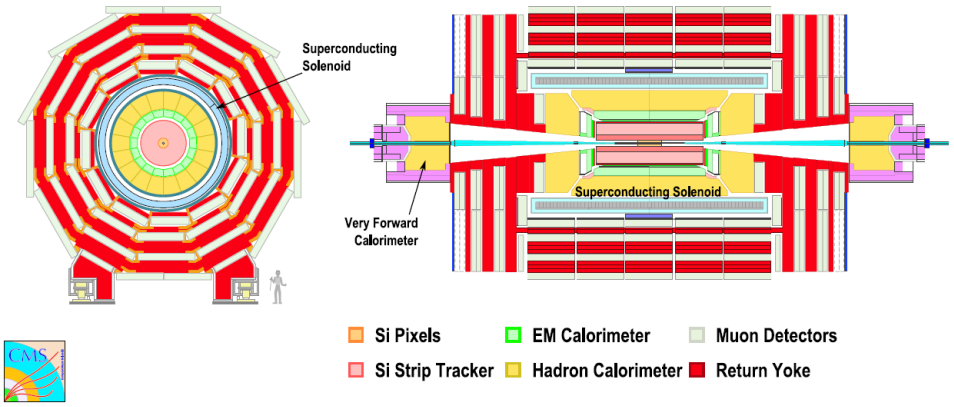


Figure 2.4: Transverse (left) and longitudinal (right) cross sections of the CMS detector showing the volumes of the different detector subsystems. The transverse cross section is drawn for the central barrel, coaxial with the beam line, while complementary end-caps are shown in the longitudinal view.

- search for new Physics in various topologies: multilepton events, multijet events, events with missing transverse energy[1] or momentum, any combination of the three above,
- study of the B-hadron rare decay channels (like $B^0_{(s)} \to \mu\mu$) and of CP violation in the decay of the B mesons (like $B^0_s \to J/\psi\phi \to \mu^+\mu^- K^+K^-$),
- search for $B^0 \to \mu^+\mu^-$ decays,
- study of QCD and jet physics at the TeV scale,
- study of top quark and EW physics.

CMS has been therefore designed as a multipurpose experiment, with particular focus on muon, photon, and displaced tracks reconstruction. Superb performances have been achieved overall, in particular in:

- primary and secondary vertex localization,
- charged particle momentum resolution and reconstruction efficiency in the tracking volume,
- electromagnetic energy resolution,
- isolation of leptons and photons at high luminosities,

[1] Missing transverse energy $\not\!\!E_T$ is the amount of energy which must be added to balance the modulus of the vector sum of the projections of the track momenta and calorimeter clusters in the plane perpendicular to beam axis. Its direction is opposite to this vector sum directions.

- measurement of the direction of photons, rejection of $\pi^0 \to \gamma\gamma$,
- diphoton and dielectron mass resolution $\sim 1\%$ at 100GeV,
- measurement of the missing transverse energy $\not{E}_T$ and dijet mass with high resolution,
- muon identification over a wide range of momenta,
- dimuon mass resolution $\sim 1\%$ at 100 GeV,
- unambiguously determining the charge of muons with p_T up to 1 TeV,
- triggering and offline tagging of τ leptons and b jets.

The reference frame used to describe the CMS detector and the collected events has its origin in the geometrical centre of the solenoid. Different types of global coordinates measured with respect to the origin[2] are used:

- cartesian coordinate system, $\hat{x}$ axis points towards the centre of LHC, $\hat{y}$ points upwards, perpendicular to LHC plane, while $\hat{z}$ completes the right-handed reference,
- polar coordinate system, directions are defined with an azimuthal angle $\tan\phi = y/x$ and a polar angle $\tan\theta = \rho/z$, where $\rho = \sqrt{x^2 + y^2}$,
- polar coordinate system, with instead of the polar angle the rapidity y and the pseudorapidity η, obtained for any particle from

$$y = \frac{1}{2}\ln\left(\frac{E + p_z}{E - p_z}\right) \quad ,$$

$$\eta = -\ln\left(\tan\frac{\theta}{2}\right) \quad ,$$

  where E is the particle energy and p_z the component of its momentum along the beam direction.

2.2.1 Magnet

The whole CMS detector is designed around a $\sim$ 4 T superconducting solenoid [25] 12.5 m long and with inner radius of 3 m. The solenoid thickness is 3.9 radiation lengths and it can store up to 2.6 GJ of energy.

The field is closed by a 10^4 t iron return yoke made of five barrels and two end-caps, composed of three layers each. The yoke is instrumented

[2]Global coordinates are measured in the CMS reference frame while local coordinates are measured in the reference frame of a specific sub-detector or sensitive element.

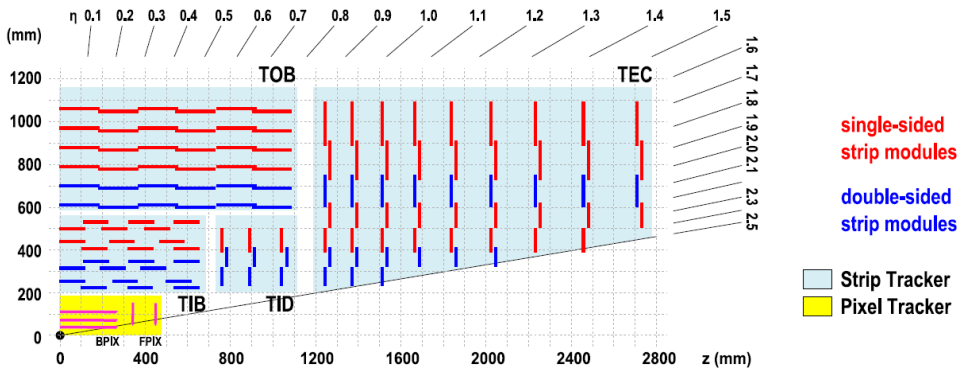


Figure 2.5: Layout of the CMS silicon tracker showing the relative position of hybrid pixels, single-sided strips and double-sided strips. Figure from [21].

with four layers of muon stations. The coil is cooled down to 4.8 K by a helium refrigeration plant, while insulation is given by two pumping stations providing vacuum on the 40 m^3 of the cryostat volume.

The magnet was designed in order to reach precise measurement of muon momenta. A high magnetic field is required to keep a compact spectrometer capable to measure 100 GeV track momentum with percent precision. A solenoidal field was chosen because it keeps the bending in the transverse plane, where an accuracy better than 20 μm is achieved in vertex position measurements. The size of the solenoid allows efficient track reconstruction up to a pseudorapidity of 2.4. The inner radius is large enough to accommodate both the Silicon Tracking System and the calorimeters. During the 2012 acquisitions the magnet was operated at 3.8 T.

2.2.2 Tracking System

The core of CMS is a Silicon Tracking System [26] with 2.5 m diameter and 5.8 m length, designed to provide a precise and efficient measurement of the trajectories of charged particles emerging from LHC collisions and reconstruction of secondary vertices.

The CMS Tracking System is composed of both silicon Pixel and Strip Detectors, as shown in figure 2.5. The Pixel Detector consists of 1440 pixel modules arranged in three barrel layers and two disks in each end-cap as in figure 2.6. The Strip Detector consists of an inner tracker with four barrel layers and three end-cap disks and an outer tracker with six barrel layers and nine end-cap disks, housing a total amount of 15148 strip modules of both single-sided and double-sided types. Its active silicon surface of about 200 m^2 makes the CMS tracker the largest silicon tracker ever built.

The LHC physics programme requires high reliability, efficiency and precision in reconstructing the trajectories of charged particles with transverse

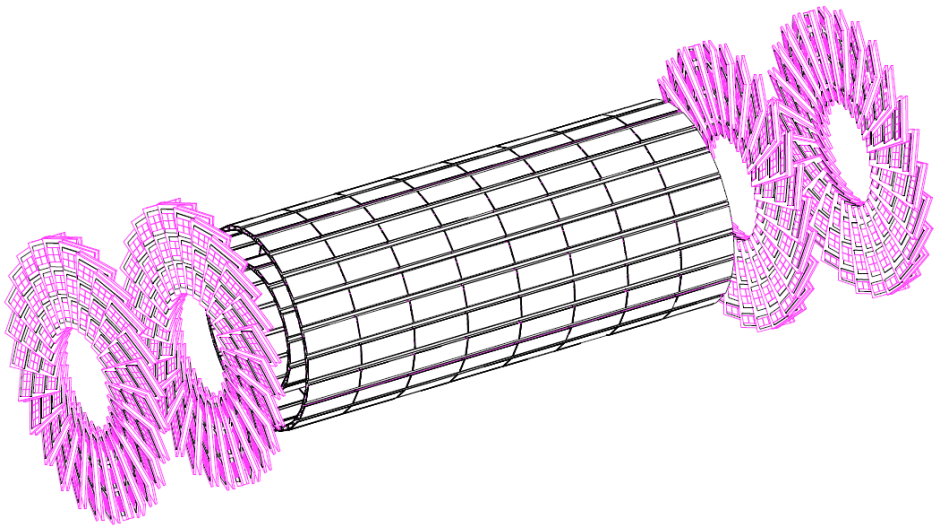


Figure 2.6: Layout of the current CMS Pixel Detector. Figure from [26].

momentum larger than 1 GeV in the pseudorapidity range $|\eta| < 2.5$. Heavy quark flavours can be produced in many of the interesting channels and a precise measurement of secondary vertices is therefore needed. The tracker completes the functionalities of ECAL and Muon System to identify electrons and muons. Also hadronic decays of tau leptons need robust tracking to be identified in both the one-prong and three-prongs topologies. Tracker information is heavily used in the High Level Trigger of CMS to help reducing the event collection rate from the 40 MHz of bunch crossing to the 100 Hz of mass storage.

Silicon Pixel Detector

The large number of particles produced in 25 pile-up events[3], at nominal LHC luminosity, results into a hit rate density of 1 MHz mm^{-2} at 4 cm from the beamline, decreasing down to 3 kHz mm^{-2} at a radius of 115 cm. Pixel detectors are used at radii below 10 cm to keep the occupancy below 1%. The chosen size for pixels, $0.100 \times 0.150\,\text{mm}^2$ in the transverse and longitudinal directions respectively, leads to an occupancy of the order of 10^{-4}. The layout of the Pixel Detector consists of a barrel region (BPIX), with three barrels at radii of 4.4, 7.3 and 10.2 cm, complemented by two disks on each side (FPIX), at 34.5 and 46.5 cm from the nominal interaction point. This layout provides about 66 million pixels covering a total area of about 1 m^2 and measuring three high precision points on each charged particle trajectory up to $|\eta| = 2.5$. Detectors in FPIX disks are tilted by 20°

[3]Events that occur in the same bunch crossing.

in a turbine-like geometry to induce charge sharing and achieve a spatial resolution of about 20 μm.

Silicon Strip Tracker

In the inner Strip Tracker, which is housed between radii of 20 and 55 cm, the reduced particle flux allows a typical cell size of $0.080 \times 100\,\mathrm{mm}^2$, resulting in a 2% occupancy per strip at design luminosity. In the outer region, the strip pitch is increased to $0.180 \times 250\,\mathrm{mm}^2$ together with the sensor thickness which scales from 0.320 mm to 0.500 mm. This choice compensates the larger capacitance of the strip and the corresponding larger noise with the possibility to achieve a larger depletion of the sensitive volume and a higher charge signal.

The Tracker Inner Barrel and Disks (TIB and TID) deliver up to 4 (r, ϕ) measurements on a trajectory using 0.320 mm thick silicon strip sensors with strips parallel to the beamline. The strip pitch is 0.080 mm in the first two layers and 0.120 mm in the other two layers, while in the TID the mean pitch varies from 0.100 mm to 0.141 mm. Single point resolution in the TIB is 0.023 mm with the finer pitch and 0.035 mm with the coarser one. The Tracker Outer Barrel (TOB) surrounds the TIB/TID and provides up to 6 $r - \phi$ measurements on a trajectory using 0.500 mm thick sensors. The strip pitch varies from 0.183 mm in the four innermost layers to 0.122 mm in the outermost two layers, corresponding to a resolution of 0.053 mm and 0.035 mm respectively. Tracker End-Caps (TEC) enclose the previous sub-detectors at 124cm $< |z| <$ 282cm with 9 disks carrying 7 rings of microstrips, 4 of them are 0.320 mm thick while the remaining 3 are 0.500 mm thick. TEC strips are radially oriented and their pitch varies from 0.097 mm to 0.184 mm.

As shown in figure 2.5, the first two layers and rings of TIB, TID and TOB, as well as three out of the TEC rings, carry strips on both sides with a stereo angle of 100 milliradians to measure the other coordinate: z in barrels and r in rings. This layout ensures 9 hits in the silicon Strip Tracker in the full acceptance range $|\eta| < 2.4$, and at least four of them are two-dimensional. The total area of Strip Tracker is about 198 m^2 read out by 9.3 million channels.

Trajectory Reconstruction

Due to the magnetic field charged particles travel through the tracking detectors on a helical trajectory which is described by 5 parameters: the curvature κ, the track azimuthal angle ϕ, the pseudorapidity η, the signed transverse impact parameter d_0 and the longitudinal impact parameter z_0. The transverse (longitudinal) impact parameter of a track is defined as the transverse (longitudinal) distance of closest approach of the track to the primary ver-

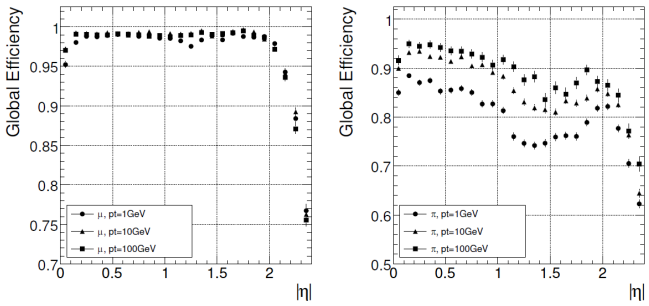


Figure 2.7: Global track reconstruction effciency as a function of track pseudorapidity for muons (left) and pions (right) of transverse momenta of 1, 10 and 100 GeV. Figures from [21].

tex. The main standard algorithm used in CMS for track reconstruction is the Combinatorial Track Finder (CFT) algorithm [27] which uses the reconstructed positions of the passage of charged particles in the silicon detectors to determine the track parameters. The CFT algorithm proceeds in three stages: track seeding, track finding and track fitting. Track candidates are best seeded from hits in the pixel detector because of the low occupancy, the high efficiency and the unambiguous two-dimensional position information. The track finding stage is based on a standard Kalman filter pattern recognition approach [28] which starts with the seed parameters. The trajectory is extrapolated to the next tracker layer and compatible hits are assigned to the track on the basis of the χ^2 between the predicted and measured positions. At each stage the Kalman filter updates the track parameters with the new hits.

