Online selection of fully-hadronic $t\bar{t}$ decays with the ATLAS detector

Author:
Menelaos Tsiakiris

MSc Thesis
Supervisor:
Prof. Dr. Stan Bentvelsen

Daily Supervisor:
Dr. Sander Klous

Second Reviewer:
Dr. Marcel Vreeswijk
ABSTRACT

One of the main goals of the ATLAS detector, being built at the Large Hadron Collider in CERN, is to discover the Higgs particle. Theoretically, the Higgs particle can be produced in association with a top and antitop ($t\bar{t}$) pair and thus understanding the $t\bar{t}$ decay channels is an important step on the road to discovery. The final states of the $t\bar{t}$ are distinguished as: semi-leptonic, di-leptonic and full-hadronic state. The fully-hadronic final state consists of six jets, from which two jets originate from the hadronization of the $b$-quark. This channel has the largest branching ratio (44%) and because of its distinguishable final state it can be fully reconstructed. The major disadvantage, however, is that it encounters a large multijet background coming from QCD processes. The motivation of examining the feasibility of analysing fully-hadronic $t\bar{t}$ events leads to the investigation of the jet reconstruction at the online selection level (trigger). In this thesis, a validation study on jet trigger reconstruction at the High-Level Trigger (HLT) of ATLAS is performed. The efficiency for accepting jets with a certain energy, at various trigger thresholds and multiplicities, is calculated by means of the so-called “turn-on” curves. Also, the efficiency of accepting fully-hadronic $t\bar{t}$ events with respect to the trigger rate of QCD multijet background events, is calculated. Finally, a proof of principle study is made for comparing the validation turn-on curves with the turn-on curves based on real data.
1 Introduction

2 The Standard Model and the top quark
   2.1 Overview of the Standard Model
   2.2 The Standard Model Higgs
      2.2.1 Spontaneous Symmetry Breaking and Higgs mechanism
      2.2.2 Theoretical limits for the SM Higgs
      2.2.3 Production cross sections and branching ratios
   2.3 The top quark
      2.3.1 Top quark production and decay
      2.3.2 The fully-hadronic channel
      2.3.3 Backgrounds to top-antitop decay
      2.3.4 The associated Higgs production

3 The LHC and the ATLAS detector
   3.1 The LHC machine in a nutshell
   3.2 Physics goals of the ATLAS experiment
   3.3 The detector’s layout
      3.3.1 The Inner Detector
      3.3.2 The Muon Spectrometer
      3.3.3 The Calorimetry system
   3.4 The Trigger and Data Acquisition system of ATLAS
      3.4.1 Architecture of the three levels Trigger
      3.4.2 The Calorimeter and the Jet trigger
## 4 Online selection of fully-hadronic top pair decays

<table>
<thead>
<tr>
<th>Subsection</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>4.1 Analysis Data Object - AOD</td>
<td>35</td>
</tr>
<tr>
<td>4.2 Analysis tools</td>
<td>36</td>
</tr>
<tr>
<td>4.3 The Jet Trigger Slice</td>
<td>37</td>
</tr>
<tr>
<td>4.4 Jet Trigger quality based on fully-hadronic events</td>
<td>38</td>
</tr>
<tr>
<td>4.4.1 Matching procedure</td>
<td>39</td>
</tr>
<tr>
<td>4.4.2 Analysis scheme</td>
<td>40</td>
</tr>
<tr>
<td>4.4.3 Jet signatures and trigger acceptance</td>
<td>40</td>
</tr>
<tr>
<td>4.4.4 Background study with QCD multijet events</td>
<td>46</td>
</tr>
<tr>
<td>4.5 Cross-correlation with semi-leptonic events</td>
<td>48</td>
</tr>
</tbody>
</table>

## 5 Summary and conclusions

<table>
<thead>
<tr>
<th>Subsection</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Summarize and conclude the analysis</td>
<td>53</td>
</tr>
</tbody>
</table>
CHAPTER 1

Introduction

The top quark is the heaviest elementary particle yet discovered. Its mass, of about 175 GeV, is of the same order as the electroweak scale and roughly twice that of the $W^\pm$ and $Z^0$ bosons. In the Standard Model, because of its large Yukawa coupling to the Higgs boson, and hence to the mechanism of electroweak symmetry breaking, the top quark is expected to have unique dynamics.

The main top production channel, for the Large Hadron Collider energies, will be through $t\bar{t}$ pair production. Understanding the top quark decays is of great interest for the ATLAS collaboration. The $t\bar{t}$ events are related to, e.g., the potential discovery channel of the Higgs boson, through the associated production of $t\bar{t}H$. Particularly interesting is the $t\bar{t}$ fully-hadronic decay channel because of its large branching ratio and its without missing-$E_T$ final state which can be fully reconstructed.

The drawback of the fully-hadronic channel is that it experiences an enormous QCD background, thus an efficient selection procedure capable of identifying this specific channel is necessary. Due to the “only-jets” final state the main interest lies in the jet trigger efficiency. It is the purpose of this thesis to examine the status of the Jet Trigger Algorithms in the High-Level Trigger.

In Chapter 2, an introduction is given focusing on the theoretical aspects of the top quark and the related Higgs physics. In addition, the different production and decay channels of the $t\bar{t}$ events are addressed with main interest on the fully-hadronic decay channel and its expected background.

In Chapter 3, a brief overview of the Large Hadron Collider and the

\footnote{The natural units system is used throughout the thesis ($c = \hbar = 1$).}
ATLAS detector layout is given, concentrating mainly on the Calorimeter properties that determine the quality of the jet reconstruction. Also, the Trigger system of ATLAS and the procedure of selecting the events is explained. Special attention is given to those aspects related to the Jet Trigger, such as the Calorimeter Trigger reconstruction for jets. Finally, the HLT jet reconstruction algorithms are discussed briefly.

In Chapter 4, a quality study of the Jet Trigger follows. The goal is to examine the jet trigger quality using fully-hadronic events, especially at the HLT level. Since the analysis is based on MonteCarlo generated events, a small section of the chapter addresses the specifications of these samples and the analysis tools. The validation methodology is explained in detail and the results are discussed. Additionally, the trigger efficiency for fully-hadronic $t\bar{t}$ is evaluated with respect to the QCD multijet background rate.

Finally, extracting the trigger efficiency as a function of jet energy is studied based only on information that will be available in real data. This is done with a cross-correlation between the jet trigger and the lepton trigger in semi-leptonic $t\bar{t}$ events. Also, the relation between the information during the data taking period with the information from the validation study is examined.
The Standard Model is the accepted and experimentally well-tested theory of strong, electromagnetic and weak interactions. For more than twenty years it has provided the guidelines for the discoveries of the, \( W^\pm \) and \( Z^0 \) bosons, and recently for the top quark. The Standard Model predicts the existence of the Higgs particle. One of the main challenges of the Large Hadron Collider experiments is to find this particle.

2.1 Overview of the Standard Model

One of the basic elements of theoretical High Energy Physics is the principle of symmetry and from all the kinds, gauge symmetry is playing the most important role. Gauge symmetries are formulated by requiring the Lagrangian of a system to be invariant under gauge transitions. For a field \( \phi \) this would require invariance under the following transformation:

\[
\phi \rightarrow e^{i\alpha(x)} \phi, \text{ where } \alpha(x) \text{ is an arbitrary function.} \tag{2.1.1}
\]

The Standard Model (SM) is the outcome of lot of research in the field of High Energy Physics. It was introduced in the late 60's by Glashow, Salam and Weinberg (GSW) and is a quantum field theory model of point-like fermions and bosons, defined by the gauge symmetry group: \( SU(3)_C \otimes SU(2)_L \otimes U(1)_Y \). It describes the interactions between the three generations of quarks and leptons (spin-1/2 particles and one Higgs field. The constituent symmetries of the group are: \( SU(3)_C \), the Quantum Chromo-Dynamics (QCD) color symmetry describing the strong interactions, and \( SU(2)_L \otimes U(1)_Y \) the isospin and hypercharge symmetry describing the electroweak interactions.
These symmetries, introduce a total of 12 vector bosons (spin-1 particles): the 8 gluons from the QCD theory, mediators of the strong force, and 3 + 1 from the electroweak which are the $W^\pm$, $Z^0$, mediators of the weak force, and $\gamma$, mediator of the electromagnetic force. Together with the three “families” of quarks and leptons (Table 2.1), as well as the theoretically predicted Higgs boson, a spin-0 particle, we have the pieces that constitute the SM puzzle.

<table>
<thead>
<tr>
<th>Particle symbol</th>
<th>Mass (MeV)</th>
<th>Electric Charge</th>
</tr>
</thead>
<tbody>
<tr>
<td>First generation</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Quarks</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$u$</td>
<td>1.5 − 4.0</td>
<td>2/3</td>
</tr>
<tr>
<td>$d$</td>
<td>4 − 8</td>
<td>−1/3</td>
</tr>
<tr>
<td>Leptons</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$e^-$</td>
<td>0.511</td>
<td>−1</td>
</tr>
<tr>
<td>$\nu_e$</td>
<td>≤ 2.2 × 10^{-6}</td>
<td>0</td>
</tr>
<tr>
<td>Second generation</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Quarks</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$c$</td>
<td>1150 − 1350</td>
<td>2/3</td>
</tr>
<tr>
<td>$s$</td>
<td>80 − 130</td>
<td>−1/3</td>
</tr>
<tr>
<td>Leptons</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\mu^-$</td>
<td>105.7</td>
<td>−1</td>
</tr>
<tr>
<td>$\nu_\mu$</td>
<td>≤ 0.17</td>
<td>0</td>
</tr>
<tr>
<td>Third generation</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Quarks</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$t$</td>
<td>170900 ± 1800</td>
<td>2/3</td>
</tr>
<tr>
<td>$b$</td>
<td>4100 − 4400</td>
<td>−1/3</td>
</tr>
<tr>
<td>Leptons</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\tau^-$</td>
<td>1784.1</td>
<td>−1</td>
</tr>
<tr>
<td>$\nu_\tau$</td>
<td>≤ 15.5</td>
<td>0</td>
</tr>
<tr>
<td>Gauge Bosons</td>
<td>Mass (GeV)</td>
<td>Electric Charge</td>
</tr>
<tr>
<td>gluons (g)</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>photon ($\gamma$)</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>$W$ boson</td>
<td>80.398 ± 0.025</td>
<td>±1</td>
</tr>
<tr>
<td>$Z$ boson</td>
<td>91.1876 ± 0.0021</td>
<td>0</td>
</tr>
</tbody>
</table>

Table 2.1: Overview of the three generations of quarks and leptons, and the gauge bosons. Antiparticles have opposite electric charges. Top quark and $W$ boson mass from the LEP Electroweak Group. Tau neutrino mass from combined results of ALEPH and CLEO experiments. The rest from the Particle Data Group.

As predicted from the GSW model, the weak force carriers have mass. However, the presence of mass terms for gauge field destroys the gauge invariance of the Lagrangian. Thus, the question that raises is: how do these masses originate? The solution to this problem comes from the Higgs mechanism.
2.2 The Standard Model Higgs

The Higgs mechanism is based on the concept of Spontaneous Symmetry Breaking (SSB) and the modification of the vacuum state energy. It successfully gives mass to the weak force carriers but also introduces a “new” particle, the Higgs particle. Its existence is not only based on the theoretical approach but also from experimental results such as the measurement of the electroweak parameter $\rho$. This parameter is experimentally confirmed to be close to unity, in agreement with SM theoretical predictions. These predictions follow from the particular form of the Higgs field (see [1], page 278).

