PhD Thesis

M. De Mattia

August 2006

Contents

Introduction

The Standard Model of fundamental interactions (SM) has been extensively tested in many particle experiments during the last 25 years and it has proven to be extremely successful up to the energy scale typical of the weak force. Nevertheless, there is still no experimental evidence of the Higgs boson, one of the key components of the SM, responsible for the breaking of the electroweak symmetry and for the masses of the fermions and of the weak bosons.
The Large Hadron Collider (LHC) is scheduled to provide the first proton on proton collision in 2008 at the center of mass energy of 14 Tev, an energy one order of magnitude higher than the regime explored so far.
The CMS experiment is an omni-purpose experiment that will operate at the LHC, it will give insight into Standard Model physics and search for physics beyond the Standard Model.

In this work we consider the usage of pixel detector information in the reconstruction of hadronic jets in events collected by the CMS detector under high luminosity running conditions, to add a small piece of information to best answer the question of the improvements possible on the physics output of CMS in case of an upgrade of its triggering capabilities.

In the first chapter we give a review of the Standard Model and its Minimal Supersymmetric extension. We then summarize the Higgs boson search channels at the LHC and present an update on the current status of Higgs searches at the Tevatron Collider.

The second chapter introduces the phenomenology of proton-proton interactions at the LHC and it gives a description of LHC and the CMS experiment, with particular attention to the pixel detector.

Chapter three describes a possible application of the pixel detector for triggering at the first level of the CMS trigger.
We will first introduce the software frameworks we used for all the simulations, with some details on the tools used in the production of the data samples.
In the second part we describe a pixel-based first level trigger and we derive optimized cuts for the selection of the $t\bar{t}$H signal.

In chapter four we consider a search channel of a Standard Model Higgs boson which could benefit from a reduction of energy thresholds of jet triggers, with the double aim of showing what can be obtained with such search using a sufficient statistics collected by CMS in the running phase of high luminosity, and to show what benefit the channel could receive from a trigger using jets plus pixels.

Finally, chapter five gives a summary of the results.

Chapter 1

The Standard Model and its possible extensions

The standard model is a quantum field theory based on a local gauge symmetry described by the group

$SU(3)_c \otimes SU(2)_L \otimes U(1)_Y$.

The matter fields are represented by Dirac spinors associated to two distinct sets of point-like particles of spin-$\frac{1}{2}$ (fermions). Each set is composed by 6 elements organized into three families. The interactions between the matter fields are described by the coupling with the gauge vector bosons which arise from the spontaneous breaking of the gauge symmetry $SU(2)_L \otimes U(1)_Y$. These gauge bosons are defined: W^+, W^- and Z for weak interactions and photon γ for electromagnetic interactions. Also a new particle is introduced in the theory, the Higgs boson. Including the strong interactions related to the $SU(3)_c$ color symmetry and mediated by eight colored massless gluons. The fermion and boson masses can be described by their couplings to the Higgs field.

In the following we take $\hbar = c = 1$.

1.1 Elementary Particles

In the standard model the elementary spin-$\frac{1}{2}$ fermions are the leptons, the quarks and their anti-particles. They are classified in three families (genera-

tions):

$$\begin{pmatrix} \nu_e \\ e \end{pmatrix} \quad \begin{pmatrix} \nu_\mu \\ \mu \end{pmatrix} \quad \begin{pmatrix} \nu_\tau \\ \tau \end{pmatrix}$$

$$\begin{pmatrix} u \\ d \end{pmatrix} \quad \begin{pmatrix} c \\ s \end{pmatrix} \quad \begin{pmatrix} t \\ b \end{pmatrix}$$

The constituents of matter are leptons and hadrons. Hadrons can be distinguished in mesons (formed by one quark and one anti-quark) and baryons (formed by combinations of three quarks). There is no experimental evidence of the existence of a fourth family. The interactions between the elementary particles are mediated by bosons, as follows:

Gravity:	graviton G
Strong:	8 gluons g
Weak:	W^+, W^-, Z
Electromagnetic:	photon γ

In the standard model weak and electromagnetic forces are originated from the electroweak interaction through the mechanism of spontaneous symmetry breaking.

1.2 Electroweak interaction

Electromagnetic and weak interactions are but two aspects of the same interaction. Sheldon Glashow, Abdus Salam, and Steven Weinberg[1, 2, 3] developed a theory based on the group $SU(2)_L \otimes U(1)_Y$. It is a chiral theory in the sense that left-handed and right-handed components behave differently under its gauge group. Weak interactions are a V-A theory and act only on the left components. For the fermion fields there is no mass term and the helicity is a good quantum number. Dirac spinors are introduced with well defined helicity states, left-handed (L) and right-handed (R). For the first generation of fermions:

$$L_L = \begin{pmatrix} \nu_\ell \\ \ell \end{pmatrix}_L \tag{1.1}$$

$$L_R = \begin{pmatrix} \ell \end{pmatrix}_R \tag{1.2}$$

where $\ell = (e, \mu, \tau)$. The right-handed neutrino is not included in the theory because, being massless[1], it has no interactions. For the quarks, the

[1]The neutrinos are not massless. This is indicated by their oscillations [4], however they have small masses and in first approximation we will take them to be zero.

only difference is the presence of two right-handed singlets. Helicity states can be represented using two quantum numbers: *weak isospin* T and *weak hypercharge* Y defined as

$$Y = 2(Q_{\rm em} - T_3) \ , \tag{1.3}$$

where T_3 is the third axis projection of the weak isospin and $Q_{\rm em}$ is the electric charge expressed in units of elementary charge [2]. We can then express the lepton states in the following notation

$$L_L \in (2, -1) \tag{1.4}$$

$$L_R \in (1, -2) \tag{1.5}$$

where the first number is the dimension of the $\mathrm{SU(2)_L}$ group representation and the second number is the hypercharge.

Fermions	Quantum Numbers			
	$\mathrm{Q_{em}}$	T	$\mathrm{T_3}$	Y
$(\nu_\ell)_L$	0	$\frac{1}{2}$	$+\frac{1}{2}$	-1
$(\ell)_L$	-1	$\frac{1}{2}$	$-\frac{1}{2}$	-1
$(\ell)_R$	-1	0	0	-2
$(u)_L$	$+\frac{2}{3}$	$\frac{1}{2}$	$+\frac{1}{2}$	$+\frac{1}{3}$
$(d)_L$	$-\frac{1}{3}$	$\frac{1}{2}$	$-\frac{1}{2}$	$+\frac{1}{3}$
$(u)_R$	$+\frac{2}{3}$	0	0	$+\frac{4}{3}$
$(d)_R$	$-\frac{1}{3}$	0	0	$-\frac{2}{3}$

Table 1.1: Electric charge $\mathrm{Q_{em}}$ (in unit of e), weak isospin T and third axis projection T_3 and weak hypercharge Y quantum numbers of lepton ($\ell = e, \mu, \tau$) and quark ($[u, d] = $ [u,d], [c,s], [t,b]) helicity states.

1.2.1 Leptons

A simple Lagrangian for the lepton fields without gauge interactions can be written in the form

$$\mathcal{L} = L_L^\dagger \tilde{\sigma}^\mu \imath \partial_\mu L_L + L_R^\dagger \sigma^\mu \imath \partial_\mu L_R \ , \tag{1.6}$$

[2] e = 1.602176462(63)× 10^{-19} C is the absolute value of the electric charge of the electron. The electron charge value is -e [5]

L_L transforms under SU(2) as

$$L \to L' = UL \ , \tag{1.7}$$

where U is any element of SU(2), and L_R is left invariated. Requiring the Lagrangian to be invariant under SU(2) leads to the introduction of the gauge derivative

$$\mathcal{D}_\mu = \partial_\mu + \imath(g_2/2)\vec{W}_\mu + \imath(g'/2)(Y)B_\mu \ , \tag{1.8}$$

where $\vec{W}_\mu$ and B_μ are the boson vector fields associated to the SU(2)$\otimes$U(1) symmetry, g_2 and g' are coupling constants and $\vec{W}_\mu = W^k_\mu(x)\tau^k \quad \tau^k$ are the Pauli matrices. Under SU(2)$\otimes$U(1) local gauge transformation the fields transform

$$\begin{aligned} B_\mu(x) &\to B'_\mu(x) = B_\mu(x) + (2/g_1)\partial_\mu\theta(x) \\ \vec{W}_\mu(x) &\to \vec{W'}_\mu(x) = U(x)W_\mu U^\dagger(x) + (2\imath/g_2)(\partial_\mu U(x))U^\dagger(x) \ . \end{aligned} \tag{1.9}$$

Again g_1 is an arbitrary constant.

The coupling constant g_2 of the SU(2) gauge theory must be the same for all leptons. This derives from the non-Abelian nature of SU(2) group and is called lepton universality.

The lepton part of the U(1)$\otimes$SU(2) invariant Lagrangian is

$$\mathcal{L}_{lepton} = L_L^\dagger \tilde{\sigma}\imath\mathcal{D}_\mu L_L + L_R^\dagger \sigma_\mu \imath\mathcal{D}_\mu L_R \ . \tag{1.10}$$

The dynamical part of the boson fields Lagrangian may be written in a gauge invariant form as

$$\mathcal{L}_{dyn} = -\frac{1}{4}B_{\mu\nu}B^{\mu\nu} - \frac{1}{4}\vec{W}_{\mu\nu}\cdot\vec{W}^{\mu\nu} \ , \tag{1.11}$$

with

$$\begin{aligned} B_{\mu\nu} &= \partial_\mu B_\nu - \partial_\nu B_\mu \ , \\ \vec{W}_{\mu\nu} &= \partial_\mu \vec{W}_\nu - \partial_\nu \vec{W}_\mu + g_2\left(\vec{W}_\mu \times \vec{W}_\nu\right) \ . \end{aligned} \tag{1.12}$$

1.2.2 Quarks

Comparing β decay

$$n \to p + e^- + \bar{\nu}_e \ , \tag{1.13}$$

or, at quark level

$$d \to u + e^- + \bar{\nu}_e \ , \tag{1.14}$$

to muon decay

$$\mu^- \rightarrow \nu_\mu + e^- + \bar{\nu}_e \ , \tag{1.15}$$

suggests that the left-hand components U_L and D_L of the quark fields should be put together in an SU(2) doublet

$$Q_L = \begin{pmatrix} U_L \\ D_L \end{pmatrix} \ . \tag{1.16}$$

U_R and D_R act like L_R, unchanged by SU(2) transformations. Again, the coupling to the W_μ field is uniquely determined by SU(2) symmetry and the coupling to the B_μ field by the quark electric charges: $2e/3$ for the u quark and $-e/3$ for the d quark. The Lagrangian invariant under U(1)$\otimes$SU(2) transformations is

$$\mathcal{L}_{quark} = Q_L^\dagger \tilde{\sigma}^\mu \imath \mathcal{D}_\mu Q_L + U_R^\dagger \sigma^\mu \imath \mathcal{D}_\mu U_R + D_R^\dagger \sigma^\mu \imath \mathcal{D}_\mu D_R \ . \tag{1.17}$$

This model is then extended to a gauge field theory $\mathrm{SU(3)_c \otimes SU(2)_L \otimes U(1)_Y}$ with the introduction of strong interactions between quarks, called quantum chromodynamics (QCD). In QCD we have three fields for each flavor of quark. This are put in a so called color triplet. For example

$$U = \begin{pmatrix} U_r \\ U_g \\ U_b \end{pmatrix} \ , \tag{1.18}$$

where U_r, U_g and U_b are four component Dirac spinors. To make the theory invariant under a local SU(3) transformation a 3x3 matrix gauge field G_μ is introduced and it is taken to be Hermitian and traceless. Therefore it can be expressed in terms of the eight Gell-Mann matrices λ_a

$$\vec{G}_\mu = \frac{1}{2} \sum_{a=1}^{8} G_\mu^a \lambda_a \ , \tag{1.19}$$

where the coefficients G_μ^a correspond to eight real independent gluon gauge fields.
The gluon dynamical contribution to the Lagrangian is

$$\mathcal{L}_{gluon} = -\frac{1}{4} \sum_{a=1}^{8} G_{\mu\nu}^a G^{\mu\nu a} \ , \tag{1.20}$$

where $G_{\mu\nu}$ is defined as

$$G_{\mu\nu} = \partial_\mu G_\nu - \partial_\nu G_\mu + ig(G_\mu G_\nu - G_\nu G_\mu) \ , \tag{1.21}$$

$G^a_{\mu\nu}$ is its expression by the G_μ components and g is the strong coupling constant.
The quark color interaction Lagrangian is

$$\mathcal{L} = \sum_{f=1}^{6} \bar{q}_f i\gamma^\mu \left(\partial_\mu + igG_\mu\right) q_f \ , \tag{1.22}$$

the sum is over all flavors of quark.

1.2.3 Spontaneous symmetry breaking

A simple lagrangian invariant under SU(2)⊗U(1) for a scalar field ϕ is

$$\mathcal{L}_\phi = \left(\mathcal{D}_\mu\phi\right)^\dagger \mathcal{D}^\mu\phi - V(\phi^\dagger\phi) \ , \tag{1.23}$$

where the gauge derivative $\mathcal{D}_\mu$ is defined in (1.8).
The scalar field is taken to be a complex doublet

$$\phi = \begin{pmatrix} \phi^+ \\ \phi^0 \end{pmatrix} \in (2, +1) \ , \tag{1.24}$$

and it is called the Higgs field [6].
The most general form of the potential $V(\phi^\dagger\phi)$ that allows renormalizability of the theory is

$$V(\phi^\dagger\phi) = \frac{m^2}{2\phi_0^2}\left[(\phi^\dagger\phi) - \phi_0^2\right] \ , \tag{1.25}$$

with a mass term m and a real parameter ϕ_0.
The vacuum state of the system is obtained for $\phi^\dagger\phi = \phi_1^2 + \phi_2^2 + \phi_3^2 + \phi_4^2 = \phi_0^2$ and is degenerate. We have three a^k parameters which specify an element of SU(2) to select the gauge. We adopt a gauge in which $\phi_A = 0$ and ϕ_B is real. Then

$$\phi_{ground} = \begin{pmatrix} 0 \\ \phi_0 \end{pmatrix} \ . \tag{1.26}$$

This vacuum state breaks the initial U(1)⊗SU(2) symmetry, but a U(1) symmetry is still present. This can be interpreted as the $\mathrm{U(1)_{em}}$ group of electromagnetic interactions, since the quantum numbers of the vacuum state are $T = 1/2$, $T_3 = -1/2$ and $Y = +1$ and thus $Q_{em} = (1/2)Y + T_3 = 0$.

Excited states of the Higgs field can be written as

$$\phi = \begin{pmatrix} 0 \\ \phi_0 + h(x)/\sqrt{2} \end{pmatrix} \ , \tag{1.27}$$

where $h(x)$ is a real function. Thus the mass terms in (1.23) for the boson fields are

$$\frac{g_2^2}{2} W_\mu^+ W^{+\mu} \phi_0^2 + \frac{g_1^2 + g_2^2}{4} Z_\mu Z^\mu \phi_0^2 \ , \tag{1.28}$$

where we conveniently defined

$$\begin{aligned} W_\mu^\pm &= \frac{W_\mu^1 \mp i W_\mu^2}{\sqrt{2}} \ , \\ Z_\mu &= W_\mu^3 \cos\theta_W - B_\mu \sin\theta_W \ , \\ \tan\theta &= \frac{g_1}{g_2} \end{aligned} \tag{1.29}$$

and θ_W is the Weinberg angle.
The orthogonal component to Z_μ is $A_\mu = W_\mu^3 \sin\theta_W + B_\mu \cos\theta_W$, it is massless and describes the photon.

Equation (1.28) represents the mass terms for three boson vectors. The $W^\pm$ have mass $M_W = \phi_0 g_2$ and the Z has mass $M_Z = \phi_0 (g_1^2 + g_2^2)^{1/2}$. Those vectors have three polarizations each.
The vacuum expectation value of the Higgs field is $\phi_0 = 246$ GeV.

1.2.4 Lepton masses

A gauge and Lorentz invariant contribution to the lagrangian which leaves the neutrino massless is

$$\mathcal{L}_{mass}^{lepton} = -G_\ell \left[L_L^\dagger \phi L_R + L_R^\dagger \phi^\dagger L_L \right] \ . \tag{1.30}$$

After symmetry breaking

$$L_{mass}^{lepton} = -G_\ell \phi_0 (L_L^\dagger L_R + L_R^\dagger L_L) - \frac{G_\ell h(x)}{\sqrt{2}} (L_L^\dagger L_R + L_R^\dagger L_L) \ , \tag{1.31}$$

Thus the charged lepton mass is $m_\ell = G_\ell \phi_0$.

1.2.5 Quark masses

For D quarks we proceed like electrons

$$\mathcal{L}_{Higgs}(D) = -\sum \left[G_{ij}^D (L_i^\dagger \phi) D_{Rj} + G_{ij}^{d*} D_{Rj}^\dagger (\phi^\dagger L_i) \right] \ , \tag{1.32}$$

which, after symmetry breaking, gives the mass term

$$\mathcal{L}_{mass}(D) = -\phi_0 \sum \left[G_{ij}^D D_{Li}^\dagger D_{Rj} + G_{ij}^{D*} D_{Rj}^\dagger D_{Li} \right] \ , \tag{1.33}$$

where G_{ij} is an arbitrary 3x3 complex matrix. In general it is possible to write $\phi_0 G_{ij} = D_L^\dagger m_D D_R$, where D_L and D_R are unitary and m^D is real and diagonal.

For the U quarks a suitable expression which we can construct from the doublets ϕ and L_i is $\epsilon_{ab}\phi_a L_{ib}$ where $\epsilon_{11} = \epsilon_{22} = 0$ and $\epsilon_{12} = -\epsilon_{21} = 1$. The contribution to the lagrangian is

$$\mathcal{L}_{Higgs}(U) = \sum \left[G^U_{ij}(\epsilon_{ab} L^\dagger_{ib} \phi^\dagger_a) U_{Rj} + G^{U*}_{ij} U^\dagger_{Rj} (\epsilon_{ab}\phi_a L_{ib}) \right] \ , \tag{1.34}$$

with G^U_{ij} another 3x3 complex arbitrary matrix.
Again after symmetry breaking we obtain

$$\mathcal{L}_{mass}(U) = -\phi_0 \sum \left[G^U_{ij} U^\dagger_{Li} U_{Rj} + G^{U*}_{ij} U^\dagger_{Rj} U_{Li} \right] \tag{1.35}$$

and like before $\phi_0 G^\mu = U_L^\dagger m^\mu U_R$. We can define 'true' quark fields as those that diagonalize the G_{ij} matrices

$$\begin{aligned} d_{Li} &= D_{Lij} D_{Lj} \ , \quad d_{Ri} = D_{Rij} D_{Rj} \ , \\ u_{Li} &= U_{Lij} U_{Lj} \ , \quad u_{Ri} = U_{Rij} U_{Rj} \ . \end{aligned} \tag{1.36}$$

Thus

$$\mathcal{L}_{mass(quark)} = -\sum_{i=1}^{3} \left[m^d_i (d^\dagger_{Li} d_{Ri} + d^\dagger_{Ri} d_{Li}) + m^u_i (u^\dagger_{Li} u_{Ri} + u^\dagger_{Ri} u_{Li}) \right] \ . \tag{1.37}$$

From (1.17) the off-diagonal terms of W_μ mix U and D quarks. When rewritten in terms of true quark fields they are

$$\mathcal{L}_{qW} = -\frac{e}{\sqrt{2}\sin\theta_W} (u^\dagger_L, c^\dagger_L, t^\dagger_L) \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} \begin{pmatrix} \tilde{\sigma}^\mu d_L \\ \tilde{\sigma}^\mu s_L \\ \tilde{\sigma}^\mu b_L \end{pmatrix} W^+_\mu + \textbf{H.conj.} \ , \tag{1.38}$$

where $V = U_L D_L^\dagger$ is a 3x3 unitary matrix. Five of the 9 parameters in V are non-physical phases and can be eliminated. The resulting matrix is called the Cabibbo-Kobayashi-Maskawa (CKM) matrix.

The CKM matrix is responsible for CP violation. CP is conserved if and only if $V_{ij} = V^*_{ij}$ (V_{ij} is real), which is not the case experimentally.

1.3 Beyond the Standard Model

Numerous experimental and theoretical reasons lead to believe that the Standard Model is only an effective theory valid up to the Electroweak scale.

A set of critical questions arises when the quantum corrections to the free parameters of the Higgs Lagrangian are computed. By solving the renormalization group equation for λ (the self interaction term in the Higgs potential), one obtains theoretical bounds on m_H as a function of the renormalization scale Λ: requiring the Higgs potential to be bounded from below is equivalent to impose $\lambda(\mu)$ to stay positive. For large values of m_H, $\lambda(\mu)$ increases with μ (see fig. 1.3), passing the perturbative limit ($\lambda \lesssim 1$) and eventually hitting the Landau singularity. The consistency of the model till the Planck scale ($M_p \sim 10^{19}$ GeV) is obtained only if m_H belongs to the range $[150 - 180]$ GeV.
The Higgs mass receives quantum corrections proportional to the cut-off scale squared. Since there are no (weakly broken) symmetries recovered when $m_H \to 0$ that "protect" m_H (like the chiral symmetry for the fermions' masses), either one has to Λ, i.e. the validity range of the theory, smaller than the TeV, or, if is required to be $m_H \ll \Lambda$, in the context of the SM only a huge fine-tuning is required. A scenario where $\Lambda \lesssim 1$ TeV has been ruled out by the precision tests of the SM, which exclude the contribution from new physics at such a close scale [7].
Another still open question is how the gravitational interactions can be quantized and unified with the other forces. General Relativity is a classical theory with high dimension operators in the Lagrangian which prevent it to be renormalizible. Superstring theories provide an appealing approach to gravitational field quantization, but they still lack predictive power, due to the difficulties in linking the gravity perturbative regime with the SM phenomena.
Many different models try to connect different interactions in view of an unified theory. Among the most significant, the SUSY models [8] introduce a new symmetry that associates to each SM particle a partner with opposite spin-statistic, thus canceling the fermion contribution in the loop of the Higgs self energy. The minimal supersymmetric extension to the SM will be discussed in more detail in the next section. Thecnicolor [9] theories follow a different approach. They consider the Higgs as a composite object, made of strongly interacting fermions (in analogy with the π mesons in QCD). In models with warped extra dimensions [10], the SM particles live in a four dimensional brane of a higher dimensional space. The gravity is strong as the other interactions (shifting its scale down to the SM one) and it seems weaker only because, being the only field which propagates though the "bulk", its component on the brane is small.
These models usually introduce many free parameters, which make them difficult to falsify.
From the analysis of the anysotropy power spectrum in the cosmic microwave

background, the total density parameter of the universe Ω is measured to be equal to 1 with an error of 2% [11]. Considering that the known matter, described by the SM, accounts only for few percents, the main components are still unknown. In particular, astronomical observations (e.g. of galactic rotation curves) implies the existence of non-luminous and non-absorbing matter, which cannot be explained in terms of standard matter. This dark matter contributes with a density of $\Omega_{DM} \sim 25\%$. Most of the SM extensions like SUSY, provide candidates for the dark matter, imposing symmetries to the predicted new particles, making at least one of them stable. The rest of the density component needed to make the universe flat is known as Dark Energy and it comes from a cosmological constant in the Einstein equation. It can be thought as the energy of the vacuum and has a negative pressure. Up to now there is no clue about the origin and the nature of this component. From cosmological observations comes the evidence that the antimatter component of the universe is negligible with respect to the matter one. This can be caused by CP-violating interactions, but the weak charged currents alone are not enough to explain the matter-antimatter asymmetry in the universe. The SM explains extremely well the experimental data up the scale of 100 GeV, but there are numerous arguments of paramount importance which still need to be understood and tested. The Large Hadron Collider has been designed to investigate the energy scales of the order of few TeV.

1.4 Minimal supersymmetric extension of the Standard Model

Supersymmetry (SUSY), a symmetry under interchange of bosonic and fermionic degrees of freedom, provides an elegant solution to the hierarchy problem, and has long been considered one of the most promising new physical scenarios among various possibilities. In SUSY, each particle in the spectrum has a "superpartner" with spin differing by half a unit. Quadratically divergent radiative corrections to the Higgs mass arising from a given particle are cancelled by the corresponding superpartner contributions. The fine tuning required to maintain an electroweak scale Higgs mass is no longer needed, thus solving the "hierarchy problem".

In the minimal supersymmetric extension to the Standard Model (MSSM), each Standard Model particle is accompanied by a superpartner with the same gauge quantum numbers as given in Tab. 1.4. The symbols for SM superpartners are the same as for the SM particles, but with an additional tilde. In order to write a theory that is supersymmetric, one needs two Higgs

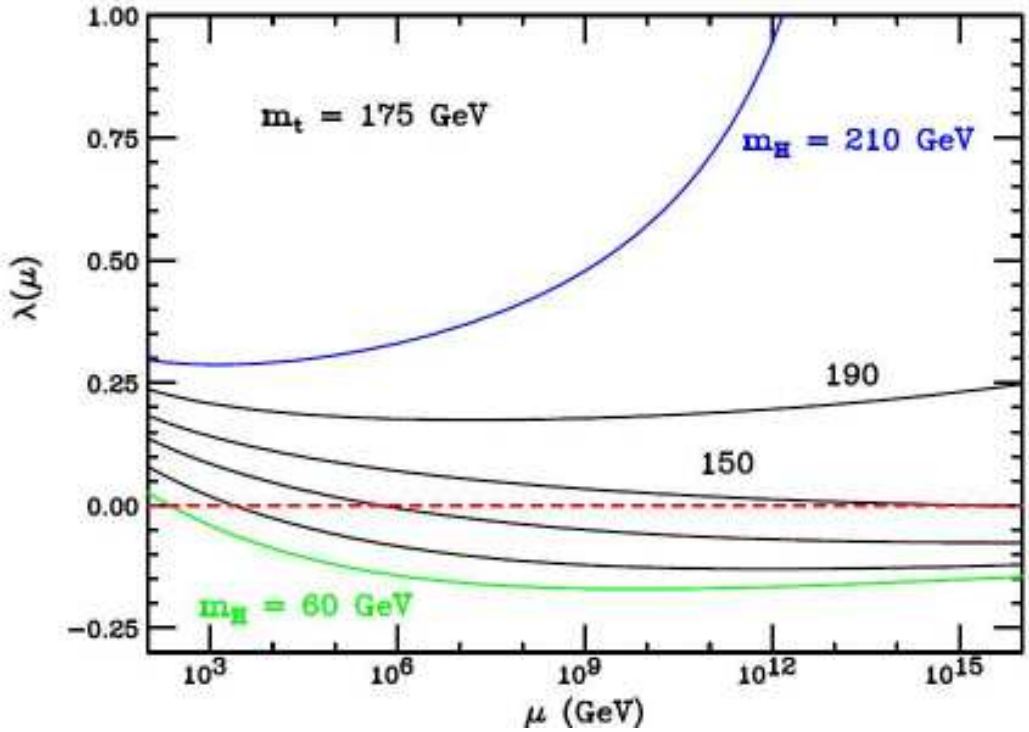


Figure 1.1: Higgs self interaction term λ dependency on the renormalization scale μ.

doublets with hypercharges 1 and -1. In order for the "hierarchy problem"

spin	quark sector			lepton sector		Higgs sector		gauge sector		
	Q	$\bar{u}$	$\bar{d}$	L	$\bar{e}$	H_u	H_d			
0	$(\tilde{u}_L, \tilde{d}_L)$	$\tilde{u}_R^*$	$\tilde{d}_R^*$	$(\tilde{\nu}, \tilde{e}_L)$	$\tilde{e}_R^*$	(H_u^+, H_u^0)	(H_d^0, H_d^-)			
1/2	(u_L, d_L)	$u_R^\dagger$	$d_R^\dagger$	(ν, e_L)	$e_R^\dagger$	$(\tilde{H}_u^+, \tilde{H}_u^0)$	$(\tilde{H}_d^0, \tilde{H}_d^-)$	$\tilde{g}$	$\tilde{W}^\pm, \tilde{W}^0$	$\tilde{B}^0$
1								$\tilde{g}$	$\tilde{W}^\pm, \tilde{W}^0$	$\tilde{B}^0$

Table 1.2: Particle contents for MSSM

to be solved, however, SUSY can be broken in masses but not in couplings (soft breaking), so that the coefficient of the canceling contributions remain equal and opposite.

The most general MSSM Lagrangian contains interactions that violate conservation of lepton/baryon number. To avoid having an unacceptably high decay rate for the proton a new symmetry, called R-parity, is invoked and the corresponding quantum number

$$P_R = (-1)^{3(B-L)-2s} \ , \quad s \text{ is the particle spin.} \tag{1.39}$$

All SM particles have $P_R = 1$ while all superpartners have $P_R = -1$. Conservation of P_R has some important consequences:

- The lightest supersymmetric particle (LSP) is stable
- Superpartners are pair-produced in collider experiments
- Superpartners contribute to SM observables only via loop corrections

1.5 Higgs searches

In this section we will review the status and prospects of the searches for the Higgs boson. In the first part, we will discuss the Higgs phenomenology at LHC and the channels of interest for the different ranges of its mass. The second part will give a brief review of searches for MSSM Higgs at LHC.

1.5.1 SM Higgs searches at LHC

At LHC a number of different Higgs production and decay channels can be studied. The preferred channels for $m_h \lesssim 200GeV$ are

$$\begin{aligned} & gg \to H \to \gamma\gamma \ , \\ & gg \to H \to VV^{(*)} \ , \\ & qq \to qqV^{(*)}V^{(*)} \to qqH \ , \quad H \to \gamma\gamma \ , \ \tau^+\tau^- \ , \ VV^{(*)} \ , \\ & gg, \ q\bar{q} \to t\bar{t}H \ , \quad H \to b\bar{b} \ , \ \gamma\gamma \ , \ WW^{(*)} \ , \end{aligned} \tag{1.40}$$

where $V = W$ or Z. The gluon-gluon fusion mechanism is the dominant Higgs production mechanism at the LHC.
Another appreciable Higgs production channel comes through the VV electroweak gauge boson fusion, which can be separated from the gluon fusion process by employing a forward jet tag and central jet vetoing techniques. Finally, the cross-section for $t\bar{t}H$ production [12] can be significant for $m_H \lesssim 200$ GeV, although this cross-section falls faster with Higgs mass compared to the gluon and gauge boson fusion mechanism.
Note that for $2m_W \lesssim m_H \lesssim 2m_Z$, the Higgs branching ratio to ZZ^* is quite suppressed with respect to WW. Hence, in this mass window, $H \to W^+W^- \to \ell^+\nu\ell^-\bar{\nu}$ is the main Higgs discovery channel [13], as shown in fig 1.5.1. For $m_h \lesssim 150GeV$, the expected accuracies of Higgs couplings to W^+W^-, $\gamma\gamma$, taus and gg can be determined to an accuracy in the range of 5-15% [15]. These results are obtained assuming that the partial Higgs widths to W^+W^+ and ZZ are fixed by electroweak gauge invariance, and the ratio to the partial Higgs widths to $b\bar{b}$ and $\tau^+\tau^-$ are fixed by the universality of Higgs couplings to down-type fermions. Assume also that the remaining unobserved modes, both in the SM and beyond, possess small branching ratios of order 1%. The resulting accuracy anticipated is in the range of 10-25%, depending on the Higgs mass.

The main Higgs production channel at LHC is the gg fusion. Other processes with distinctive signatures are

- WW and ZZ fusion with a cross section of about 20% of the gg fusion at low masses and becomes approximately equal for $m_H \approx 1TeV$.

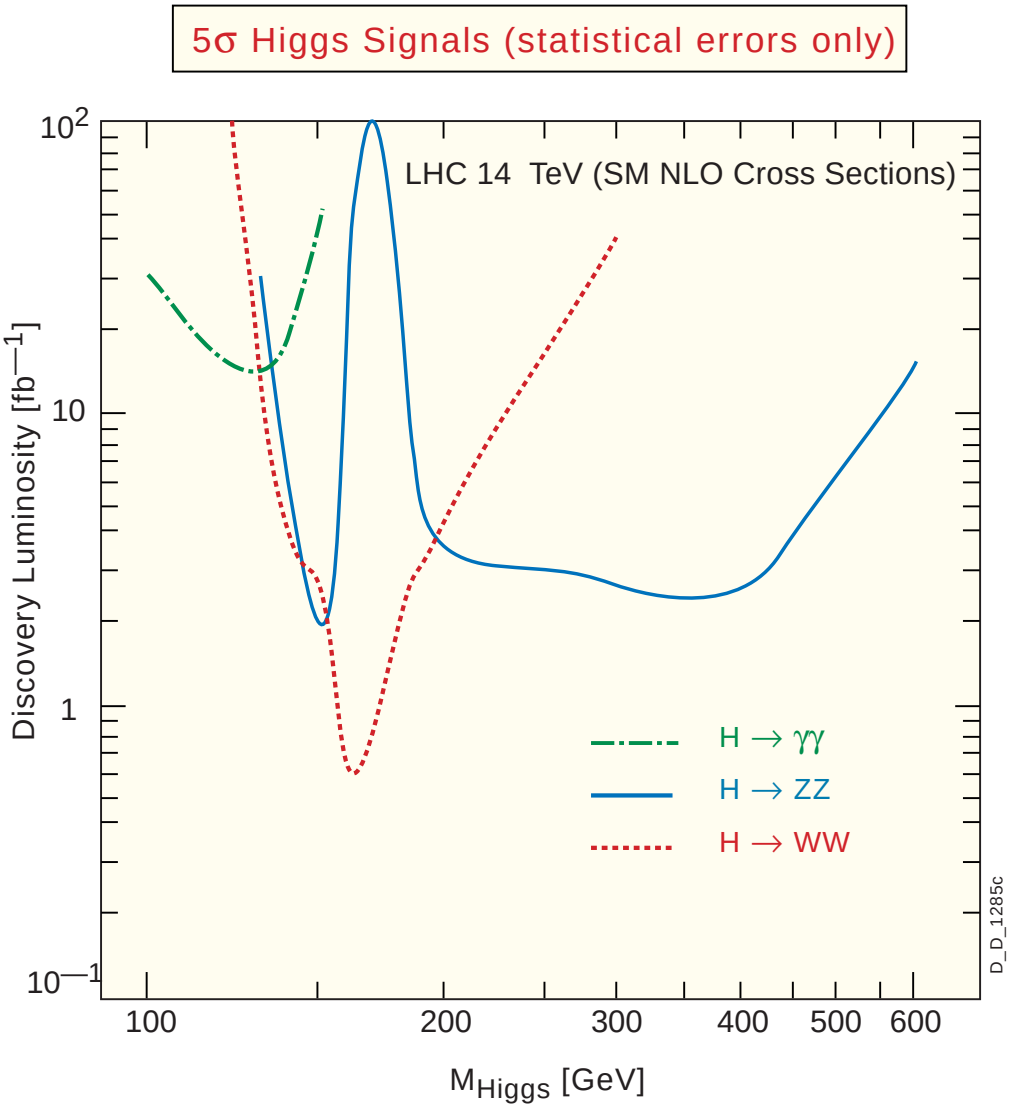


Figure 1.2: Expected 5σ discovery luminosity requirements for the SM Higgs boson at the LHC for one experiment, based on a study performed with CMS fast detector simulation, assuming statistical errors only [14].

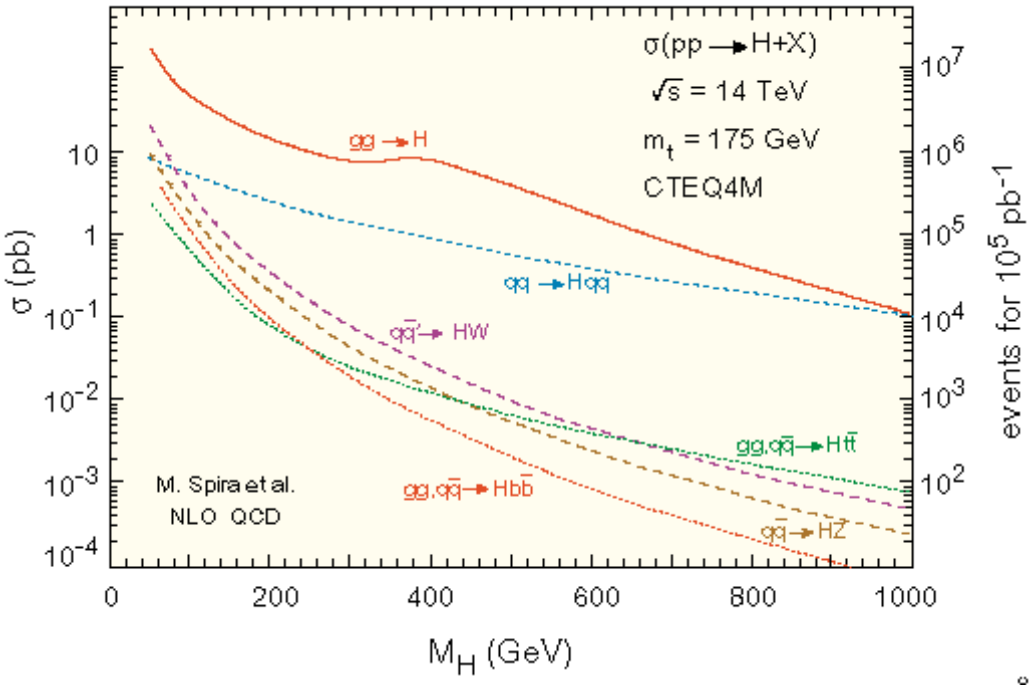


Figure 1.3: Higgs production cross section for the center of mass value of $\sqrt(s) = 14$ TeV of LHC.

- The $t\bar{t}H$ process where a Higgs boson is produced in association with a $t\bar{t}$ pair.
- The WH and ZH processes, where the Higgs boson is radiated by a vector boson.

The production cross sections for the different processes are shown in figure 1.5.1 as a function of the Higgs boson mass. The main decay channels for the Higgs boson are $H \to b\bar{b}$ and $H \to \tau^+\tau^-$ for low masses, $m_H \lesssim 135GeV$, but the backgrounds to these channels are too large and additional signatures are needed. For the $H \to b\bar{b}$ for example the production in association with a $t\bar{t}$ pair and for the $H \to \tau^+\tau^-$ channel WW or ZZ fusion can be used. For higher Higgs masses the WW and ZZ decay channels dominate and the leptonic decays of the vector bosons can be used.

For low masses, $100GeV < m_H < 150GeV$, the $H \to \gamma\gamma$ channel can be used. Although it has a small cross section, it has a clear signature.
For $m_H \approx 170GeV$, $H \to WW$ has a branching ratio of almost 100% and is

the only detectable channel.
Table 1.5.1 gives a summary of the search channels [16] and figure 1.5.1 shows the relative branching ratios as a function of the Higgs boson mass.
In the following we discuss in some detail some of the above signatures.

Higgs mass range	Decay	Production			
		Incl.	qqH	WH/ZH	ttH
Low mass, $m_H < 200GeV$	$H \to \gamma\gamma$	Yes	Yes	Yes	Yes
	$H \to b\bar{b}$				Yes
	$H \to \tau^+\tau^-$		Yes		
	$H \to WW^* \to \ell\nu\ell\nu$	Yes	Yes	Yes	
	$H \to ZZ^* \to 4\ell$	Yes			
Intermediate mass, $200GeV < m_H < 700GeV$	$H \to WW \to \ell\nu\ell\nu$	Yes			
	$H \to ZZ4\ell$	Yes			
High mass, $m_H > 700GeV$	$H \to WW \to \ell\nu q\bar{q}$		Yes		
	$H \to ZZ \to \ell^+\ell^-\nu\bar{\nu}$		Yes		

Table 1.3: Most sensitive production and decay channels for the SM Higgs boson search. $\ell = e, \mu$

1.5.2 $H \to \gamma\gamma$

The branching ratio for this channel is small, of order 10^{-3}. It is characterized by the presence of two isolated photons with large transverse energy that allow the reconstruction of a narrow mass peak. This allows separation from a large irreducible background due to photon pairs from gluon-gluon fusion and quark annihilation.
The background from misidentified jets and π^0 can be reduced with isolation cuts and photon identification requirements.

The mass measurement in this channel can be made very accurate, with a foreseen precision of $\approx 1\%$ for CMS and $\approx 1.5\%$ for ATLAS[16]
For $100GeV < m_H < 150GeV$ this channel alone can give a 5σ discovery with $\sim 30fb^{-1}$ [].

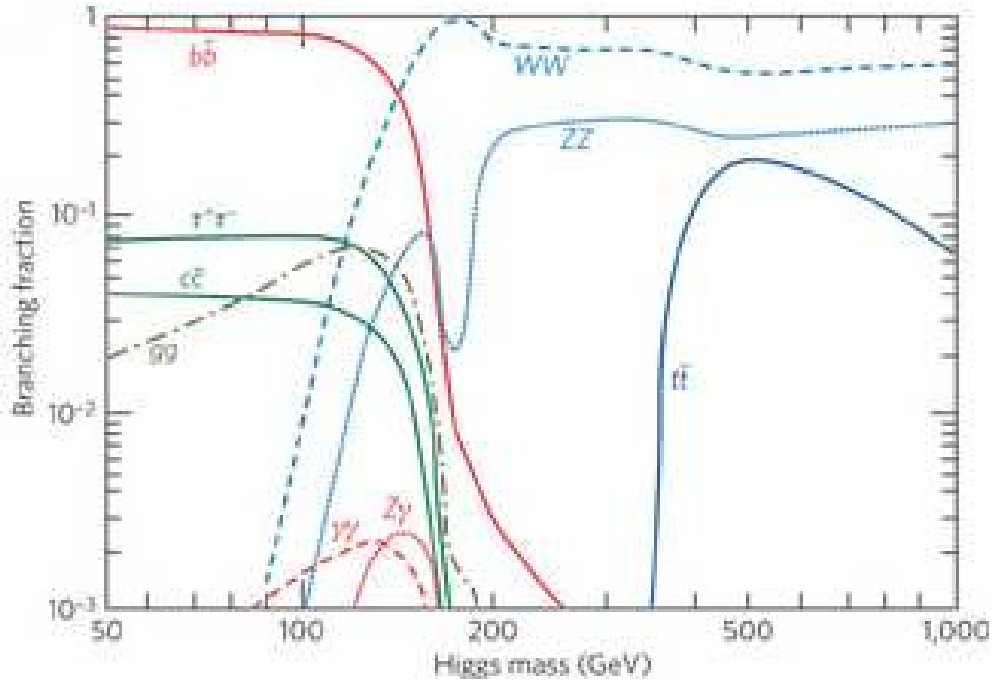


Figure 1.4: Branching ratio of the Standard Model Higgs boson as a function of its mass.

1.5.3 $H \to ZZ^{(*)} \to 4\ell$

The only irreducible background for this channel comes from the ZZ continuum production and the main reducible backgrounds are $t\bar{t}$ and $Zb\bar{b}$ with b quarks decaying semileptonically
Like in the previous channel the mass resolution is very good, $\approx 1\%$ for both CMS and ATLAS and the background can be easily estimated fitting the sidebands of the invariant mass distribution of data. The $\sigma\times$ BR for this channel is lower for $m_H < 2M_Z$, but increases when both Z bosons are real.

1.5.4 $H \to b\bar{b}$

This decaying mode can be exploited for small masses, $m_H < 130GeV$. In order to reduce the enormous background from $b\bar{b}$ coming from QCD, the channel $ttH \to \ell\nu qqbbbb$ is used and the signal events are triggered using the charged lepton from top decay. The jets must be matched to the originating partons by a kinematic fit in order to identify the correct $b\bar{b}$ combination corresponding to the Higgs boson decay.

The relative mass resolution on the reconstructed Higgs boson cannot

be reduced below about 10%, but it is possible to separate the signal from backgrounds using the presence of four b quarks in the final state.

1.5.5 Weak boson fusion

ATLAS has performed a study on the channels $qqH \rightarrow qqWW^*$ and $qqH \rightarrow qq\tau^+\tau^-$.
In qqH events two additional jets are produced which can be detected by forward calorimeters and give a distinctive signature. A jet veto can also be applied in the central region.
The main sources of background are the processes $Z+$ jets, $t\bar{t}$ and $WW+$ jets.

ATLAS[17] showed that the weak boson fusion channels provide a sensitivity that is similar to the $H \rightarrow \gamma\gamma$ and $H \rightarrow ZZ^*$ and that, in the MSSM model, it is less sensitive to a possible reduction of gluon-gluon fusion cross section.
The mass resolution, however, is not excellent and, in the case of the $H \rightarrow WW^*$ decays, only the transverse mass can be measured. This makes a background estimation from data harder. In addition the background estimation and the signal efficiency will be affected by large systematics uncertainties.

1.5.6 SM Higgs summary

The SM Higgs boson is expected to be found at the LHC with several decay channels over the full expected mass range in the CMS and ATLAS detectors. In the region 130 GeV $\lesssim$ M_H $\lesssim$ 500 GeV the discovery is possible already with an integrated luminosity of 10 fb^{-1} or less with the H$\rightarrow$WW*/WW and H$\rightarrow$ZZ*/ZZ decay channels. The light SM Higgs boson with $M_H \lesssim 150$ GeV is expected to be found in the inclusive H$\rightarrow \gamma\gamma$ channel, in the H$\rightarrow b\bar{b}$ in the associated $t\bar{t}$H production and the weak boson fusion channels with decays to H$\rightarrow \gamma\gamma$, H$\rightarrow$WW* and H$\rightarrow \tau^+\tau^-$ [18].

1.5.7 MSSM Higgs searches at the LHC

In the MSSM five physical Higgs bosons remain after eliminating unphysical phases: two neutral scalar h and H, one neutral pseudo-scalar A and two charged scalars $H^{\pm}$. At tree level all masses and couplings in the Higgs sector are determined by two independent parameters and the mass of the lightest Higgs boson h is predicted to be below M_Z. Radiative corrections modify those predictions, but the lightest Higgs boson h is still below $\approx$ 130 GeV for

any choice of Supersymmetric parameters. In addition for $M_A > 150$ GeV the three heavier Higgs bosons H, A and $H^{\pm}$ are approximately degenerate in mass.
For $M_A > 200$ GeV we are in the so called decoupling limit, where the h boson is very similar to the SM Higgs and the SM Higgs searches can be directely applied. Since the h boson mass is less then ≈ 130 GeV, the searches for the light SM Higgs mass apply. On the other hand, for $M_A = O(M_Z)$ and large $\tan\beta$ the H boson behaves like the SM Higgs boson and is also light. For large $\tan\beta$, in all other cases, the couplings of h and H to WW and ZZ are suppressed, while $A \to WW, ZZ$ is never allowed at the tree level, and h, H and A are produced in association with a $b\bar{b}$ pair and decay with almost 100% BR into $b\bar{b}$ and $\tau^+\tau^-$. Finally for large M_A and small $\tan\beta$ H and A predominantly decay into $t\bar{t}$ but for masses around 200-300 GeV we can also have the decays $H \to hh$ and $A \to Zh$.
The SM Higgs searches are effective in a large portion of the MSSM parameter space to find the scalar Higgs bosons h or H. In case a SM-like Higgs boson will be discovered, it will not be possible to distinguish it from a SM Higgs boson, but in a large portion of the parameter space, especially for large $\tan\beta$, it will be possible to detect also other MSSM Higgs bosons. On the other hand, in the decoupling region, it may be hard to disentangle the Standard Model from the MSSM.

The supersymmetric structure of the theory imposes constraints on the Higgs sector. As a result, all Higgs sector parameters at tree-level are determined by two free parameters, which may be taken to be $\tan\beta$ and m_A. One first consequence of this result is that there is a tree-level upper bound to the mass of the light CP-even Higgs boson, h: $m_h \leq m_Z|\cos 2\beta| \leq m_Z$.
If no Higgs boson is discovered at the Tevatron, the LHC will cover the remaining unexplored regions of the $m_A - \tan\beta$ plane, as shown in fig. 1.5.7. That is, in the maximal mixing scenario (and probably in most regions of MSSM Higgs parameter space), at least one of the Higgs bosons is expected to be discovered at the LHC. A large fraction of the parameter space can be covered in the search for a neutral CP-even Higgs boson by employting the SM Higgs search techniques, where the SM Higgs boson is replaced by h or H with the appropriate rescaling of the couplings. Moreover, figure 1.5.7 illustrates that in some regions of the parameter space, both h and H can be simultaneously observed, and additional Higgs search techniques can be employed to discover A and/or $H^{\pm}$. At LHC it will be either possible to exclude the entire $m_A - \tan\beta$ plane, eliminating the MSSM Higgs sector as a viable model, or achieve a 5σ discovery of at least the lightest of the MSSM Higgs bosons. In this case, however, in a significant region of the parameter

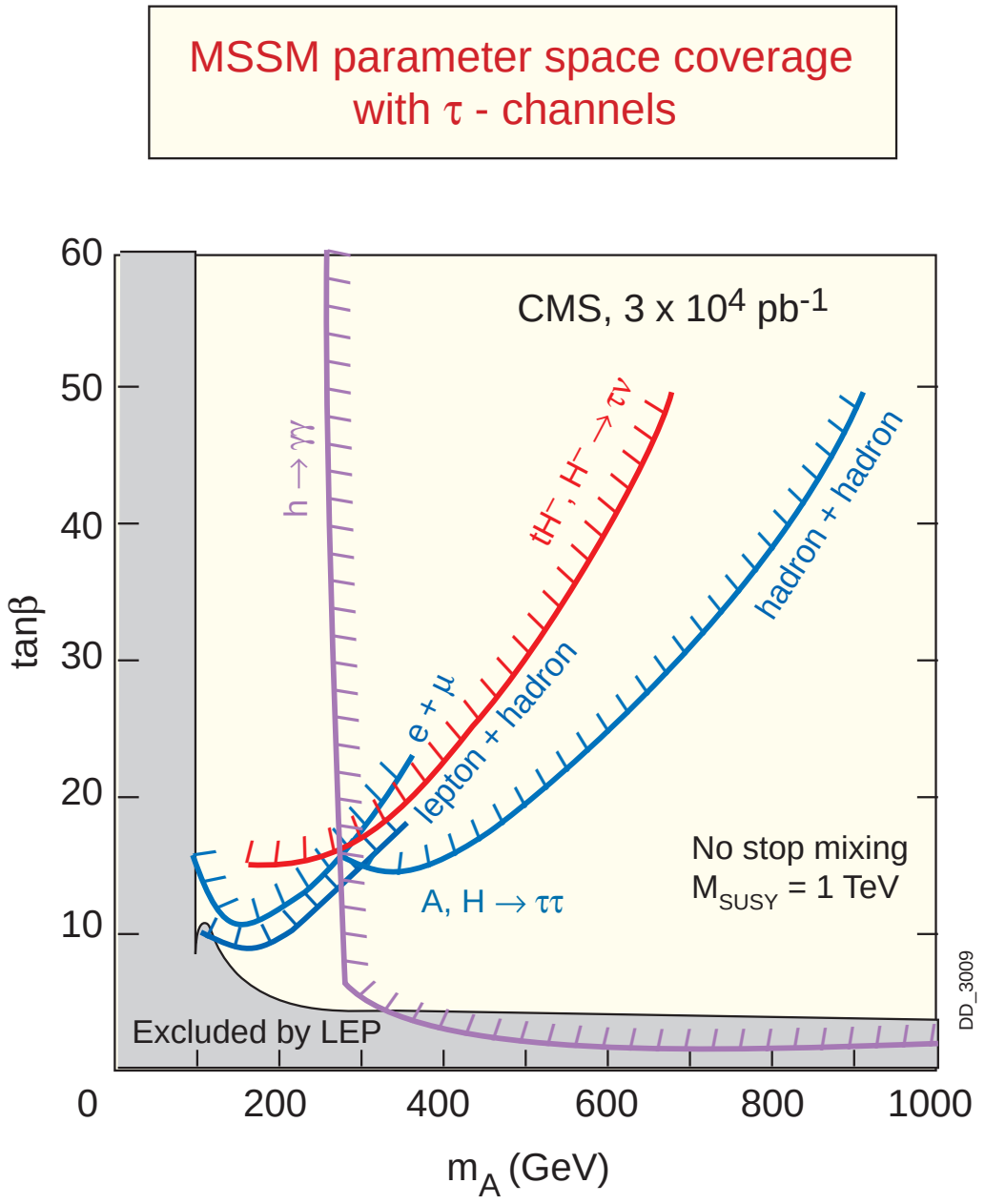


Figure 1.5: Expected 5σ discovery contours for MSSM Higgs boson detection in various channels in the $m_A - \tan\beta$ plane, in the maximal mixing scenario, assuming an integrated luminosity of $L = 30 fb^{-1}$ for the CMS detector [19].

space the heavier Higgs bosons cannot be discovered at the LHC. In this parameter region, only the lightest CP-even Higgs boson can be discovered, and its properties are nearly indistinguishable from those of the SM Higgs boson.