The tracks are assigned a quality based on the χ^2 and the number of missing hits and only the best quality tracks are kept for further propagation. Ambiguities between tracks are resolved during and after track finding. In case two tracks share more than 50% of their hits, the lower quality track is discarded. For each trajectory the finding stage results in an estimate of the track parameters. However, since the full information is only available at the last hit and constraints applied during trajectory building can bias the estimate of the track parameters, all valid tracks are refitted with a standard Kalman filter and a second filter (smoother) running from the exterior towards the beam line. The expected performance of the track reconstruction is shown in figure 2.7 for muons, pions and hadrons. The track reconstruction efficiency for high energy muons is about 99% and drops at $|\eta| > 2.1$ due to the reduced coverage of the forward pixel detector. For pions and hadrons the efficiency is in general lower because of interactions with the material in the tracker.

The material budget is shown in figure 2.8 as a function of pseudorapidity, with the different contributions of sub-detectors and services.

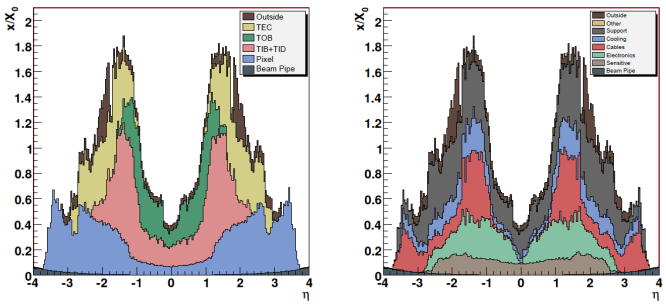


Figure 2.8: Material budget of the current CMS Tracker in units of radiation length X_0 as a function of the pseudorapidity, showing the different contribution of sub-detectors (left) and functionalities (right). Figures from [21].

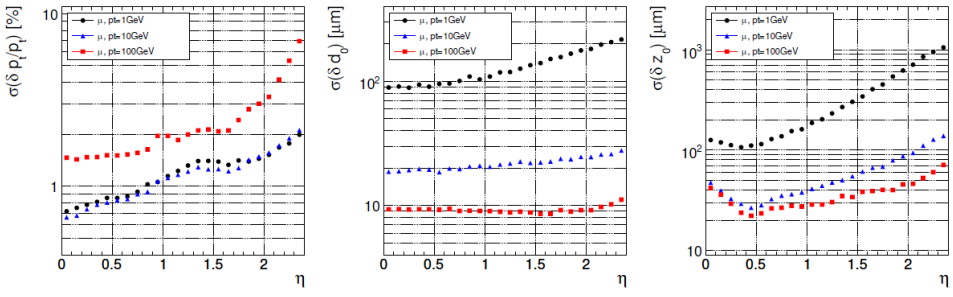


Figure 2.9: Resolution of several track parameters as a function of track pseudorapidity for single muons with transverse momenta of 1, 10 and 100 GeV: transverse momentum (left), transverse impact parameter (middle) and longitudinal impact parameter (right). Figures from [21].

The performance of the Silicon Tracker in terms of track reconstruction efficiency and resolution, of vertex and momentum measurement, are shown in figure 2.7 and figure 2.9 respectively. The first one, in particular, shows the difference in reconstruction efficiency for muons and pions, due to the larger interaction cross section of pions, which cannot be assumed to be minimum-ionizing particles and therefore are much more degraded by the amount of material.

Vertex Reconstruction

The reconstruction of interaction vertices allows CMS to reject tracks coming from pile-up events. The primary vertex reconstruction is a two-step process. Firstly the reconstructed tracks are grouped in vertex candidates and their z coordinates at the beam closest approach point are evaluated, retaining only tracks with impact parameter respect to the vertex candidate less than 3 cm. Vertices are then reconstructed through a recursive method for parameter estimation through a Kalman filter [29] algorithm. For a given event, the primary vertices are ordered according to the total transverse momentum of the associated tracks, $\sum p_T$. The vertex reconstruction efficiency is very close to 100% and the position resolution is of the order of $\mathcal{O}(10)\,\mu$m in all directions.

It is also possible to reconstruct the secondary vertices, for example those from b-quark decays. The secondary vertex reconstruction uses tracks associated to jets applying further selection cuts: the transverse impact parameter of the tracks must be greater than 100 μm, to avoid tracks coming from the primary vertex, and the longitudinal impact parameter below 2 cm, to avoid tracks from pile-up events.

2.2.3 Muon Spectrometer

Detection of muons at CMS exploits different technologies and is performed by a "Muon System" rather than a single detector [30]. Muons are the only particles able to reach the external muon chambers with a minimal energy loss when traversing the calorimeters, the solenoid and the magnetic field return yoke. Muons can provide strong indication of interesting signal events and are natural candidates for triggering purposes. The CMS Muon System was designed to cope with three major functions: robust and fast identification of muons, good resolution of momentum measurement and triggering.

The Muon System is composed of three types of gaseous detectors, located inside the empty volumes of the iron yoke, and therefore arranged in barrel and end-cap sections. The coverage of Muon System is shown in figure 2.10.

In the barrel region the neutron-induced background is small and the

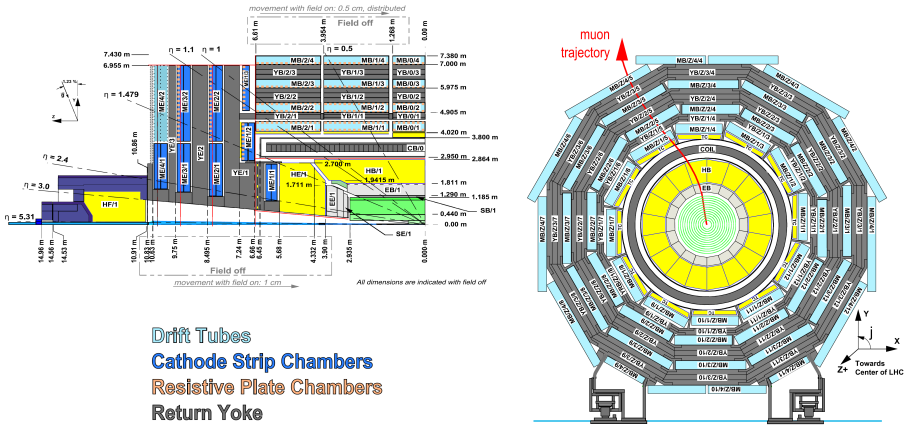


Figure 2.10: Transverse and longitudinal cross sections of the CMS detector showing the Muon System with particular emphasis on the different technologies used for detectors; the ME/4/2 CSC layers in the end-cap were included in the design but are not currently installed. Figures from [21].

muon rate is low; moreover, the field is uniform and contained in the yoke. For these reasons, standard drift chambers with rectangular cells are used. The barrel Drift Tubes (DT) cover the $|\eta| < 1.2$ region, are divided in five wheels in the beam direction and are organized in four stations housed among the yoke layers. The first three stations contain 12 cell planes, arranged in two superlayers providing measurement along $r\phi$ and one superlayerlayer measuring along z, each of them containing four layers. The fourth station provides measurement only in the transverse plane.

Both the muon rates and backgrounds are high in the forward region, where the magnetic field is large and non uniform. The choice for muon detectors fell upon cathode strip chambers (CSC) because of their fast response time, fine segmentation and radiation tolerance. Each end-cap is equipped with four stations of CSCs. The CSCs cover the $0.9 < |\eta| < 2.4$ pseudorapidity range. The cathode strips are oriented radially and provide precise measurement in the bending plane, the anode wires run approximately perpendicular to the strips and are read out to measure the pseudorapidity and the beam-crossing time of a muon. The muon reconstruction efficiency is typically $95 - 99\%$ except for the regions between two barrel DT wheels or at the transition between DTs and CSCs, where the efficiency drops.

Both the DTs and CSCs can trigger on muons with a Level 1 p_T (see section 2.2.5) resolution of 15% and 25%, respectively. Additional trigger-dedicated muon detectors were added to help measured the correct beam-crossing time. These are Resistive Plate Chambers (RPC), gaseous detector operated in the avalanche mode, which can provide independent and fast trigger with high segmentation and sharp p_T threshold over a large portion of the pseudorapidity range. The overall p_T resolution on muons is shown

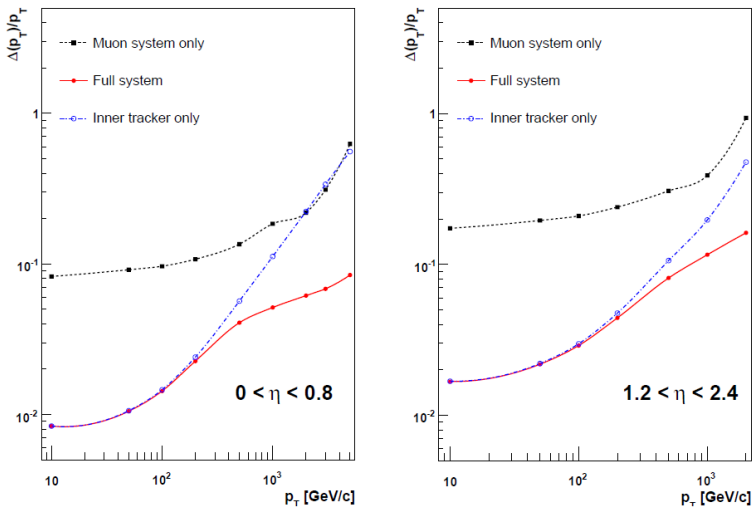


Figure 2.11: Resolution on p_T measurement of muons with the Muon System, the Silicon Tracker or both, in the barrel (left) and end-caps (right). Figures from [21].

in figure 2.11, with emphasis on the different contribution from the Muon System and the Silicon Tracker.

Muon Reconstruction

Muon detection and reconstruction play a key role in the CMS physics program, both for the discovery of New Physics and for precision measurements of SM processes. CMS has been designed for a robust detection of muons over the entire kinematic range of the LHC and in a condition of very high background. The muon system allows an efficient and pure identification of muons, while the inner tracker provides a very precise measurement of their properties. An excellent muon momentum resolution is made possible by the high-field solenoidal magnet. The steel flux return yoke provides additional bending power in the spectrometer, and serves as hadron absorber to facilitate the muon identification. Several muon reconstruction strategies are available in CMS, in order to fulfil the specific needs of different analyses. The muon reconstruction consists of three main stages:

1. local reconstruction: in each muon chamber, the raw data from the detector read-out are reconstructed as individual points in space; in CSC and DT chambers, such points are then fitted to track segments;

2. stand-alone reconstruction: points and segments in the muon spectrometer are collected and fitted to tracks, referred to as "stand-alone muon tracks";

3. global reconstruction: stand-alone tracks are matched to compatible tracks in the inner tracker and a global fit is performed using the whole set of available measurements: the resulting tracks are called "global muon tracks".

Muon identification represents a complementary approach with respect to global reconstruction: it starts from the inner tracker tracks and flags them as muons by searching for matching segments in the muon spectrometer. The muon candidates produced with this strategy are referred to as "tracker muons". After the completion of both algorithms, the reconstructed stand-alone, global and tracker muons are merged into a single software object, with the addition of further information, like the energy collected in the matching calorimeter towers. This information can be used for further identification, in order to achieve a balance between efficiency and purity of the muon sample.

2.2.4 Calorimetry

Identification of electrons, photons, and hadrons relies on accurate calorimetry, which is a destructive measurement of the energy of a particle. As in most of the particle physics experiments, a distinction is made between electromagnetic calorimetry and hadron calorimetry. Electromagnetic calorimetry is based on the production of EM showers inside a high-Z absorber, while hadron calorimetry measures the effects of hadron inelastic scattering with heavy nuclei, including production of photons from neutral pions and muons, and neutrinos from weak decays. Calorimetry must be precise and hermetic also to measure any imbalance of momenta in the transverse plane which can signal the presence of undetected particles such as high-p_T neutrinos.

The electromagnetic calorimeter of CMS, ECAL, is a homogeneous calorimeter, where the absorber material is the same as the sensitive one [31]. ECAL is composed of 61200 lead tungstate ($PbWO_4$) crystals in the barrel region and 7324 crystals in the end-caps, as shown in figure 2.12. The crystal cross-section is $22 \times 22\,\mathrm{mm}^2$ at the front face, while the length is 230 mm. End-caps are equipped with a preshower detector. Lead tungstate was chosen because of its high density, 8.28 g cm^{-3}, short radiation length, 0.89 cm, and small Molière radius, 2.2 cm. This way, the calorimeter can be kept compact with fine granularity, while scintillation and optical properties of $PbWO_4$ make it fast and radiation tolerant. Signal transmission exploits total internal reflection. Scintillation light detection relies on two different technologies. Avalanche photodiodes (APD) are used in the barrel region, mounted in pairs on each crystals, while vacuum phototriodes (VPT) are used in the end-caps. The preshower detector is a sampling calorimeter composed of lead radiators and silicon strips detectors, and it is used to identify neutral pions in the forward region. The nominal energy resolution,

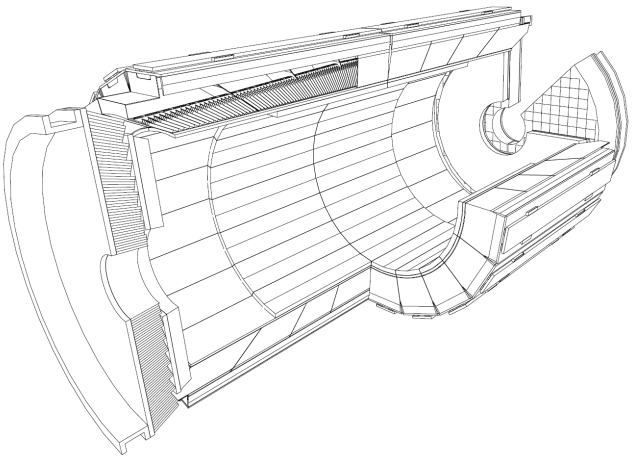


Figure 2.12: Cut-away view of the CMS ECAL showing the hierarchical structure of crystals arranged in supercystals and modules and the orientation of crystals whose major axis is always directed to the origin of the reference frame.

measured with electron beams having momenta between 20 and 250 GeV, is

$$\left(\frac{\sigma_E}{E}\right)^2 = \left(\frac{2.8\%}{\sqrt{E}}\right)^2 + \left(\frac{0.12}{E}\right)^2 + (0.30\%)^2 \quad ,$$

where all the energies are in GeV and the different contributions are respectively: the stochastic one (S), due to fluctuations in the lateral shower containment and in the energy released in the preshower, that due to electronics (N), digitization and pile-up, and the constant term (C), due to intercalibration errors, energy leakage from the back of the crystal and non-uniformity in light collection.