2.2.1 Spontaneous Symmetry Breaking and Higgs mechanism

In this and the following sections the formulation follows from [1], [2] and [3]. To formulate the Higgs mechanism we consider the Lagrangian for a complex scalar field $\phi$:

$$L(\phi) = \frac{1}{2} (\partial_\mu \phi)^*(\partial^\mu \phi) - V(\phi)$$

(2.2.1)

where $V(\phi)$ is the potential:

$$V(\phi) = -\mu^2 (\phi^* \phi) + \lambda (\phi^* \phi)^2$$

(2.2.2)

Where $\lambda$ is a positive constant and $\mu^2$ is the mass term of scalar field theory [4]. For $\lambda > 0$ and $\mu^2 > 0$ the Lagrangian is transformed to the QED equivalent Lagrangian of a scalar particle with mass $\mu$. If however $\lambda > 0$ and $\mu^2 < 0$, the minimum energy is taken by requesting:

$$\frac{\partial V}{\partial \phi} = 0$$

(2.2.3)

and results to a “mexican hat” potential. A profile view of the potential is shown in figure 2.1. The potential now has a circle of minima which is given by:

$$\phi = \frac{\mu}{\sqrt{2\lambda}} e^{i\theta}, 0 \leq \theta < 2\pi.$$  

(2.2.4)

Due to gauge invariance we can do perturbation theory around any point of minimum energy. This way we force the system to select a ground state (SSB). Choosing $\theta = 0$ gives a vacuum expectation value (VEV) of $u = \mu/\sqrt{2\lambda}$. It is possible to rewrite the Lagrangian as a function of two real scalar fields $\eta$ and $\xi$ using the relation $\phi - u = \eta + i\xi$. After SSB the $\eta$ field

1The electroweak parameter is defined as the ratio of neutral and charged current strength at zero-momentum transfer.

2With this form both fields have zero VEV.
is massive with \( m_\eta = \sqrt{2u\lambda} = 2\mu \) while \( \xi \) is massless.

We can apply the same process to a simple U(1) gauge theory. In this case, after requesting the Lagrangian to be invariant under unitary gauge transformations, we get one real massive scalar field (the \( \eta \)) and one massive vector field (denoted as \( A_\mu \)) while the massless \( \xi \) field has been consumed. This process is known as the Higgs mechanism and applying it to the electroweak gauge group \( SU(2)_L \otimes U(1)_Y \) leads to the appearance of the Higgs particle.

Initially the Lagrangian will contain a complex scalar doublet, four massless vector bosons and zero-mass fermions. The complex scalar doublet is coupled to the fermions with Yukawa couplings of the form \( g_f \bar{\psi}_f \phi \psi_f \). With SSB the Yukawa couplings give the mass terms of the fermions. Moreover, applying the Higgs mechanism results in one real scalar (Higgs boson), three massive vector vector (\( W^\pm \) and \( Z^0 \)) and one massless vector boson (photon). Thus, the masses of the elementary particles are explained.

### 2.2.2 Theoretical limits for the SM Higgs

Although the Higgs mechanism predicts the existence of the Higgs boson, the mass term of the Higgs remains a free parameter. Nevertheless, there are theoretical limits in the mass range.

#### Unitarity

One of the main requirements in the Standard model is unitarity. Unitarity suggests that the sum of the scattering amplitudes of Feynman diagrams has to be unitary or else the theory is non-perturbative and no quantitative predictions are possible. In fact, this implies conservation of probability.

Considering \( 2 \to 2 \) scattering, it is possible to decompose any scattering amplitude into spin-L partial waves (\( \alpha_L \)) making sure that they all obey unitarity (\( \alpha_L < 1 \)). In this case, the total cross section of the scattering can
be expressed as the sum on all the partial waves. Applying this methodology to the $W^+W^- \rightarrow W^+W^-$ boson scattering, as proposed from [6], gives an upper limit to the Higgs mass, $m_H < 800$ GeV. This limit, however, suggests only the breakdown of perturbation theory and it does not exclude the existence of a Higgs with a mass above this limit.

**Triviality and Vacuum stability**

In addition to unitarity, there are two requirements that suggest the existence of a Standard Model Higgs in a region below 800 GeV. These are the triviality and the vacuum state stability. The first suggests that the theory is perturbative until an upper “cut-off” energy, while the second requires that $\lambda$ remains positive. Both of them are a result of the running coupling $\lambda$. The derivation follows from [3].

Consider the potential $V(\phi^*\phi) = \mu^2(\phi^*\phi) + \lambda(\phi^*\phi)^2$, where $\lambda$ is the self-coupling of the Higgs field. In all gauge interactions, the coupling constant “runs” with the energy. The same happens with $\lambda$ which runs with energy scale $\mu$. Using renormalization group scaling, $\frac{1}{\lambda(\Lambda)} = \frac{1}{\lambda(u)} + (...)\ln \left(\frac{\Lambda}{u}\right)$, this “running” can be expressed as a beta-function [7]:

$$\frac{d\lambda}{dt} = \beta_\lambda$$  \hspace{1cm} (2.2.5)

where $t = \ln(\frac{\Lambda^2}{\mu})$ relating $\lambda(\Lambda)$ to $\lambda(u)$, with $\Lambda$ the cutoff energy up to which we accept the SM. The top quark due to its large mass, has non-negligible radiative corrections and must be taken into account, hence (2.2.5) becomes:

$$\frac{d\lambda}{dt} = \frac{1}{32\pi^2} \left\{ 24\lambda^2 - (3g'^2 + 9g^2 - 12h_t^2)\lambda + \frac{3}{8}g'^4 + \frac{3}{4}g^2g'^2 + \frac{9}{8}g^4 - 6h_t \right\}$$ \hspace{1cm} (2.2.6)

Here, $h_t = \sqrt{2}(m_t/u)$ is the Yukawa coupling of the top quark to the Higgs boson and $g, g'$ are the $SU(2)_L \times U(1)_Y$ gauge couplings. In our definition above there are two regimes which can be examined: either $\lambda > h_t, g, g'$ or $\lambda < h_t, g, g'$.

**Triviality**

The first regime, $\lambda > h_t, g, g'$ originates from the fact that in a renormalizable theory containing a massive scalar field, such as the SM, the strength of...
the quartic self-interaction terms reaches infinity at some scale — i.e. they are inversely proportional to the strength of the interaction. Requiring the validity up to a given scale, bounds the self-coupling and the Higgs mass. The only valid theory for all scales is the trivial one, hence no interactions are present. Mathematically, this writes the beta-function \( \beta_{\lambda} \) as:

\[
\frac{d\lambda}{dt} = \frac{3}{4\pi^2} \lambda^2 \Rightarrow \lambda(\Lambda) = \frac{\lambda(u)}{1 - \frac{3\lambda(u)}{4\pi^2} \ln \left( \frac{\Lambda^2}{u^2} \right)} \quad (2.2.7)
\]

as only the first term is important. Because the \( \lambda(\Lambda) \) is increasing with increasing \( \Lambda \), at some point it reaches infinity. Since we have related the coupling constant between the two different energy scales, \( \Lambda \) and \( u \), the point where \( \lambda(\Lambda) \) reaches infinity should place a bound to \( \lambda(u) \). This is expressed from the following:

\[
\frac{3\lambda(u)}{4\pi^2} \ln \left( \frac{\Lambda^2}{u^2} \right) < 1 \quad (2.2.8)
\]

using the relation \( m_h^2 = 2\lambda u^2 \), derived from the Lagrangian with the Higgs potential, we reach the relation:

\[
(m_h^{\text{max}})^2 < \frac{8\pi u^2}{3} \ln \left( \frac{\Lambda^2}{u^2} \right) \quad (2.2.9)
\]

which for the Grand Unified Theories scale \( \Lambda = 10^{16} \) GeV gives \( m_h \leq 175 \) GeV and for the Planck scale \( \Lambda = 10^{19} \) GeV gives \( m_h \leq 155 \) GeV.

**Vacuum Stability**

The second regime, \( \lambda < h_t, g, g' \), comes from the necessity of having the correct structure of the vacuum. Radiative correction, mainly from the top quark, may drive the values of \( \lambda \) to negative values in the high energy regime. To avoid having \( \lambda(\Lambda) < 0 \) it is needed to start with a high \( \lambda(u) \). At the minimum value, it gives a lower limit to \( m_h \). In this regime the important parameters are the last ones from \( 2.2.7 \), and the beta-function \( \beta_{\lambda} \) can be written as follows:

\[
\frac{d\lambda}{dt} = \beta_{\lambda} = -\frac{3h_t^4}{16\pi^2} + \frac{3}{8}(g'^4 + 2g^2g' + 3g^4) \quad (2.2.10)
\]

In addition it can be written as a contribution\(^4\) from \( M_W, M_Z \) and \( M_t \). We get:

\[
\lambda(\Lambda) - \lambda(u) = \beta_{\lambda} \ln \left( \frac{\Lambda^2}{u^2} \right) \quad (2.2.11)
\]

\(^4\)Using: \( h_t^4 = 4\frac{m_t^2}{G_F} \), \( M_W = \frac{1}{2} u g \), \( M_Z = \frac{1}{2} u \sqrt{g'^2 + g^2} \)
from which, if we require $\lambda(\Lambda) > 0$ and we use $m_h^2 = 2\lambda u^2$, we conclude that:

\[
(m_h^{\text{min}})^2 > \frac{3}{8\pi^2 u^2}(2M_W^4 + M_Z^4 - 4M_t^4)\ln\frac{\Lambda^2}{u^2}
\] (2.2.12)

However, this result is always true (the $M_t$ is much larger than the $M_W$ and $M_Z$) and it does not introduce any limit to the Higgs mass. The method needs further refining by introducing two loop corrections which drive the coupling $\lambda$ further down. At some point a limit for the Higgs is visible. This was explicitly shown in [8] but also in [9,10]. In the end, for the Planck scale of $\Lambda = 10^{19}$ GeV, we have $m_h > 130$ GeV.

The mass regions of the Higgs are shown in figure 2.2. The region above the upper line is theoretically forbidden by triviality and the region below the lower line theoretically forbidden from vacuum stability. Also, direct searches at LEP[11] have excluded a large part of the allowed region giving a lower limit of 114.4 GeV at 95% C.L. Furthermore, an upper limit of 144 GeV at the 95% C.L. is implied based on electroweak precision measurements (March 2007 - based on updated measurements of the top quark and $W$ masses [12]).

![Fig. 2.2: The theoretical Higgs mass limitations and the uncertainty for the upper and lower boundaries calculated with $m_t = 175$ GeV and $\alpha_s(m_Z) = 0.118$ [13].](image)

### 2.2.3 Production cross sections and branching ratios

The construction of the Large Hadron Collider (LHC) and of the ATLAS experiment was to a large extent motivated for probing the underlying physics of the Higgs boson. The decay channels of the Higgs boson, for a mass region of 80 GeV up to 1 TeV, are of interest for the detector design.
Typically, for the LHC, the production cross-section of an SM Higgs with a mass around 500 GeV is 1 pb, derived from QCD calculations [14] (see also figure 2.6). This value corresponds to about $10^5$ Higgs bosons produced in one year of LHC running at high luminosity. The cross section increases rapidly in case of a lower Higgs mass. The main production of the Higgs happens through gluon fusion, $gg \rightarrow H$. However, processes such as $q\bar{q} \rightarrow HW$, $q\bar{q} \rightarrow HZ$ or $gg, q\bar{q} \rightarrow t\bar{t}H$ and $gg, q\bar{q} \rightarrow b\bar{b}H$ are also contributing (figure 2.3) [14], and although they have cross section lower by a factor of about 100, in certain cases they could give better signal to background ratios. The Higgs production cross-sections are shown in figure 2.4 as a function of the Higgs mass, for proton-proton collisions with $\sqrt{s} = 14$ TeV.

