DISSERTATION/DOCTORAL THESIS

Titel der Dissertation/Title of the Doctoral Thesis

“Upgrade of the Global Muon Trigger for the Compact Muon Solenoid experiment at CERN”

verfasst von/submitted by
Mag. Dinyar Sebastian Rabady

angestrebter akademischer Grad/in partial fulfilment of the requirements for the degree of
Doktor der Naturwissenschaften (Dr. rer. nat.)

Wien, im Jänner 2018/Vienna, in January 2018

Studienkennzahl lt. Studienblatt/ degree programme code as it appears on the
student record sheet:
A 796 605 411

Studienrichtung lt. Studienblatt/ field of study as it appears on the
student record sheet:
Physik

Betreut von/Supervisor:
Dipl.-Ing. Dr. Claudia-Elisabeth Wulz
Hon.-Prof. Dipl.-Phys. Dr. Eberhard Widmann
Für meinen Großvater.
The Large Hadron Collider is a large particle accelerator at the CERN research laboratory, designed to provide particle physics experiments with collisions at unprecedented centre-of-mass energies. For its second running period both the number of colliding particles and their collision energy were increased. To cope with these more challenging conditions and maintain the excellent performance seen during the first running period, the Level-1 trigger of the Compact Muon Solenoid experiment — a sophisticated electronics system designed to filter events in real-time — was upgraded. This upgrade consisted of the complete replacement of the trigger electronics and a full redesign of the system’s architecture. While the calorimeter trigger path now follows a time-multiplexed processing model where the entire trigger data for a collision are received by a single processing board, the muon trigger path was split into regional track finding systems where each newly introduced track finder receives data from all three muon subdetectors for a certain geometric detector slice and reconstructs fully formed muon tracks from this. In contrast, the system in operation during the first data taking period was built around subdetector-specific track finders whose output tracks were merged in the Global Muon Trigger. This restructuring of the muon trigger required a novel system to receive muon objects from the track finding layer, remove any duplicate tracks, and forward the best-reconstructed objects to the global decision layer.

The upgraded Global Muon Trigger (μGMT) was designed to satisfy the requirements for such a system. It receives muon track data from all regional track finder processors as well as energy information from the final stage of the calorimeter trigger. These muon tracks are ranked and sorted, while in parallel their azimuthal coordinate is corrected and duplicate muon tracks are removed. An isolation variable using the calorimetric information is computed before the eight highest-ranked muon tracks are sent to the upgraded Global Trigger (μGT). The μGT then determines whether a readout process shall be initiated by applying complex algorithms on the muon tracks as well as on data received from the calorimeter trigger chain.

This thesis presents the design and development of the μGMT, outlining the requirements as well as challenges encountered during development of the upgraded trigger. Furthermore, the excellent performance of the new system is demonstrated in the context of studies on quarkonia: Both the impact of the improved duplicate removal system as well as of the correction of the azimuthal coordinate are evaluated and shown to improve upon the performance demonstrated by the original Level-1 trigger. Additionally, a simple scheme that increases the purity of dedicated trigger algorithms to select \( J/\psi \) mesons by requiring the presence of two oppositely charged muons is presented and assessed.

Das neue Global Muon Trigger (μGMT) System wurde entworfen um die obigen Anforderungen zu erfüllen. Es erhält Myonen-Daten von allen regionalen Spurenfindern, wie auch Energie-Information von der finalen Stufe des Kalorimeter-Triggers. Die Myonen-Spuren werden gereiht und sortiert, während zeitgleich deren azimutale Koordinate korrigiert und Duplikate entfernt werden. Mit Hilfe der Kalorimeter-Informationen wird eine Isolationsvariable berechnet, bevor die acht hochst-gereihten Spuren an das neue Global Trigger (μGT) System geschickt werden. Der μGT entscheidet mit Hilfe komplexer Algorithmen die auf die erhaltenen Objekte angewendet werden, ob ein Auslesevorgang des Detektors initiiert werden soll.

**Contents**

Abstract ........................................... i

Kurzfassung ........................................ iii

1 Introduction ...................................... 1
   1.1 Overview of the thesis ....................... 1
   1.2 Personal contributions ....................... 2

2 Quarkonium physics ................................ 3
   2.1 Introduction .................................. 3
   2.2 Quarkonium production ....................... 4
   2.3 Quarkonium polarisation ..................... 5
   2.4 Challenges in the trigger system .......... 7

3 The Large Hadron Collider ......................... 9
   3.1 Introduction .................................. 9
   3.2 The injection chain ......................... 11
   3.3 Luminosity .................................. 11
   3.4 Operational periods of the LHC ............ 13
   3.5 Experiments at the LHC ..................... 15
      3.5.1 ALICE .................................. 15
      3.5.2 ATLAS .................................. 16
      3.5.3 CMS ..................................... 17
      3.5.4 LHCb ..................................... 18

4 The Compact Muon Solenoid experiment ........... 21
   4.1 Overview ..................................... 21
      4.1.1 Geometry ................................ 21
      4.1.2 Coordinate system ....................... 22
   4.2 Inner tracking ................................ 23
      4.2.1 The pixel detector ....................... 24
      4.2.2 The silicon strip tracker ............... 25
   4.3 Calorimetry .................................. 26
      4.3.1 Electromagnetic calorimeter ............ 26
      4.3.2 Hadron calorimeter ....................... 27
## Contents

6.6 Monitoring and control of the upgraded Global Muon Trigger  .................................................. 102
  6.6.1 Monitoring with triggered events .......................................................................................... 102
  6.6.2 Monitoring via online software ............................................................................................ 103
6.7 Technological challenges ........................................................................................................... 105
  6.7.1 Distributed firmware development ......................................................................................... 105
  6.7.2 Routing on large FPGAs ....................................................................................................... 106
  6.7.3 Ensuring stability of the firmware ........................................................................................ 107

7 Using the upgraded Level-1 Trigger for physics .............................................................................. 111
  7.1 Monte Carlo event samples ....................................................................................................... 111
  7.2 Charge tagging performance .................................................................................................... 112
    7.2.1 Study of opposite charge requirement in the Level-1 trigger for LHC Run-2 .......................... 115
    7.2.2 Analysis of 2015 data ....................................................................................................... 117
  7.3 Removal of duplicate muon candidates ....................................................................................... 118
    7.3.1 Estimating the spatial separation of duplicates in the Level-1 trigger ................................. 120
    7.3.2 Double muon trigger rates ................................................................................................. 125
    7.3.3 Double muon efficiency .................................................................................................... 131
  7.4 Invariant mass computation ...................................................................................................... 137
    7.4.1 Corrections on azimuthal coordinates .................................................................................. 137
    7.4.2 Invariant mass resolution for the $J/\psi$ meson after extrapolation of the azimuthal coordinate ......................................................................................................................... 141
    7.4.3 Rates and efficiency of an invariant mass $J/\psi$ trigger ..................................................... 141

8 Conclusion ...................................................................................................................................... 145

Acknowledgements .......................................................................................................................... 149

Bibliography ...................................................................................................................................... 151

Glossary ............................................................................................................................................. 161

Curriculum vitae ............................................................................................................................. 171
Chapter 1

Introduction

The Compact Muon Solenoid (CMS) experiment at the CERN research facility near Geneva, Switzerland is a general purpose particle physics experiment with a wide range of physics goals, ranging from precision measurements in the Standard Model of particle physics, to searches for new phenomena, such as hints of super-symmetry, dark matter, or microscopic black holes. To this end, it profits from collisions provided by the Large Hadron Collider (LHC), a circular particle accelerator that collides bunched beams of protons or heavy ions in four interaction regions which house the experiments.

The LHC collides these particles with a rate of 40 MHz. It is infeasible to record the data of all collisions provided by the LHC due to limitations in read-out, storage, and analysis capabilities. To ensure that enough relevant data for both precision studies as well as searches for physics beyond the Standard Model are recorded, a filter and selection system, called trigger has been implemented. CMS utilises a two-tiered trigger system, made up of a hardware-based Level-1 trigger and a software-based event filter system, called High-Level trigger (HLT). While the HLT uses sophisticated reconstruction algorithms on the full detector information with a total processing delay of about 100 ms for a given event, the Level-1 trigger is restricted in computing time to less than 4 µs per event. As a result, the use of less complex algorithms that can be implemented in fast hardware-based systems is necessary.

After a successful first data taking period from 2010 to 2013, the LHC underwent an upgrade that increased both its collision energy as well as the instantaneous luminosity, a number proportional to the number of colliding particles in a certain time. To optimally cope with these more challenging conditions, the Level-1 trigger system of the CMS experiment was replaced by an upgraded system at the end of 2015, exploiting significant improvements in data transmission and semiconductor technology.

1.1 Overview of the thesis

This thesis describes the upgrade of a central component of the CMS Level-1 trigger, the upgraded Global Muon Trigger (µGMT), and demonstrates its improved performance in the context of studies on quarkonia, bound states of a quark and its anti-quark. The µGMT represents the final stage of the muon trigger chain, receiving reconstructed muon track objects from three track finder systems which each operate on specific detector regions. The input objects are ranked by their reconstruction quality and transverse momentum before being sorted. In parallel, duplicates that can arise due to the structure of the track finding systems are detected and removed while energy sums received from the calorimeter trigger can be used to compute an
Chapter 1 Introduction

isolation variable for the eight highest ranked muon objects. The μGMT furthermore corrects
the azimuthal coordinate of each muon track to neutralise the effect of magnetic bending. The
eight highest ranked muon track objects are then sent to the upgraded Global Trigger (μGT).

The thesis is structured to motivate the study of heavy quarkonia as well as describe dif-
ficulties in efficiently collecting events with quarkonia that decay to two close-by muons in
chapter 2. Following this, the LHC and its pre-accelerator chain are described in chapter 3. This
chapter additionally gives an overview of the experiments utilising the accelerator complex at
CERN. The CMS experiment is described in greater detail in chapter 4, completing the overview
of the experimental setup.

In chapter 5 the upgraded Level-1 trigger is introduced and contrasted with the legacy
system, in use from 2010 to 2015. The technical design of the μGMT is presented in chapter 6:
This chapter outlines the requirements for the upgraded system before describing details of
the implementation and finally discussing techniques used to overcome or mitigate obstacles
encountered during the development of the μGMT firmware. Chapter 7 is concerned with the
physics performance of the upgraded Level-1 muon trigger in the context of quarkonium physics.
The chapter presents analyses of several novel algorithms designed to improve the performance
of the muon trigger: An opposite charge requirement for double muon trigger algorithms
was designed to reduce the number of erroneously recorded events collected with algorithms
dedicated to the detection of J/ψ mesons. Furthermore, the performance of a system to remove
duplicate muon tracks, as well as the impact of the correction of the azimuthal coordinate by
the μGMT on the invariant mass computation by the Level-1 trigger were examined.

Finally, a summary together with an outlook are given in chapter 8.

1.2 Personal contributions

My work on the Level-1 trigger comprised the development of the μGMT as well as the introduc-
tion of an invariant mass-based trigger algorithm aimed at selecting J/ψ particles decaying to
two muons. In the development of this trigger algorithm I could exploit my knowledge of both
the μGT and μGMT systems, thereby improving the performance of the algorithm significantly.

What concerns the μGMT, I was the principal firmware developer and also performed studies
of the Level-1 trigger double muon efficiency, with an emphasis on the J/ψ → µµ detection
performance. Furthermore, I completed a study for the B physics group to estimate the behaviour
of a trigger algorithm exploiting an opposite charge requirement in order to more efficiently
detect muons from the decay of J/ψ mesons.

Finally, I was deeply involved in the commissioning of the upgraded Level-1 trigger and acted
as an on-call expert during extended periods.
In this chapter the Standard Model of particle physics is briefly introduced before giving an overview of quarkonium physics. The importance of studying J/ψ mesons for an improved understanding of quantum chromodynamics (QCD) is highlighted and examples of important measurements are given. Finally, the challenges for the muon trigger system are outlined.

References [1, 2] are extensively used.

2.1 Introduction

In the Standard Model of particle physics forces as well as matter constituents are represented by particles. While force-mediating bosons carry integer spin, particles constituting matter — called fermions — carry half-integer spin values. Fermions are then further split into the strongly interacting quarks, as well as leptons that do not experience the strong force (tables 2.1 and 2.2).

<table>
<thead>
<tr>
<th>Leptons</th>
<th>Symbol</th>
<th>Electric charge (e)</th>
<th>Spin</th>
<th>Affected by</th>
</tr>
</thead>
<tbody>
<tr>
<td>Electron</td>
<td>e</td>
<td>-1</td>
<td>1/2</td>
<td>EM, weak</td>
</tr>
<tr>
<td>Electron neutrino</td>
<td>νₑ</td>
<td>0</td>
<td>1/2</td>
<td>weak</td>
</tr>
<tr>
<td>Muon</td>
<td>µ</td>
<td>-1</td>
<td>1/2</td>
<td>EM, weak</td>
</tr>
<tr>
<td>Muon neutrino</td>
<td>νₘ</td>
<td>0</td>
<td>1/2</td>
<td>weak</td>
</tr>
<tr>
<td>Tau</td>
<td>τ</td>
<td>-1</td>
<td>1/2</td>
<td>EM, weak</td>
</tr>
<tr>
<td>Tau neutrino</td>
<td>νₜ</td>
<td>0</td>
<td>1/2</td>
<td>weak</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Quarks</th>
<th>Symbol</th>
<th>Electric charge (e)</th>
<th>Spin</th>
<th>Affected by</th>
</tr>
</thead>
<tbody>
<tr>
<td>Up</td>
<td>u</td>
<td>+2/3</td>
<td>1/2</td>
<td>EM, weak, strong</td>
</tr>
<tr>
<td>Down</td>
<td>d</td>
<td>-1/3</td>
<td>1/2</td>
<td>EM, weak, strong</td>
</tr>
<tr>
<td>Charm</td>
<td>c</td>
<td>+2/3</td>
<td>1/2</td>
<td>EM, weak, strong</td>
</tr>
<tr>
<td>Strange</td>
<td>s</td>
<td>-1/3</td>
<td>1/2</td>
<td>EM, weak, strong</td>
</tr>
<tr>
<td>Top</td>
<td>t</td>
<td>+2/3</td>
<td>1/2</td>
<td>EM, weak, strong</td>
</tr>
<tr>
<td>Bottom</td>
<td>b</td>
<td>-1/3</td>
<td>1/2</td>
<td>EM, weak, strong</td>
</tr>
</tbody>
</table>

Both leptons and quarks are categorised into three generations, based on their masses. For the leptons, the electron (0.511 MeV\(^1\)) and its associated neutrino form the first generation, while

---

\(^1\)In this work natural units are used, where \( c = h = 1 \) and eV is the unit of energy. Consequently energies, masses,
Chapter 2 Quarkonium physics

the muon (approximately 105.7 MeV) and tau (approximately 1777 MeV) leptons are associated with the second and third generation, respectively. Neutrinos are massless in the Standard Model of particle physics, however the observation of so-called neutrino oscillations implies that neutrinos are in fact massive particles.