The hadron calorimeter of CMS, HCAL, is a sampling calorimeter employed for the measurement of hadron jets and neutrinos or exotic particles resulting in apparent missing transverse energy [32]. A longitudinal view of HCAL is shown in figure 2.13. The hadron calorimeter size is constrained in the barrel region, $|\eta| < 1.3$, by the maximum radius of ECAL and the inner radius of the solenoid coil. Because of this, the total amount of the absorber material is limited and an outer calorimeter layer is located outside of the solenoid to collect the tail of the showers. The pseudorapidity coverage is extended in the $3 < |\eta| < 5.2$ by forward Cherenkov-based calorimeters. The barrel part, HB, consists of 36 wedges, segmented into 4 azimuthal sectors each, and made out of flat brass absorber layers, enclosed between two steel plates and bolted together without any dead material on the full radial extent. There are 17 active plastic scintillator tiles interspersed between the stainless steel and brass absorber plates, segmented in pseudorapidity

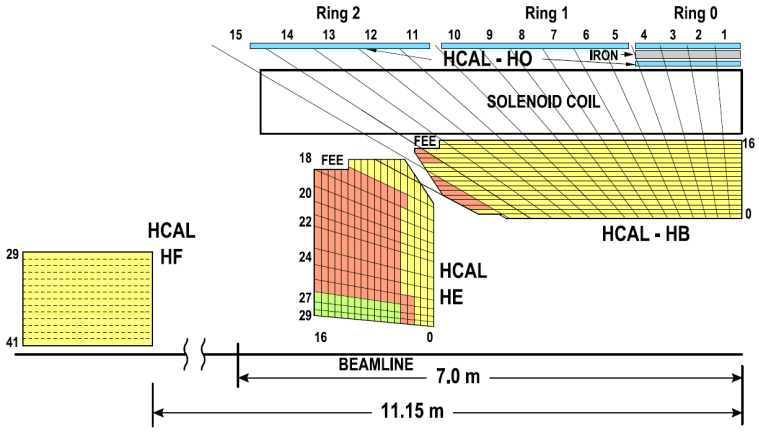


Figure 2.13: Cross section of the CMS HCAL showing the tower segmentation. Figure from [32].

to provides an overall granularity of $\Delta\phi \times \Delta\eta = 0.087 \times 0.087$. The same segmentation is maintained in end-cap calorimeters, HE, up to $|\eta| < 1.6$, while it becomes two times larger in the complementary region. The maximum material amount in both HB and HE corresponds to approximately 10 interaction lengths λ_I. The energy resolution on single electron and hadron jets is shown in figure 2.14.

2.2.5 Trigger and Data Acquisition

High bunch crossing rates and design luminosity at LHC correspond to approximately 20–25 superimposed events every 25 ns, for a total of 10^9 events per second. The large amount of data associated to them is impossible to store and process, therefore a dramatic rate reduction has to be achieved. This is obtained with two steps: the Level 1 Trigger [33] and the High Level Trigger, HLT [34].

The Level 1 Trigger is based on custom and programmable electronics, while HLT is a software system implemented on a $\sim$ 1000 commercial processors farm. The maximum allowed output rate for Level 1 Trigger is 100 kHz, which should be even kept lower, about 30 kHz, for safe operation. Level 1 Trigger uses rough information from coarse segmentation of calorimeters and Muon Detectors and holds the high-resolution data in a pipeline until acceptance/rejection decision is made. HLT exploits the full amount of collected data for each bunch crossing accepted by Level 1 Trigger and is capable of complex calculations such as the off-line ones. HLT algorithms are those expected to undergo major changes in time, particularly with increasing luminosity. Configuration and operation of the trigger components are handled by a software system called Trigger Supervisor.

The Level 1 Trigger relies on local, regional and global components. The

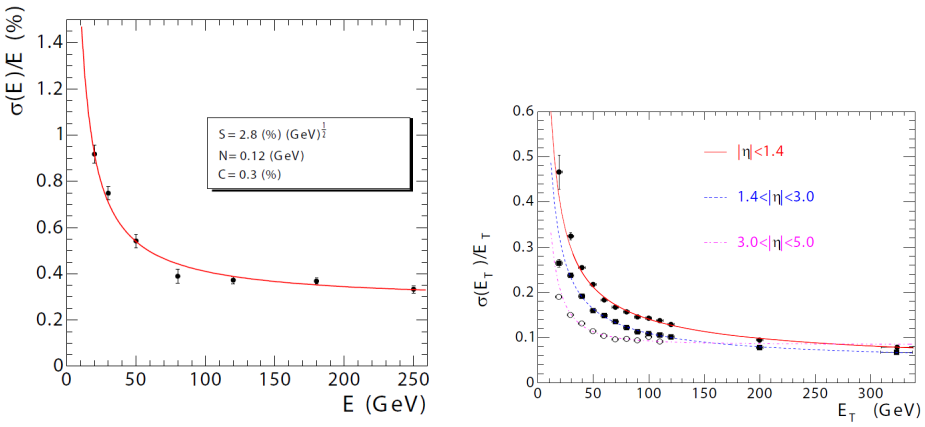


Figure 2.14: Left: ECAL energy resolution as a function of the electron energy as measured from a beam test. The energy was measured in a 3×3 crystals array with the electron impacting the central one. The stochastic, noise and constant terms are given. Right: the jet transverse energy resolution as a function of the transverse energy for barrel jets, end-cap jets and very forward jets reconstructed with an iterative cone algorithm with cone radius $R = 0.5$. Figures from [21].

Global Calorimeter and Global Muon Triggers determine the highest-rank calorimeter and muon objects across the entire experiment and transfer them to the Global Trigger, the top entity of the Level 1 hierarchy. The latter takes the decision to reject an event or to accept it for further evaluation by the HLT. The total allowed latency time for the Level 1 Trigger is $3.2\,\mu$s. A schematic representation of the Level 1 Trigger data flow is presented in figure 2.15.

Muon Trigger

All Muon Detectors – DT, CSC and RPC – contribute to the Trigger. Barrel DTs provide Local Trigger in the form of track segments in ϕ and hit patterns in η. End-cap CSCs provide 3-dimensional track segments. Both CSCs and DTs provide also timing information to identify the bunch crossing corresponding to candidate muons. The Local DT Trigger is implemented in custom electronics. BTIs, Bunch and Track Identifiers, search for coincidences of aligned hits in the four equidistant planes of staggered drift tubes in each chamber superlayer. From the associated hits, track segments defined by position and angular direction are determined. TRACOs, Track Correlators, attempt to correlate track segments measured in the two ϕ superlayers of each DT chamber, enhancing the angular resolution and producing a quality hierarchy.

The requirement of robustness implies redundancy, which introduces,

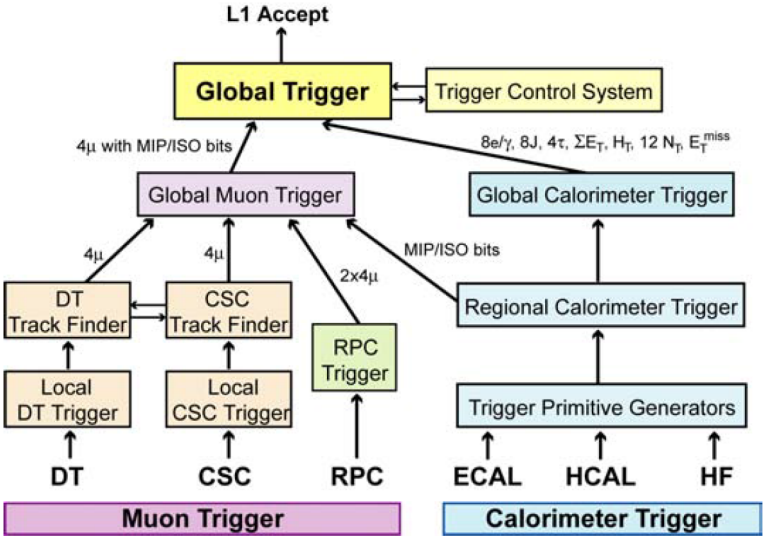


Figure 2.15: Schematic representation of the Level 1 Trigger data flow.

however, a certain amount of noise or duplicate tracks giving rise to false Triggers. Therefore the BTIs, the TRACOs and the different parts of the Local Trigger contain complex noise and ghost reduction mechanisms. The position, transverse momentum and quality of tracks are coded and transmitted to the DT regional Trigger, called the Drift Tube Track Finder (DTTF), through high-speed optical links.

The Global Muon Trigger (GMT) combines the information from DTs, CSCs and RPCs, achieving an improved momentum resolution and efficiency compared to the stand-alone systems. It also reduces the Trigger rate and suppresses backgrounds by making use of the complementarity and redundancy of the three Muon Systems. The Global Muon Trigger also exploits MIP/ISO bits[4] from the Regional Calorimeter Trigger. A muon is considered isolated if its energy deposit in the calorimeter region from which it emerged is below a defined threshold. DT and CSC candidates are first matched with barrel and forward RPC candidates based on their spatial coordinates. If a match is possible, the kinematic parameters are merged. Several merging options are possible and can be selected individually for all track parameters, taking into account the strengths of the individual Muon Systems. Muons are back-extrapolated through the calorimeter regions to the vertex, in order to retrieve the corresponding MIP and ISO bits, which are then added to the GMT output and can be taken into account by the Global Trigger (GT). Finally, the muons are sorted by transverse momentum

[4]The MIP bit is set if the calorimeter energy is consistent with the passage og a minimum ionizing particle, the isolation bit is set if a certain energy threshold in the trigger towers surrounding the muon is not exceeded.

and quality to deliver four final candidates to the GT. The Muon Trigger is designed to cover up to $|\eta| < 2.4$.

Global Trigger

The Global Trigger takes the decision to accept or reject an event at Level 1, based on candidate e/γ, muons, jets, as well as global quantities such as the sums of transverse energies (defined as $E_T = E \sin\theta$), the missing transverse energy and its direction, the scalar transverse energy sum of all jets above a chosen threshold (usually identified by the symbol H_T), and several threshold-dependent jet multiplicities. Objects representing particles and jets are ranked and sorted. Up to four objects are available and characterized by their p_T or E_T, direction and quality. Charge, MIP and ISO bits are also available for muons. The Global Trigger has five basic stages implemented in Field-Programmable Gate-Arrays (FPGAs): input, logic, decision, distribution and read-out. If the Level 1 Accept decision is positive, the event is sent to the Data Acquisition stage.

High Level Trigger and Data Acquisition

The CMS Trigger and DAQ system is designed to collect and analyse the detector information at the LHC bunch crossing frequency of 40 MHz. The DAQ system must sustain a maximum input rate of 100 kHz, and must provide enough computing power for a software filter system, the High Level Trigger (HLT), to reduce the rate of stored events by a factor of 1000. In CMS all events that pass the Level 1 Trigger are sent to a computer farm (Event Filter) that performs physics selections, using faster versions of the offline reconstruction software, to filter events and achieve the required output rate. The various subdetector front-end systems store data continuously in 40 MHz pipelined buffers. Upon arrival of a synchronous Level 1 Trigger Accept via the Timing, Trigger and Control System (TTCS) the corresponding data are extracted from the front-end buffers and pushed into the DAQ system by the Front-End Drivers (FEDs). The event builder assembles the event fragments belonging to the same Level 1 Trigger from all FEDs into a complete event, and transmits it to one Filter Unit (FU) in the Event Filter for further processing. The DAQ system includes back-pressure from the filter farm through the event builder to the FEDs. During operation, Trigger thresholds and pre-scales will be optimized in order to fully utilize the available DAQ and HLT throughput capacity.

2.3 Monte Carlo Event Generator

Monte Carlo (MC) event generators provide an event-by-event prediction of complete hadronic final states based on QCD calculation. They allow

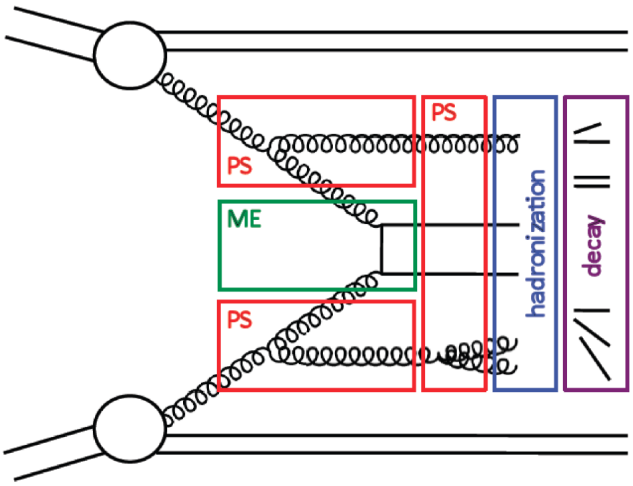


Figure 2.16: Schematic view of the subsequent steps of a MC event generator: matrix element (ME), parton shower (PS), hadronization and decay.

to study the topology of events generated in hadronic interactions and are used as input for detector simulation programs to investigate detector effects. The event simulation is divided into different stages as illustrated in figure 2.16. First, the partonic cross section is evaluated by calculating the matrix element in fixed order pQCD. The event generators presently available for the simulation of proton-proton collisions provide perturbative calculations for beauty production up to NLO. Higher order corrections due to initial and final state radiation are approximated by running a parton shower algorithm. The parton shower generates a set of secondary partons originating from subsequent gluon emission of the initial partons. It is followed by the hadronization algorithm which clusters the individual partons into colour-singlet hadrons. In a final step, the short lived hadrons are decayed. In the framework of the analysis presented here, the MC event generator PYTHIA 6.4 [35] is used to compute efficiencies, kinematic distributions, and for comparisons with the experimental results. This programs were run with its default parameter settings, except when mentioned otherwise.