1.6 Higgs searches at the Tevatron

While the LHC experiments are certain to discover a Higgs boson with at most a few years of running, one must not forget that two other experiments are already searching that particle with all their means, and still have a chance to reach the goal before CMS and ATLAS are allowed to get into the game.

Many results of Higgs boson searches have been published by now by CDF and D0 using datasets corresponding to one or two inverse femtobarns of proton-antiproton collision at 1.96 TeV. These results are still unable to exclude any value for the Higgs mass, but they are quickly approaching that sensitivity. To understand what can be expected by the Tevatron in the next couple of years it is maybe best to compare present results with past expectations.

In 2003, the Higgs Sensitivity Working Group at the Tevatron conducted a study on the required integrated luminosity for CDF and D0 experiments to discover the Higgs boson, or exclude its presence. The study[20] was an update of a previous one[21] which had been based on simplified Monte Carlo simulations of the "average" Tevatron detector. In the 2003 study the availability of real data from CDF and D0 enabled a more careful determination of the discovery reach in the most sensitive search channels, specifically for light Higgs masses: the $WH \to l\nu b\bar{b}$ and the $ZH \to llb\bar{b}$, $\to \nu\nu b\bar{b}$ signatures. The results of the study confirmed the earlier claims that a combination of Tevatron results with a few inverse femtobarns of integrated luminosity could have enough sensitivity to discover a light Higgs or exclude it in the full mass range up to 180 GeV (see Fig. 1.6.

If we take the most recent combined limits[**?**] obtained by CDF and D0 for the 2008 winter conferences (shown in Fig. **??**, we are now in the position to verify their consistency with the early claims of discovery reach.

The following table shows the integrated luminosity of each search considered in the average, and the mass range where it was applied.

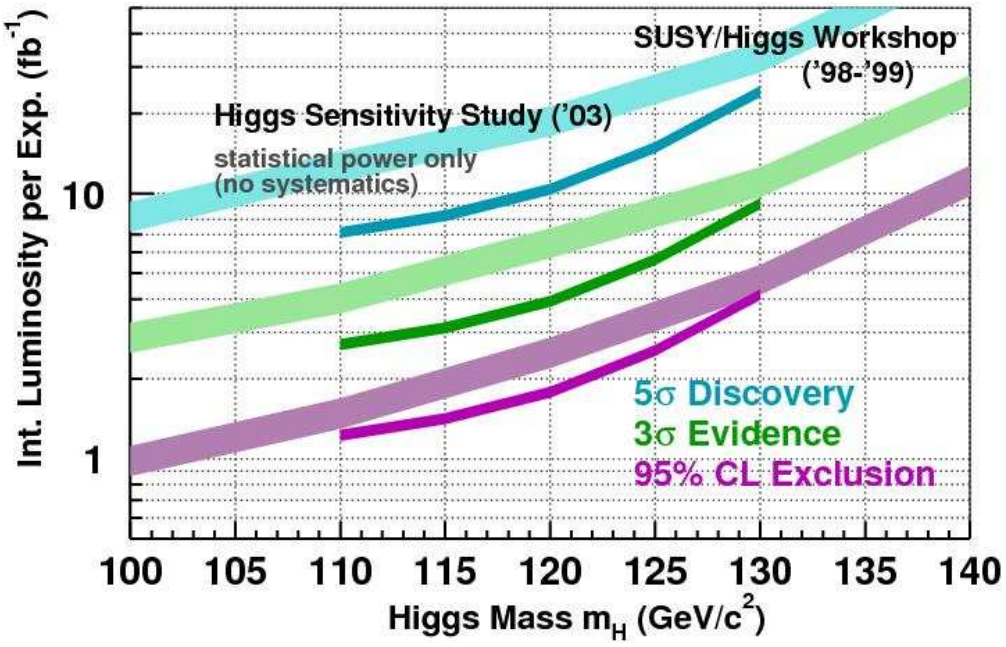


Figure 1.6: Reach of the SM Higgs search at the upgraded Tevatron. Shown are the integrated luminosities delivered per experiment which are required to exclude the SM Higgs boson at 95% CL, observe it at the 3σ level or discover it at the 5σ level, as a function of Higgs mass. In the low-mass Higgs region, below 140 GeV, the curves shown are the result of combining the W+Higgs and Z+Higgs channels (where the Higgs decays in $b\bar{b}$ and the W and Z decay leptonically), using a neural network based selection, and the statistical power of both experiments. In the high mass region, above 140 GeV, the curves are the result of combining various channels in which the Higgs boson decays in WW (where one W may be virtual). The lower edge of the bands is the calculated threshold; the bands extend upward from these nominal thresholds by 30% as an indication of the uncertainties in b-tagging efficiency, background rate, mass resolution, and other effects.[20].

Exp.	WH	$ZH \to \nu\nu b\bar{b}$	$ZH \to l\ b\bar{b}$	$H \to WW$	$WH \to WWW$
CDF	1.7/fb	1.7/fb	1.0/fb	1.9/fb	
	110-150	100-150	110-150	110-200	
D0	1.7/fb	0.9/fb	1.1/fb	1.7/fb	1.1/fb
	100-145	100-135	105-145	120-200	120-200

Given the higher sensitivity of the WH search at low mass and the $H \to WW$ one at high mass, one can average those numbers with an average luminosity per experiment of 1.4/fb at masses up to 140 GeV and 1.8/fb at higher masses. Now, if we instead examine Fig. 1.6 again, we extract the integrated luminosity per experiment predicted in 1999 (and confirmed in 2003) to exclude at 95% CL the Higgs boson at the Tevatron: we find the values listed in the fourth column in the table below. By dividing these values by the two "averages" of 1.4/fb and 1.8/fb one obtains the multiplicative factors in the fifth column. The square root of these are a prediction based on the 1999/2003 study of what "times-SM" ratio the Tevatron experiments would be expected to set with today's analyzed luminosity. These can be directly compared to the numbers in the second column, which are the expected numbers obtained from the summer 2007 combination (actual limits have been slightly different, but it is more consistent to compare 2003 expectations with 2007 expectations rather than with 2007 limits).

M_H(GeV)	2007 lim/SM	$L_{eff}(fb^{-1})$	HSWG 95% (fb^{-1})	R	$\sqrt{(}R)$	discrepancy
110	x3.8	1.4	1.5	x1.1	x1.05	3.5
120	x5.0	1.4	2.5	x1.8	x1.4	3.5
140	x4.2	1.4	10	x7.0	x2.6	1.6
160	x1.9	1.8	4	x2.2	x1.5	1.3
180	x2.9	1.8	8	x4.4	x2.1	1.4

From the exercise one learns that at low mass (110,120 GeV) the Tevatron is doing 3.5 times worse than it predicted, while at higher masses it is only doing marginally worse. This is not a surprise: all low-mass Higgs searches require the continuous refining of several critical tools (b-tagging, jet energy resolution, lepton acceptance improvements, more refined analysis strategies). The Tevatron experiments have shown in the past that time is a very important factor: as an example, the most precise top mass measurements in Run I have only been published six years after the end of data taking. For that reason, the early predictions of the Higgs Sensitivity Working Group cannot be dismissed too quickly, even in light of the rough comparison we carried out; moreover, it is to be noted that the curves in Fig. 1.6 are only the av-

erage of many pseudoexperiments which have a large variance: luck is also a factor in these searches! We therefore hold that it remains possible that the Tevatron still competes with the LHC in the forthcoming years, especially if the latter is delayed in starting data-taking. In the next few years we will most likely assist at a very interesting race for the Higgs particle.

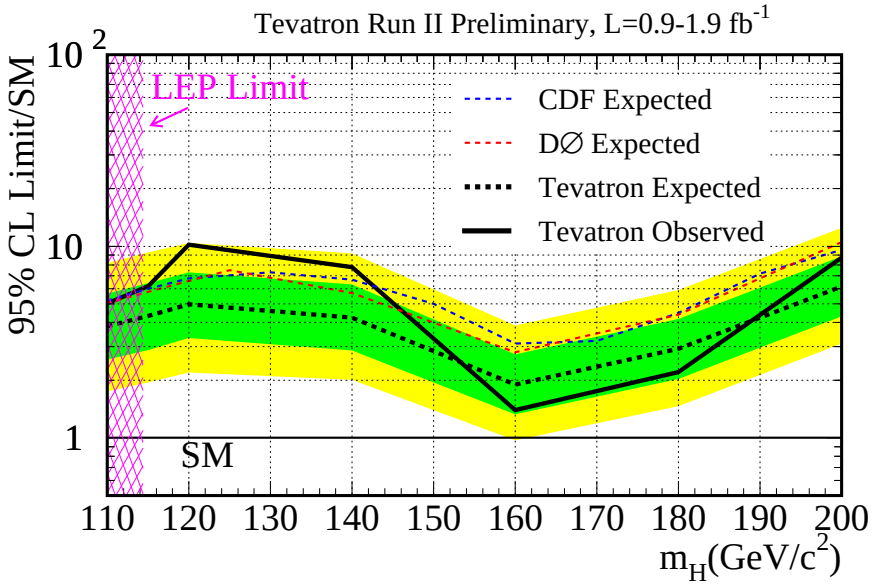


Figure 1.7: Combined limits on SM Higgs boson production by Tevatron experiments[**?**]

Chapter 2

The CMS experiment at LHC

The high energies and luminosities required to extend our knowledge of particle physics at the TeV scale can presently only be reached by hadron colliders. Hadron colliders produce a large number of low energy particles resulting in a not clean environment if compared to lepton colliders. However they allow to study a wide energy range and they provide high production rates. For these reasons, after the end of the successful LEP II program, CERN started the construction of the Large Hadron Collider (LHC) and the associated experiments.

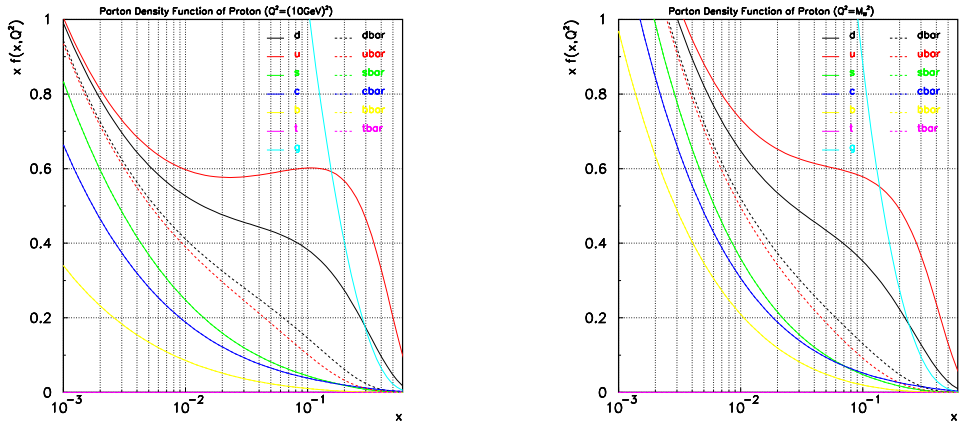
2.1 Phenomenology of proton-proton collisions

Protons are not elementary particles. Unlike leptons, they are composed with three valence quarks (uud) surrounded by a "sea" of quarks and gluons produced mainly by gluon radiation from valence quarks followed by gluon splitting $g \to q\bar{q}$.
In inelastic proton collisions the interaction involves partons inside the proton. If $\sqrt{s}$ is the total center of mass energy for the proton-proton collision, the two partons will have x_a and x_b fractions of it and the parton-parton effective center of mass energy is $\sqrt{\tilde{s}} = \sqrt{x_a x_b s}$. Since x_a and x_b are not fixed, the effective center of mass energy varies in every collision giving the possibility to explore a wide energy range.
The total cross section of a generic hard interaction can be written in the form

$$\sigma = \sum_{a,b} \int dx_a dx_b f_a(x_a, Q^2) f_b(x_b, Q^2) \tilde{\sigma}(x_a, x_b) \ , \tag{2.1}$$

where the sum is over all the partons a and b of the two protons and $\tilde{\sigma}(x_a, x_b)$ is the cross section of the elementary interaction between the partons a and b.

$f_i(x_i, Q^2)$ is the probability of finding the i parton carrying a fraction x_i of the proton momentum with exchanged four momentum Q^2 during the interaction and is called Parton Density Function (PDA) [22]. PDFs are different for the gluons, u and d valence quarks and low momentum quark-antiquark pairs of the sea. They are function of the exchanged four momentum Q^2 since for higher values shorter distances are probed and the contribution of gluons and sea quarks increases. Figure 2.2.2 shows the PDFs for two different values of Q^2.



(a) (b)

Figure 2.1: Parton Density Function for a proton with $Q^2 = (10\text{GeV})^2$ (a) and $Q^2 = M_W^2$ (b).

As the fraction of energy carried by each parton is not necessarily the same, the effective center of mass is boosted along the beam direction, hence it is more convenient to use variables invariant under boosts along the beam direction to describe the dynamics of the interaction. The most important variables are the transverse momentum P_T defined as the magnitude of the momentum $\vec{p}$ projection on a plane perpendicular to the beam axis, and the rapidity which, taking z as the beam axis, is given by

$$y = \frac{1}{2} \ln \frac{E + p_z}{E - p_z} \ , \tag{2.2}$$

where E is the energy and p_z the momentum of the particle along the beam axis. The rapidity distribution dN/dy is invariant under boosts along the z axis. In the ultrarelativistic approximation $m/|\vec{p}| \ll 1$, the rapidity can be

expanded to

$$y = \frac{1}{2}\ln\left[\frac{1+\cos\theta+\frac{1}{2}\left(\frac{m}{|\vec{p}|}\right)^2+o\left(\left(\frac{m}{|\vec{p}|}\right)^2\right)}{1-\cos\theta+\frac{1}{2}\left(\frac{m}{|\vec{p}|}\right)^2+o\left(\left(\frac{m}{|\vec{p}|}\right)^2\right)}\right] \simeq -\ln\tan\left(\frac{\theta}{2}\right) \equiv \eta \ , \quad (2.3)$$

where $\cos\theta = \frac{p_z}{|\vec{p}|}$. We have defined the pseudorapidity η, which in the ultra-relativistic limit is approximately equal to the rapidity and has the advantage that it depends only on θ and can also be used for particles of unknown mass. The use of protons as colliding particles poses a major challenge to the capability of calculating total and differential cross sections and consequently on the accuracy of the simulation of the physical process with Monte Carlo programs. The asymptotic freedom, i.e. the decreasing of the α_s with the increasing of the scale, allows to factorize the hard interaction as an incoherent scattering of the single partons, but the long range phenomena cannot be treated perturbatively. To avoid infra-red soft and collinear singularities producing logarithmic divergent terms, one has to consider only inclusive quantities. In particular, when considering hadronic initial states the arising singularities have to be absorbed by the parton density function of the proton (PDF), which is dependent on the exchanged momentum (Q^2).
The proton PDF have been computed rather precisely fitting deep inelastic scattering and Drell-Yan production data, but the (x, Q^2) range explored by the LHC extends by orders of magnitude the one covered by other colliders as shown in figure 2.1.

In a $p-p$ collision, long range interactions are important in both the initial and the final state. In the initial state, in addition to the hard scattering, other soft interactions can occur between the spectator partons originating the so called "underlying event", which receives also a contribution from initial state radiation from the hard interaction. In the final state, each colored parton hadronizes, i.e. stabilizes into color singlet hadrons (jets). All these processes cannot be computed perturbatively and have to be modeled on the data.

2.2 LHC

The LHC resides in the 26.7 km LEP [23, 24] tunnel placed $\sim$ 100 meters underneath the country area around CERN in Geneva. It will accelerate two proton beams up to 7 TeV producing collisions in four interaction points along the circle with a center of mass energy ($\sqrt{s}$) of 14 TeV. It will be fed by existent accelerators. In a circular collider of radius ρ the energy loss per

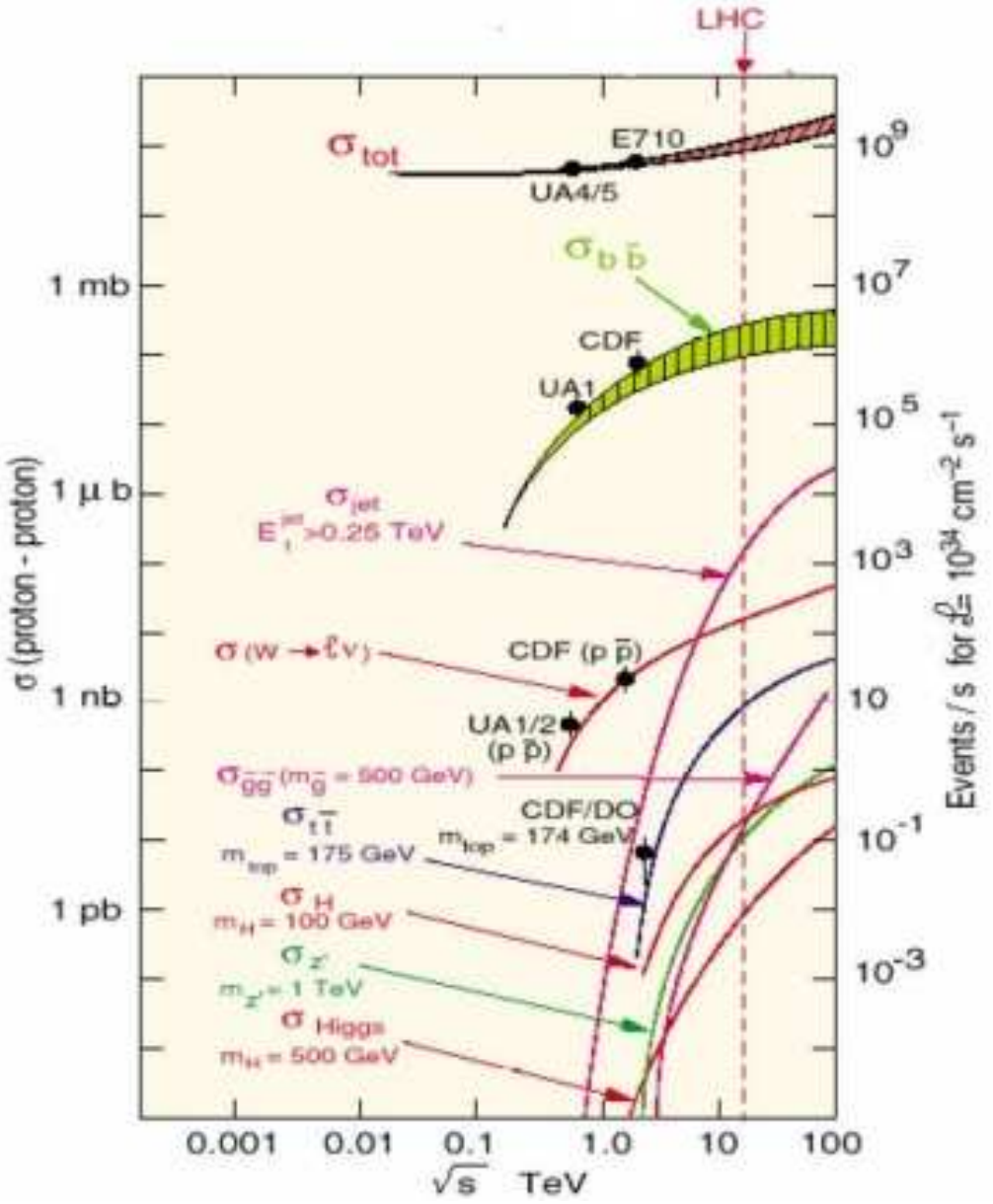


Figure 2.2: Cross sections for various processes with protons in the initial state as a function of mass energy.

turn due to synchrotron radiation is proportional to $E^4/(m^4\rho)$, where E and m are the energy and mass, respectively, of the accelerated particle. The 200 GeV reached by the electron beam in LEP were the limit for the given radius of 4.3 km of the tunnel. To reach higher center of mass energies protons have to be used, since the synchrotron radiation for them is $(m_p/m_e)^4 \sim 10^{13}$

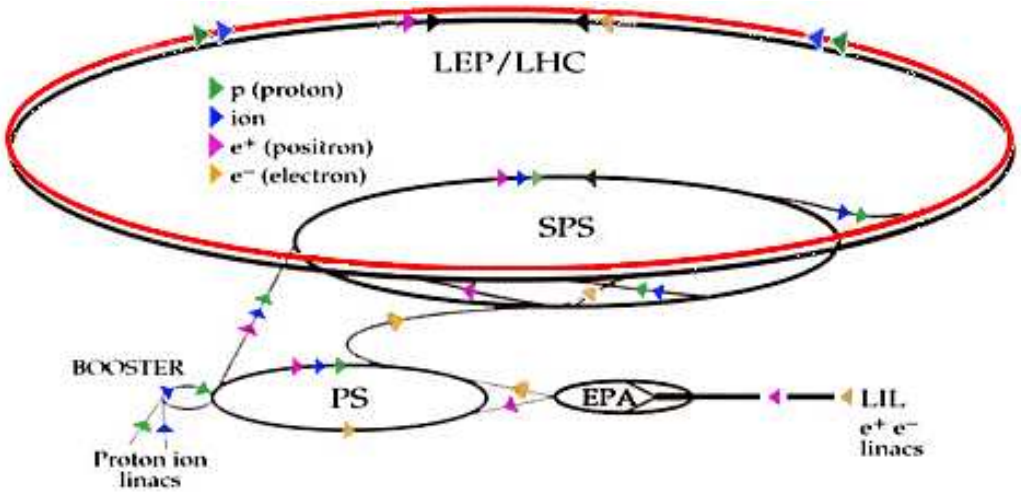


Figure 2.3: LHC schematic view.

smaller than for electrons. LHC will have very few radio-frequency cavities, placed in only one place along the machine and providing 0.5 MeV/turn to the beams. Furthermore, the high beam intensities implied by a luminosity of $\mathcal{L} = 10^{34} cm^{-2} s^{-1}$ exclude the use of antiproton beams and one common vacuum and magnet system for both circulating beams (as it is done in Tevatron) and implies the use of two proton beams. This is not a disadvantage from a physics point of view, given the higher center of mass energy; on the contrary, it allows an easy and fast population of both beams with a high number of particles.
For these reasons the LHC is designed as a proton-proton collider with separate magnet fields and vacuum chambers in the main arcs and with common sections only in the insertion regions where the experimental detectors are located. The proton beams will have a center of mass energy of 14 TeV. The LHC will also be able to collide heavy ions such as Pb at a center of mass energy of 2.76 TeV/u.

To reach the $\sqrt{s} = 14$ TeV value the protons will be accelerated and brought up to 50 MeV by a linear accelerator Linac2. A proton synchrotron booster then raises the beam energy up to 1.4 GeV and injects the proton beams into the old circular accelerator PS (proton synchroton). The last acceleration is due to the super proton synchrotron (SPS) which finally injects 450 GeV proton beams in the LHC.

The value of center of mass energy is limited by the magnetic field needed to bend the beams and the fixed radius of the LEP tunnel ($\rho = 4.3km$). The LHC tunnel is $26.659km$ long with 8 curvilinear sections $2.804km$ and 8 rec-

tilinear sections, where the beams collide.
In the curvilinear sections 1232 main superconducting dipoles operating at 1.9K and generating a magnetic field up to 8.33 T will be used to bend the beams. 386 quadrupoles, 360 sextupoles and 336 octupoles will provide stability control.
The acceleration to the particles will be provided by superconducting radiofrequency cavities placed in the linear sections, which will have electric fields from 3 MV^{-1} at injection to 16 MV^{-1}.

Circumference	26.659	
Maximum Dipole field	8.33 T	
Magnetic Temperature	1.9 K	
	p - p	$^{82}_{208}$Pb - $^{82}_{208}$Pb
Beam energy at injection	450 GeV	73.8 TeV
Beam energy at collision	7 TeV	574 TeV (2.76 ATeV)
Maximum Luminosity	$1 \times 10^{34} cm^{-2} s^{-1}$	$2 \times 10^{27} cm^{-2} s^{-1}$
Number of Bunches	2808	608
Bunch spacing	7.48 cm	5.3 cm
Bunch separation	24.95 ns	124.75 ns
Number of particles per bunch	1.1×10^{11}	8×10^{7}
Total crossing angle	300 μrad	$< 100 \mu$rad
Bunch length (r.m.s.)	7.55 cm	7.5 cm
Transverse beam size at Impact Point	16.7 μm	15 μm
Luminosity lifetime	10 h	4.2 h
Filling time per ring	4.3 min	9.8 min
Energy loss per turn	7 keV	
Total radiated power per beam	3.8 kW	
Stored energy per beam	362 MJ	

Table 2.1: LHC, technical parameters

The machine luminosity depends only on the beam parameters and, in the case of gaussian shaped bunches, it can be written as [25]

$$\mathcal{L} = F \frac{\nu n_b N_1 N_2}{4\pi \sigma_x \sigma_y} \ , \tag{2.4}$$

where ν is the revolution frequency, n_b the number of bunches per beam. $F = 0.9$ is a correction factor due to non-zero crossing angle (lower than 300 μrad),

N_1 and N_2 are the number of protons in the two colliding bunches, σ_x and σ_y the beam profiles in horizontal and vertical directions at the interaction point. The number of events for a selected "i" process generated in the LHC collisions is given by

$$N_i = \sigma_i \int \mathcal{L} dt \ , \tag{2.5}$$

where σ_i is the cross section for the process under study and $\int \mathcal{L}$ is called integrated luminosity and is measured in b^{-1}, $1b = 10^{-24} cm^2$. For the low luminosity period it is expected to collect $20 fb^{-1}$ per year for three years, and the high luminosity period will then last at least five years giving $500 fb^{-1}$.

2.2.1 Cross sections and luminosity requirements

The interesting processes at LHC have cross sections of order of fb and to be able to detect them it is necessary to increase the statistics as much as possible. The event rate for a given process i, R_i, can be written as

$$R_i = \sigma_i \cdot \mathcal{L} \ , \tag{2.6}$$

where σ_i is the cross section and $\mathcal{L}$ the instantaneous luminosity. The cross section depends on the center of mass energy as shown in figure 2.1 and it grows rapidly for the Higgs production while it remains almost constant for minimum bias production. This suggests to take the biggest center of mass energy possible. To raise $\mathcal{L}$ affects all the processes, however, it must be taken big enough to have good statistics for the rare processes.
At $\sqrt{s} = 14$ TeV the total inelastic cross section estimated is $\sigma_{pp}^{in} = 55mb$, leading to $5.5 \times 10^8 evs^{-1}$ ($1.1 \times 10^8 evs^{-1}$) at high (low) luminosity $\mathcal{L} = 10^7 mbs^{-1}$ ($\mathcal{L} = 2 \times 10^6 mbs^{-1}$). The average number of inelastic scatterings per event is then given by

$$\mathcal{N}_{int} = \frac{R}{f(1-e)} = \frac{\sigma_p^{in} p \cdot \mathcal{L}}{f(1-e)} \ , \tag{2.7}$$

where f is the beam collision frequency and is equal to $(1/25)ns = 40$MHz and $e = 0.2$ is the fraction of empty bunch crossings. This means an average of 17.2 (3.4) inelastic interactions per each bunch crossing at high (low) luminosity. Every minimum bias event will bring on average seven charged particles in the central pseudorapidity region with a mean transverse momentum $p_T = 0.5$ GeV and an almost flat distribution between +6 and -6 η. In addition, an average of 8.3 primary photons per unit of rapidity will be produced.

This demands finely space-segmented detectors and also high time resolution. To reduce minimum bias interactions keeping $\mathcal{L}$ constant it is necessary to reduce the number of particles in each bunch and increase the collision frequence. This requires very fast detectors to reduce the pile-up of events from consecutive bunch crossings.

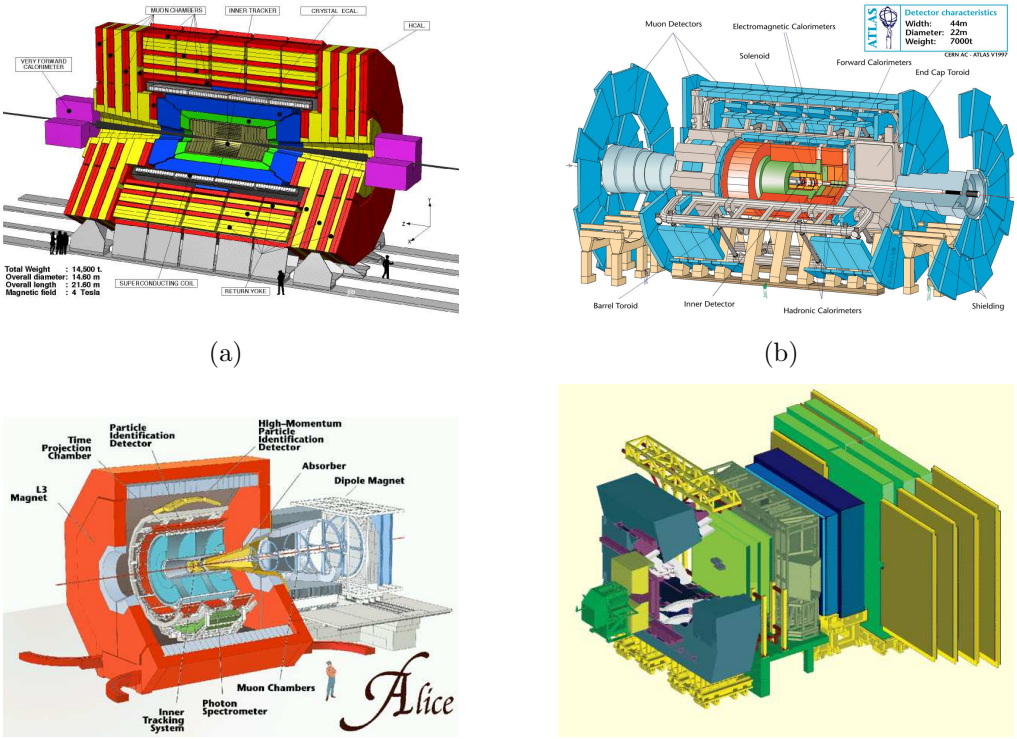
2.2.2 Experiments at the LHC

Four experiments will be installed in the caverns surrounding the interaction points[1]. Two of them are devoted to specific topics: ALICE (A Large Ion Collider Experiment) [26] will investigate heavy ion collisions and LHC-b (LHC Beauty experiment) [27] will study b-physics and will perform precision measurements of CP violation. The other two are general purpose experiments. Their design differs significantly in the configuration of the magnetic field. ATLAS (A Toroidal LHC ApparatuS) [28] uses a toroidal configuration, produced by three sets of air-core toroids complemented by a small inner solenoid, while CMS (Compact Muon Solenoid) [29] uses a solenoidal field generated by a big superconducting solenoid.
A toroidal magnet guarantees a constant resolution in pt as a function of pseudorapidity. A very large air-core toroid leads to a very good resolution even without the inner tracker, but it requires an excellent alignment of the detectors which must also be very precise. The iron-core solenoid design, on the other hand, can generate a very intense field, allowing to have a compact magnet and insert the calorimeters inside it. This improves the detection and energy measurement of electrons and photons. The constant field within the magnet and inside the return yoke allow precise tracking. Multiple scattering, however, degrades the resolution of the muon system.

2.3 CMS

The Compact Muon Solenoid (CMS) is a general purpose experiment that will operate at the LHC. The first requirement for CMS was that of a compact design, which leads to the choice of a strong magnetic field, obtained through a 4 T superconducting solenoid.
The design priorities of CMS are a good and redundant muon system, a good electromagnetic calorimeter and a high quality central tracking. The solenoid

[1] A fifth experiment called TOTEM, dedicated to luminosity measurement and diffractive physics is also present. It is placed close to CMS, with some of the detectors inside the CMS forward hadronic calorimeters.



(a) (b)

(c) (d)

Figure 2.4: 3D representations of the 4 experiments installed at LHC: (a) CMS, (b) ATLAS, (c) ALICE and (d) LHC-b.

can accommodate the calorimeters, allowing precision measurements of electrons and photons. The 4 T magnetic field allows precision tracking in the all-silicon inner tracker and it also reduces pile-up from soft hadrons in the muon system installed in the magnet return yoke.

CMS is designed to identify the very energetic particles emerging from the proton-proton collisions, and to measure as efficiently and precisely as feasible their trajectories and momenta [30]. Weekly interacting particles like neutrinos or some Dark Matter candidates cannot be directly detected and their presence must be inferred from the balance in the total energy-momentum of the reconstructed event. The detector acceptance, therefore, should be as high as possible, in order to have a precise estimation of the

missing transverse energy (E_t^{miss}), in particular the hadronic calorimeters should reach up to $|\eta| \sim 5$. The cracks in the subdetectors (needed for the read-out and high/low voltage cables and the cooling pipes) should also be as limited as possible.
High resolutions and granularities in the electromagnetic calorimeters are required to be able to observe sharp invariant mass peaks over continuous background. To extract the Higgs boson signal in the $\gamma\gamma$ decay channel over the $\gamma\gamma$ irreducible background, for example, an accuracy on two-photon invariant mass, $m_{\gamma\gamma}$, smaller than 1% is needed.
High particle fluxes pose a significant challenge to the central tracking. The adoption of a strong magnetic field and the use of silicon pixel detectors (closest to the beam axis) and silicon microstrip detectors for the inner tracker, which provide high granularity at all radii, allow to keep the occupancy under 2-3% even at high luminosity and the impact of pile-up is therefore minimal.
One of the most difficult constrains on the detectors comes from the LHC bunch crossing rate of 40 MHz. To avoid the integration of signals coming from different bunch crossings, each subdetector must have a response time of the order of 25 ns. This constraints the detection principles (no large Time Projection Chambers or no organic scintillators can be used) and requires also sophisticated and highly-performing front-end electronics.

CMS has the typical structure of general purpose experiments designed for a collider. It has a cylindrical shape made by several cylindrical layers coaxial to the beam direction, referred to as barrel layers, closed at both ends by detector disks orthogonal to the beam pipe, the endcaps, to ensure detector hermeticity. The full length of the detector is 21.6 m, with a diameter of 15 m and the total weight is $\sim$ 12500 t. A schematic view is shown in figure 2.3 and 2.3.

The CMS coordinate frame is a right-handed cartesian system with the x axis pointing to the center of the LHC ring, the z axis coincident with the CMS cylinder axis and the y axis directed almost upwards[2] along the vertical.
Due to the cylindrical symmetry of CMS, however, the algorithms use a "pseudo-cylindrical" coordinate system based on the distance r from the z axis, the azimuthal angle ϕ with respect to the y axis and the pseudo-rapidity η defined in 2.3. CMS subdetectors are installed radially from inside out:

- **Tracker r $<$ 1.2 m $|\eta| < 2.5$** Silicon pixel vertex is the innermost detector and is essential for precise vertex reconstruction and b-tagging and has to deal with a very high track density. In the baseline de-

[2]Since the beams are 1.23% inclined with respect to a plane perpendicular to the direction of the gravity force vector, the y axis is not exactly parallel to the vertical.

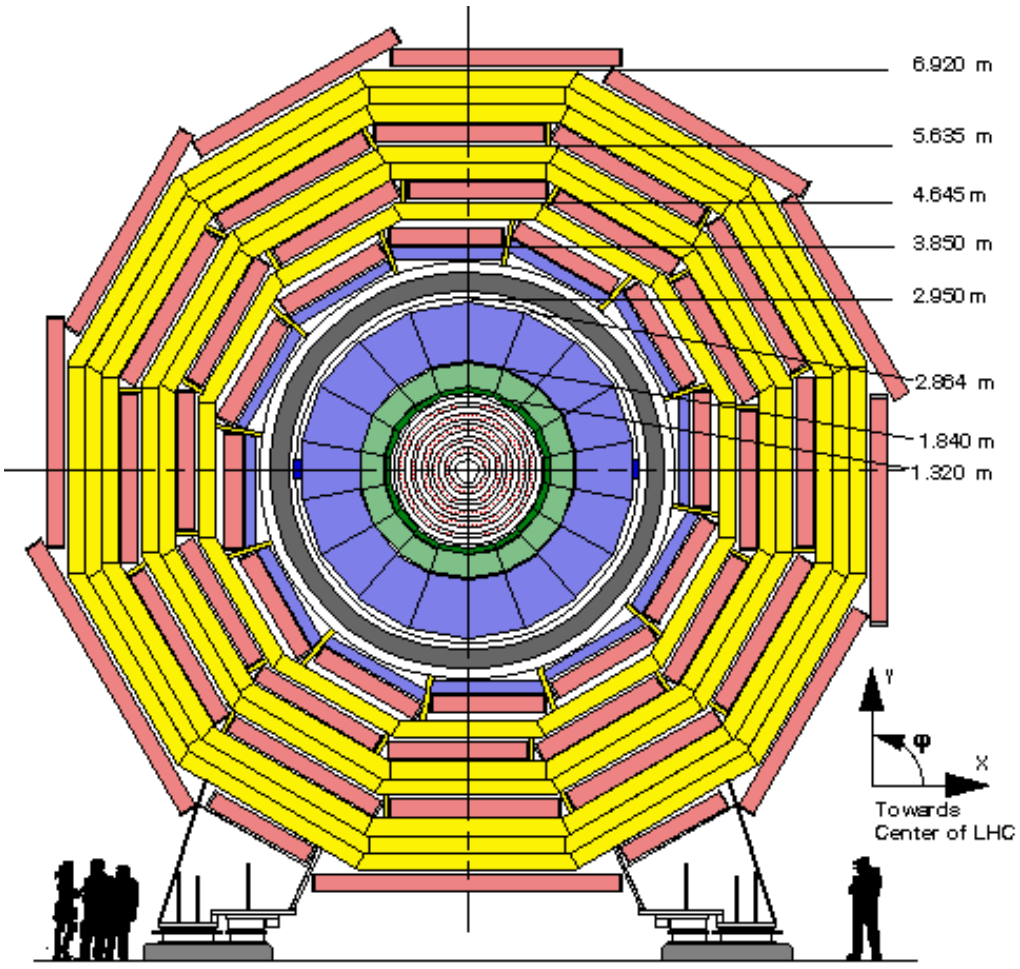


Figure 2.5: Transverse view of the CMS detector.

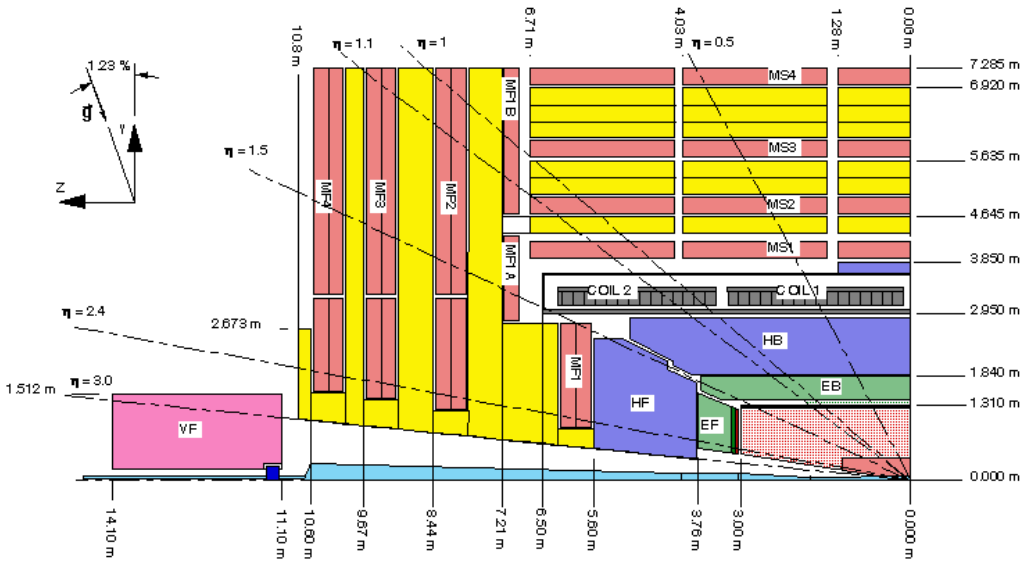


Figure 2.6: Longitudinal view of the CMS detector.

sign it consists of 3 barrel layers and 2 forward disks and it extends up to $r \sim 11$ cm. Outside the pixel detector, a silicon strip detector is installed allowing charged tracks reconstruction with up to 12 measurement points.

- **ECAL 1.2 m < r < 1.8 m $|\eta| < 3$** The electromagnetic calorimeter (ECAL) is composed by $PbWO_4$ scintillating crystals covering the region $|\eta| < 3.0$ (EB,EE). In the endcaps, it will be supplemented by a lead/silicon preshower detector, to improve the resolution in the determination of electron and photon direction and help pion rejection.

- **HCAL 1.8 m < r < 2.9 m $|\eta| < 5.3$** Jet and energy impbalance are measured by a sampling hadronic calorimeter (HCAL) installed just before the coil. It is composed of a copper alloy and stainless steel instrumented with plastic scitillators. The $|\eta|$ coverage of the barrel and endcap components (HB,HE) are the same as the ECAL, but they are complemented by a very forward calorimeter (HF), which extends the coverage up to $|\eta| < 5.3$ and enhances the hermeticity of the detector and its ability to measure missing transverse energy.

- **Muon System 4.0 m < r < 7.4 m $|\eta| < 2.4$** The muon system is placed inside the magnet yoke and is composed by drift tubes detectors in the barrel region and cathode strip chambers in the endcaps (up to $|\eta| < 2.4$), complemented by a system of resistive plate chambers with

coverage of $|\eta| < 2.1$.

In the following sections CMS subdetectors are described in more detail.

2.3.1 Tracker

The silicon tracker [31, 32] has the goal to reconstruct high-pt charged tracks in the region $|\eta| < 2.5$ with high efficiency and momentum resolution, to measure their impact parameter and to reconstruct secondary vertices. This is obtained with semiconductor detectors made of silicon covering the largest ever designed Si detector surface of 198 m^2. The pixel detector is the closest to the interaction region and the most crucial for impact parameter measurement. For these reasons it is finely segmented and uses thin detectors and overdepleted silicon bulks to cope with the fast primary charge collection required. The pixel detector will be described in greater detail in the next chapter.
The pixel detector is surrounded by the single-sided and double-sided microstrip detectors extending up to $r = 110$ cm and $|z| = 270$ cm and collectively called silicon microstrip detector. The inner part of this detector consists of four barrel layers and three small forward disks. The outer part consists of six barrel layers and nine forward disks. The redundancy is guaranteed by the overall design which allows many measured points per track within an acceptable material budget not to impairing too much the electromagnetic calorimeter performance. in this way an average of 12-14 hits per track are guaranteed to allow a high tracking efficiency and a low rate (10^{-3} or less) of fake tracks, which are reconstructed tracks not corresponding to any real track.
The full tracker consists of about 15000 microstrip detectors, with a pitch size ranging from 80 to 180 μm. As shown in figure 2.7 there are layers with single sided and layers with double sided the modules. The double sided modules are composed by two detectors mounted back-to-back with the strips rotated by 100 mrad. These double-sided ("stereo") modules will also provide a measurement in the coordinate along to the strips.
The high particle density and the high flux in the tracking volume of neutrons evaporated from nuclear interactions in the material of the electromagnetic calorimeter, make radiation damage to the tracker sensors a serious issue. To contrast the malfunctioning caused by radiation damage, both pixels and microstrips detectors have to be kept at a working temperature of -10^o C for the whole tracker volume.

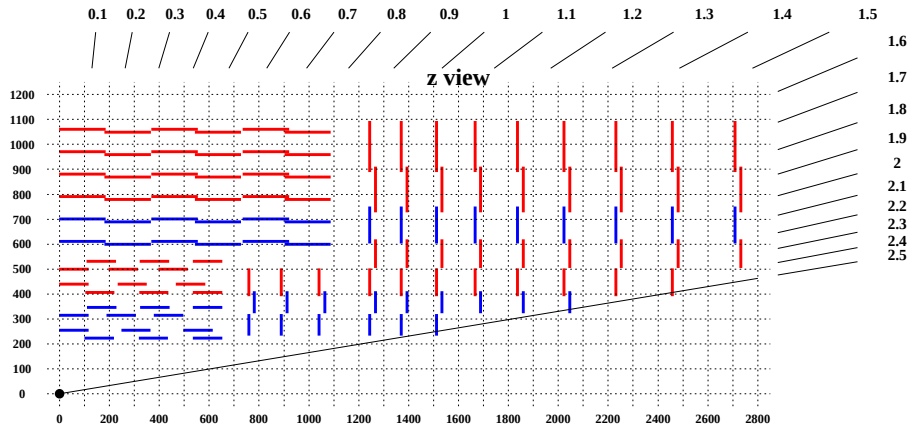


Figure 2.7: Sketch of the tracker layout (1/4 of the z view). Red lines represent single modules, blue lines double modules.

2.3.2 ECAL

In a general purpose LHC experiment it is fundamental to have a high performance electromagnetic calorimeter for precise measurements of electrons and photons. In CMS the ECAL design [33] has been driven by the requirement of a 1% two-photon mass resolution in order to allow the observation of a low-mass Higgs ($\leq$ 150 GeV) in the $\gamma\gamma$ channel and the possibility to measure the electrons and positrons from the decays of Ws and Zs originating from the $H \rightarrow ZZ^{(*)}$ and $H \rightarrow WW$ decay chain for 140 GeV $\leq m_H \leq$ 700 GeV.
A homogeneous calorimeter has therefore been chosen, composed with about 80000 finely segmented lead tungstate ($PbWO_4$) crystals. Lead tungstate is a radiation-hard, fast scintillator characterized by a small Molière radius[3] (21.9 mm) and a short radiation length[4] (X_0 8.9 mm), that allows good shower containment in the limited space available for the ECAL. Moreover, the lead tungstate has also a short scintillator decay time $\tau = 10ns$ that allows to collect 85% of the light in the 25 ns interval between two pp collisions. Scin-

[3]The Molière radius is the transversal dimension length scale of an electromagnetic shower evolving within a calorimeter.

[4]The radiation length X_0 is defined as the distance over which an high energy electron loses on average all but $1/e$ of its energy within a material.

tillator light is collected by silicon avalanche photo-diodes (APD) in the case of barrel crystals, and vacuum photo-triodes (VPT) for the endcaps crystals. The calorimeters is composed by trapezoidal crystals with a square face of 22×22 mm^2 in the barrel and 30×30 mm^2 in the endcaps matching the Molière radius. The length of the crystals is 230 mm in the barrel and 220 mm in the endcaps, corresponding to 25.8 and 24.7 radiation lengths respectively.
The granularity is $\Delta\eta \times \Delta\phi = 0.0175 \times 0.0175$, high enough to efficiently separate π^0 from γ.

The endcap crystals are shorter, but in front of them a preshower with 3 radiation lengths in the two regions $1.65 < |\eta| < 2.6$ is forseen. Each preshower is composed with two lead radiators and two planes of silicon microstrips to increase the π^0 rejection power in the highly irradiated forward regions, which are also affected by the reduced performance due to the increase in granularity at higher $|\eta|$, with a maximum of $\Delta\eta \times \Delta\phi = 0.05 \times 0.05$ in the very forward crystals.

Within the range $25 < E[\text{GeV}] < 500$, which is valid for $H \to \gamma\gamma$ decay, the electromagnetic energy resolution of a calorimeter can be parameterized as

$$\left(\frac{\sigma}{E}\right)^2 = \left(\frac{a}{\sqrt{E}}\right)^2 + \left(\frac{\sigma_n}{E}\right)^2 + c^2 \ , \tag{2.8}$$

where the first term $a/\sqrt{E}$ is referred to as the stochastic term and parameterizes the effects of fluctuations in photo-statistics and shower containment. The second term σ_n/E describes the noise from the electronics and pile-up and c is a constant term related to the calibration of the calorimeter. The different contributions to the calorimeter resolution are shown in figure 2.8.