For quarks, the up and down quarks (about 2.2 MeV and 4.7 MeV) constitute the first generation, charm and strange (about 1280 MeV and 96 MeV) form the second generation, and top and bottom quarks (about 173.1 GeV and 4.18 GeV) make up the third generation. [1]

Table 2.2: Table of bosons. Values taken from [1].

<table>
<thead>
<tr>
<th>Particle</th>
<th>Symbol</th>
<th>Electric charge (e)</th>
<th>Spin</th>
<th>Interaction mediated</th>
</tr>
</thead>
<tbody>
<tr>
<td>Photon</td>
<td>(\gamma)</td>
<td>0</td>
<td>1</td>
<td>EM</td>
</tr>
<tr>
<td>W boson</td>
<td>(W^\pm)</td>
<td>1</td>
<td>1</td>
<td>weak</td>
</tr>
<tr>
<td>Z boson</td>
<td>(Z^0)</td>
<td>0</td>
<td>1</td>
<td>weak</td>
</tr>
<tr>
<td>Gluon</td>
<td>(g)</td>
<td>0</td>
<td>1</td>
<td>strong</td>
</tr>
<tr>
<td>Higgs boson</td>
<td>(H)</td>
<td>0</td>
<td>0</td>
<td>—</td>
</tr>
</tbody>
</table>

While the electromagnetic and weak interactions are well understood and have been unified in the electroweak theory, the strong force — described by QCD in the Standard Model — poses computational challenges due to the self interaction of the gluons. Although short range or high energy interactions in QCD can be accurately modelled using the technique of perturbative QCD, long range/low energy interactions cannot easily be computed because the strong coupling constant grows exponentially with decreasing energy, a phenomenon called asymptotic freedom [2].

Among the simplest systems governed by the strong force are the quarkonia\(^2\) \(J/\psi = c\bar{c}\) and \(\Upsilon = b\bar{b}\). The study of quarkonia is crucial to test the current understanding of hadron formation because it is the only way to experimentally probe the interaction between short-range and long-distance components of the strong force. Among other measurements, this includes the determination of quarkonium production cross-sections as well as the mesons’ polarisation.

Due to its excellent tracker and muon detector system, the Compact Muon Solenoid (CMS) experiment is well suited for a study of quarkonia by making use of the double muon decay channel. The provision of efficient trigger algorithms for a large range of meson transverse momenta is therefore crucial for collecting an adequate number of events containing quarkonia.

2.2 Quarkonium production

The production of quarkonium states is modelled in two parts. In a first step the production of an intermediate \(Q\bar{Q}\) pair at short distance is calculated within perturbative QCD. This is followed by hadronisation, where the intermediate pair forms a QCD bound quarkonium state.

\(^2\)Quarkonia are bound states of a heavy quark and its respective antiquark.

and momenta are all given in units of eV.
2.3 Quarkonium polarisation

The hadronisation step takes place in the non-perturbative realm of QCD and can thus not be computed easily.

The timescale of the production of the $Q\bar{Q}$ intermediate pair is on the order of $1/m_Q$ while the timescale for the hadronisation is on the order of $1/(m_Q \cdot v^2)$ [3]. In the previous term, $m_Q$ stands for the quark mass and $v$ stands for the relative quark velocity in the bound state, where $v^2 = 0.1$ holds for bottomonium while $v^2 = 0.3$ holds for charmonium [4]. As this allows a reasonable separation of the two production steps even for the charmonium state, quarkonia can be considered uniquely suitable for the study of hadron formation.

As shown in the previous paragraph, the relative quark velocity in the bound state is sub-relativistic and the perturbative production step can be dealt with separately from the non-perturbative hadronisation. These two features allow quarkonium production to be modelled by non-relativistic QCD (NRQCD) [3], an effective field theory that arises from full QCD in the limit of $m_Q \to \infty$.

In NRQCD the cross-section $\sigma(\Omega)$ of the quarkonium $\Omega$ is given by [3]
\[
\sigma(\Omega) = \sum_n S [Q\bar{Q}(n)] \cdot O^\Omega(n),
\]
where the perturbative production of the initial $Q\bar{Q}$ pair in quantum state $n$ is described by the short-distance coefficients $S [Q\bar{Q}(n)]$ while the expression $O^\Omega(n)$ stands for the long-distance matrix elements pertaining to the non-perturbative evolution of the $Q\bar{Q}$ pair into a bound quarkonium $\Omega$.

An individual term $\sigma(n) = S [Q\bar{Q}(n)] \cdot O^\Omega(n)$ is called a partial cross-section for an intermediate state $n$. The sum in eq. (2.1) runs over all intermediate $Q\bar{Q}$ states $n$, including in particular colour-octet and colour-singlet configurations. In this formalism colour-octets can exist as intermediate states that transform into the colour-singlet quarkonium bound state under emission of soft gluons.

As previously mentioned, the short-distance coefficients $S [Q\bar{Q}(n)]$ can be calculated in perturbative QCD and account for the kinematic dependence of the decay distributions and cross-section. They are therefore different for various processes, in particular dependent on the collision system and centre-of-mass energy.

In contrast to the short-distance coefficients, the long-distance matrix elements $O^\Omega(n)$ are constants that are independent of the $Q\bar{Q}$ kinematics and are assumed to only depend on the initial $Q\bar{Q}$ state $n$ as well as the final quarkonium state $\Omega$. They can be interpreted as being proportional to the probability of an intermediate $Q\bar{Q}$ in state $n$ to transition into the quarkonium state $\Omega$. These long-distance matrix elements can currently not be computed from theory, but have to be derived from experimental data.

2.3 Quarkonium polarisation

The polarisation of the quarkonium state corresponds to the preferred spin alignment of the system, caused by symmetries of the electroweak and strong interaction as well as basic conservation laws. This preferred spin alignment influences the angular distribution of the decay to two leptons in the quarkonium rest frame when considering decays such as $J/\psi \to \mu\mu$. 

5
Chapter 2 Quarkonium physics

Figure 2.1: Illustration of the decay angles in the quarkonium rest frame. The production plane is constructed by combining the momentum vector of the colliding beams with the momentum vector of the quarkonium in the laboratory frame. [5]

An unpolarised quarkonium state would correspond to an isotropic angular decay distribution while a polarised state would lead to anisotropies in the decay distribution.

The measurement of the angular distribution is performed relative to a polarisation reference frame in the quarkonium rest frame. The azimuthal and polar decay angles, \( \varphi \) and \( \vartheta \), are defined as the angles of the positive lepton in the reference frame and are described in fig. 2.1. The \( x-z \) plane in the reference frame is the production plane in proton-proton collisions, which is constructed by combining the momentum vector of the colliding beams with the momentum vector of the quarkonium in the laboratory frame in the following way: The direction of the \( y \) axis is defined as \( \vec{b}_{r1} \times \vec{b}_{r2} \), where \( \vec{b}_{r1} \) and \( \vec{b}_{r2} \) are the momentum vectors of the colliding beams in the quarkonium rest frame.

Finally, the quantisation axis \( z \) within the production plane along which the angular momentum is specified must be chosen to fully define the reference frame. There are several physically motivated choices that can be made, among these is the helicity axis that is aligned with the flight direction of the quarkonium.

The angular decay distribution of a vector particle’s parity-conserving double lepton decay can be written as [5]

\[
W(\cos \vartheta, \varphi | \vec{\lambda}) \propto \frac{1}{3 + \lambda_\vartheta} \cdot (1 + \lambda_\vartheta \cos^2 \vartheta + \lambda_\varphi \sin^2 \vartheta \cos 2\varphi + \lambda_{\vartheta \varphi} \sin 2\vartheta \cos \varphi). \tag{2.2}
\]

The decay distribution is parameterised by the so-called polarisation parameters or anisotropy parameters \( \vec{\lambda} = (\lambda_\vartheta, \lambda_\varphi, \lambda_{\vartheta \varphi}) \). The polar anisotropy is described by the parameter \( \lambda_\vartheta \), the azimuthal anisotropy by \( \lambda_\varphi \), while \( \lambda_{\vartheta \varphi} \) represents the change of the azimuthal anisotropy with respect to the polar angle.

It is possible to calculate the polarisation parameters with respect to a given quantisation axis by defining the short-distance coefficients that correspond to the projections of the \( Q\bar{Q} \) spin \( S \) on to the quantisation axis \( z \), \( S_{ij}[Q\bar{Q}(n)] \) (with \( i, j = 0, \pm 1 \)) [6].
where $S_{00}$ is the longitudinal short-distance coefficient and $S_{11}$ the transverse short-distance coefficient. The total short-distance coefficient can then be written as

$$S_{Q\bar{Q}}(n) = S_{00} \left[ Q\bar{Q}(n) \right] + 2 \cdot S_{11} \left[ Q\bar{Q}(n) \right].$$

### 2.4 Challenges in the trigger system

The selection of quarkonia in the CMS Level-1 trigger system is particularly challenging: At low transverse momenta of the quarkonium under study the two decay muons are difficult to distinguish from the high rate background of soft muons originating from hadronic jets. Conversely, at high transverse momenta the decay products are badly separated spatially for low mass quarkonia such as the $J/\psi$ meson. Two close-by muons can then be misidentified as a single muon by the Level-1 trigger system’s duplicate removal system.

A well tuned Level-1 trigger system is therefore imperative for a successful quarkonia physics program.
This chapter describes the Large Hadron Collider (LHC) and its injection chain together with their associated experiments. After an overview of the accelerator complex, the concept of luminosity is introduced. To conclude the sections concerned with the accelerator, an outline of the LHC’s running periods is presented.

Finally, a short description of the experiments at the LHC itself is provided.

3.1 Introduction

The LHC at the European Organisation for Nuclear Research (CERN) is a large accelerator that utilises the tunnel formerly occupied by the Large Electron–Positron Collider (LEP). It was designed to provide experiments with unprecedented collision energies and instantaneous luminosities\(^1\) and is thus designated a “discovery machine”. Its physics program relies mostly on proton-proton collisions except for about one month of each year that is dedicated to heavy ion runs in which ionised atoms are made to collide either with each other or with protons. [7, 8, 9]

For proton-proton physics the LHC design parameters foresee an ultimate centre-of-mass energy of 14 TeV as well as an instantaneous luminosity of \(1 \times 10^{34} \text{ cm}^{-2} \text{s}^{-1}\) while for heavy ion collisions an energy of 2.76 TeV per nucleon and a luminosity of \(1 \times 10^{27} \text{ cm}^{-2} \text{s}^{-1}\) were planned.

A driving force for the construction of the LHC was the expected discovery of the Higgs boson in order to complete the Standard Model, as well as the search for physics beyond the Standard Model. Such new physics is strongly hinted at by phenomena such as dark matter and dark energy, as well as the large number of free parameters in the current theory. Additionally, the LHC’s environment is suitable for precision measurements of Standard Model physics, in particular using \(b\) and \(t\) quarks as well as \(W^\pm\) and \(Z^0\) vector bosons that are produced in large quantities at the LHC. Finally, a portion of the physics program is dedicated to heavy ion physics that probes a novel state of hadronic matter, the so-called quark-gluon plasma under conditions similar to those in the early universe.

Many processes indicative of the above-mentioned phenomena are expected at the TeV scale but are characterised by very low cross-sections. A hadron collider has several advantages over lepton colliders: Cross-sections of several predicted processes are higher in strong interactions

\(^1\) Instantaneous luminosity is proportional to the number of colliding particles, the collision frequency, and inversely proportional to the transverse beam sizes in the horizontal and vertical direction.
and a wide range of centre-of-mass energies can be probed simultaneously. This advantage arises because in lepton accelerators the colliding particles have a well defined energy while hadron colliders accelerate protons to a known energy, with the interacting particles — quarks and gluons — carrying a variable fraction of the proton’s total energy. Finally, a hadron collider’s centre-of-mass energy can be significantly larger than that of a lepton collider, as it suffers from less energy loss due to synchrotron radiation.

Synchrotron radiation power loss scales with \( \left(\frac{E}{m}\right)^4 \) where \( E \) is the accelerated particle’s energy, \( \rho \) stands for the bending radius of the circular collider and \( m \) for the accelerated particle’s mass. It reached 18 MW at LEP2 when it operated at a centre-of-mass energy of 200 GeV, proving a formidable challenge for the vacuum components and requiring good control of the experiments’ background. [10] From the above equation it follows that more massive particles can compensate for an increase in the centre-of-mass energy. As the proton is heavier than the electron by a factor of \( \sim 2000 \) it is a good candidate particle for use in a discovery machine.

Historically hadron colliders have served as discovery machines that broke energy records, while lepton accelerators provided high-precision measurements in lower energy regimes. A graph of the development of particle colliders is given in fig. 3.1.