PYTHIA

In the PYTHIA program, the matrix elements are calculated in LO pQCD and convoluted with the proton PDF, chosen herein to be CTEQ6L1. The mass of the b-quark is set to $m_b = 4.8$ GeV. The underlying event is simulated with the D6T tune [**?**]. Pile-up events were not included in the simulation. The parton shower algorithm is based on a leading-logarithmic approximation for QCD radiation and a string fragmentation model (implemented in JETSET [**?**]) is applied. The longitudinal fragmentation is de-

scribed by the Lund symmetric fragmentation function [**?**] for light quarks and by the Peterson fragmentation function for charm and beauty quarks, that is

$$f(z) \propto \frac{1}{z\left[1 - \frac{1}{z} - \frac{\varepsilon_Q}{(1-z)}\right]^2} \quad ,$$

where z is defined as

$$z = \frac{(E + p_\parallel)_{\text{hadron}}}{(E + p)_{\text{quark}}} \quad ,$$

$(E + p_\parallel)_{\text{hadron}}$ is the sum of the energy and momentum component parallel to the fragmentation direction carried by the primary hadron, $(E + p)_{\text{quark}}$ is the energy-momentum of the quark after accounting for initial state radiation, gluon bremsstrahlung and photon radiation in the final state. The parameters of the Peterson fragmentation function are set to $\epsilon_c = 0.05$ and $\epsilon_b = 0.005$. In order to estimate the systematic uncertainty introduced by the choice of the fragmentation function, samples generated with different values of ϵ_b are studied. The hadronic decay chain used in PYTHIA is also implemented by the JETSET program. For comparison, additional event samples are generated where the EvtGen program is used to decay the b-hadrons. EvtGen is an event generator designed for the simulation of the physics of b-hadron decays, and in particular provides a framework to handle complex sequential decays and CP violating decays.

Chapter 3

Data analysis

This thesis work is part of an analysis which aims at the measurement of the time integrated mixing probability, $\overline{\chi}$, of b-hardons produced in pp collision at centre of mass energy $\sqrt{s} = 8$ TeV. The peculiarity of this analysis is that only the b-hadrons generated in the decay of a top quark, in events where a top quark pair is produced, are used.

This kind of events provides a self-tagged sample of b-hadrons. Indeed, if a semileptonic decay of at least one of the top quarks occurs, the charge of the produced lepton, ℓ_{hard}, provides information about their flavour and thus about the b-hadron flavour at the production. In events where also the b-hadron decays semileptonically, the charge of the lepton produced in this decay, ℓ_{soft}, tags the flavour of the hadron at the decay.

The event topology is shown in figure 3.1. Only events in which one top quark decays semileptonically and the other one hadronically are considered. This because in that kind of event it is easier to reconstruct the top decay topology.

The conveniences of using the b-hadrons from top quark decay instead of the events in which a b pair is produced are several:

- The absence of uncertainties introduced by the opposite-side tagging used in $b\bar{b}$ events, that is the tagging of a b-hadron at the production reconstructing the flavour of the other hadron, which introduce a mistag ratio of about 30%.

- The signal purity of the sample after the application of *ad hoc* selection criteria, thanks to the production of two high-energy muons and four high-energy jets.

Furthermore, this is a mandatory step in order to improve the analysis for the measurement of the semileptonic asymmetry, $\mathcal{A}_{\text{sl}}$, with the data that will be collected during RUN2, as proposed by Gedalia et al. in [36].

The disadvantage of this kind of analysis is the low statistic, due to the smaller cross section of top pair events, with respect to the $b\bar{b}$ one, and to

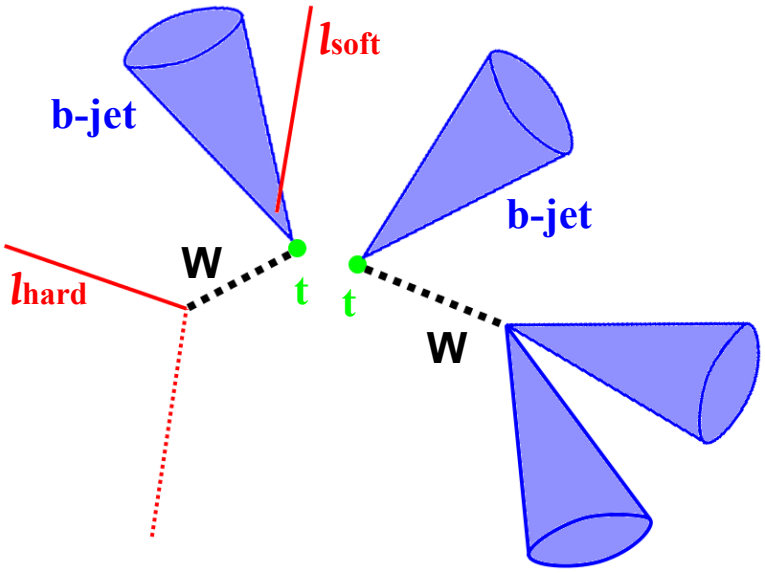


Figure 3.1: Topology of a typical event used in the analysis.

the request of two semileptonic decays.

The b-hadron averaged time-integrated mixing probability is obtained from

$$\overline{\chi} = \frac{N_{\rm lep}^{\pm\pm} + N_{\rm had}^{\pm\mp}}{N_{\rm lep}^{\pm\pm} + N_{\rm had}^{\pm\pm} + N_{\rm lep}^{\pm\mp} + N_{\rm had}^{\pm\mp}} \quad ,$$

where $N_{\rm lep}^{\pm\pm}$ ($N_{\rm lep}^{\pm\mp}$) is the number of events that have $\ell_{\rm hard}$ and $\ell_{\rm soft}$ with equal (opposite) charge, if the b-hadron is produced by the top quark that decays semileptonically, and $N_{\rm had}^{\pm\pm}$ ($N_{\rm had}^{\pm\mp}$) is the number of events that have $\ell_{\rm hard}$ and $\ell_{\rm soft}$ with equal (opposite) charge, if it is produced by the full-hadronic top quark.

In order to enhance the statistic, also the leptons produced in charmed-hardon semileptonic decays are considered when the b-hadron decays hadronically. Indeed the average oscillation probability of the charmed-hadrons is negligible and then the charge of these leptons is related to the flavour of the b-hadron.

The challenging steps of this analysis are:

- the choice of selection cuts that allow to discriminate signal events from background events or, as an alternative, the optimization of a method intended to separate them;
- the discrimination between events where the b-hadron is produced by the semileptonic-decayed top quark (SLt) and events where it is produced by the hadronic-decayed one (Ht);

- the discrimination between the events where ℓ_{soft} is produced by the b-hadron decay (*direct* ℓ_{soft}) and the events where it is produced by the charmed-hadron decay (*cascade* ℓ_{soft}).

In the last two steps and in the eventual separation method of the first step we choose to use Multi-Variate Analysis (MVA) methods, which provide a discriminating output parameter, calculated as function of a set of input event variables. In order to optimize these methods we made use of the ROOT-based toolkit TMVA [37]. A complete discussion about this method implementation is reported in appendix A.

Since also this well-performing discriminating methods do not provide an adequate separation to perform a single-cut classification, we choose to construct bi-dimensional PDFs of the MVA output parameter distribution. In order to do that, we use Monte Carlo simulated data to construct the bi-dimensional distribution for each different type of event (SLt *direct* ℓ_{soft}, SLt *cascade* ℓ_{soft}, Ht *direct* ℓ_{soft}, Ht *cascade* ℓ_{soft} and background[1]). Then we fit the data distribution with this five components in order to obtain the relative abundance of each class. So, calculating the MVA output values for a real event, one knows its probability to belong to each class and its charge information is split in the different $\overline{\chi}$ contributions.

For this analysis we use only events in which ℓ_{hard} and ℓ_{soft} are muons, because they have a better track parameter resolution and a easier particle identification compared to the electron one. In the future we will extend the analysis to use also electrons and τ leptons that decay leptonically.

3.1 Datasets and Event Selection

3.1.1 Top pair production cross section

At tree level, top pair production is mediated by quark pair annihilation and gluon fusion diagrams, as shown in figure 3.2, which contributes respectively to 13% and 87% of the top quark pair production cross section, in pp collision at centre of mass energy of $\sqrt{s} = 8$ TeV. This cross section has been theoretically calculated at next-to-next-to-leading order and next-to-next-to-leading logarithm (NNLO+NNLL) to be, at this energy scale,

$$\sigma_{t\bar{t}} = 245.8\,^{+6.2}_{-8.4}(\text{scales})\,^{+6.2}_{-6.4}(\text{pdf})\,\text{pb} \quad ,$$

that for an integrated luminosity of $\sim 20\,\text{fb}^{-1}$ predicts the production of about $4.9 \cdot 10^6$ events.

[1] Note that also signal events in which we tag the wrong ℓ_{hard} or ℓ_{soft} are categorised as background.

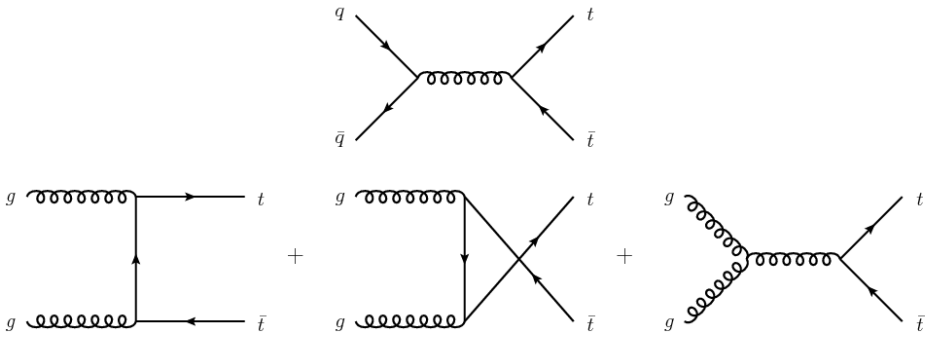


Figure 3.2: Feynman diagrams that contribute at tree level to top pair production in hadronic collision.

Dataset	Triggers	L [fb^{-1}]
`/DobleMu/Run2012A-22Jan2013-v1`	1 340 460	0.876
`/DobleMuParked/Run2012B-22Jan2013-v1`	12 251 486	4.412
`/DobleMuParked/Run2012B-22Jan2013-v1`	13 832 540	7.017
`/DobleMuParked/Run2012B-22Jan2013-v1`	12 536 714	7.369
Total	**39 961 200**	19.67 ± 0.51

Table 3.1: Number of events selected by the `HLT_Mu17_Mu8` trigger and integrated luminosities for the datasets taken in the 2012 run.

We required that exactly one top quarks and at least one b-hadron or charmed hadron decay semileptonically producing two muons. The branching ratio for this decay channel is ~ 0.048, for a total of about $2.4 \cdot 10^5$ events.

An estimated statistical error on the measurement of $\overline{\chi}$ is then about 0.003 (0.002 if also electrons are considered).

3.1.2 Datasets

Data samples used in this analysis have been collected in the 2012 runs, at centre-of-mass energy of 8 TeV. Actually we are using all four data-taking periods of `DoubleMuParked` datasets, for a total integrated luminosity of $19.7 \pm 0.5\,\text{fb}^{-1}$. The analysed events has been selected by the trigger `HLT_Mu17_Mu8`. It requires two muons with p_T grater than 17 and 8 GeV, respectively, with $|\eta| < 2.4$. The four datasets are presented in table 3.1.

3.1.3 Simulated data

For the simulation of the signal sample, top quark pair production where only one of them decays semileptonically (`ttSL` in the following), the `MADGRAPH` [38] event generator is used. The value of the top-quark mass is fixed to

$m_t = 172.5$ GeV and the proton structure is described by the parton density functions (PDF) CTEQ6L1 [39]. The generated events are subsequently processed with PYTHIA [35] for parton showering and hadronisation, the MLM prescription [40] is used for the matching of the jets with parton showers and the CMS detector response is simulated using GEANT4 [41].

Standard-model background samples are simulated with MADGRAPH, POWHEG [42] or PYTHIA, depending on the process. Top quark pair production with two semileptonic decays (ttFL in the following), top quark pair with additional boson production (ttWj and ttZj), W-boson production with additional jets (Wjet), Drell-Yan production of lepton pairs (DY when the two lepton invariant mass is greater than 50 GeV, DYj10 when it is between 10 GeV and 50 GeV), single-top-quark production, in s-, t- and tW-channel(t_s, t_t, t_tW for the t quark and tb_s, tb_t, tb_tW for the $\bar{t}$ quark), di-boson (WW, WZ and ZZ) and QCD multijet events are considered as background processes. The Wjet sample is simulated with MADGRAPH with up to four partons in the final state. POWHEG is used for single-top-quark production, while PYTHIA is used to simulate di-boson and QCD multijet events. Parton showering and hadronisation are also simulated with PYTHIA in all the background samples. The PYTHIA Z2 tune [43] is used to characterise the underlying event in both the signal and the background samples.

The normalisation of the above samples is taken from NNLO (W+jets, Z+jets), NLO+NNLL (single-top-quark [44]), NLO (diboson [45]), and leading-order (LO) (QCD multijet [35]) calculations. In the analysis all the simulated samples are rescaled to the acquired-data integrated luminosity of $19.67\,\mathrm{fb}^{-1}$.

All the simulated dataset are summarized in table 3.2.

3.1.4 Selection cuts

The selection criteria that I used in order to reduce strongly the background contamination in the data sample are divided in two steps: the first one is a pure inclusive selection, based on a list of event requests, which do most of the work of background reduction; the second step is a top pair decay reconstruction with few requirements on event topology, in order to provide the best candidate for each event element (ℓ_{hard}, ℓ_{soft}, the two b-jets and the two jets from the hadronic W decay).

Selection 1

This selection is an adaptation of the standard ttSL selection criteria. It is composed by the following requests:

- at least one muon with $p_T > 20$ GeV and isolation[2] (Iso) less than

[2]Isolation has been calculated as the total energy depositions in calorimeters in a cone of $\Delta R = 0.3$ around the muon track divided by the muon p_T.

Dataset	Triggers	L [fb^{-1}]
ttSL	1 400 511	237.20
ttFL	2 053 296	474.38
ttWj	17 067	845.03
ttZj	20 358	1021.68
Wjet	33 071	1.62
DY	4 576 826	8.70
DYj10	219 205	3.42
t_s	4 417	68.59
t_t	1 078	1.99
t_tW	16 700	44.67
tb_s	2 361	77.76
tb_t	21 845	67.19
tb_tW	16 624	44.30
WW	101 618	250.64
WZ	237 837	406.35
ZZ	446 206	725.92
QCD	280 343	0.16

Table 3.2: Number of events selected by the HLT_Mu17_Mu8 trigger and equivalent integrated luminosities for the simulated datasets.

0.15;

- no more than one muon with Iso < 0.15;
- at least two muons with $p_T > 8$ GeV;
- at least two jets with $p_T > 40$ GeV;
- at least four jets with $p_T > 25$ GeV;
- no opposite-sign muon pairs with invariant mass included in a 30 GeV wide interval centred in the $m_Z = 91.1876$ GeV.

Note that there are not requests on the jet b-tag parameter (CSV), in order to avoid any bias on the mixing probability. Indeed, the b-tag supports the events where the b-hadron has a longer life and it will increase the average mixing probability.