2.3.3 HCAL

The hadron calorimeter [34] is responsible, together with the ECAL, to measure the energy and direction of jets, the transverse energy E_T and the imbalance of tranverse energy, or missing transverse energy, $\not{E}_T$. For this purpose high hermeticity is required, which is obtained complementing the barrel and endcaps parts of the HCAL with a very forward calorimeter (HF) placed outside the magnet yokes, ± 11 m away along the beam direction from the nominal interaction point, giving a total coverage of $|\eta| < 5.3$.
A schematic view of the HCAL is shown in figure 2.9.
The barrel and endcaps HCAL is a sampling calorimeter with 3.7 mm thick active layers of plastic scintillators alternated with 5 cm thick brass plate absorbers and read out by wavelength-shifting fibres. The first layers read

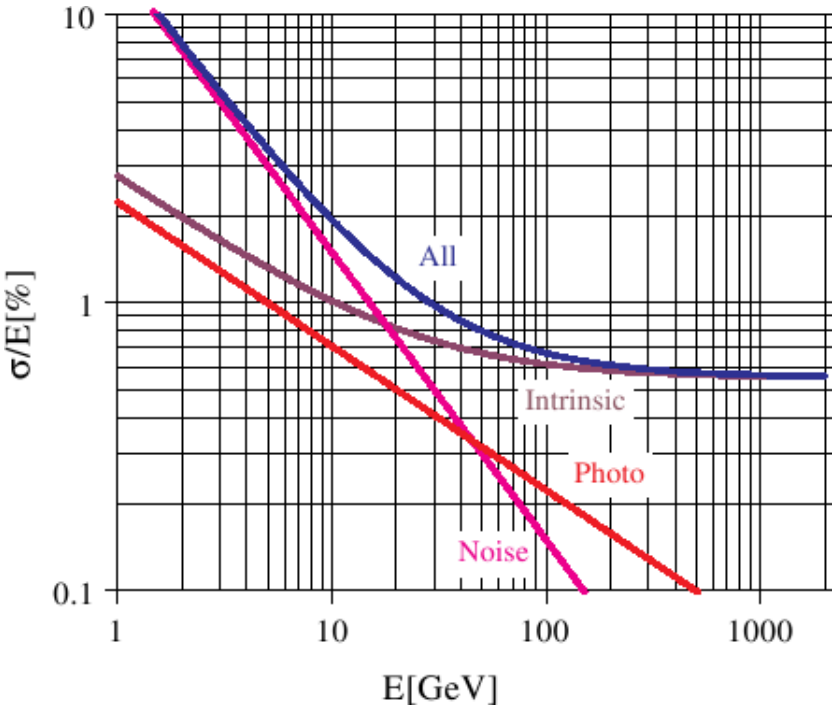


Figure 2.8: Different contributions to the energy resolution of the $PbWO_4$ calorimeter.

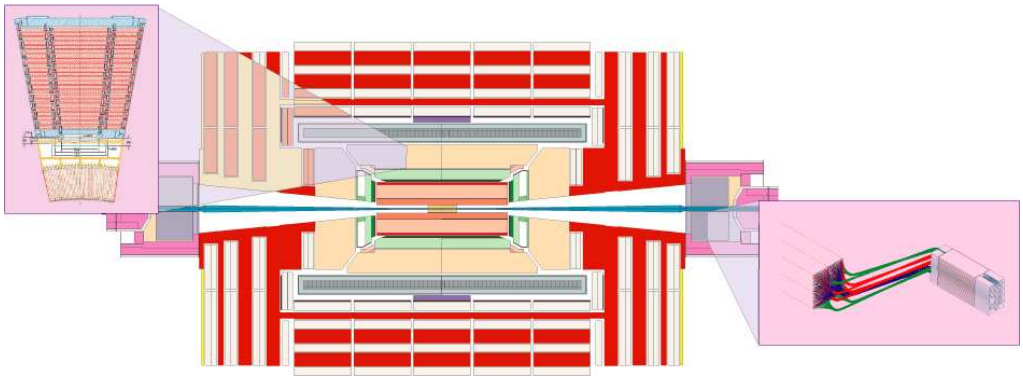


Figure 2.9: Schematic view of the hadronic calorimeter with details of the barrel and the HF components.

out separately, while all others are read out together. Both barrel and end-caps are read-out in towers with a size of $\Delta\eta \times \Delta\phi = 0.087 \times 0.087$, which is fine enough to allow an efficient di-jet separation. The thickness goes from

8.9 to 10 interaction lengths[5]. The thinner part is in the barrel region and is not sufficient to contain the full length of very energetic showers, thus an additional "tail catcher" made of scintillators tiles is placed outside the magnet.

The active elements of the HF are quartz fibres parallel to the beam, inserted in steel absorber plates.

The energy resolution is $\sigma/E \sim 65\%\sqrt{E} \oplus 5\%$ in the barrel; $\sigma/E \sim 85\%\sqrt{E} \oplus 5\%$ in the endcaps and $\sigma/E \sim 100\%\sqrt{E} \oplus 5\%$ (E in GeV) in the HF.

2.3.4 Muon system

The muon system [35] is used to identify muons and allows, in combination with the inner tracker, an accurate measurement of their transverse momenta. It is placed outside the magnet, embedded in the iron return yoke to make full use of the 1.8 T magnetic return flux. Since high p_T muons are clear signatures of many physics processes, the muon system plays an important role in the CMS trigger system and it is also used to measure precisely the bunch crossing time [36]. The muon system is divided in three independent subsystems: in the barrel, where the track occupancy is relatively low (< 10 Hz/cm^2), drift tubes (DT) detectors are installed, while in the endcaps where the particle rates are higher (> 100 Hz/cm^2) and a larger residual magnetic field within the yoke plates is present cathode strip chambers (CSC) are used. These systems are arranged in a multilayer structure to allow efficient rejection of single hits produced by low range particles, and they cover the $|\eta| < 2.4$ region.

In the $|\eta| < 2.1$ region redundancy is provided by resistive plate chambers (RPC), which, despite a limited space resolution, have a fast response and excellent time resolution, less than 3 ns. They are used for unambiguous bunch crossing identification and also complement the DT+CSC measurement of p_T at trigger level, since they are finely segmented.

A schematic representation of the muon system is shown in figure 2.3.4. Drift tubes are made of parallel aluminum plates insulated from perpendicular "I" shaped aluminum cathodes by polycarbonate plastic profile. The anodes are 50 μm diameter stainless steel wires placed between the "I" cathodes. A binary mixture of 80% Ar and 20% CO_2 at atmospheric pressure is used to fill the internal volume, because this gas is nonflammable and can be safely used in the underground locations of CMS. The resolution is about 100 μm

[5]The nuclear interaction length λ_0 is the mean free path for a hadron before having a nuclear interaction inside a material.

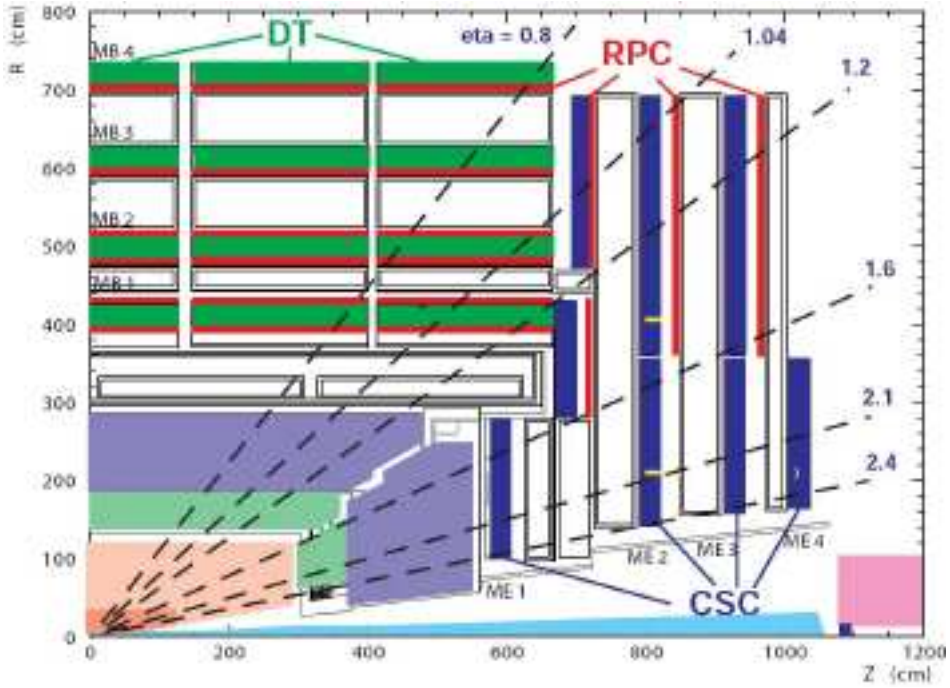


Figure 2.10: Longitudinal view of one quarter of the CMS muon systems.

both in $r\phi$ and rz views.

Cathode strip chambers are composed with arrays of anode wires between pair of cathode planes, segmented into strips perpendicular to the wires. The gas used to fill the gaps is a mixture of 30% Ar, 50% CO_2 and 20% CF_4. The measurement of the ϕ coordinate is performed interpolating the signal of neighboring strips, this allows a 50μm resolution.

Resistive plate chambers are structured in planes of phenolic resin (bakelite) with a bulk resistivity of $10^{10} \div 10^{11}\Omega$ cm, separated from aluminum strips by an insulating film. A nonflammable gas mixture of 94.5% freon ($C_2H_2F_4$) and 4.5% isobutane (i-C_4H_10) fills the gaps.

2.3.5 Trigger System

The maximal event rate at CMS is 40 MHz and the raw event size is about 1 Mb. It is not possible, both for practical and technical difficulties, to permanently store information for all the events. As we have seen in the previous chapter, the rate of interesting events is considerably small, with the exception of b physics, this poses the challenge of selecting these rare events among all those produced every 25 ns.
The maximum rate at which it is possible to store event informations is 100

Hz, therefore it is necessary to implement a trigger system to reduce the incoming rate of a factor of 4×10^5.
The CMS trigger system is a two-levels system dividing the selection in a first step, called level-1 trigger, in which the rate is reduced to 100 kHz (50 kHz in the low luminosity case), and a subsequent high-level trigger (HLT) which brings the rate down to 100 Hz.
High p_T leptons alone cannot be used as a good selection criteria, since $W \to \ell\nu$ and $Z \to \ell\ell$ would saturate the channel. Therefore, the trigger must be able to select events on the base of **their physics**, and online selection algorithms must have a level of complication comparable to that of offline reconstruction.

The first level selection must be very quick, having to cope with a 40 MHz rate of incoming events. The 25 ns avaliable to the level-1 trigger, however, are not even sufficient to read the raw data from all the detectors. For these reasons, the level-1 trigger is implemented on dedicated programmable hardware and makes use only of the calorimeter and muon data, but with coarse granularity.

In case of positive decision from the level-1 trigger, data is temporarily stored and passed to the HLT which uses commercial processors, organized in a farm of personal computers. Dedicated algorithms will run to select events on physics basis and will represent the first step of physics analysis selection. Using a parallel processing scheme up to 100 kHz of events can be processed, 500 ms/ev being the estimated maximum allowed processing time on a single CPU.

At startup the data acquisition system (DAQ) will be able to handle an event rate of up to 50 kHz, which will be increased to 100 kHz when the full LHC design luminosity is reached. Only one third of this bandwith is allocated, the rest being used as safety margin accounting for all uncertainties in the simulation of the basic physics processes.

2.3.5.1 Level-1 Trigger

The level-1 trigger [37] uses an asynchronous pipelined structure of processing elements, each taking less than 25 ns to complete its computations, to run dead-time free. At every bunch crossing, each processing element passes its output to the next element and receives a new event to analyse. During this process, the complete event is stored in pipeline memories, whose depth is technically limited to 128 bunch crossings. The level-1 decision is taken after a fixed time of 3.2 μs, including the transmission time between the detector and the counting room (up to 90 m of cables each way) and, in the case of drift tubes detectors, the electron drift times (up to 400 ns). Thus,

the effective time available for calculations can be as low as 1 μs. Figure 2.11 shows a logical view of the level 1 trigger structure.

The two elements of the level-1 trigger, calorimetry and muon system,

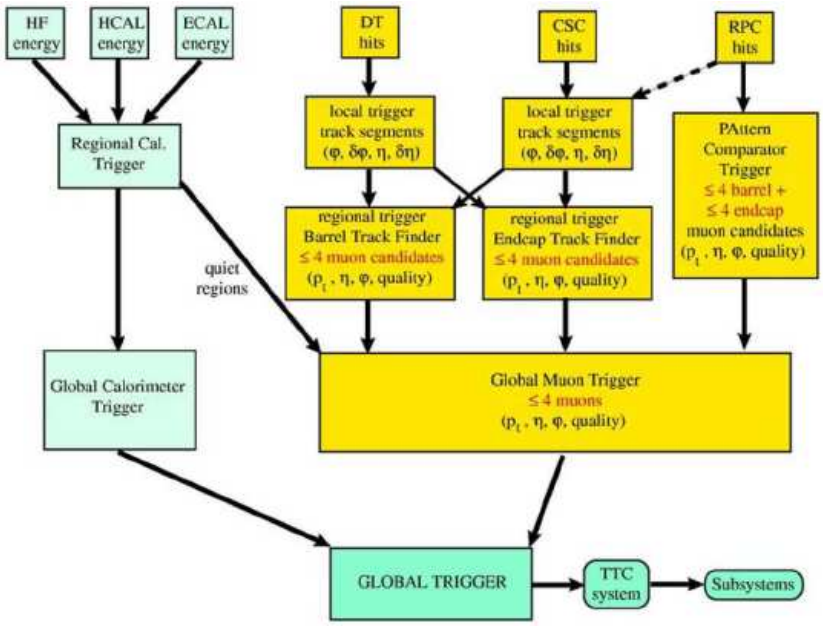


Figure 2.11: Overview of CMS level 1 trigger.

work in parallel and analyse the data locally. These triggers identify "trigger objects" of different types, they are: isolated and non-isolated electrons/photons; forward, central and τ-jets; and muons. The four best candidates of each type, together with the measurement of their position, transverse energy or momentum and a quality word, are passed to the global trigger. The global trigger receives also the total and missing transverse energy measurement from the calorimeter trigger.
It then selects events according to programmable thresholds. Topological conditions and correlations between objects can also be required. A total of 128 conditions can be tested in parallel and each can be pre-scaled to accept only a fraction of the events passing the cuts.

Level-1 Calorimeter Trigger The calorimeter towers read by the calorimeter trigger have a size fo $\Delta\phi \times \Delta\eta = 0.087 \times 0.087$ up to $\eta \sim 2$. At higher pseudorapidity values $\Delta\eta$ increases up to 0.35. Trigger towers match the size of the HCAL towers up to $\eta > 1.74$ and above this value physical calorimeter towers have twice the ϕ dimension of the trigger tower. In the barrel ECAL,

each tower corresponds to 5×5 crystals, while in the endcaps they are arranged in a $x - y$ geometry, and a variable number of crystals is grouped, matching as much as possible the HCAL trigger tower boundaries. For the HF the trigger towers size is $\Delta\phi \times \Delta\eta = 0.348 \times 0.5$. Calorimeter regions are defined by 4×4 trigger towers, with a size of about $\Delta\phi \times \Delta\eta = 0.35 \times 0.35$. The HF does not participate in the electron/photon trigger, therefore it does not have a small ϕ binning. It has 18 HF ϕ divisions which exactly match the trigger boundaries of the 4×4 trigger tower regions in the rest of the hadronic calorimeter. The ϕ divisions are exactly four times the size of the towers in the hadronic barrel and endcap and the η divisions are approximately the size of outer hadronic endcap divisions. With this structure, the overlapping jet trigger extends seamlessly to $\eta = 5$. The missing transverse energy is computed using ϕ divisions of 0.348 for the entire (η, ϕ) plane.

The first element to process ECAL and HCAL data is the trigger primitive generator which is embedded in the calorimeter readout electronics. A peak finder algorithm provides bunch crossing identification, and for each tower the so-called trigger primitives are evaluated. They are the sum of the tranverse energy and a fine grain bit. In the ECAL the fine grain bit provides information on the lateral extension of the electromagnetic shower, and is used in the rejection of backgrounds by the electron trigger. The HCAL fine grain bit is used to denote the compatibility of the deposit with the passage of a minimum ionizing particle, and is set if the HCAL energy before conversion to the transverse scale is within a programmable range, of the order of [1.5,2.5] GeV.
The objects reconstructed by the calorimeter trigger are: isolated electrons/photons, non-isolated electrons/photons, central jets, forward jets and tau jets. For the four most energetic objects in each category the measured position and transverse energy is sent to the global trigger, together with a measurement of the total missing transverse energy. The calorimeter trigger also sends informations about the activity in all calorimeter regions to the muon trigger. Those informations are used to determine if the energy deposit is compatible with the passage of a muon (MIP bit) and if it is below a programmable threshlod (Quiet bit).

The algorithms reconstructing these objects are described in the following.

Photon and electron trigger Electron and photon candidates are indistinguishable at level-1, therefore are treated together. They are reconstructed with a sliding window algorithm using a 3×3 towers sized window. Identification requires a large energy deposition in one or two adjacent trigger

towers. Requirements on the lateral and longitudinal profile of the shower are set through the fine grain bit and the ratio of ECAL to HCAL energy deposits, respectively.
A candidate is labeled isolated on the basis of the energy deposits and ECAL fine grain bits in the eight towers around the center of the 3×3 window.
In each calorimeter region the highest-E_T isolated and non-isolated candidates are selected. The four most energetic objects of each type are sent by the global calorimeter trigger to the global trigger.

Jets and τ-jets The jet trigger is based on the sum of ECAL and HCAL transverse energy. The level-1 candidates are defined by the trigger towers with energy deposit higher than the neighboring eight. The measured jet transverse energy E_T is corrected online depending both on measured transverse energy and pseudorapidity values. The correction is performed using a second degree polynomial, whose coefficients are listed in some calibration tables [38]. In the following we will refer to corrected values of E_T whenever thresholds on jet transverse energy applied.

It can happen that low E_T jets from hard scattering can be mismeasured as high E_T jets. A possible reason is particles or a small jet from a pile-up interaction impact the calorimeter near one of the low E_T jets, creating a "fake jet" with higher E_T. Fake jets can also arise when particels from different interactions impact the calorimeter too close-by.
These are compared to jets with a good matching in (η, ϕ) plane with jets made at generator level. The number of fake jets becomes very small when only jets with $E_T \geq 30$ GeV are taken.

τ-jets are identified by their narrow profile. A jet candidate is identified as τ-jet if each of the nine regions of the window contain no more than two towers above a programmable threshold.
Jets are searched for separately in the central region ($|\eta| < 3$) and in the forward region ($3 < |\eta| < 5$), while τ-jets are only identified in the central one. The global calorimetric trigger sends the four most energetic candidates of each type to the global trigger, together with the number of jets above a programmable threshold.

Total and missing transverse energy The total energy is evaluated as the sum of the transverse energy of all ECAL and HCAL calorimeter regions. The missing transverse energy is determined from the sum of the E_x and E_y components of the deposit in each region, obtained from the E_T deposits using the coordinates of the center of the region. The total energy and the absolute value and ϕ direction of the missing energy are to then sent

to the global trigger.

Quiet and MIP bits For each calorimeter region, a "Quiet" bit (Isolation bit) is set if the transverse energy deposit in ECAL plus HCAL is below a programmable threshold. The MIP bit is set if the calorimeter energy is consistent with the passage of a minimum ionizing particle. Both bits are used in the Global Trigger to suppress background and to improve selectivity.

Level-1 Muon Trigger The level-1 muon trigger uses the three detector systems described in 2.3.4 to identify muons, reconstruct their position and transverse momentum and provide bunch crossing assignment with high purity and efficiency. The good spatial resolution of drift tubes and cathode strip chambers and the excellent time resolution of resistive plate chambers complement themselves and the redundancy provided allows a robust trigger with high efficiency and good background rejection.

In the case of DTs and CSCs, the information of each chamber is first processed independently by a local trigger step, where track segments are reconstructed. The regional track finders match the segments of the different stations, reconstruct muon tracks and estimate their p_T. The candidates found are then sent to the global muon trigger, with a word to indicate their quality. In the overlap region, DT and CSC segments are used by both track finders to allow for reconstruction of full tracks in each of the subsistems.
In the case of RPCs, the hits are collected by a pattern comparator trigger (PACT), which looks for predefined patterns. The PACT provides an estimate of the p_T of the muon, its position and a word to indicate its quality.

Each subsystem reconstructs up to four muon candidates and sends them to the global muon trigger. There they are matched and cuts on the MIP and Quiet bits in the corresponding calorimeter regions are applied.

The global muon trigger sends the four muons with the highest p_T to the global trigger.

2.3.5.2 High-level trigger

The CMS high-level trigger (HLT) [39] consists of dedicated software running on a farm of commercial processors. The high-level trigger needs to reduce the about 100 kHz rate incoming from the level-1 trigger to 100 Hz mass storage with "fast" algorithms. It refines the level-1 measurement of jets, leptons and photons through intermediate steps. The algorithms will reconstruct better the raw level-1 physics objects together with streams dedicated to particular physics channel identification, for instance related to b physics, selected including the information of other subdetectors. The tracker signals

and the pixel hits after zero-suppression are used, allowing track finding and primary vertex reconstruction.
The output rate to be stored on disks is divided into different categories (see table 2.3.5.2 for the initial low luminosity run) including selections following the level-1 triggering scheme but also HLT dedicated algorithms, which make use of b-tagging techniques and conditional track finding within silicon tracker.
The HLT system will have to process, on average, one event every 10 μs and to reduce by a factor 1000 the amount of data. Since each recorded event has an average size of 1 MB, the DAQ system must be able to feed data from the front-ends to the PC farm at a sustained bandwidth up to 100 kHz $\cdot$ 1 MB = 100 GBs^{-1}.
The required HLT computing power at start-up with a level-1 rate of 50 kHz has been roughly estimated to be 15000 1GHz Pentium III CPUs.
Assuming a total time of 20 hours data taking per day, a total disk space of 10 TB will be filled, per day, at full luminosity.

2.4 Pixel detector for the level-1 trigger

The CMS silicon strip tracker, built from a high number of channels and performing no zero suppression, cannot participate in the level-1 trigger decision since too much data would have to be processed in too little time. The CMS Pixel detector has an even higher number of channels but it performs zero suppression and could possibly be used at level-1 trigger.

The CMS tracker is 5.6 meters long, has a diameter of 2.4 meters and it is

Physics Object	CPU time per Level-1 event [ms]	Level-1 trigger rate [kHz]	Weighted CPU time [ms]
e/γ	160	4.3	44
μ	710	3.6	164
τ	130	3.0	25
Jets & $\not{E}_T$	50	3.4	11
e & jet	165	0.8	8
Inclusive b-jets	300	0.5	16

Table 2.2: High-level trigger low luminosity selection timing.

the most central part of the CMS detector. The original design was based on MSGC[6] plus silicon detectors (both pixel and strip) but was later changed to an all silicon tracker. The pixel detector is the closest to the interaction point, starting at a radius of about 4 cm, with three pixel barrel layers and two pixel forward disks.
The very high particle fluxes at such a small distance suggests the use of pixels for reconstruction of space point information with high precision. The capacity to deliver two or more hits per track over the full acceptance of the detector is fundamental for precise vertex determination and therefore for tagging of long lived objects such as b or c quarks and τ leptons and to distinguish them over the large background of light jets.
These are important requirements for detecting Higgs and SUSY decay final states and of course for conventional B-hadron physics and top physics.
The pixel detector will also play a crucial role in the high luminosity phase, where it will help in pattern recognition of the many tracks present in each bunch crossing. Track segments from the tracker will be confirmed or rejected by the pixel detector and the selected tracks will be extrapolated with precision to the vertex.

2.4.1 Pixel detector overview

The CMS pixel detector is composed of three layers (pixel barrel) and two end-disks (pixel forward endcaps). The three barrel layers are located at mean radii 4.4, 7.3 and 10.2 cm and are 53 cm long. The two disks are placed at 34.5 and 56.5 cm from the interaction point. The barrel has a total of 720 modules with 16 Read-Out Chips each and the endcaps have 672 modules with from 2 to 10 Read-out Chips. It is planned to have a two layer barrel only inserted at start-up (LHC low luminosity phase). The two layers will be the internal layers while the additional external one will be inserted for LHC high luminosity phase. At this luminosity the $\sim$4 cm layer will be replaced every two years, due to the $\sim$10 times increase in radiation level. The high radiation level decreases the performance of the silicon sensors. The sensors will have to work with reduced depletion depth after irradiation. The pixel Read-out Chip (ROC) was changed from 150μm $\times$ 150μm (DMILL design) to 150 μm $\times$ 100 μm (Deep Submicron (DSM) design). This should lower the depletion depth of the irradiated sensors, in particular for the $\sim$4 cm layer. Each pixel barrel layer is composed of two half-cylinders, which are made of ladders. Each ladder consists of eight pixel modules and each module contains sixteen pixel ROCs. The ROC is made of 52 x 80 pixel

[6]Micro-Strip Gas Counters

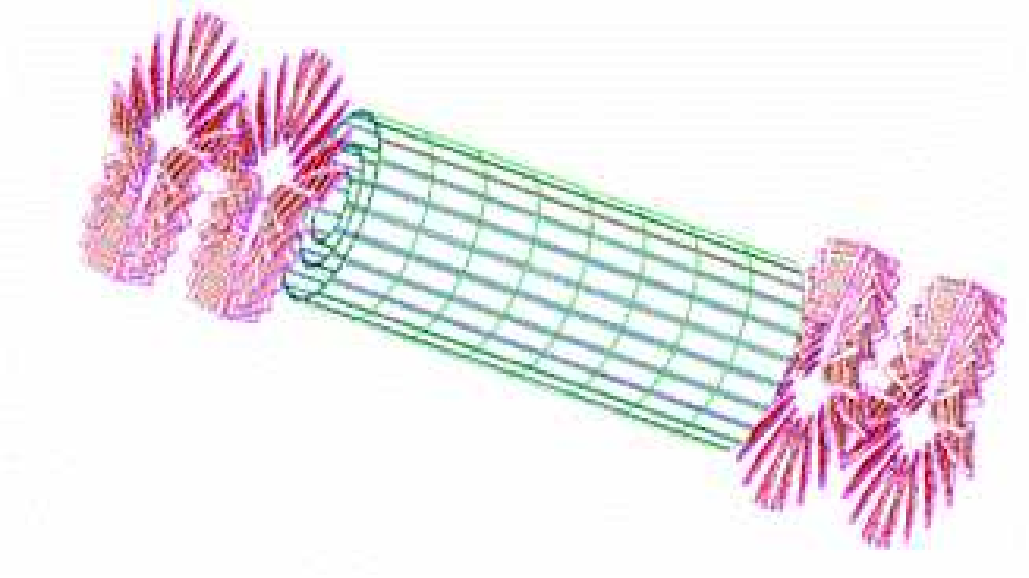


Figure 2.12: 3D view of the pixel detector in the final configuration.

unit cells (PUCs). The number of channels in the pixel barrel is 50 million

Pixel Barrel	Radius [mm]	Ladders	Modules	Chips	Pixels
Layer 1	41 - 45	18	144	2304	9.58×10^6
Layer 2	70 - 74	30	240	3840	16.0×10^6
Layer 3	107 - 112	46	368	5888	24.5×10^6
Overall		94	752	12032	50.1×10^6

Table 2.3: Main parameters of the CMS pixel barrel detector. Note that at the junction of the two half cylinders, two half ladders are counted together, and so are half-modules.

units with the new PUC geometry, while the complete pixel detector has 68 million channels. Analog pulse height information is recorded from each hit pixel. Offline, the pulse height information enhances the impact parameter reconstruction resolution by charge sharing exploitation. The geometry of the pixel detector is therefore tuned to maximize charge sharing effects. In the barrel, the modules will not be tilted so that the charge sharing due to the high magnetic field is important. In the end-disks, the detectors will be tilted by 20^o in order to enhance E $\times$ B effects. The pixel barrel module

has a sensor approximately 6.4 cm long, 1.6 cm wide and 300 μm thick. The silicon baseplate provides a stiff structure to the assembly. The 16 ROCs are bump-bonded to the silicon sensor in one PUC to one sensor unit cell correspondence scheme. The capton cable brings all signals to the module and routes the signals/data out of the modules. The power supplies are fed to the module from a power cable. The High Density Interconnect (HDI) provides the routing of the powers and of the signals at the module level. The capacitors filter the high frequency noise in the ROCs and the Token Bit Manager (TBM) handles the communication of the signals to the ROCs and organizes the readout of the data from the ROCs. The Module Trigger Chip (MTC) is a chip dedicated to the participation of the pixel detector to CMS First Level Trigger.

2.4.2 Pixel Read-Out Chip overview

The pixel ROC will operate in a high radiation environment; it is therefore built in radiation-hard technology. Leakage current in the sensor after irradiation must be absorbed by the analog circuit. Post-irradiation noise contributions must also be handled. The pixel size on the sensor is limited by the minimal PUC area achievable on the ROC. The reconstruction precision depends on the pixel size. To reach the small 150 μm $\times$ 100 μm PUC size, a high integration of the ROC electronics is needed. To achieve this, the array of 52 $\times$ 80 PUCs is organized in 26 Double-Columns of 160 PUCs. Each PUC samples the charge collected on the sensor. The sampling is performed in the analog part of the PUC. If the amount of charge is large enough, the comparator communicates to the digital part of the PUC that a hit is recorded. The digital part of the PUC initiates the immediate transfer of the PUC data to the Double-Column periphery. Communications from the PUCs to the Double-Column periphery occur through buses located between the two columns. The first signal sent from the hit PUCs to the Double-Column periphery is the IColOR, which notifies the Double-Column periphery that it should take action to record PUCs data. The PUC data are then stored into the Data Buffer, in the periphery, and wait either to be confirmed by a level-1 trigger, or to be disregarded. If the level-1 trigger accepts the event the data are transferred from the buffers to the ROC output. One hit pixel data consists of two analog levels to code the Double-Column number, three analog levels to code the PUC number inside the Double-Column and an analog pulse height information.

The hit resolution achieved in the barrel pixel detector is about 10 μm in ϕ and from 10 μm to 20 μm in z.

Offline the pixel detector can be used in studies requiring displaced ver-

tices, such as channels involving τ-jets and b-jets. For the High Level Trigger, the pixel seeding is essential for fast and accurate track reconstruction (see [40, 41]). It can also provide an efficient trigger on B decays (in combination with one silicon strip layer [42]).

If there is no level-1 trigger confirmation issued after 3.2 μs delay, the data are overwritten in all subdetectors.

2.4.3 Overview of the pixel first level trigger

The level-1 trigger must give an answer within its latency time of about 3.2 μs (about 128 bunch crossings $\times$ 25 ns). Subdetectors participating in the level-1 decision must provide relevant information within the trigger time latency.
Silicon strip detectors in CMS do not perform zero-suppression and generate a quantity of data too large to be processed in the short level-1 trigger time. Pixel detectors, instead, perform zero-suppression, but have high hit rates and an elevated number of channels, which cannot be handled by the DAQ at level-1.
A smaller number of channels can be analyzed, however, dropping the fine granularity of single pixels and looking at the ROCs. The ROC can be considered as an array of double columns, in which case it looks like a digital strip detector with 300 μm pitch, 8 mm long pseudo-strips. Figure 2.4.3 shows one such double column and the basic level-1 trigger logic implemented on the ROC.
Using double-columns the information is coarser, but can be handled at level-1 trigger. When a hit is recorded in a PUC, the digital part of the PUC initiates the transfer of the data from the PUC to the double-column periphery by sending a signal through the IColOR bus. A copy of the IColOR is made for level-1 trigger purposes and it is the basis of the level-1 trigger logic in each double-column.
At ROC level two variables can be defined: the total number of hit double-columns in the considered bunch crossing N_{DC} and the total number of clusters of hit double-columns in the considered bunch crossing N_{Cl}. A cluster of double-columns is defined as a group of adjacent hit double-columns.
These two variables are transmitted to the output of the ROCs and are collected at the module level. When the threshold on one of this variables is exceeded in a ROC, the ROC address is transmitted. Then, the level-1 trigger information, composed now of ROC addresses, reaches the reconstruction hardware through the optical fibres of the three layers of the pixel barrel detector. Figure 2.4.3 shows a sketch of the pixel level-1 trigger reconstruction.

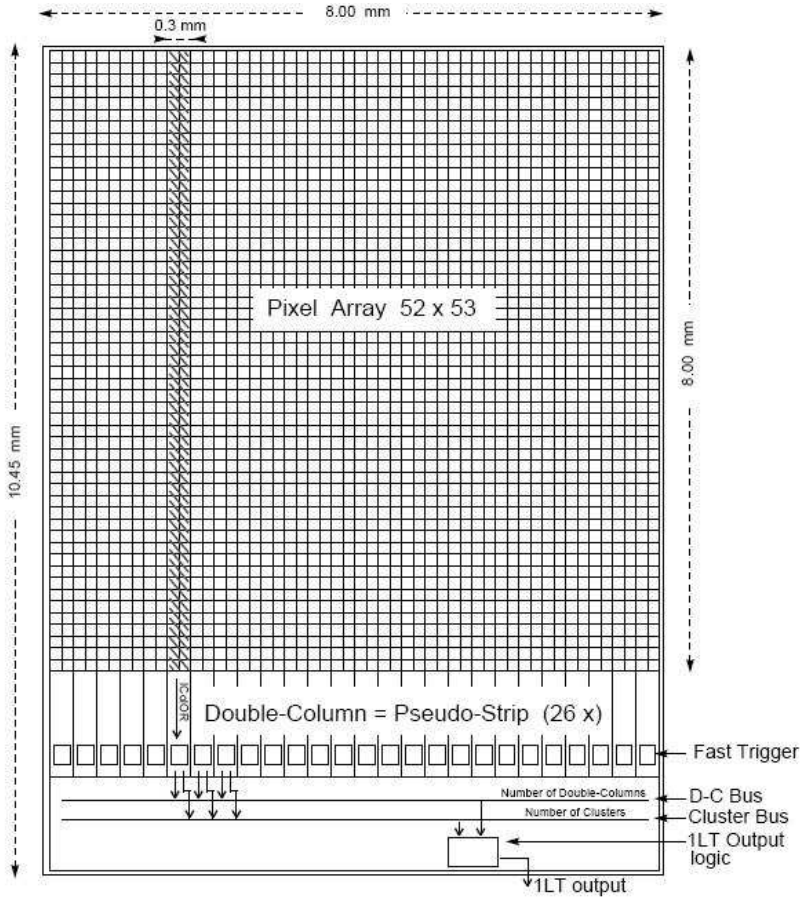


Figure 2.13: Sketch of the level-1 trigger logic in the ROC. From the "IColOR" signal, two variables are generated in the "Fast Trigger" cells. These are the "number of double-columns" and the "number of clusters" of double-columns taking part in the event. The ROC in the figure is the old DMILL technology model, the only difference with the final DSM technology version is the number of pixels in the vertical direction.

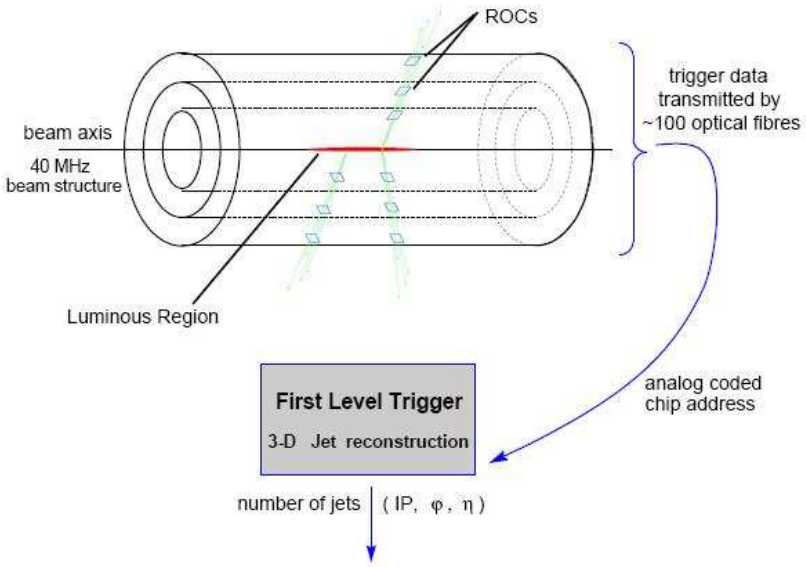


Figure 2.14: An example of a three jet event in the pixel barrel and the sketch of the pixel level-1 trigger reconstruction. Analog coded ROC addresses are passed to the reconstruction hardware with a high bandwidth.

Chapter 3

Pixel-based level 1 trigger

3.1 Introduction

In this chapter we will describe a possible application of the pixel detector for triggering at the first level of the CMS trigger. We will first introduce the software frameworks we used for all the simulations, with some details on the tools used in the production of the data samples. In the second part we will deal with the pixel trigger itself and we will derive optimized cuts for the selection of the $t\bar{t}$H signal, showing that the efficiency in the selection of channels with a high multiplicity of jets in the final state can be significantly increased keeping a fixed QCD rate from the trigger.

3.2 Software framework

In modern experiments using complex detectors it is important to have a detailed simulation of the detectors to study and optimize them for the data taking runs. In CMS the initial project used a group of packages to perform all the simulation from event generation to detector signals digitization.

The development of a framework for the simulation of the CMS detector and the reconstruction up to the final objects used in the analysis has started in 1998. The name of the object oriented framework is called COBRA and the collection of reconstruction and reconstruction code developed in this framework is called ORCA [43].
In older versions of ORCA, the simulation package, called CMSIM, was based on GEANT3 [44]. The transition to GEANT4 [45] was completed in 2003 and the new simulation package was called OSCAR.
All the simulations done with ORCA for this thesis are made using CMSIM. The acronym ORCA stands for **O**bject Oriented **R**econstruction for **C**MS

Analysis and it is an object oriented system for which C++ has been chosen as programming language.
The design of ORCA is based on CARF, the **C**MS **A**nalysis and **R**econstruction **F**ramework, which was developed to prototype reconstruction methods, initially for testbeam applications. All the parts in common between ORCA, OSCAR (**O**bject oriented **S**imulation for **C**ms **A**nalysis and **R**econstruction) and IGUANA (**I**nteractive **G**raphical **U**ser **ANA**lysis) are included via COBRA (**C**oherent **O**bject-oriented **B**ase for **R**econstruction, **A**nalysis and simulation), in particular CARF.
ORCA has two logical layers, *subsystems* and then within a subsystem *packages*. The packages within a subsystem typically produce a single library to cope with a certain aspect of the corresponding sub-detector or reconstruction utility.
The ORCA versions used in this thesys for the production of the samples are 8_6_0, 8_7_3, 8_7_1 and 8_13_3 (the last released version). ORCA versions have retrocompatibility, but the opposite is not true. Furthermore there was a significant change in the compilation instruments and Buildfile structure from version ORCA_8_7_3 to version ORCA_8_13_3 which forced the user to modify and rewrite significant portions of code.

Recently, the framework and the services it provides and the model of data storage underwent a comprehensive set of changes. The overall collection of software is now called CMSSW and we have used this for most of the high luminosity studies.
CMSSW uses again C++ as the main language and it has a more modular approach with respect to ORCA
The version of CMSSW used for all the simulations and reconstructions is 1.7.0, with the exception of some packages from different tags. At the time of starting the samples production, the software was in fact still in a quickly evolving form and we had to settle on a version which allowed to have all the components we needed. The simulation of the level 1 calorimetric jets, for example, was not in the Fast Simulation before versione 1.7.0 and was developed and included during the pre-releases of this version. The forward jets in particular were improved in the final released version, so that we managed to use them in the multijet trigger simulation.
The full list of the tags used is the following:

- JetMETCorrections/Configuration/data/MCJetCorrections152.cff − version 1.3
- FastSimulation/CaloRecHitsProducer − tag CMSSW_1_7_0_pre11

- FastSimulation/Configuration/data/FamosSequences.cff − tag CMSSW_1_7_0_pre11
- EgammaAnalysis/EgammaIsolationProducers − tag V00−00−03
- RecoEgamma/EgammaIsolationAlgos − tag V00−00−03

For CMSSW a hybrid simulation approach was chosen. Due to the very high computation time requirements of performing a full simulation and reconstruction of events with high luminosity pile-up, the Fast Simulation was chosen as the main simulation and reconstruction tool. This is explained in more detail in the following section.

3.2.1 FastSimulation

The CMS Fast Simulation is a tool to perform simulation and reconstruction of many events in a short amount of time, while still remaining accurate. The full (slower) simulation of the CMS detector is GEANT based and it is very detailed. The Fast Simulation have been tuned with the full simulation, together with the results of test beam data. It manages to be very quick by making some simplifying assumptions, parameterizations and also using dedicated reconstruction algorithms.
Many of the algorithms of the standard reconstruction can run directly on Fast Simulation data, since the final objects are the same. In some cases some of the values required by an algorithm are not produced and a dedicated version must be used. Dedicated algorithm are also used when the standard version is too slow.
We have used the default Fast Simulation present in version 1.7.0 of CMSSW, with the notable exception of the CaloRecHitsProducer which is from version 1.7.0 pre11. Also the FastSimulation/Configuration/data/FamosSequences.cff file, which defines all the modules configurations, was brought from version 1.7.0 pre11, but we modified it to make it work with CRAB[1].

The full chain to get the final reconstructed events starts with the production of events with a generator (for example Pythia) which are then injected in the Fast Simulation for the simulation of the detector response and the reconstruction of the final objects. The vertex smearing is performed internally and also the pile-up is added using minimum bias samples generated with Fast Simulation and available in the CMSSW release.
The track reconstruction algorithms used in the full simulation are too slow to be used without change. For this reason a fast tracking algorithm is used,

[1]CRAB stands for CMS Remote Analysis Builder and it is a tool used to submit, monitor and retrieve the output of jobs on the grid.

which starts from the simulated track hits to reconstruct the track. This leads to having no fake tracks, a good approximation with tracker tracks and low luminosity when much less than 1% of the tracks are fakes. The reconstructed hits are made smearing the simulated hit position, thus they do not have any information of the digitized hit, needed by the full reconstruction tracking algorithms.
In the barrel part of the pixel detector, the presence of only three hits per track is expected to give a significant number of fake pixel-tracks, especially in the high luminosity scenario, which would not be simulated by the Fast Simulation approach.
We decided to use the full simulation for the pixel detector while the Fast Simulation is used for the rest of the CMS detector. This appears to be the best tradeoff between computation time and accuracy of the reconstructed events, especially in the part of the detector crucial for this study.

3.2.2 Simulation and reconstruction of the event samples

The two fundamental requirements when simulating big event samples of high complexity, as in the case of high luminosity events, are computation time and space to store the events. In our case, the first step to reduce computation time was to use the Fast Simulation for most of the detector parts. The second step consisted in the high parallelization of the productions by use of CERN Batch Farm[2] and of the grid via CRAB. The jobs were split in blocks short enough to use the eight-hour queue of the batch system. The same event sizes were used when submitting to the grid.
It was not possible to store all the reconstructed event information, because the space required for one event would be $\sim$ 2.5 Mbytes, which for a million events leads to $\sim$ 2.5 Tbytes. We developed specific data formats to hold the needed information and reduce the space for one event to an acceptable $\sim$ 10Kbytes, thus giving a total of about $\sim$ 10 Gbytes. All the files were then stored on CASTOR[3].

For the low luminosity dataset the study was performed using ORCA and the datasets were either simulated locally or taken from the official CMS

[2]the CERN Batch Farm consists of 3000 machines providing computing power for data analysis and simulation tasks. It is based on a queue system with different queues for the various time requirements of the submitted jobs.

[3]CASTOR stands for CERN Advanced STORage manager and it is a hierarchical storage management system developed at CERN used to store physics production files and user files.

productions. All was done in full simulation.

The framework was later changed to CMSSW, and this was used for the high luminosity part. Going from no pileup to low-luminosity pileup conditions the increase in occupancy, for example, in the ECal is about 12-15%. Moving to high-luminosity condition brings an increase of about 110% in occupancy for the same detector. The time to fully simulate and reconstruct the events depends on the particular subdetector, and it is in general not linear, leading to a very long time needed. For example in CMSSW_1_3_1, to go from generation to full reconstruction on 30 events of $t\bar{t}H$ with high luminosity pileup and the standard Spring07 configuration file (without the roadsearch tracking algorithm) took more than 10 hours.

For the rest of the detector the fast simulation is used, since a good agreement is found with the full simulation and the time for production is more then three orders of magnitudes smaller. For high luminosity events this is especially true, on a 2GHz processor with the fast simulation and the fully simulated pixel part the rate of events produced is ~ 10/minute.

Pile-up is simulated adding on average 5/25 events for low/high luminosity. They are added with poissonian statistic to the main event. In the ORCA case 3000 events were simulated and used for the local productions. In the CMSSW case Fast Simulation uses one million events by default, which are stored in the official release area, allowing to run grid productions as well as local productions.

Tables 3.2.2 and 3.2.2 show the total number of events produced in ORCA and CMSSW respectively. In the case of CMSSW more samples were produced, but were not used in the level 1 trigger study and are described in more detail in the next chapter.

ORCA					
Sample	Interval(GeV)	Number of events	pile-up luminosity	ORCA version	online/local
$t\bar{t}$		10496	low	8_7_3	local
ttH		4996	low	8_13_3	online
QCD	>70	5000	low & high	8_6_0	local
	>120	5000	low & high	8_6_0	local
	>170	5000	low & high	8_6_0	local
	30-50	1800	low	8_13_3	online
	50-80	3800	low	8_13_3	online
	80-120	16400	low	8_13_3	online
	120-170	9600	low	8_13_3	online
	170-230	7000	low	8_13_3	online
	230-300	1900	low	8_13_3	online
	300-380	2000	low	8_13_3	online
Wjetjet	50-85	16900	low	8_13_3	online
	85-150	12400	low	8_13_3	online

Table 3.1: **ORCA samples** - Number of events in each sample and corresponding luminosity.

CMSSW		
Sample	Interval(GeV)	Number of events
$t\bar{t}$H		1634000
$t\bar{t}$	+ 0 jets	59900
	+ 1 jet	66000
	+ 2 jets	98159
	+ 3 jets	14768
	+ 4 jets	5352
$t\bar{t}$	inclusive	1302000
qcd	30-50	300000
	50-80	364000
	80-120	386000
	120-170	407100
	170-230	536000
	230-300	447100
	300-380	530000
	380-incl	508000

Table 3.2: **CMSSW high luminosity samples** - Number of events in each sample. All the events have high luminosity pile-up.

CMSSW		
Sample	Interval(GeV)	Number of events
W + 0 jets		88000
W + 1 jet 0	pt < 100 100 < pt < 300	40000 100500
W + 2 jet	pt < 100 100 < pt < 300	99520 105300
W + 3 jet	pt < 100 100 < pt < 300	107900 86260
W + 4 jet	pt < 100 100 < pt < 300	83040 30800
W + 5 jet	pt < 100 100 < pt < 300	59020 41860
$Z \rightarrow \nu\nu$	120-170 GeV	29900
	170-230 GeV	25600
t$\rightarrow e\nu$ b		92000
t$\rightarrow \mu\nu$ b		94000
t$\rightarrow \tau\nu$ b		94000

Table 3.3: **CMSSW high luminosity samples** - Number of events in each sample. All the events have high luminosity pile-up.

3.2.3 Event multiplication

To study the selection capacity of offline cuts, which must extract few signal events from a big background, it is necessary to have a large number of QCD events. Unfortunately it was not possible to generate such a sample locally, and it was still difficult to retrieve online samples. However, since we are interested in calorimetric quantities, the detector resolution plays an important role. By parameterizing the resolution of offline jets with respect to jets reconstructed from the partons, we can simulate new events with different energies of the offline jets starting from the same Monte Carlo events: a single process leading to n partons in the final state is exploited several times. The missing Et will then be changed too, according to its own distribution.

The resolution functions depend on the Pt of the offline jet (non linearity of calorimeter response) and on the η (the calorimeter is not symmetrical in η).
For these reasons the offline jets are divided in bins of η and Pt and are associated to jets reconstructed from the partons.

The association [4] is performed requiring a minimum Pt of 10 GeV on both type of jets and a minimum ΔR of 0.3 for the association to Monte Carlo jets.
Figure 3.2.3 shows the resolution functions in the different bins of η and Pt obtained from QCD samples.

3.2.4 Event multiplication in ORCA

In ORCA the few events at our disposal forced us to generate also level 1 jet energies. The algorithm used performs the following actions:

- associates offline jets to jets made from partons;

[4] All the associations between jets or between tracks are made using our private algorithm. This is an exclusive associator, but does not associate always the two closest objects. It starts from the first element in the first collection and associates the closest element of the second collection to it. It then removes this element from the second collection and iterates.
The association requires that the (η,ϕ) distance between the two objects is less than a defined threshold.
It must be noted that if the element in the second collection were closest to one of the following elements in the first collection, we would have not taken the exact closest match. This is a simplified approach used to reduce the number of combinations to evaluate. We always provide the two collections sorted in decreasing Pt, so that the higher Pt elements are compared first and have higher priority on being associated.
This guarantees that the association is quite accurate at high energy, where resolution effects play a bigger role in determining the tail in the missing Et distribution.