![Graph showing the development of particle colliders](image)

**Figure 3.1**: The development of particle colliders plotted over time. The maximal centre-of-mass energy has increased almost exponentially until the end of the past century. [11]
3.2 The injection chain

The LHC machine is contained in an almost circular tunnel measuring 27.6 km in circumference, about 100 m below the surface between the Jura mountain range and lake Geneva. Two beams are injected into the ring in opposite directions, crossing each other at four interaction points that house particle physics experiments. The LHC forms the last stage of an accelerator chain (fig. 3.2) that begins with the Linear accelerator 2 (LINAC2) and the Proton Synchrotron Booster (PSB) for proton beams, and the Linear accelerator 3 (LINAC3) in conjunction with the Low Energy Ion Ring (LEIR) for heavy ions. The beam is accelerated in PSB from 50 MeV that are provided by LINAC2 to 1.4 GeV and in LEIR to 72.2 MeV from the 4.2 MeV achieved in LINAC3.

The PSB is implemented as four stacked rings, that each inject a bunch into the Proton Synchrotron (PS). The PS is filled with 6 such bunches that are then each split in 12 to form a train of 72 bunches which is accelerated to 25 GeV before being injected into the Super Proton Synchrotron (SPS). LEIR captures 4 to 5 pulses of heavy ions from LINAC3 and forms two bunches that are injected into PS. Using the same mechanism as for proton bunches, PS splits the two heavy ion bunches in half to provide four bunches with 5.9 GeV to SPS.

For proton beams in normal operating conditions the SPS is filled three or four times with trains made of 72 bunches by the PS and accelerates them to 450 GeV. In heavy ion mode the SPS is filled 13 times by the PS for a total of 52 bunches that are accelerated to 177 GeV.

Finally, the LHC is filled 12 times for each beam direction by the SPS, reaching a nominal peak energy of 7 TeV for protons and 5.1 TeV for heavy ions. [13, 14, 15, 16]

PSB, PS, as well as SPS additionally provide physics beams to smaller facilities such as the Isotope mass Separator On-Line facility (ISOLDE) in the case of PSB, the Antiproton Decelerator (AD) ring for PS, and several experiments in the so-called North Area as well as the Proton Driven Plasma Wakefield Acceleration Experiment (AWAKE) for the SPS.

3.3 Luminosity

The rate of a certain event \( R_{\text{event}} \) is defined as

\[
R_{\text{event}} = L \cdot \sigma_{\text{event}}
\]

where \( \sigma_{\text{event}} \) is the energy-dependent cross-section for the event under study and \( L \) is the machine luminosity. For events of interest in collider experiments this cross-section generally increases with rising energy. The machine luminosity only depends on the beam parameters and can be written as

\[
L = \frac{N_b^2 n_b f_{\text{rev}} \gamma_r}{4\pi \epsilon_n \beta^*} \cdot F
\]

for a Gaussian beam distribution. In the numerator, \( N_b \) stands for the number of particles per bunch, \( n_b \) is the number of bunches per beam, \( f_{\text{rev}} \) the revolution frequency, and \( \gamma_r \) the relativistic gamma factor. In the denominator, \( \epsilon_n \) is the normalised transverse beam emittance,
Figure 3.2: Illustration of the CERN accelerator complex. The LHC is supplied with protons or heavy ions by the SPS via two transfer lines which additionally provides beams to several non-LHC experiments. The SPS is fed with preaccelerated protons by the PS. PS was CERN’s first synchrotron and it continues to supply smaller experiments with protons. The first circular accelerator in the proton-proton injection chain is PSB which is filled by the linear accelerator LINAC2. For beams of heavy ions the injection chain begins with LINAC3 which feeds PS via the LEIR. (©CERN, 2013–2016) [12]
3.4 Operational periods of the LHC

and $\beta^*$ the beta function at the collision point$^2$. Finally, $F$ stands for the geometric luminosity reduction factor due to the crossing angle at the interaction point. It depends on the full crossing angle, the bunch length, and the transverse beam size at the interaction point. It thus follows from eqs. (3.1) and (3.2) that the exploration of rare events requires both high beam energies as well as intensities. [13, 17]

As can be seen in fig. 3.3, the instantaneous luminosity provided by LHC has been significantly increased from the beginning of the LHC operation to the present. Similarly, the so-called integrated luminosity, defined as $L = \int L \cdot dt$ significantly grew. The integrated luminosity is a measure for the amount of data that were recorded over a given time and is counted in units of inverse cross-section, e.g. $1/b$ where $b$ is a barn and is defined as $1\, b = 1 \times 10^{-24}\, cm^2$. Similarly, the instantaneous luminosity can be given in $cm^{-2}\, s^{-1}$ as well as $Hz/b$.

3.4 Operational periods of the LHC

The operation of the LHC has been split into so-called running periods, separated by long shutdowns of multiple years in which major consolidation and upgrade work is performed (fig. 3.4).

While the LHC was designed to ultimately be able to provide collisions with a centre-of-mass energy of $14\, TeV$ and an instantaneous luminosity of $1 \times 10^{34}\, cm^{-2}\, s^{-1}$ to the experiments, it was initially operated at significantly lower energies and luminosities.

During the first running period from 2010 to 2012, called Run-1 the centre-of-mass energy was increased in steps up to $8\, TeV$ and the instantaneous luminosity eventually reached almost $8 \times 10^{33}\, cm^{-2}\, s^{-1}$. Furthermore, the LHC was operated with a bunch spacing of $50\, ns$ instead of the planned $25\, ns$. During the subsequent Long Shutdown 1 (LS1) the interconnections between the superconducting magnets were consolidated in order to allow operation at a centre-of-mass energy of $13\, TeV$, and the LHC experiments underwent upgrades in order to prepare for an instantaneous luminosity of up to $2 \times 10^{34}\, cm^{-2}\, s^{-1}$ which significantly exceeds the LHC’s nominal luminosity. Furthermore, the collision frequency was doubled to the nominal value of $40\, MHz$.

LHC Run-2 began in 2015 and will continue until end of 2018 with an ultimate instantaneous luminosity of $1.6 \times 10^{34}\, cm^{-2}\, s^{-1}$. It will culminate in Long Shutdown 2 (LS2) for which upgrades to the LHC injectors as well as further extensive upgrades to the experiments are planned, allowing the instantaneous luminosity for Run-3 (foreseen to begin in 2020) to reach $2 \times 10^{34}\, cm^{-2}\, s^{-1}$. The Long Shutdown 3 (LS3) (2023 to 2024) will be used to upgrade LHC to the High Luminosity Large Hadron Collider (HL-LHC) providing an unprecedented instantaneous luminosity of $5 \times 10^{34}\, cm^{-2}\, s^{-1}$. [19]

$^2$ $\beta^*$ is the distance from the interaction point to the point where the beam width is twice as wide as at the interaction point.

$^3$ The maximum instantaneous luminosity provided by the LHC surpassed $2.0 \times 10^{34}\, cm^{-2}\, s^{-1}$ in 2017, however using a non-standard operational mode that led to a very high number of overlapping particle collisions during one bunch crossing. During this mode of operation the instantaneous luminosity was lowered to almost $1.5 \times 10^{34}\, cm^{-2}\, s^{-1}$ by de-focusing the particle beams at the interaction point.
(a) The development of the instantaneous luminosity provided by LHC. As can be seen, large increases in instantaneous luminosity were achieved during the first LHC running period from 2010 to 2012. In 2015, the first year of the second running period, significant time was spent on commissioning of the upgraded LHC machine while 2016 was designated a physics production year which was dedicated to provide as much data to the experiments as possible.

(b) The total integrated luminosity recorded by the experiments at the LHC each year can be explained both by the machine’s instantaneous luminosity as shown in the plot above as well as the general machine availability over the year. For instance, in 2015 significantly less data was recorded than in 2011 with similar instantaneous luminosity due to the later start of physics data production in that year.

Figure 3.3: Plots showing both the instantaneous and integrated luminosities recorded by the Compact Muon Solenoid (CMS) experiment at the LHC. [18]
3.5 Experiments at the LHC

The beams accelerated by LHC cross in four so-called interaction regions, around which large experiments have been constructed. While A Toroidal LHC Apparatus (ATLAS) and Compact Muon Solenoid (CMS) are designed as general purpose experiments, tasked with discovering new physics as well as completing the Standard Model by finding the Higgs boson, A Large Ion Collider Experiment (ALICE) is intended primarily to be used during heavy ion runs, and LHC beauty (LHCb) specialises in heavy flavour physics and probes of the matter-antimatter symmetry. These four large experiments are complemented by smaller ones sharing their interaction point.

3.5.1 ALICE

The ALICE experiment was optimised to cope with the event conditions produced in heavy ion collisions. Its advanced Time Projection Chamber (TPC) can track individual particle tracks in very high multiplicity environments while introducing only a minimum of dense matter into their paths. As individual heavy ion collisions occur at lower energies than proton-proton collisions a low material budget is crucial in order to be able to measure low-momentum tracks with great precision. The problem of long drift times in a TPC is mitigated by the lower rate of collisions during heavy ion runs. During proton-proton collisions the instantaneous luminosity at the ALICE interaction point is artificially reduced by a process known as luminosity levelling. [20, 21]

During LS2 the ALICE experiment will undergo significant upgrades in order to prepare for the challenges of HL-LHC. The inner tracking system will be complemented with a silicon pixel
Chapter 3 The Large Hadron Collider

detector to improve the tracking precision and the ALICE read-out system will be upgraded to be able to accept all events at an interaction rate of 50 kHz during heavy ion runs, representing a 100-fold increase in read-out rate. [22]

3.5.2 ATLAS

The main distinguishing feature of the ATLAS experiment is its toroid magnets, complementing a thin superconducting solenoid that generates a 2 T magnetic field in the interaction point. It is the largest of the four LHC experiments, measuring 25 m in diameter and 44 m in length.

Similar to CMS, the ATLAS experiment is a general purpose detector with the goal of using the collisions provided by LHC to discover new physics as well as to complete or extend the Standard Model by discovering and studying the Higgs boson. In order to allow discoveries by ATLAS or CMS to be confirmed by the other experiment care was taken to build the detectors using different detector technologies such that they complement each other. [23]

The ATLAS collaboration has decided to split the upgrades leading to a detector suitable for the conditions provided by HL-LHC into two phases, to be installed over the three long shutdowns. The Phase-1 is to be completed over LS1 and LS2 while the Phase-2 upgrades will be installed in LS3. Such a scheme allows the experiment to profit from the LHC’s instantaneous luminosity exceeding the design value already in Run-2.

During LS1 a fourth layer for the pixel detector, the so-called insertable B-layer was added between the original pixel detector and a new beam pipe with smaller diameter [24]. In LS2 an upgrade of the muon spectrometer as well as the trigger system is planned. For the trigger it is foreseen to increase the granularity in the calorimeter path as well as to add information from the tracking system between the Level-1 and Level-2 trigger stages using a dedicated hardware trigger system. [25, 26]

For LS3 a complete replacement of the existing tracking system, currently comprised of the pixel detector, a silicon strip tracker, and a Transition Radiation Tracker (TRT), is foreseen. The new system will be a full silicon tracker, composed of silicon pixels close to the interaction region and silicon strip detectors further outward. Furthermore, the trigger system will be upgraded such that the functionality in the current Level-1 trigger will move to a new stage, named the Level-0 trigger. Level-1 will gain functionality and will receive information from the tracking systems. The trigger rates will be 500 kHz at Level-0 and 200 kHz at Level-1. [27]

LHCf

The LHC forward (LHCf) detector is a forward experiment, profiting from collisions taking place in ATLAS’s interaction region. It is placed at 140 m on either side of the ATLAS collision point and is designed to detect particles moving almost parallel to the beam-line in order to simulate cosmic rays under laboratory conditions. The detector is used to measure the forward production spectra of photons and neutral pions and the leading particle spectrum. These measurements are used to calibrate the Monte Carlo nuclear interaction model that is applied for several important measurements in cosmic ray physics such as the derivation of the cosmic ray energy spectra. Previously available data only reached up to an energy of $1 \times 10^{14}$ eV, while
3.5 Experiments at the LHC

the LHC provides laboratory equivalent energies up to $1 \times 10^{17}$ eV allowing the exploration of interesting energies in high energy cosmic ray interactions. [28, 29]

3.5.3 CMS

The CMS detector is designed as the sister experiment to the ATLAS detector. The main distinguishing feature of the CMS experiment is the large superconducting solenoid that contains the complete inner tracking system as well as most of the calorimeters and produces a magnetic field of 3.8 T in the interaction region. The CMS detector measures 14.6 m in diameter and 21.6 m in length with a total weight of 12 500 t compared to ATLAS’s 7000 t. [30]

Analogous to the upgrade plan of the ATLAS collaboration, CMS has split the detector upgrade campaign into two phases, the first comprising the period from LS1 to LS2, with the upgrades for the second phase being installed during LS3.