Selection 2

The aim of this selection is to find the best candidates for ℓ_{hard}, ℓ_{soft}, the b-jet from SLt and the one from Ht and the two jets from the decay of the W boson produced by Ht. For this purpose all the possible combinations

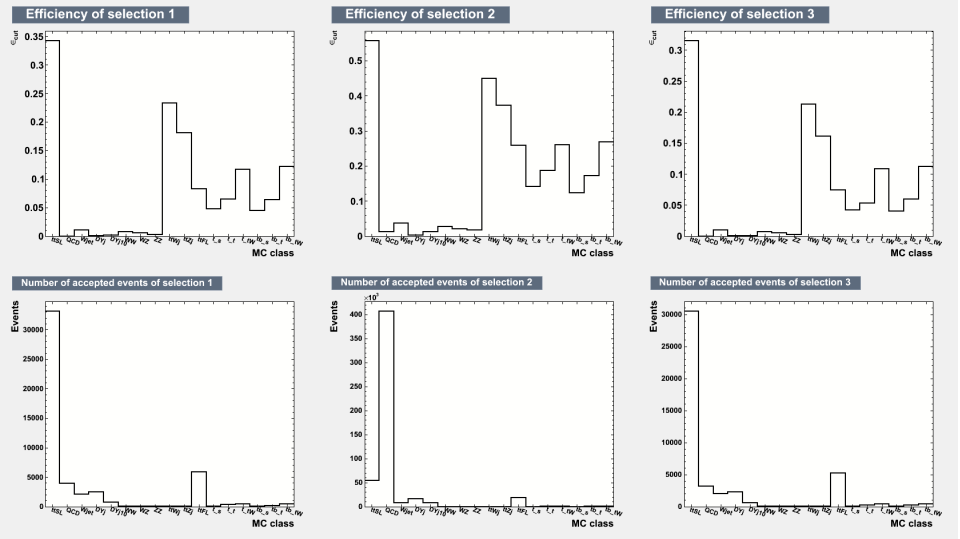


Figure 3.3: Efficiencies of the selections and number of selected events expected in a sample with integrated luminosity of $19.67\,\mathrm{fb}^{-1}$, as function of the type of event. The *Selection 3* is obtained applying both *Selection 1* and *Selection 2*.

are taken into account and that one with the lowest χ^2, which measures the compatibility of the reconstructed top quark masses and the hadronic W boson one, is used.

Applying both these selections a good background rejection is achieved. The efficiencies and number of selected events expected in the data sample for each type of event, obtained applying the selections to the simulated datasets, is plotted in figure 3.3.

With this selection the identification of the b-jets and the assignment of each of them to the correct top (Ht or SLt) is performed. However, this is a preliminary event reconstruction and a more efficient method is going to be improved.

3.2 Muon discrimination

The main goal of this thesis work is to optimize a method to discriminate events where the reconstructed ℓ_{soft} muon is produced by a b-hadron decay (*direct* ℓ_{soft}) from those where it is produced by a charmed hadron decay (*cascade* ℓ_{soft}). For this purpose a multi-variate analysis has been performed.

A Multi-Variate Analysis classification method is an algorithm that receives in input a set of event variables and gives as output a parameter that gives information about how much the event is signal-like or background-

like (in our case *direct*-like or *cascade*-like). Several methods exist for this purpose, with completely different input-variable treatments. A complete description of some of these methods, as well as an explanation of the method parameter tuning, training and application procedures and an introduction on the use of the ROOT-based TMVA toolkit are reported in appendix A.

3.2.1 Input variables

Firstly, I chose a set of input variables on the base of their discriminating power. More the shapes of a variable distribution for *direct* and *cascade* muons are different, more the variable is useful in the analysis. The effects of a low-discriminating-power input variable on the analysis are different for each MVA method. Some of them, like the likelihood or Boost Decision Tree (BDT) methods, have a small performance gain (or no gain) from them, while other methods, like Multilayer Perceptrons (MLP), could have a little performance loss. For this reason, a set of standard variables has been chosen and a performance test of the MVA methods is performed, excluding different subsets of them. The top-performing set of variables, common for all of the methods used, is the following:

- the transverse momentum of the muon with respect to the beam, p_T;
- the isolation of the muon;
- the distance between the muon direction and the axis of the jet that contains it, expressed in terms of $\Delta R = \sqrt{\Delta\eta^2 + \Delta\phi^2}$;
- the transverse momentum of the muon, with respect to the axis of its jet, recalculated without the contribution of the muon, $p_{T\text{out}}$;
- the ratio between the energy of the muon and the total energy of its jet;
- the jet-charge of the jet containing the muon, defined as the charge of the muon times the weighted average over the jet particle flow components of their charge, with, as weights, a power of their transverse momentum with respect to the beam,

$$Q_{\text{jet}} \equiv q_\mu \cdot \sum_i \frac{(p_T)_i^\alpha q_i}{(p_T)_i^\alpha} \quad ,$$

  where the best performance are obtained for $\alpha = 1.5$ and summing only over the particle flow components that have $\Delta R < 0.4$ with respect to the jet axis.

The normalized distribution of the input variables are shown in figures 3.4, 3.5 and 3.6.

3.2.2 MVA methods

The most simple method is the Log-Likelihood Ratio (LLR). With this approach, two binned Probability Density Functions (PDFs) are created for each input variable (one for *direct* muons, one for *cascade* muons), using the events available in the simulated signal dataset. Then, for each event in the acquired sample, the LLR value is calculated, using the definition

$$L = \sum_i \log \frac{S_i(x_i)}{B_i(x_i)} \quad ,$$

where the sum is performed over the set of input variable, S_i and B_i are the PDFs of *direct* and *cascade* muons, respectively, for the variable i and x_i is the value of this variable in the analysed event. However, this discriminating parameter performances can be improved, since this method does not take into account the correlations between the input variable.

Better performances are obtained using the methods provided in the TMVA toolkit. For this analysis I chose to optimize a BDT method and a MLP one, because they are two of the most performing methods. For more details see appendix A and the toolkit user guide [37]. Since the performance evaluations provided by the TMVA macros are related to a single-cut classification, I performed the analysis with both of them, in order to test which one provides the best fit performance.

Firstly, the setting parameters are optimized in order to obtain the best separation for the output values distribution. The resulting parameters are:

- for **BDT**: 800 trees in the forest, each with maximum depth of 4 nodes and leaves with a minimum size of 2% of the training sample. Each leaf is classified with the *Gini Index* and the weights boosting is performed by the *AdaBoost* algorithm, with $\beta = 0.5$. No sample bagging nor tree pruning is performed.

- For **MLP**: input variables normalized, 800 training cycles, a neuron structure composed by three hidden layers with, in order, 7, 6 and 5 neurons each. The activation function is the hyperbolic tangent, for all the neurons. The *Regulator* tool is used in order to reduce over-training. Event sampling is applied, with 80% of the total events used each cycle.

All the settings not specified here are left in the default state.

The methods have always better performances when applied to the sample used for training, since they are related to its statistical fluctuation (the method is over-trained). Thus, its use to build the fit PDFs could introduce a bias. For this reason I chose to train the methods with only a sub-sample of the available simulated signal dataset, which, later, has been excluded

from the analysis. In order to have sufficient training events to avoid overtraining, but leave enough events to build precise PDFs, the 17% of the `ttSL` events that fulfil the selection criteria has been used.

The MVA output normalized distributions are shown in figure 3.7.

3.3 Fit optimization

In order to extract the fractions of *direct* ℓ_{soft}, *cascade* ℓ_{soft} and background events in the data it is necessary to perform a fit of the MVA output distribution, with the PDFs built with the simulated data. This is only a preliminary step: once also the SLt-Ht separation will be optimized, a fit will be performed with an *ad-hoc*-built likelihood method on the bi-dimensional distribution of the two MVA outputs, in order to separate the five categories listed at the begin of this chapter.

Before fitting the data, it is necessary to optimize a fit mechanism by fitting a simulated dataset with known event fractions. For this purpose, the simulated datasets have been randomly split in two statistically independent subsets with the same size, one used to build the PDFs, the other used, without event distinction, to simulate the fitted data.

Firstly, I attempted to use the ROOT-based package `RooFit` [46]. This fitting toolkit performs an unbinned-likelihood fit; it has also a function to build the PDFs from MC datasets with an adaptive kernel estimator and provides a complete treatments of statistical and PDF shape uncertainties. However it is not possible to attribute weights to the events in the fitted dataset, so I could not use `RooFit` since my dataset is composed by events with different normalization.

The ROOT-based `TFractionFitter` package [47] is designed to perform a fit of a dataset distribution with two or more MC templates, in order to extract the relative abundances of each event class in the data. This fit is performed with the binned likelihood method; the MC histograms are used as templates, without the construction of analytic PDFs, and their bin heights are varied independently according to their statistic errors. The fitter also provide a complete treatments of the uncertainties.

Then, I constructed the three templates with the MVA output distributions for *direct*, *cascade* and background muons from the first data subset and the data histogram with the second subset. The fit is tested with the output of both of the MVA methods. The results are plotted in figures 3.8 and 3.9.

As one can see, the background amount is systematically underestimated. This is due both to the low statistic of some background classes, that produces some shape features that are different in the two subsets, and to the shape of the MVA output distribution. Indeed, the MVA method

is trained to discriminate *direct* and *cascade* events and the background distribution results to have a shape similar to the sum of a direct-like and a cascade-like shape, so that the fitter is not able to extract the correct fractions.

To overcome this problem, I chose to optimize a second MVA method in order to discriminate signal events from background ones and to perform the fit taking into account both the MVA output distributions.

3.3.1 Signal-background discrimination

For this discrimination I used again the BDT and MLP methods provided TMVA package. The input variables are:

- p_T and η of the ℓ_{hard} candidate;
- p_T and η of the ℓ_{soft} candidate;
- energy of the candidate b-jet that contains ℓ_{soft};
- energy and CSV of the candidate b-jet that does not contain ℓ_{soft};
- missing transverse energy, $\not{E}_T$;
- reconstructed mass of the top quark that decays hadronically and its χ^2;
- reconstructed mass of the top quark that decays semileptonically[3] and its χ^2;

The normalized distributions of the input variables are shown in figures 3.10, 3.11, 3.12, 3.13, 3.14 and 3.15.

The MVA parameters chosen to improve the output separation are:

- for **BDT**, random 60% of the sample is used in each iteration, the `Grad` method is used for boosting, nodes with a minimum size of 2% of the training sample, and 30 tested cuts for each variable.
- For **MLP**, the input variable are gaussianized.

All the settings not specified here are left in the default state.

For the signal simulated dataset the events used for training are the same used in the first method trainings. For the backgrond, the size of these subsamples varies for different datasets and depends on the number of simulated events that fulfilled the selection cuts. I chose to use 50% of the data for the `ttWj`, `ttZj`, `Wjet`, `DYj10`, single top and di-boson datasets and 17% of the

[3]For the reconstruction of the semileptonic top quark mass it is necessary to consider the $\not{E}_T$ contribution and impose the mass of the W boson to overcome the lack of information about the neutrino.

data for the `ttFL` and `DYj` datasets. No multijet QCD events are inserted in the training sample for two reason: on one hand they have an extremely low statistic and only 27 simulated events are selected, so a further event subtraction will increase substantially the PDF errors; on the other hand a method trained with few overweighted events results strongly over-trained, so that it is useless, if not damaging, inserting them in the training sample.

The MVA output distributions are shown in figure 3.16.

3.3.2 Results

In order to perform a combined fit to the two MVA outputs, a simultaneous one-dimensional fit to different profiles of the bi-dimensional distribution has been done. Once the bi-dimensional distribution of the two MVA outputs is built, the signal-background MVA output axis range has been divided in 8 bins and for each of them a profile distribution of the direct-cascade MVA output is built. These 8 distributions are then placed side by side in the same histogram. Fitting this histogram is now equivalent to perform a bi-dimensional binned fit, with the advantage that it is not necessary to build bi-dimensional PDFs.

The dataset and the MC templates are then prepared according to this distribution, using again the two random subsamples of the simulated datasets. The fit is performed with `TFractionFitter` package and the best performance is obtained using the BDT method in both discriminations. The results of this fit are shown in figure 3.17.

The resulting fractions are now quite compatible with the relative abundances in the fitted sample. However the best fit performance and a more transparent error treatment will be provided when an *ad-hoc* likelihood fitting methods will be performed.

3.4 Data - MC comparison

Before performing a fit on the data, it is necessary to test the compatibility between the data distributions of the MVA input variables and simulated event ones. The latter ones are normalized to a integrated luminosity of $19.67\,\mathrm{fb}^{-1}$. In figures 3.18-3.26 the variable distributions with the data/MC event ratios are reported.

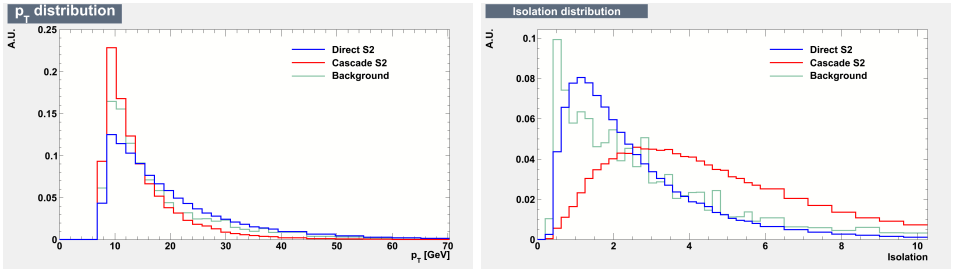


Figure 3.4: Normalized distributions of muon variables for *direct*, *cascade* and background muons. Left: p_T distribution. Right: Iso distribution.

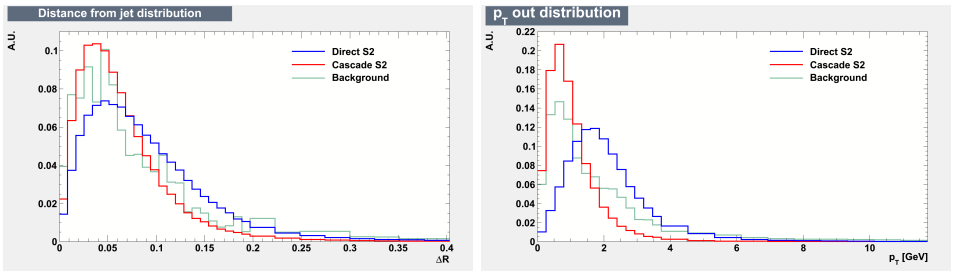


Figure 3.5: Normalized distributions of muon variables for *direct*, *cascade* and background muons. Left: distance from jet distribution. Right: $p_{T\text{out}}$ distribution.