After production the Higgs will decay relatively fast. Its decay process can serve as a tool to determine its mass. For different mass regimes, different decay processes have to be taken into account as can be realized from figure 2.5. We can distinguish three possible cases for the Higgs particle: the Light Higgs case, where $m_H < 200$ GeV, the Intermediate mass Higgs, with $200$ GeV $< m_H < 800$ GeV, and the Heavy Higgs regime, with $m_H > 800$ GeV.

The branching ratios of different decay channels with respect to the Higgs mass, are shown in figure 2.5. Taking into account the cross-sections for $pp$ (or $p\bar{p}$) collisions in figure 2.6, the main background channels are known. The
combined information can lead to selecting the important decay processes for each regime.

In the intermediate mass region the most promising channel is the $H \to ZZ^* \to 4l$. It has long been known to be a major discovery channel with the ATLAS detector \cite{17} and extensive studies have been performed \cite{18,19}. In this region, the decay width, $\Gamma_H$, is small and the background originates mostly from fast muons, hadrons that passed through the calorimeters, as well as neutrons and secondary particles coming from shower effects. Thus, a narrow peak is expected to be visible on top of the broad background.

For the Heavy Higgs ($m_H > 800$ GeV) case, the dominant decay channel is $H \to ZZ \to l^+l^-\nu\bar{\nu}$, six times larger than the four lepton channel. It can be detected by selecting events with two leptons with high $p_t$ and high missing $E_t$ \cite{20}. Also of interest could be the channels, $H \to WW$ and $ZZ \to l^\pm\nu + 2\text{ jets}$ \cite{21}.

Very interesting is the case of the light Higgs regime. In the region below 150 GeV channels such as the $H \to \gamma\gamma$ or even better, the associated $WH \to l\nu\gamma\gamma$ channel, can contribute significantly to the Higgs discovery, but it requires an excellent calorimeter \cite{17}. However, as seen in figure 2.5, in most of the cases, the Higgs decays into a pair of $b$-quarks ($H \to b\bar{b}$). Unfortunately, a huge QCD background is expected as the cross section for $b\bar{b}$ production is very large (figure 2.6), which makes this channel useless. The associated production of $t\bar{t}H$ with the Higgs going to a pair of $b$-quarks can be used instead. The top quarks can decay, producing lepton(s) and(or) jets. This can eventually give an enhanced signal significance, for integrated

\begin{figure}[h]
\centering
\includegraphics[width=0.5\textwidth]{higgs_branching_ratios.png}
\caption{Higgs branching ratios w.r.t its mass \cite{16}.}
\end{figure}
luminosity of 100 fb\(^{-1}\) (LHC 1 year run at high luminosity) as shown in figure 2.7. Related studies have been made in [17], [22] and [23]. Because of its increased statistic, this channel can prove an excellent candidate for Higgs discovery. For this reason determining the top quark properties (mass, cross-section etc.) as well as understanding the physics related to the top (production, decay etc.), is an important step.

Fig. 2.6: Cross sections for processes in \(p\bar{p}\) collisions in the Tevatron regime and \(pp\) collisions in the LHC regime [14].

Fig. 2.7: Signal Significance for various channels (SM Higgs) with integrated luminosity 100 fb\(^{-1}\) [17].

2.3 The top quark

The top quark has envisaged a new window to modern physics. The theoretical motivation for its existence was that the cancellation of anomalies in currents that couple to gauge fields, require complete families [1]. Predicted with a mass close to the electroweak scale, it was expected to play an important role in the electroweak symmetry breaking and therefore in the origin of fermion masses. In addition, its large mass introduces radiative correction in the SM of non-negligible amount. The most recent result for the top mass comes from the Tevatron: 170.9 ± 1.8 GeV/c\(^2\) [12].
2.3 The top quark

2.3.1 Top quark production and decay

Hadron colliders, such as the LHC, are considered major top quark factories. For the LHC at $\sqrt{s} = 14$ TeV, having the cross-section at around 830 pb (QCD calculations, figure 2.6) it is expected to produce 8 million $t\bar{t}$ events at integrated luminosity of $\int L dt = 10 fb^{-1}$. Pair production follows from the following QCD processes (figure 2.8):

- $q\bar{q} \rightarrow t\bar{t}$
- $gg \rightarrow t\bar{t}$

For LHC energies, the contribution according to calculations, is about 13% for the first one and 87% for the second [24].

Fig. 2.8: Lowest order processes for QCD $t\bar{t}$ production.

The top quark, once produced, decays fast through a $W$ boson and a $b$-quark ($t \rightarrow Wb$) [25], [26]. The $W$ can either hadronize by decaying into $q\bar{q}$ pairs which are followed by jets (light jets) or decay into a lepton and a neutrino ($W \rightarrow l\nu$). The $b$-quark also hadronizes to a $B$-meson which has a long enough lifetime before it decays and thus, the produced $b$-jet can be distinguished by its displaced vertex if boosted. The process of selecting the $b$-jet, from the rest of the jets, is called $b$-tagging and it plays a crucial role in modern detectors.

The main decay channels of $t\bar{t}$ are illustrated in figure 2.9. There are three major categories:

- Di-lepton decay, where both $Ws$ decay through a lepton and a neutrino, with a branching ratio of 5%,
• Semi-leptonic decay, were one $W$ decays in a lepton and a neutrino while the other decays in two quarks, with a branching ratio of 30%, and

• Full-hadronic decay, were both the $W$s decay hadronically with a branching ratio of around 45%.

2.3.2 The fully-hadronic channel

The fully-hadronic channel has a final state topology of, six jets, no high-$p_T$ leptons, and small missing $E_T$. The small missing $E_T$ makes it the most kinematically constrained of all the other $t\bar{t}$ topologies as it can be fully reconstructed. The large branching ratio gives a cross-section of 369 pb providing us with around 3.7 million $t\bar{t}$ events in this final state for an integrated luminosity of 10 $fb^{-1}$ (LHC 1 year low luminosity run), making it an interesting channel to investigate.

In figure 2.10, the mass spectra of reconstructed top quarks from fully-hadronic decays and semi-leptonic decays (hadronic-side) are shown. The selection was made by requesting 4 untagged jets and 2 $b$-jets for the fully-hadronic case, and 2 untagged jets and 2 $b$-jets for the semi-leptonic case. The top quark from the fully-hadronic case has a broader distribution with respect to the semi-leptonic case. However, the fully-hadronic top has not introduced any ambiguity due to missing $E_T$ and is sensitive to systematic errors originating from different sources than in the case of the semi-leptonic decay. This information can be used to understand the systematic errors,
improve the resolution of the reconstruction and eventually increase the reconstruction efficiency of the top quark.

![Graphs showing top mass and combinatorial background](image)

Fig. 2.10: Top mass and combinatorial background by comparing three jet invariant mass (MonteCarlo signal samples): fully-hadronic(left), semi-leptonic(right). The low top mass (159 GeV) is probably due to jet energy scale calibration [27].

### 2.3.3 Backgrounds to top-antitop decay

The main background in the $t\bar{t}$ fully-hadronic channel originates from parton showers produced by the QCD $2 \rightarrow 2$ parton processes (e.g. $q_i q_j \rightarrow q_i q_j$, $gg \rightarrow gg$, $gg \rightarrow q\bar{q}$ etc.) with cross-section of the order of $1 \mu b$. Especially for the case of the fully-hadronic decay channel, the QCD multijet background is overwhelming.

Preliminary studies for the ATLAS detector included simple kinematic cuts and requirement for two b-tagged jets, leading to a $S/B \approx 1/57$ [28], [17]. However, with $\chi^2$ constraints the $S/B$ improved to 1/8 for the mass window of 130-200 GeV. Additional increase to the $p_T$ threshold for selecting jets leads to further improvement, e.g. $p_T \geq 25$ GeV gives $S/B \approx 1/6$. This is still very low but the CDF and D0 experiments in Tevatron showed that it is possible to isolate this particular channel with the use of dedicated triggers for these events, $b$-tagging and sophisticated multivariate analysis with Neural Networks [29], [30], [31].

Another major background, but mostly for the semi-leptonic and di-leptonic $t\bar{t}$ decays, comes from the $W/Z +$ jets signal. In the low luminosity phase of the LHC ($10^{33}cm^{-2}s^{-1}$), for every pair of top quarks produced there will be 200 $W$-bosons and 50 $Z$-bosons.
2.3.4 The associated Higgs production

The significance of the top quark decay can be illustrated by considering cases such as the associated Higgs production. Having a cross section of about 1 pb (LHC, $M_H \geq 115$ GeV, figure 2.4) and a distinctive final state it can be considered a possible discovery channel for $M_H \leq 130$ GeV.

The $t\bar{t}H(H \rightarrow b\bar{b})$ channel is considered one of the most interesting ones for the search of a SM Higgs in LHC. With only a small value of $t\bar{t}Z, Z \rightarrow b\bar{b}$ cross-section, with respect to the signal one, it does not compete with background from the $Z$-resonance. Also, by having only the $H \rightarrow b\bar{b}$, the final state of this channel consists of 2 $b$-jets coming from the Higgs, this alone has to deal with an enormous QCD background which turns the signal undetectable, but the associated production with a $t\bar{t}$ pair can significantly increase the signal to background ratio.

As mentioned before, the $t\bar{t}$ pairs can end up in three major final states: the di-leptonic, the semi-leptonic and the fully-hadronic state. For the fully hadronic case, the final state will consist of at least eight reconstructed jets of which 4 jets originate from $b$-quarks and 4 from hadronic decay of the $W$-bosons. A good $b$-tagging efficiency and an excellent calorimeter can provide a very good signal to background ratio for this channel and considering the large branching ratio, the probability of finding the Higgs can be doubled. It is of great importance to be able to detect those events and for this reason an efficient selection procedure must be developed.
The Large Hadron Collider (LHC) is a two-ring, quasi-circular superconducting accelerator located at CERN, the European Organization for Nuclear Research, in Geneva - Switzerland. The project was approved by the CERN council in December 1994. With four collision points, it will provide the necessary environment for detecting “new” physics in the four large detectors (ATLAS, CMS, LHCb and ALICE) that will accompany the machine. The first collisions are expected in medio 2008.

3.1 The LHC machine in a nutshell

The LHC [3] is the most challenging collider that has ever been built. Located at the old LEP ring several meters below the surface and with a circumference of 27 km, it will accelerate bunches of mainly protons with a very high luminosity \(10^{34} \text{cm}^{-2}\text{s}^{-1}\). The protons will be accelerated in opposite directions and will collide with a center-of-mass energy of 14 TeV. This energy opens a whole new window on physics in the Standard Model and beyond. A schematic view of the LHC is show in figure 3.1. At the nominal luminosity a rate of \(10^9\) collisions per second is expected. The bunches of about \(1.15 \times 10^{11}\) protons collide at a cross rate of 40 MHz. This accounts for 23 proton-proton interactions per bunch crossing.

The LHC will make use of the existing accelerator chain (Linac, Booster, PS, SPS) to inject the protons in the main ring. The Linac provides protons of 50 MeV to the Booster, the Booster accelerates those to around 1.4 GeV before they are inserted to the Proton Synchrotron(PS). The Booster and the PS are being upgraded extensively to cope with the new requirements.
for the LHC. Eventually, the PS provides protons of 26 GeV to the Super Proton Synchrotron (SPS) which in turn will insert them with an energy of 450 GeV to the main LHC ring. The insertion will happen in Point 2 (near ALICE detector) for the clock-wise beam and in Point 8 (near the LHCb detector) for the opposite beam.

![Schematic view of the LHC and the supporting accelerators.](image)

The construction of the accelerator is based on that of the LEP collider. However, protons are much heavier than electrons and they need a much stronger magnetic field (around 8.3 Tesla) to keep them into track. For this reason the development of modern superconducting magnets was necessary. Moreover, for the two beams to collide, because they are of the same charge, they must be accelerated in different beam pipes with opposite magnetic fields. The new dipole magnets contain two cylindrical cavities and two coils creating opposite magnetic fields, housed inside a common cryostat. A cross-section of such a magnet is shown in figure 3.2. This is only one of the many technological challenges this project has faced.