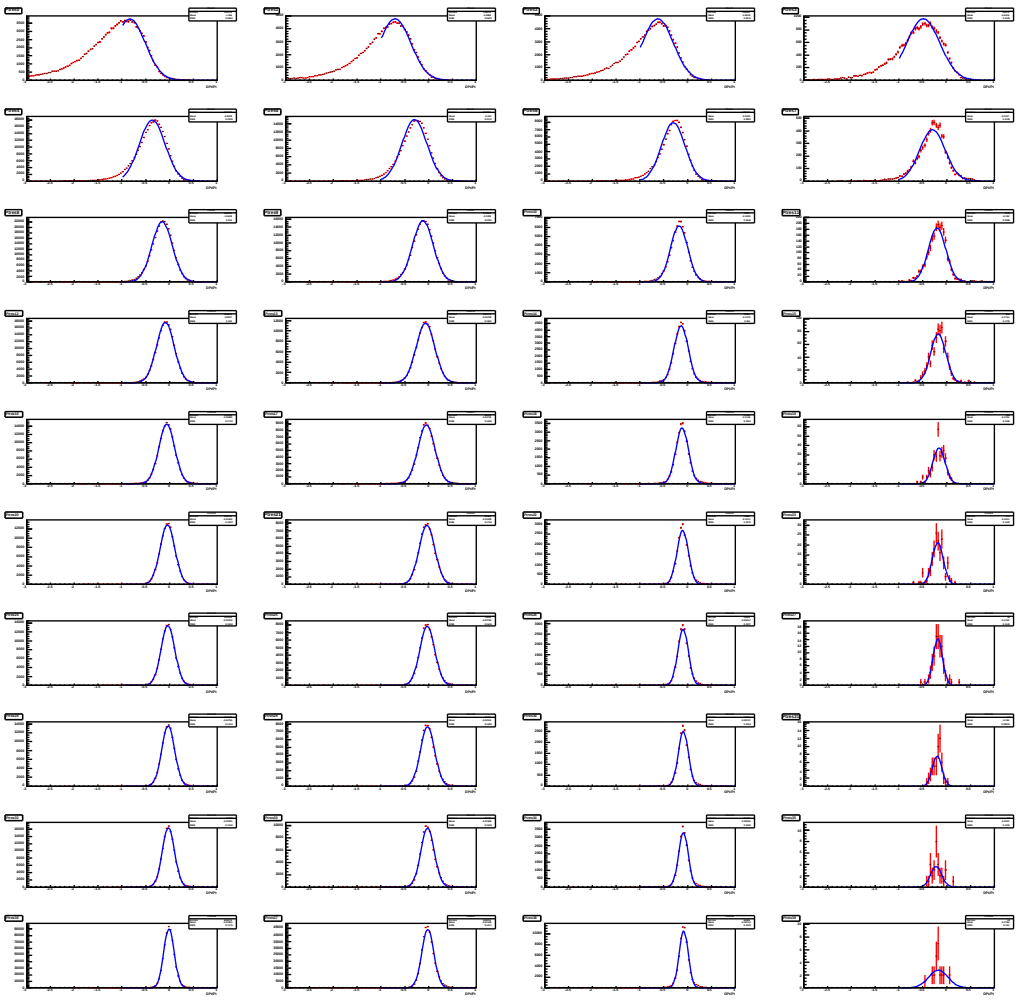


Figure 3.1: Resolution functions and gaussian fits for different bins in Et and η. Et bins are large 20 GeV and go from top (0−20 GeV), to bottom (>180 GeV). Bins in η are large 1.2 and go from left (0−1.2) to right (> 3.6).

- randomly generates a new transverse energy for the offline jet according to the corresponding distribution;
- level-1 calorimetric jets are associated to offline jets and are varied according to the new offline jet transverse energy with their own resolution functions.

The offline missing transverse energy must be changed too. Being built directly from the calorimetric towers means that a scaling of the energy fluctuations of offline jets is needed prior to missing transverse energy correction. Offline jets energy variations $\alpha \times \Delta\vec{E}$ are summed to $\not\!\!\vec{E}_T$ and the value of α is determined by a Kolmogorov test [46] between the distribution of original $\not\!\!E_T$ and the modified $\not\!\!E_T$. The best value is found to be $\alpha = 0.6$.

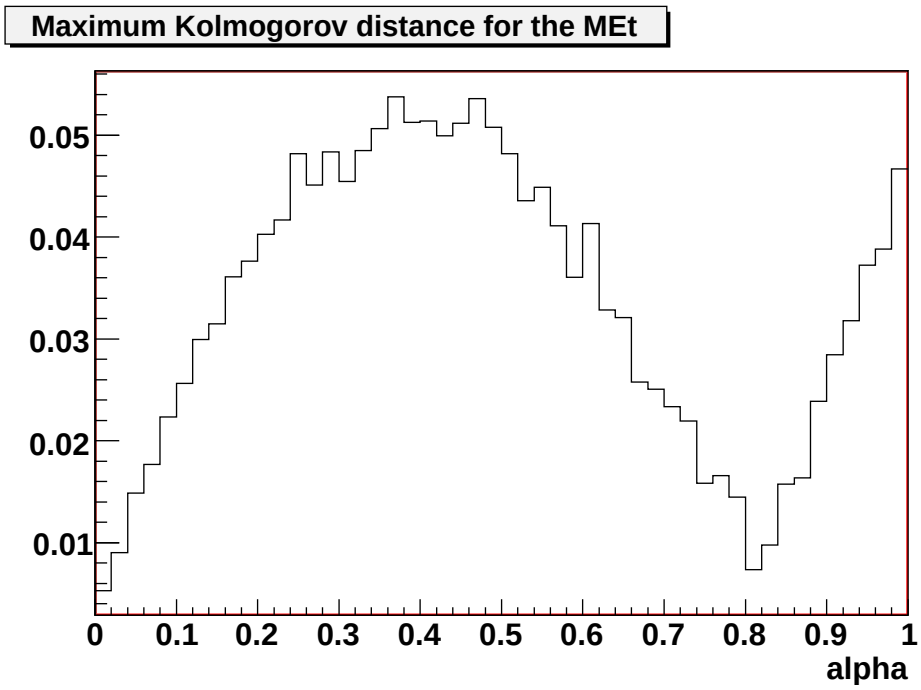


Figure 3.2: Maximum Kolmogorov distance for histogram containing the missing transverse energy Et_x and Et_y before and after the regeneration in CMSSW. The value of α is chosen as 0.8 in order to minimize the differences between the distributions.

After the procedure is complete the number of jets in the event above any given threshold, the missing transverse energy distribution and the distribution of the minimum distance in ϕ between the missing transverse energy and the offline jets are in good agreement, as can be seen from Figures 3.2.4 and 3.2.4.

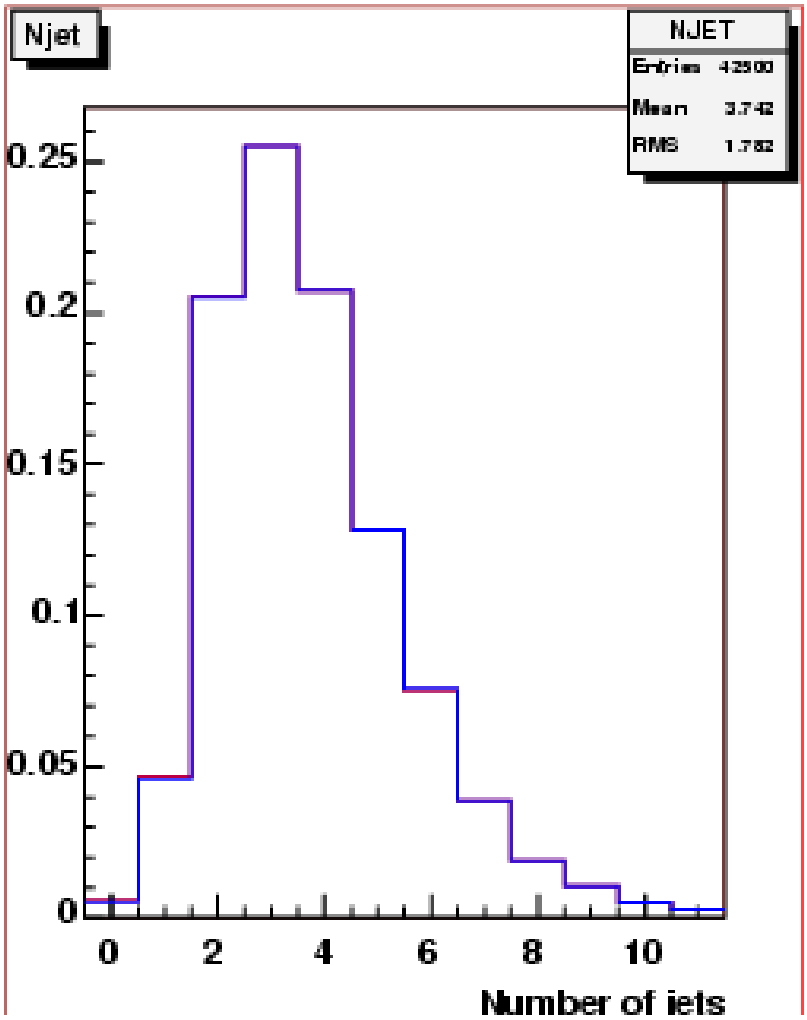
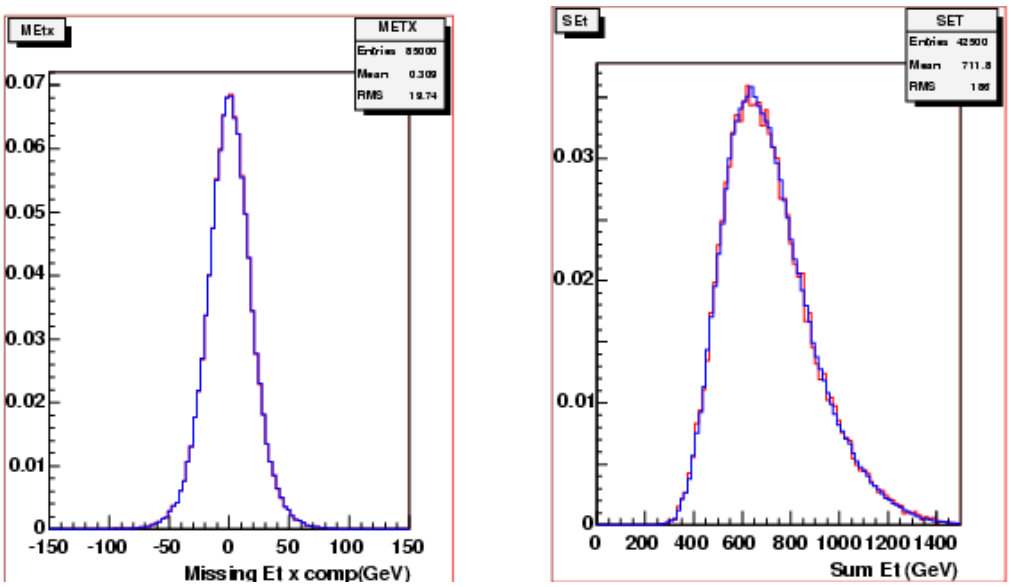


Figure 3.3: Offline jet multiplicity for original and modified events.

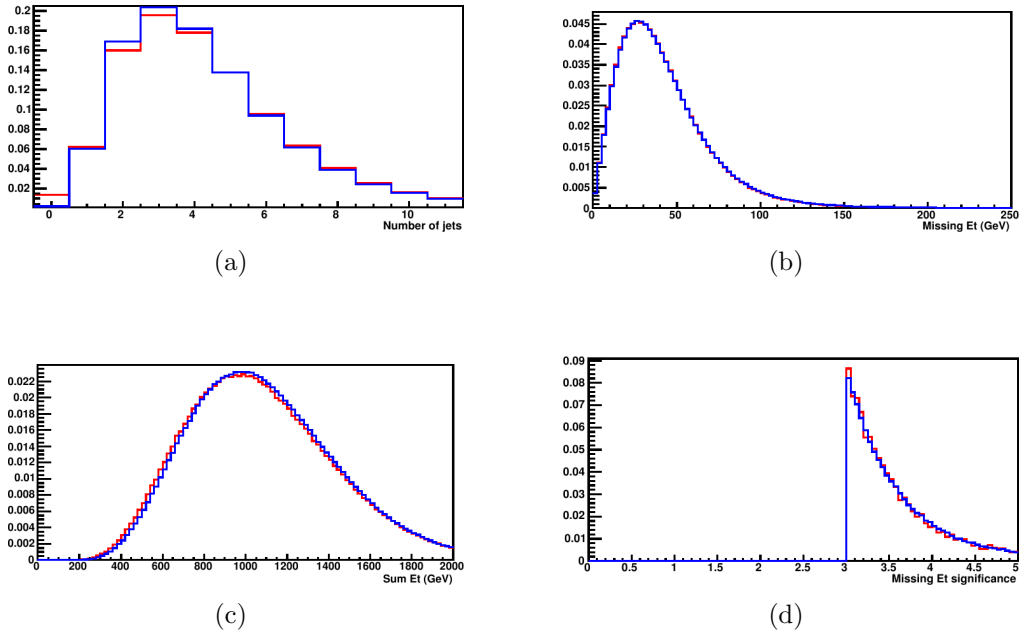


(a) (b)

Figure 3.4: (a) Missing transverse energy along the x direction for original and modified events; (b) Minimum $\Delta\phi$ between the missing transverse energy direction and the offline jets before and after modifications of jet energies.

3.2.5 Event multiplication in CMSSW

Moving to the new framework and using the FastSimulation of the event reconstruction, allowed to produce much bigger samples than in the ORCA case. For this reason, in CMSSW we did not modify the level 1 jet energies, we instead changed only the offline jet Et and correspondingly the missing transverse energy.
Again we find that a scale factor is needed for the missing Et, and with a Kolmogorov test its value is found to be 0.8, as shown in figure 3.2.4. Figure 3.2.5 shows the agreement between events before and after the regeneration.



(a) (b)

(c) (d)

Figure 3.5: Comparison of the values of modified variables before and after regeneration of the events. In red are the original events, in blue the regenerated. Number of jets (a), missing Et (b), sum Et (c). The plot in (d) shows the missing Et significance after the cuts on requiring at least 5 jets with $et > 25$ GeV and missing Et significance > 3.0. It shows that the tail of the generated distribution is still in good agreement with the original events.

3.3 Pixel Trigger

In this section we will describe a possible use of the pixel detector for triggering on multijet events. The study is made assuming to have the full pixel

QCD bin (GeV)	Generated	Generator efficiency (%)	Written events	Cut efficiency (%)
30-50	11229443	96.1631±0.0057	61292	0.5458±0.0022
50-80	26347818	99.2101±0.0017	194451	0.7380±0.0017
80-120	24971508	99.77571±0.00095	240496	0.9631±0.0020
120-170	29514603	99.94124±0.00045	452564	1.5334±0.0023
170-230	40608575	99.99196±0.00014	936749	2.3068±0.0024
230-300	32268604	99.997682±0.000085	1146776	3.5538±0.0033
300-380	37943909	100±0	1918910	5.0572±0.0036
380-incl	33232000	100±0	3295042	9.9153±0.0052

Table 3.4: Summary of the regenerated events. Here "generated" means events for which at least one of the offline jets was successfully associated to a monte carlo jet and its energy was changed. "generator efficiency" shows the percentage of events actually regenerated, it is close to 100% for all the samples, the only one with a slightly lower efficiency being the lowest Et bin. "written events" show the number regenerated events that pass the offline selection of having at least 5 offline jets with $et > 25$ GeV and $|\eta| < 3$. a missing Et significance of at least 3. Finally, "cut efficiency" represents percentage of events passing the cuts with respect to the generated events.

granularity available at the first level of trigger. Only the barrel part of the pixel detector will be considered, limiting the pseudorapidity coverage to $|\eta| < 1.5$, since we are interested in identifying events from hard interactions with high Pt and reducing the number of channels at level 1 is certainly an important requirement. In the last chapter 5, we will review the proposal made for a first level trigger with pixels for super LHC, which suggests to bring the high level trigger algorithms to the first level trigger with as small changes as possible and to use them for track reconstruction.

The level-1 trigger uses only informations from calorimeters and muon systems. Thus, final states with energetic muons or a few energetic jets are selected.
However, a number of interesting channels, involving also the Higgs production, are characterized by several low-energy hadronic jets and do not have a dedicated trigger. The pixel detector can help in the selection of such processes, since it is able to detect the charged component of low-energy jets that are seen as groups of charged tracks pointing at the same vertex.
The pixel detector can discriminate jets coming from primary and pile-up vertices, and matching these informations with the calorimeter contribution

it is possible to significantly increase trigger efficiency in multihadronic jets final states.
In this section the fundamental objects built from pixel hits are described. These objects will later be used to implement a pixel based trigger. We will only consider the barrel part of the pixel detector.
We assume to have access to the full pixel granularity at the first trigger level. This is done in order to study the possible improvements in the best possible scenario. Our objective is to determine how much it is possible to improve the level 1 trigger efficiency in the selection of signal events by the inclusion of the pixel detector and with the proposed trigger strategy.
We have also (together with Livio Fan and Michele Pioppi from "Universit di Perugia") developed a simulation in ORCA of the ROC response with the possibility to vary the dimension of the pixels to study the degradation in the resolution [47].
In this section we will, however, concentrate only on full granularity conditions.

3.3.1 Pixel tracks

In the final configuration the pixel detector provides three points in space along the particle path within the geometrical coverage. These points are used to reconstruct a track from which the transverse momentum and the longitudinal and transverse impact parameters are computed.
Tracks are reconstructed using the procedure implemented in the official CMS reconstruction [48]. The track is fitted as a circle in the transverse plane and a line in the $r-z$ plane. In the transverse plane the curvature radius R and the transverse momentum $\mathrm{p}_T = 3/1000 \times \mathrm{B} \times R$ (p_T in GeV, B in T and R in cm) are evaluated. The estimate is accurate only for p_T values below ~ 10 GeV, because of the small lever arm provided by the pixel detector. Figure 3.3.1 shows the track resolution with respect to p_T. This has been evaluated matching pixel-tracks to monte carlo tracks with a cone in (η,ϕ) of radius 0.05. It can be seen that up to $\mathrm{p}_T \sim 1$ GeV the resolution is better than 1%. It degrades to 22% when $\mathrm{p}_T \sim 10$ GeV.
The transverse impact parameter $\mathrm{IP}_{r\phi}$ is defined as the distance of closest approach to the beam axis. It can be determined with the center coordinates (x_c, y_c) and the radius R of the circle passing though the three hits:

$$\mathrm{IP}_{r\phi} = |\sqrt{x_c^2 + y_c^2} - R|. \tag{3.1}$$

For high Pt values, though, this reduces to a difference between big numbers which can lead to inaccuracies. Therefore a parabola approximation is used.

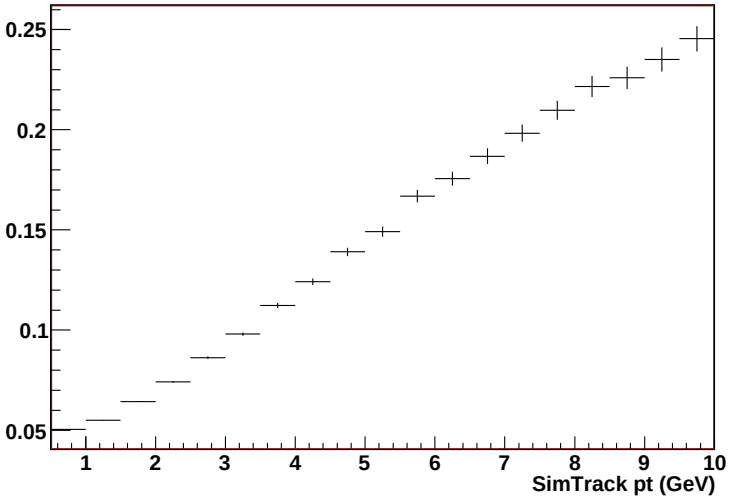


Figure 3.6: Linear behavior of $\sigma(pt)/pt$ as a function of Pt for pixel-tracks matched to simulated tracks.

The transverse impact parameter resolution for tracks with $pt > 6GeV$ is around 80 μm.
The accuracy of the longitudinal impact parameter z_{IP} is crucial in the determination of the interaction vertex. A full helix parameterization is used, allowing much improved resolution than a linear fit. The three pixel hits are projected onto the (ψ, z) plane, where ψ is the azimuthal angle difference between the hit and the point of closest approach around the circle defined by the three hits. The trajectory projected in this plane is expected to be a straight line, up to uncertainties due to multiple scattering in the detector material and hit position measurement errors. The longitudinal impact parameter is defined as the intercept of the line passing by the first two hits $(\psi_{1,2}, z_{1,2})$ and the z axis:

$$z_{\mathrm{IP}} = z_1 - \frac{\psi_1}{\psi_1 - \psi_2}(z_1 - z_2). \tag{3.2}$$

To quantify the quality of the tracks used, the χ^2 of the linear fit in the (r, z) plane is used.
Only tracks with $pt > 0.1GeV$, a transverse impact parameter < 1 and a $\chi^2 < 1000$ are selected for further processing.
The efficiency to reconstruct a pixel track with three pixel hits is $\epsilon_3 > 90\%$ for $|\eta| < 2.2$.

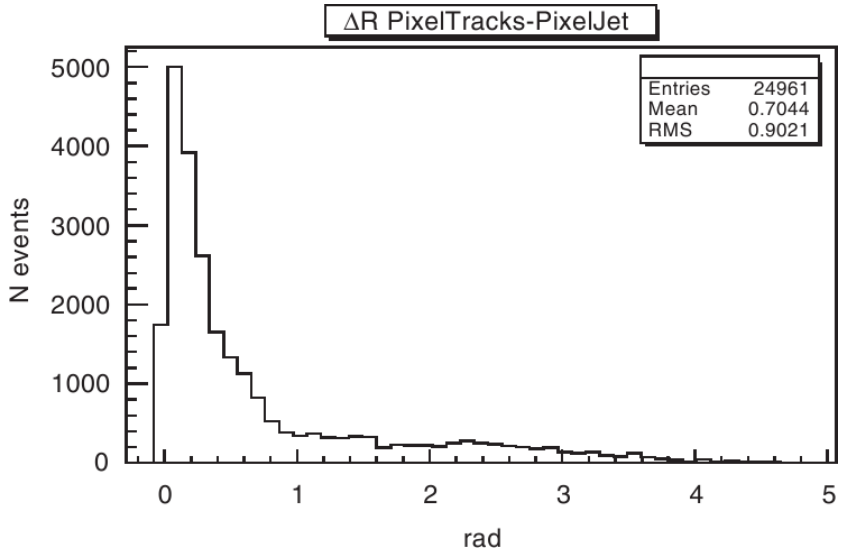


Figure 3.7: ΔR between the parton and the associated pixel-tracks.

3.4 Pixel-Jets

In order to help in the selection of events with high multiplicity of jets in the final state, pixel tracks are used to reconstruct jets (here called Pixel-Jets) which are the basic objects used by the proposed pixel trigger.
Pixel-Jets are defined as a group of pixel tracks inside an $\eta - \phi$ region.
The Pixel-Jet reconstruction algorithm has been chosen to be simple enough for a possible implementation at a future level-1 trigger, yet capable to bring a sufficient resolution on the Pixel-Jets vertices for accurate primary vertex identification.

In the following paragraphs the algorithm is described and the characteristics of the Pixel-Jets are shown in the case of full granularity.

3.4.1 Pixel-Jets reconstruction algorithm

The algorithm used for the reconstruction of Pixel-Jets is a simple iterative cone algorithm, in which the seeds are the highest Pt tracks.
To reconstruct a Pixel-Jet a cone of $\Delta R = \sqrt{\Delta\eta^2 + \Delta\phi^2} = 0.7$ is chosen, because more than 70% of the tracks associated with a parton are within the cone, as can be seen from figure 3.4.

Algorithm description

- the pixel tracks are sorted in decreasing Pt
- the first track in the list is taken and distance R in the η-ϕ plane with respect to the other tracks is evaluated
- each time R < 0.7 the second track is put in a Pixel-Jet defined by the first track and removed from the initial list.
- after all tracks have been compared to the first, the procedure is iterated to the next track in the new list.
- for all the Pixel-Jets:
  - The z coordinate of the vertex of the Pixel-Jet is defined as the pt-weighted mean of the longitudinal impact parameters of the tracks.
  - The Pt is defined as sum of the Pt of all the constituent tracks.
  - (η and ϕ are evaluated from sum of the pixel-tracks momentum vectors.)

The last step is not required for the trigger itself, but is needed to be able to associate Pixel-Jets to monte carlo jets and evaluate resolutions.

Without making any requirement on the number of pixel-tracks forming a Pixel-Jet, the mean number of reconstructed Pixel-Jets per event is shown in figure 3.4.1.

		$t\bar{t}$	ttH	qcd
Mean Pixel-Jets	Low Luminosity	10.3	11.8	6.8
	High Luminosity	16.9	17.5	16.3
Mean pixel-tracks per Pixel-Jet	Low Luminosity	5.1	5.7	4.0
	High Luminosity	7.0	7.8	6.0

Table 3.5: Mean number of Pixel-Jets and corresponding mean number of pixel tracks in a Pixel-Jet for low and high luminosity.

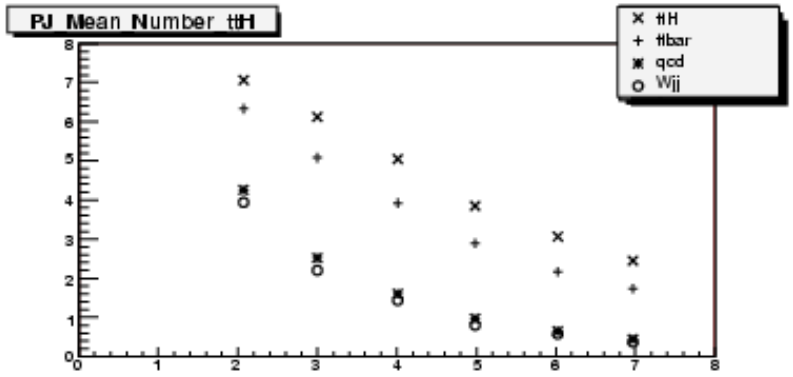


Figure 3.8: Mean number of reconstructed Pixel-Jets in $t\bar{t}$ and $t\bar{t}H$ inclusive events with respect to pure qcd background events and Wjj events in the low luminosity pile-up scenario. On the x axis the minimum number of tracks in each Pixel-Jet is reported.

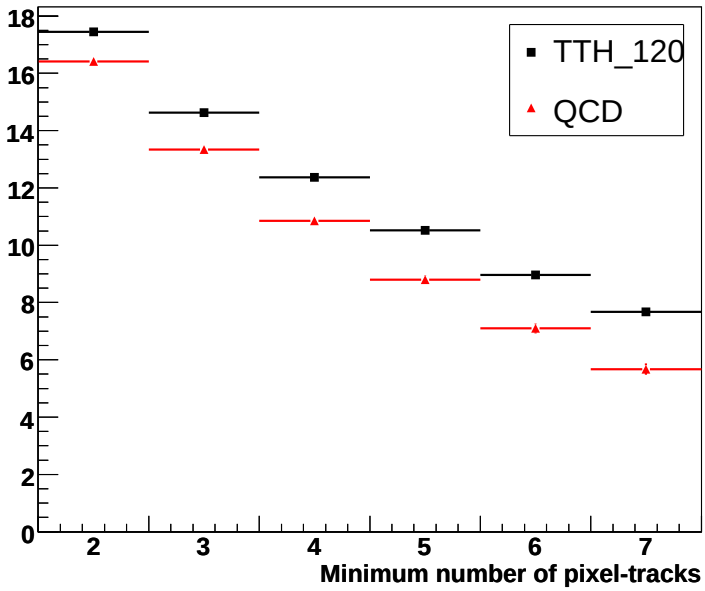


Figure 3.9: Mean number of reconstructed Pixel-Jets vs the minimum number of pixel-tracks required, in the high luminosity scenario. Only $t\bar{t}H$ and QCD are shown.

3.4.2 Pixel-Jets characterization

To evaluate Pt resolutions of the Pixel-Jets, they have been compared to jets reconstructed with the simulated partons.
In ORCA we did not have any suitable jet collection and we built our own "parton-jets". These parton-jets are reconstructed with a cone taking all the partons in the event coming directly from the primary partons. Partons with Pt > 1 GeV are taken as seeds to form clusters, a seed must have $\Delta R > 0.7$ from the next closest seed.
Clusters are then formed with all the partons within $R = 0.5$ from a seed. The final clusters are the parton-jets, with the Pt and the components of the momentum defined as the sum of the corresponding values of the constituent partons. The vertex z is evaluated as the weighted mean in Pt of the z of the constituent partons. The ϕ and η coordinates of the parton-jets are evaluated by the P_x and P_y components of the momentum of the parton-jet.
Many low transverse momentum parton-jets are reconstructed by this method, as shown in Figure 3.4.2. Since we are interested in high Pt jets, only parton-jets with Pt > 10 GeV are taken. The Pt resolution of Pixel-Jets is found to go from 20% to 40% as the Pt increases. This is expected since the Pt resolution of constituent pixel-tracks decreases with increasing Pt, although the effect is not linear since higher Pt Pixel-Jets have also a higher average number of pixel-tracks. Figure **??** shows the resolutions for three different cuts in minimum number of pixel-tracks for Pixel-Jets, together with level1-jets. In CMSSW we used the "GenJets" collection for the matching.

3.5 Primary vertex reconstruction with Pixel-Jets

The Pixel-Jets can be used to reconstruct interaction vertices in the event. The knowledge of the primary and pile-up vertices, as well as secondary vertices, can be used to reduce pile-up and better select signal samples. From this point on we will concentrate on the high luminosity scenario.
In order to have a fast reconstruction algorithm, the vertex reconstruction is reduced to a one-dimensional search along the z axis, where the spread due to the bunch-length is $\pm$ 15 cm. In the following we introduce the simple algorithm used for vertex reconstruction, its performance and an application to the level-1 trigger.

The PixelTrigger uses Pixel-Jets to identify events with a high number of high Pt jets coming from the same vertex. The algorithm used is a simple

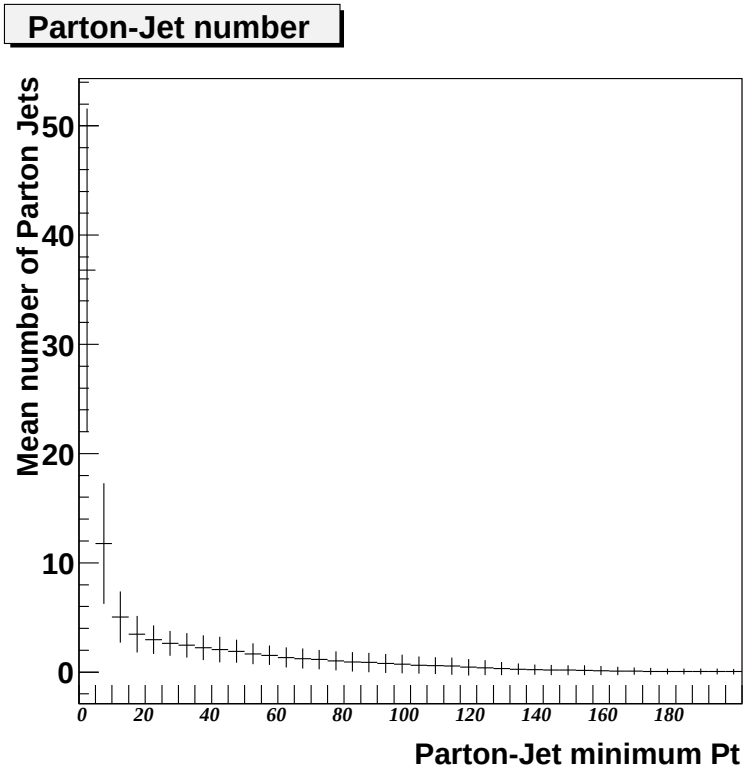


Figure 3.10: Mean number of parton-jets versus minimum Pt of the jets for 10496 events of $t\bar{t}$ inclusive with pile-up at low luminosity. The step is 5GeV.

divisive algorithm capable to determine the primary vertex quickly. It has only one free parameter, which will be determined in the next section.

Algorithm description

- The collection of Pixel-Jets is arranged using the z of their vertex, in increasing z.
- A divisive algorithm is used to construct vertices:
  - The search starts from the smallest z and moves to the next pixeljet.
  - If the two are closer than Δz they are associated and a new vertex is formed.
  - If their distance is greater than Δz, a new vertex is created and the second pixeljet is assigned to it.

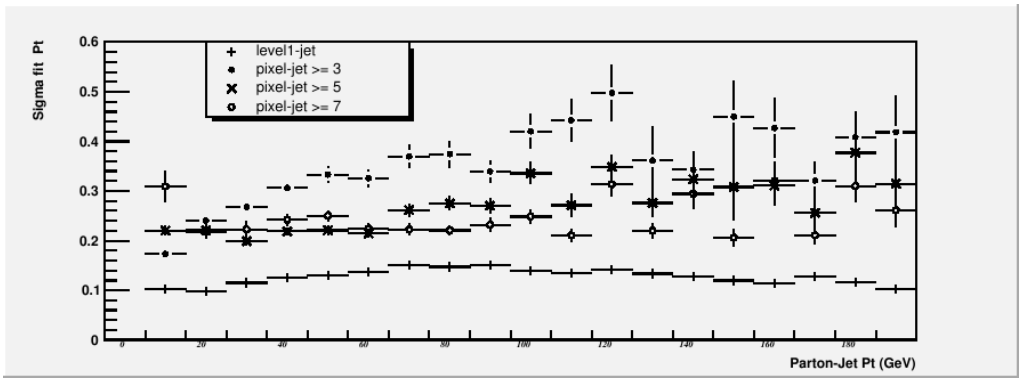


Figure 3.11: Pt resolutions of Pixel-Jets and level1-jets evaluated with parton-jets. Three different cuts on the number of pixel-tracks in a Pixel-Jets are considered. The association with parton-jets is made with a radius of 0.5. 68200 events coming from $t\bar{t}$, unmultiplied qcd, Wjetjet and $t\bar{t}$H were used.

  - The search continues iteratively.

- For each vertex the z coordinate is the pt-weighted sum of the z of the component Pixel-Jets.
- Each vertex has a Pt defined as the sum of the Pts of all component Pixel-Jets.
- The vertices are ordered in Pt and the on with the highest value is taken as the primary vertex.

The average number of reconstructed vertices in different samples is shown in figure 3.13. It can be noted that this number does not differ very much moving from a high jet multiplicity sample like $t\bar{t}H$ to samples with lower multiplicities. As it has been shown for the number of reconstructed Pixel-Jets, this can again be considered an effect of the pile-up.

The reconstructed primary vertex has a resolution of the order of the centimeter on the simulated primary vertex, as can be seen from figure 3.5. Figure 3.5 shows the average Pt of the primary vertex as a function of the Δz and figure 3.5 for some choices of values of Δz. The characterizing variable is the Pt, as the difference in the number of Pixel-Jets between different samples is not very distinctive. Furthermore, it quickly degrades when increasing the Δz due to the scarce z resolution and consequent inglobation of wrong Pixel-Jets.

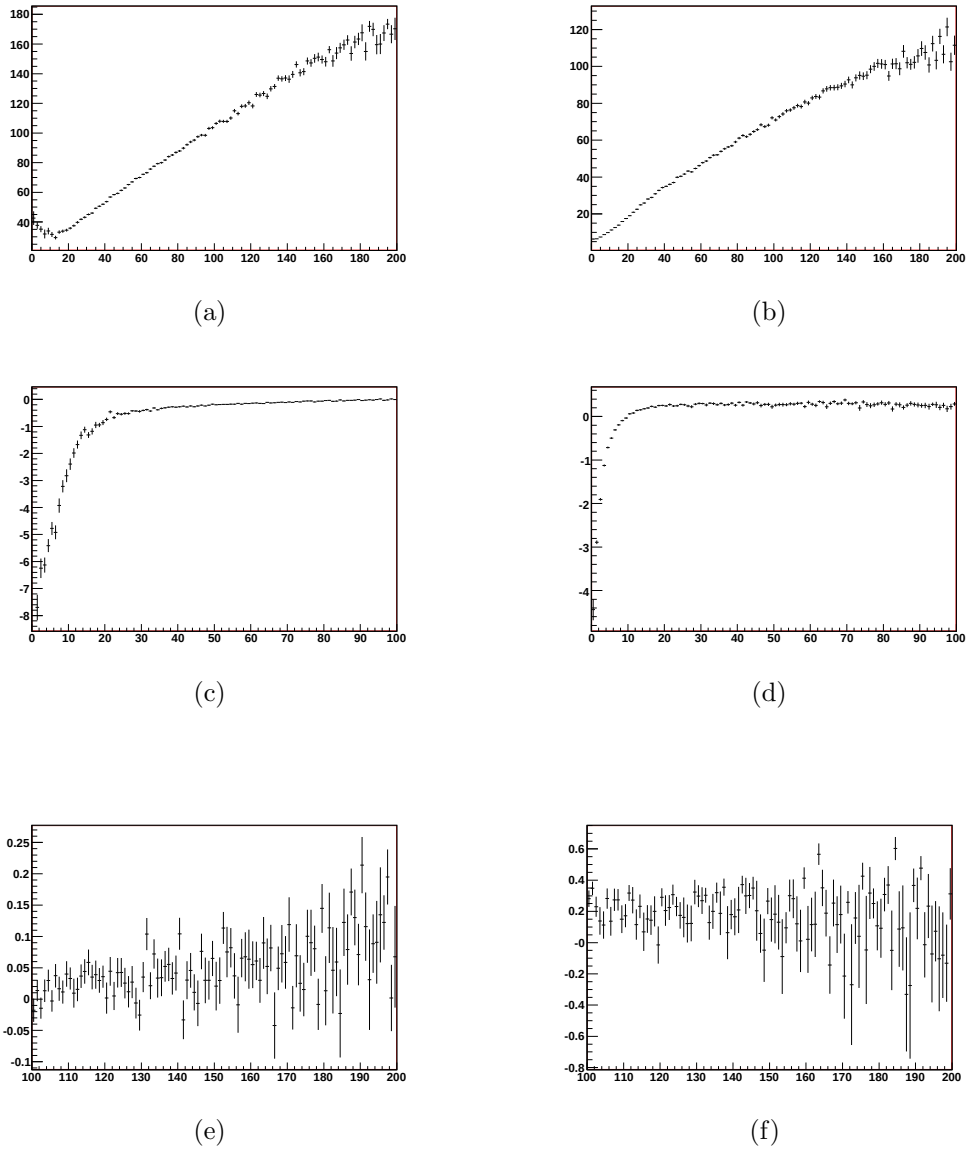


Figure 3.12: The first row of plots shows the Et of level 1 jets (a) and Pixel-Jets (b) versus GenJets Et. The remaining two rows show: (GenJet$_E t$ − level1Jet$_E t$)/GenJet$_E t$ on the left side and the same for Pixel-Jets on the right side. The histograms are divided in intervals of Pt: [0-100]GeV for (c) and (d) and (100-200]GeV for (e) and (f). The association with GenJets is performed with a cone in (η,ϕ) of radius 0.5. All are in the high luminosity case. As expected Pixel-Jets, that measure only the charged component of the jet, have a lower Et than GenJets.

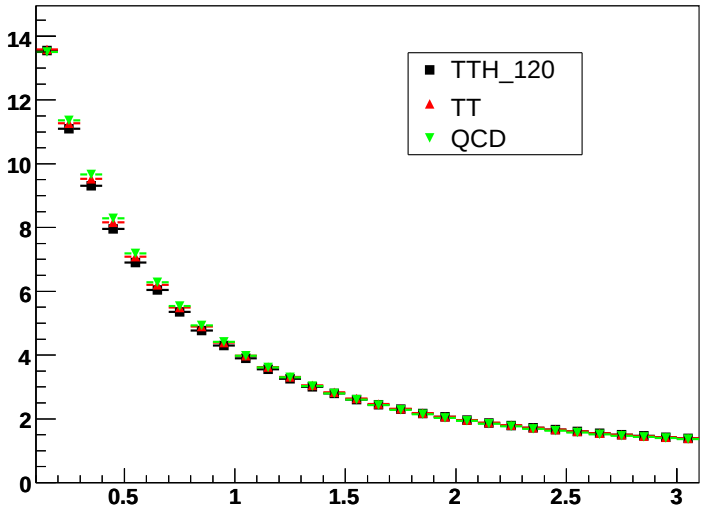


Figure 3.13: Mean number of vertices reconstructed from Pixel-Jets for different event samples in the high luminosity case.

3.6 Level 1 trigger with pixels

The reconstruction of the primary vertex using jets made of pixels can help in selecting high multiplicity events, that will have many jets coming from the interaction vertex.
The $t\bar{t}H$ channel, for low Higgs masses, is a very challenging channel and it has proven to be difficult even for the highest purity lepton based analyses. At high luminosities the standard trigger has only about 40% efficiency on this channel, coming mainly from the multijet trigger which accounts for 32% alone.
The pixel-based trigger response is required to be in logical AND with the standard multijet trigger. The cuts of both pixel and standard trigger will then be optimized to have the highest ttH efficiency, but maintaining an acceptable trigger rate.
This approach has the advantage that no intermediate steps in the trigger process are required. The pixel trigger can process the informations in parallel with the other triggers and send it to the Global Trigger for the final event selection.
The parameters characterizing the Pixel Trigger are the number of Pixel-Jets belonging to it (N) and the Δz used in the reconstruction. Another parameter we leave to determine in the following optimization is the mini-

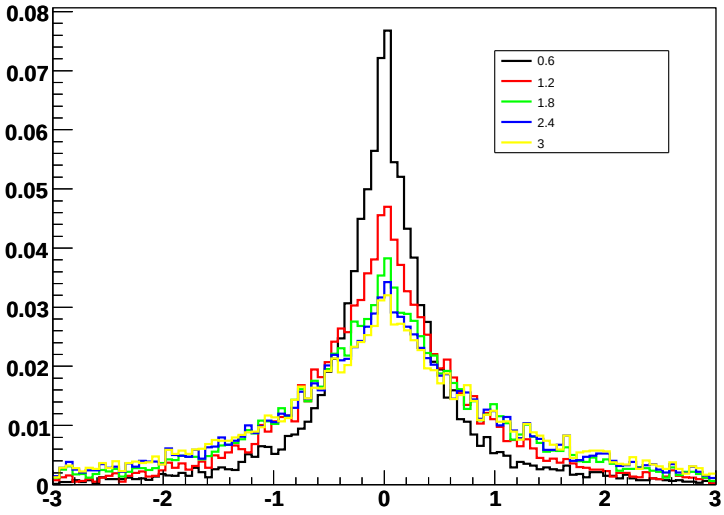


Figure 3.14: z difference between the primary vertex reconstructed with the Pixel-Jets and the simulated primary vertex.

mum number of pixel tracks required to form a Pixel-Jet.
The optimization of the trigger cuts is performed on Pixel-Jets reconstructed with different requirements on the minimum number of pixel tracks. The Δz used in the vertex reconstruction is also left as a free parameter. The interval used for the number of tracks is 2-13, because asking for more tracks would lower too much the starting number of Pixel-Jets and consequently the starting efficiency for the $t\bar{t}$H. The interval for Δz was chosen as wide as possible since the resolution in z is not very good.
The multijet Et cuts are varied around their default values.
Finally, the Pt is taken in an interval around 100 GeV, which is where $t\bar{t}H$ and QCD distributions show the biggest differences.
We first evaluated the rate for the standard mulitjet trigger. The high luminosity cut values reported in the Trigger TDR [37] have been used in the evaluation of the rate of the multijet trigger.
As we have seen in the previous chapter, the level 1 calorimeter trigger passes the four most energetic jets in the central region ($|\eta| < 3$), in the forward region ($|\eta| > 3$) and the four most energetic τ jets to the global trigger. The four jet energy thresholds are applied on the three incoming collections separately and the response of the trigger is the logical or of all the single Et cuts.
In table 3.6, the rates for the single et cuts are reported, as well as the rate

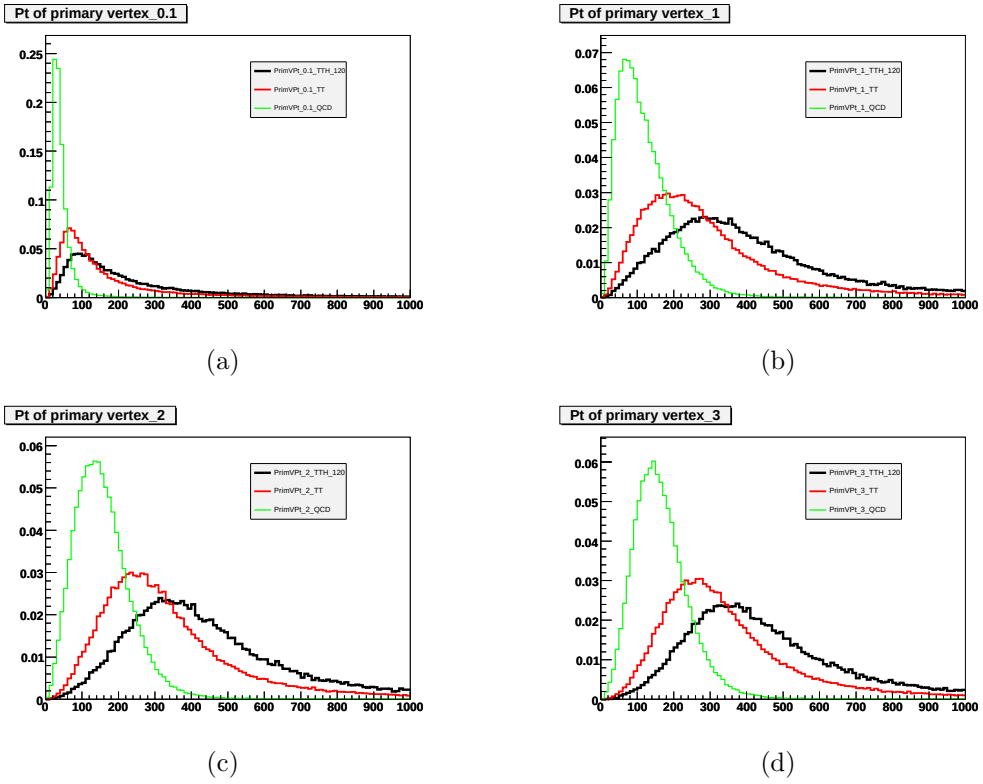
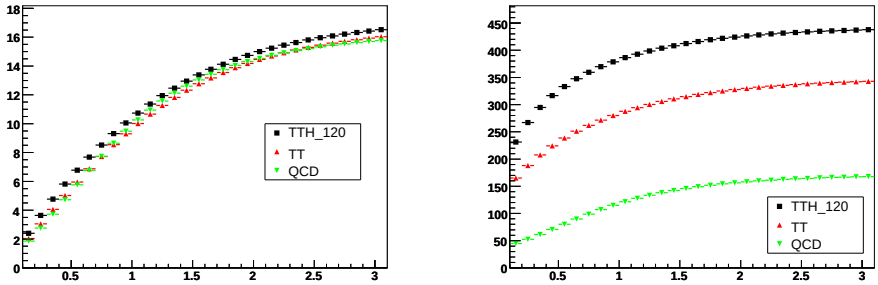


Figure 3.15: Primary vertex reconstructed with the Pixel-Jets. Pt in GeV is reported in the x axis. The histograms refer to different Δz values: (a) 0.1 cm, (b) 1.0 cm, (c) 2.0 cm and (d) 3.0 cm.

for the cumulative cuts. The Et cut values are also shown. We have also examined the missing Et + jet trigger, but we found a rate of 6.45 $\pm$ 0.23 kHz, which is much higher than the expected 0.1 kHz in the Trigger TDR. We noticed that QCD from lower Q^2 bins up to 80-120 GeV is responsible for almost all the rate. Given the big discrepancy with the TDR and the fact that the level 1 objects were still in a development phase, we decided not to use the missing Et + jet trigger and concentrate only on the multijet trigger.

After the optimization process, the following best values are found:

- minimum number of pixel tracks in a Pixel-Jet = 5
- Δz = 0.4 cm



(a) (b)

Figure 3.16: Mean number of Pixel-Jets in the primary vertex (a) and Pt of the primary vertex (b) as a function of Δz.

	1st	2nd	3rd	4th	Cumulative
Et cut	250 GeV	200 GeV	100 GeV	80 GeV	
rate	1240 ± 55 Hz	302 ± 19 Hz	310 ± 28 Hz	60.7 ± 3.9 Hz	1603 ± 34 Hz

Table 3.6: Rates of the single Et cuts for the standard multijet trigger with high luminosity thresholds.

- multijet trigger:
  - first jet minimum Et = 230 GeV
  - second jet minimum Et = 200 GeV
  - third jet minimum Et = 70 GeV
  - fourth jet minimum Et = 70 GeV
- minimum number of Pixel-Jets in the primary vertex = 1
- minimum Pt primary vertex = 170 GeV

The resulting rates and efficiencies for $t\bar{t}$H and $t\bar{t}$ are reported in table 3.6. Figure 3.6 shows the best value of the efficiency for some combinations of the cuts on minimum number of Pixel-Jets and Pt of the primary vertex once all the other parameters and cuts have been optimized. Figure 3.6 shows the Pt of the reconstructed primary vertex using the optimized values for the minimum track number and the Δz. It can be seen that the cut on the number of pixel jets in the vertex is not very effective.

	QCD Rate	$t\bar{t}$ eff %	$t\bar{t}H$ eff %
Standard multijet trigger	1603 ± 34 Hz	16.339 ± 0.090	32.140 ± 0.037
Pixel+multijet trigger	1536 ± 46 Hz	22.88 ± 0.11	46.878 ± 0.039

Table 3.7: Comparison between standard multijet trigger and pixel-multijet trigger on the QCD rate and efficiencies for $t\bar{t}H$ and $t\bar{t}$.

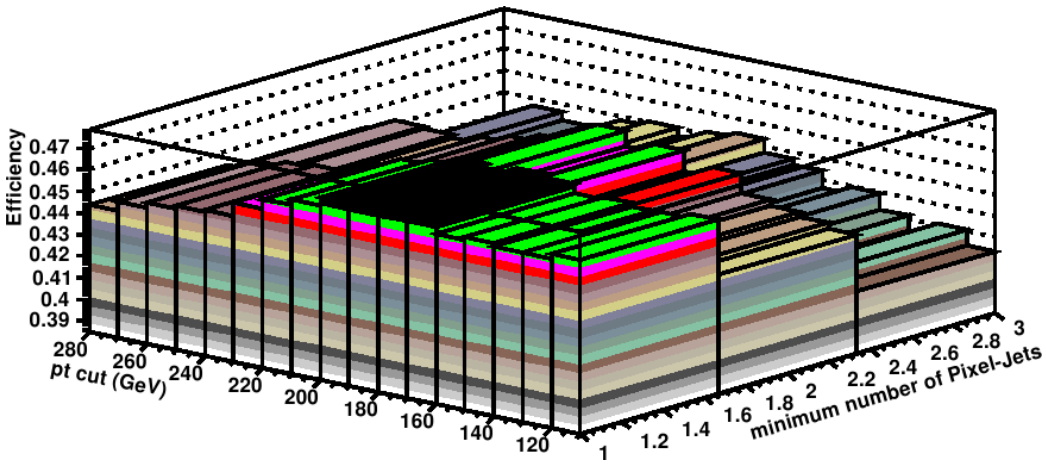


Figure 3.17: Best $t\bar{t}H$ efficiency for different combinations of cuts. In each bin is reported the best value once Δz, number of pixel-tracks in a Pixel-Jet and the multijet trigger cuts have been optimized with the constraint of not exceeding the standard level 1 rate.