In LS1 CMS underwent a significant upgrade of the trigger and data acquisition (DAQ) system as well as work on the tracking system, replacements for the hadronic calorimeter (HCAL), and an extension of the endcap muon chambers. The schedule for the upgrade of the trigger system foresaw to also include the first year of LHC Run-2 in order to validate the new system on data in parallel to the operation of the legacy system while not endangering data taking. Furthermore, the beam pipe was replaced with a thinner tube, in order to accommodate a future replacement of the pixel detector. This replacement of the pixel detector was originally scheduled for LS2, however it was possible to install it during the end-of-year extended technical stop 2016/2017 due to faster than anticipated progress. [31, 32]

For LS2 upgrades of the HCAL front-end and photo-detectors are foreseen. Furthermore, there is maintenance work planned on the solenoid as well as further work on the beam pipe. [33]

During LS3 the entire tracker will be replaced due to radiation damage to the current system. The granularity of both the pixel and strip modules will be increased by about a factor four in order to cope with the expected high multiplicities. Additionally, the strip modules will be able to provide track stub information at 40 MHz to the Level-1 trigger which allows the introduction of a dedicated Level-1 track trigger. Similarly, the endcap electromagnetic and hadronic calorimeters will have to be exchanged due to significant radiation damage. They will be replaced by the High Granularity Calorimeter (HGC), which consists of electromagnetic and hadronic sections and can provide detailed three dimensional images of showers. In the endcap muon system the pseudorapidity range of $1.5 < |\eta| < 2.4$ will be augmented with the addition of layers of Gas Electron Multiplier (GEM) and resistive plate chamber (RPC) muon detectors. Finally, the trigger and DAQ system will be significantly upgraded as well. The Level-1 trigger latency will be more than tripled to 12.5 $\mu$s, allowing to reconstruct tracks in hardware. Additionally, the maximal read-out rate will be increased from 100 kHz to 750 kHz. The DAQ system will be entirely replaced in order to accommodate the increased read-out rate and event size, as well as the increased complexity of event reconstruction due to the high pile-up. [34]

CMS and the upgrades done in LS1 are described in detail in chapter 4. The CMS Level-1 trigger and its Phase-1 upgrade are described in chapter 5.
Chapter 3 The Large Hadron Collider

TOTEM

The TOTal cross section, Elastic scattering and diffraction dissociation Measurement at the LHC (TOTEM) experiment shares its interaction point with the CMS detector. It is designed to measure the total proton-proton cross-section as well as allow physicists to gain a deeper understanding of the proton structure by studies of elastic scattering with large momentum transfers. These measurements require a good acceptance for particles travelling close to the beam.

TOTEM is comprised of two telescopes on either side of the interaction point designed to detect inelastically produced charged particles complemented by detectors in special movable beam-pipe insertions, called Roman Pots. The first telescope, placed 9 m from the interaction region uses cathode strip chambers (CSCs) while the second telescope, at a distance of 13.5 m from the interaction point is comprised of GEM detectors. In contrast, the Roman Pots are located 147 m and 220 m from the interaction point and detect leading protons at only a few mm from the beam centre. They are silicon devices, custom designed in order to reduce the insensitive area at the edges of the sensors which allows to maximise the experiment’s acceptance for particles scattered elastically or diffractively at polar angles of a few µrad at the interaction point. [35, 36]

In order to allow common data taking between TOTEM and CMS the TOTEM DAQ was designed to be compatible with the CMS central DAQ. This integration was achieved during LS1 with the CMS-TOTEM Precision Proton Spectrometer (CT-PPS) project which integrates the TOTEM detectors together with additional detectors close to the beam line into the CMS DAQ system. This integration allows the study of so-called central exclusive production in proton-proton collisions, represented by the process \( pp \rightarrow pXp \), during which the protons do not dissociate, but interact via photon or colour-singlet exchange to produce the system \( X \) in the central region. CT-PPS can detect such events by measuring the system \( X \) (which can be e.g. \( \gamma\gamma \rightarrow W^+W^- \)) in the CMS detector while detecting the only very slightly scattered protons in the Roman Pot detectors at a distance of about 200 m from the interaction region on either side of the CMS detector. [37]

3.5.4 LHCb

The primary goal of the LHCb experiment is the search for indirect evidence of new physics in rare decays of beauty and charm hadrons or CP violation. These studies require the measurement of the rare decays of \( B_d \), \( B_s \), and D mesons which are produced with a large cross-section at the LHC.

In order to exploit this large number of b hadrons, the LHCb detector requires a robust and flexible trigger that is able to select events with high efficiency in the harsh hadronic environment of the LHC. Furthermore, superior vertex and momentum resolution are crucial to study the rapidly oscillating \( B_s - \bar{B}_s \) meson system. Lastly, the LHCb DAQ has to provide powerful online data processing.

To allow very high resolution measurements of the vertex position, LHCb is equipped with the highly sophisticated Vertex Locator (VELO), a silicon tracker that can be moved to only 7 mm distance from the proton beams once beam conditions are declared stable. [38]
In contrast to the remaining LHC experiments, the LHCb detector is a single-arm spectrometer. This design decision was justified by the fact that b hadrons are predominantly produced in the forward or backward cone, which allows the experiment to ignore one side of the interaction point. Furthermore, by modifying the LHC optics to move the interaction point by 11.25 m to the side of the cavern, LHCb can make maximal use of the available space. [39]

The LHCb experiment was already operating at double its design luminosity of $2 \times 10^{32} \text{cm}^{-2}\text{s}^{-1}$ during the last year of Run-1. For Run-3 it is foreseen to increase this even further to a constant instantaneous luminosity of $2 \times 10^{33} \text{cm}^{-2}\text{s}^{-1}$, making significant upgrades during LS2 to both the detector itself as well as its read-out system necessary.

In the current read-out system, a Level-0 trigger is implemented in the hardware of the individual sub-detectors, reducing the event rate from the nominal LHC bunch crossing rate of 40 MHz to about 1 MHz based on transverse energy and momentum measured by the calorimeters and muon system. The events selected by the Level-0 trigger are then filtered by the software-based High-Level trigger (HLT) which writes events to disk at a rate of 5 kHz. For the upgrade it is foreseen to remove the hardware-based trigger and reconstruct events in the HLT with the full inelastic event rate of 30 MHz. The upgraded HLT will have an increased output rate of 20 kHz to 100 kHz. [33, 40, 41]

Furthermore, the VELO will be upgraded with a new pixel detector featuring thinner sensors, smaller cells for increased granularity, and a reduction of distance from the beam to 3.5 mm. Tracking stations outside the VELO will be upgraded with improved silicon micro-strip sensors that provide finer granularity and were reduced in thickness lowering the material budget. The remaining part of the tracking system will be replaced by layers of scintillating fibres with a diameter of 250 µm. Finally, the particle identification system will undergo significant upgrades in order to maintain the current performance under the increased instantaneous luminosity. [40, 41]

MoEDAL

The Monopole and Exotics Detector At the LHC (MoEDAL) experiment is designed to find signs of new physics by detecting signatures of highly-ionising particles. A primary goal of the experiment is the search for magnetic monopoles, however it is also sensitive to massive, slow-moving, stable or long-lived charged particles.

The detector is located inside the LHCb experiment’s cavern, constructed around the interaction point. It is a largely passive detector, featuring Nuclear Track Detector (NTD) stacks that show characteristic damage when traversed by highly ionising particles as well as Magnetic Monopole Trappers (MMTs), a paramagnetic trapping volume that can capture electrically and magnetically charged particles in order to further study them at a remote detector facility. Finally, an array of pixel detectors provides monitoring of highly ionising backgrounds in the cavern. [42, 43]
In this chapter a detailed description of the Compact Muon Solenoid (CMS) detector is given. After describing the coordinate system used by CMS, the individual detector subsystems are introduced, beginning with the innermost Pixel detector and finally reaching the muon system, placed outside the large superconducting solenoid. Finally, the CMS Trigger and data acquisition (DAQ) system is introduced.

4.1 Overview

The CMS detector is one of the large experiments at the Large Hadron Collider (LHC) in Geneva, Switzerland. It is installed at one of the four LHC interaction points, 100 m underground close to the French village of Cessy. The detector is intended as a general-purpose particle physics experiment, with emphasis put on superior performance in detecting the then-predicted Higgs boson during the design phase. This is reflected in the advanced calorimetry and muon systems which were foreseen to measure the Higgs decay channels containing muons, electrons, or photons.

The unique environment provided by the LHC implies certain challenges for the detector’s design. At the LHC’s design energy and instantaneous luminosity of $\sqrt{s} = 14$ TeV and $1 \times 10^{34} \text{ cm}^{-2} \text{s}^{-1}$ for proton-proton collisions an event rate of approximately $10^9$ inelastic events per second was expected. As only a fraction of these data can be stored, the online event selection system (trigger) needs to efficiently select about 1000 events per second. The LHC bunch spacing of 25 ns leads to additional requirements on the design of the read-out and trigger systems. In a single bunch crossing an average of 20 inelastic collisions was expected for the LHC Run-1 period. This implies that about 1000 charged particles will traverse the detector every 25 ns. To mitigate the effects of this so-called pile-up, CMS uses high-granularity detectors with good timing resolution, in particular in the areas close to the interaction point where occupancies are higher.

4.1.1 Geometry

CMS’ dominating feature is a large superconducting solenoid capable of creating a 3.8 T magnet field in the interaction zone. It has a diameter of 6 m making it possible to contain both the inner tracking system and most of the calorimetry. In order to deal with large track
multiplicities, CMS employs high precision silicon pixel detectors surrounded by several layers of silicon microstrip detectors for tracking. The inner tracking system is surrounded by the electromagnetic calorimeter (ECAL) which uses lead tungstate (PbWO₄) crystals with coverage in pseudorapidity up to $|\eta| < 3.0$. A preshower system is installed in front of the endcap ECAL for $\pi^0$ rejection. Encompassing ECAL is a brass/scintillator sampling hadronic calorimeter (HCAL) with coverage matching that of ECAL. It is complemented by the iron/quartz-fibre forward hadron calorimeter (HF) calorimeter, covering the range $3.0 < |\eta| < 5.0$. The solenoid’s magnetic field lines run parallel to the beam direction in the central region. The flux is closed by a return yoke made from iron, interspersed with four layers of muon stations. Each muon station consists of several layers of aluminium drift tubes (DTs) in the barrel region, or cathode strip chambers (CSCs) in the endcap region, complemented by resistive plate chambers (RPCs). A cutaway diagram of CMS is shown in fig. 4.1.

Figure 4.1: Cutout view of the CMS detector. The large superconducting solenoid, surrounded by the red return yoke and the muon chambers (white) encloses the HCAL (yellow) and ECAL (turquoise), as well as the inner tracking system (light blue). [44]

### 4.1.2 Coordinate system

CMS’s coordinate system is centred at the nominal interaction point of the LHC particle beams, with the y-axis pointing vertically upward, and the x-axis pointing radially inward toward the centre of the LHC. Thus, the z-axis points along the beam direction.
Due to CMS’s cylindrical shape it is useful to define an azimuthal angle $\phi$ which is measured from the x-axis and a polar angle $\theta$ which is measured from the z-axis. However, it is common to use the so-called pseudorapidity ($\eta$) instead of the polar angle for collider experiments. It is defined as

$$\eta = -\ln \left( \tan \left( \frac{\theta}{2} \right) \right) \quad (4.1)$$

which is approximately the relativistic rapidity in z-direction

$$\xi = \frac{1}{2} \ln \left( \frac{E + p_z}{E - p_z} \right) \quad (4.2)$$

for either small masses or big momenta when $E \approx p$ holds. Pseudorapidity is a useful quantity as particle production is constant as a function of rapidity and differences in pseudorapidity are invariant under Lorentz transformations.

The momentum and energy transverse to the beam direction, denoted by $p_T$ and $E_T$, respectively, are computed from the x- and y-components. The imbalance of momentum measured in the transverse plane is denoted by $E_T^{\text{miss}}$.

### 4.2 Inner tracking

The CMS inner tracking system is designed to allow a highly accurate reconstruction of particle tracks in order to be able to precisely measure momenta and to distinguish primary and secondary vertices\(^1\). During the conception phase it was expected that an average of about 1000 particles from more than 20 overlapping proton-proton interactions would be traversing the tracker for each bunch crossing at the LHC design luminosity of $1 \times 10^{34}$ cm\(^{-2}\) s\(^{-1}\). Given such a large particle flux in particular close to the interaction point a detector technology providing high granularity and fast response times is required, so that particle trajectories can be identified reliably and assigned to the correct bunch crossing. These demands however imply a high power density due to on-detector electronics which requires an efficient cooling system. The above has to be balanced against the desire to limit the amount of material in the particles’ paths in order to minimise multiple scattering, bremsstrahlung, nuclear interactions, and photon conversion which have a detrimental effect on tracking efficiency. A further constraint is the requirement to guarantee the continued operation of the inner tracker within a high radiation environment until its foreseen replacement after ten years. The combination of the requirements on speed, granularity, and radiation hardness outlined above led to the design of a tracker entirely based on silicon detector technology.

The total tracking volume is given by a cylinder of 5.8 m length and 2.6 m diameter. To cope with the extremely high occupancy at the LHC’s design luminosity, the inner tracking layers are realised as a silicon pixel tracker. It provides excellent spatial resolution and can cope with a very high track multiplicity environment, but requires 66 million read-out channels. The outer

---

\(^1\)Primary vertices are the direct result of the interaction of two protons during a bunch crossing. In contrast, a secondary vertex is the consequence of the decay of a collision product that was itself produced in a primary vertex.
layers of the tracker are implemented as silicon microstrip detectors which provide only one spatial coordinate per detector element, but require far fewer read-out channels (fig. 4.2).

![Diagram of CMS inner tracker with pixel detector](image)

Figure 4.2: Schematic illustration of the CMS inner tracker, surrounding the pixel detector. In the central region it is made up of 4 layers in the inner barrel (TIB), and 6 in the outer barrel (TOB). In the endcap there are 3 inner disks (TID) inside the TOB, and 9 disks (TEC) outside it. Each line indicates a detector module. Double lines represent so-called stereo modules capable of measuring both spatial coordinates. [30]

### 4.2.1 The pixel detector

The pixel detector forms the part of the tracking system closest to the interaction region and covers a pseudorapidity range of $0 < |\eta| < 2.5$. It is comprised of the barrel tracker’s three inner layers with radii of 4.4 cm, 7.3 cm, and 10.2 cm respectively, as well as the endcap tracker’s two inner disks at distances of 34.5 cm and 46.5 cm from the nominal interaction point. The layout of the three barrel layers and the forward pixel disks on each side allows for three tracking points over almost the entire $\eta$-range. There are a total of 66 million pixels, each 100 $\mu$m $\times$ 150 $\mu$m in size. The total active region is therefore about 1 m$^2$.