### 3.2 Physics goals of the ATLAS experiment

The ATLAS detector (A Toroidal LHC ApparatuS) is a general purpose experiment for recording the properties of particles produced by the proton-proton collisions at the LHC. Constructed in a cavern 100m below the surface
3.2 Physics goals of the ATLAS experiment

it will be a valuable tool for detecting interesting physics events. As a part of the LHC, it has to cope with the enormous amount of data produced from the proton-proton collisions. Eventually, ATLAS will have to deal with a number of physical phenomena, some of which are beyond the SM but theoretically expected. They could be summarized as follows:

- Standard model physics, including top quark physics and b-physics.
- Supersymmetry.
- Higgs boson.
- Exotic phenomena as extra dimensions and microscopic blackholes.

The search for the Higgs boson is closely related with most of the other research areas, such as top-quark physics. As mentioned in chapter 1, research could focus on a number of potential channels for the discovery of the Higgs boson in the SM and the SUSY regime. A look on those decay channels shows, that identification of leptons, photons and jets, at the same time, is very important (figure [2.5]). ATLAS has to take tracking, calorimetry and spectrometry methods into account for effective identification of these events.

In addition, by looking at figure [2.6] one sees that the pp interactions have a large hadronic jet cross-section which dominates the spectrum, while the Higgs boson production has a cross-section significantly lower than that. Clearly, those deep inelastic scattering effects interfere with the potential
discovery of the Higgs boson. As an example, for a Higgs mass around 150 GeV, only one will be produced for every $10^{10}$ pp-interactions. This justifies the high-luminosity LHC wants to achieve but also it suggests the need for a highly restrictive and yet accurate event selection.

3.3 The detector’s layout

The ATLAS detector\footnote{Technical details presented in this chapter, especially for the detector parts, are available in the “ATLAS Detector and Physics Performance” Technical Design Report (TDR)\cite{17} as well as in other specified TDRs.} consists of three detector parts: the Inner Detector, the Electromagnetic and Hadronic Calorimeters, and the Muon spectrometers (figure 3.3). In addition, it can be divide into two major areas for which, in some cases, a different implementation of each type of detector parts exists: the central (or Barrel) region and the forward (or Endcap) regions.

Fig. 3.3: Overview of the ATLAS detector.
3.3 The detector’s layout

3.3.1 The Inner Detector

The Inner Detector is the core of the experiment. It is contained inside a cylinder with a length of 6.80 m and a radius of 1.15 m covering the range $|\eta| \leq 2.5$. Its main purpose is to determine the tracks of the particles created after a $pp$ collision. Furthermore, it plays a role in momentum measurements, vertex recognition and electron identification.

For momentum measurements, the cylinder is covered by a solenoid magnet that creates a magnetic field along the axis of the beamline in order to move the trajectories of the charged particles. Their tracks are then determined by three different sub-detectors at different distances from the interaction point. The three subdetectors, shown in figure 3.4, are: the innermost Pixel detector, the Silicon Tracker (SCT) and the outermost Transition Radiation Tracker (TRT).

![Image of the Inner Detector]

Fig. 3.4: The Inner Detector of ATLAS.

The Pixel detector

Located in the inner side of the Inner Detector, it is designed to provide very-high granularity measurements as close to the interaction point as possible. Its main purpose is to measure the impact parameter resolution and to detect short-lived particles such as B-hadrons and $\tau$-leptons. It consists of three concentric barrels at increasing radii in the central region and of three disks in each of the forward regions (figure 3.5). The innermost barrel layer, the

\[2\] The ATLAS coordinate system is a right-handed system with the $x$-axis pointing to the center of the LHC ring, the $z$-axis following the beam direction with the positive side towards Point 8 and the $y$-axis pointing upwards. The azimuthal angle $\phi = 0$ corresponds to the positive $x$-axis and $\phi$ increases clockwise looking to the positive $z$ direction. The polar angle $\theta$, is measured from the positive axis $z$. In this system, pseudorapidity ($\eta$) is defined by $\eta = -\log (\tan \frac{\theta}{2})$.
B-layer, plays a very important role in secondary vertex measurements but due to its proximity to the beam pipe, the radiation damage is expected to be extensive. For this reason, it is designed to be removable for maintenance.

The pixels of the detector are made of silicon with a size of $50 \times 400 \mu m$. The detector has a high granularity with about 140 million pixels, segmented in modules. Each module contains around 61500 pixels and is read by 16 chips with short-memory buffers to store information during the trigger latency.

![Internal structure of the Pixel detector.](image1)

**The SCT detector**

The SCT surrounds the Pixel detector (figure 3.6). With four concentric barrels and nine wheels at each end it provides full tracking coverage with four spatial points per track. It contributes to the momentum, impact parameter and vertex measurements. The SCT consists of micro-silicon strip detectors of 80 $\mu m$ pitch, based on a $6.4 cm^2$ wedge shaped silicon wafer. In total 16000 wafers are built in 4000 modules. Each module contains two strip detectors and each strip detector contains two silicon wafers. Each module also has associated on-detector readout electronics.

![The Silicon Tracker detector.](image2)
The TRT detector

The TRT is located in the outer shell of the Inner Detector. It is a combined straw tracker and transition radiation detector. Each straw is a small cylindrical drift tube of 4 mm diameter. Straw trackers are able to operate at the very high rates expected at the LHC, because of their small diameter and the isolation of the wires within individual gas volumes. The TRT is designed to contribute to the increase of momentum resolution but mainly to provide good pattern recognition from continuous tracking with 36 two-dimensional points for each track. In addition it will be part of the Level-2 Trigger (section 3.4.1) of ATLAS. The resolution for a charged particle at $|\eta| < 2.5$ and $p_T > 0.5$ GeV is calculated to be better than 0.150 mm.

3.3.2 The Muon Spectrometer

A vital part of the ATLAS detector is the Muon Spectrometer [35], as muon identification is very important in high energy interactions. The muon system of ATLAS has a dual role: it serves both as a trigger system for muons with high energy, and as a spectrometer based on the magnetic deflection of the muons trajectories using an air-core toroid magnet. A cross-section of the left upper quarter of the system is shown in figure 3.7.

Fig. 3.7: Sideview of a quarter of the Muon system.

The magnetic field is created by the barrel and end-cap toroids. In the barrel region, the system consists of three cylindrical shells surrounding the Inner Detector (stations). Precision measurements are made from the Monitored Drift Tubes (MDTs). The trigger system is implemented by means of Resistive Plate Chambers (RPCs), which are mounted on both sides of
the intermediate MDT station and the outer side of the outer MDT station ($|\eta| < 2.4$). In the end-cap regions, the system consists of four disks of MDTs for tracking purposes and Thin Gap Chambers (TGCs) as trigger elements, which are very similar to the RPCs. Only in the very forward directions and close to the interaction point Cathode Strip Chambers (CSCs) are used to withstand the large particle flux. For a detailed description of the different technologies (MDTs, RPCs, TGCs and CSCs) see [35].

3.3.3 The Calorimetry system

The Calorimetry system of ATLAS consists of two components: the Electromagnetic (EM) calorimeter and the Hadronic calorimeter (figure 3.8). The Electromagnetic Calorimeter should identify electrons and photons, and measure their energy content while the Hadronic Calorimeter’s purpose is to identify and measure the energy of jets. In both cases the energy resolution is very important and is given by the direct summation of three terms:

$$\sigma_E / E = k \sqrt{E} \oplus b / E \oplus c\%$$ \hspace{1cm} (3.3.1)

where $k$ is a stochastic term (representing statistical fluctuations, given in $%GeV^{1/2}$), $b$ is a noise term (given in GeV) and $c$ is a constant term which represents uncertainties due to miscalibration, cracks etc. and becomes important for high energies (given in %). Especially for the Electromagnetic LAr calorimeter, the goal is to keep this term below 0.7% [36].

Fig. 3.8: Sideview of the Calorimetry system.
The calorimeters can also give information related to particles that do not get absorbed, such as muons and neutrinos. Muons do not cause showers but they leave an ionizing trace, while neutrinos do not leave any signal but their presence can be determined by transverse momentum conservation. The measured total transverse momentum is usually called the missing transverse energy ($E_T^{\text{miss}}$). In the low luminosity regime, the resolution of $E_T^{\text{miss}}$ was calculated using fully-simulated $\tau\tau$ events and it follows from the relation ([17], page 289):

$$\sigma_{\text{miss}} = 0.46\sqrt{\sum E_T}\text{GeV}$$  \hspace{1cm} (3.3.2)

where $\sum E_T$ is the total transverse energy measured from the calorimeter. In ATLAS, two types of calorimeters are used: the Liquid Argon calorimeter and the Tile calorimeter. Both types are sampling calorimeters, made from more than one substance, the active medium and the passive medium.

**The Liquid Argon calorimeter (LArg)**

The LArg covers the pseudorapidity area of $|\eta| < 4.9$, where the barrel region ($|\eta| < 1.4$) and the end-cap regions ($1.4 < |\eta| < 3.2$) constitute the Electromagnetic calorimeter (ECAL). In the large pseudorapidity region ($3.2 < |\eta| < 4.9$), the forward calorimeter is placed (FCAL).

The ECAL detector has accordion shaped electrodes with Pb(lead) absorber (figure 3.9) and it is expected to provide excellent resolution in energy and position measurements. In addition to that a liquid Argon layer is used for the “presampler”, a layer before the main body of the calorimeter, and at region $|\eta| \leq 1.8$, which allows to compensate for dead material in front of the calorimeter.

The FCAL consists of a metal frame with cylindrical holes where rods are placed, leaving a gap of about 250-400 $\mu$m. The gap is then filled with liquid Argon creating a “tube” of active material. For the EM FCAL the absorber is from Cu while for the Hadronic FCAL is from Tungsten.

Liquid Argon is also used for the Hadronic calorimeter in the end-cap region, where high levels of radiation are expected. For $1.5 < |\eta| < 3.2$ the end-cap hadronic calorimeter (HEC) is using a simpler design (without the accordion shape) than that for the ECAL, having Cu(copper) plates as an absorber, placed perpendicular to the beam axis.

In total, the energy resolution of the Electromagnetic Calorimeter gives 1.3% to 1.5% depending on $\eta$ [37].
The Tile Calorimeter (TileCal)

The TileCal [38] covers the barrel region of the hadronic calorimeter in the region of $|\eta| < 1.7$. It is cylindrical with a length of 6.10 m, and it’s placed outside the LAr at a radius of 4.25 m. It is divided in three major parts: the central cylinder and two cylindrical extensions. It consists of plastic scintillator tiles placed vertically to the beam axis and embedded in an iron absorber. The width of the TileCal in the radial direction, at $|\eta| = 0$, is about nine interaction lengths[3]. The signal by the scintillators is collected from photomultiplier tubes (PMTs) located outside the calorimeter.

---

[3] The mean free path of a particle before undergoing an interaction that is neither elastic nor quasi-elastic (diffractive), in a given medium, usually designated by $\lambda$. 

---

Fig. 3.9: The accordion shaped ECAL.
The TileCal plays an important role in the detection and energy determination of jets. It has a very good resolution in measuring the jet energy as well as the missing transverse energy [37]:

\[
\frac{\sigma}{E_T} = \frac{50\%}{\sqrt{E(\text{GeV})}} \oplus 3\% \text{, for } |\eta| < 3 \text{, TileCal and HEC.}
\]

\[
\frac{\sigma}{E_T} = \frac{100\%}{\sqrt{E(\text{GeV})}} \oplus 10\% \text{, for } 3 < |\eta| < 5 \text{, Hadronic FCAL.}
\] (3.3.3)

The calorimeters (including the presamplers) are divided into cells except for the FCAL which consist of tubes. The barrel ECAL has in total 110000 cells, the endcap ECAL has a total of 64000 cells while the EM FCAL consists of 12000 tubes each being a readout channel. Thus, in total around 200000 channels are read from the Electromagnetic calorimeters. The barrel region of the Hadronic Calorimeter, the TileCal, has a total number of 10000 readout cells while the Hadronic End-cap contains a total of 4000 cells. The Hadronic FCAL has a total of 18000 tubes [37].