The evaluation of the best values was performed using a subsample of the events because of the long computation time required. However, in table 3.6 the results evaluated on the full statistic are shown. More detailed efficiencies for the different decay channels are given in section **??**, together with a description of all the generated and reconstructed samples.

3.7 Conclusions

In this chapter we have discussed a possible application of the pixel detector at the first level of the CMS trigger. Assuming full granularity and using only the barrel part of the detector, pixel-tracks are reconstructed. Those tracks are then used to build Pixel-Jets that have been shown to have a linear Pt response up to about 200 GeV of the corresponding Monte Carlo jet.

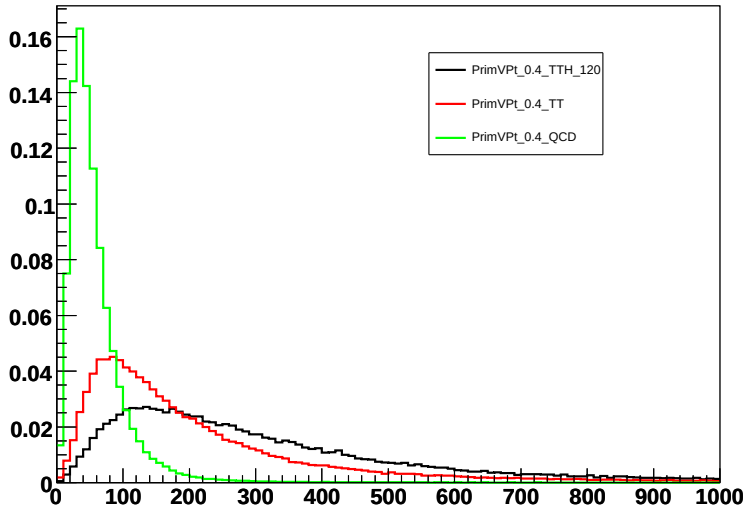


Figure 3.18: Primary vertex Pt for Pixel-Jets with five tracks and a cut on $\Delta z < 0.4$ cm.

Pixel-Jets are used to reconstruct vertices using only the z coordinate, with a divisive algorithm.
A trigger based on a cut on the Pt of the primary vertex was used together (in logical AND) with the multijet trigger and this combination was shown to be able to better select high jet multiplicity events than the multijet trigger alone. Optimized cuts for the pixel and multijet triggers were given, bringing an increase of $\sim$ 46% in the efficiency of $t\bar{t}$H and of $\sim$ 40% for the $t\bar{t}$ but maintaining the same QCD rate.
In the next chapter we will present a selection of the $t\bar{t}$H signal with a likelyhood based analysis. The results of this selection will be given both for the standard multijet trigger and for the pixel based trigger.

Chapter 4

Study of the observability of the $pp \to t\bar{t}\mathrm{H} \to$ jets + missing E_t process

4.1 Introduction

4.1.1 Goals of the study

In the former chapter we saw how the availability of information from pixel layers would allow a more precise identification of events with many hadronic jets, with a benefit in terms of reduction of energy thresholds needed at level 1 to keep the data acquisition rate at an acceptable level for triggers based on one, two, three, or four jets. In particular, the reduction of the threshold on the fourth jets' energy due to the rejection of events where jets do not all come from the same interaction vertex may allow the level 1 trigger to collect events of high interest with higher efficiency.

In this chapter we consider a search channel of a Standard Model Higgs boson which could benefit from a reduction of energy thresholds of jet triggers, with the double aim of showing what can be obtained with such search using a sufficient statistics collected by CMS in the running phase of high luminosity, and to show what benefit the channel could receive from a trigger using jets plus pixels, using values optimized in Ch. 3.

Our analysis does not seek to yield a definitive answer to the question, albeit interesting, of what could be the amount of experimental data necessary to observe a significant signal of the $t\bar{t}$H associated production process in the missing Et+jets final state with CMS. We believe that a necessarily more definitive and precise answer may be given only after a sufficient statistics of data will be collected in high luminosity running conditions, because

simulations of backgrounds –in particular, obviously, the one due to inclusive QCD processes– cannot be considered free from systematic uncertainties which may have a large impact in the tiny corner of phase space that a $t\bar{t}$H search is considering.

Because of the limited precision of the preliminary works illustrated in this chapter, due to the use of trigger thresholds not finalized and subjected to modifications (because of the uncertainty in their cross section and to their secondary importance in the global level 1 recipe in CMS), besides our use of a imperfect simulation of the detector (FastSim, which has however demonstrated a noteworthy precision in reproducing many of the characteristics simulated with the standard CMSSW simulation**??**), and to the practical impossibility to generate sufficient quantities of background events (in particular, inclusive QCD and top-antitop samples), we have limited the level of detail of our analysis. For example, we took in consideration only one mass value for the Higgs boson, the one of maximum interest for CMS in the context of this analysis; furthermore, we chose a simplified method for signal extraction from the mixture of data passing our selection criteria. Despite of these caveats, the main goal of the work here presented has been reached with success, as will be clear in the following.

4.1.2 Synopsis of the chapter

Associated production of Higgs boson and a pair of top quarks has been described in Ch.1. It is a rather rare process, which has at the 14 TeV c.m. energy provided by LHC a cross section of just 0.667 picobarns. Initially considered a promising channel to identify the Higgs boson for masses below 130 GeV [25, 49], it has been gradually derated from its status of discovery channel, and is now only cited without too much emphasis –and only in the context of $H \to \gamma\gamma$ decay signatures– in the summarizing documents on Higgs search prospects [50]. As a matter of fact, an accurate study of radiative processes producing a large multiplicity of additional jets in events containing heavy quarks, as described for instance by the AlpGen generator, has clarified that the identification of $t\bar{t}$H production is hard even with integrated luminosities in excess of a few tens of inverse femtobarns. Nowadays, the interest of the associated production process lays mostly with the chance of directly testing the Higgs-top quark coupling. Because of that, the original enthusiasm for the signature of eight hadronic jets, or six jets plus a lepton and missing energy, has been replaced in recent years by a more pragmatic interest for the signature of $t\bar{t}$H decays involving the $H \to \gamma\gamma$ decay, a spectacular but extremely rare process.

In this chapter we do not aim at modifying the picture described above,

but at adding a smal bit of information to it. In fact, among the possible experimental signatures of associated $t\bar{t}$H production with a decay into pairs of b-quark jets of the Higgs boson, there exist detailed studies of all identification channels of top quark pairs [51] except one. We are talking about the signature including hadronic jets and missing energy, one which at CDF has surprisingly demonstrated to yield a considerable sensitivity, providing at a certain point during Tevatron Run II the third best measurement of top pair production cross section [52].

In Sec. 4.2 we summarize briefly the search of top pair signatures in CDF including missing energy and hadronic jets, discussing in particular those aspects that prove important in a search of the same process in CMS, and a similar signature for $t\bar{t}$H production. In Sec. 4.3 we discuss the Monte Carlo samples we used for our study, and the many background processes with which we have to deal in the definition of an event selection strategy. In Sec. 4.4 we discuss the main characteristics of the searched events, and we lay the bases for the search by defining selection criteria on the triggered dataset. Two scenarios are considered: one when the analysis is carried out on events collected by the existing trigger menu for high luminosity running, and one when the trigger including a full-granularity pixel detector is included in the Level 1 decision. In Sec. 4.5 we discuss the identification of jets originated from b-quarks by the algorithm called "track counting", and we define the quantities which will be later used to select the signal, which nominally produces as many as four jets from b-quarks in the final state. In Sec. 4.6 we offer a study of the kinematical characteristics of selected events, identifying the variables which result to be most discriminant for $t\bar{t}$H events with respect to the dominant backgrounds, and in particular inclusive QCD production. In Sec. 4.6.4 we combine the most useful variables in a global discriminator, optimizing its construction. In Sec. 4.8 we discuss the extraction of the signal and its significance with a statistics equal to 100 inverse femtobarns, and the hypothetical gain in the case of a combined calorimeter-pixel detector Level 1 trigger. In Sec. 4.9 we provide some concluding remarks concerning this study.

4.2 The top-antitop production signal in events with missing energy and jets

As is well known, the searches for the top quark at the Tevatron have always considered a broad classification of the final state of top pair production in three categories: the hadronic channel, which identifies top-antitop pairs

where both W bosons emitted in top quark decay produce in turn pairs of quarks; the single lepton channel, which includes events where one of the two W bosons has produced a electron-neutrino or a muon-neutrino pair; and the dilepton channel, characterized by a decay to electron or muon of both W bosons.

The three main categories described above do not include the totality of final states into which top-antitop pairs may decay, not even if one assumes the branching fraction of top to Wb to be 100%. In fact, the three categories exclude from classification the states including tau leptons, which are considered a non-univocal signature, not easily classifiable: tau leptons have a rich phenomenology of decays, and their indentification is a very interesting but tough problem. Even more important is the observation that the three search channels, which together cover nominally 64/81ths of the total of final states, collect in reality a much smaller percentage of the actual decays, since the acceptance of the CDF and D0 detectors to electrons and muons of high momentum is limited to the central region of those apparata, to small values of rapidity; and even for hadronic jets a selection is normally applied which avoids considering those with rapidity above 2.0 (CDF) or 2.5 (D0). These limitations result in a sizable loss of top-antitop candidates, potentially identifiable with less tight criteria than the classical categorization into "dilepton", "single lepton", and "all hadronic" signature.

Of course, the active search for tau leptons –undertaken by CDF and D0 in Run II by accepting signatures consisting in single stiff tracks or narrow jets [53] – only partly covers the deficit, since even in that case the defects of hermeticity of the detectors remain, besides the low tau-identification efficiency.

An important step toward increasing the global efficiency of the search for top-antitop events consists instead in electing missing transverse energy to the main signature of the leptonic decay of a W boson: if a significant unbalancing in the transverse energy read in the calorimeters constitutes a sufficient signature for a W boson even in the absence of an additional charged lepton, the possibility of defining new independent experimental signatures arises.

Let us therefore turn the problem upside down, and let us consider a final state including significant missing transverse energy and four or five hadronic jets with high energy. Such an event may be the result of the decay of a top-antitop pair into electron, neutrino, and four jets, where the electron has missed the identification criteria in the detector, or has escaped detection because of high-rapidity emission. In the former case the electron is to all effects exchanged for a jet, in the second case it is absent. If the four quarks emitted in the final state are all identified as hadronic jets of high transverse

energy in the calorimeter system, one will obtain in the first case an event with five jets and missing Et, in the second case an event with four jets and missing Et; of course, initial and final state radiation, jets emitted at high rapidity, or jet merging can all modify the jet multiplicity of such an event, but the nominal signature of missing Et and four or five jets is certainly destined to collect a sizable number of events of the kind just discussed.

As far as tau-neutrino W decays are concerned, the same considerations above apply, since even tau leptons can manifest themselves as hadronic jets, with a branching ratio $\tau \rightarrow$hadronic of $\sim 65\%$. For the muon-neutrino W decay, instead, the final state will nominally contain only four hadronic jets and missing transverse energy. In all cases, one may reasonably expect a sizable acceptance of top-antitop events with single lepton decay, where the lepton has failed the identification or has escaped the detector unseen.

The search performed by CDF for these events [54] has started from a multi-jet trigger which required a minimum of four hadronic jets with a online-measured transverse energy above 15 GeV, besides a cut on total transverse energy measured in the calorimeters above 125 GeV. Such a trigger results sub-obtimal for top events where the leptonic signature of the W is delegated to missing transverse energy, since simulations do show that a sizable fraction of events with four jets reconstructed off-line would not pass the trigger criteria, because of the insufficient precision in the reconstruction and measurement of jets and $\sum$Et at trigger level 2.

The CDF trigger efficiency in CDF on inclusive top-antitop events results of about 60%, a number which is however sufficiently large to allow a reasonable measurement of top decays in the multijet-triggered data. A point to note is that the multijet trigger of CDF is not optimized for top events with a leptonic W decay: on the contrary, it was originally designed, and always tuned, to collect all-hadronic top pair decays: the two additional jets of that final state provide ample margin to the trigger, and in fact on all-hadronic top-antitop decays the trigger has globally an efficiency well above 90%. One may deduce that in CDF, associated $t\bar{t}$H production with a semileptonic decay of one of the two top quarks would still yield a sufficient number of jets to guarantee a high trigger efficiency.

In CMS, however, the enormous QCD cross section forces trigger level 1 to set very high Et thresholds for jets, and as we saw in Ch.3 the efficiency for the semileptonic $t\bar{t}$H production is less than satisfactory for high-luminosity running conditions.

The extraction of top-antitop signal with single lepton decay from multijet-triggered data required a first selection of events with four offline-reconstructed hadronic jets. The jet algorithm in CDF is of the interative cone type, with a default radius of 0.4 radians in the $\eta - \phi$ space. A very stringent cut on

the significance of missing transverse energy – the quantity obtained by dividing the missing Et measurement by its estimated uncertainty – $S > 4.0$. After that requirement, the dominant background is still made up of QCD events where one or more jets has been badly measured, yielding an imprecise measurement of the transverse energy flux in the direction of the jet or in the opposite direction in azimuth. In fact, the maximum part of QCD events surviving the cut on missing Et significance includes a jet lying close in azimuthal angle to the direction of missing energy. A cut on the minimum angle between missing Et and one of the selected jets, $DP_{min} > 0.4$, results extremely performant in reducing further the backgrounds.

After the selection of events with significant missing Et not pointing along the azimuth of a leading jet, remaining backgrounds are in sizable part constituted by electroweak processes, where a W boson has been produced in association with hadronic jets. Such a contamination is a clear indication that missing Et alone is a sensible way to select W decays. However, in CMS the same selection strategy would probably be less effective in reducing QCD backgrounds, both because of the worse resolution on missing energy of the LHC experiment, and because of the much larger cross section of multi-jet processes in the new environment, where the increase is larger for strong interactions than for electroweak processes.

Data selection in the CDF missing Et+multijet analysis includes the request that at least a jet contains a secondary vertex, a SecVtx b-tag[55]. Such a requirement is almost mandatory to characterize top pair decays, since they produce two b-jets while all background processes are usually devoid of them. Among events with one or more b-tags, CDF finds 688 events in 311 inverse picobarns of integrated luminosity.

The prediction in the number of selected background events containing one b-tag operated by CDF in the analysis described here is based on the parametrization of the probability that a jet contains a secondary vertex, as a function of the jet and event characteristics. The parametrization is performed on a sample independent from the one selected by the kinematical cuts, and its effectiveness is verified on control samples which are orthogonal to the one where the top signal is sought. The effectiveness of a parametrization of the probability for a generic jet to contain a b-tag is proven by years of tests within CDF, but before using a similar approach in LHC it will be necessary to study the applicability of the method with real data, given the uncertainties still present in our understanding of heavy flavor production in events with high jet multiplicity. In any case, one can say that depending on the purity of the b-tagging algorithm that LHC experiments will use, the theoretical uncertainty on heavy flavor production will be important or less so, since a high-efficiency, loose algorithm with a sizable rate of false positives

(fake b-tags) is less affected by such uncertainties, since b-tags in background processes are in the latter case dominated by resolution effects which an appropriate parametrization is capable of picturing with effectiveness.

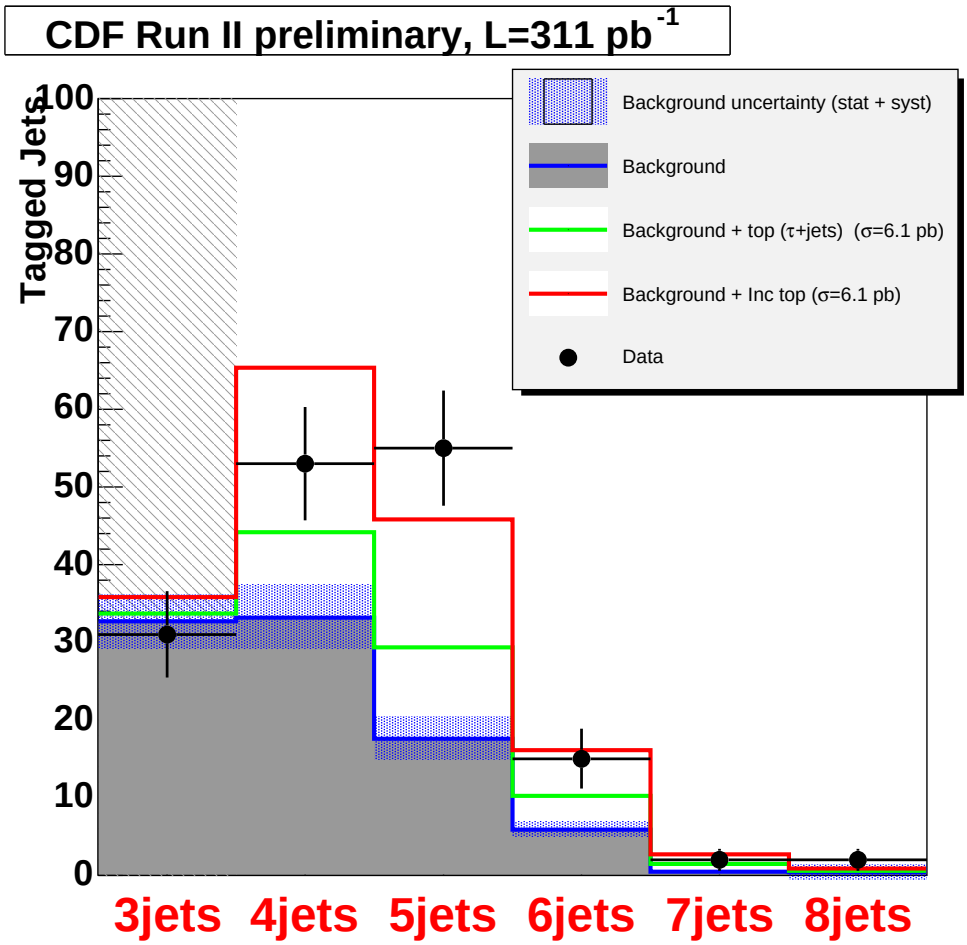
The result of the CDF search is summarized in the following table and the corresponding graph. One observes a significant excess of jets with b-tags in events with 4, 5, and 6 jets after the kinematic selection, with respect to tag matrix predictions. Such an excess is well described by the presence of top-antitop events with a single lepton, as shown in the graph (red histogram) 4.2. Given the efficiency of the selection on simulated signal events, the excess observed corresponds to a cross section of $\sigma_{t\bar{t}} = 6.7 \pm x \pm y$ picobarns, in very good agreement with theoretical predictions at the Tevatron for a top mass of 175 GeV. Such a measurement has recently become imprecise with respect to the most recent ones of CDF and D0, which have been obtained with samples of data six times larger in the standard decay channels; despite that, at the time of publication, the measurement has surprisingly resulted the third best determination of $\sigma_{t\bar{t}}$ by CDF: a clear indication of a success of the signal extraction method.

The above success teaches us two things. The first is that the neutrino –or rather, missing transverse energy- by itself constitutes a sufficient signature of an electroweak decay. The second is that analyses keeping an inclusive character in the search of rare processes has an indisputable advantage over exclusive searches that strive to achieve the maximum purity in the definition of the final state.

Given the above observations, we consider that the attempt at measuring the sensitivity of a search for $t\bar{t}$H events based on the signature of missing transverse energy and hadronic jets constitutes a valid contribution to the physics that CMS will investigate in the forthcoming years. Clearly, we do not expect to obtain upsetting results, capable of turning the tables on the possibility of extracting the associated production signal; we however consider important to complete the picture of possible searches, with an eye also to the simple definition of a sample enriched in top-antitop events collected in a way which is effectively independent from the more classical signatures with charged leptons or hadronic jets alone.

4.3 Monte Carlo samples and background processes

In this chapter we describe the experimental samples used in our study on the possibility to extract the $t\bar{t}$H associated production signal (with Higgs



Number of positive SecVtx tagged jets as a function of the jet multiplicity in the sample after the optimized kinematical selection. Data are represented by dots, the solid grey histogram shows the tagging matrix background prediction while the red and green lines represent the signal + background expectation for the inclusive and tau+jets ttbar decays respectively normalized to the theoretical cross section of 6.1 pb.

boson mass at 120 GeV) from a sample with significant missing *et* and many hadronic jets.

4.3.1 $t\bar{t}$H signal

The associated production of top-antitop pairs and the Higgs boson has at CMS a cross section of 0.667 pb for the Higgs mass value of 120 GeV, which we consider as the reference value for our study. The process, simulated with pythia version 6.409 [56] is not filtered in generation requiring any particular decay mode, respecting the inclusive approach of our research.
We have generated 1634000 events and filtered them with a standard trigger simulation, or alternatively with the multijet+pixel trigger discussed in chapter 3. The sample composition in term of the different decay channel contributiong after the trigger selection is shown in tables **??**. We can see, as expected, that the multijet trigger selects with maximum efficiency the hadronic decay channel of $t\bar{t}$ pairs, but the semileptonic decays also give a significant contribution. Looking at the Higgs boson decays, we notice that the decays in b quark coules help the selection of events with single lepton decay of the $t\bar{t}$ couple, giving two additional jets the trigger can select.

In the case of the pixel based trigger, tables **??** shows that the efficiency increases mainly in channels with the semileptonic decay of the $t\bar{t}$ pairs, which benefit most from the reduction of the threshoulds of the level 1 jets *et*. In total, the channels containing at least one energetic neutrino are selected witha an efficiency of (27.780 $\pm$ 0.047)% by the multijet trigger and (40.817 $\pm$ 0.052)% by the multijet+pixel trigger. These will be the majority of the events selected from the requirement of significant missing *et*, as will be discussed in the next sections.

	Standard multijet trigger			
	$H \to bb$		$H \to cc$	
	N_{sel}/N_{tot}	%	N_{sel}/N_{tot}	%
$t\bar{t} \to jjbjjb$	188297/495783	37.980 ± 0.069	10788/27399	39.37 ± 0.30
$t\bar{t} \to e\nu bjjb$	54618/159460	34.25 ± 0.12	3042/8543	35.61 ± 0.52
$t\bar{t} \to \mu\nu bjjb$	41550/158849	26.16 ± 0.11	2391/8928	26.78 ± 0.47
$t\bar{t} \to \tau\nu bjjb$	46344/159452	29.07 ± 0.11	2700/8632	31.28 ± 0.50
$t\bar{t} \to e\nu be\nu b$	4126/12685	32.53 ± 0.42	246/704	34.94 ± 1.80
$t\bar{t} \to e\nu b\mu\nu b$	5620/25191	22.31 ± 0.26	342/1397	24.48 ± 1.15
$t\bar{t} \to e\nu b\tau\nu b$	6523/25680	25.40 ± 0.27	335/1388	24.14 ± 1.15
$t\bar{t} \to \mu\nu b\mu\nu b$	1935/12816	15.10 ± 0.32	126/741	17.00 ± 1.38
$t\bar{t} \to \mu\nu b\tau\nu b$	4428/25443	17.40 ± 0.24	288/1396	20.63 ± 1.08
$t\bar{t} \to \tau\nu b\tau\nu b$	2553/12734	20.05 ± 0.35	148/722	20.50 ± 1.50
Total	355994/1088093	32.717 ± 0.045	20406/59850	34.09 ± 0.19

Table 4.1: Number of events selected by the standard multijet trigger, and relative percentages, for different principal decay channels of the Higgs boson and for the principal final states of the $t\bar{t}$ pairs.

	Pixel Trigger			
	$H \to bb$		$H \to cc$	
	N_{sel}/N_{tot}	%	N_{sel}/N_{tot}	%
$t\bar{t} \to jjbjjb$	273850/495358	55.283 ± 0.071	15315/27376	55.94 ± 0.30
$t\bar{t} \to e\nu bjjb$	79586/159326	49.95 ± 0.13	4368/8538	51.16 ± 0.54
$t\bar{t} \to \mu\nu bjjb$	63026/158704	39.71 ± 0.12	3628/8919	40.68 ± 0.52
$t\bar{t} \to \tau\nu bjjb$	67335/159316	42.27 ± 0.12	3785/8620	43.91 ± 0.53
$t\bar{t} \to e\nu be\nu b$	5795/12673	45.73 ± 0.44	337/703	47.94 ± 1.88
$t\bar{t} \to e\nu b\mu\nu b$	8704/25171	34.58 ± 0.30	534/1397	38.23 ± 1.30
$t\bar{t} \to e\nu b\tau\nu b$	9495/25656	37.01 ± 0.30	510/1385	36.82 ± 1.30
$t\bar{t} \to \mu\nu b\mu\nu b$	2943/12808	22.98 ± 0.37	196/740	26.49 ± 1.62
$t\bar{t} \to \mu\nu b\tau\nu b$	6809/25419	26.79 ± 0.28	404/1392	29.02 ± 1.22
$t\bar{t} \to \tau\nu b\tau\nu b$	3717/12724	29.21 ± 0.40	210/722	29.09 ± 1.69
Total	521260/1087155	47.947 ± 0.048	29287/59792	48.98 ± 0.20

Table 4.2: Number of events selected by the multijet+pixel trigger, and relative percentages, for different principal decay channels of the Higgs boson and for the principal final states of the $t\bar{t}$ pairs.

	Standard multijet trigger			
	$H \to \tau\tau$		$H \to WW$	
	N_{sel}/N_{tot}	%	N_{sel}/N_{tot}	%
$t\bar{t} \to jjbjjb$	20806/59771	34.81 ± 0.19	36948/105322	35.08 ± 0.15
$t\bar{t} \to e\nu bjjb$	6389/19358	33.00 ± 0.34	10733/33714	31.84 ± 0.25
$t\bar{t} \to \mu\nu bjjb$	4471/19402	23.04 ± 0.30	7876/33567	23.46 ± 0.23
$t\bar{t} \to \tau\nu bjjb$	4971/19423	25.59 ± 0.31	8658/33602	25.77 ± 0.24
$t\bar{t} \to e\nu be\nu b$	525/1586	33.10 ± 1.18	798/2670	29.89 ± 0.89
$t\bar{t} \to e\nu b\mu\nu b$	649/3101	20.93 ± 0.73	1044/5466	19.10 ± 0.53
$t\bar{t} \to e\nu b\tau\nu b$	815/3222	25.30 ± 0.77	1288/5323	24.20 ± 0.59
$t\bar{t} \to \mu\nu b\mu\nu b$	190/1537	12.36 ± 0.84	368/2622	14.04 ± 0.68
$t\bar{t} \to \mu\nu b\tau\nu b$	464/3062	15.15 ± 0.65	813/5348	15.20 ± 0.49
$t\bar{t} \to \tau\nu b\tau\nu b$	270/1569	17.21 ± 0.95	453/2719	16.66 ± 0.71
Total	39550/132031	29.95 ± 0.13	68979/230353	29.945 ± 0.095

Table 4.3: Number of events selected by the standard multijet trigger, and relative percentages, for different principal decay channels of the Higgs boson and for the principal final states of the $t\bar{t}$ pairs.

	Pixel Trigger			
	$H \to \tau\tau$		$H \to WW$	
	N_{sel}/N_{tot}	%	N_{sel}/N_{tot}	%
$t\bar{t} \to jjbjjb$	29589/59726	49.54 ± 0.20	53913/105229	51.23 ± 0.15
$t\bar{t} \to e\nu bjjb$	8863/19346	45.81 ± 0.36	15714/33680	46.66 ± 0.27
$t\bar{t} \to \mu\nu bjjb$	6692/19386	34.52 ± 0.34	12067/33536	35.98 ± 0.26
$t\bar{t} \to \tau\nu bjjb$	7081/19411	36.48 ± 0.35	12803/33577	38.13 ± 0.27
$t\bar{t} \to e\nu be\nu b$	663/1585	41.83 ± 1.24	1137/2668	42.62 ± 0.96
$t\bar{t} \to e\nu b\mu\nu b$	928/3101	29.93 ± 0.82	1616/5464	29.57 ± 0.62
$t\bar{t} \to e\nu b\tau\nu b$	1103/3221	34.24 ± 0.84	1795/5317	33.76 ± 0.65
$t\bar{t} \to \mu\nu b\mu\nu b$	300/1535	19.54 ± 1.01	549/2621	20.95 ± 0.79
$t\bar{t} \to \mu\nu b\tau\nu b$	705/3062	23.02 ± 0.76	1237/5342	23.16 ± 0.58
$t\bar{t} \to \tau\nu b\tau\nu b$	401/1568	25.57 ± 1.10	695/2717	25.58 ± 0.84
Total	56325/131941	42.69 ± 0.14	101526/230151	44.11 ± 0.10

Table 4.4: Number of events selected by the standard multijet trigger, and relative percentages, for different principal decay channels of the Higgs boson and for the principal final states of the $t\bar{t}$ pairs.

	Standard multijet trigger			
	$H \rightarrow$ others		Total	
	N_{sel}/N_{tot}	%	N_{sel}/N_{tot}	%
$t\bar{t} \rightarrow jjbjjb$	21256/56347	37.72 ± 0.20	278095/744622	37.347 ± 0.056
$t\bar{t} \rightarrow e\nu bjjb$	6227/18133	34.34 ± 0.35	81009/239208	33.866 ± 0.097
$t\bar{t} \rightarrow \mu\nu bjjb$	4702/18260	25.75 ± 0.32	60990/239006	25.518 ± 0.089
$t\bar{t} \rightarrow \tau\nu bjjb$	5137/17922	28.66 ± 0.34	67810/239031	28.369 ± 0.092
$t\bar{t} \rightarrow e\nu be\nu b$	463/1449	31.95 ± 1.23	6158/19094	32.25 ± 0.34
$t\bar{t} \rightarrow e\nu b\mu\nu b$	645/2891	22.31 ± 0.77	8300/38046	21.82 ± 0.21
$t\bar{t} \rightarrow e\nu b\tau\nu b$	761/2935	25.93 ± 0.81	9722/38548	25.22 ± 0.22
$t\bar{t} \rightarrow \mu\nu b\mu\nu b$	234/1426	16.41 ± 0.98	2853/19142	14.90 ± 0.26
$t\bar{t} \rightarrow \mu\nu b\tau\nu b$	525/2897	18.12 ± 0.72	6518/38146	17.09 ± 0.19
$t\bar{t} \rightarrow \tau\nu b\tau\nu b$	288/1413	20.38 ± 1.07	3712/19157	19.38 ± 0.29
Total	40238/123673	32.54 ± 0.13	525167/1634000	32.140 ± 0.037

Table 4.5: Number of events selected by the standard multijet trigger, and relative percentages, for different principal decay channels of the Higgs boson and for the principal final states of the $t\bar{t}$ pairs.

	Pixel Trigger			
	$H \rightarrow$ others		Total	
	N_{sel}/N_{tot}	%	N_{sel}/N_{tot}	%
$t\bar{t} \rightarrow jjbjjb$	30188/56783	53.16 ± 0.21	402855/744472	54.113 ± 0.058
$t\bar{t} \rightarrow e\nu bjjb$	8748/18121	48.27 ± 0.37	117279/239011	49.07 ± 0.10
$t\bar{t} \rightarrow \mu\nu bjjb$	6907/18246	37.85 ± 0.36	92320/238791	38.66 ± 0.10
$t\bar{t} \rightarrow \tau\nu bjjb$	7289/17910	40.70 ± 0.37	98293/238834	41.16 ± 0.10
$t\bar{t} \rightarrow e\nu be\nu b$	628/1446	43.43 ± 1.30	8560/19075	44.88 ± 0.36
$t\bar{t} \rightarrow e\nu b\mu\nu b$	956/2887	33.11 ± 0.88	12738/38020	33.50 ± 0.24
$t\bar{t} \rightarrow e\nu b\tau\nu b$	1020/2932	34.79 ± 0.88	13923/38511	36.15 ± 0.24
$t\bar{t} \rightarrow \mu\nu b\mu\nu b$	317/1426	22.23 ± 1.10	4305/19130	22.50 ± 0.30
$t\bar{t} \rightarrow \mu\nu b\tau\nu b$	737/2895	25.46 ± 0.81	9892/38110	25.96 ± 0.22
$t\bar{t} \rightarrow \tau\nu b\tau\nu b$	379/1411	26.86 ± 1.18	5402/19142	28.22 ± 0.33
Total	57169/124057	46.08 ± 0.14	765567/1633096	46.878 ± 0.039

Table 4.6: Number of events selected by the standard multijet trigger, and relative percentages, for different principal decay channels of the Higgs boson and for the principal final states of the $t\bar{t}$ pairs.

4.3.2 top-antitop inclusive events production

In the setting of the search for the Higgs boson, the process of inclusive production of $t\bar{t}$ pairs is a problematic background - and infact the most difficult to reduce, being very similar to the signal -, the selection of a significant quantity of top decays in a sample in much part orthogonal to all the others usually considered, can be considered an added value to the study.
As we have seen in section 4.2, CDF has demonstrated that the missing Et plus jets sample contains a significant quantity of semileptonic decays of the top-anitop pairs, with which measurements of interesting potentials are possible. We will not discuss in this chapter the characteristics of this sample as a possible signal to extract, we will instead search for variables usefull to remove it efficiently from our selection.
The top-antitop production we used is generated with AlpGen version 2.12 and Pythia version 6.409, and consists in five different samples, each one obtained selecting a fixed number of energetic partons, above the two heavy quarks, in the final state: 0,1,2,3, and at least 4 independent partons.
The cross section of the five processes are respectlively: 619, 176, 34, 6, 1.5 pb.
AlpGen is used for the generation of the events and Pythia for the showering part. In order to have truly independent samples for the different number of jets bins, the MLM [57] mechanism is used: The problem is in the use of two different mechanisms in the event generation. Alpgen uses matrix elements to generate the events, so the processes can be completely selected at this point. The showering part, however, is a fenomenological description and it can thus produce new jets independently. Those new jets must be kept into consideration to avoid double counting in the evaluation of the cross section and consequent rescaling of the different jet number bins.
The actual MLM mechanism works in the following way:

- produce the events with the minimum Pt cut on partons from matrix elements
- do the showering (with Pythia in this case)
- reconstruct the parton jets with a simple cone algorithm and associated them to the partons
- all the partons must be associated, if there are extra jets, discard the event

In the sample used the settings for the cone algorithm were $\Delta R < 0.7$ and $Pt > 70$ GeV. Since the offline algorithm for jet reconstruction we are using

is different (iterative cone with $\Delta R < 0.5$), the same parton multiplicity will give rise to a spectrum of jet multiplicity. For this reason, even if interested only in certain jet numbers all the samples must be used. The MLM mechanism guarantees that the sum of all the different $t\bar{t} + Njets$ bins gives the full inclusive $t\bar{t}$ sample without double counting.
Earlyer samples were produced with a 20 GeV cut, which was considered unreliable by QCD experts for high jet multiplicities.
Tables **??** shows the relative contribution of the various decays after the multijet trigger selection is performed.
Tables **??** shows the same informations for the events that would be selected by the multijets-pixel trigger.
Similarly to the $t\bar{t}$H case, the efficiency on the selection of events with at least one neutrino is increased from $(13.74 \pm 0.11)\%$ to $(16.44 \pm 0.13)\%$.

	Standard multijet trigger			
	$t\bar{t}$ + 0 jets		$t\bar{t}$ + 1 jets	
	N_{sel}/N_{tot}	%	N_{sel}/N_{tot}	%
$t\bar{t} \to jjbjjb$	2921/26745	10.92 ± 0.19	11057/29272	37.77 ± 0.28
$t\bar{t} \to e\nu bjjb$	793/8897	8.91 ± 0.30	3350/9805	34.17 ± 0.48
$t\bar{t} \to \mu\nu bjjb$	448/8698	5.15 ± 0.24	2681/9867	27.17 ± 0.45
$t\bar{t} \to \tau\nu bjjb$	541/8884	6.09 ± 0.25	2889/9780	29.54 ± 0.46
$t\bar{t} \to e\nu be\nu b$	61/720	8.47 ± 1.04	272/836	32.54 ± 1.62
$t\bar{t} \to e\nu b\mu\nu b$	54/1481	3.65 ± 0.49	387/1617	23.93 ± 1.06
$t\bar{t} \to e\nu b\tau\nu b$	68/1536	4.43 ± 0.52	444/1647	26.96 ± 1.09
$t\bar{t} \to \mu\nu b\mu\nu b$	10/746	1.34 ± 0.42	151/792	19.07 ± 1.40
$t\bar{t} \to \mu\nu b\tau\nu b$	37/1480	2.50 ± 0.41	335/1566	21.39 ± 1.04
$t\bar{t} \to \tau\nu b\tau\nu b$	17/713	2.38 ± 0.57	190/818	23.23 ± 1.48
Total	4950/59900	8.26 ± 0.11	21756/66000	32.96 ± 0.18

Table 4.7: Number of events selected by the standard multijet trigger, and relative percentages, for the different samples of $t\bar{t}$ pairs and for their principal final states.

	Pixel Trigger			
	$t\bar{t}$ + 0 jets		$t\bar{t}$ + 1 jets	
	N_{sel}/N_{tot}	%	N_{sel}/N_{tot}	%
$t\bar{t} \to jjbjjb$	4976/26745	18.61 ± 0.24	14539/29272	49.67 ± 0.29
$t\bar{t} \to e\nu bjjb$	1357/8897	15.25 ± 0.38	4346/9805	44.32 ± 0.50
$t\bar{t} \to \mu\nu bjjb$	720/8698	8.28 ± 0.30	3450/9867	34.97 ± 0.48
$t\bar{t} \to \tau\nu bjjb$	884/8884	9.95 ± 0.32	3549/9780	36.29 ± 0.49
$t\bar{t} \to e\nu be\nu b$	93/720	12.92 ± 1.25	328/836	39.23 ± 1.69
$t\bar{t} \to e\nu b\mu\nu b$	93/1481	6.28 ± 0.63	463/1617	28.63 ± 1.12
$t\bar{t} \to e\nu b\tau\nu b$	122/1536	7.94 ± 0.69	503/1647	30.54 ± 1.14
$t\bar{t} \to \mu\nu b\mu\nu b$	17/746	2.28 ± 0.55	164/792	20.71 ± 1.44
$t\bar{t} \to \mu\nu b\tau\nu b$	54/1480	3.65 ± 0.49	368/1566	23.50 ± 1.07
$t\bar{t} \to \tau\nu b\tau\nu b$	34/713	4.77 ± 0.80	200/818	24.45 ± 1.50
Total	8350/59900	13.94 ± 0.14	27910/66000	42.29 ± 0.19

Table 4.8: Number of events selected by the standard multijet trigger, and relative percentages, for the different samples of $t\bar{t}$ pairs and for their principal final states.

	Standard multijet trigger			
	$t\bar{t}$ + 2 jets		$t\bar{t}$ + 3 jets	
	N_{sel}/N_{tot}	%	N_{sel}/N_{tot}	%
$t\bar{t} \to jjbjjb$	29655/43629	67.97 ± 0.22	5648/6492	87.00 ± 0.42
$t\bar{t} \to e\nu bjjb$	9258/14538	63.68 ± 0.40	1930/2256	85.55 ± 0.74
$t\bar{t} \to \mu\nu bjjb$	8163/14520	56.22 ± 0.41	1671/2127	78.56 ± 0.89
$t\bar{t} \to \tau\nu bjjb$	8543/14541	58.75 ± 0.41	1809/2250	80.40 ± 0.84
$t\bar{t} \to e\nu be\nu b$	740/1217	60.80 ± 1.40	154/186	82.80 ± 2.77
$t\bar{t} \to e\nu b\mu\nu b$	1164/2361	49.30 ± 1.03	253/345	73.33 ± 2.38
$t\bar{t} \to e\nu b\tau\nu b$	1286/2420	53.14 ± 1.01	283/377	75.07 ± 2.23
$t\bar{t} \to \mu\nu b\mu\nu b$	514/1188	43.27 ± 1.44	126/174	72.41 ± 3.39
$t\bar{t} \to \mu\nu b\tau\nu b$	1165/2484	46.90 ± 1.00	271/380	71.32 ± 2.32
$t\bar{t} \to \tau\nu b\tau\nu b$	600/1261	47.58 ± 1.41	135/181	74.59 ± 3.24
Total	61088/98159	62.23 ± 0.15	12280/14768	83.15 ± 0.31

Table 4.9: Number of events selected by the standard multijet trigger, and relative percentages, for the different samples of $t\bar{t}$ pairs and for their principal final states.

	Pixel Trigger			
	$t\bar{t}$ + 2 jets		$t\bar{t}$ + 3 jets	
	N_{sel}/N_{tot}	%	N_{sel}/N_{tot}	%
$t\bar{t} \to jjbjjb$	33629/43629	77.08 ± 0.20	5857/6492	90.22 ± 0.37
$t\bar{t} \to e\nu bjjb$	10484/14538	72.11 ± 0.37	1992/2256	88.30 ± 0.68
$t\bar{t} \to \mu\nu bjjb$	9553/14520	65.79 ± 0.39	1824/2127	85.75 ± 0.76
$t\bar{t} \to \tau\nu bjjb$	9604/14541	66.05 ± 0.39	1917/2250	85.20 ± 0.75
$t\bar{t} \to e\nu be\nu b$	812/1217	66.72 ± 1.35	151/186	81.18 ± 2.87
$t\bar{t} \to e\nu b\mu\nu b$	1347/2361	57.05 ± 1.02	275/345	79.71 ± 2.17
$t\bar{t} \to e\nu b\tau\nu b$	1385/2420	57.23 ± 1.01	315/377	83.55 ± 1.91
$t\bar{t} \to \mu\nu b\mu\nu b$	609/1188	51.26 ± 1.45	140/174	80.46 ± 3.01
$t\bar{t} \to \mu\nu b\tau\nu b$	1304/2484	52.50 ± 1.00	302/380	79.47 ± 2.07
$t\bar{t} \to \tau\nu b\tau\nu b$	672/1261	53.29 ± 1.41	134/181	74.03 ± 3.26
Total	69399/98159	70.70 ± 0.15	12907/14768	87.40 ± 0.27

Table 4.10: Number of events selected by the standard multijet trigger, and relative percentages, for the different samples of $t\bar{t}$ pairs and for their principal final states.

	Standard multijet trigger	
	$t\bar{t}$ + 4 jets	
	N_{sel}/N_{tot}	%
$t\bar{t} \to jjbjjb$	2235/2363	94.58 ± 0.47
$t\bar{t} \to e\nu bjjb$	785/838	93.67 ± 0.84
$t\bar{t} \to \mu\nu bjjb$	702/778	90.23 ± 1.06
$t\bar{t} \to \tau\nu bjjb$	698/759	91.96 ± 0.99
$t\bar{t} \to e\nu be\nu b$	72/81	88.89 ± 3.49
$t\bar{t} \to e\nu b\mu\nu b$	130/147	88.44 ± 2.64
$t\bar{t} \to e\nu b\tau\nu b$	95/116	81.90 ± 3.58
$t\bar{t} \to \mu\nu b\mu\nu b$	51/64	79.69 ± 5.03
$t\bar{t} \to \mu\nu b\tau\nu b$	125/144	86.81 ± 2.82
$t\bar{t} \to \tau\nu b\tau\nu b$	52/62	83.87 ± 4.67
Total	4945/5352	92.39 ± 0.36

Table 4.11: Number of events selected by the standard multijet trigger, and relative percentages, for the different samples of $t\bar{t}$ pairs and for their principal final states.

	Pixel Trigger	
	$t\bar{t}$ + 4 jets	
	N_{sel}/N_{tot}	%
$t\bar{t} \to jjbjjb$	2230/2363	94.37 ± 0.47
$t\bar{t} \to e\nu bjjb$	784/838	93.56 ± 0.85
$t\bar{t} \to \mu\nu bjjb$	719/778	92.42 ± 0.95
$t\bar{t} \to \tau\nu bjjb$	713/759	93.94 ± 0.87
$t\bar{t} \to e\nu be\nu b$	74/81	91.36 ± 3.12
$t\bar{t} \to e\nu b\mu\nu b$	139/147	94.56 ± 1.87
$t\bar{t} \to e\nu b\tau\nu b$	103/116	88.79 ± 2.93
$t\bar{t} \to \mu\nu b\mu\nu b$	59/64	92.19 ± 3.35
$t\bar{t} \to \mu\nu b\tau\nu b$	130/144	90.28 ± 2.47
$t\bar{t} \to \tau\nu b\tau\nu b$	60/62	96.77 ± 2.24
Total	5011/5352	93.63 ± 0.33

Table 4.12: Number of events selected by the standard multijet trigger, and relative percentages, for the different samples of $t\bar{t}$ pairs and for their principal final states.

4.3.3 Inclusive QCD events

Collisions mediated by strong interaction, and in particular $gg \rightarrow gg$ scattering, dominate the cross section of all jet triggers at LHC. As we have seen in the previous chapters, their selection at the first trigger level is reduced through quite stringent requirements on the jets energy. Considering that the aquisition frequency of the multijet trigger is of the order of $\sim$ 3 kH, which, at the instantaneous luminosity of 10^{34} cm^{-2}s^{-1}, correspond to 300 nanobarns, it immediately becomes clear the entity of the problem in front of us: in order to have an inclusive QCD statistic equivalent to that of $t\bar{t}H$, we would need 400.000 events passing the trigger requests for each $t\bar{t}H$ event, a statistic totally out of our range.
Like for the $t\bar{t}H$ case, pythia is used for the simulation of the events.
We will use two different approximations to develop an suitable background description with the about 3.5 millions of QCD events we have generated. The first one is based on the parametrerization of a three-dimensional matrix describing the probability for a jet to satisfy the b-tagging requirements for the chosen algorithm, with a technique similar to the one used at CDF for the background estimate in the search for top-antitop pairs in the missing Et + jets channel: this will allow us to use many times a QCD event, each time assigning the b-tags to different jets with a different total probability. The second exploits the intstrumental nature of the missing transverse energy in QCD events, modulating the parton energy measurement on their own energy resolution functions, and generating with the right probability events with high missing transverse energy. These techniques will be described in more detail in section 4.5. Here it is enough to say that they allow to treat, even if in an approximate way, a problem otherwise difficult to face with limited computational power and time.
In the following table it is shown the number of events generated and filtered with the multijet trigger for eight different strong interaction simulations, divided by the interval of the momentum of final partons in the $2 \rightarrow 2$ tree level process.

From the tables 4.13 and 4.14 it appears clear that a significant portion of the events accepted by the multijet trigger comes from scattering events of low energy, surviving the trigger thanks to an insufficient energetic resolution available for calorimetric clusters at level 1 and the help, sometimes important, of additional interactions. The comparison with the pixel trigger shows a reduction in the acceptance of events from high pt bins. The requirement from the pixels, removes events with high energy in one single jet which were passing the cut on the first jet Et, the main contributor to the rate of the multijet trigger.

Standard multijet trigger				
Pt bin	Nsel/Ntot	% accepted	σ generation	L1 σ
30-50 GeV	35/300000	0.0117 ± 0.0020	0.163 mb	0.0000190 ± 0.0000032 mb
50-80 GeV	260/364000	0.0714 ± 0.0044	21.6 μb	0.01543 ± 0.00096 μb
80-120 GeV	4725/386000	1.224 ± 0.018	3.08 μb	0.03770 ± 0.00055 μb
120-170 GeV	25620/407100	6.293 ± 0.038	494 nb	31.09 ± 0.19 nb
170-230 GeV	153691/536000	28.674 ± 0.062	101 nb	28.960 ± 0.062 nb
230-300 GeV	350056/447100	78.295 ± 0.062	24.5 nb	19.182 ± 0.015 nb
300-380 GeV	512921/530000	96.778 ± 0.024	6.24 nb	6.0389 ± 0.0015 nb
380-incl GeV	506215/508000	99.6486 ± 0.0083	2.821 nb	2.81109 ± 0.00023 nb

Table 4.13: Generated QCD events and percentage of filtered events after the level 1 multijet trigger.

Pixel trigger				
Pt bin	Nsel/Ntot	% accepted	σ generation	L1 σ
30-50 GeV	67/300000	0.0223 ± 0.0027	0.163 mb	0.0000364 ± 0.0000044 mb
50-80 GeV	437/364000	0.1201 ± 0.0057	21.6 μ b	0.0259 ± 0.0012 μb
80-120 GeV	3392/386000	0.879 ± 0.015	3.08 μ b	0.02707 ± 0.00046 μb
120-170 GeV	18102/407100	4.447 ± 0.032	494 nb	21.97 ± 0.16 nb
170-230 GeV	107381/536000	20.034 ± 0.055	101 nb	20.234 ± 0.055 nb
230-300 GeV	262622/447100	58.739 ± 0.074	24.5 nb	14.391 ± 0.018 nb
300-380 GeV	428025/530000	80.759 ± 0.054	6.24 nb	5.0394 ± 0.0034 nb
380-incl GeV	461606/508000	90.867 ± 0.040	2.821 nb	2.5634 ± 0.0011 nb

Table 4.14: Generated QCD events and percentage of filtered events after the level 1 multijet+pixel trigger.

4.3.4 Events of W+jets productions with leptonic decay of the W boson

Among the processes generating true missing transverse energy, the production of W with leptonic decays is the one with the highest cross section: at LHC will be produced, in the high luminosity phase, about 400 W events per second. Nevertheless, the requirement of many hadronic jets in the final state makes this a small contribution to the sample, but not totally negligible, because of the extreme rarity of the signal we are looking for.
The W+jets samples were generated with AlpGen and Pythia using the same versions as for the $t\bar{t}$+jets described before [58]. They are divided in subsamples each with a different number of partons generated together with the W in the final state and a different Pt of the W boson. We have used the totality of the events at our disposal for a complete reconstruction of the pixel detector (requiring compatibility with version 1.7.0 of CMSSW), using, as for the other samples, the full simulation of the pixel detector and a fast simulation for the rest of the detector with FastSim.
Table 4.18 gives a detailed view of the simulations used and the events satisfying the standard multijet trigger. In table 4.16, instead, the multijet-pixel trigger case is considered.