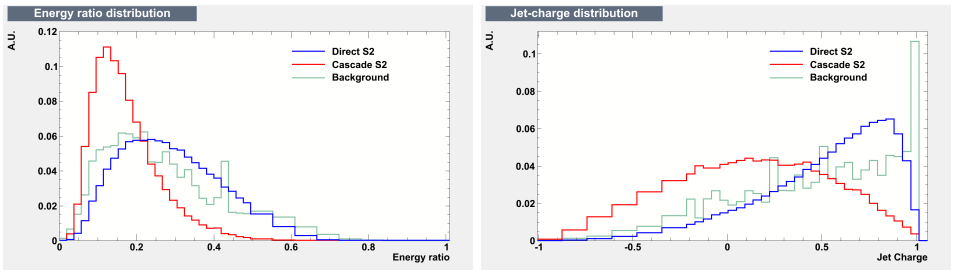


Figure 3.6: Normalized distributions of muon variables for *direct*, *cascade* and background muons. Left: Energy ratio distribution. Right: Jet-charge distribution.

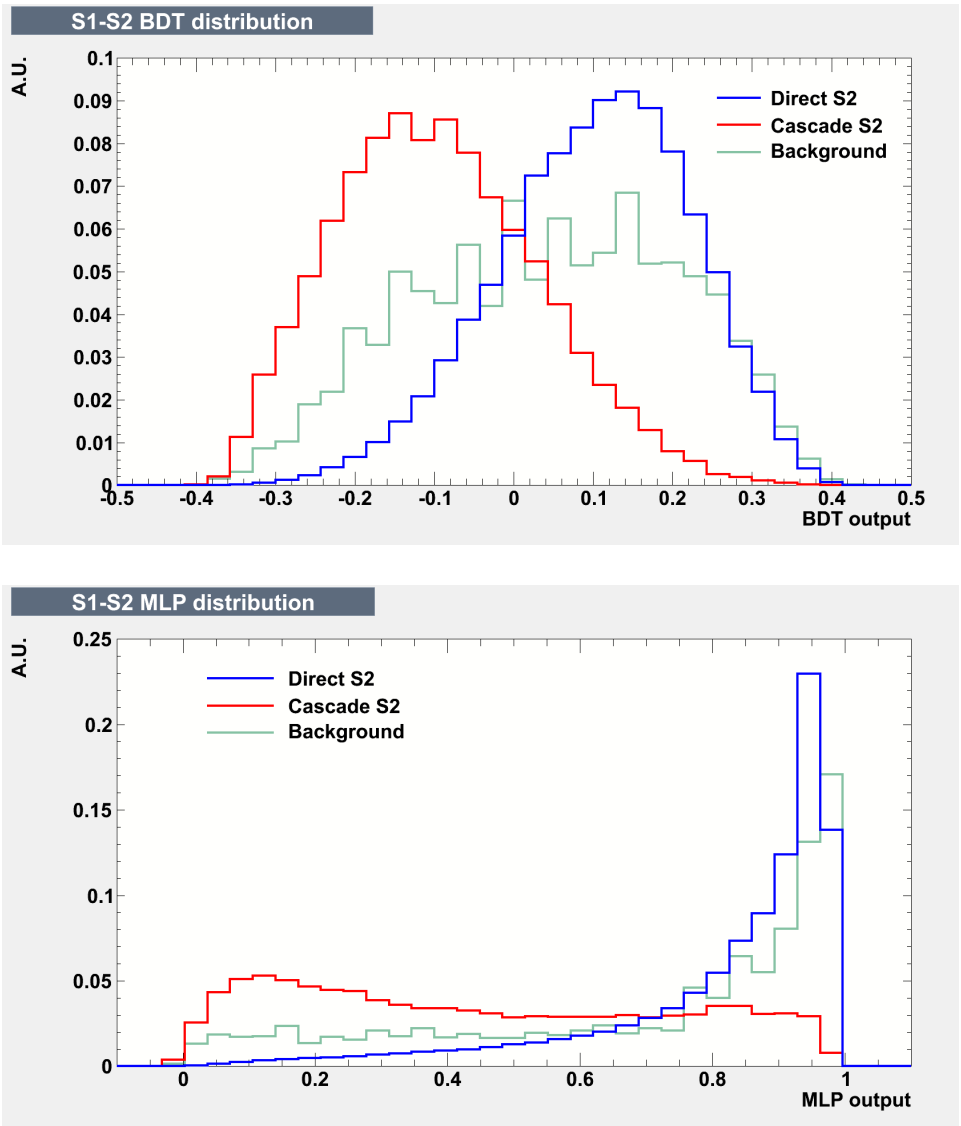
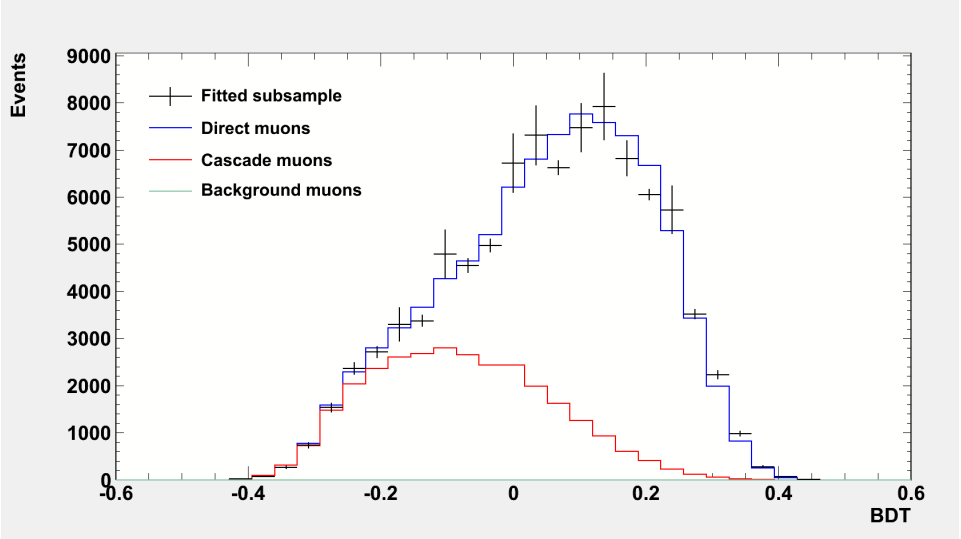
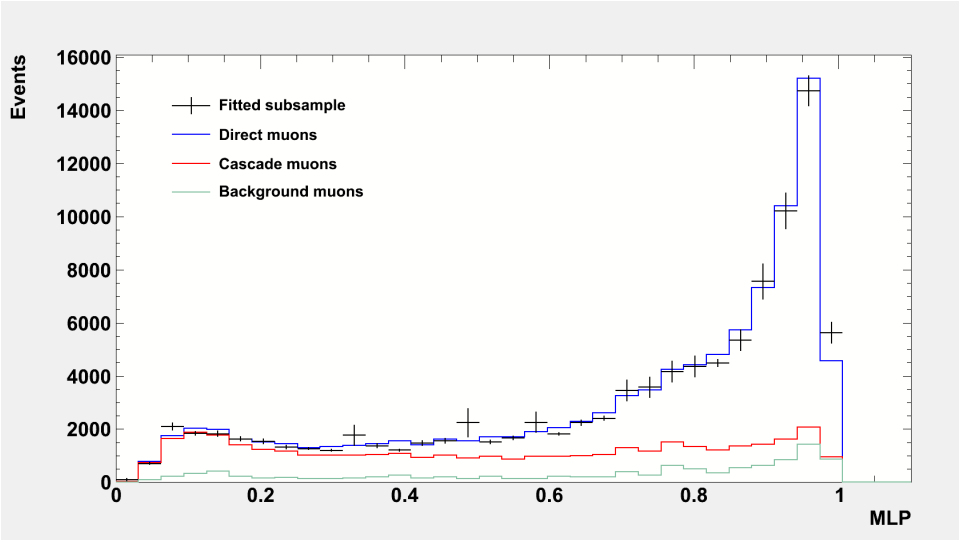


Figure 3.7: Normalized distributions of MVA outputs for *direct*, *cascade* and background muons. Top: BDT output distribution. Bottom: MLP output distribution.



Class	Real fraction	Fitted fraction
Direct	0.460	0.670 ± 0.007
Cascade	0.179	0.329 ± 0.006
Background	0.361	0.001 ± 0.014

Figure 3.8: Results of the fit of the BDT output distribution.



Class	Real fraction	Fitted fraction
Direct	0.460	0.616 ± 0.011
Cascade	0.179	0.275 ± 0.008
Background	0.361	0.109 ± 0.017

Figure 3.9: Results of the fit of the MLP output distribution.

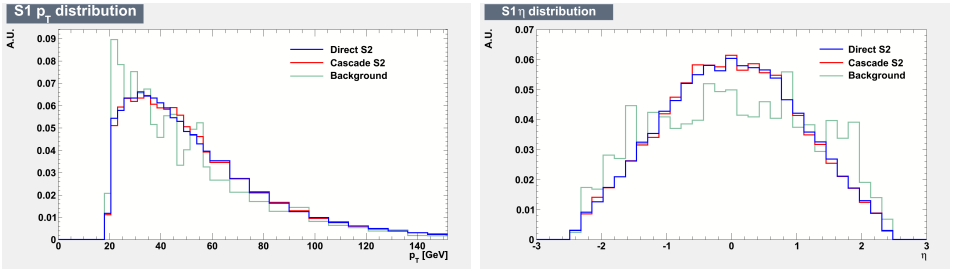


Figure 3.10: Normalized distributions of muon variables for *direct*, *cascade* and background muons. Left: ℓ_{hard} p_T distribution. Right: ℓ_{hard} η distribution.

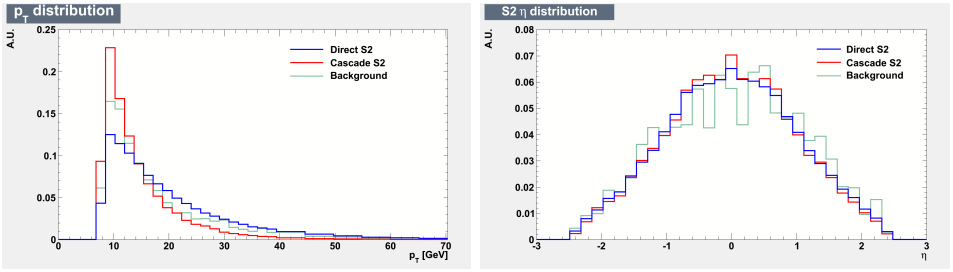


Figure 3.11: Normalized distributions of muon variables for *direct*, *cascade* and background muons. Left: ℓ_{soft} p_T distribution. Right: ℓ_{soft} η distribution.

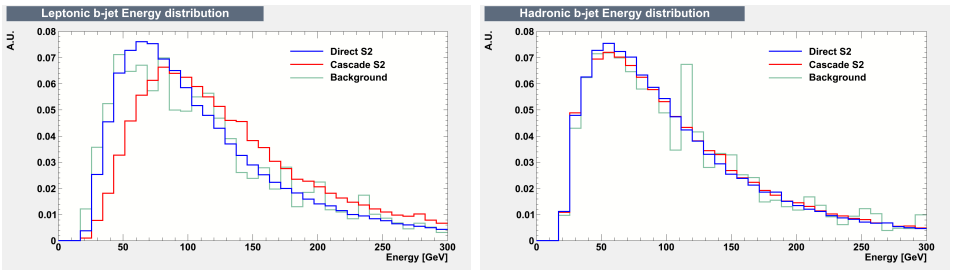


Figure 3.12: Normalized distributions of muon variables for *direct*, *cascade* and background muons. Left: energy distribution of the b-jet containing ℓ_{soft}. Right: energy distribution of the b-jet that does not contain ℓ_{soft}.

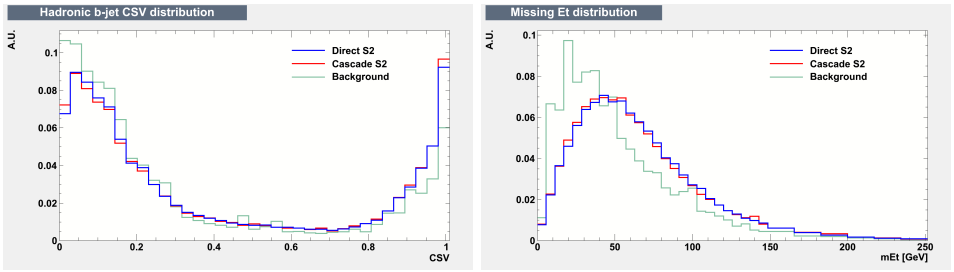


Figure 3.13: Normalized distributions of muon variables for *direct*, *cascade* and background muons. Left: CSV distribution of the b-jet that does not contain ℓ_{soft}. Right: missing transverse energy distribution.

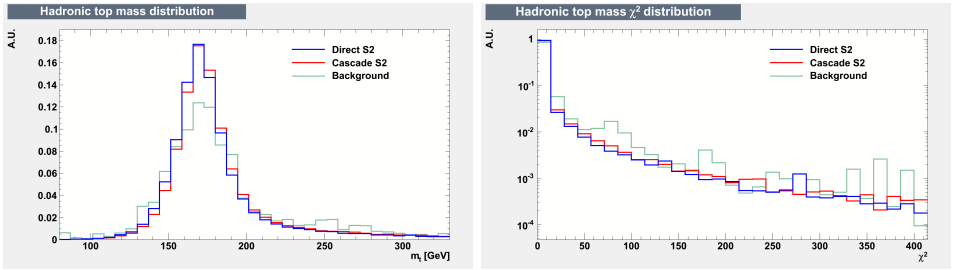


Figure 3.14: Normalized distributions of muon variables for *direct*, *cascade* and background muons. Left: Ht mass distribution. Right: χ^2 distribution of its reconstructed mass.

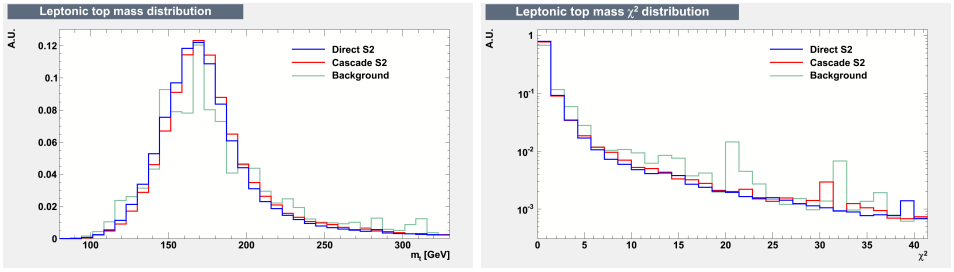


Figure 3.15: Normalized distributions of muon variables for *direct*, *cascade* and background muons. Left: SLt mass distribution. Right: χ^2 distribution of its reconstructed mass.