The pixel detector provides precise tracking hits in the $r$-$\phi$ and $z$ directions, and therefore is relied on heavily for good secondary vertex reconstruction.

A pixel sensor is implemented as a reverse-biased diode, detecting a charged particle by the ionisation trail it causes. In the barrel layers the magnetic field induces a Lorentz angle which increases the charge sharing between neighbouring pixels. Charge sharing allows to achieve sub-pixel resolution, which leads to the pixel tracker’s final resolution of about 10 $\mu$m to 20 $\mu$m. The pixel detectors in the forward direction are tilted at an angle of 20° to induce charge sharing and arrive at similar improvements to the resolution.

With growing integrated luminosity the required bias voltage that is needed to reach a depletion depth corresponding to full hit efficiency decreases at first, then increases due to changes in the effective doping caused by particles traversing the tracker. This effect can be
4.2 Inner tracking

counteracted by lowering its operating temperature. The pixel detector was operated at 0 °C during LHC Run-1. After Long Shutdown 1 (LS1) this temperature was lowered to −15 °C. [30, 45, 46, 47]

Upgrade of the pixel detector

The current pixel detector’s read-out electronics were designed for the data rates and occupancies expected at the LHC design luminosity with 25 ns bunch spacing. For these beam parameters there is a dynamic inefficiency of 4 % at the nominal CMS read-out rate from the current read-out chip in the innermost layer. This data loss in the inner layer will increase to 16 % at the instantaneous luminosity of $2 \times 10^{34} \text{cm}^{-2} \text{s}^{-1}$ expected during Run-3. In a possible scenario for the LHC operation after LS1 with 50 ns bunch spacing the losses would have increased almost exponentially to nearly 50 % in the inner layer.

In order to maintain or exceed the excellent performance of the current pixel detector in a high luminosity environment the existing detector was replaced with a new high efficiency, low mass detector with four barrel layers and three endcap disks, which can provide four hit coverage up to pseudorapidities of $|\eta| < 2.5$. In order to place the inner pixel layer closer to the interaction region a new smaller diameter beam pipe was required. The installation of the new beam pipe was completed during LS1. The upgraded system is foreseen to withstand the high luminosity environment of the LHC through Long Shutdown 3 (LS3) as well as continue to provide high tracking performance at pile-up up to and exceeding 50.

The CMS detector has been designed to allow fairly easy access to the central detector systems, the pixel detector in particular being designed with the ability to be removed and reinstalled in CMS during a LHC year-end technical stop. This allowed the schedule for the pixel detector upgrade to include provisions for an early replacement before Long Shutdown 2 (LS2), which was realised in the year-end technical stop of 2016/2017. This earlier replacement gives CMS the opportunity to fully profit from the higher LHC luminosity provided before LS2 with a higher performance pixel detector. [32]

4.2.2 The silicon strip tracker

The CMS silicon strip tracker is segmented into 9.6 million p+ strips that are implanted into n-type bulk with a thickness of 320 µm in the inner layers and 500 µm in the outer layers and n-type back side. In the barrel there are 10 tracking layers placed at radii of 25 cm to 110 cm while the endcaps is equipped with 12 disks placed between 70 cm and 280 cm from the nominal interaction point in the z-direction (fig. 4.2). Four layers in the barrel and several disks in the endcaps are equipped with stereo modules that are constructed from two silicon modules glued back-to-back with their strips aligned at an angle of 100 mrad, thereby allowing local 2D measurement.

The strip size varies between $81 \mu m \times 10 \text{cm}$ and $183 \mu m \times 25 \text{cm}$ with resolution generally decreasing as a function of distance from the interaction point. The spatial resolution reaches $15 \mu m$ to $40 \mu m$ in the barrel area of the silicon strip detector.

As described in section 4.2.1, the tracker suffers significant radiation damage with growing luminosity that can be counteracted by operating it at lower temperature. During LHC Run-1
the silicon strip tracker was operated at 4 °C. The temperature was lowered to −15 °C after LS1. [46, 47, 48, 49]

4.3 Calorimetry

The CMS calorimeters measure the energy of leptons and photons, as well as hadrons. The CMS ECAL is optimised for the precise measurement of the energy of traversing electrons and photons, as well as a fast response time. Surrounding it, the HCAL is constructed as a sampling calorimeter with large mass in order to stop hadrons at the expected energies at the LHC.

4.3.1 Electromagnetic calorimeter

The primary benchmark while designing the CMS ECAL was the $H \rightarrow \gamma\gamma$ channel. This decay process produces a sharp peak above the background for Higgs boson masses of 140 GeV and lower, which leads to a strict constraint of 1% at 100 GeV on the ECAL energy resolution. Additional requirements are good angular resolution as well as accurate pion/photon separation.

The CMS ECAL is a hermetic, homogeneous calorimeter, constructed from 61,200 lead tungstate (PbWO$_4$) crystals in the barrel and 7,324 crystals in the endcaps. The ECAL barrel (EB) covers a pseudorapidity region of $|\eta| < 1.479$, which the endcaps extend to $|\eta| < 3$.

The barrel is split into 360 segments in $\phi$ and $2 \times 85$ segments in $\eta$. The crystals are mounted in a way so that their axes make an angle of $3^\circ$ with respect to the vector from the nominal interaction point in both $\eta$ and $\phi$ projections to avoid inefficiencies due to cracks between the crystals aligning with particle trajectories. The inner face of each crystal in the barrel measures $22 \text{ mm}^2 \times 22 \text{ mm}^2$ while the outer face measures $26 \text{ mm}^2 \times 26 \text{ mm}^2$. Each crystal is 23 cm long (about 26 radiation lengths).

The endcaps consist of identically shaped PbWO$_4$ crystals grouped in units of $5 \times 5$ crystals, called supercrysials. The crystals have an inner face area of $28.62 \text{ mm}^2 \times 28.62 \text{ mm}^2$ and an outer face area of $30 \text{ mm}^2 \times 30 \text{ mm}^2$, and are arranged to point at a focus 1,300 mm beyond the interaction point, leading to off-pointing angles from 2° to 8°. They are 22 cm long, corresponding to almost 25 radiation lengths. A schematic of CMS’s ECAL system is given in fig. 4.3.

Lead tungstate was chosen as scintillation material as it possesses high density, small Molière radius, and small radiation length which allow to construct a compact calorimeter with fine granularity. Furthermore, the scintillation decay time is of the same order of magnitude as the LHC bunch crossing period. The comparatively small light light yield provided by PbWO$_4$ crystals can be mostly mitigated by modern photo diodes. The light yield is highly dependent on the operating temperature, requiring constant cooling to 18°C.

Finally, radiation damage to the PbWO$_4$ crystals leads to a wavelength-dependent loss of light transmission, however without changes to the scintillation mechanism. This effect can thus be measured and corrected for by monitoring the optical transparency with injected laser light. [50, 51, 52]
4.3 Calorimetry

The preshower detector

The preshower detector is a sampling calorimeter consisting of two layers: lead radiators induce electromagnetic showering from incoming photons and electrons while silicon strip sensors behind each radiator determine the deposited energy as well as the transverse shower profiles. It is placed in front of the ECAL endcaps in the pseudorapidity region of $1.65 \lesssim |\eta| \lesssim 2.6$ to identify neutral pions and minimum ionising particles. Additionally, it contributes to the identification of electrons against minimum ionising particles and provides position information for electrons and photons with high granularity. [53]

4.3.2 Hadron calorimeter

The hadronic calorimeter (HCAL) is particularly important for the measurement of hadronic jets as well as the detection of neutrinos or exotic particles that manifest in the detector as an apparent missing transverse energy. For this reason the CMS HCAL has a large hermetic coverage as well as fine lateral segmentation. It is implemented as a brass/scintillator sampling calorimeter with coverage up to $|\eta| < 3$, split into a barrel (HB) and an endcap (HE) part (fig. 4.4).

In the barrel, the space between the ECAL and the superconducting solenoid is not sufficient to completely absorb the most energetic hadrons, the magnet itself is thus used as absorption material by the outer barrel part (HO) which consists of one layer of scintillators placed immediately before the barrel muon system as a tail-catcher. This design ensures that hadronic showers are sampled with nearly 11 hadronic interaction lengths.
Chapter 4 The Compact Muon Solenoid experiment

Figure 4.4: The CMS HCAL system consists of the barrel part (HB), the outer barrel part (HO), the endcap calorimeter (HE), and the forward hadron calorimeter (HF). This provides coverage up to very high pseudorapidities. [30]

Coverage is extended up to $|\eta| < 5.0$ by the HF. It is implemented as an iron/quartz-fibre calorimeter, due to the very challenging radiation environment. Cherenkov light is produced in the quartz fibres which is then measured in hybrid photodiodes. [30, 54]

4.4 Forward detectors

The CMS very forward detectors work in tandem with TOTal cross section, Elastic scattering and diffraction dissociation Measurement at the LHC (TOTEM) (described in section 3.5.3) in order to provide coverage in very high $\eta$ ranges.

4.4.1 Castor

The Centauro and Strange Object Research (Castor) detector is a electromagnetic and hadronic sampling calorimeter, similar in construction to HF. It consists of tungsten and quartz layers, positioned around the beam pipe in the very forward region of the CMS detector, behind HF at 14.385 m from the nominal interaction point and covering a pseudorapidity range of $5.1 < |\eta| < 6.55$ (fig. 4.5). The Cherenkov light produced in the quartz scintillators is collected by photomultipliers through air-core light guides.

Castor is designed to complement the CMS nucleus-nucleus physics program as well as the diffractive and low-x physics in proton-proton collision runs. [30, 55]
4.4.2 Zero degree calorimeters

Two zero degree calorimeters (ZDCs) are placed on both sides of CMS at a distance of 140 m from the interaction point and cover the pseudorapidity region of $|\eta| \geq 8.3$. They are placed behind the LHC beam separator which means that only neutral particles reach it.

The ZDCs are used to complement CMS’s very forward region for heavy ion and proton-proton diffractive studies. Each ZDC is split into two independent sections: an electromagnetic and a hadronic part (fig. 4.6). Both parts of the ZDCs are implemented as tungsten plates and quartz fibres similar to HF and Castor. [56]

4.5 The superconducting solenoid

The CMS experiment’s superconducting solenoid provides a magnetic field with field lines parallel to the beam-line and a field strength of 3.8 T at the interaction point. It measures 6 m in diameter and 12.5 m in length, with a total weight of 220 t. A distinctive feature is the 4-layer winding made from a stabilised, reinforced NbTi conductor that provides an energy density of 11.6 kJ/kg which leads to a deformation of 0.15 % during energising.
Figure 4.6: Side view of the ZDC. On the left side is the electromagnetic section (EM) followed by the hadronic section (HAD). [56]
4.6 The muon system

As explained in section 4.1, the muon system is of central importance to the CMS physics goals. For instance, the $H \rightarrow ZZ \rightarrow 4\ell$ decay channel is considered gold-plated if all the leptons in it are muons. This follows from the relative ease of muon identification outside the superconducting solenoid as well as the high 4-particle mass resolution that can be achieved when all leptons are muons due to their low radiative losses in the tracking system when compared to electrons. [57]

Amongst the main design goals for the CMS muon system were accurate muon identification and momentum resolution over a wide range of transverse momenta and angles, a good double muon mass resolution, as well as the ability to determine the charge of muons with $p_T < 1$ TeV with high precision. A further requirement was the possibility to include information from the muon system in the Level-1 trigger. Finally, high radiation tolerance — in particular in the endcaps — and low cost due to the large area covered were desired features.

The CMS muon system is embedded in the return yoke of the superconducting solenoid and is composed of three distinct detectors: DTs in the barrel, and CSCs in the endcaps are complemented by an RPC system. DTs were chosen in the barrel region where the neutron-induced background is small due to their lower cost and very high position resolution. To mitigate their low timing accuracy due to drift times on the order of 400 ns, the DT system is complemented by RPCs with a response time of about 1 ns. Due to the harsh radiation environment and inhomogeneous magnetic field in the forward region, CSCs were chosen as high precision detectors in the endcaps (fig. 4.7).

![Figure 4.7: The muon detectors in a longitudinal view in the configuration used before LS1. DT detectors in the barrel and a CSC system in the endcaps are complemented by RPCs. The fourth disk in the endcaps was not fully instrumented for Run-1, but retrofitted during LS1.][58]

---

[57]: Reference
[58]: Reference
The muon system primarily improves transverse momentum resolution for muons with $p_T > 200$ GeV in the reconstruction stage (fig. 4.8), provides strong muon identification, and supplies track data to the Level-1 trigger.

![Figure 4.8: Transverse momentum reconstruction resolution for barrel and endcap region. The muon system significantly improves the resolution for high transverse momentum muons.][30]

### 4.6.1 Drift Tubes

Due to the low muon rate and neutron-induced background as well as the uniform magnetic field, drift tube (DT) detectors can be used to reconstruct muon tracks with high spatial precision. The DT system covers a pseudorapidity range of $|\eta| < 1.2$ and is segmented into five wheels along the $z$-direction, each in turn divided into twelve $30^\circ$ sectors. Each sector contains four layers of muon stations, with each muon station containing both the DT and RPC detectors.

Four layers of DT cells are grouped to a superlayer (SL) which comprises the smallest independent unit of the DT detector design. The three stations closest to the interaction region contain 3 SLs; two measuring the coordinate in the $r$-$\phi$ bending plane and one providing a measurement in the $z$-direction. The outermost station does not contain the SL to measure the $z$-coordinate. The two SLs measuring the $\phi$ coordinate are separated maximally in radial distance in order to achieve the best angular resolution (fig. 4.9).