### 3.4 The Trigger and Data Acquisition system of ATLAS

With an interaction rate at around 1 GHz the detector will eventually have to deal with an incredible amount of data. However, a large part of those interactions is of no interest for the physics purposes of ATLAS. In addition, technical but also budget limitations suggest that it is impossible to store all the information produced in the collisions. For this reason, ATLAS implements an efficient online selection system (Trigger) which will be able to reduce the incoming data rate to the specified limits [39].

The online reconstruction and selection capability of the calorimeter is of great importance specifically for the $t\bar{t}$ decay channels. In particular, the fully-hadronic channel, since it involves only jets, is heavily affected by the performance and the limitations of the jet trigger. The trigger menu has to be optimized carefully, especially when the large QCD multijet background is taken into account.

#### 3.4.1 Architecture of the three levels Trigger

The Trigger system consists of three levels, the Level 1 (LVL1), the Level 2 (LVL2) and the Event Filter (EF). The LVL1 is a hardware implemented trigger while the LVL2 and the EF, commonly referred as High-Level Trigger (HLT), are software based. The DAQ system must provide event data from...
the LVL1 to the LVL2 and to the EF and eventually make them available to the mass storage system.

In figure 3.10 an overview of the Trigger system is shown. The initial 40 MHz bunch crossing rate has to be reduced to 200 Hz due to limited storage capacity. At this rate, events of about 1.5 Mb in size will be written to the mass storage system. This accounts to only 0.001% of the initial information size but still it gives around 1 Petabyte in 1 year of data collection.

![Trigger system overview](image)

**Fig. 3.10: Schematic of the ATLAS Trigger system.**

The LVL1 Trigger

The LVL1 trigger \cite{40, 41} is the first selection layer responsible for the reduction of the bunch crossing rate to the input rate of the LVL2 trigger. It consists of roughly four parts: the muon trigger, the calorimeter trigger, the Central Trigger Processor (CTP) and the Timing, Trigger and Control (TTC) system (figure 3.11). The Inner Detector is not used at LVL1 due to its complexity. The LVL1 uses special-purpose processors which act only on reduced granularity data from the muon and calorimeter triggers, due to time limitations, which are handled by the CTP. The CTP makes a decision which is then passed to the TTC system which distributes it to the Front-End electronics. The latency time of the LVL1 is around 2 $\mu$s while the
design limit is implemented around 2.5 $\mu$s, leaving a safety margin of 0.5 $\mu$s. The bunch-crossing time however is around 25 ns, a lot less than the LVL1 latency. For this reason, during the LVL1 processing time, the information from multiple bunch-crossings is stored in pipeline memories of \( \frac{2.5 \, \mu s}{25 \, \text{ns}} = 100 \) cells deep.

The maximum output rate for LVL1 is limited to 100 kHz due to bandwidth limitations as well as LVL2 specifications.

The LVL2 Trigger

The LVL2 trigger works on the output rate of LVL1. It uses information from the Inner Detector in addition to full-granularity data from the muon and calorimeter trigger. However, it does not process the full event but only those regions of the detector that the LVL1 has identified as containing interesting parts of an event. These regions are called the Regions of Interest (RoIs) and are constructed by the RoI Builder (see figure 3.11). Thus, the LVL2 is guided (or “seeded”) by the geometrical RoI information. In figure 3.12 an overview of LVL2 is shown.

If the LVL1 accepts an event then the data are pushed from the pipeline memories to the Readout drivers (RODs), located off detector. At this point, a LVL1 ID number and bunch crossing information are added. The data are then sent via Readout Links (ROLs) to Readout Buffers (ROBs) housed in
the Readout Subsystem (ROS, process 1a). The data remain stored in the ROBs until a decision is made by the LVL2 trigger system.

In parallel, the LVL1 sends the RoI information to the RoI Builder (process 1b). The RoI Builder is processing the LVL1 information and sends a RoI description to the LVL2 Supervisor (L2SV). The L2SV assigns the event to one of the processors of the LVL2 farm (process 2), the LVL2 Processing Unit (L2PU). The L2PU acts as an interface to retrieve any additional information from the ROS when algorithms request it (process 3). The decision of the LVL2 is sent back to the L2SV (reverse process 2) which forwards it to the Dataflow Manager (DFM, not shown in the picture).

![Fig. 3.12: The LVL2 Architecture.](image)

If an event is rejected, the DFM requests the ROS to delete the data. If the event is accepted then the event is assigned to an Event Builder (EB) node (process 4). The EB acts as an intermediate layer between the LVL2 and the EF. It works based on the LVL2 messaging system and its task is to fully build an event (full ATLAS event) from all the available ROS segments. When the event is built, it is passed to an Event Filter Sub-Farm through an SFI (Sub-Farm Input). The latency time of the LVL2 trigger is from 1 to 10 ms with an accept rate in the order of 1 kHz.

**The Event Filter**

The Event Filter works on the output rate of the LVL2. In contrast to LVL2 it processes the full ATLAS event in order to reach a decision. In principle, the EF is more properly defined as a filter as it is not limited by response
3.4 The Trigger and Data Acquisition system of ATLAS

3.4.1 The Trigger and Data Acquisition system of ATLAS

The Trigger and Data Acquisition system of ATLAS consists of a Trigger, where events are generated, and a Data Acquisition system, where the full event is stored on disk and reconstructed with offline algorithms. Accepted events are written to mass storage space, rejected events are deleted.

Once the SFI receives the full event it passes it to an EF sub-farm. The sub-farm consists of three major parts: the Distributor which receives events from the SFI, the Processing Tasks which perform selection and any other transformation on the event, and the Collector which injects selected events into the mass storage. Each of these parts is itself divided into several other pieces (figure 3.13). D1 performs a sorting operation and stores the event on the local disk. D2 further distributes the event according to the event type. Additionally, D3 passes the event to different processing tasks. The Processing Task makes the actual algorithmic selection using the offline reconstruction algorithms including vertex reconstruction and track fitting. C1 collects the event in a process symmetrical to that of D3. Finally, C2 collects all the events from C1 and pushes them to a mass storage device.

In general a seeded – from LVL2 information – reconstruction is most suitable for EF, but full event reconstruction is also possible within the specified limits. The EF reduces the event rate down to 200 Hz while its event processing time is of the order of 1 - 10 seconds. Both the LVL2 and EF selection software make use of software algorithms written in C++ and use the ATHENA offline analysis framework of ATLAS.

3.4.2 The Calorimeter and the Jet trigger

As mentioned in section 3.3.3, the calorimeters consist of cells. Clusters of cells form the building blocks of the calorimeter reconstructed objects. The input to the LVL1 algorithms are a set of Trigger Towers (TTs) which are formed by analog summation of the calorimeter cells on a straight line from the interaction point. TTs have a fixed granularity of 0.1 \( \times \) 0.1 (in terms of \( \Delta \eta \times \Delta \phi \)).

Jet reconstruction in the trigger is based on “jet elements”. Jet elements have two types: the trigger towers(TTs), and the topological clusters. Topological clusters do not have a fixed size but their borders are based on weighted energy, position and shape variables [17]. The jet trigger only discriminates based on the \( E_T \) and multiplicity of jets above a certain \( E_T \) threshold. Analogue signals from the TTs are sent to the front-end pre-processor which calculate the \( E_T \) by summing the TTs. The RoIs which contain jet

---

4Filtering algorithms taken from the offline reconstruction. Note that calibration of these algorithms is different in the online system.
clusters are flagged for analysis by the LVL2. In total eight programmable sets of thresholds are provided for each calorimeter sub-trigger.

The jet trigger algorithm, is based on a “window” of jet elements. Jet elements have a granularity of $0.2 \times 0.2$ in terms of $\Delta \eta \times \Delta \phi$ ($2 \times 2$ in terms of TT) and are summed in depth between the EM and the Hadronic calorimeters. The algorithm consists of two components, a $2 \times 2$-element cluster ($\Delta \eta \times \Delta \phi = 0.4 \times 0.4$) which is used to identify the position of candidate jet RoIs based on local $E_T$ maximum, and a trigger cluster which is used to measure the jet $E_T$. The trigger cluster slides in steps of 0.2 (or one element at a time) in both $\eta$ and $\phi$ directions. This is called the sliding window technique, figure 3.14. The trigger cluster varies and can be a $2 \times 2$, $3 \times 3$ or $4 \times 4$ jet elements, which in terms of $\Delta \eta \times \Delta \phi$ translates to $0.4 \times 0.4$, $0.6 \times 0.6$ or $0.8 \times 0.8$ respectively. This works only for the region where $|\eta| \leq 3.2$ which is the limit of the end-cap acceptance.

In the HLT the jet trigger mainly tries to reduce the number of events containing jets without throwing away the interesting ones. This is achieved with a more efficient energy calibration and jet definition, and by using the full-granularity of the detector. The LVL2 jet finding is guided by the RoIs selected from the LVL1 while the EF looks across the whole $\eta$ region with better jet calibration.

Two types of reconstruction algorithms may be used: Cone and $k_T$. Cone is based on geometrical arguments in order to define a jet, while $k_T$ is based
on fragmentation parameters. In principle, the Cone algorithm merges all the “jet elements” (particles) that follow the relation:

\[
\sqrt{(\eta_i - \eta_j)^2 + (\phi_i - \phi_j)^2} < R
\]  

(3.4.1)

where \((\eta_i, \phi_i)\) are the angular coordinates of the particle, \((\eta_j, \phi_j)\) the coordinates of the center of the cone, typically defined as the jet axis, and \(R\) the predefined parameter which is either 0.7 or 0.4. On the other hand, for the \(k_T\) each particle\((i)\) is defined with a parameter \(d_{ii} = (E^{2}_{T,i})\), also for each particle pair\((i,j)\) the following parameters are defined:

\[
\Delta R_{ij} = \sqrt{(\Delta \phi_{ij})^2 + (\Delta \eta_{ij})^2}
\]

\[
d_{ij} = \min \left[ E^{2}_{T,i}, E^{2}_{T,j} \right] \frac{(\Delta R_{ij})^2}{D^2}
\]  

(3.4.2)

where \(D\) is the predefined parameter which typically is chosen as 0.6 or 0.4. If the minimum from all the \(d_{ii}\) and \(d_{ij}\) is \(d_{ij}\) then the pair of particles are combined, else \(i\) is considered a jet. A schematic view of the Cone and \(k_T\) algorithms is in figure 3.15 showing how the jet constituents are combined. The standard algorithm for ATLAS is Cone.

The online reconstruction algorithms are part of a large number of HLT algorithms which are used by the trigger. All those HLT algorithms are

\footnote{ATLAS is using “fastKt” an implementation of \(k_T\) algorithm based on the \(k_T\) finder algorithm from [43] and the “fastjet” idea from [44] which improves the computing time.}
based on two principles: being “seeded” algorithms – the selection procedure is “seeded” by the results of the preceding trigger level –, and performing on a “step-wise” procedure where at each step the hypothesis is evaluated and is rejected if not confirmed (“early-reject” concept). The main advantage of these two principles is that they allow fast and light-weight HLT decisions[45]. They can be divided into two major classes:

- The Feature Extraction Algorithms(FEX).

- The Hypothesis Algorithms(HYPO).

The FEX algorithms process either raw data directly from the detector or data received from a previous algorithm. Next, the extracted features are used by the HYPO algorithms to make a decision for a certain trigger element.