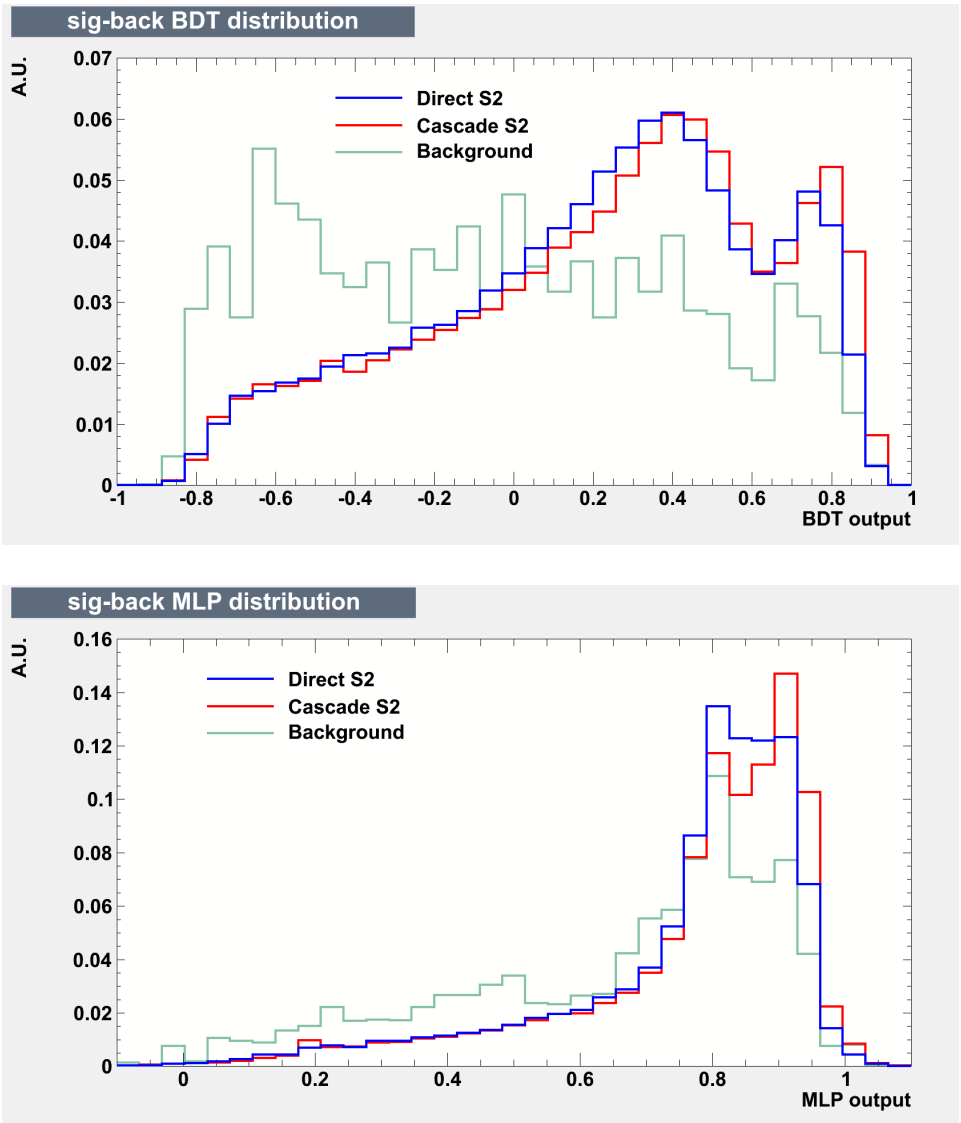
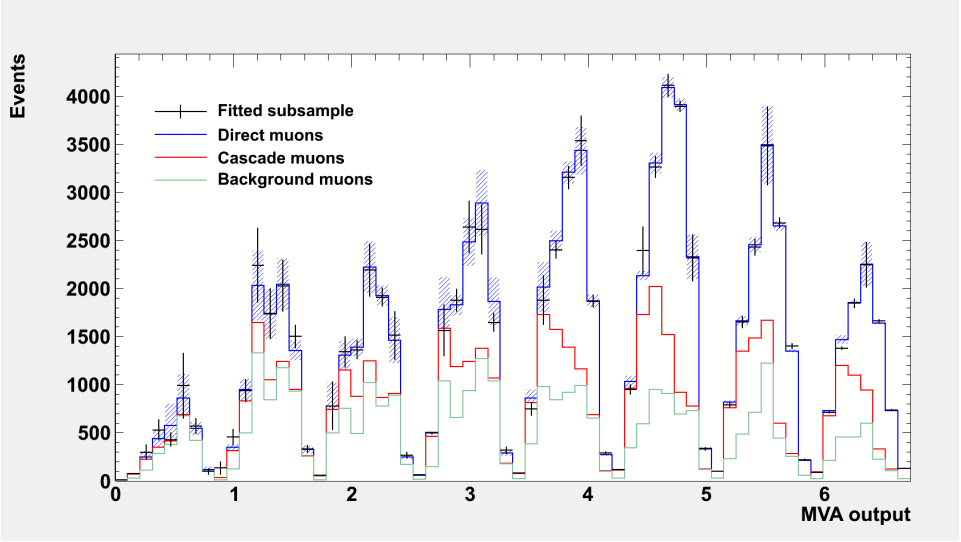
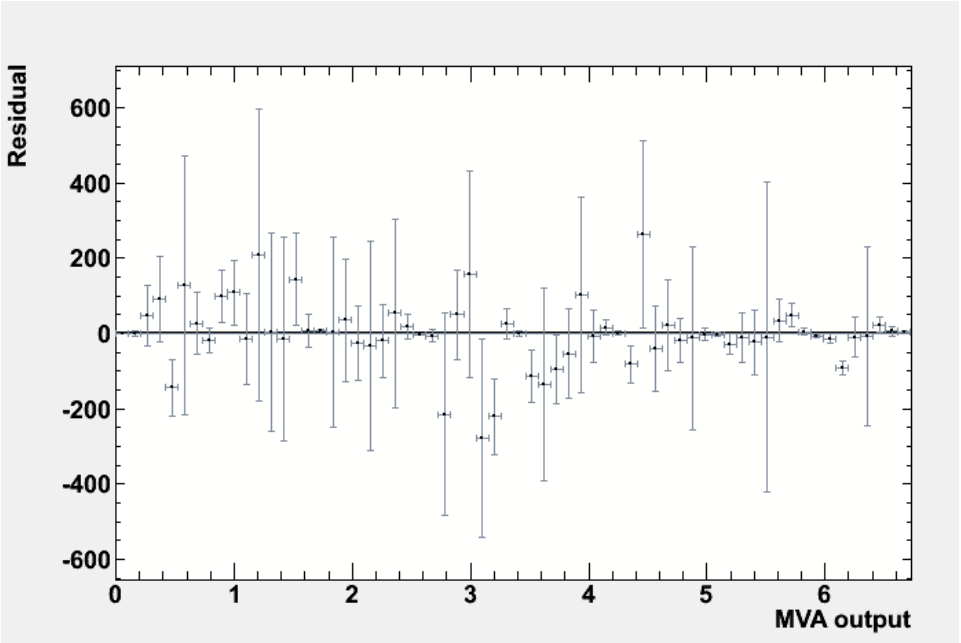


Figure 3.16: Normalized distributions of signal-background MVA outputs for *direct*, *cascade* and background muons. Top: BDT output distribution. Bottom: MLP output distribution.





Class	Real fraction	Fitted fraction
Direct	0.444	0.438 ± 0.006
Cascade	0.181	0.191 ± 0.005
Background	0.375	0.370 ± 0.007

Figure 3.17: Results of the MVA output fit.

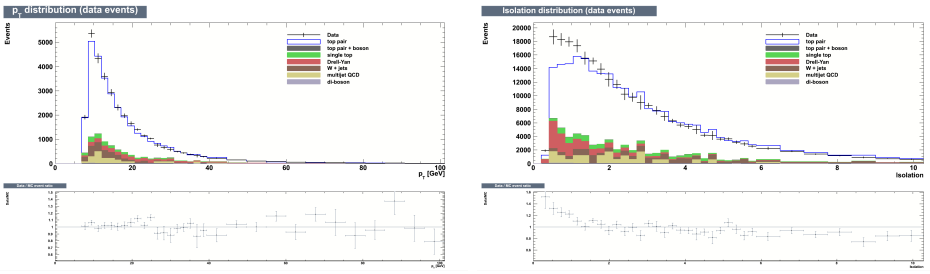


Figure 3.18: Distributions of MVA input variables for different event types. Left: p_T distribution. Right: Iso distribution.

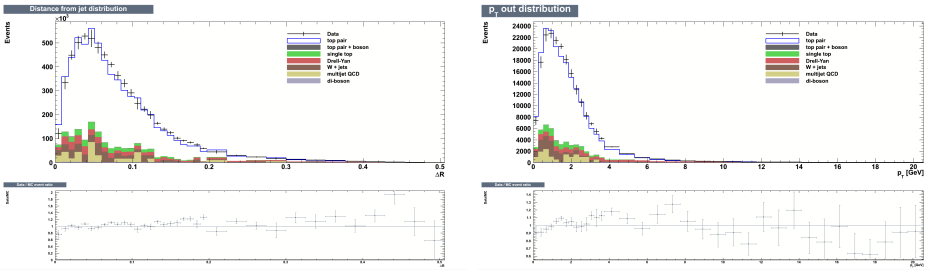


Figure 3.19: Distributions of MVA input variables for different event types. Left: distance from jet distribution. Right: $p_{T\text{out}}$ distribution.

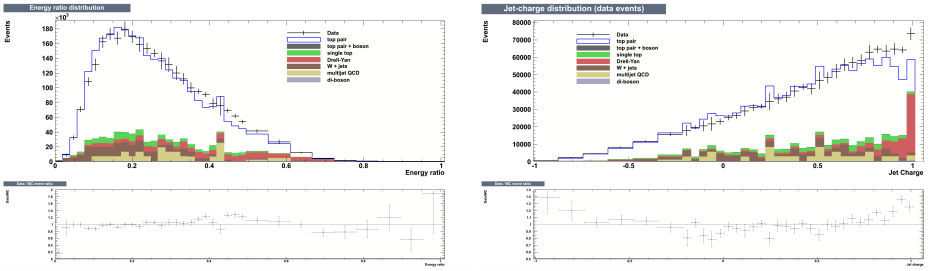


Figure 3.20: Distributions of MVA input variables for different event types. Left: Energy stack distribution. Right: Jet-charge distribution.

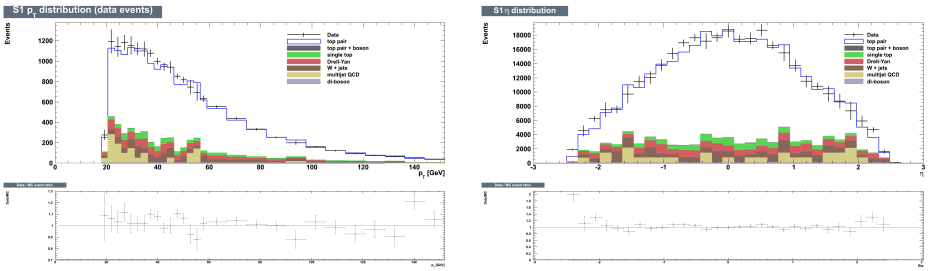


Figure 3.21: Distributions of MVA input variables for different event types. Left: ℓ_{hard} p_T distribution. Right: ℓ_{hard} η distribution.

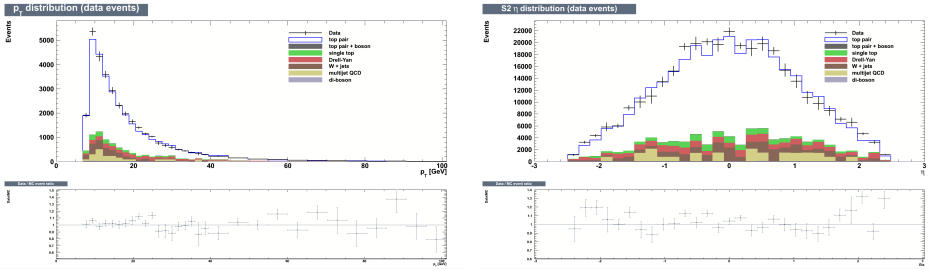


Figure 3.22: Distributions of MVA input variables for different event types. Left: ℓ_{soft} p_T distribution. Right: ℓ_{soft} η distribution.

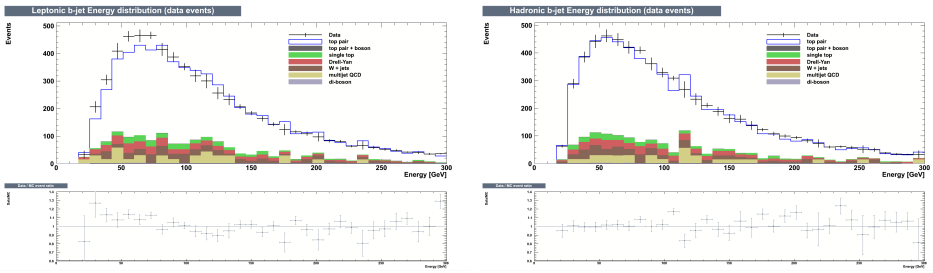


Figure 3.23: Distributions of MVA input variables for different event types. Left: energy distribution of the b-jet containing ℓ_{soft}. Right: energy distribution of the b-jet that does not contain ℓ_{soft}.

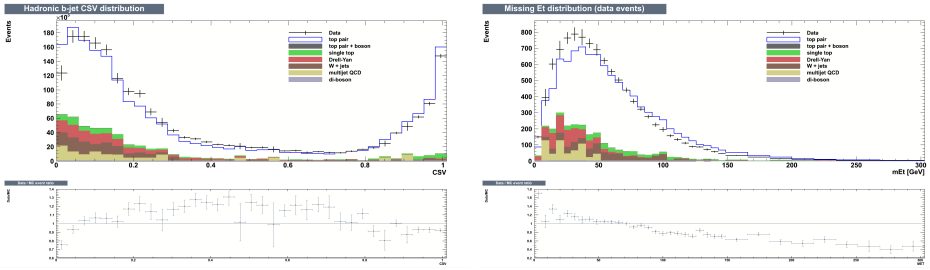


Figure 3.24: Distributions of MVA input variables for different event types. Left: CSV distribution of the b-jet that does not contain ℓ_{soft}. Right: missing transverse energy distribution.