The DTs contained in the muon stations are 2.5 m long, their face measuring $42\,\text{mm} \times 13\,\text{mm}$. They are operated with an Ar/CO$_2$ mixture at atmospheric pressure. The single-wire resolution
4.6 The muon system

Figure 4.9: Depiction of the DT system in the CMS detector. A muon station (or DT chamber) consists of three SLs that are made of 4 layers of DT cells. A SL measuring the z-coordinate (in green) and a honeycomb spacer (red) are positioned between two \( \phi \)-measuring SLs (blue). [59]

for a DT cell is about 250 \( \mu \text{m} \) which is improved upon by utilising data from neighbouring wires to arrive at a resolution of approximately 100 \( \mu \text{m} \). [30, 58]

4.6.2 Cathode Strip Chambers

In the forward regions of the CMS detector the muon rates and neutron-induced background are significantly higher than in the barrel region. In addition the magnetic field is stronger and inhomogeneous, making drift time calculations very complicated. For this reason cathode strip chambers (CSCs) are used as high-precision spatial tracking detectors in the pseudorapidity region of \( 0.9 < |\eta| < 2.4 \).

A given CSC consists of an anode wire plane within a gas volume and enclosed by two copper cathodes. One of the cathodes is segmented in strips that run orthogonal to the anode wires in order to provide a complementary position measurement. CSCs feature a fast response time due to a short drift path, can be finely segmented, and are able to tolerate the non-uniformity of the magnetic field. [60]

Each endcap consists of four muon stations that can contain both CSC and RPC detectors, similar to the barrel region. The stations are attached to the faces of the endcap steel disks, perpendicular to the beam direction. The outermost disk was not fully instrumented during LHC Run-1 with the station furthest from the interaction region missing (fig. 4.7). A CSC is constructed from six layers, each measuring the muon track in two coordinates. The cathode strips are arranged to run radially outward in order to provide precision measurements in the
Chapter 4 The Compact Muon Solenoid experiment

r-\(\phi\) plane while the anode wires measure the muon track’s coordinate in the radial direction. They are ganged into groups to reduce the number of read-out channels (fig. 4.10).

Figure 4.10: Left: Illustration of a CSC showing the six layers and the orientation of the strips and wires. Right: View of the gas gap in a CSC showing the cathode strips and the anode wires, as well as an illustration of the charge distribution induced by the gas ionisation avalanche on the cathode strips. [58]

The chambers in the inner ring of station 1, named ME1/1 are split in half at \(|\eta| = 2.1\), essentially acting as two independent chambers for read-out and trigger purposes. The region closer to the beampipe is called ME1/1a while the outer region is called ME1/1b. In Run-1 the cathode strips in ME1/1a were ganged in groups of three in order to reduce the number of read-out channels from 48 to 16. During LS1 this ganging was reversed in order to improve the spatial resolution of these chambers.

The spatial dimensions of the CSC detectors vary throughout the system, but range from 1.7 m to 3.4 m in length and subtend an angle from 10° to 20°. The chambers use an Ar/CO\(_2\)/CF\(_4\) gas mixture.

Upgrade of the CSC system during long shutdown 1

As mentioned in section 4.6.2 the outer ring of the ME4 disk had not been instrumented for LHC Run-1. In LS1 these stations were installed together with an additional layer of yoke disks that shield the chambers from environmental noise (fig. 4.11). The addition of a further layer of muon stations in the pseudorapidity region of 1.2 < \(|\eta|\) < 1.8 increases redundancy and enables the trigger to successfully build a high-quality track more often.

A further upgrade was the replacement of the read-out electronics for station ME1/1a, allowing the full detector granularity to be used. This improved the precision of the trigger primitives found in the pseudorapidity range of 2.1 < \(|\eta|\) < 2.4. [61, 62]
4.6 The muon system

4.6.3 Resistive Plate Chambers

The resistive plate chamber (RPC) system provides CMS with a complementary muon system up to a pseudorapidity of $|\eta| < 1.6$. Its excellent timing resolution of about 1 ns allows unambiguous bunch crossing identification, however the position resolution is coarser than that of the DT or CSC detectors.

In the barrel there are six layers of RPCs installed. Two layers in each of the two innermost stations and one layer in each of the two outer stations. The duplication closer to the interaction region allows an accurate measurement even for low transverse momentum muons that may not reach the outer stations. During LHC Run-1 the endcap contained three layers of RPC detectors, complementing the three inner CSC chambers (fig. 4.7).

The RPCs are double-gap chambers that are operated in avalanche mode, ensuring good operation at high rates. Two 2 mm thick bakelite plates separated by 2 mm form a gas gap filled with a mixture of 95.2 % $\text{C}_2\text{H}_4\text{F}_4$, 4.5 % $\text{C}_6\text{H}_{10}$, and 0.3 % $\text{SF}_6$. The outer surface of the plates is covered with a conductive graphite layer in order to apply a voltage of 9.6 kV. Read-out strips are placed between two adjacent gas gaps. When a muon crosses an RPC, the gas in both gaps is ionised which leads to an avalanche due to the high electric field. This field is then measured by the read-out strips (fig. 4.12). [58, 63]

Upgrade of the RPC system during long shutdown 1

The RPC system profited from the same upgrade as the CSCs, described in section 4.6.2. An additional layer of RPC detectors was added in the endcap, significantly improving reconstruction efficiency and background rejection in the Level-1 trigger (fig. 4.11). [61, 62]
Chapter 4 The Compact Muon Solenoid experiment

4.7 Trigger and data acquisition

At standard running conditions the CMS detector produces 1 MB to 2 MB of data per event after zero suppression. If every such event were to be read out at the nominal LHC bunch crossing rate of 40 MHz this means at least 40 TB/s data rate which cannot be stored or feasibly analysed with currently available computing and storage capacities. Furthermore, the detector front-ends only support a maximal read-out rate of 100 kHz for the full event information. For these reasons an online event selection system, called the trigger is introduced.

The vast majority of produced events at the LHC are comprised of well-known physics processes or so-called minimum bias events — events with almost no interaction between the protons in the colliding bunches. Conversely, the cross-sections for interesting events vary over orders of magnitude. While the production of $b \bar{b}$ has a cross-section on the order of $1 \times 10^{-3}$ b, the cross-sections for the production of the Higgs boson in the relevant mass region is on the order of $1 \times 10^{-12}$ b (fig. 4.13).

To efficiently record useful physics data it is imperative to provide a very flexible trigger system that can not only be programmed to accept events with widely varying signatures but also record relatively common event types at a limited rate.

The CMS experiment uses a two-tiered trigger system to select events (fig. 4.14). The Level-1 trigger, built from specialised, low latency hardware provides a first selection within less than 4 µs. It receives coarse data from the muon and calorimetry systems as well as simple pulses.

![Figure 4.12: Schematic view of a generic RPC layer. Two gas gaps are separated by a read-out strip.][58]
Figure 4.13: The cross-sections for processes studied at the LHC vary over orders of magnitude. The right hand side gives the interaction rates at nominal luminosity. [64]
from other systems, such as the Beam Pickup Timing for the eXperiment (BPTX) system that indicates the presence of LHC hadron bunches in the interaction point. The event data are stored in ring buffers on the detector or in the underground counting room for the duration of the Level-1 trigger latency and are only read out if the event is determined to be useful. The read-out rate has to be kept below 100 kHz due to limitations in the detector front-ends as well as restricted bandwidth and computing power in the High-Level trigger (HLT) system. A detailed overview of the Level-1 trigger system is given in chapter 5.

![Diagram of the CMS read-out chain](image)

**Figure 4.14:** The CMS read-out chain. The Level-1 trigger reduces the incoming event rate of 40 MHz provided by LHC to about 100 kHz. Read-out units (RUs) receive event fragments at this rate and provide them to the builder units (BUs) which build full events. There events are then operated on by the CMS filter systems (HLT). Event flow is directed by the event manager (EVM). Red arrows represent instructions while blue, dashed arrows stand for status information. The output event rate of the filter units is less than 1 kHz which is stored on tape. [30]

The HLT event filter is implemented as a large computing farm, operating on the event data with algorithms very close to the ones used for offline analysis. In each event, objects such as electrons, muons, and jets are reconstructed and only those events which are of interest for physics analyses are stored.

During Run-1 event fragments for each bunch crossing accepted by the Level-1 trigger were read out by the data acquisition (DAQ) system from the front-end modules and assembled into super fragments using Myrinet switches in a first event building stage before being delivered to read-out units (RUs). Builder units (BUs) that assembled complete events received the super fragments via a large switching fabric implemented in Gigabit Ethernet (GbE). The event manager (EVM) managed the flow of data between RUs and BUs based on trigger information. The filter units (FUs) which executed the HLT algorithms shared the same physical computer as their corresponding BU (fig. 4.14). The CMS DAQ for Run-1 was logically split into eight slices that could record data independently of each other.
4.7 Trigger and data acquisition

Event building on these combined BU/FU units proceeded by events being preassembled in the processor memory of the BU. An independent process that ran the HLT algorithms analysed each event and determined whether it should be accepted for storage on tape. In addition to filtering the Level-1 output stream, the FUs also provided online data-quality monitoring (DQM) that allowed operators to identify and diagnose problems almost in real-time during runs.

The HLT components were controlled by the central run-control (RCMS) via a hierarchical structure of finite state machines. [65]

4.7.1 Upgrade of the DAQ and HLT during long shutdown 1

During LS1 a significant upgrade of the CMS DAQ system was performed in order to cope with the increased number of read-out channels as well as larger data fragments expected as the result of higher instantaneous luminosity.

The first level of event building was reimplemented using 10/40 GbE instead of the Myrinet-based [66] solution employed during LHC Run-1. Similarly, the event builder network connecting the RUs to the BUs, based on GbE during Run-1 has been reimplemented using InfiniBand [67]. InfiniBand offers 56 times the performance of Gigabit Ethernet links, however at a significantly larger cost. For this reason the BU and FU functionalities were split in the upgraded DAQ in order to keep the InfiniBand network comparatively small: 62 BU nodes are connected to the event builder network, writing complete events to a large RAMdisk. The FUs are then connected via a large Ethernet network based on 40 GbE switches that provide 10 GbE breakout. Legacy nodes with 1 GbE interfaces can be connected via 540-port 1/10 GbE switches. The allocation of FUs to a BU is done statically according to the relative performance of the FUs. [68]
Chapter 5

The Level-1 Trigger

This chapter contains a detailed description of both the legacy and upgraded Level-1 trigger systems. After giving an introduction to the fundamentals of online event selection, the legacy trigger system — in operation from 2009 until the end of 2015 — is described. Subsequently the need for the upgraded Level-1 trigger system is established and a system description is given while contrasting it to the legacy system.

5.1 Introduction

The Level-1 trigger of the Compact Muon Solenoid (CMS) experiment is responsible for reducing the event rate from the nominal Large Hadron Collider (LHC) bunch-crossing frequency of 40 MHz to about 100 kHz which is the maximum read-out frequency for the full event information due to limitations in detector front-ends. Currently only calorimeter and muon detector data are used in the Level-1 trigger. Information from the two detector systems is processed separately until the highest level of the Level-1 trigger.

Due to the low-latency requirement for the Level-1 trigger, relatively simple algorithms are run in parallel on custom hardware (field-programmable gate arrays (FPGAs)\(^1\) and application-specific integrated circuits (ASICs)\(^2\)). Logic, memory, and bandwidth constraints, even in large FPGAs and ASICs, require the Level-1 trigger to operate in a “divide and conquer” fashion: So-called trigger primitives are found locally in early stages of the trigger chain and successively combined to physics objects such as muons, electrons or photons, jets, as well as various energy sums like the total transverse energy. These physics objects are subsequently received by a global stage, where they are processed further before the highest-ranked of these objects are sent to the Global Trigger (GT). The GT issues a read-out signal based on programmable algorithms working on the received physics objects. Each of these algorithms is used to select a certain class of physics events. Furthermore, the Level-1 trigger is equipped with a mechanism to throttle the trigger accept rate for a given algorithm by a factor \(n\), called prescale. These features make it possible to formulate trigger algorithms in a natural way — e.g. to record only one in every 50 events that contain two muons and one jet, each above certain transverse momentum thresholds and fulfilling certain quality criteria.

As mentioned before it is not feasible to read out the full detector information at 40 MHz. For this reason the bulk of the detector information is buffered on-detector or in the underground.

---

\(^1\)An integrated circuit that can be configured by the user after manufacturing.  
\(^2\)An integrated circuit customised during manufacturing for a particular use, as opposed to general-purpose usage.
counting room until a read-out decision is received from the Level-1 trigger. The maximal latency for the read-out decision to be made is determined by the tracker and preshower detector buffers that contain space for the detector data collected during 160 bunch crossing intervals, which translates to a time of less than 4 µs. The actually available time for processing the data is reduced even further after considering delays introduced due to the drift time in the drift tube (DT) chambers and the signal propagation time due to the distance between the detector and the counting room, where a large part of the trigger electronics is located. The time required for the signal propagation sums up to about 1 µs. [69]

5.2 The Level-1 trigger for LHC Run-1

The Level-1 trigger system used during LHC Run-1 and the first year of Run-2, is based on Versa Module Europa (VME) technology [70] utilising mainly FPGAs as well as ASICs with mostly galvanic parallel links for inter-card communication. It is synchronised to the LHC clock of 40 MHz and works in a fully-pipelined mode. Initially trigger primitives are found independently in the calorimeter and muon system — i.e. track segments or energy clusters — which are then merged in regional trigger stages to produce trigger objects. Finally, these objects are received and operated on in the Global Calorimeter Trigger (GCT) and Global Muon Trigger (GMT), which forward the best candidates to the GT (fig. 5.1).

5.2.1 The calorimeter trigger

The calorimeter trigger identifies the four most energetic physics objects from each of isolated and non-isolated electrons or photons, jets in the forward and central regions, and τ-like jets (table 5.1).