From the tables we can see that the majority of level 1 passing events in this sample come from events in which the W boson is produced with high transverse momentum and a sufficient number of hadronic jets. The pixel trigger reduces the overall acceptance in all the channels.

Standard multijet trigger					
sample	Pt bin (GeV)	Nsel/Ntot	% accepted	σ	L1 σ
W + 0 jets		7/88000	0.0080 ± 0.0030	45 nb	0.0036 ± 0.0014 nb
W + 1 jet 0	pt < 100 100 < pt < 300	27/40000 7512/100500	0.068 ± 0.013 7.475 ± 0.083	9.2 nb 250 pb	0.0062 ± 0.0012 nb 18.69 ± 0.21 pb
W + 2 jet	pt < 100 100 < pt < 300	536/99520 10228/105300	0.539 ± 0.023 9.713 ± 0.091	2.5 nb 225 pb	0.01346 ± 0.00058 nb 21.85 ± 0.21 pb
W + 3 jet	pt < 100 100 < pt < 300	2109/107900 12286/86260	1.955 ± 0.042 14.24 ± 0.12	590 pb 100 pb	11.53 ± 0.25 pb 14.24 ± 0.12 pb
W + 4 jet	pt < 100 100 < pt < 300	4333/83040 6278/30800	5.218 ± 0.077 20.38 ± 0.23	125 pb 40 pb	6.522 ± 0.096 pb 8.153 ± 0.092 pb
W + 5 jet	pt < 100 100 < pt < 300	6421/59020 11828/41860	10.88 ± 0.13 28.26 ± 0.22	85 pb 40 pb	9.25 ± 0.11 pb 11.302 ± 0.088 pb

Table 4.15: W + jets generations produced with AlpGen and fractions of events passing the multijet trigger.

Pixel trigger					
sample	Pt bin (GeV)	Nsel/Ntot	% accepted	σ	L1 σ
W + 0 jets		8/88000	0.0091 ± 0.0032	45 nb	0.0041 ± 0.0014 nb
W + 1 jet	pt < 100 100 < pt < 300	17/40000 2852/100500	0.042 ± 0.010 2.838 ± 0.052	9.2 nb 250 pb	0.00391 ± 0.00095 nb 7.09 ± 0.13 pb
W + 2 jet	pt < 100 100 < pt < 300	354/99520 5994/105300	0.356 ± 0.019 5.692 ± 0.071	2.5 nb 225 pb	0.00889 ± 0.00047 nb 12.80 ± 0.16 pb
W + 3 jet	pt < 100 100 < pt < 300	1999/107900 9064/86260	1.853 ± 0.041 10.51 ± 0.10	590 pb 100 pb	10.93 ± 0.24 pb 10.51 ± 0.10 pb
W + 4 jet	pt < 100 100 < pt < 300	4356/83040 5436/30800	5.246 ± 0.077 17.65 ± 0.22	125 pb 40 pb	6.557 ± 0.097 pb 7.060 ± 0.087 pb
W + 5 jet	pt < 100 100 < pt < 300	6960/59020 11233/41860	11.79 ± 0.13 26.83 ± 0.22	85 pb 40 pb	10.02 ± 0.11 pb 10.734 ± 0.087 pb

Table 4.16: W + jets generations produced with AlpGen and fractions of events passing the multijet+pixel trigger.

4.3.5 Z + jets events production with invisible Z decay

Another process contributiong with true missing transverse energy to our sample is the one in which the Z boson, produced in association with hadronic jets, decays in two neutrinos. The branching ratio of the Z in neutrinos is important (20%), but the Z production cross section is inferior of a factor 3 to the W production. Therefore, we expect a lower contribution from this channel to the sample selected by the trigger, once a significant missing transverse energy will be required.

Standard multijet trigger					
sample	Pt bin	Nsel/Ntot	% accepted	σ	L1 σ
Z$\rightarrow \nu\nu$	120-170 GeV 170-230 GeV	832/29900 4522/25600	2.783 ± 0.095 17.66 ± 0.24	51.47 pb 15.52 pb	1.432 ± 0.049 pb 2.741 ± 0.037 pb

Table 4.17: $Z \rightarrow \nu\nu$ generations produced with Pythia and fractions of events passing the multijet trigger.

Pixel trigger					
sample	Pt bin	Nsel/Ntot	% accepted	σ	L1 σ
$Z \rightarrow \nu\nu$	120-170 GeV 170-230 GeV	295/29900 2067/25600	0.987 ± 0.057 8.07 ± 0.17	51.47 pb 15.52 pb	0.508 ± 0.029 pb 1.253 ± 0.026 pb

Table 4.18: $Z \rightarrow \nu\nu$ generations produced with Pythia and fractions of events passing the multijet+pixel trigger.

4.3.6 Electroweak production events of single top quark

Even if the electroweak production of top quark has a much lower cross section than top-antitop strong interaction production, this process can contribute to our sample even after the requirement of missing transverse energy and high jet multiplicity. In table 4.19 we show the cross section of the considered simulated events and the number of events passing the standard level 1 multijet trigger selection. In table 4.20 we show the same informations when the multijet trigger is confirmed with the Pixel-Jets.

Standard multijet trigger				
sample	Nsel/Ntot	% accepted	σ	L1 σ
t$\rightarrow e\nu$ b	8/92000	0.0087 $\pm$ 0.0031	27.43 pb	0.00239 $\pm$ 0.00084 pb
t$\rightarrow \mu\nu$ b	8/94000	0.0085 $\pm$ 0.0030	26.97 pb	0.00230 $\pm$ 0.00081 pb
t$\rightarrow \tau\nu$ b	8/94000	0.0085 $\pm$ 0.0030	28.71 pb	0.00244 $\pm$ 0.00086 pb

Table 4.19: $Z \rightarrow \nu\nu$ generations produced with Pythia and fractions of events passing the multijet trigger.

Pixel trigger				
sample	Nsel/Ntot	% accepted	σ	L1 σ
t$\rightarrow e\nu$ b	3/92000	0.0033 $\pm$ 0.0019	27.43 pb	0.00089 $\pm$ 0.00052 pb
t$\rightarrow \mu\nu$ b	3/94000	0.0032 $\pm$ 0.0018	26.97 pb	0.00086 $\pm$ 0.00050 pb
t$\rightarrow \tau\nu$ b	3/94000	0.0032 $\pm$ 0.0018	28.71 pb	0.00092 $\pm$ 0.00053 pb

Table 4.20: $Z \rightarrow \nu\nu$ generations produced with Pythia and fractions of events passing the multijet+pixel trigger.

4.4 Definition of a preliminary event selection

The first step in the selection of a sample where the component of $t\bar{t}$H events may be put in evidence by our analysis is that of defining jets and missing transverse energy, which we will later study to picture the event kinematics.

To identify hadronic jets we base on the iterative cone algorithm in use in CMS, which adopts a cone with radius R=0.5. The algorithm is dubbed "IC5" in the following. The choice of the IC5 algorithm appears the most promising among the many available, especially given the high luminosity conditions in which we ideally place our search. At high luminosity, a cone with a wider radius would be negatively influenced by the energy flux coming from two dozen additional collisions. Similarly, a successive recombination algorithm such as Kt is expected to be less stable, due to the variability in the extension of energy deposits in the eta-phi plane in the presence of a large amount of noise.

We select jets with transverse energy larger than 25 GeV contained in the rapidity range $-3.0 < \eta < 3.0$. With such a choice, we maximize the fraction of signal events of the single lepton kind $(t\bar{t}H \rightarrow (jjb)(l\nu\bar{b})(b\bar{b})$ where all hadronic top partons and Higgs decay partons are correctly matched to a jet in the list. Fig. 4.4 shows the fraction of correct matches (defined with a maximum $\eta - \phi$ distance of 0.4) for the five quarks. The figure also shows that the number of jets correctly matching a parton among those emitted by the $t\bar{t}H$ decay is maximized for the choice of cuts above. Further, the average eta-phi distance shows a minimum in the region corresponding to our choice of a jet collection, and so does the number of lost matches. Fig.4.4 shows separately the fraction of associated partons from higgs and top decay. Again, the choice [25, 3.0] for the boundaries defining the jet collection seems sound.

The choice of a wide range of rapidity –wider than those used in other similar high-Pt analyses in CMS- is justified by the plots shown above, but we must keep in mind that by considering jets with high rapidity we open the door to QCD backgrounds, which are notoriously less central and rich with high-rapidity jets. We will show later, in Sec. 4.7, how it is possible to discard the events accepted by this pre-selection, with variables explicitly picturing the centrality of the event.

Fig. 4.4 shows the distribution of jet multiplicity satisfying the mentioned cuts, for signal and for the main background processes, after the simulation of the standard multijet trigger. One notes that the $t\bar{t}H$ signal typically shows 7 jets with the required characteristics; inclusive top production has usually two jets less, as expected.

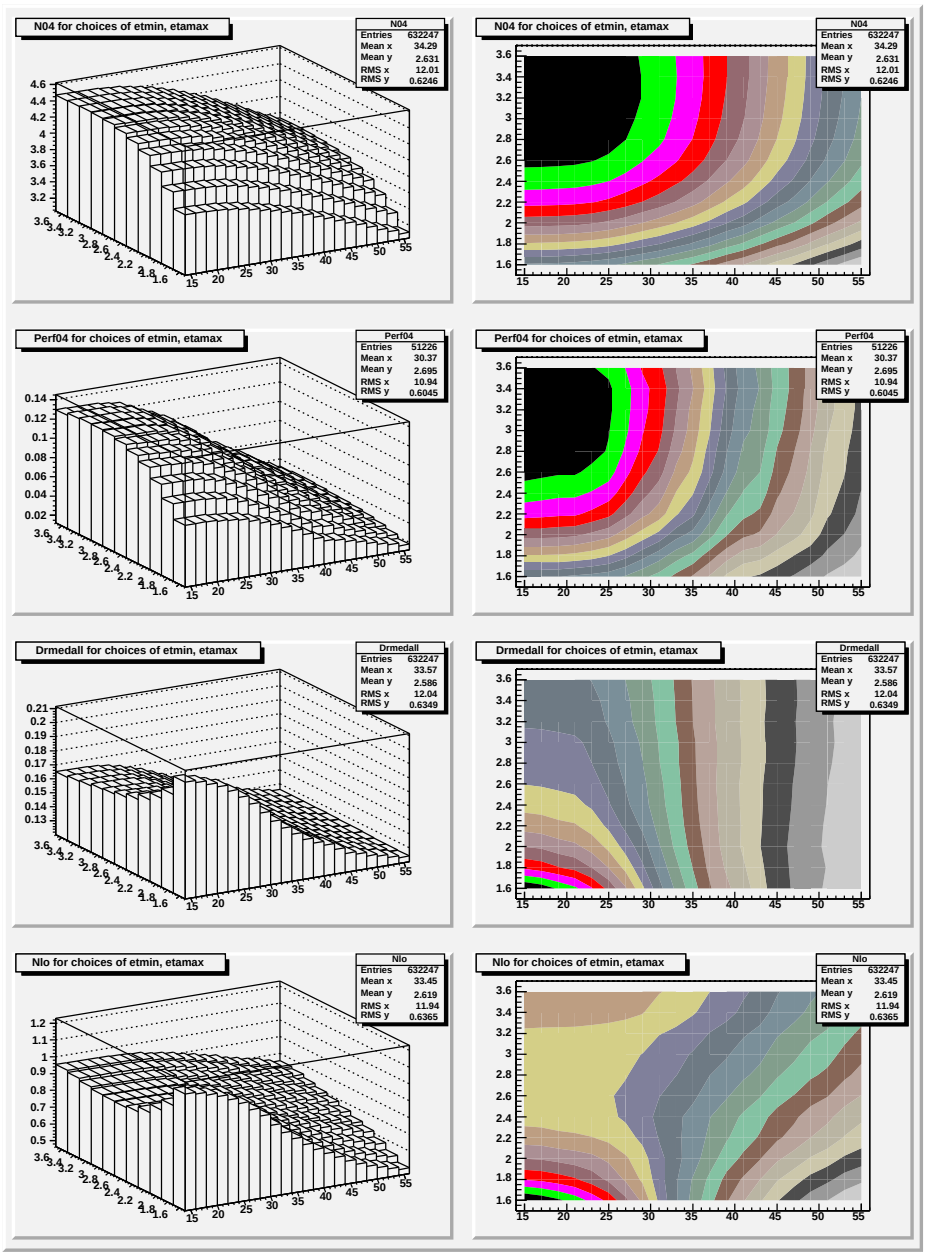


Figure 4.1: *Optimization of the definition of the jet collection.*

As far as missing transverse energy is concerned, there exists a standard definition in CMS for computing its value, and a routine in the CMSSW

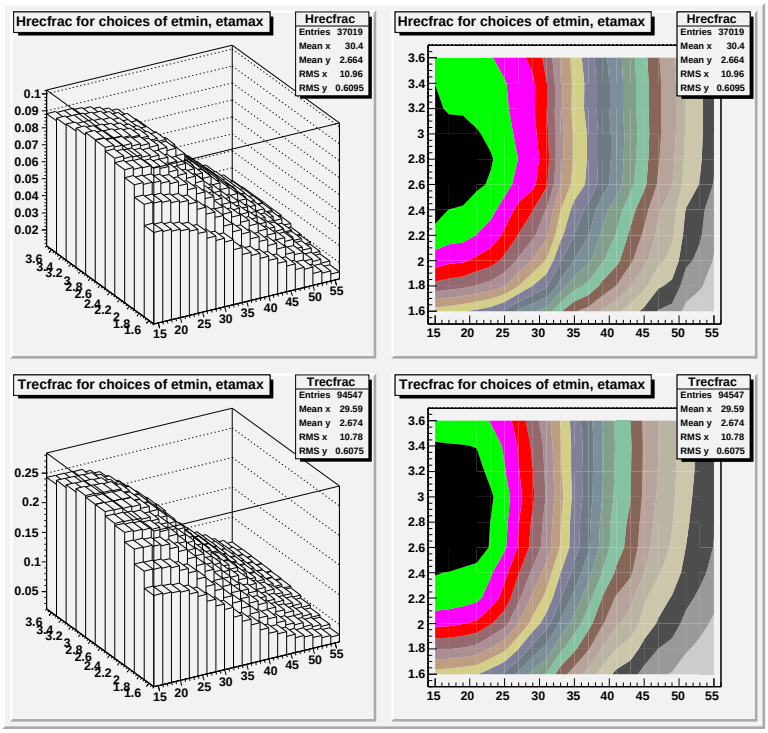


Figure 4.2: *Optimization of the definition of the jet collection: the fraction of higgs (top) and top (bottom) decays with all partons matched to jets belonging to a collection is shown as a function of the minimum jet Et and maximum absolute rapidity which define the collection.*

software package provides magnitude and direction in the transverse plane, besides the sum of transverse energy with which missing Et was computed. The significance of missing Et is then given by the ratio between its magnitude and the square root of the sum of transverse energy.

Despite the availability of a standard missing E_t, it is useful to verify the effectiveness of its definition by studying alternative definitions. It is in fact well known that as a function of the physical process one is interested in reconstructing, the optimal operative definition with which missing energy is computed varies. In general, the parameters subjected to optimization are several, but those it is interesting to focus on are the maximum rapidity η^{max} within which to consider energy deposits in the calorimeter, the minimum energy they have to be considered in the computation E_t^{min}, and the level of energy correction applied to jets. We studied many different definitions of missing transverse energy by varying these parameters, obtaining distributions of the width of the missing energy along the x or y directions

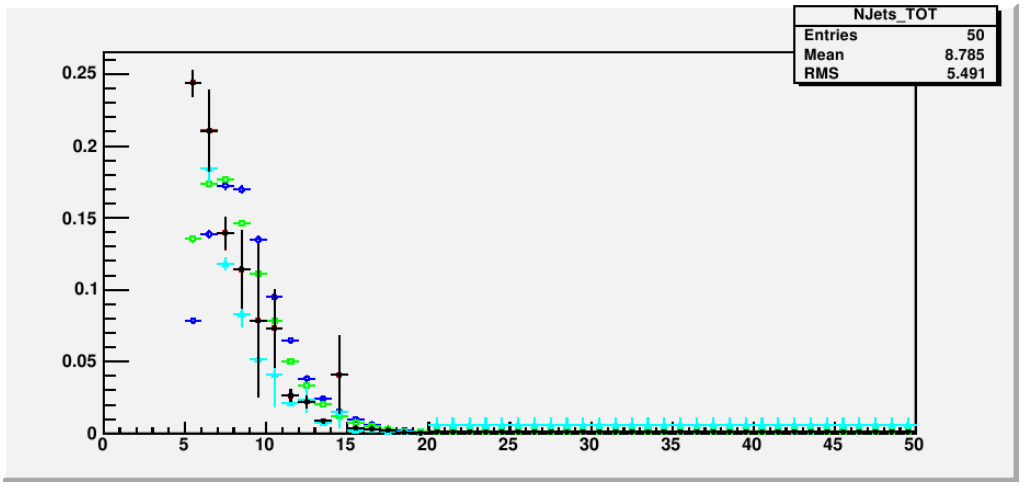


Figure 4.3: *Multiplicity distribution of IC5 jets with* $E_t > 25GeV$ *and* $|\eta| < 3.0$ *for* $t\bar{t}H$ *signal (blue), QCD inclusive production (in red), inclusive top production (in green) and W+jets production (in cyan). All simulations have been filtered with standard multijet trigger simulation. The distributions are normalized to unit area.*

as a function of the sum of transverse energy considered in the sums. In Fig. 4.4 we observe the distribution of missing energy along x for QCD inclusive events (which contain a negligible amount of real penetrating particles such as muons or neutrinos) in different intervals of total transverse energy. The widths of gaussian fits to those distributions are shown in the figure in the lower right corner of 4.4 as a function of the total transverse energy, for different choices of the η_{max} and E_t^{min} parameters mentioned above.

From that study results that the standard definition of missing E_t is the one providing the best resolution even in events with a high jet multiplicity, in a range of total E_t typical for events like the ones we seek. The missing E_t resolution is observed to follow with pretty good accuracy the simple law

$$\sigma(ME_t) = 0.753(\Sigma E_t)^{0.509}, \tag{4.1}$$

as shown in Fig. 4.4. The scaling with a simple 0.5 power of the Sum Et is typical of the dominant measurement errors of calorimetric deposits, whose determination has a uncertainty which also scales with the square root of its value. In view of the above considerations, we consider valid a definition of missing Et significance as

$S = ME_t/\sqrt{\Sigma E_t}$

In Fig. 4.4 is shown a comparison between missing Et significance in QCD inclusive events and $t\bar{t}H$ signal. A selection requiring $S > 3.0$ promises an

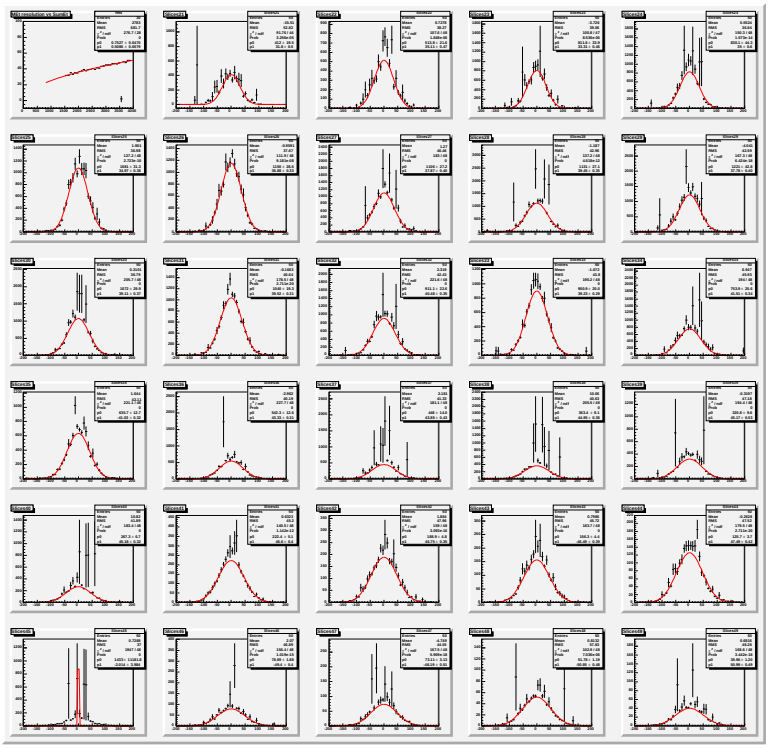


Figure 4.4: x component of the missing transverse energy for QCD events in differnet intervals of total transverse energy with corresponding gaussian fits. The plot on the upper left corner is displayed zoomed in **??** and represents the widths of the gaussian fits in the previous distributions.

increase of the significance of signal, besides defining the sought final state as "missing Et + jets", as decided from the outset in this study.

The preliminary event selection we apply to the data therefore requires Nj(IC5)>= 5 , $S > 3.0$. Tables **??** show the efficiency of the selection on the different samples. Tables **??** show the same information in the case of a trigger using also pixel information. We do not show the single top because with our samples, only 1 event passes the multijet trigger and no event passes the pixel trigger.

Standard multijet trigger		
sample	type	efficiency (%)
$t\bar{t}H$		9.208 ± 0.023
$t\bar{t}$	+ 0 jets	2.205 ± 0.060
	+ 1 jets	9.99 ± 0.12
	+ 2 jets	19.61 ± 0.13
	+ 3 jets	28.02 ± 0.37
	+ 4 jets	33.05 ± 0.64

Table 4.21: $t\bar{t}H$ and $t\bar{t}$ efficiency of the selection up to Nj≥5 and missing Et significance > 3, for events selected with the multijet trigger.

Pixel trigger		
sample	type	efficiency (%)
$t\bar{t}H$		12.588 ± 0.026
$t\bar{t}$	+ 0 jets	3.010 ± 0.070
	+ 1 jets	11.34 ± 0.12
	+ 2 jets	21.21 ± 0.13
	+ 3 jets	29.50 ± 0.38
	+ 4 jets	33.39 ± 0.64

Table 4.22: $t\bar{t}H$ and $t\bar{t}$ efficiency of the selection up to Nj≥5 and missing Et significance > 3, for events selected with the pixel trigger.

Standard multijet trigger	
P_t bin	efficiency (%)
30-50 GeV	0.00118 ± 0.00083
50-80 GeV	0.0110 ± 0.0037
80-120 GeV	0.0408 ± 0.0038
120-170 GeV	0.203 ± 0.015
170-230 GeV	0.941 ± 0.031
230-300 GeV	2.809 ± 0.052
300-380 GeV	4.571 ± 0.062
380-incl GeV	7.281 ± 0.081

Table 4.23: QCD efficiency of the selection up to Nj≥5 and missing Et significance > 3, for events selected with the multijet trigger.

Pixel trigger	
P_t bin	efficiency (%)
30-50 GeV	0.00059 ± 0.00059
50-80 GeV	0.0037 ± 0.0021
80-120 GeV	0.0145 ± 0.0023
120-170 GeV	0.100 ± 0.010
170-230 GeV	0.630 ± 0.025
230-300 GeV	2.004 ± 0.044
300-380 GeV	3.709 ± 0.056
380-incl GeV	6.562 ± 0.078

Table 4.24: QCD efficiency of the selection up to Nj≥5 and missing Et significance > 3, for events selected with the pixel trigger.

Standard multijet trigger		
sample	Pt bin (GeV)	efficiency (%)
W + 0 jets		0.0011 ± 0.0011
W + 1 jet	pt < 100 100 < pt < 300	0.0175 ± 0.0066 2.879 ± 0.053
W + 2 jet	pt < 100 100 < pt < 300	0.114 ± 0.011 4.227 ± 0.062
W + 3 jet	pt < 100 100 < pt < 300	0.411 ± 0.019 7.342 ± 0.089
W + 4 jet	pt < 100 100 < pt < 300	1.275 ± 0.039 12.07 ± 0.19
W + 5 jet	pt < 100 100 < pt < 300	2.702 ± 0.067 17.42 ± 0.19

Table 4.25: W+jets efficiency of the selection up to Nj≥5 and missing Et significance > 3, for events selected with the multijet trigger.

Pixel trigger		
sample	Pt bin (GeV)	efficiency (%)
W + 0 jets		0.0011 ± 0.0011
W + 1 jet	pt < 100 100 < pt < 300	0.0050 ± 0.0035 1.223 ± 0.035
W + 2 jet	pt < 100 100 < pt < 300	0.0512 ± 0.0072 2.512 ± 0.048
W + 3 jet	pt < 100 100 < pt < 300	0.311 ± 0.017 5.232 ± 0.076
W + 4 jet	pt < 100 100 < pt < 300	1.031 ± 0.035 10.07 ± 0.17
W + 5 jet	pt < 100 100 < pt < 300	2.431 ± 0.063 16.1 4± 0.18

Table 4.26: W+jets efficiency of the selection up to Nj≥5 and missing Et significance > 3, for events selected with the pixel trigger.

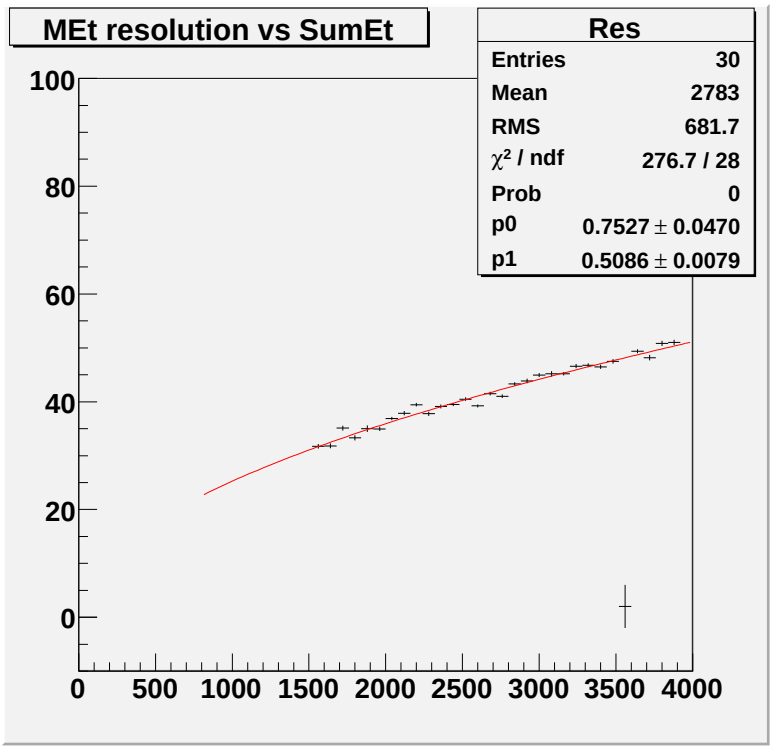


Figure 4.5: widths of the gaussian fits in the distributions in 4.4.

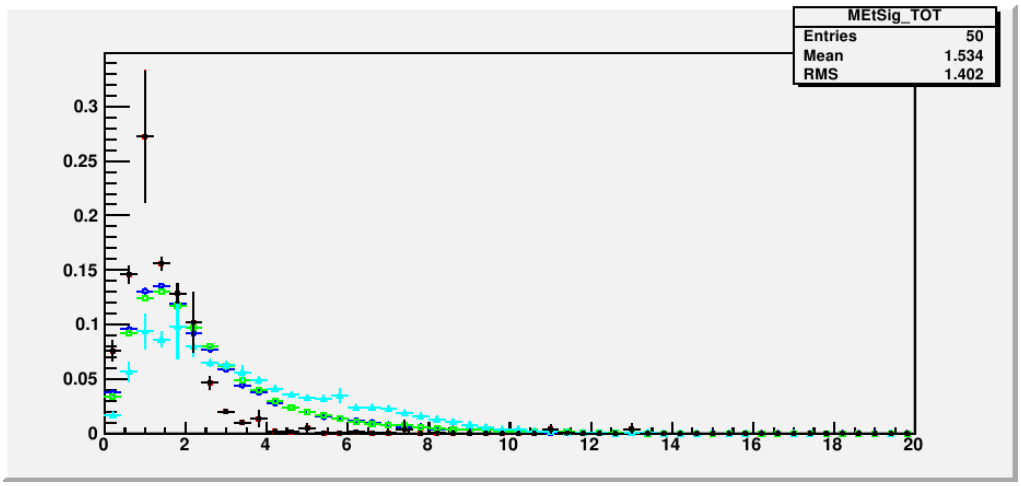


Figure 4.6: Comparison between missing Et significance in QCD inclusive events and $t\bar{t}$H signal.

Standard multijet trigger		
sample	Pt bin	efficiency (%)
Z$\rightarrow \nu\nu$	120-170 GeV 170-230 GeV	1.308 $\pm$ 0.066 10.76 $\pm$ 0.19

Table 4.27: Z$\rightarrow \nu\nu$ efficiency of the selection up to Nj$\geq$5 and missing Et significance > 3, for events selected with the multijet trigger.

Pixel trigger		
sample	Pt bin	efficiency (%)
Z$\rightarrow \nu\nu$	120-170 GeV 170-230 GeV	0.672 $\pm$ 0.047 5.39 $\pm$ 0.14

Table 4.28: Z$\rightarrow \nu\nu$ efficiency of the selection up to Nj$\geq$5 and missing Et significance > 3, for events selected with the pixel trigger.

4.5 B-tagging jets with track counting

The identification of jets with heavy flavour represents the most powerful handle for the extraction of top decays in events with hadronic jets, because generic jets produced from strong interactions have a small probability of fragmenting bottom or charm quarks. In the case of the signal we seek, b-tagging is even more important, because the final state of $t\bar{t}H$ production with a $H \to b\bar{b}$ decay promises not two, but up to four jets with b-quarks. For background processes, and in particular inclusive QCD and W+jets production, it is extremely rare to possess as many as four b-jets. Because of that, an algorithm that guarantees high efficiency of b-jet identification, even at the cost of a non negligible fake rate for jets with light quarks or gluons, is to be preferred to one offering high purity at the price of a smaller efficiency.

We focused our attention on an algorithm studied in good detail in CMS, which for its simplicity should be less affected than more complex others (such as are methods involving the direct reconstruction of secondary vertices in jets) by an imperfect simulation of the effective experimental conditions of data taking at high luminosity. It is the algorithm called "track counting" [59], which exploits the precise measurement of track coordinates thanks to the pixel detector and the many layers of silicon strip detectors present in the central part of CMS. The measurement of impact parameter, for central tracks ($|\eta| < 2.40$) with transverse momentum larger than $10GeV$, has an uncertainty of 30 micrometers, an appropriate value for the task of discriminating in-flight decays of heavy objects.

The track counting algorithm produces a list of tracks passing quality requirements, ordered by the parameter $S = d/s_d$ computed between the impact parameter d and its predicted error s_d. The impact parameter of the n-th track in the list is then used as a discriminating quantity. For instance, the so-called "Loose tagging" requires that $S(2) > 2.3$; "Medium tagging" requires $S(2) > 5.3$; and "Tight tagging" requires $S(3) > 4.8$. These values have been optimized on samples of b-jets and light-quark and gluon jets in QCD simulations. The following table details the expected tagging rates for jet mixtures.

4.5.1 B-tagging requirement

Tables **??** and **??** the efficiency for $t\bar{t}H$ signal and each of the background sources after the selection of events with at least five jets, missing Et significance larger than three, and a minimum number of b-tags (none, two, three, and four) after the multijet and the pixel-multijet trigger respectively. The b-tags are counted only in the leading 8 jets: the rationale of this choice is to

	b-efficiency	c-mistag	uds-mistag	g-mistag	udsg-mistag
Loose (S(2)>2.3)	70.49	32.33	8.64	10.43	9.98
Medium (S(2)>5.3)	50.30	10.77	0.92	0.98	0.96
Tight (S(3)>4.8)	31.94	2.93	0.10	0.11	0.10

Table 4.29: *Rates of b-tagging in different jet mixtures, for the three operating points (S(2)>2.3 , S(2)>5.3, S(3)>4.8) Numbers are from Ref. [60].*

Standard multijet trigger				
sample	type	2 b-tags (%)	3 b-tags (%)	4 b-tags (%)
$t\bar{t}H$		4.611 ± 0.016	1.6141 ± 0.0099	0.3125 ± 0.0044
$t\bar{t}$	+ 0 jets	0.574 ± 0.031	0.0467 ± 0.0088	0.0017 ± 0.0017
	+ 1 jets	2.688 ± 0.063	0.321 ± 0.022	0.0167 ± 0.0050
	+ 2 jets	5.535 ± 0.073	0.833 ± 0.029	0.0835 ± 0.0092
	+ 3 jets	8.35 ± 0.23	1.314 ± 0.094	0.108 ± 0.027
	+ 4 jets	9.92 ± 0.41	1.94 ± 0.19	0.206 ± 0.062

Table 4.30: Efficiency of the cuts up to 2, 3 or 4 b-tags for $t\bar{t}H$ and $t\bar{t}$ selected with the multijet trigger.

avoid events with very high jet multiplicity from background processes from passing our selection thanks to their cumulative probability of having a large number of fake b-tags. Fig. 4.5.1 shows that indeed, in $t\bar{t}H$ events, when three b-tags are required in the leading 8 jets only a small fraction of events get discarded: little is gained in acceptance by counting b-tags in additional soft jets.

The request of three or more b-tags sizably decreases the amount of signal, but is the one which yields the highest ratio between number of selected signal events and square root of the number of expected background events, which constitutes a reasonable estimator of the statistical significance. We will therefore enforce the presence of at least three b-tags in our final selection, although in the following section we study the signal kinematics before applying this restrictive cut, in order to retain sufficient statistics to model with reasonable accuracy the kinematical variables.

Pixel trigger				
sample	type	2 b-tags (%)	3 b-tags (%)	4 b-tags (%)
$t\bar{t}H$		6.432 ± 0.019	2.266 ± 0.012	0.4332 ± 0.0051
$t\bar{t}$	+ 0 jets	0.888 ± 0.038	0.077 ± 0.011	0.0100 ± 0.0041
	+ 1 jets	3.361 ± 0.070	0.414 ± 0.025	0.0242 ± 0.0061
	+ 2 jets	6.207 ± 0.077	0.919 ± 0.030	0.0907 ± 0.0096
	+ 3 jets	8.88 ± 0.23	1.388 ± 0.096	0.108 ± 0.027
	+ 4 jets	10.13 ± 0.41	2.02 ± 0.19	0.206 ± 0.062

Table 4.31: Efficiency of the cuts up to 2, 3 or 4 b-tags for $t\bar{t}H$ and $t\bar{t}$ selected with the pixel trigger.

Standard multijet trigger			
P_t bin	2 b-tag (%)	3 b-tags (%)	4 b-tags
30-50 GeV	0	0	0
50-80 GeV	0	0	0
80-120 GeV	0.00035 ± 0.00035	0	0
120-170 GeV	0.00208 ± 0.0015	0	0
170-230 GeV	0.021 ± 0.004582	0.001 ± 0.001	0
230-300 GeV	0.0833 ± 0.0090	0.0127 ± 0.0035	0
300-380 GeV	0.160 ± 0.012	0.0188 ± 0.0041	0.00089 ± 0.00089
380-incl GeV	0.315 ± 0.018	0.0275 ± 0.0052	0.0039 ± 0.0020

Table 4.32: Efficiency of the cuts up to 2, 3 or 4 b-tags for QCD selected with the multijet trigger.

Pixel trigger			
P_t bin	2 b-tag (%)	3 b-tags (%)	4 b-tags
30-50 GeV	0.00059 ±0.00059	0	0
50-80 GeV	0	0	0
80-120 GeV	0.00035 ± 0.00035	0	0
120-170 GeV	0.0042 ± 0.0021	0	0
170-230 GeV	0.0250 ± 0.0050	0.0020 ± 0.0014	0
230-300 GeV	0.0716 ± 0.0084	0.0078 ± 0.0028	0
300-380 GeV	0.144 ± 0.011	0.0161 ± 0.0038	0.00089 ± 0.00089
380-incl GeV	0.295 ± 0.017	0.0265 ± 0.0051	0.0039 ± 0.0020

Table 4.33: Efficiency of the cuts up to 2, 3 or 4 b-tags for QCD selected with the multijet trigger.

Standard multijet trigger				
sample	Pt bin (GeV)	2 b-tag (%)	3 b-tags (%)	4 b-tags
W + 0 jets		0	0	0
W + 1 jet	pt < 100 100 < pt < 300	0 0.0070 ± 0.0026	0 0	0 0
W + 2 jet	pt < 100 100 < pt < 300	0 0.0314 ± 0.0055	0 0.0048 ± 0.0021	0 0.00095+-0.00095
W + 3 jet	pt < 100 100 < pt < 300	0.0074 ± 0.0026 0.0707 ± 0.0091	0.00093 ± 0.00093 0.0116 ± 0.0037	0 0
W + 4 jet	pt < 100 100 < pt < 300	0.0157 ± 0.0043 0.123 ± 0.020	0.0012 ± 0.0012 0.0130 ± 0.0065	0 0.0032 ± 0.0032
W + 5 jet	pt < 100 100 < pt < 300	0.0169 ±0.0054 0.198 ±0.022	0 0.0119 ± 0.0053	0 0.0024 ± 0.0024

Table 4.34: Efficiency of the cuts up to 2, 3 or 4 b-tags for W+jets selected with the multijet trigger.

Pixel trigger				
sample	Pt bin (GeV)	2 b-tag (%)	3 b-tags (%)	4 b-tags
W + 0 jets		0	0	0
W + 1 jet	pt < 100 100 < pt < 300	0 0.0040 ± 0.0020	0 0	0 0
W + 2 jet	pt < 100 100 < pt < 300	0.0010 ± 0.0010 0.0190 ± 0.0042	0 0.0067 ±0.0025	0 0.00095 ± 0.00095
W + 3 jet	pt < 100 100 < pt < 300	0.0056 ± 0.0023 0.0591 ± 0.0083	0.00093±0.00093 0.0104 ± 0.0035	0 0
W + 4 jet	pt < 100 100 < pt < 300	0.0181 ± 0.0047 0.146 ± 0.022	0.0024 ± 0.0017 0.0162 ± 0.0073	0 0
W + 5 jet	pt < 100 100 < pt < 300	0.0169 ± 0.0054 0.229 ± 0.023	0 0.0143 ± 0.0059	0 0.0024 ±0.0024

Table 4.35: Efficiency of the cuts up to 2, 3 or 4 b-tags for W+jets selected with the pixel trigger.

Standard multijet trigger				
sample	Pt bin	2 b-tags (%)	3 b-tags (%)	4 b-tags (%)
Z→ $\nu\nu$	120-170 GeV 170-230 GeV	0.0100 ± 0.0058 10.76 ± 0.19	0 0	0 0

Table 4.36: Efficiency of the cuts up to 2, 3 or 4 b-tags for $Z \rightarrow \nu\nu$ selected with the multijet trigger.

Pixel trigger				
sample	Pt bin	2 b-tags (%)	3 b-tags (%)	4 b-tags (%)
Z→ $\nu\nu$	120-170 GeV 170-230 GeV	0.0167 ± 0.0075 0.0664 ± 0.016	0 0	0 0

Table 4.37: Efficiency of the cuts up to 2, 3 or 4 b-tags for $Z \rightarrow \nu\nu$ selected with the pixel trigger.

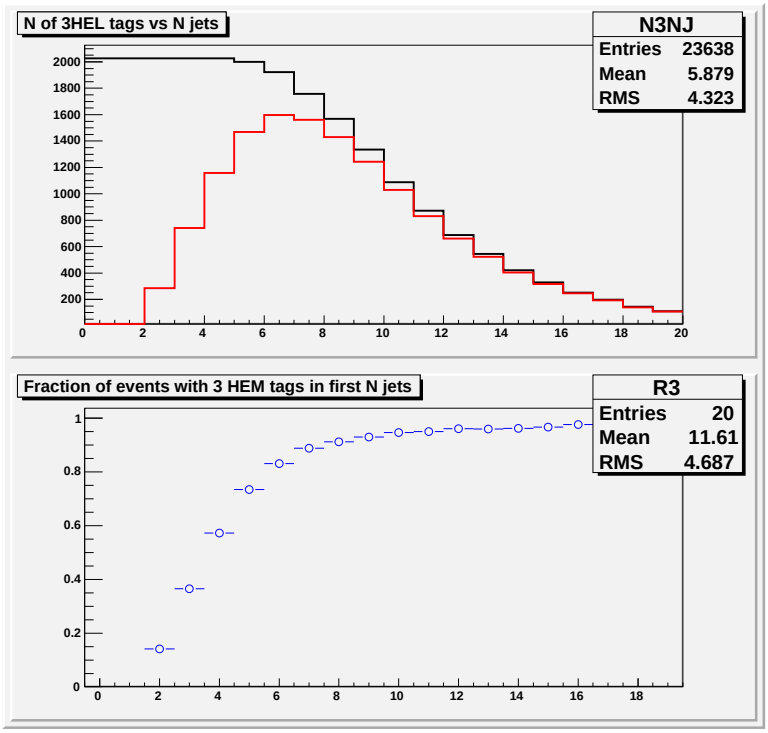


Figure 4.7: *Number of jets with 3 b-tags and fraction within the first N*

4.5.2 B-tag matrix parametrization

By a quick examination of tables **??** and **??** it is evident that we have a problem: the requirement of three or four b-tags so reduces our background samples that the rate estimates become rather imprecise, and we still need to study the kinematics of the surviving events, which would require a large statistics of the selected datasets. One good way to solve this problem is to rely on rate estimates based on the raw event counts, and to model the kinematics using the method described in Sec. 3.2.5, entailing a multiplication of events by modifying the jet energies according to their expected resolution functions. Here, however, we exploit another possibility. Inclusive QCD events and W+jet events have a b-tagging rate which is dominated by fakes. By modeling the b-tagging probability of generic jets as a function of their Et, rapidity and number of tracks, we can use several times the same events, weighting the b-tag configuration of the jets by the combined probability of their b-tag.

We select jets passing the standard cuts $E_t > 25GeV$, $|\eta| < 3.0$ in our QCD simulation, regardless of trigger accept or other selections, and con-

struct a probability matrix $P(E_t, \eta, N_{tr})$ depending on the corrected jet E_t, jet rapidity, and number of tracks with significance larger than 1.0. The latter choice is determined by the way the b-tag algorithm is constructed – track lists are made only for tracks passing that loose criterion.

$$P(E_t, |\eta|, N_{tr}) = N_{tag}(E_t, |\eta|, N_{tr})/N_{jet}(E_t, |\eta|, N_{tr}). \quad (4.2)$$

We construct the matrix with 8 bins in E_t: 25-50 GeV, 50-70 GeV, 70-90-110-150-200-300-500; 4 bins in $|\eta|$: 0-0.5, 0.5-1.0, 1.0-2.0, and 2.0-3.0; and 8 bins in the number of tracks: N_{tr} = 2, 3, 4, 5, 6, 7, 8, and 9 or more. With the large statistics of QCD Monte Carlo before any selection, the matrix has typical uncertainties of less than 1% in each bin. A picture of the variations of b-tagging probability as a function of one of the three variables in turn is shown in Fig. 4.5.2.

It is quite clear that such a rough construct cannot model all the subtleties in the rate of multiple b-tags in our background events: the correlation between b-tagging of different jets will spoil our rate estimates. What is worse, the very selection of events with significant missing Et does bias the tag rate, because of the larger probability that QCD events with heavy flavor create jets with neutrinos in them. However, we believe the matrix is good enough to allow a modeling of the event kinematics. A check of the resulting kinematics can in any case be performed with the multiplication strategy discussed in Ch.3.

To use multiple times simulated QCD events and W+jets events we operate as follows. In an event passing our trigger, $N(jet) > 5$, and missing $E_t > 3.0$ requirements, we take every jet and extract the probability P that the jet is tagged based on its Et, rapidity and number of tracks. We can then estimate a probability for every combination of tagged/untagged jets, as clarified in the following table.

P1,P2,P3,P4, and P5 are defined by the Et, eta, and number of tracks of each of the five jets. By combining the probabilities, we can use the same event $2^(max(N_j, 8))$ times, by adopting our prescription that we only consider b-tags in the leading 8 jets. Each combination is given a weight

$$W = \Sigma_{i=1,Ntags}[P_i(E_t, |\eta|, N_{tr})] \times \Sigma_{i=1,Nuntagged}[(1 - P_i(E_t, |\eta|, N_{tr}))] \quad (4.3)$$

in all the distributions of kinematical quantity we can compute after selecting events based on the number of tags –or, in this case, of "possible tags".

A special tretment is needed for all variables which are directly biased by the b-tag requirement of the jet. It is the case of the sum of jet tagging discriminants, which are computed using the significance of the second most

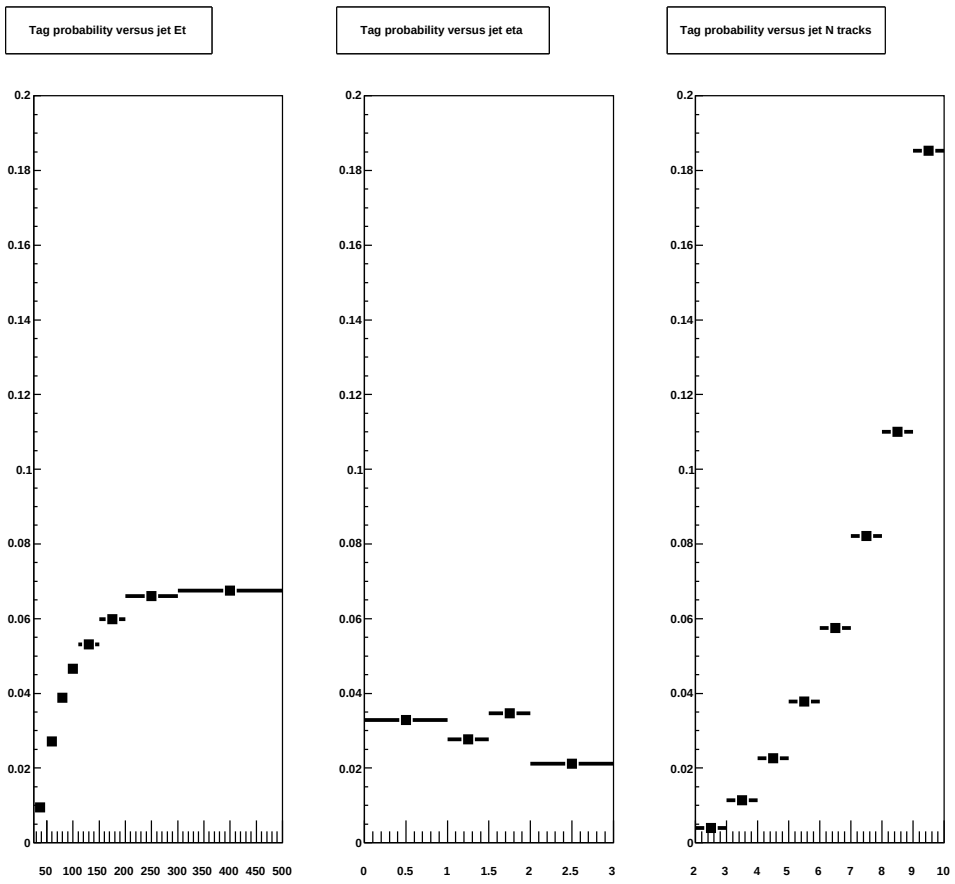


Figure 4.8: *Tag rate as a function of jet E_t (left), rapidity (middle), and number of tracks with significance larger than 1.0 (right)*

displaced track in each jet. Of course, a b-tagged jet has a large (and > 5.3) value of this quantity. To mimic correctly the sum of discriminants in tag matrix-weighted QCD and W+jets events, we assign to each jet which we wish to call a b-tag (with probability $P(E_t, |\eta|, N_{tr})$) a discriminant value taken randomly from a distribution extracted from real b-tagged jets having the same track multiplicity N_{tr}. This procedure effectively "unbiases" the resulting distribution of sum of discriminants.

The same procedure is performed to compute, for matrix-weighted QCD events and W+jets events, the track mass of each jet which is assigned a b-tag. Again, we pick random values distributed according to the track mass

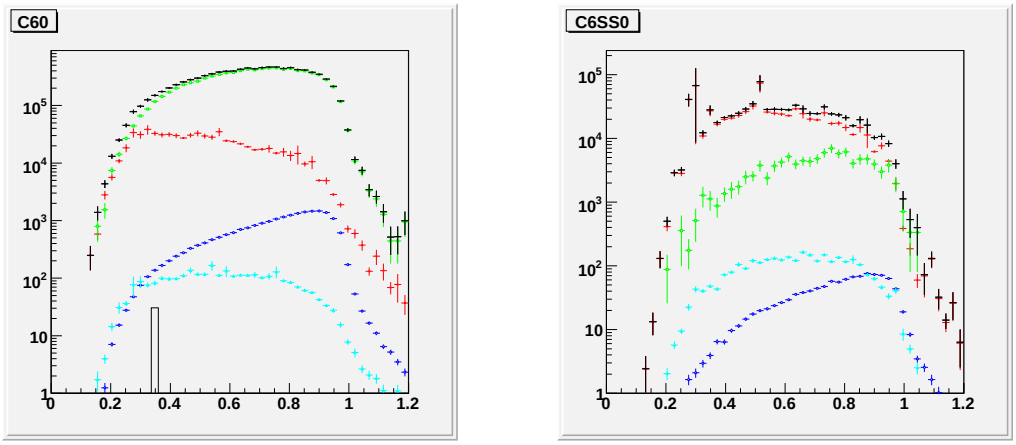
permutation	probability	number of b-tags
00000	(1-P1)(1-P2)(1-P3)(1-P4)(1-P5)	No tags
10000	P1(1-P2)(1-P3)(1-P4)(1-P5)	1 tag
...		1 tag
00001	(1-P1)(1-P2)(1-P3)(1-P4)P5	1 tag
11000	P1P2(1-P3)(1-P4)(1-P5)	2 tags
...		2 tags
00011	(1-P1)(1-P2)(1-P3)P4P5	2 tags
11100	P1P2P3(1-P4)(1-P5)	3 tags
...		3 tags
00111	(1-P1)(1-P2)P3P4P5	3 tags
11110	P1P2P3P4(1-P5)	4 tags
...		4 tags
01111	(1-P1)P2P3P4P5	4 tags
11111	P1P2P3P4P5	5 tags

Table 4.38: *Combinations of tags and probabilities*

distribution of b-tagged jets with the correct number of tracks, extracted from real b-tags in the QCD simulation. The resulting distribution is shown in Fig. 4.6.4. The comparison with the same shape extracted from a different set of QCD events shows that the procedure correctly extracts the expected distribution of sum of track masses. We can therefore reliably use this variable and the sum of discriminants in the definition of a global event discriminant to separate $t\bar{t}H$ events from all remaining backgrounds, as explained below. Figure 4.5.2 shows the effect of the requirement of two b-tags on the centrality computed on the first six jets in the Et sorted collection.