---

6A trigger element can be considered as everything that can be reconstructed by an algorithm and can lead to the acceptance or rejection of a specific event.
The trigger selection efficiency of the fully-hadronic $t\bar{t}$ events is directly correlated to the quality of the jet reconstruction in the online system. In this chapter, the performance of the jet reconstruction is investigated based on fully-hadronic $t\bar{t}$ events.

4.1 Analysis Data Object - AOD

The analysis presented in this chapter is based on MonteCarlo simulated samples. These samples are called Analysis Objects Data (AODs) and their production follows a chain of steps. It starts with generating particle four-vectors from specified physics processes. The resulting particles are then inserted in a GEANT4 simulation of the ATLAS detector and produce “hits” in the detector. The hits are then translated to “digits” (times and voltages) to correspond to the actual “raw” data objects (RDO) that the detector will produce. Next, the RDOs are processed and the digits are reconstructed into tracks and energy depositions. The result is the Event Summary Data (ESD). Finally, the AOD is built and corresponds to a summary of the reconstructed information of the ESD.

For the analysis 6 different AOD samples were used, all of them from the official ATLAS production. These are: 8500 events of fully-hadronic $t\bar{t}$ decays (denoted as 5204), 8500 events of semi(di)-leptonic $t\bar{t}$ decays (5200), 5000 events from the QCD 3 jets exclusive sample (5061), 4900 events from the QCD 4 jets exclusive sample (5062), 4750 events from the QCD 5 jets exclusive sample (5063) and 4250 events from the QCD 6 jets inclusive sample (5064). All of the samples were reconstructed AODs processed using Atlas.
4. Online selection of fully-hadronic top pair decays

Offline Release (ATHENA) 12.0.6 (v120006) and with the geometry ATLAS-CSC-01-02-00 which introduces misaligned geometry with material distortion (misal1), for detailed description see [46]. Also, they have been reconstructed with the trigger activated (trig1) using the CSC-06 Trigger configuration (see [47]). The relevant items for the investigation of the jet trigger are: jet160, 2jet120, 3jet65, 4jet50 where the name gives the energy threshold of the signature and the number of jets required to be above this threshold in order for the signature to fire (multiplicity), e.g. jet160 requires one jet with threshold above 160 GeV, 2jet120 requires two jets with 120 GeV or more etc.

The 5204 and 5200 sample have been used for the study of the signal behavior in the jet trigger. The 5061, 5062, 5063 and 5064 samples have been used for studying the background rate with respect to the signal efficiency. Unfortunately, they are not the best to be used for this study. That is because they contain rather tight cuts in the generation of the jets and so they do not “simulate” exactly the expected QCD background. However, they are the only ones available in the current ATLAS production. The optimal solution for studies in this channel would be additional QCD multijet samples produced with a smaller cut in the leading jet $p_T$ in order to avoid a bias in the acceptance of the low threshold triggers.

4.2 Analysis tools

The tools used in the analysis are the following:

- The ATHENA offline analysis framework of ATLAS (release 12.0.6) [48].
- The EventView Analysis package [49].
- The TriggerRateTools package [50].

ATHENA

The ATHENA framework is the baseline of the ATLAS offline analysis. It has a component-based architecture, specifically designed for physics data-processing applications. It supports modular analysis and provides a common analysis skeleton based on standard tools. ATHENA is able to process the standard ATLAS analysis data sets (Analysis Object Data - AOD) and will eventually play an important role for the validation of official ATLAS results before publication.

\[ \text{The leading jet at } p_T > 80 \text{ GeV and 4 jets with } p_T > 40 \text{ GeV.} \]
EventView

EventView (EV) is an analysis package built on top of the ATHENA framework. Analysis with EventView is based on a set of modular pieces of code (“tools”) which perform specific analysis tasks and do not interfere with each other. The tools are generic and configurable enough to be reusable according to specific needs. Additionally, personal tools for more explicit analysis can be built on top of it. Several ATLAS physics groups have adopted the EV framework to build packages with tools related to the physics requirements of the group e.g. TopView for top physics, SUSYView for Supersymmetry physics, EV Trigger for trigger aware analysis, and more.

For the analysis in this thesis, a small privately developed package is created (MeniosView) with additional analysis tools built on top of the EventView tools (including those from the sub-packages). This allows integration of the latest stable EV tools directly from the groupArea (a pre-compiled area of EV) with the personally developed components. In the present analysis except from the generic EV tools, also other tools were used from the EV Trigger package as well as from the TopView package.

TriggerRates Tool

The TriggerRates Tool is a small package containing default tools for calculating the trigger rates of the various signatures.

4.3 The Jet Trigger Slice

A Trigger Slice defines the processing sequence for the selection of events in the online system based on the presence of an interesting physics related feature, such as high energy jets or leptons, missing $E_T$, etc. The trigger slice contains the selection criteria of all trigger levels (LVL1, LVL2 and EF). The Jet Trigger Slice is shown in figure 4.1.

The LVL1 identifies the jet RoIs (JETROI) and passes this information to the LVL2 (see sections 2.4.1 and 2.4.2). The LVL2 is using FEX algorithms and reconstructs the RoIs provided by LVL1 (T2CaloJet). By default the reconstruction algorithm in LVL2 is Cone. The result of the FEX is then handled by the HYPO algorithm (TrigL2JetHypo) which decides, based on energy and multiplicity cuts, to accept or reject the event.

The EF runs a more refined reconstruction based on several FEX algorithms, on events accepted by LVL2. The TrigCaloCellMaker unpacks the data from the calorimeter cells which are then used from TrigCaloTowerMaker to build the trigger towers, TrigJetRec is performing the reconstruc-
4. Online selection of fully-hadronic top pair decays

Fig. 4.1: The Jet Trigger Slice.

Each item corresponds to a sequence of signatures. These are shown in table 4.1.

<table>
<thead>
<tr>
<th>Trigger item name</th>
<th>LVL1 signature</th>
<th>LVL2 signature</th>
<th>EF signature</th>
</tr>
</thead>
<tbody>
<tr>
<td>jet160</td>
<td>L1_J45</td>
<td>L2_jet160</td>
<td>EF_jet160</td>
</tr>
<tr>
<td>2jet120</td>
<td>L1_2J45</td>
<td>L2_2jet120</td>
<td>EF_2jet120</td>
</tr>
<tr>
<td>3jet65</td>
<td>L1_3J45</td>
<td>L2_3jet65</td>
<td>EF_3jet65</td>
</tr>
<tr>
<td>4jet50</td>
<td>L1_4J45</td>
<td>L2_4jet50</td>
<td>EF_4jet50</td>
</tr>
</tbody>
</table>

Table 4.1: Signature sequence of trigger items.

Each signature holds the name of the trigger level to which it corresponds (L1, L2 or EF), as well as the multiplicity of jets required to pass the hypothesis (2jet, 3jet etc.), the last number defines the energy threshold, set for each jet.

4.4 Jet Trigger quality based on fully-hadronic events

The quality of the Jet Trigger Slice can be studied with the behavior of individual trigger signatures and their efficiency at different jet energies. These plots, showing the trigger efficiency as a function of energy, are called “turn-on” curves. At this part of the analysis this quality study is made with respect to the true energy of a jets, this is possible with MonteCarlo data, but with real data this analysis has to use offline reconstructed information. This analysis is considered as a validation study for the trigger, which makes a comparison to the truth energy information relevant.

There are many ways to construct turn-on curves. The relation between the trigger efficiency and energy is affected by the method used. Here, the
construction of the turn-on curve for the jet triggers starts with the identification of “trigger objects” and “truth jets”.

- **Trigger “objects”**: Trigger objects are created by FEX algorithms. They contain information about the physics related features such as momentum, angular coordinates etc. In our case we are interested in the jet trigger objects.

- **“Truth” jets**: A truth jet refers to the actual information content of the jet as this was deposited on the detector. The input truth jets for this analysis originate from Cone4TruthParticleJets. These are jets formed by using the Cone4 algorithm on the generated truth particles (Generation level) in order to form what it would be seen in reality as a jet. The reconstruction is only considering the truth particles after fragmentation and hadronization.

### 4.4.1 Matching procedure

The matching between the truth jets and the trigger objects plays an important role in this analysis. The tool for performing this association is part of the EventView Trigger package. By default it is iterating through the truth jets and relates it to the closest, in terms of $\Delta R$, trigger object. The $\Delta R$ is defined by the relation:

$$
\Delta R = \sqrt{(\Delta \eta)^2 + (\Delta \phi)^2}
$$

(4.4.1)

where:

$$
\Delta \eta = \eta_{\text{truth}} - \eta_{\text{trigger}}
$$

$$
\Delta \phi = \phi_{\text{truth}} - \phi_{\text{trigger}}
$$

(4.4.2)

The default cut in the matching modules is a very loose one, requiring a matching within a $\Delta R \leq 1$. In this analysis it is replaced by $\Delta R \leq 0.3$ to assure that the truth jet, reconstructed with Cone4 ($\Delta R = 0.4$), will overlap with the associated trigger object, also reconstructed with Cone4.

After the matching procedure one or more trigger objects in an event can be unmatched. These objects may fire the trigger, but since no truth jet has been assigned to them they cannot be taken into account for the efficiency calculation as a function of true jet energy. For this reason, events that do not contain the necessary number of matched trigger objects with energy above the threshold (in our hypothetical case 1 object with more than 50 GeV), are excluded. The unmatched jets will be discussed in a later stage.
4. Online selection of fully-hadronic top pair decays

4.4.2 Analysis scheme

To illustrate the applied methodology let’s consider a hypothetical signature with threshold at 50 GeV and multiplicity of one jet. For the HYPO algorithm to accept the event there must be at least one trigger object with energy above the threshold.

The first step is to select only those events that passed any preceding trigger level. The next step is to associate the truth with the trigger objects (matching procedure). The result is a collection of “matched” truth jets. The trigger objects that were not associated to a truth jet are unmatched. These are ignored for now but they are discussed later on.

We can now plot two distributions from the events that fulfill the requirements of a specific signature only based on trigger objects that were matched to a truth jet. The first distribution contains all the matched truth jets (denoted as distA), and the second one contains the truth matched to a trigger objects that passed the threshold (denoted as distB). Dividing \( \frac{\text{distB}}{\text{distA}} \) and plotting it with respect to the truth energy gives the turn-on curve. In this particular case, this translates into the percentage of jets that pass the specific energy threshold of the trigger (efficiency) as a function of their true energy. Ideally, the turn-on curve would be a step function as shown in (figure 4.2).

Based on such a turn-on curve one would conclude that there are no truth jets with energy below 50 GeV that pass this specific threshold whilst truth jets with energies above 50 GeV are always accepted. In this hypothetical case, most events of interest contain at least one jet above 50 GeV, and the rest is background. Hence, such a trigger is very efficient in the selection of interesting events. Unfortunately, the shape of the turn-on curve is affected by many different experimental effects. In practice, the turn-on curve will be S-shaped, as is shown in the next section.

It is important to note that in the above method all the jets are treated on an equal footing. For fully-hadronic \( t \bar{t} \) events, this assumption is not entirely correct. These events contain 2 \( b \)-jets with slightly different energy scales and a slightly different jet shape. Nevertheless, these effects are small enough to ignore in the trigger level reconstruction algorithms.

4.4.3 Jet signatures and trigger acceptance

Using the method described above, the jet160, 2jet120, 3jet65 and 4jet50 turn-on curves for each trigger level are constructed (figures 4.3, 4.4, 4.5 and 4.6). The error bars on the \( x \) – axis correspond to the bin width and are equal to \( \pm 6 \) GeV. The error on the \( y \) – axis is originating from statistics
4.4 Jet Trigger quality based on fully-hadronic events

Fig. 4.2: An ideal turn-on curve should look as a step function.

and is propagated from the number of entries in the current bin of \textit{distA} and \textit{distB}. This error calculation is not correct. In fact, the content of \textit{distA} and \textit{distB} is correlated. This correlation has to be taken into account.