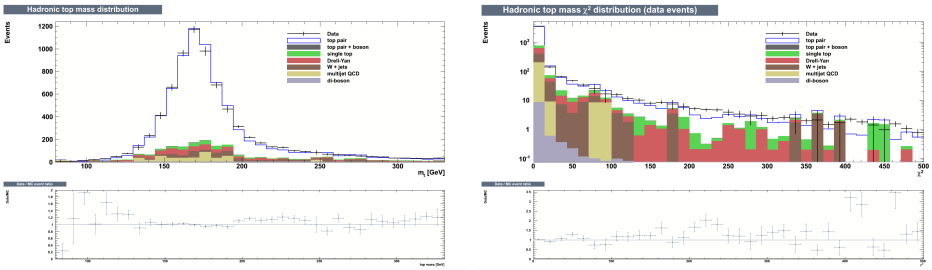


Figure 3.25: Distributions of MVA input variables for different event types. Left: Ht mass distribution. Right: χ^2 distribution of its reconstructed mass.

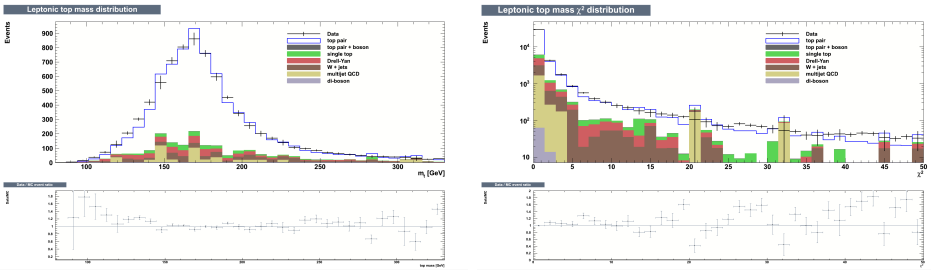


Figure 3.26: Distributions of MVA input variables for different event types. Left: SLt mass distribution. Right: χ^2 distribution of its reconstructed mass.

Chapter 4

Conclusions

In this work I optimized a Multi-Variate Analysis method to discriminate muons that originates from the semileptonic decay of b-hadrons and muons that come from charmed hadrons decay in b-jets produced in top quark decays. This method provides a very good separation between the two classes.

Also selection cuts and a MVA method to discriminate top quark pair events with only one top semileptonic decay from other types of events have been provided. The selection criteria provide a powerful background reduction and the MVA method has a discrete performance.

Furthermore, a fitting procedure has been performed in order extract the relative abundances of b-hadron muons, charmed hadron muons and background muons in a sample of data. This procedure has been tested on a simulated sample with known abundances and has given good results.

The agreement of the data and Monte Carlo distributions for the set of event variables used as input in the MVA methods has been tested.

The next steps for this analysis are optimizing the selection criteria, providing a method to appoint the hadronic-decaying top and the semileptonic-decaying one to the correct b-jets. Moreover, it will be necessary to provide an *ad-hoc*-defined likelihood method to perform the MVA outputs fit and extract the relative abundances and, finally, the application of this discrimination to the data to measure the time integrated mixing probability will be possible.

Appendix A

The TMVA toolkit

The Toolkit for Multivariate Data Analysis (TMVA) [37] provides a ROOT-integrated framework for the processing and parallel evaluation of many different multivariate classification and regression techniques. The classification is done in terms of two event categories, e.g. signal and background. The idea of TMVA is to integrate a large variety of powerful multivariate classifiers in one common environment with a single interface allowing the user to compare all classification techniques for any given problem. TMVA offers convenient preprocessing possibilities for the data prior to feeding them into any of the classifiers. Auxiliary information about the data is provided such as the correlations between the input variables, their separation power and ranking, various classifier specific validations and finally efficiency versus background rejection curves for all trained classifiers. These criteria allow the user to choose the optimal classifier for the given problem. The package currently includes implementations of:

- multi-dimensional rectangular cut optimisation using a genetic algorithm or Monte Carlo sampling;
- projective likelihood estimation;
- multi-dimensional likelihood estimation (k-nearest neighbour (k-NN) and probability density estimator range-search (PDERS));
- linear and non-linear discriminant analysis (Fisher, H-Matrix, Functional Discriminant Analysis);
- artificial neural networks (three different multilayer perceptron (MLP) implementations);
- Support Vector Machine;
- boosted/bagged decision trees (BDT) with pruning;
- predictive learning via rule ensembles.

A detailed description of the individual classifiers including the configuration parameters available for their tuning is given in the TMVA Users Guide [37]. TMVA provides training, testing and performance evaluation algorithms, visualisation scripts and auxiliary tools such as parameter fitting and variable transformations.

A.1 Data Preprocessing, Training and Testing

Training and testing of the classifiers is performed with user-supplied datasets with known event classification. This data is given in form of either ROOT trees or ASCII text files. The data sets are divided into statistically independent samples of training and testing data, omitting here an independent validation sample. Individual event weights may be attributed when specified in the data set. All classifiers see the same data sets and use the same prescription for the evaluation allowing for an objective comparison between them. A Factory class organises the interaction between the user and the TMVA analysis steps including preanalysis and preprocessing of the training data.

During the preanalysis, a preliminary ranking of the input variables is provided and their linear correlation coefficients are displayed. The variable ranking is later superseded by the ranking provided for each of the classifiers.

Preprocessing of the dataset includes the application of conventional preselection cuts that are common for all classifiers. In addition one can apply two different variable transformations, decorrelation via the square-root of the covariance matrix and via a principal component decomposition. These transformations can be individually chosen for any particular classifier. Removing linear correlations from the data sample may be useful for classifiers that intrinsically do not take into account variable correlations as for example rectangular cuts or projective likelihood.

After the training, each classifier writes the entire information needed for its later application to weight files. The classifiers are then tested and evaluated to assess their performance. The optimal classifier to be used for a specific analysis strongly depends on the problem at hand and no general recommendations can be given. To simplify the choice, TMVA computes and displays for each classifier a number of benchmark quantities such as:

- The signal efficiency and background rejection obtained from cuts on the classifier output. The area of the background rejection versus signal efficiency function is used for ranking the different classifiers.

- The separation$\langle S^2 \rangle$ of a classifier y, defined by the integral

$$\langle S^2 \rangle = \frac{1}{2} \int \frac{(\hat{y}_S(y) - \hat{y}_B(y))^2}{\hat{y}_S(y) + \hat{y}_B(y)} \, \mathrm{d}y \quad ,$$

where $\hat{y}_S$ and $\hat{y}_B$ are the signal and background PDFs of y, respectively. The separation is zero for identical signal and background shapes, and it is one for shapes with no overlap.

- The discrimination significance of a classifier, defined by the difference between the classifier means for signal and background divided by the quadratic sum of their root-mean-squares.

In addition, the variable distributions, correlation matrices and scatter plots, overtraining validation plots, as well as classifier specific information such as likelihood reference distributions, the neural network architecture and decision trees are conveniently plotted using ROOT macros executed via a graphical user interface that comes with TMVA.

A.2 Classifier Application

The application of the trained classifiers to the selection of events from a data sample with unknown signal and background composition is handled via a Reader object. It reads and interprets the weight files of the chosen classifier and can be included in any C++ executable, ROOT macro or python analysis job.

A.3 Boosted/bagged decision trees

A decision tree is a classifier that is structured as a binary tree. For each test event, repeated left/right (yes/no) decisions are performed on a single variable at a time until the event reaches a so called leaf node which classifies it as being either signal or background. The collection of leaf nodes split the phase space into many disjunct regions that are classified as being either signal or background like. The tree structure is defined during the training (tree building) phase, when starting from the whole training sample, consecutive binary splits are determined using the variable and cut value that allows maximum separation between signal and background at the time. When the splitting is finished, the node is classified as either signal or background depending on the majority of training events that end up in it.

In TMVA, the stop criteria for the splitting during the training phase is given by the minimum number of events which is demanded for a leaf node. Small numbers of events in leaf nodes are able to capture small features in the phase space discriminating signal from background. However this may easily result in over-training, i.e. the capture of statistical fluctuations in the training sample rather than genuine features of the underlying PDFs. A pruning algorithm with adjustable prune strength can be applied after the

tree building to remove statistically insignificant nodes from the decision tree.

Boosted decision trees represent an extension of a single decision tree. The classification of individuals of an ensemble of decision trees are combined to form a classifier which is given by a (weighted) majority vote of the classification from the individual decision trees. The individual trees are derived from the same training sample by reweighting the events. In TMVA, the standard AdaBoost [48] algorithm is implemented, which calculates the boost weight used in the next tree building depending on the number of misclassified training events in the previously trained tree.

Rather than performing boosting in order to create an ensemble of decision trees, bagging as defined in [49] uses randomly drawn events taken from the original training sample with replacement to construct different training samples. An ensemble of decision trees is then constructed from the collection of training samples derived by this re-sampling. A variant of this idea implemented in TMVA uses random event weights to create different training samples from which the decision trees in the ensemble are constructed.

Both bagging and boosting stabilise the response of the decision trees with respect to fluctuations in the training sample.

A.4 Multilayer Perceptron

Artificial neural networks (ANNs) are classifiers that feed the weighted input variables of a test event into a set of nodes also called neurons. Each node generates an output as response of the input according to its (non linear) activation function which can either be fed as input into consecutive nodes or used as an output, i.e. the final response of the network. With each node acting as a basis function, an appropriate choice of arrangement and number of nodes and their interconnecting weights in principle allows to approximate any decision boundary. For a given node arrangement (architecture) the training of a neural network consists of finding the interconnecting weights such that the separation between background and signal event is optimised.

There are three different implementations of ANNs in TMVA, all being feed-forward networks. This means that the nodes are arranged in an array with connections only in one direction, i.e. forward, from the input nodes (through the hidden layers) towards the output nodes without cycles or loops.

The CFANN was adapted from a FORTRAN code developed at the Université Blaise Pascal in Clermont-Ferrand and which uses random Monte Carlo sampling for the weight fitting during training. The other two networks both use standard back-propagation during the weight optimisation. One is an interface to the network already previously implemented in ROOT

and the other is our own development offering additional flexibility concerning the choice of different activation functions.

Bibliography

[1] H. Albrecht *et al.* (ARGUS Collab.). *Phys. Lett.*, B324:249, 1994.

[2] J. Bartelt *et al.* (CLEO Collab.). *Phys. Rev. Lett.*, 71:1680, 1993.

[3] B.H. Behrens *et al.* (CLEO Collab.). *Phys. Lett.*, B490:36, 2000.

[4] J. Beringer *et al.* (Particle Data Group). *Phys. Rev.*, D86:010001, 2012.

[5] D. Buskulic *et al.* (ALEPH Collab.). *Z. Phys.*, C75:397, 1997.

[6] J. Abdallah *et al.* (DELPHI Collab.). *Eur. Phys. J.*, C28:155, 2003.

[7] P. Abreu *et al.* (DELPHI Collab.). *Z. Phys.*, C76:579, 1997.

[8] M. Acciarri *et al.* (L3 Collab.). *Eur. Phys. J.*, C5:195, 1998.

[9] G. Abbiendi *et al.* (OPAL Collab.). *Phys. Lett.*, B493:266, 2000.

[10] K. Ackerstaff *et al.* (OPAL Collab.). *Z. Phys.*, C76:401, 1997.

[11] B. Aubert *et al.* (BaBar Collab.). *Phys. Rev.*, D73:012004, 2006.

[12] K. Abe *et al.* (Belle Collab.). *Phys. REF.*, D71:072003, 2005.

[13] T. Affolder *et al.* (CDF Collab.). *Phys. Rev.*, D60:112004, 1999.

[14] V.M. Abazov *et al.* (DØ Collab.). *Phys. Rev.*, D74:112002, 2006.

[15] R. Aaij *et al.* (LHCb Collab.). *Phys. Lett.*, B709:177, 2012.

[16] R. Aaij *et al.* (LHCb Collab.). *Phys. Lett.*, B709:177, 2012.

[17] L3 OPAL ALEPH, DELPHI and SLD Collabs.). *Phys. Reports*, 427:257, 2006.

[18] V.M. Abazov *et al.* (DØ Collab.). *Phys. Rev.*, D74:092001, 2006.

[19] E. Bruning *et al. CERN-2004-003-V-1.*

[20] ATLAS Collab. *JINST*, 3:S08003, 2008.

[21] CMS Collab. *JINST*, 3:S08004, 2008.

[22] LHCb Collab. *JINST*, 3:S08005, 2008.

[23] ALICE Collab. *JINST*, 3:S08002, 2008.

[24] CMS Collab. *Technical Design Report CMS Series*, CMS-TDR-008-1.

[25] CMS Collab. *Technical Design Report CMS Series*, CERN:CMS–TDR–001, 1997.

[26] CMS collab. *Technical Design Report CMS Series*, CERN:CMS–TDR–005, 1998.

[27] W. Adam *et al. CMS NOTE-2006*, 041.

[28] Billoir Pierre. *Computer Physics Communications*, 57:390, 1989.

[29] T. Speer *et al. CMS NOTE-2006*, 032.

[30] CMS Collab. *Technical Design Report CMS Series*, CERN:CMS–TDR–003, 1997.

[31] CMS Collab. *Technical Design Report CMS Series*, CERN:CMS–TDR–004, 1997.

[32] CMS Collab. *Technical Design Report CMS Series*, CERN:CMS–TDR–002, 1997.

[33] CMS Collab. *Technical Design Report CMS Series*, CERN:CMS–TDR–006–01, 2000.

[34] CMS Collab. *Technical Design Report CMS Series*, CERN:CMS–TDR–006–02, 2002.

[35] T. Sjostrand *et al. JHEP*, 05:026, 2006.

[36] O. *et al.* Gedalia.

[37] Andreas *et al.* Hoecker.

[38] J. Alwall *et al. JHEP*, 06:128, 2011.

[39] J. Pumplin *et al. JHEP*, 07:012, 2002.

[40] M.L. Mangano *et al. JHEP*, 01:2007, 013.

[41] S. Agostinelli *et al. Nucl. Instrum. Meth. A*, 506:250, 2003.

[42] S. Frixione *et al. JHEP*, 11:070, 1007.

[43] R. Field. *Acta Physica Polonica B*, 42:2631, 2011.

[44] N. Kidonakis *et al.*

[45] J.M.Campbell *et al. JHEP*, 07:018, 2011.

[46] W. Verkerke and D. Kirkby.

[47]

[48] Y. Freund and R.E. Schapire. *J. of Computer and System Science*, 55:119, 1997.

[49] R. Olshen L. Breiman, J. Friedman and C. Stone. Classification and regression trees. *Wadsworth*.