4.6 Study of signal kinematics

A full reconstruction of the kinematics of the $t\bar{t}H$ decay is impossible in the events targeted by our search, which implies one leptonic W decay: they not



(a) (b)

Figure 4.9: Centrality computed with the first six jets before (a) and after (b) the requirement of 2-btags.

only contain no information on the neutrino Pz component, but totally miss any information on the charged lepton. Nonetheless, it is important to try and find differences between signal and background based on what we know we have a means of reconstructing: the mass of the hadronically-decaying top, and the mass of the two b-jets associated to the Higgs boson.

4.6.1 Custom Jet corrections

Jet energies after standard jet corrections in the Fast Simulation we use show a non-negligible bias, which shifts all resonance masses upwards. This is the combined result of a incorrect tuning of the jet corrections, which do not appear to work well enough with jets produced by FastSim, and the large amount of minimum bias and pile-up events in high-luminosity simulations we are using.

Because of our wish to attempt a simple reconstruction of the mass of top quark and Higgs boson in our sample, we need custom jet energy corrections. We perform a quick study by associating final state partons from top and Higgs decay to the closest jet in eta-phi space, retaining only very clean matches ($\Delta R < 0.2$). We divide the sample into two sets: b-tagged jets and untagged jets, without checking whether the b-tag are indeed b-quarks: this is because we need to extract a function we can use "blindly", i.e. without knowing the exact nature of the jet.

The difference between parton Et and "standard" corrected jet Et shows indeed a downward slope when drawn as a function of corrected jet Et, for both b-tagged jets and non-tagged ones. We fit the profile distributions with

a sixth order polinomial, as shown in Fig. 4.10, and use those parameters to further correct jet energies.

The result of this procedure is shown in Fig. 4.6.1: the W, T, and H masses peak at values much closer to their true mass.

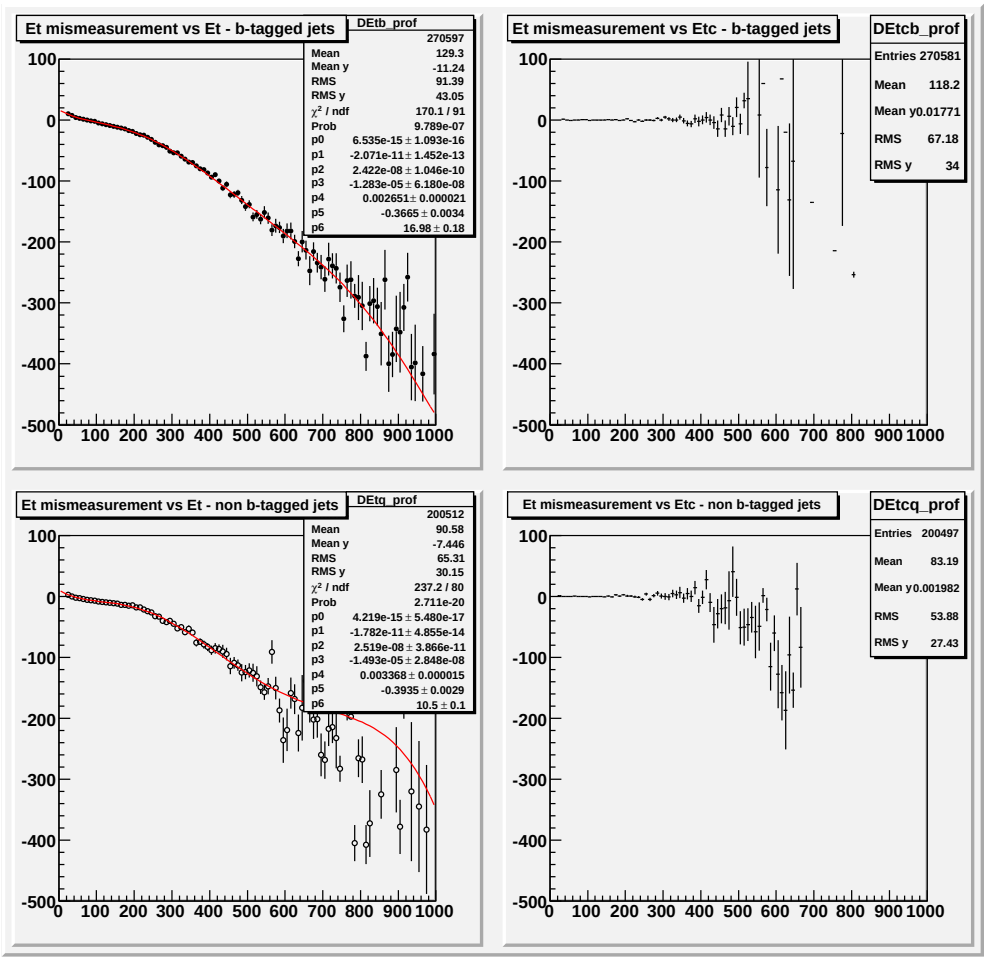


Figure 4.10: *Difference between parton and associated jet transverse energy as a function of jet E_t, for b-tagged jets (top) and non-tagged jets (bottom). The application of the correction functions extracted from the sixth-order polinomial fits on the left allow a better estimate of parton transverse energy, as shown in the corrected E_t plots on the right.*

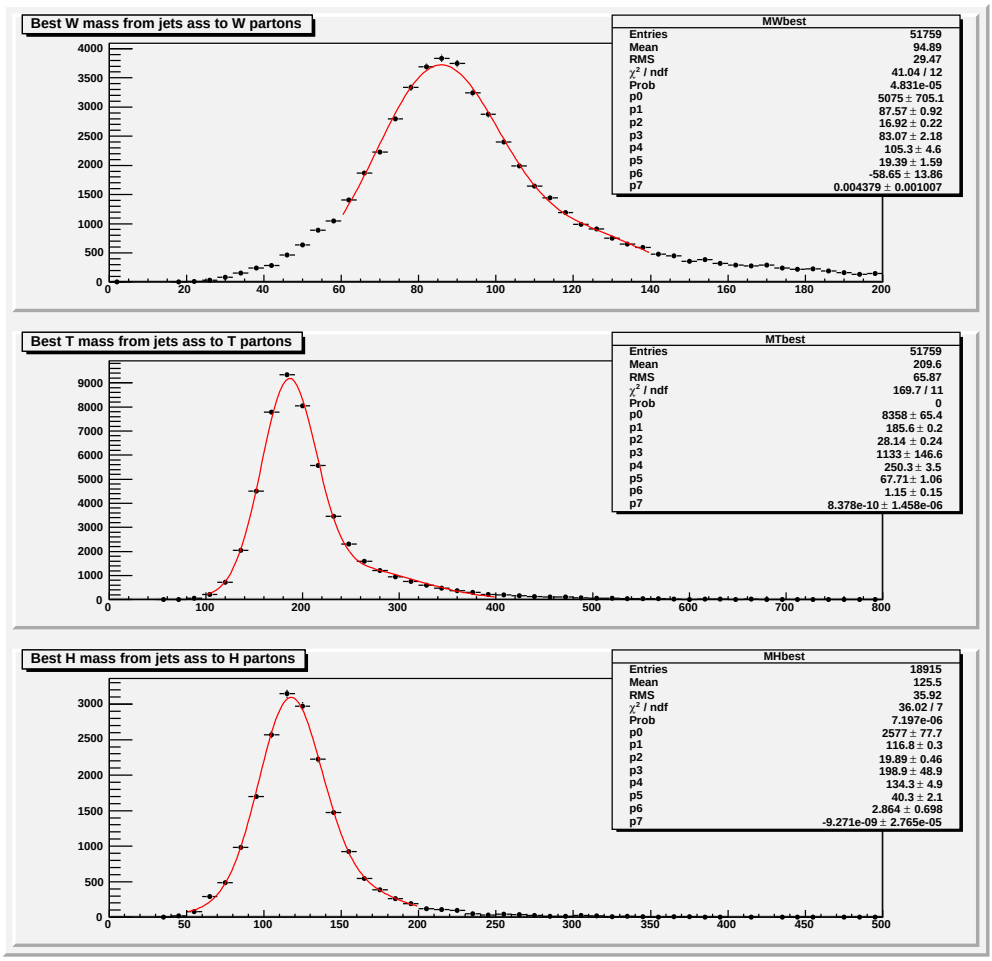


Figure 4.11: *Distribution of the invariant mass of jet pairs originated from the two W decay partons (top), the three jets from hadronic top decay (middle), and the two jets originated from the Higgs boson decay (bottom).*

4.6.2 Definition of the H→bb decay products

In the attempt at choosing the combination of two b-tagged jets which is most likely to have originated from H decay we studied the kinematics of the jets associated to the b-quarks from H. The association is performed by a simple algorithm matching to the first 8 jets in the Et-ordered collection the six partons from the single-lepton $t\bar{t}H$ final state by a minimum ΔR criterion.

We find that in the Et-ordered list of b-tagged jets, the two coming from Higgs decay are just as likely to be the first as they are to be the last ones

(see Fig. 4.6.2). Because of that, we do not include any information related to the b-jet Et in the definition of the best b-quark jet pair.

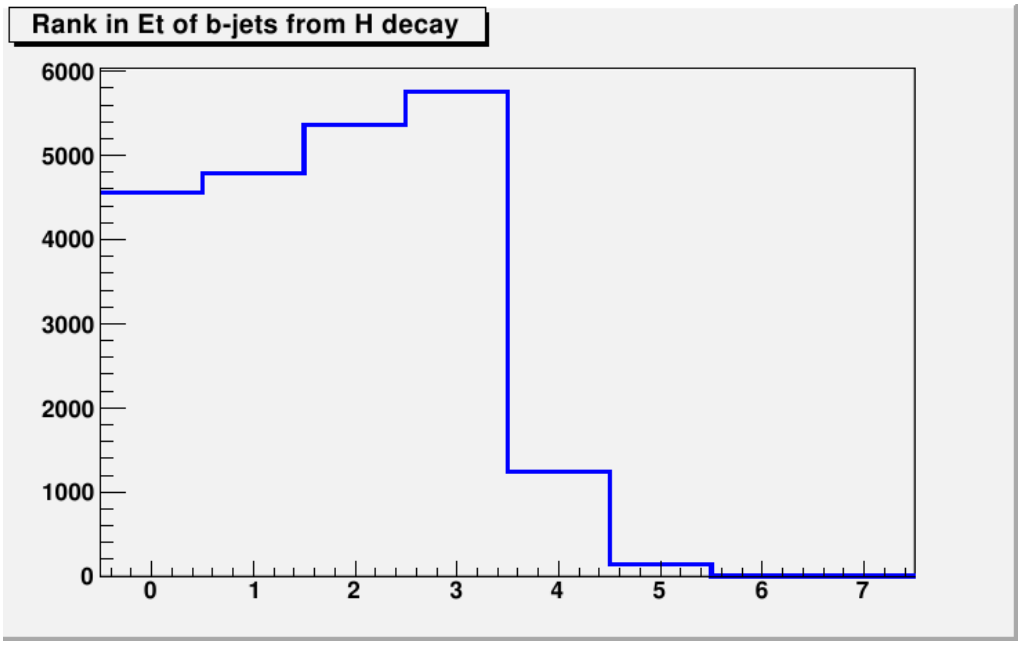


Figure 4.12: Ranking of the b-jets coming from the Higgs boson decay among the 8 jets with highest Et in the event, ordered in ascending Et (0 is the first).

Instead, we find a striking difference in the angle, in the $\eta - \phi$ plane, between the two jets associated to the two Higgs decay products, and minor differences in the Pt of the jet pair and their combined rapidity (see Fig. 4.6.2). We use those three quantities to define a probability ratio by filling two (10x10x10) matrices of the three quantities with the relative frequency of each $(\Delta R, P_t, \eta)$ combination: one, P_{right}, for the right jet pairs and one, P_{wrong}, for all incorrect combinations of jet pairs. The ratio

$$R(\Delta R, P_t, \eta) = P_{right}(\Delta R, P_t, \eta)/P_{wrong}(\Delta R, P_t, \eta) \tag{4.4}$$

can be used as a simple order parameter to select the most likely pair of tagged b-jets to use for the computation of the Higgs kinematics.

The mass of the pair picked according to the criterion described above is shown in Fig. 4.6.2, for $t\bar{t}H$ events and backgrounds as usual, at the level of selection where at least three jets contain a b-tag. One clearly sees that the criterion allows a reasonable fraction of events to contain a well-reconstructed Higgs peak. The distribution is not as narrow as the one obtained from the

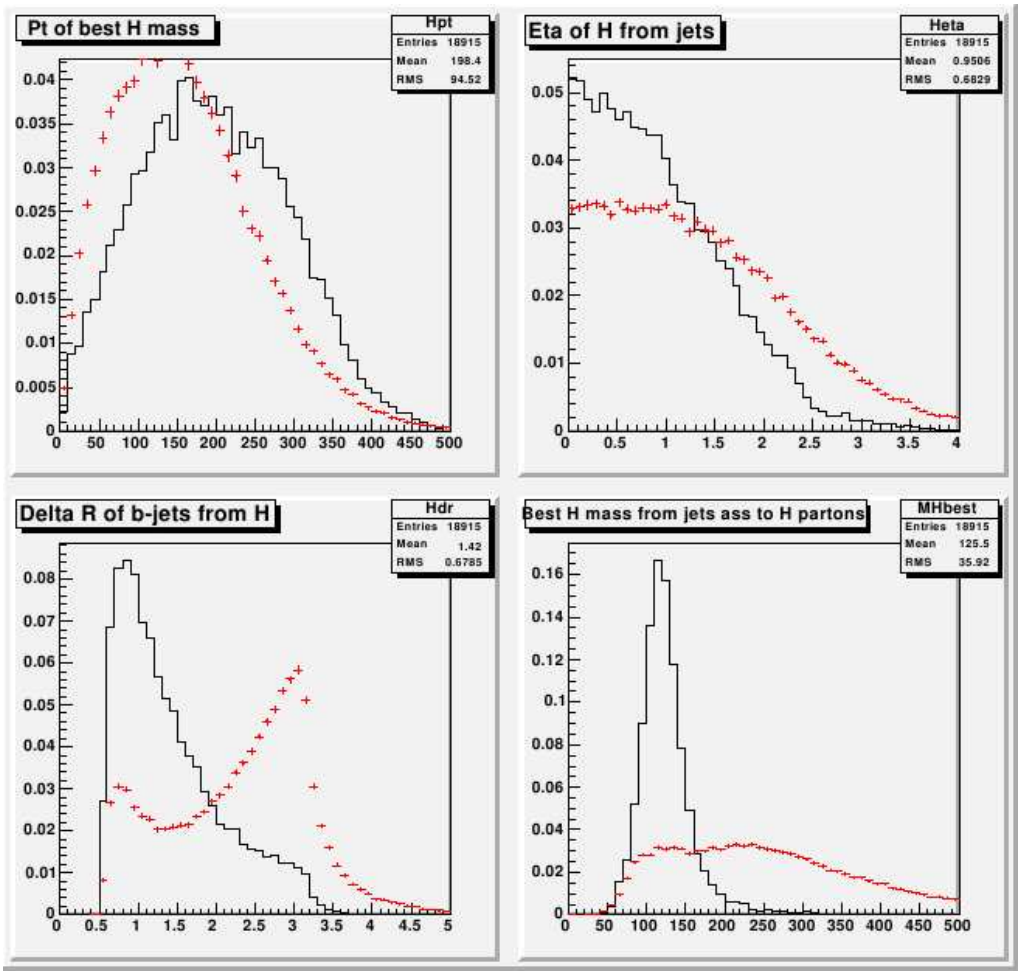


Figure 4.13: Distributions of quantities used to identify the pair of b-jets coming from the Higgs. In black are Higgs b-jets pairs and in red non-higgs b-jets pairs.

two jets associated to the true partons from H decay by a $\eta - \phi$ distance criterion, but it is peaked enough to allow some discrimination with background processes.

4.6.3 Definition of the hadronic top decay products

In a similar fashion, we study the reconstruction of the three jets coming from a hadronically-decaying top quark in our signal sample. In the list of jets selected by our cuts, once the two associated to the Higgs decay as described in the previous paragraph have been removed, we select the leading eight by

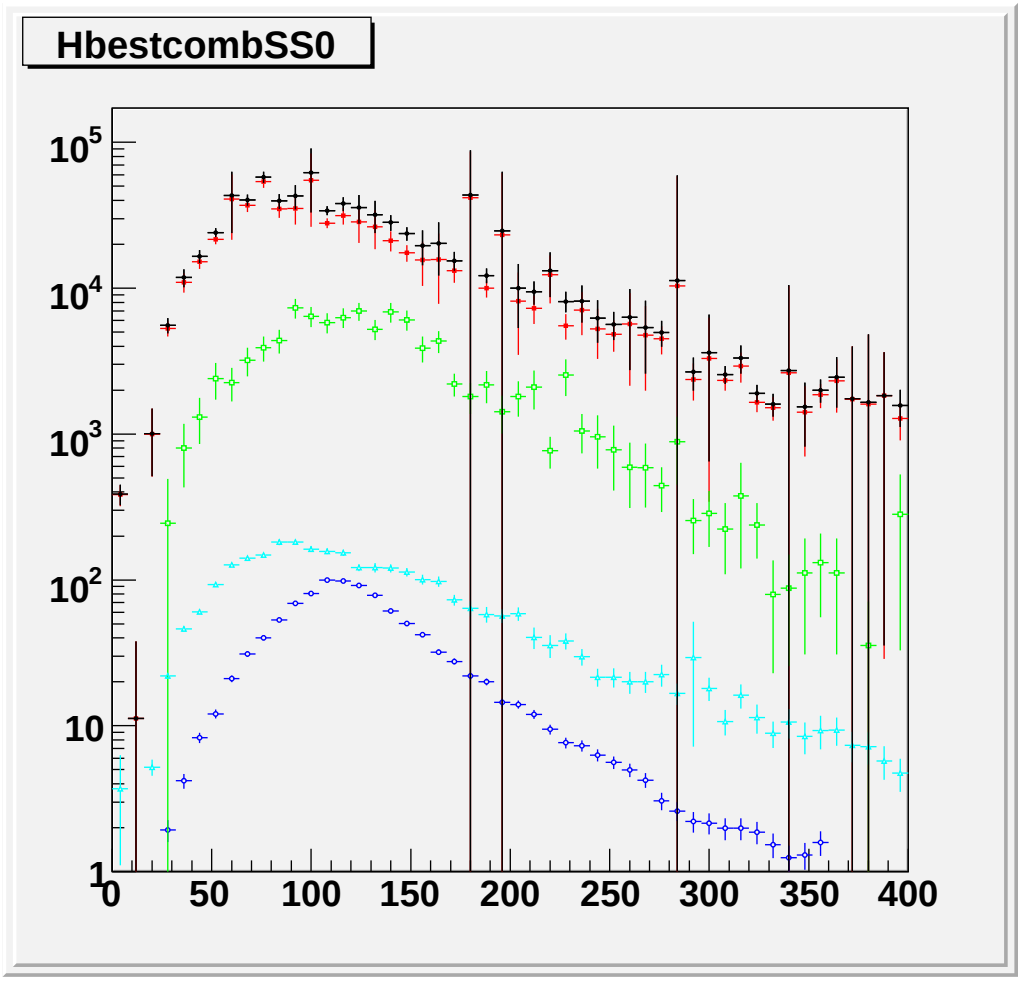


Figure 4.14: Mass of the best pair of b-jets selected with the criterion descibed in the text from events with at least three b-tags for: $t\bar{t}$H (blue), $t\bar{t}$ (green), W (cyan), QCD (red) and all processes (black).

transverse energy, and we search for triplets of jets, without for the moment enforcing the presence of one b-tag in one of them. In 5-jet events there is only one combination possible, while in 6-jet events there are four, in 7-jet events ten, and in 8-jet events twenty combinations. Among these many possible choices, we need to pick the one which is most likely to have originated from the hadronic top decay: this way, we will be able to formulate a hypothesis on the kinematics of the decay process, reconstructing angles between the higgs and the top quark, and computing other potentially discriminating variables.

We start by finding in our jet list the three jets truly originated from a hadronic top decay, in events with a single-lepton topology (exactly one $W \to jj$ decay). We use as always a $\eta-\phi$ association criterion, by looping on the Et-ordered jet list after removing the two previously found best matching the Higgs decay topology, and finding for each jet the closest parton. The reconstructed mass, combined momentum, rapidity and other kinematical characteristics are shown in Fig. 4.6.3 for the triplet matching the three partons from top decay. Overlaid to those distributions are shown the same quantities computed for triplets which have at least one jet not matching a top decay parton.

We find that the three most discriminating quantities are the top momentum, the projection of the same along the direction of the Higgs boson, and the angle between top and higgs. We thus construct again two (10x10x10) matrices with the relative frequencies of correct and wrong combinations, such that we can then define a ratio $R(P_t, \eta, Pr) = P_{right}(P_t, \eta, Pr)/P_{wrong}(P_t, \eta, Pr)$ for each triplet. We then have a prescription to guess which jets to pick as the most likely top decay products.

The mass distribution of the chosen triplets is shown in Fig. 4.6.3 (lower right plot). We note that it peaks at about **200** GeV, a value not too different from the most likely value for the matched triplets (186 GeV). This quantity discriminates well $t\bar{t}H$ signal events from background processes, as shown in Fig. 4.6.3.

4.6.4 Variables describing the kinematics of the event

Besides trying to spot the jets originated from the hadronic top decay –when one exists- and the Higgs decay to a pair of tagged b-jets, there is a large set of kinematical quantities we can study to seek differences with background processes. In this section rather than describing the process by which we came to study a particular set of variables, we barely list them one by one, discussing their particularities in turn.

1. Of course, one of the most important quantities at a hadron collider

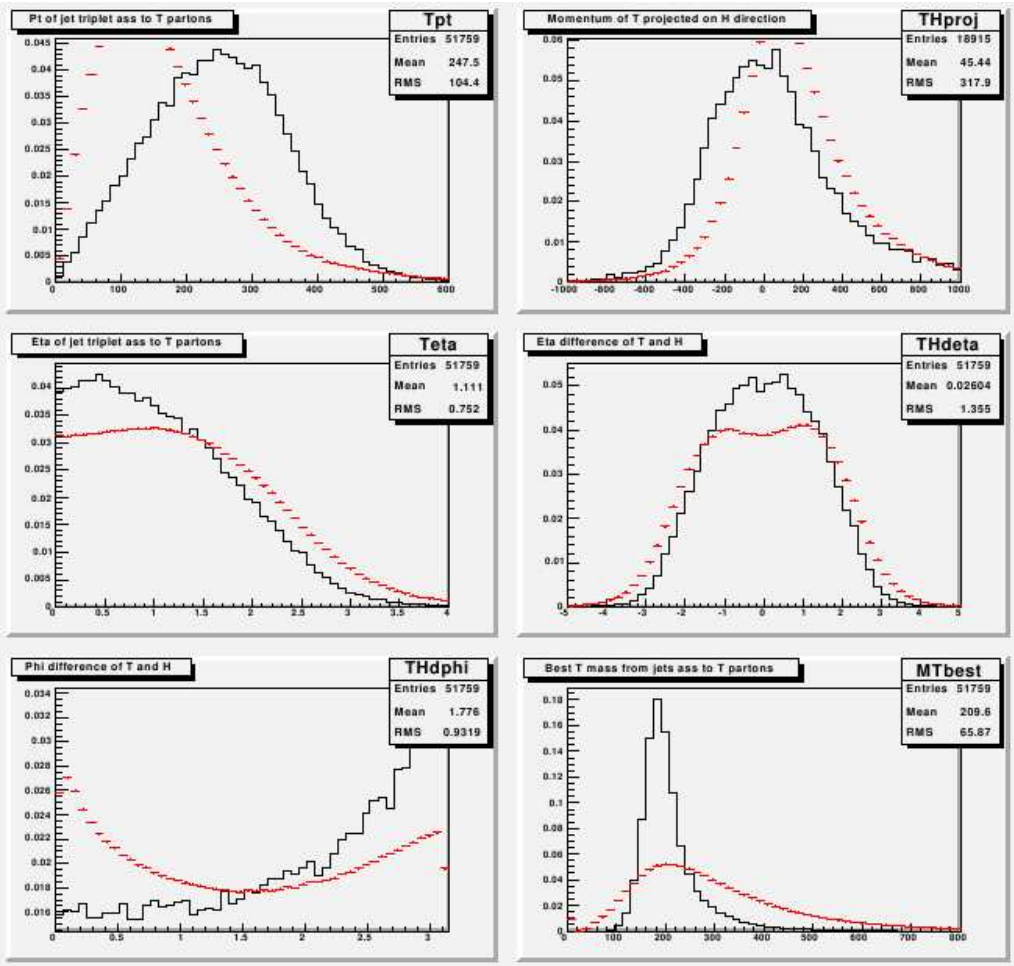


Figure 4.15: Distributions of various kinematical quantities for the triplet of jets associated to top decay partons (red) and triplets where at least one of the jets is not associated to a parton from the top.

is the total sum of transverse energies of the identified jets. This variable is directly connected to the magnitude of the acceleration of final state partons in the hard subprocess. Different production processes yield different distributions of ΣE_t, reflecting the peculiarities in the Q^2 dependence of the production. We expect the ΣE_t to discriminate $t\bar{t}H$ production from softer processes, but of course the requirement of a large number of jets makes all backgrounds – and particularly inclusive QCD – similar to the signal in this variable. Nonetheless, it is an important ingredient in any complete recipe to picture a multijet final state.

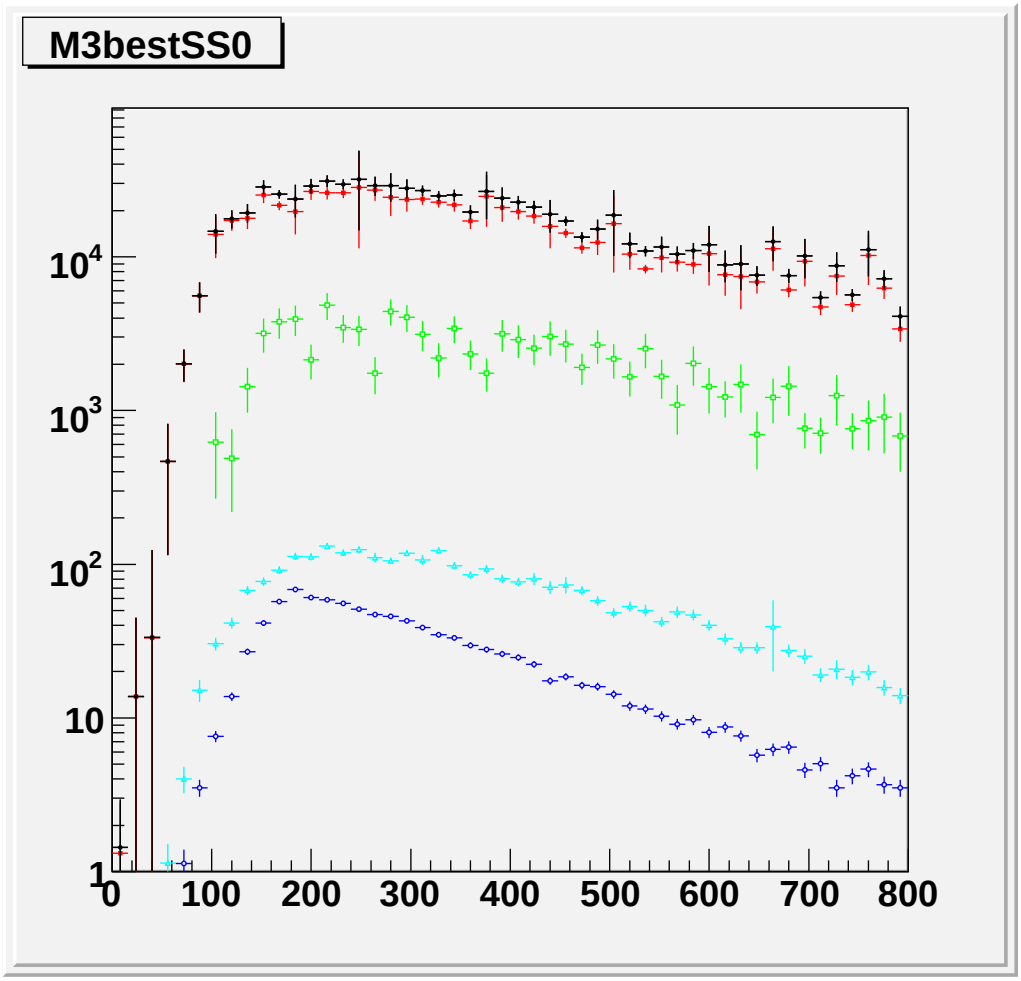


Figure 4.16: Distribution of the reconstructed top mass from the best jet triplet after removing the jets assigned to the Higgs for: $t\bar{t}$H (blue), $t\bar{t}$ (green), W+jets (cyan), QCD (red), all processes (black)

2. The missing transverse energy is tightly linked to one of the variables on which we base our preselection, the missing E_t significance. After a cut $\sigma(ME_t) > 3.0$ the missing E_t is not much different in background processes and signal, although the latter does exhibit a longer tail, because it is rich in real high-Et neutrinos, while QCD backgrounds do not usually yield any.

3. The sum of ΣE_t and missing E_t is called H_t. In the presence of charged leptons, one would add the Et or Pt of these to the H_t too, but we retain an inclusive character in our search and neglect electron and muon identification. Being the sum of two other variables, we do not expect to learn much by a study of the H_t.

4. The missing E_t significance is used in the preselection of the data, and shows longer tails for events with real neutrinos.

5. The angle in azimuth (i.e., in the plane transverse to the beam) between missing E_t direction and leading jet direction is usually capable of discriminating real and fake missing E_t: when a jet transverse energy is grossly overestimated, the missing E_t will point opposite to it. On the other hand, most QCD processes have a distinctly back-to-back topology even after radiation off initial and final state enrich the event with additional jets. Because of that, one often finds the missing E_t pointing along the leading jet direction rather than opposite to it: this happens not infrequently, when a multi-parton emission happens to recoil against a leading parton. We thus understand the double peaking structure in this variable.

6. Another interesting way to discriminate fake from real missing E_t is to find the angle between missing E_t and the second jet in the E_t-ordered list. In this case, one expect to mostly find a small angle in events like those discussed at the end of the previous description.

7. In events with only four or five jets CDF showed that the minimum azimuthal difference between missing E_t and any jet direction was the one capable to best identify the jet responsible for the fake missing E_t: by fluctuating low, a jet causes missing transverse energy to point along its direction. Unfortunately, that variable is by far less discriminant in CMS in events of the kind we are studying, which possess not just four or five, but as many as 8 jets or more. The information carried by the minimum azimuthal angle decreases sharply with the number of jets, because the phase space – the support $[0 - \pi]$ – is limited and the

resolution in missing E_t azimuth is not very good. We are thus not surprised to find virtually no discrimination power in this quantity.

8. The single-lepton $t\bar{t}H$ process nominally yields six energetic partons in the final state. The transverse energy of the sixth jet is thus a good discriminant of a hard parton emission if compared to processes where the sixth jet is mostly due to subleading radiation. We in fact relied in this quantity when we defined a very loose $E_t > 25GeV$ cut in the jet definition: the choice then was dictated by our wish to maximize our chances of correctly describing the $t\bar{t}H$ decay topology, but we kept our option of cutting harder when maximizing the signal significance.

9. Using the six highest-E_t jets (or five, if a sixth jet does not exist) one can compute a variable describing the centrality of the event. It is a powerful discriminant between heavy particle production and QCD radiation processes, constructed by taking the ratio between the E_t sum of the six jets and their six-body invariant mass. QCD background events exhibit a flatter distribution in jet rapidity, with the result of a smaller $\sum E_t$ for a given event Q^2. The event Q^2 is roughly estimated by the six-body mass, and the ratio turns out to be one of the most discriminating variables at our disposal.

10. We add to the list the mass of the six jets used in the denominator, for a deeper understanding of the event kinematics

11. The same definition of centrality can be used to compute it with eight jets (or fewer, if the event does not have enough of them).

12. The mass of the eight jets is also added to the list for the same reason mentioned above.

13. Having defined the jets most likely belonging to the hadronic top decay with the recipe summarized in 4.6.3, we compute the three-jet mass, which does show a more peaked structure at about the right mass, although a long tail is present.

14. In the triplet of jets assigned to the hadronic top, we select the two jets with no b-tag, and compute the most likely W boson mass.

15. Using the two b-tagged jets assigned to the H decay with the procedure discussed in 4.6.2, we compute their mass and find a broad peak around the correct Higgs mass value for the signal, and a broader distribution for backgrounds.

16. The total mass of the five jets assigned to hadronic top and Higgs boson is another useful quantity to study.

17. The remaining jets not assigned to top and Higgs can be used to compute their total mass.

18. If the triplet from top and the doublet from Higgs are defined, one can compute the projection of the top three-momentum on the direction of flight of the Higgs boson. This variable is an interesting complement to our other variables.

19. The rapidity difference between top and higgs reconstructed flight directions complements the former variable in describing the kinematics of these particles, for the cases when the assignment is correct.

20. With the mass reconstructed for Higgs, top and W boson using the best guess for a jet-parton assignment, one can compute a chi-squared as the sum of squares of the deviation of reconstructed and true particle mass, each weighted with their typical resolution.

21. For events with four b-tags it is possible to compute the mass of the two b-tagged jets not assigned to the Higgs decay. In the signal they most likely come from the decay of the two top quarks, while in background processes they may be the result of gluon splitting.

22. The angle between the two b-tagged jets not assigned to the Higgs decay is also useful, in events with four jets.

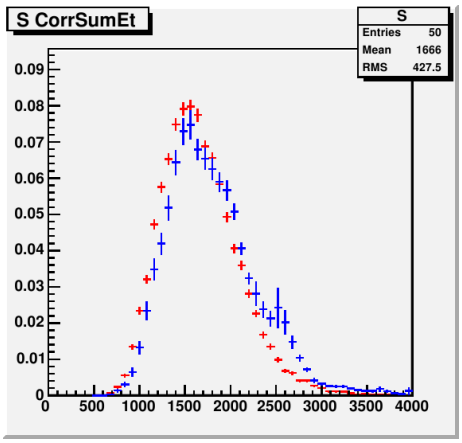
23. The b-tagging information used in our study is digital: a jet is either tagged or untagged. However, that information comes from a real number: the significance S of the impact parameter of the second-most-displaced track. This is larger in real b-quark jets than in light-quark jets, and even after the selection ($S > 4.3$), b-tagged jets from real b-quarks have a higher average of S. By summing the significances of the four highest-S jets one can thus create a discriminator capable of separating events where at least one jet is not originated from real b-quarks and events where all b-tags are due to heavy flavor. The importance of this variable is also due to its potential discrimination of top-antitop production, which is much harder to separate from $t\bar{t}H$ signal than are all other background processes.

24. The same sum discussed at the previous point can be made with the six largest-S jets. The rationale of this choice is of allowing charm-jets

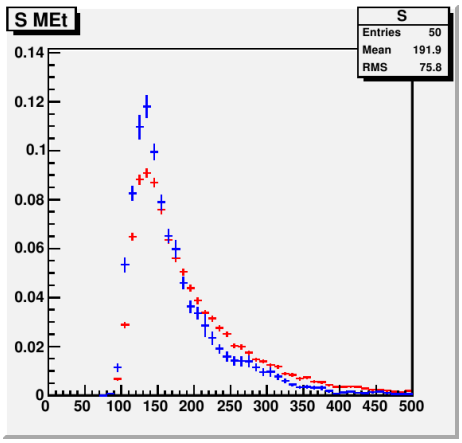
occasionally present in W decays to also contribute to the discrimination; some additional contribution may also come from jets originated from tau lepton decay.

25. Finally, one additional analogic information which is usually capable of discriminating heavy flavor b-tags from fake b-tags is the invariant mass of all tracks with significant impact parameter. By summing the invariant mass of tracks in all b-tagged jets we obtain a further discriminator of events with many b-quarks and events with fake b-tags.

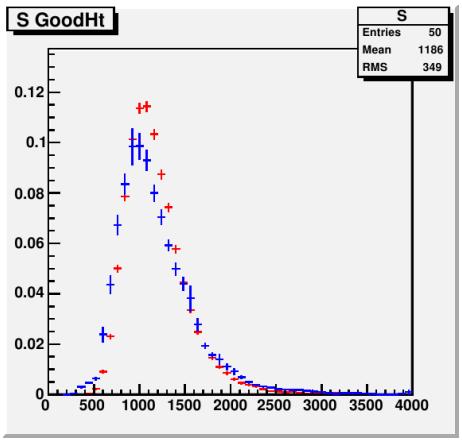
Figs. 4.6.4 to 4.6.4 show the distribution of the 25 kinematical quantities described above.



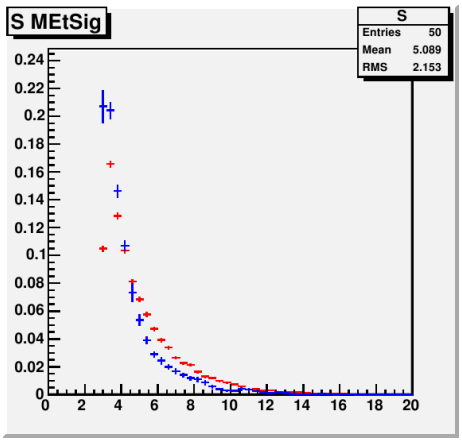
(a)



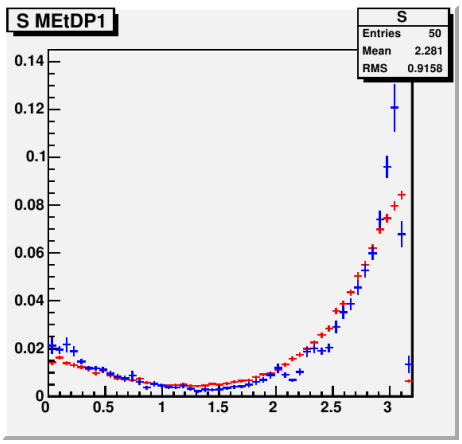
(b)



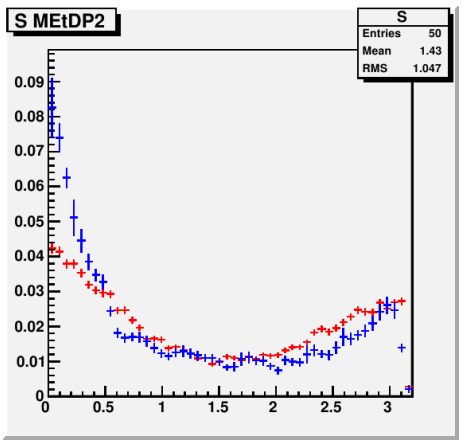
(c)



(d)

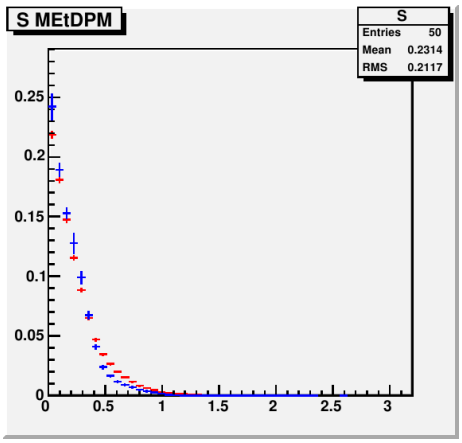


(e)

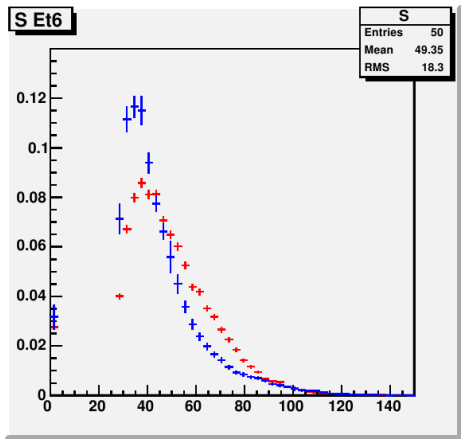


(f)

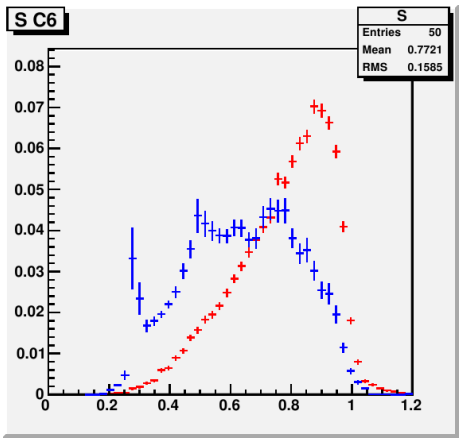
Figure 4.17: Distributions of $\sum$Et (a), missing Et (b), H_t (c), missing Et significance (d), angle between the missing Et and the first jet (e), second jet (f), ordered in Et. The signal is in red, the background in blue.



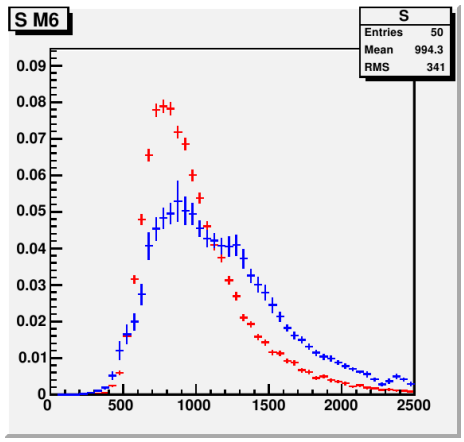
(a)



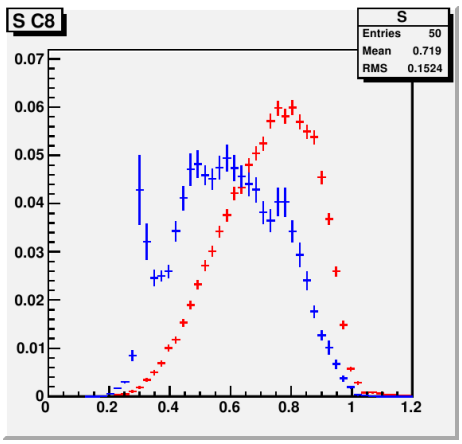
(b)



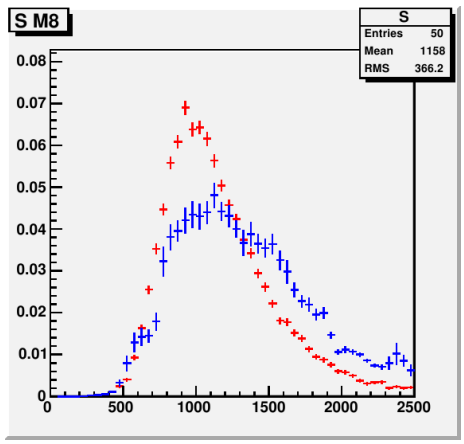
(c)



(d)

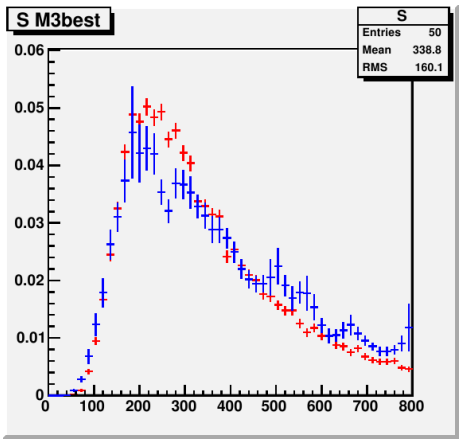


(e)

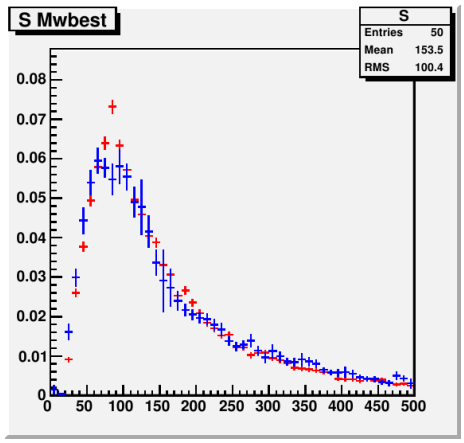


(f)

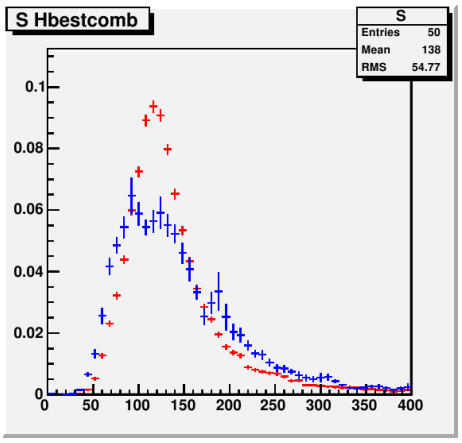
Figure 4.18: Distributions of angle between the missing Et and the closest jet (a), Et of the sixth jet (b), centrality of the first 6 jets (c), mass of the first 6 jets (d), centrality of the first 8 jets (e), mass of the first 8 jets (f). The signal is in red, the background in blue.



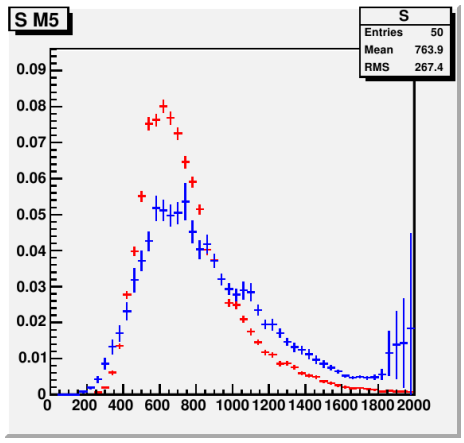
(a)



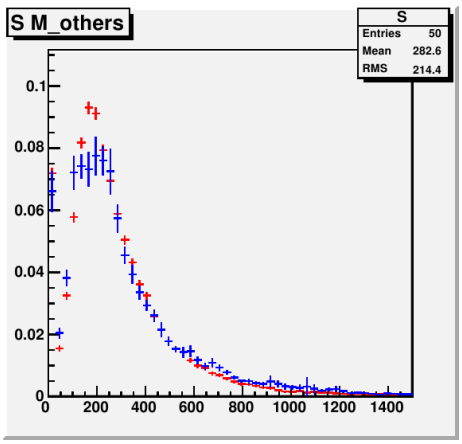
(b)



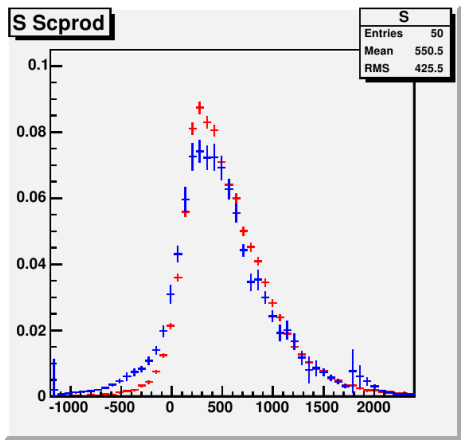
(c)



(d)

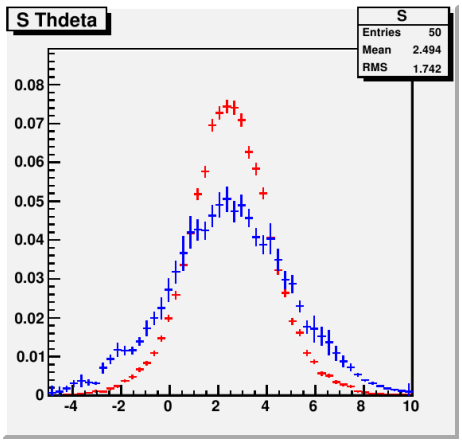


(e)

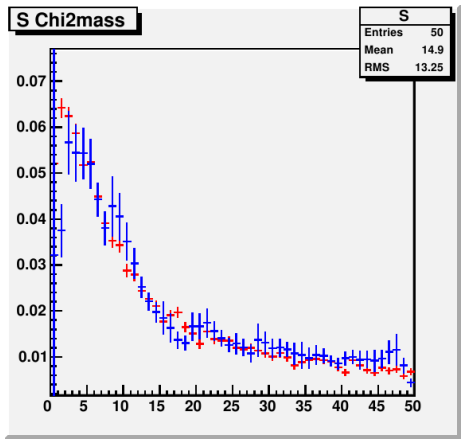


(f)

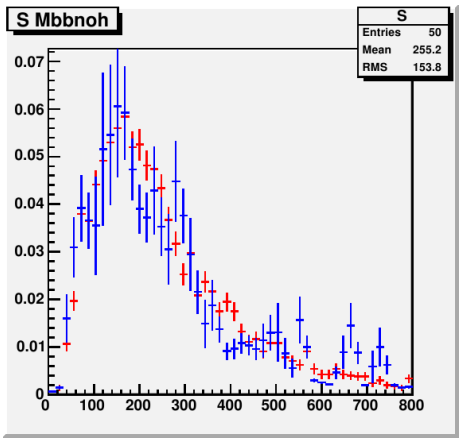
Figure 4.19: Distributions of top mass from the best triplet of jets (a), mass of the W from the best top jet pair without b-tags (b), Higgs mass from the best pair of b-tagged jets (c), Total mass of jets assigned to hadronic top and Higgs (d), Mass of the not assigned jets (e), projection of the three-momentum of the top in the direction of flight of the Higgs (f). The signal is in red, the background in blue.