The figures show that the threshold for LVL2 is slightly lower than the threshold for the EF. This is an expected behavior as the jet energy scale corrections and reconstruction in LVL2 should be less accurate that in EF. This also suggests that the expected shape of the EF should be steeper than that of LVL2. Surprisingly, this is not the case. In figure 4.3 the LVL2 starts building-up at around 120 GeV and it reaches a plateau at about 180-190 GeV, indicating a top-to-bottom difference of 60-70 GeV. The EF begins from about 130 GeV and reaches a plateau at the energy of 210 GeV, having a difference of 80 GeV. Also, at the threshold energy (160 GeV) the LVL2 has reached a 70% efficiency and the EF is at 30%, while for high energy jets the LVL2 has an efficiency of 95% and the EF 90%. The reason of this behaviour might be related to an erroneous in the jet energy calculation.

The actual size of the JETROI is smaller than the Cone and $k_T$ reconstruction area, covering only about 70% of the total. Hits on the detector that do not belong to the RoI are not taken into account for the jet energy calculation. For the Cone4, running in the EF, this leads to reconstructing only 70% of the same energy as the same algorithm would reconstruct offline [51].

The second graph (figure 4.4) shows an improvement as both triggers seem to start developing at around the same energy (90 GeV). However, the EF still experiences a much slower build-up due to the problem in the jet energy calculation and at around the threshold energy (120 GeV) the LVL2
is reaching 80% efficiency while the EF remains at 50 – 60%. Also, the steepness of the curves is improved with the LVL2 reaching a plateau at 140 GeV, for a difference from top-to-bottom of about 50 GeV, and the EF at 150 GeV, a difference of 60 GeV. This is an indication that the jet energy scale calibration improves for jets at lower energies.

For the 3jet65 (figure 4.5) the efficiency of LVL2 at threshold energy (65
Fig. 4.5: The 3jet65 turn-on curves of LVL2 and EF.

Fig. 4.6: The 4jet50 turn-on curves of LVL2 and EF.

GeV) is almost 65% while the EF has about 25%. The steepness gives a difference of 40 GeV for LVL2 and 50 GeV for EF as both start building from about 50 GeV. In the 4jet50 case the LVL2 reaches roughly 70 – 80% and the EF 35 – 45% at threshold energy (50 GeV). The steepness indicate a difference of top-to-bottom of 30-40 GeV for the LVL2, and 40-50 GeV for the EF.
However, the EF seems to experience a disturbing misbehavior with a number of low energy jets passing the trigger. The indicated $10 - 15\%$ efficiency of accepting low energy jets using these signatures may result to extremely high rates due to the increased number of jets expected below $10$ GeV. This effect is reported and still under investigation.

As a next step, the trigger acceptance for the fully-hadronic $\bar{t}t$ events is retrieved. The results of the various signatures are shown in table 4.2 with respect to the initial $8500$ events, for completeness the LVL1 acceptance is shown as well. The overall acceptance for LVL2 and EF was calculated by counting the events that had at least one of the defined signatures fired.

<table>
<thead>
<tr>
<th></th>
<th>jet160</th>
<th>2jet120</th>
<th>3jet65</th>
<th>4jet50</th>
<th>Overall</th>
</tr>
</thead>
<tbody>
<tr>
<td>LVL1</td>
<td>99.9%</td>
<td>99.8%</td>
<td>97.8%</td>
<td>87.6%</td>
<td>99.9%</td>
</tr>
<tr>
<td>LVL2</td>
<td>19.9%</td>
<td>18.8%</td>
<td>40.9%</td>
<td>39.7%</td>
<td>53.2%</td>
</tr>
<tr>
<td>EF</td>
<td>12.7%</td>
<td>11.1%</td>
<td>24.0%</td>
<td>20.5%</td>
<td>32.4%</td>
</tr>
</tbody>
</table>

Table 4.2: Signature acceptance of fully-hadronic $\bar{t}t$ events from a total of $8500$ events sample.

Unmatched jets

The acceptance of the trigger does not imply that all of the trigger objects correspond to a truth jet. For this reason, we examine two cases to study the behavior of the trigger in detail:

- **A**: Assuming that the highest energetic jet trigger objects fired a specific signature i.e. the highest for the jet160, the two highest for the 2jet120 etc. Each of these have a match to a truth jet. The number of events that passed a signature with at least one of those objects without a match to a truth jet, is calculated.

- **B**: Assuming that any trigger object matched to a truth jet could have fired the trigger. The number of the events that passed a specific signature but also fulfill the requirements of multiplicity and threshold based only on the matched objects, is calculated.

For case A and B the results are shown in table 4.3.

Consider the case of the EF_3jet65 trigger as an example, it has an acceptance of $24\%$ (2040 events from the total of 8500 events). Table A shows that $97.1\%$ of the 2040 events, fired based on at least one jet without a truth match. This is a worrying effect, since it means that the trigger was fired
### 4.4 Jet Trigger Quality Based on Fully-Hadronic Events

<table>
<thead>
<tr>
<th></th>
<th>jet160</th>
<th>2jet120</th>
<th>3jet65</th>
<th>4jet50</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>1.2%</td>
<td>5.4%</td>
<td>9.2%</td>
<td>13.5%</td>
</tr>
<tr>
<td>LVL2</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>53.3%</td>
<td>96.6%</td>
<td>97.1%</td>
<td>99.2%</td>
</tr>
<tr>
<td>B</td>
<td>99.9%</td>
<td>97.8%</td>
<td>96%</td>
<td>93.4%</td>
</tr>
<tr>
<td>LVL2</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>99.1%</td>
<td>96.8%</td>
<td>93.2%</td>
<td>89.1%</td>
</tr>
</tbody>
</table>

Table 4.3: Percentage of events that passed the trigger: with at least one unmatched highest energetic object (case A), which fulfill the requirement of multiplicity and threshold with only matched objects (case B).

by an object that is not understood. However, Table B shows that 93.2% of the 2040 events would have been triggered anyway, even if only matched jets are taken into account (case B). Apparently, there is an enormous overlap between the two cases which led to the identification of yet another problem in the EF jet trigger reconstruction.

In the EF both the $k_T$ and Cone algorithms are used. This allows studies on the performance of the algorithms at EF level. The trigger objects from both algorithms are registered in the same collection. This was verified by a fast calculation on the EF trigger objects that had an overlap within a $\Delta R \leq 0.1$ with other EF trigger objects in the same event. From a total of 93672 objects, 46719 overlapping pairs were found. This accounts to 49.9% of the total, indicating that indeed the EF collection is filled with entries from two different algorithms reconstructing the same things. This does not affect the final Trigger Decision however (figure 4.2), as for that only the Cone algorithm is taken into account.

The number of events triggered by unmatched objects is much smaller for the LVL2 trigger. Nevertheless, this number is significant and should be understood. Especially because the EF problem of jet duplication is not present at LVL2. Eventually, it turned out that this effect was caused by a problem in LVL1. Although very different in nature from the previously discussed problem, the effects are similar.

At LVL1, cases exist where a jet splits into two close JETROIs. This happens because the LVL1 RoI is too small to contain the entire jet and thus two RoIs are created to cover the whole region. These RoIs are very close to each other and this affects the HLT because the Cone algorithm is wider than the RoI. As a result it creates two times the same jet. The $k_T$ behaves similarly. Studies of this effect showed that around 10% of the events contain this type of pairs. This accounts to 5–10% increase in the jet rate for LVL2 and 3% increase for the EF [52]. In the matching procedure this leads to
only one of the two jets being matched to a truth while the other will remain unmatched.

4.4.4 Background study with QCD multijet events

Due to the reasons mentioned above, our understanding of the EF jet trigger is limited. However, we can still do a feasibility study, based on the Trigger Decision of the various signatures, by correlating their acceptances to the expected background.

As already mentioned, the major background for the fully-hadronic $t\bar{t}$ events originates from QCD multijet events. Unfortunately, not many background samples relevant to this channel were produced. For this reason, only the 4 samples, 5061, 5062, 5063 and 5064 were chosen. Their specifications are already mentioned in section 4.1.

To evaluate the effect of background, first we calculate the trigger acceptance for the various signatures of LVL2 and EF and then the event rate for accepting these events. This calculations were done using the TriggerRates Tool[50]. The rate was calculated using the following formula:

$$Rate(S) = R_s = \varepsilon_s \times L \times \sigma \times \frac{1}{P_s}$$

(4.4.3)

where $S$ is the given signature, $\varepsilon_s$ its efficiency, $L$ the instantaneous luminosity in $cm^{-2}s^{-1}$, $\sigma$ the cross-section for the given interaction in $cm^2$, and $P_s$ the prescale factor for the given signature. The resulting rate is in $s^{-1} = Hz$. In this calculation the low luminosity regime was examined ($L = 10^{33} cm^{-2}s^{-1}$).

The cross-sections used are noted in table 4.4. The results for the rates and acceptances are in tables 4.5 and 4.6.

<table>
<thead>
<tr>
<th>Sample</th>
<th>Cross-section (pb)</th>
</tr>
</thead>
<tbody>
<tr>
<td>5204 (signal)</td>
<td>369</td>
</tr>
<tr>
<td>5061</td>
<td>21200</td>
</tr>
<tr>
<td>5062</td>
<td>53300</td>
</tr>
<tr>
<td>5063</td>
<td>9900</td>
</tr>
<tr>
<td>5064</td>
<td>6440</td>
</tr>
</tbody>
</table>

Table 4.4: Cross-sections of the different samples used.

\[\text{The prescale factor is the number of the events that must be accepted for a given signature to fire once. For example a signature with a prescale factor of 2 will accept only one in every two events that pass its multiplicity and threshold requirements. The prescale factor for our signatures of interest are 1.}\]
4.4 Jet Trigger quality based on fully-hadronic events

<table>
<thead>
<tr>
<th>LVL2</th>
<th>L2_jet160</th>
<th>L2_2jet120</th>
<th>L2_3jet65</th>
<th>L2_4jet50</th>
<th>Overall</th>
</tr>
</thead>
<tbody>
<tr>
<td>5204 Rate(Hz)</td>
<td>0.0736</td>
<td>0.0692</td>
<td>0.151</td>
<td>0.147</td>
<td>0.196</td>
</tr>
<tr>
<td>5061 acceptance Rate(Hz)</td>
<td>20.8%</td>
<td>16.2%</td>
<td>27.6%</td>
<td>18.7%</td>
<td>43.4%</td>
</tr>
<tr>
<td>5062 acceptance Rate(Hz)</td>
<td>4.40</td>
<td>3.43</td>
<td>5.86</td>
<td>3.97</td>
<td>9.19</td>
</tr>
<tr>
<td>5063 acceptance Rate(Hz)</td>
<td>21.7%</td>
<td>17.7%</td>
<td>34%</td>
<td>27.6%</td>
<td>50.6</td>
</tr>
<tr>
<td>5064 acceptance Rate(Hz)</td>
<td>32.8%</td>
<td>25.5%</td>
<td>48.5%</td>
<td>48.5%</td>
<td>67.5%</td>
</tr>
</tbody>
</table>

Table 4.5: LVL2 trigger Rates and acceptances for the signal and the QCD background samples. The acceptance of 5204 is shown in table 4.2.