<table>
<thead>
<tr>
<th>Physics object</th>
<th>η region</th>
</tr>
</thead>
<tbody>
<tr>
<td>isolated electron or photon candidates</td>
<td></td>
</tr>
<tr>
<td>non-isolated electron or photon candidates</td>
<td></td>
</tr>
<tr>
<td>jets in the central region</td>
<td></td>
</tr>
<tr>
<td>jets in the forward region</td>
<td>3 &lt;</td>
</tr>
<tr>
<td>τ-like jets</td>
<td></td>
</tr>
<tr>
<td>total transverse energy</td>
<td>N/A</td>
</tr>
<tr>
<td>transverse hadronic energy</td>
<td>N/A</td>
</tr>
<tr>
<td>missing energy</td>
<td>N/A</td>
</tr>
</tbody>
</table>

Table 5.1: Physics objects found by the calorimeter trigger.

Furthermore, it computes these objects’ transverse energy and position, identifies the bunch crossing the objects were created in, and calculates the total missing energy as well as the total transverse energy in the calorimeters.

The calorimeter trigger chain is logically segmented into several stages. The Trigger Primitives Generator (TPG) sub-system which is integrated in the calorimeter read-out is primarily
5.2 The Level-1 trigger for LHC Run-1

Figure 5.1: The Level-1 trigger of the CMS experiment is built in a highly modular way. In both the muon and calorimeter trigger chains trigger primitives are built in a first step. These are sent to regional trigger stages, which construct the final trigger objects and assign values for transverse momentum or energy, quality, and spatial coordinates to them. Subsequently the Global Muon and Global Calorimeter Triggers send the best of these objects on to the Global Trigger which can issue a signal to initiate a detector read-out.
Chapter 5 The Level-1 Trigger

responsible to assign the correct bunch crossing to a given detector pulse and calculate the transverse energy sum for each so-called trigger tower.

Trigger towers are defined as \((\eta, \phi)\)-regions in both the electromagnetic calorimeter (ECAL), where they are used for electron/photon identification, and the hadronic calorimeter (HCAL). Their size is defined to be equal to that of a physics HCAL tower. This corresponds to \(0.087 \times 0.087\) in \(\eta \times \phi\) in the detector region with \(|\eta| < 2.1\). Beyond this the size increases in \(\eta\) with growing pseudorapidity. The size of these towers for the ECAL was chosen in a trade-off between background-suppression which is improved with smaller tower sizes and cost issues incurred due to a larger number of trigger channels. It corresponds to \(5 \times 5\) PbWO\(_4\) crystals in the barrel. [71]

A group of \(4 \times 4\) trigger towers is combined into a so-called calorimeter region which forms the basis for jet objects and energy values that are calculated in the Regional Calorimeter Trigger (RCT). In the forward hadron calorimeter (HF) region a calorimeter region consists of only one trigger tower (fig. 5.2).

![Figure 5.2: Trigger towers and calorimeter regions in the barrel of CMS. [69]](image)

The RCT consists of 18 processing crates. It receives the energy deposit in each trigger tower in the barrel and endcap region as well as in the HF calorimeters via the TPGs. All operations in the RCT are performed at the level of calorimeter regions. For each region in the barrel and endcap the RCT builds the total energy sum, isolated and non-isolated electron/photon candidates, as well as a so-called \(\tau\)-veto bit that is used to identify \(\tau\)-like jets. The \(\tau\)-veto bit is set if there are more than two active trigger towers in a given region. In the Level-1 trigger no distinction between electrons and photons is possible as this requires information from the inner tracker. RCT regions from the HF calorimeters are forwarded without significant processing to the GCT (fig. 5.3). [72]

In the GCT the electron/photon candidates received from the RCT are sorted while in parallel jets are reconstructed from the energy sums calculated by the RCT and sorted. Additionally, if none of the \(\tau\)-veto bits are set in the RCT regions that were used to build a given jet, the GCT tags it as \(\tau\)-like. Finally, the number of jets, the global transverse energy as well as the total missing energy are calculated and sent to the GT along with the best four isolated and non-isolated electron/photon, jet, and \(\tau\)-like objects.

The GCT additionally contributes to luminosity monitoring by providing transverse energy
5.2 The Level-1 trigger for LHC Run-1

Figure 5.3: Flow diagram for the operation of the RCT. Trigger primitives are constructed by the TPGs and used to construct electron/photon candidates as well as energy sums and a so-called $\tau$-veto in the RCT. They are then forwarded to the GCT. Energy values from HF are forwarded with minor preprocessing to GCT. The GCT sorts the received electron/photon objects, finds jets, and calculates energy sums before forwarding them to the GT. The transmission of minimum-ionising particle (MIP) and Quiet bits as indicated in this figure was not fully implemented. [73]
5.2.2 The muon trigger

The purpose of the Level-1 muon trigger is the fast reconstruction of muons with high efficiency while keeping the fake rate low. In order to accomplish this, correct bunch crossing assignment and muon identification, as well as good transverse momentum ($p_T$) resolution over a wide range of $p_T$ are critical. In particular, correct transverse momentum assignment is required in order to be able to reject the large background of low-$p_T$ muons while retaining high-$p_T$ muons that can indicate interesting physics.

The Level-1 muon trigger uses information from drift tube (DT), cathode strip chamber (CSC), and resistive plate chamber (RPC) detectors in dedicated track finding systems to reconstruct muon tracks. These tracks are subsequently combined and sorted in the GMT before being sent on to the GT (fig. 5.4).

Local triggers

Both DT and CSC detectors form local track segments prior to transmission to their respective track finders.

In the DT system, local trigger (DTLT) track segments are reconstructed in trigger boards housed in so-called mini-crates that are mounted on each of the 250 DT chambers. A pipelined architecture of ASICs performs bunch crossing assignment for each hit before constructing track segments in the $r$-$\phi$ plane and finding hits in the $\eta$ direction. In the $r$-$\phi$ plane up to two DTLTs can be reconstructed in a chamber per bunch crossing. These segments contain information on the position, the bending angle, as well as a measure of the reconstruction quality that depends on the number of DT layers that were used to build the track segment. The DTLTs are sent to the sector collector (SC) system which is installed on the balconies around the detector. Hits along the $\eta$ direction are sent to the SC system without processing as no bending in pseudorapidity is assumed.

The SC system is implemented using flash-based FPGAs because of their high tolerance against corruption of the configuration memory due to impact of low energy neutrons that dominate the radiation environment outside the yoke and is comprised of 60 modules, each covering a DT sector (see section 4.6.1). Its purpose is the synchronisation and concentration of the incoming DTLTs and hits along the $\eta$ direction before transmitting this information via one 6 Gbit/s link per sector to the underground counting room where the data are delivered to the Drift Tube Track Finder (DTTF) system. [58, 59]

The CSC local trigger operates by finding local charged-tracks (LCTs) which are comprised of the position, angle, and bunch crossing information for the candidate track. The forward region is especially challenging due to high occupancy and the considerable particle momentum parallel to the beam direction in this region of the detector. To suppress this large background cathode strip signals are digitised with half-strip width, which allows a resolution of 0.15 strip widths to be reached.

The CSCs measure radial and azimuthal coordinates separately using the anode wires and cathode strips, respectively. The anode wire electronics was specifically optimised for bunch
5.2 The Level-1 trigger for LHC Run-1

Figure 5.4: Flow diagram for the muon trigger. DT and CSC each form local trigger track segments within the detector elements to be used in their respective track finder system. In contrast, information from the RPC detector is directly combined to tracks by the Pattern Comparator Trigger (PACT). The results found in the regional triggers for each sub-system are then forwarded to the GMT which merges and sorts candidates in order to send the best four muon candidates to the GT. [69]
crossing identification while the cathode strip electronics was designed to optimally measure position. In contrast to the DT local trigger, the two track components are already combined at the local trigger level in the trigger motherboard (TMB). The TMB then forwards up to two LCTs to a muon port card (MPC). In total there are 60 MPCs housed in peripheral crates within the experimental cavern. Each MPC receives data from nine TMBs and sends the best three LCTs to the Cathode Strip Chamber Track Finder (CSCTF) via optical transmission. [58, 74]

Muon track finders

In the DT and CSC trigger systems, dedicated track finders use the track segments constructed during the local trigger stage to build muon tracks. The DTTF and CSCTF are designed to reconstruct muon tracks in the barrel and endcap and determine their transverse momentum $p_T$, their spatial coordinates $\eta$ and $\phi$, as well as compute a measure for the quality of their measurement.

The candidates are then sorted by their rank which depends on the candidate’s transverse momentum and reconstruction quality in order to send the best four muon candidates each in endcap and barrel to the GMT.

As can be seen in fig. 4.7, a substantial overlap region exists between the DT and CSC coverage for muons originating from the interaction point. In order to improve the reconstruction quality of muon tracks found within this overlap the two track finders share information from the overlapping stations.

The Drift Tube Track Finder. The Drift Tube Track Finder (DTTF) receives up to two DTLTs per sector from the DT SC. It uses these track segments to construct a full track and assign it a value for the muon’s transverse momentum, spatial coordinates as well as the quality of the reconstruction. For practical reasons the DTTF works on regions of $30^\circ$ in the $\phi$ angle, called wedges. Each wedge is itself split into six sectors in the $z$-direction. A sector correlates with the area of one detector wheel, except for the central wheel which is split into two sectors to avoid asymmetries (fig. 5.5).

As outlined in section 5.2.2 the DT system provides separate measurements in the $r$-$\phi$ plane and in the $\eta$ direction to the DTTF. The DTTF maintains this separation during track reconstruction.

Tracks in the $r$-$\phi$ plane within a sector are reconstructed by a Sector Processor (SP) using the DTLTs received for a given sector as well as additional track segments from neighbouring sectors in order to efficiently reconstruct muon tracks crossing sector boundaries. SPs for the outermost wheels additionally receive information from the CSC sub-system as explained before.

Reconstruction proceeds by extrapolating from each DTLT track segment from the two inner stations to all outer stations (i.e. from station 1 to stations 2 through 4, and from station 2 to stations 3 and 4) as well as from station 4 to station 3 by using the bending angle $\phi_b$ to find extrapolation targets; there is no extrapolation from station 3 as the bending angle in this station is negligible due to the magnetic field configuration. The extrapolation uses the position and bending angle encoded in the DTLTs and compares the position of the actual track segments in the examined station to the extrapolated one. The extrapolation is accepted as successful if an actual track segment is found within a programmable window of the extrapolated value. In this
5.2 The Level-1 trigger for LHC Run-1

Figure 5.5: The barrel region is logically split into wedges which are then again split into 6 sectors each. [69]

way all stations are linked in order to form a full track candidate (fig. 5.6). The DTTF algorithm only follows a muon one step into a neighbouring wedge, i.e. it uses at most one track segment from a neighbour wedge in track reconstruction. The $p_T$ value is then assigned based on the difference in azimuthal coordinates of the two inner track segments. Finally, a quality code is computed based on the number of track segments that could be used to construct the track.

Tracks in $\eta$ are constructed by a pattern matching algorithm: The pattern of $\eta$ hits received from the SCs is compared to predefined track patterns. These predefined patterns are basically straight-line patterns, but take into account multiple scattering and the actual magnetic field configuration.

Each SP sends up to two tracks to the wedge sorter which also receives $\eta$ tracks and attempts to match the $\eta$ and $\phi$ tracks before selecting the two best candidates based on their transverse momentum and reconstruction quality (fig. 5.5). If the matching process does not succeed, the standalone $\phi$ track is assigned a coarse $\eta$ value based on which muon stations were used for the track’s construction. The $\eta$ and $\phi$ coordinates of the selected tracks are measured at a reference plane which was defined to bisect station 2. If no track segment from station 2 is present the $\phi$ value is obtained by extrapolation from the innermost station with an available track segment.

The barrel sorter receives the best two muons from each wedge and sorts them before sending the best four tracks to the GMT. Both the wedge and barrel sorter apply ghost busting algorithms based on track addresses to incoming muon candidates in order to remove potential duplicate tracks due to sharing of information at sector and wedge boundaries, respectively. [75]
Chapter 5  The Level-1 Trigger

Figure 5.6: The SP reconstructs muon tracks in the $r$-$\phi$ plane by extrapolation, determines the best two tracks and finally calculates the transverse momentum, spatial coordinates and a quality code for the final tracks. In parallel tracks in the $\eta$ direction are found using a pattern-based approach. [69]
5.2 The Level-1 trigger for LHC Run-1

**The Cathode Strip Chamber Track Finder.** Track finding in the endcap region is highly challenging due to the large background rates, produced by low-transverse momentum muons and other particles. Furthermore, the non-uniform magnetic field in the endcaps leads to complicated particle trajectories. For this reason the track reconstruction is performed in both the $\eta$ and $\phi$ planes simultaneously.

The Cathode Strip Chamber Track Finder (CSCTF) system is segmented into six sectors in each endcap, each covering $60^\circ$ in $\phi$. Data from a sector are processed by a SP which receives up to three LCTs from each of the five MPCs for a given sector via Sector Receiver (SR) boards. The SPs do not support sharing of information between neighbouring sectors as it was not deemed necessary due to the small bending angle in the endcaps. This leads to slight inefficiencies at the sector boundaries.

The CSCTF reconstruction logic operates by performing pairwise comparisons of track segments in different detector stations, testing for compatibility in $\eta$ and $\phi$ of a muon originating from the collision point within a certain tolerance window that takes into account bending due to the magnetic field in that region. Furthermore, the track finding logic searches for track segments in different assigned bunch crossings within a sliding time window of programmable length (nominally two bunch crossings wide). This feature improves reconstruction quality in those cases when a track segment was assigned to an incorrect bunch crossing, but requires to remove duplicate tracks found in consecutive bunch crossings. The CSCTF assigns the resulting track to the bunch crossing of the second arriving track segment. Transverse momentum is assigned using a large static random-access memory (SRAM) look-up table (LUT) that takes the track type, $\eta$ coordinate, and the differences in the $\phi$ coordinates of up to three track segments, as well as the bending angle of the track segment found in the first measurement station for two-station tracks. To improve efficiency in the overlap region between DT and CSC systems the CSCTF system also receives DTLT track segments from the DTTF.