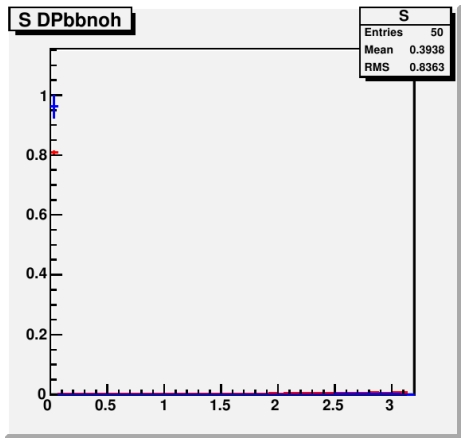
(a)



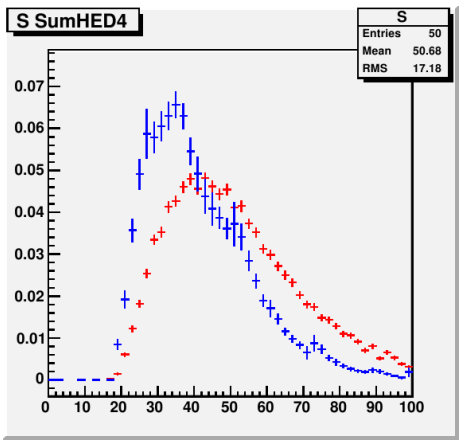
(b)



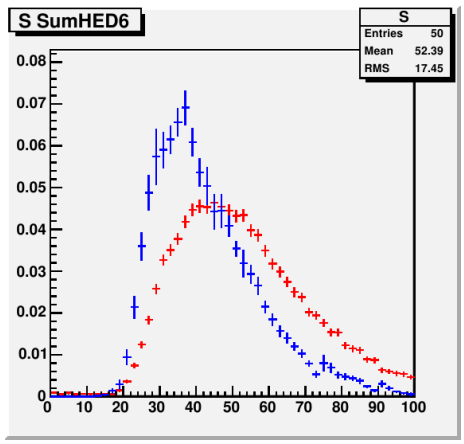
(c)



(d)



(e)



(f)

Figure 4.20: Distributions of rapidity difference between top and Higgs reconstructed flight directions (a), sum of squares of the deviation of reconstructed and true particle mass (b), mass of the two not assigned b-tagged jets (c), angle between the two b-tagged jets not assigned to the Higgs (d), sum of the significance of the impact parameter of the first 4 jets (e) and of the first 6 jets (f). The signal is in red, the background in blue.

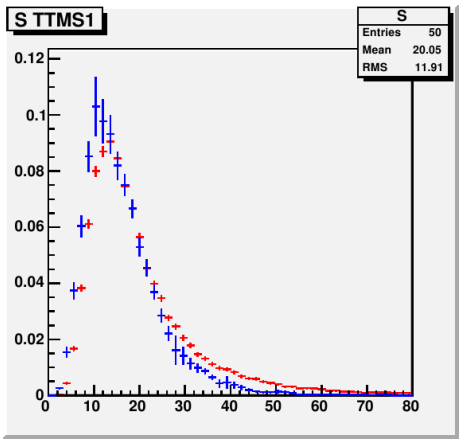


Figure 4.21: Sum of the invariant mass of tracks in all b-tagged jets.

4.7 Signal extraction

After the choice of a selection based on significant missing energy and at least three b-tags, the background samples suffer from the low statistics: in particular that is the case of the top quark samples. In order to select the variables which are most sensitive to differences between background and signal we first sum backgrounds together using their respective normalizations, and then smooth the resulting distributions to avoid single bins with a large uncertainties from spoiling the understanding of the variable under study.

The smoothing procedure used is not the one provided by the ROOT program, which we find wanting in several of its features. Our smoothing function works as follows: 1- histogram bins are ordered by increasing value of the relative uncertainty $S = \sigma_x/x$ in the bin contents. Outlier bins are chosen as those with the 20% worst uncertainties. 2- For each bin with a large relative error, a triplet of adjacent bins is chosen using the two neighbors on the sides 3- The triplet is fit with a linear interpolation, and the value of the central bin is substituted by the fit result at the same abscissa if there is a discrepancy of more than two standard deviations between bin value and fit value, using as uncertainty in the difference the one provided by the fit.

The simple prescription described above allows us to smooth background shapes quite effectively, as is shown in Figs. 4.22 to 4.7. What we get rid of is what we effectively need to level down: bins with a large content and a large error, due to the few events surviving trigger requirements in the simulations of background processes with small Q^2 (QCD) or small number of jets (W+jets samples). Of course, the normalization of the original histogram is not affected by the procedure, and it remains the reference value for our

background studies.

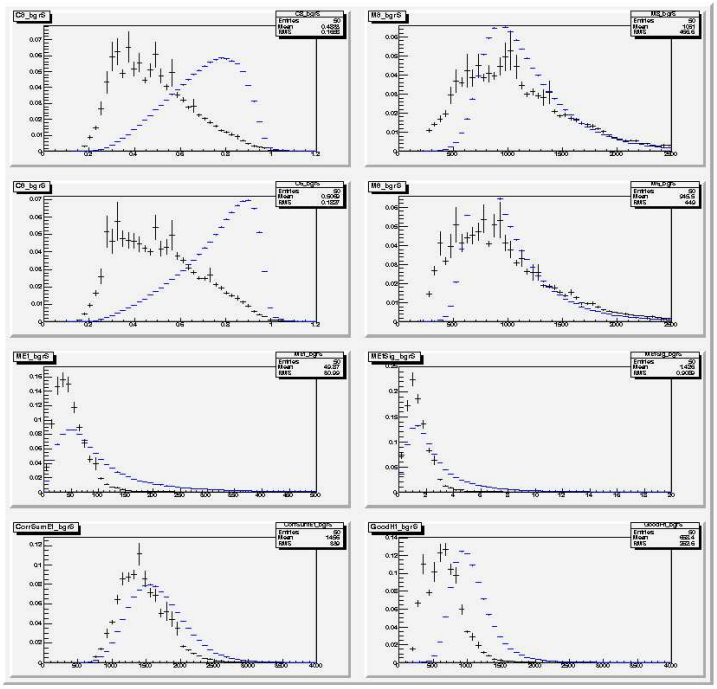


Figure 4.22: Smoothed histograms: signal is blue and background is black

A global estimator of the likelihood that an event is signal can be constructed with the most discriminating variables at our disposal, by taking the ratio of the signal p.d.f. and background p.d.f. for each variable and then combining them as follows:

$$K = \Sigma_i log(TTH(x_i)/BGR(x_i)) \tag{4.5}$$

By using K we use the full information from the distributions, rather than the belonging of the x-values to one or the other side of a fixed cut. Of course, since we are ignoring correlations among the variables in the computation of K, its discriminating power is strongly affected by the choice of variables [x]. In particular, a mixture of very discriminating and poorly discriminating quantities is not expected to fare as well as a subset of the more discriminating ones, because the power of K is "washed out" by distributions which only show minor differences between signal and background.

To help us choose a reasonable set $[x]$ among the 25 variables shown in Figs.**??** we compute the individual logarithms, by taking the N ratios $k_i = TTH(x_i)/BGR(x_i)$ between the p.d.f. TTH and BGR of each variable

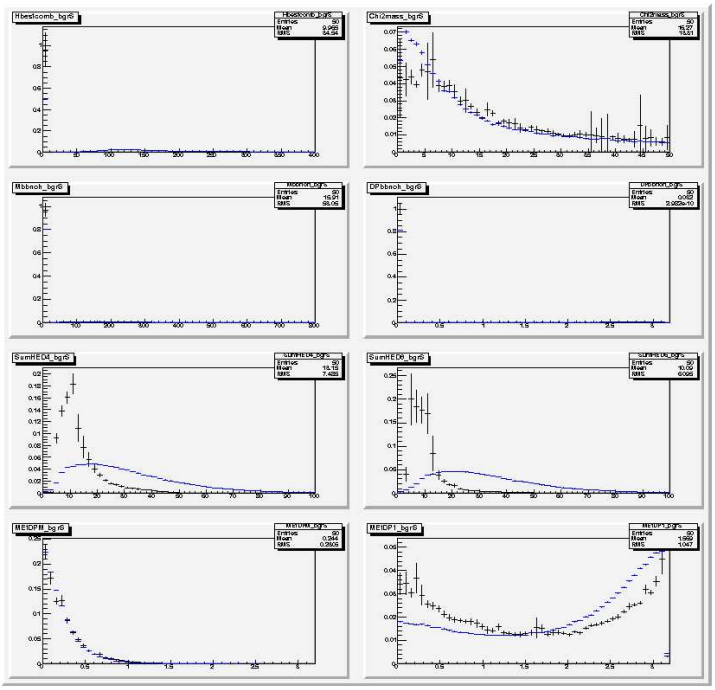


Figure 4.23: Smoothed histograms: signal is blue and background is black

in turn, and plotting it for signal events and background events. Since we chose N=50 bins in our distributions, we end up with 50 possible values of the ratio of logarithms. Signal events will have a larger value of k_i than background events, and by scanning the value of k_i one can compute the gain

$$G(t) = \int_t^\infty TTH(x_i) / \sqrt{\int_t^\infty BGR(x_i)} \tag{4.6}$$

The maximum value of G gives us a direct measure of the discriminating power of x in our recipe for K.

Table 4.7 lists the variables we used and the maximum values of G found for each. To construct K we then pick the ones with highest values of G, with some additional ad-hoc choice which results by a trial-and-error iteration. Asterisks in table 4.7 identify the variables chosen for the final computation of K.

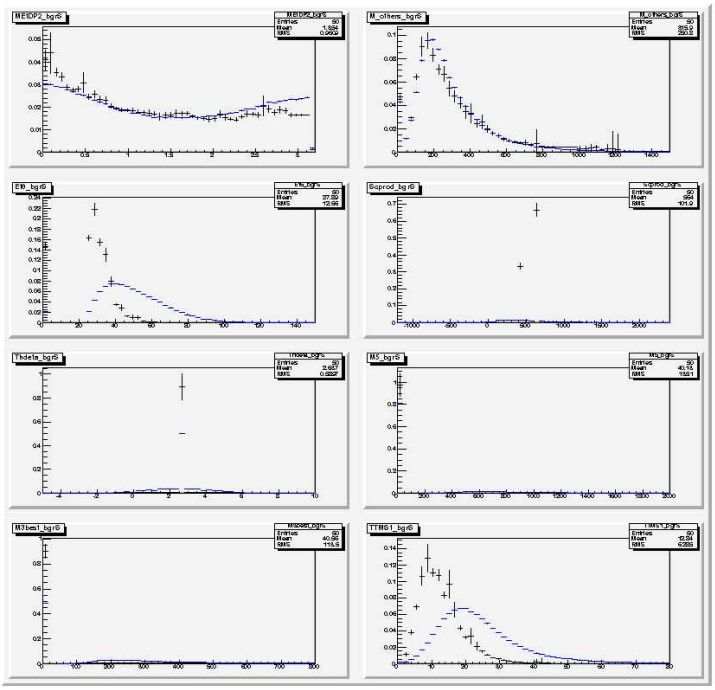


Figure 4.24: Smoothed histograms: signal is blue and background is black

Variable	maximum quality factor	likelihood ratio
1 * C8	1.28306+-0.01968	> 2
2 * M8	1.06595+-0.00884779	> 1.6
3 * C6	1.31172+-0.019505	> 2
4 - M6	1.09726+-0.0127666	> 1.8
5 - MEt	1+-0.1	> 0
6 * MEtSig	1.01299+-0.00837954	> 1.6
7 - CorrSumEt	1.02112+-0.00516084	> 1.6
8 * GoodHt	1.02403+-0.00377464	> 1.6
9 * Hbestcomb	1.05797+-0.0108383	> 1.8
10 * Chi2mass	1+-0.1	> 0
11 - Mbbnoh	1.01309+-0.0046658	> 1.6
12 - DPbbnoh	1.00422+-0.00263224	> 1.4
13 * SumHED4	1.07625+-0.00742688	> 1.4
14 * SumHED6	1.09889+-0.0169409	> 1.8
15 - MEtDPM	1+-0.1	> 0
16 - MEtDP1	1+-0.1	> 0
17 * MEtDP2	1+-0.1	> 0
18 * M_others	1.00101+-0.000971664	> 1.6
19 * Et6	1.0005+-0.000640972	> 1.6
20 * Scprod	1.02729+-0.00517903	> 1.6
21 * Thdeta	1.11361+-0.00912291	> 1.6
22 * M5	1.11609+-0.011883	> 1.8
23 * M3best	1.00846+-0.00495874	> 1.6
24 * Mwbest	1.0018+-0.000425412	> 0.4
25 * TTMS1	1.01955+-0.00387547	> 1.2

Table 4.39: Maximum quality factors for the 25 selected variables computed as explained in the text.

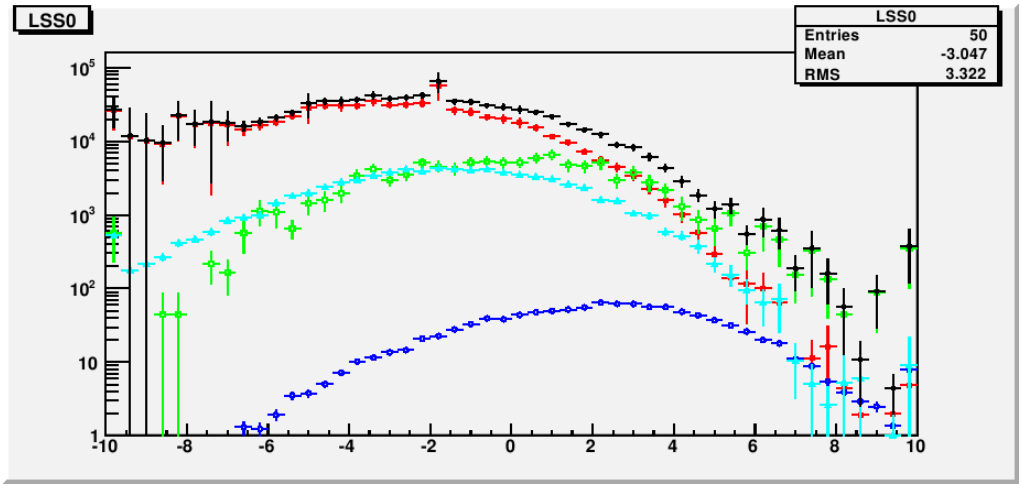


Figure 4.25: Distribution of the dicriminant G for $t\bar{t}$H (blue), $t\bar{t}$ (green), W (cyan), QCD (red) and all processes (black).

4.8 Study of signal significance

Fig. 4.8 shows the shape of the discriminant G for signal events with three b-tags, five or more jets, and a missing E_t significance $S > 3.0$. The total number of events expected from signal in 100/fb of data amounts to $N = 1076.6 \pm 6.6$. The sum of backgrounds is instead 260000+-500 . It has been computed by considering inclusive top pair production, single top production, inclusive jet production, and W+jets production; we do not include the negligible process Z+jets, for which we have no suitable simulation of the full range of transverse momentum.

Despite the use of the tag matrix approach to determine the shape of all kinematic quantities for QCD and W+jets backgrounds, the normalization of these processes is used by considering the actual number of events passing the selection. Because of that, these processes have a sizable uncertainty in their normalization. Top pair production also carries a sizable uncertainty, because of the lack of enough statistics of the Monte Carlo samples. Background uncertainty in a real analysis is going to be limited by the availability of suitable control samples more than it is in the study described here, but the relative contribution of the three main processes may indeed constitute a sizable systematic uncertainty. We return to this problem below.

In order to extract the signal from a distribution of G one can rely on a simple two-component fit. We use pseudo-experiments to obtain a distribution of fitted number of events in 10000 pseudo-data distribution obtained by choosing randomly 260000 background events from the background template,

and 1000 signal events from the signal template. These numbers are fluctuated by their respective poisson uncertainty for each pseudo-experiment. The resulting pseudo-data distributions of G are fit with the same templates used for the generation, obtaining a measure of the fit bias ($N_{fit} - N_{input}$), the relative bias (or pull) $(N_{fit} - N_{input})/\sigma_{N_{fit}}$, and the significance of the fitted signal $S = N_{fit}/\sigma_{N_{fit}}$. The distribution of pulls allows to verify the average bias and the width of the pull. The former has to be close to zero, while the latter is close to one if errors are computed accurately in the fit procedure. We observe a negative bias in the mean, while the sigma is consistent with unity.

The distribution of signal significance is shown for a few different values of the relative normalizations of the two main components of the background, i.e. QCD and top pair production. We take nine situations by considering -1 sigma and +1 sigma variations in the normalizations of these two processes, and extract from nine different sets of pseudoexperiments the distributions shown in Fig.xx. We observe that the significance of the $t\bar{t}$H signal can be estimated to be roughly equal to three, with a wide distribution. If one assumes that the relative normalization of QCD and $t\bar{t}$ is not known with precision in the data, then signal extraction becomes troublesome, as shown by a further pseudoexperiment where the background template used for the fit is not the one used in the pseudo data generation, but is obtained by varying by + or - one sigma the top content. Table 4.40 summarizes the results for the different classes of plseudoexperiments. It can be seen that the significance obtained assuming 100fb^{-1} is more than 3 σ, with the worst case (when both QCD and $t\bar{t}$ backgrounds fluctuate of $+1\sigma$) of $(3.187 \pm 0.031)\sigma$. Figure 4.8 shows the significance distributions for the 9 classes of pseudoexperiments.

We provide in Fig. 4.8 the same distributions of signal significance discussed above for events collected by the trigger using pixel information and table 4.41 provides a summary of the results. The significance is on average larger by about 30%: the availability of pixels at trigger level 1 corresponds roughly to an equivalent luminosity increase of about 50%.

4.9 Concluding remarks

This Chapter describes a first attempt at isolating $t\bar{t}H$ events from the huge background by relying mostly on the presence of a large missing E_t, indicating the escape of a energetic neutrino and thus, most likely, the leptonic decay of a W boson. The flavor and kinematics of the associated multijet system provide further handles for the task. In the analysis presented here we have

fluctuations	pull mean	pull sigma	significance
default	-0.133 ± 0.033	0.970 ± 0.027	3.282 ± 0.031
QCD+1σ	-0.173 ± 0.033	0.990 ± 0.028	3.226 ± 0.031
$t\bar{t}$+jets+1σ	-0.091 ± 0.033	0.968 ± 0.025	3.262 ± 0.031
QCD-1σ	-0.047 ± 0.036	1.016 ± 0.035	3.317 ± 0.033
$t\bar{t}$+jets-1σ	-0.109 ± 0.031	0.943 ± 0.025	3.323 ± 0.031
QCD+1σ, $t\bar{t}$+jets+1σ	-0.189 ± 0.033	0.964 ± 0.026	3.305 ± 0.031
QCD-1σ, $t\bar{t}$+jets-1σ	-0.173 ± 0.033	0.963 ± 0.028	3.187 ± 0.031
QCD+1σ, $t\bar{t}$+jets-1σ	-0.220 ± 0.033	0.983 ± 0.027	3.207 ± 0.031
QCD-1σ, $t\bar{t}$+jets+1σ	-0.124 ± 0.031	0.919 ± 0.025	3.221 ± 0.031

Table 4.40: Nine cases of fluctuations in the QCD and/or the $t\bar{t}$+jets samples selected with the multijet trigger.

chosen to completely disregard any information, positive or negative, on the presence of charged leptons, to retain an inclusive character in the search; of course, the explicit selection of events with electrons or muons is bound to increase the signal significance, and on the other hand an explicit veto on events with these signals –needed for a really orthogonal result, which is then much easier to combine to leptonic channel searches– would decrease it. An estimate of the significance which could be obtained in those cases is not necessary, however, since all we intended to prove was that the final state we have selected has some potential in the search of a light Higgs boson, despite the hindrance of huge backgrounds.

It is clear that the analysis we have carried out can be improved in many ways. We list here a few, to indicate the additional potential of a refined search.

1. First of all, as we noted in Chapter 3, we could not include in our datasets the events passing the missing E_t plus jet trigger, which operates at Level 1 with a predicted bandwidth of 0.1 kHz in high-luminosity running conditions, with thresholds at 100 GeV on Missing Et and 80 GeV on the leading jet's transverse energy. Our simulation indicates a much higher rate than the one present in the high-luminosity trigger table, and we decided to disregard it. The increase in signal efficiency is not very large if one includes the data accepted by such a trigger, but of course it would provide some gain in the expected significance of the search.

2. Our choice of a track-counting b-tagging algorithm was dictated by

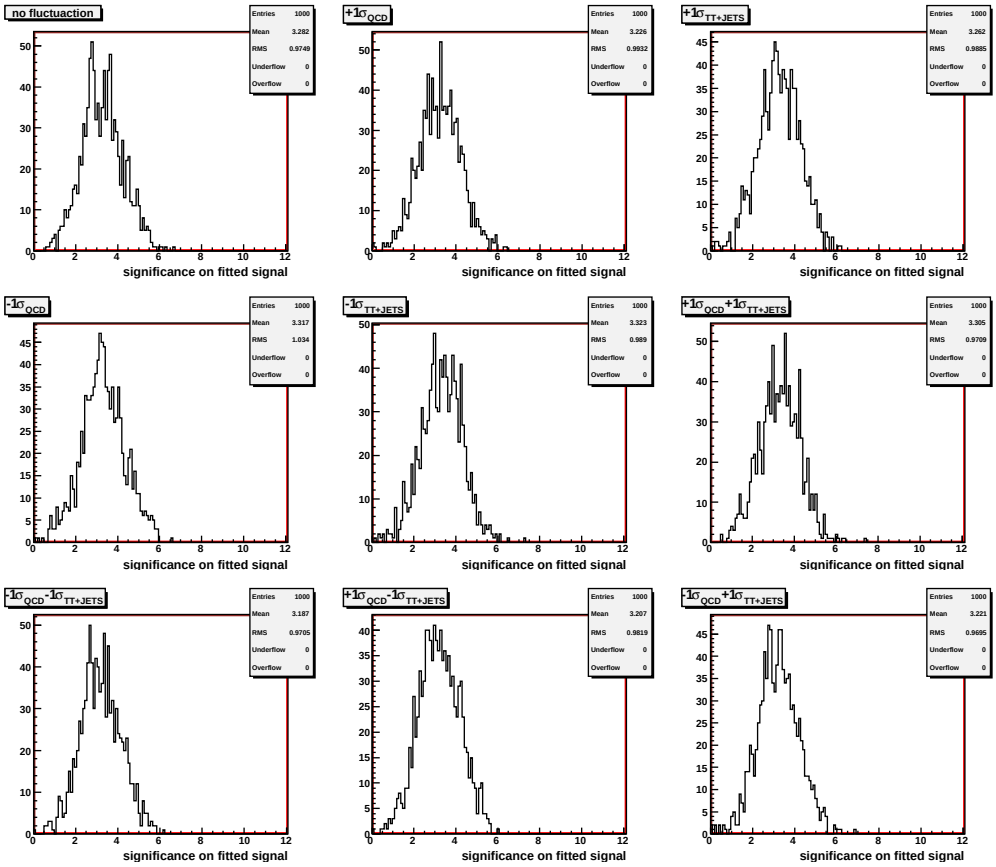


Figure 4.26: Significance distributions for the pseudoexperiments for events selected by the multijet trigger.

simplicity, and by the fact that the algorithm has been studied in some detail recently, and its output on fast-simulated monte carlo events appears in excellent agreement with that of a full simulation of the CMS detector. Nevertheless, it is certain that a more accurate study of an optimized b-tagger, tuned for events with several b-quark jets, could dramatically increase the performance of data selection, given that any b-tag efficiency increase factor $\alpha > 1$ would to first order result in a event efficiency increase of α^3. Furthermore, past experience with similar searches in CDF and D0 has shown that a combination of different b-tagging algorithms may win a lot in added discrimination power, and b-tagging algorithms based on neural networks have already started to produce improved results, as in light Higgs boson searches[61]. We also note that the use of a b-tagging algorithm of higher efficiency and/or purity is of benefit also for the kinematical selection, because a better

fluctuations	pull mean	pull sigma	significance
default	-0.143 ± 0.034	1.009 ± 0.028	4.290 ± 0.031
QCD+1σ	-0.161 ± 0.033	0.967 ± 0.026	4.295 ± 0.031
$t\bar{t}$+jets+1σ	-0.139 ± 0.032	0.954 ± 0.026	4.247 ± 0.031
QCD-1σ	-0.099 ± 0.033	0.983 ± 0.027	4.341 ± 0.032
$t\bar{t}$+jets-1σ	-0.134 ± 0.032	0.952 ± 0.025	4.365 ± 0.030
QCD+1σ, $t\bar{t}$+jets+1σ	-0.110 ± 0.033	0.975 ± 0.027	4.421 ± 0.031
QCD-1σ, $t\bar{t}$+jets-1σ	-0.196 ± 0.034	1.012 ± 0.027	4.208 ± 0.031
QCD+1σ, $t\bar{t}$+jets-1σ	-0.181 ± 0.033	0.969 ± 0.027	4.328 ± 0.031
QCD-1σ, $t\bar{t}$+jets+1σ	-0.197 ± 0.032	0.966 ± 0.026	4.201 ± 0.030

Table 4.41: Nine cases of fluctuations in the QCD and/or the $t\bar{t}$+jets samples selected with the pixel trigger.

definition of the kinematics of the final state is possible if jets are more often and more correctly assigned to the originating partons.

3. A feature we totally disregarded in our analysis was the identification and removal of jets originated by primary vertices separated from the one producing the multijet final state. This is a potentially very useful handle in high-luminosity events, as indeed is shown by the reduction in rate provided by the pixel jet requirements discussed in Chapter 3. If a search for $t\bar{t}H$ events is performed with the multijet trigger in CMS, without any track information for the level 1 decision, then an offline veto on jets pointing to vertices separated in z from the leading one is a very good idea.

4. Although we considered a very large number of kinematical variables in the attempt at finding a discrimination of signal and backgrounds –the ones we were able to discuss in this Chapter are only a subset of those we studied–, the possibility that other strong discriminant are found by a more careful investigation does exist. Furthermore, despite the lack of an identified charged lepton makes the single-lepton $t\bar{t}H$ final state very hard to reconstruct with a kinematic fit, the progress in technologies using unconstrained final states to extract the top mass from dileptonic datasets achieved by CDF and D0 suggest that this might still be a promising avenue for a higher separation of signal and backgrounds.

5. The use of an advanced algorithm to increase jet energy resolution promises to provide additional discrimination between signal and back-

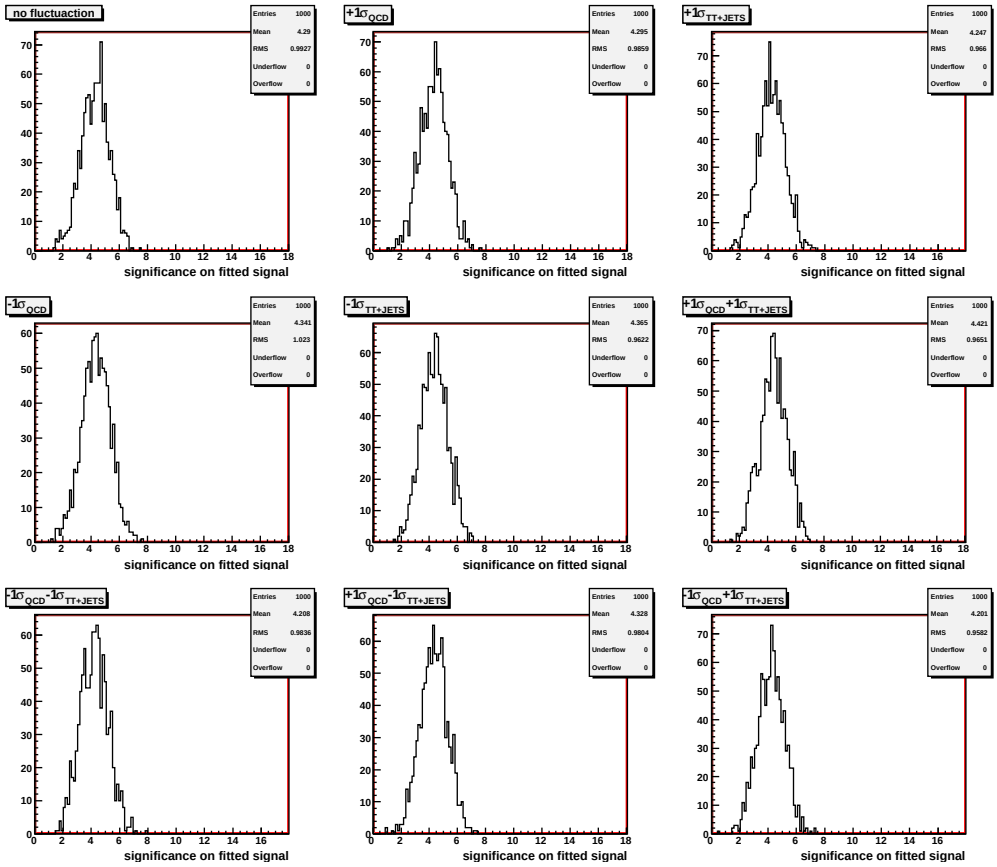
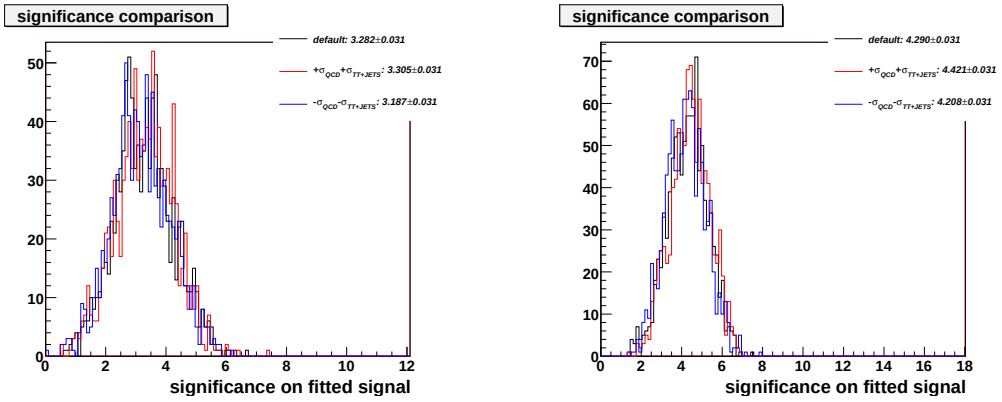


Figure 4.27: Significance distributions for the pseudoexperiments for events selected by the pixel trigger.

ground in all the kinematical variables connected to the reconstruction of the decay kinematics, and in particular the masses of Higgs and hadronically-decaying top quark.

6. Finally, we must not forget that our simplified signal extraction technique, based on the construction of a single discriminant G, is definitely sub-optimal. A multi-dimensional approach, such as a neural network classification or a multi-dimensiona likelihood, are certain to gain a sizable increase in sensitivity.

We are not in the position of being able to quantify meaningfully the possible gains in sensitivity that each of the above improvements may bring to the search. However, past experience with low-mass Higgs boson searches has shown that the gain offered by several of the above mentioned ingredients –in particular, items 2,5, and 6– is quite real. We therefore venture to conclude



(a) (b)

Figure 4.28: Comparison between the pseudoexperiments: for the standard, signal and backgrounds are extracted according to the default distributions. In red and blue the backgrounds are fluctuated lower and higher by one σ respectively. The left histogram (a) is for events taken with the multijet trigger and on the right histogram (b) is for the pixel-multijet trigger.

that 100 inverse femtobarns could indeed allow CMS to observe and measure, with four to five units of significance, the associated $t\bar{t}H$ production signal with multijet data. The fact that the analysis of the other possible final states of $t\bar{t}H$ decay can be added to the one considered here, providing a further large increase in significance, only demonstrates that associated production of top and Higgs with $H \rightarrow b\bar{b}$ decay has been a bit too hastily removed by CMS and Atlas from their menu of discovery channels.

Chapter 5

Conclusions and outlook

In this brief concluding chapter we provide the conclusions of the present study, and describe the possible developments of a further investigation of the matter.

5.1 Using the pixel detector at trigger level 1

In this work we have considered the usage of pixel detector information in the reconstruction of hadronic jets in events collected by the CMS detector under high luminosity running conditions, to add a small piece of information to best answer the question of the improvements possible on the physics output of CMS in case of an upgrade of its triggering capabilities.

It is clear from the outset that an effort to instrument the tracker with the necessary hardware to enable the processing of silicon information at trigger level 1 cannot be justified by a mere reconstruction of hadronic jets: if a justification exists, it certainly lays in the reduction of fake rates of muon triggers, which saturate the level-1 bandwidth if instantaneous luminosity exceeds $10^{34} cm^{-2} s^{-1}$ and no improvement is made of the muon P_T resolution (see Fig. 5.1). A secondary justification of pixel information in trigger level 1 decision is provided by the measurement of track impact parameter, which could enable the collection of datasets with a sizable amount of heavy flavor jets. Jet reconstruction, of course, is an ancillary bonus.

The work we have carried out was indeed motivated by the need of answering the question of what additional benefits could be gained with a pixel trigger. The question is not totally academic, as we have tried to demonstrate by presenting the reader with the analysis of multijet data containing large missing transverse energy, a sample which could provide meaningful

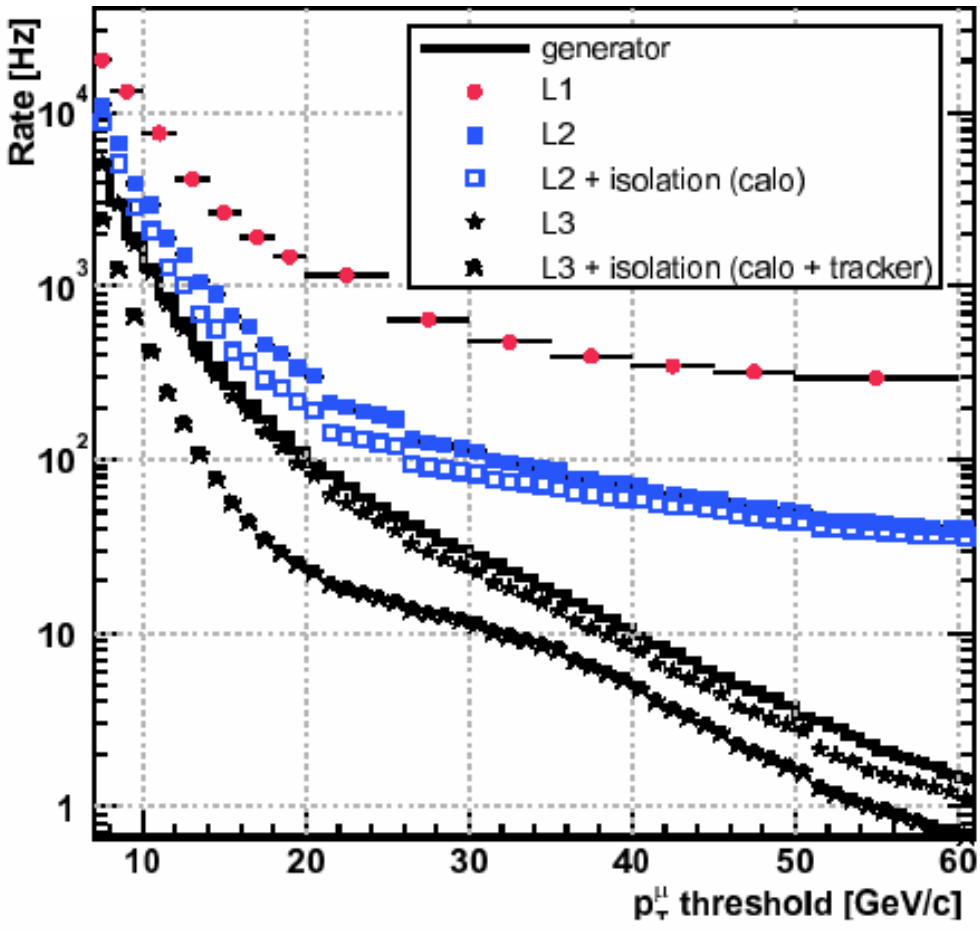


Figure 5.1: First level trigger and HLT single muon trigger rates as a function of the Pt threshold at LHC. The rates are shown separately for level 1 and HLT (level2 and level 3), with and without isolation applied at levels 2 and 3. The rate generated in the simulation is also shown.

information on the associated production process $t\bar{t}H$ with any decays where charged leptons did not provide a clear signature.

In order to understand the full potential of pixel information for a level 1 decision we developed a jet reconstruction algorithm based solely on pixel tracks. The reconstruction of jets of course suffers from the absence of neutrals, and from the very limited resolution which can be obtained on track momentum by the use of three-point fits to the track helices. Nonetheless, jets can indeed be found and their characteristics measured with meaningful resolution. In particular, the interaction point along the z axis can be measured with a resolution of a few millimeters, enabling the discrimination of jets originated by the triggering interaction and the rejection of those created by the scores of additional hard and semi-hard interactions that each bunch crossing will originate at high luminosity.

The use of pixel information for jet reconstruction can indeed, by virtue of the identification of the true vertex, significantly reduce the jet energy thresholds required to maintain the rate of inclusive jet triggers below a reasonable threshold.

Of course, what we considered in this study was an ideal case when the full granularity of the pixel detector becomes magically available to the trigger decision at level 1. This has been demonstrated to be unfeasible with the existing hardware in the tracker of CMS, and quite hard, if not similarly impossible, with an upgraded detector retaining a similar geometry. A much more realistic approach would be that of considering "tiles" of double columns of pixels: information on their occupancy still allows the reconstruction of the coordinate of the primary interaction vertex along z. Double columns could in principle be read with optical fibres and the existing detectors, but that option, too, has remained hypothetical so far.

5.1.1 Tracker upgrade for SLHC and tracking trigger

In the last few years the possible design of an upgraded tracker for the CMS experiment, to be used for a "Super-LHC" phase of running (with expected luminosity reaching $10^{35} cm^{-2} s^{-1}$), has been studied by at least two independent groups. The design is required to yield the triggering capabilities which are needed to suppress the rate of low-momentum muon tracks, as well as to support the identification of events of interest still provided by the other detector components.

The RD program for the upgrade of the tracker for Super-LHC (SLHC) [62] has been recently approved and sets the guidelines for the development of a tracking trigger. The design of the upgraded tracker and trigger systems will be closely linked from the start, and their performace must be optimised in

parallel, since the requirements of triggering and precision tracking using the same sub-detector, potentially conflict.
Among the proposals for a tracking trigger, the use of pixels for triggering at SLHC has been studied in some detail [63]. At SLHC $\sim$ 100 proton-proton collisions every 12.5 ns are expected to occur, in contrast to only $\sim$ 25 for normal LHC operations. Each proton bunch crossing will generate much more than the present 1 MB of data, and the expected data transfer to disk will not exceed the rate of about $\sim$ 100 Hz. Hence the increased luminosity will lead to a drastic revision of the existing trigger strategies.
The rate of the single muon trigger in figure 5.1 shows that it is not sufficient to increase the cut on the Pt of the muon because of the poor transverse momentum resolution at high Pt of the muon trigger alone. When it is used together with the tracker, the gain in resolution helps to lower the trigger rate.
Due to the shape of the underlying background distribution the required first level trigger rate reduction for SLHC cannot be achieved by increasing the redesigned first level trigger thresholds, even if one is willing to pay the price of cutting useful physics signals. Furthermore, the increase in minimum bias events of about a factor of 5 will degrade the efficiency and purity to select Higgs and other exotic signals.
One possibility is to try and use the HLT tracking algorithm with as few changes as possible at the first level trigger with full pixel granularity. A regional approach, defining trigger towers in the detector, can be used to reduce the combinatorics.

The task of reading silicon detector information, processing it and performing a suitable track reconstruction, however, remains formidable, despite the successful example of the Silicon Vertex Tracker which has operated so well in the Run II of the CDF experiment at the Tevatron. It is presently debated whether the use of associative memories containing track patterns is the best option to reconstruct charged tracks, when used in conjunction with muon detector primitives which can also identify tracks but yield a much worse momentum resolution.

It is thus clear that in such a fluid situation any study of physics prospects with the Super-LHC can provide useful information. If the improvement of muon momentum resolution is the main goal, on the other hand one needs to also provide a physics case justifying the effort. The study of the characteristics of the Higgs boson, its production properties, its quantum numbers, and a direct measurement of its coupling to fermion fields, are all issues which have to be taken in consideration. The associated production of top quarks and a Higgs boson may not only be an important ingredient in the Higgs

searches at low mass, in case the extraction of a diphoton mass peak turned out to be harder than expected: it is in fact a process which can and should be measured by the full potential of the CERN facilities.

5.2 Conclusions on associated top-Higgs production

In this study we have tried to demonstrate that the search for associated production of a top quark pair and a Higgs boson is not hopeless with CMS. We considered a final state which has several drawbacks with respect to the more classical signatures of top decay: a low trigger efficiency, the use of missing transverse energy – which is problematic in CMS due to its poor resolution–, and the study of a final state which prevents a kinematical fit to reconstruct fully the decay kinematics. We have, so to speak, taken the opposite approach with respect to our optimistic study of full pixel granularity for level 1 triggering, because we considered the least promising situation.

The fact that with 100fb^{-1} of data a signal of $t\bar{t}$H production could emerge from CMS data –provided a more careful study is carried out than the one we have been able to present in this thesis– should motivate further studies on this channel, which seems instead to have been abandoned recently. In Chapter 4 we have discussed several improvements to the analysis which could achieve a lot in terms of sensitivity increases.

We remain optimistic about the prospects of a $t\bar{t}$H search in CMS, especially since global fits to electroweak observables point to a light mass for the Higgs boson. If the Higgs exists, it now appears likely that it will be discovered by LHC only through a painstaking analysis of all low-mass signatures.

Bibliography

[1] S. L. Glashow. Partial symmetries of weak interactions.

[2] A. Salam. *Elementary Particle Theory.*

[3] S. Weinberg. A model of leptons.

[4] S.M. Bilenky W.M. Alberico.

[5] Particle Data Group. Review of particle physics. *Phys. Rev.*, D66, 2002.

[6] P. W. Higgs. Broken symmetries, massless particlees and gauge fields.

[7] R. Barbieri and A. Sturmia. Constraints on new physics at tev scale. *arXiv:hep-ph/0007265.*

[8] S. P. Martin. A super-symmetry premier. *arXiv:hep-ph/9709356.*

[9] E.H. Simmons C. T. Hill. Strong dynamics and electroweak symmetry breaking. *arXiv:hep-ph/0203079.*

[10] L. Randall and R. Sundrum. A large mass hierarchy from a small extra dimension. *Phys. Rev. Lett.*, (83):3370, 1999.

[11] D. N. Spergel et al. [WMAP Collaboration]. First year wilkinson microwave anysotropy probe (wmap) observations: Determination of cosmological parameters. *Astrophys. J. Suppl.*, (148):175, 2003.

[12] Beenakker W., Dittmaier S., Kramer M., Plumper B., Spira M., and Zerwas P.M. Higgs radiation off top quarks at the tevatron and the lhc. *Phys. Rev. Lett.*, 87(20), 2001.

[13] M. Dittmar and H. Dreiner. How to find a higgs boson with a mass between 155 and 180 gev at the cern lhc. *Phys. Rev. D*, 55(1):167–172, Jan 1997.

[14] Dittmar M. and A.S. Nicollerat. Higgs studies using $gg \rightarrow h_{SM}$ and $qq \rightarrow qqh_{SM}$ at the lhc. *CMS-NOTE*.

[15] D. Zeppenfeld, R. Kinnunen, A. Nikitenko, and E. Richter-Waş. Measuring higgs boson couplings at the cern lhc. *Phys. Rev. D*, 62(1):013009, Jun 2000.

[16] Pieri M. Searches for higgs bosons at lhc. *CMS conference report*, (031).

[17] S. Asai et al. Prospects for the search for a standard model higgs boson in atlas using vector boson fusion.

[18] R. Kinnunen. Lhc potential for the higgs boson discovery. (58), 2004.

[19] Denegri D. et al. Summary of the cms discovery potential for the mssm susy higgses. *CMS-NOTE*, (032).

[20] CDF and D0 Collaborations. Results of the tevatron higgs sensitivity study. (320).

[21] M. S. Carena et al. Report of the tevatron higgs working group. (349), 2000.

[22] Pumplin J., Stump D.R., Huston J., Lai H.L., Nadolsky P., and Tung W.K. New generation of parton distributions with uncertainties from global qcd analysis. *JHEP*, 0010017(012):0000, 2002.

[23] LEP Injector Study Group. *LEP Design Report*. 1984.

[24] LEP Injector Study Group. *LEP Design Report Volume III*. CERN-LEP, 1996.

[25] ATLAS Collaboration. *Atlas Detector and Physics Performance - Technical Design Report Vol. I*. CERN/LHCC 99-14, 1999.

[26] ALICE Collaboration. *A Large Ion Collider Experiment Technical Proposal*. CERN/LHCC 95/71, LHCC/P3, 1995.

[27] LHCb Collaboration. *A Large Hadron Collider Beauty Experiment fro Precision Measurements of CP-Violation and Rare Decays LHCb Technical Proposal*. CERN/LHCC 98/04, LHCC/P4, 1998.

[28] ATLAS Collaboration. *Atlas: A Toroidal LHC ApparatuS Techinal Proposal*. CERN/LHCC 94/42, LHCC/P2, 1994.

[29] CMS Collaboration. *The Compact Muon Solenoid Technical Proposal*. CERN/LHCC 94-38, LHCC/P1, 1994.

[30] D. Froidevaux and P. Sphicas. General-purpouse detectors for the large hadron collider. *Ann. Rev. Nucl. Part. Sci.*, (56):375.

[31] CMS Collaboration. *CMS: The Tracker Project Technical Design Report*. CERN/LHCC 98-06, CMS TDR 5, 1998.

[32] CMS Collaboration. *CMS: Addendum to the CMS Tracker TDR*. CERN/LHCC 2000-016, CMS TDR 5 Addendum, 2000.

[33] CMS Collaboration. *CMS: The Electromagnetic Calorimeter Project Technical Design Report*. CERN/LHCC 97-33, CMS TDR 4, 1997.

[34] CMS Collaboration. *CMS: The Hadron Calorimeter Project Technical Design Report*. CERN/LHCC 97-31, CMS TDR 2, 1997.

[35] CMS Collaboration. *CMS: The Muon Project Technical Design Report*. CERN/LHCC 97-32, CMS TDR 3, 1997.

[36] RD5 Collaboration. Bunch crossing identification at lhc using a mean timer technique. *Nucl. Instrum. Meth.*, 0000(0000):91–97 A336, 1993.

[37] CMS Collaboration. *CMS: The Trigger and Data Acquisition Project, Volume I: The Level-1 Trigger Technical Design Report*. CERN/LHCC 2000-038, CMS TDR 6.1, 2000.

[38] http://cmsdoc.cern.ch/cms/prs/jetmet/jet_calibration/.

[39] CMS Collaboration. *CMS: The Trigger and Data Acquisition Project, Volume II: The Data Acquisition and High-Level Trigger Technical Design Report*. CERN/LHCC 2000-26, CMS TDR 6.2, 2002.

[40] S. Cucciarelli et al. Track parameter evaluation and primary vertex finding with the pixel detector. *CMS Note*, 0000(026), 2003.

[41] D. Kotlińsky. Track reconstruction and primary vertex finding using the pixel detector data. *CMS IN*, 0000(022), 2000.

[42] D. Kotlińsky and A. Starodumov. High level tracker triggers for cms. *NIM A*, 0000(501):222–228, 2003.

[43] CMS Software and Computing Group. Object oriented reconstruction for cms analysis. *CMS-IN*, 001(0000):0000, 1999.

[44] M. Maire A. C. McPherson R. Brun, F. Bruyant and P. Zanarini. Geant3. (84).

[45] S. Agostinelli et al. Gimage4 a simulation toolkit. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 506(3):250–303, 2003.

[46] S. A. Teukolsky W. T. Vetterling W. H. Press, B. P. Flannery. *Numerical Recipes in C*. Cambridge University Press.

[47] M. Pioppi et al. First performance studies of a pixel-based trigger in the cms experiment. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, (570):271.

[48] S. Cucciarelli, M. Konecki, D. Kotliński, and T. Todorov. Track parameter evaluation and primary vertex finding with the pixel detector. *CMS-NOTE*, 036(0000):0000, 2003.

[49] CMS Collaboration. *CMS Physics Techincal design report Volume 2: Physics performance*. CERN/LHCC 2006-21, CMS TDR 8.2, 2006.

[50] K. S. Cranmer. Potential for higgs physics at the lhc and super-lhc.

[51] S. Cucciarelli et al. Search for h$\rightarrow b\bar{b}$ in association with a $t\bar{t}$ pair at cms. *CMS-NOTE*, (119), 2006.

[52]

[53] Alexei N. Safonova and CDF Collaboration. Physics with taus at cdf. (144):323, 2005.

[54] A Search for $t \rightarrow \tau\nu q$ in $t^{-}t$ Production. (99), 2007.

[55] J. Incandela and the CDF Collaboration. Top decay to lepton+jets: Cdf b tags and cross section.

[56] S. Mrenna T. Sjostrand and P. Skands. Pythia 6.4 physics and manual. *JHEP05*, (026), 2006.

[57] M. Spiropulu M. Pierini. Study $t\bar{t}$ + jets as a function of jet p_t and jet multiplicity for 2007 1 fb^{-1} monte carlo production at cms.

[58] M. Spiropulu M. Pierini. Study w + jets in bins of boson p_t and jet multiplicity for 2007 1 fb^{-1} monte carlo production at cms.

[59] F. Palla G. Segneri. Lifetime based b-tagging with cms. *CMS-NOTE*, (46), 2002.

[60] F. P. Schilling. B-tagging in the 2007 an notes.

[61] on behalf of the CDF B. Kilminster and D0 Collaboration. Searches for the standard model higgs at the tevatron. 2007.

[62] Imperial College London Rutherforf Appleton Laboratory University of Bristol, Brunel University. R&d in preparation for an upgrade of cms for the super-lhc by uk groups. 2007.

[63] M. D'Alfonso et al. First level trigger using pixel detector for the cms experiment.