<table>
<thead>
<tr>
<th>EF</th>
<th>EF_jet160</th>
<th>EF_2jet120</th>
<th>EF_3jet65</th>
<th>EF_4jet50</th>
<th>Overall</th>
</tr>
</thead>
<tbody>
<tr>
<td>5204 Rate(Hz)</td>
<td>0.0468</td>
<td>0.0411</td>
<td>0.0884</td>
<td>0.0758</td>
<td>0.120</td>
</tr>
<tr>
<td>5061 acceptance Rate(Hz)</td>
<td>12.8%</td>
<td>9.3%</td>
<td>10.8%</td>
<td>5.54%</td>
<td>21.7%</td>
</tr>
<tr>
<td>5062 acceptance Rate(Hz)</td>
<td>2.72</td>
<td>1.96</td>
<td>2.28</td>
<td>1.17</td>
<td>4.60</td>
</tr>
<tr>
<td>5063 acceptance Rate(Hz)</td>
<td>13.1%</td>
<td>9.4%</td>
<td>14.5%</td>
<td>7.9%</td>
<td>24.7</td>
</tr>
<tr>
<td>5064 acceptance Rate(Hz)</td>
<td>21.2%</td>
<td>14.8%</td>
<td>25.1%</td>
<td>19.3%</td>
<td>39.5%</td>
</tr>
</tbody>
</table>

Table 4.6: EF trigger Rates and acceptances for the signal and the QCD background samples. The acceptance of 5204 is shown in table 4.2.

In figure 4.7 the signal efficiency is shown with respect to the QCD background rate. The QCD samples have been combined using the TriggerRate Tool. From this graph two conclusions can be drawn:

- Looking at the pairs: L2(EF)_jet160, L2(EF)_2jet120 and L2(EF)_3jet65, L2(EF)_4jet50, it is seen that increasing the multiplicity and decreasing the threshold leads to a small decrease in the signal efficiency but a relatively larger decrease in the background rate.

- The LVL2 and the EF seem to have a prescaled relation. This is especially clear for the jet160 signature where the virtual line connecting the LVL2 point with the (0,0) point one can realize that the reduced efficiency of EF (from 20% in LVL2 down to 11.5% in EF) reduces the
background rate to about a half (from 22 Hz in LVL2 down to 14 Hz for EF). This is the same effect as prescaling the LVL2 with a factor of \( \frac{20\%}{11.5\%} \approx \frac{22}{14} = 1.8 \), hence accepting 1 every 1.8 events. This is an indication of the EF performing poorly.

![Signal Efficiency vs Background Rate](image)

**Fig. 4.7: Background Rate Vs Signal Efficiency.**

### 4.5 Cross-correlation with semi-leptonic events

Until this point the analysis was mainly a validation study of the trigger showing the trigger efficiency for the various truth jet energies. The shape of the curves helps us understand the behavior of the jet trigger reconstruction in the trigger level. Deviations from the expected ideal turn-on curve (which is a step-function turning on at the trigger threshold) can be related to trigger reconstruction effects. Because in data taking period we rely only on offline reconstructed information it is useful to see if the same effects derived from the quality study can be seen with only offline reconstructed jets. In this part, the sample 5200 is used which contains semi-leptonic \( t\bar{t} \) events. The jet
trigger turn-on curves are built from these events. The final state consists of 4 jets and a charged lepton with its corresponding (anti)neutrino.

Although we are primarily interested in the jet trigger efficiency in fully hadronic $t\bar{t}$ events, the semi-leptonic $t\bar{t}$ events also provide information about the jets in a multijet environment. In an early stage of data taking, the semi-leptonic events are easier to study due to the possibility to trigger on the electron or muon.

The approach for the construction of turn-on curves from real data is similar as before with two modifications in the previously applied method. First of all, instead of the truth jets the offline reconstructed jets are associated to the trigger objects. This corresponds to actual data taking where we will base the analysis only on the reconstructed data as no truth information can be available. Secondly, the lepton trigger is requested to have fired, this is done to assure that we cover the entire jet energy spectrum with a sample that is not biased by the jet trigger selection itself. So in this case the events that passed the lepton trigger are used. The $\text{distA}$ contains all the offline reconstructed jets that could be matched to trigger objects. The $\text{distB}$ contains only the matched reconstructed jets that had trigger objects above the threshold. Dividing again the $\frac{\text{distB}}{\text{distA}}$ and plotting it with respect to the offline reconstructed energy gives the turn-on curve.

Additional effects that might be introduced from the offline reconstructed algorithms (e.g. calibration) can be investigated. To do this, the turn-on curve based on reconstructed jet energy information will be compared with the validation (quality study) turn-on curve. Especially in the EF case, where algorithms similar to the offline reconstruction are used, it is possible for an error to be introduced in both levels (EF and offline).

Consider the following example, if the EF reconstructs an event that contains a single high-energetic jet, of e.g. 160 GeV, erroneously with only 20 GeV, this jet will be rejected by the trigger. Unfortunately, the same error could also be present in the offline reconstruction and will not appear as an anomaly in the turn-on curve. However, in the validation study for the trigger using truth information, as made in the previous section, these low energetic jets would appear and lead to obvious differences when comparing the two graphs. This example indicates how the comparison can help to improve the understanding of the jet trigger reconstruction and its relation with the offline reconstruction. For the present analysis only the LVL2 was investigated. The EF is subject to the duplicate jets feature and thus no solid conclusions can be drawn.

The resulting LVL2 turn-on curves are shown for each signature and they

---

3The tau is a special, again difficult case.
are overlayed with the corresponding LVL2 ones from the validation study of the previous chapter (figures 4.8, 4.9, 4.10, 4.11). With the star dots the LVL2 validation turn-on curve is shown and the scale on the x-axis corresponds to the $p_T$ of truth jets as it was evaluated with the fully-hadronic sample. The square dots correspond to the turn-on curve made based on the semi-leptonic $t\bar{t}$ events and the scale on the x-axis corresponds to the $p_T$ of the offline reconstructed jets.

![Turn-on Curve J160 (Truth & Reco)](image)

Fig. 4.8: The LVL2 validation and data taking turn-on curves for the jet160.

There is an obvious similarity between the two types of turn-on curves, which may indicate that the same observations as made in the validation study can be made with the data taking method. Only small variations are observed especially above the threshold energies. The turn-on curve based on offline reconstructed information rises faster to higher efficiencies compared to the validation turn-on curves while in energies much higher than the threshold the validation turn-on curve at some points seems to have a slightly smaller efficiency than the data taking turn-on curve. Nevertheless, the similarity is quite surprising as it would be expected that effects in the reconstruction process would alter the shape of the turn-on curve.

The main problem with the present analysis lies on the different energy scale used for comparing the two types of turn-on curves. One approach could be to use a kinematic constraint on the $W$ mass which would improve the energy of the measured jets. In addition, the experimental errors should be applied on the truth $p_T$ (smearing). This way the expected behavior
4.5 Cross-correlation with semi-leptonic events

Fig. 4.9: The LVL2 validation and data taking turn-on curves for the 2jet120.

Fig. 4.10: The LVL2 validation and data taking turn-on curves for the 3jet65.

of the turn-on curves with improved offline reconstruction energies can be determined. However, the feasibility of this technique should be investigated in terms of the expected combinatorial background.
Fig. 4.11: The LVL2 validation and data taking turn-on curves for the 4jet50.
CHAPTER 5

Summary and conclusions

The ATLAS experiment has to cope with the enormous amount of data produced by the \( pp \)-collisions. Only a small fraction of these collisions can be accepted for storage and further analysis. It is essential to select those events that are relevant to the physics goals of ATLAS. The trigger system is optimized for selecting these events. One such event type is the associated production of a Higgs particle with a \( t\bar{t} \) pair (\( t\bar{t}H \)). This is a potential discovery channel for a light Higgs boson. As such, understanding the physics related to top quarks is important.

Pair production of top quarks follows from the quantum chromodynamics (QCD) processes: \( q\bar{q} \rightarrow t\bar{t} \) and \( gg \rightarrow t\bar{t} \). This will be the dominant production mechanism at the Large Hadron Collider (LHC). Within the Standard Model, the dominant decay of a top quark is the electroweak process: \( t \rightarrow W + b \), with a branching ratio of nearly 100%.

A \( t\bar{t} \) final state contains two \( t \rightarrow W + b \) decays. The \( b \)-quarks will hadronize to jets that contain the decay products of a \( B \)-meson, while the \( W \) boson will decay into either a lepton and a neutrino or into a pair of jets. The case where both \( W \)s decay into jets has a branching ratio of about 44% and consists of a final state with 4 light jets and 2 \( b \)-jets. Thus, no ambiguity is introduced due to missing energy from neutrinos\(^1\). Unfortunately this channel has a very large background coming from QCD multijets events. This thesis addressed the issues related to the event selection of fully-hadronically decaying \( t\bar{t} \) events. Because of its only-jets final state the jet trigger slice was studied.

\(^1\)In fact in 14% of \( B \)-meson decays, a neutrino is emitted but the amount of missing \( E_T \) is small comparing with the high energetic neutrino from \( W \) decays.
The behavior of the jet trigger slice can be visualized in the so-called turn-on curves where the efficiency of the jet trigger is shown with respect to the energy of the jets. In an ideal case for a jet trigger at a given threshold, the turn-on curve is a step-function with zero efficiency for jets with energies below the threshold and 100% efficiency for jets with energies above the threshold.

In the first part of the analysis the turn-on curves were based on the truth energy of the jets to determine the quality of the jet trigger. The study was done for both the LVL2 and EF. The plots showed that the LVL2, as expected, has a lower threshold than the EF. For signatures with increased jet multiplicity and lower thresholds the shape was closer to the ideal turn-on curve. Also, the EF curve is expected to come closer to the ideal step-function behavior than LVL2, resulting from better calibration and more refined reconstruction of the jets. However, the results showed the opposite. Eventually, this was accounted to a jet energy miscalculation, where the EF algorithms only take into consideration deposited energy inside specific regions called Regions of Interest (RoIs). These RoIs are regions in the detector identified by the LVL1 and in principle they do not contain the full jet information.

Additionally, a large number of triggers were fired by unknown objects. It was shown that this was the result of two problems. The first, originated from erroneous splitting of a wide jet into two close JETROIs by the LVL1 trigger. This effect propagates to the HLT were in the end the same jet is reconstructed twice in both levels (LVL2 and EF). The second problem, is only present at the EF and originates from the fact that two reconstruction algorithms are used in that level. Although only one is taken into account for the Trigger Decision, the reconstructed objects reside in the same collection, making them indistinguishable. This effect obviously biases all the results presented for the EF, hence no solid conclusions can be made on its performance.

The first part of the analysis was concluded with a feasibility study, calculating the signal efficiency with respect to the background rate. For the background several existing samples were used. Although these samples are not the best choice for this study, due to pre-selection criteria on jet energies that will bias the result, they provide an indication of LVL2 trigger and EF performance.

For the last part of the analysis, a feasibility study was made using semi-leptonic $t\bar{t}$ events. This time the turn-on curves of the signatures in LVL2 were built based on the offline reconstructed information, as it would be in the case of real data. The EF was not investigated due to the duplicate jet problem. The LVL2 turn-on curves were compared to the LVL2 quality
study turn-on curves. Comparing the two types of turn-on curves provides an additional validation step for that trigger level as differences between them may indicate features in the trigger and offline jet reconstruction. However, in the present study safe conclusions cannot be made as different energy scale is used for the truth jets and the offline reconstructed jets. For further analysis and for overcoming this problem different methods using kinematic constraints can be considered.


57


[46] ATLAS twiki: The ATLAS Geometry Database
https://twiki.cern.ch/twiki/bin/view/Atlas/AtlasGeomDBTags, [link]

[47] ATLAS twiki: Trigger Menu Versions
https://twiki.cern.ch/twiki/bin/view/Atlas/TriggerMenuVersions, [link].
[48] ATLAS twiki: The Athena Framework
   https://twiki.cern.ch/twiki/bin/view/Atlas/WorkBookAthenaFramework
   [link].


   https://twiki.cern.ch/twiki/bin/view/Atlas/TriggerRateTools, [link].

[51] Private communication with Cibran Santamarina Rios.