A sorter card receives the best three muons containing their transverse momenta, spatial coordinates, charge, and quality from each SP for each bunch crossing. This card finally forwards the best four candidates to the GMT (fig. 5.7). [69, 76]

**The resistive plate chamber trigger.** The resistive plate chamber (RPC) trigger covers both the barrel and endcap regions of the CMS experiment up to a pseudorapidity of $|\eta| < 1.6$. It provides a complementary muon measurement to the DT and CSC detectors with excellent time resolution. The RPC detectors do not produce trigger primitives as in the DT and CSC detectors, but hits are directly transmitted to link boards (LBs) located in the experimental cavern which perform bunch crossing assignment, zero suppression, and multiplexing. Hit information is then transmitted via optical fibres to 84 trigger boards located in the underground counting room.

In contrast to the DT and CSC triggers which rely on track finding algorithms, the RPC trigger uses a pattern matching algorithm, implemented in the Pattern Comparator Trigger (PACT) system. The algorithm compares the particle tracks with strip patterns that are arranged in segments of approximately $0.1$ in $|\eta|$ and $2^\circ$ in $\phi$. A muon track can cross several segments and thus be reconstructed multiple times. In order to remove these duplicate tracks the RPC ghost buster (GB) operates interleaved with the reconstruction logic, finding duplicate muon...
tracks and selecting the one with the highest quality code. In case of draws the muon with the highest transverse momentum is selected.

After sorting the muons in four stages the RPC trigger finally sends the best 4+4 muons from the barrel and endcap regions to the GMT. [77, 78]

The Global Muon Trigger

The Global Muon Trigger (GMT) forms the final stage of the Level-1 muon trigger. It receives a total of 16 muon candidates from the regional triggers: four each from the DT, and RPC systems in the barrel as well as four each from the CSC, and RPC systems in the endcaps.

The GMT takes advantage of the overlapping coverage of the regional muon trigger systems to improve the trigger’s transverse momentum resolution by finding and merging muon candidates that were detected in two complementary sub-systems (i.e. DT+RPC or CSC+RPC). This improvement in resolution can be achieved due to a sophisticated merging system that exploits redundancies between DT or CSC and RPC muons. Simultaneously, the muon rate due to fake or improperly measured candidates can be controlled as unconfirmed tracks can be suppressed for certain trigger algorithms based on a combination of their pseudorapidity $\eta$ and the quality assigned to them by the track finder system.

Furthermore, the GMT is equipped with dedicated cancel-out units that remove duplicate candidates produced by both the DT and CSC sub-systems due to shared stations.

An additional function of the GMT is the correlation of muon tracks with calorimeter regions in order to both check a muon candidate for isolation as well as perform a confirmation. The muon track is extrapolated back to the interaction region or calorimeter using its coordinates $\eta$ and $\phi$ as well as its charge and transverse momentum. The GMT is designed to receive so-called MIP and Quiet bits from the calorimeter trigger which it correlates to muon tracks using their position at the calorimeter or interaction region, respectively. If the Quiet bit in the region
5.2 The Level-1 trigger for LHC Run-1

under consideration is set, an isolation tag is added to the muon indicating that it was not part of a jet. If the MIP bit is set in the given region this indicated that the calorimeter trigger detected a signature consistent with the passing of a muon within that region. Both isolation and MIP confirmation can be used in the GT as additional criteria for trigger algorithms. The transmission of MIP and Quiet bits was however never commissioned, due to insufficient resolution in the calorimeter system.

Finally, the GMT ranks and sorts the muon candidates according to the quality assigned by the regional trigger as well as the muon’s transverse momentum and position. The highest quality value is assigned to candidates measured in two complementary systems. The four highest ranking candidates are then propagated to the GT. [79, 80]

5.2.3 The Global Trigger

The Global Trigger (GT) forms the final stage of the CMS Level-1 trigger. It consists of several processing boards housed in a single 9U VME crate together with the GMT and Trigger Control System (TCS).

At each bunch crossing, the GT computes independent accept or reject decisions for 128 programmable algorithms that use inputs from both the calorimeter and muon triggers as well as so-called external conditions that consist of a single bit to signal an interesting condition, for instance the presence of colliding bunches at a given moment. Furthermore, each algorithm can be programmed to only fire every \textit{n}th detected event, a feature called prescaling. If either of these algorithms returns an accept decision, the GT sends a final OR (FinOR) signal to the TCS initiating the read-out process. Additionally, the GT receives 64 technical trigger signals that can be used to either veto or trigger a FinOR signal independently from the algorithm triggers. These technical triggers are mostly used for monitoring and calibration of several subdetectors, including the Level-1 trigger itself.

The GT receives four muon candidates containing information on transverse momentum, spatial coordinates, charge, reconstruction quality, as well as bits indicating isolation and compatibility with a muon in the calorimeter from the GMT. The GCT transmits four non-isolated and four isolated e/γ objects, four central and four forward jets, as well as four tau jets, each containing information on transverse energy ($E_T$), location, and reconstruction quality as well as the total transverse energy sum ($E_T^{\text{tot}}$), the missing transverse energy ($E_T^{\text{miss}}$), and the corresponding sums over all clustered jets found ($H_T$ and $H_T^{\text{miss}}$). These data are synchronised to each other and to the LHC orbit at the inputs before being sent to the Global Trigger Logic (GTL) card via the backplane. The GTL additionally receives the external conditions signals that can be used in the 128 trigger algorithms. Each algorithm can apply combinations of certain conditions on the properties of the input objects, such as $p_T$ and $E_T$ being above a certain threshold, or the pseudorapidity or azimuthal angle being within a certain range. In contrast to earlier Level-1 trigger systems which operated purely on such cuts and counts of objects, the GTL can also compute both the difference in pseudorapidity or azimuthal angle being within a certain range, as well as so-called correlation conditions which depend on the difference in pseudorapidity and azimuth between two objects of different kinds.

The outputs of all 128 algorithms and 64 technical triggers are combined to a FinOR signal which is propagated to the TCS system. This system can then issue a Level-1 Accept (L1A)
signal to the complete detector if certain rules are met and the detector itself is ready to be read out. Once the L1A signal has been issued the various read-out electronics transmit the event data of the triggering bunch crossing to the High-Level trigger (HLT) via the data acquisition (DAQ) system.

The TCS furthermore generates various control signals to start and stop runs, resynchronise and reset counters of front-end electronics, or trigger calibration of the CMS sub-detectors and distributes them via the Timing, Trigger, and Control (TTC) system. [81, 82]

5.2.4 The Timing, Trigger, and Control system

In order to be able to read out the CMS detector, it is of paramount importance to be able to distribute trigger, control, and clock signals to all detector components in a synchronised fashion. The provided clock must be in phase with the LHC clock, each tick corresponding to a possible bunch crossing. This synchronisation both between all sub-detectors as well as to the LHC bunch crossings is performed by the Timing, Trigger, and Control (TTC) system which was designed by experts from all LHC experiments and from the accelerators themselves.

At every LHC experiment the TTC machine interface (TTCmi) system receives the LHC clock together with an orbit signal that marks the beginning of the LHC orbit and has a frequency of 11.246 MHz from the LHC control room in Prevessin via underground fibres. It distributes these signals to 32 TTC VME interface (TTCci) boards that each form the root of a TTC distribution tree called detector partition. In CMS each detector partition can be operated independently for testing and commissioning purposes by injecting control signals at the TTCci-level. [82, 83]

5.3 The upgraded Level-1 trigger

The LHC’s expected instantaneous peak luminosity during Run-2 exceeds the original design specification of $1 \times 10^{34}$ cm$^{-2}$ s$^{-1}$. The number of interactions per bunch crossing (pile-up) already surpassed the expected number of 25 interactions per bunch crossing in the 2012 run.

The Level-1 trigger, however, must continue to support a physics program that both allows searches at the TeV scale and maintains sensitivity to electroweak scale physics. This goal needed to be accomplished with an increased centre-of-mass energy up to 14 TeV, a pile-up of up to 50, as well as an instantaneous luminosity exceeding $2 \times 10^{34}$ cm$^{-2}$ s$^{-1}$. This cannot be achieved by just increasing the Level-1 accept rate as several detectors would require major upgrades in order to accommodate the required read-out rate. Without upgrading the Level-1 trigger the anticipated increase in trigger rates due to the larger instantaneous luminosity and pile-up would require a large increase in trigger thresholds which would adversely affect the physics acceptance at the electroweak scale.

In order to avoid raising the trigger thresholds the CMS Level-1 trigger upgrade is comprised of the following new features:

- Introduction of pile-up subtraction for both jet finding and electromagnetic object isolation algorithms
- Improved hadronic $\tau$ identification
5.3 The upgraded Level-1 trigger

- Improved position and transverse momentum resolution for muon tracks in particular in challenging detector regions
- Muon isolation using pile-up subtracted energy sums provided by the calorimeter trigger
- Increased number of possible algorithms in the Level-1 trigger menu
- Introduction of invariant-mass calculation at the upgraded Global Trigger (μGT)

To support the implementation of the above mentioned improvements to the trigger, its architecture was completely overhauled. The calorimeter trigger was redesigned to consist of two processing layers, named Layer-1 and Layer-2 while the muon trigger was redesigned to perform track finding using all detectors in a given geometrical region (fig. 5.8). [84]

![Calorimeter trigger diagram](image)

Figure 5.8: The upgraded Level-1 trigger of the CMS experiment. The calorimeter trigger was reorganised into a time-multiplexing trigger, allowing the full event information to be processed in the Layer-2 processors, after preprocessing and multiplexing in Layer-1. In the muon trigger the paradigm changed from a subdetector-based track finding system to a system reconstructing all tracks in a given geometric region. The upgraded calorimeter and muon triggers are described in more detail in sections 5.3.2 and 5.3.3.

The Level-1 trigger was upgraded during Long Shutdown 1 (LS1) and commissioned during the first year of LHC Run-2. It entered operation in 2016.
5.3.1 Level-1 trigger hardware

To accommodate these new features the electronics for the calorimeter and muon trigger as well as the GT were upgraded.

While the legacy system was largely composed of custom VME processing boards for each task within the trigger, the upgrade relies heavily on common multi-purpose processing boards developed within the Level-1 trigger community. These boards follow the Micro Telecommunications Computing Architecture (MicroTCA) standard[85] which was chosen as the base technology for the CMS experiment’s Phase-1 upgrade. A direct benefit of following this common standard is the ability to take advantage of developments from other CMS subdetector groups such as the AMC13[86] read-out board, developed as a generic interface to DAQ by the HCAL Phase-1 upgrade project. A MicroTCA system furthermore supplies system-level health management, redundant power supplies and cooling as well as a high-capacity back plane which support the trigger systems in providing a high availability, high throughput system.

In the upgraded Level-1 trigger three different multi-purpose processing boards are used:

The Imperial Master Processor 7 (MP7). The MP7 is a high performance stream processing board developed by Imperial College. It supplies 72+72 10 Gbit/s channels for receiving and sending data via optical links. Multi-fibre Termination Push-on (MTP) connectors are used for I/O, in which one connector bundles 36 fibres. The MP7 board is used by the Layer-2 calorimeter trigger, the upgraded Global Muon Trigger (μGMT), as well as the μGT. Infrastructure firmware providing a common link protocol, the read-out block, etc. is shared among the users of the MP7 and developed in a public source code repository. [87]

The Modular Track Finder Processor 7 (MTF7). The MTF7 is a double-board design with the optical module providing 84 10 Gbit/s optical input receivers and 28 10 Gbit/s optical transmitters, while a mezzanine card contains a large 1 GB LUT in external random-access memory (RAM). It was developed by the University of Florida and is used in the Overlap Muon Track Finder (OMTF) and Endcap Muon Track Finder (EMTF) systems. [88]

The Calorimeter Trigger Processor 7 (CTP7). Designed by the University of Wisconsin for use in the first layer of the upgraded calorimeter trigger, the CTP7 provides 48 output channels at 10 Gbit/s and 67 input channels at 4.8 Gbit/s to 6.4 Gbit/s. It is furthermore equipped with 14 bi-directional backplane links with a bandwidth of 10 Gbit/s for intra-board communication within a MicroTCA crate. [89]

All boards are equipped with a Xilinx Virtex-7 VX690T [90] FPGA as processing chip. Furthermore, for all systems the backplane is used to communicate with the AMC13 card, that propagates fast control signals from the Trigger Control and Distribution System (TCDS) (described in section 5.3.5) as well as collects read-out information from all cards in a crate upon receipt of a Level-1 Accept signal before forwarding this information to the DAQ.
5.3 The upgraded Level-1 trigger

5.3.2 The calorimeter trigger

For the calorimeter trigger upgrade the concept of a time-multiplexed trigger (TMT) was developed: The Layer-1 calorimeter trigger, composed of 18 preprocessor cards, receives energy values measured in trigger towers in the ECAL and HCAL and forwards these in a time-multiplexed fashion to an array of 9 processing boards that constitute the Layer-2 calorimeter trigger. Each processing board in the Layer-2 calorimeter trigger receives the full information for every 9th event in a round-robin fashion (fig. 5.9). By concentrating all event information in one processor, sophisticated algorithms such a pile-up subtraction can be applied to at an early stage in object reconstruction. [91]

Figure 5.9: Concept of a time-multiplexed trigger. Detector data for one bunch crossing are spread over multiple preprocessor cards before they are combined in the second layer. The second layer consists of multiple processors, each operating on the full event information for a given bunch crossing. [92]
Chapter 5  The Level-1 Trigger

The Layer-1 calorimeter trigger

The Layer-1 calorimeter trigger consists of 18 CTP7 processing boards. It receives the trigger primitive data for each tower of the calorimeters (ECAL, HCAL, and HF) on 576 optical fibres at 4.8 Gbit/s and 576 optical fibres at 6.4 Gbit/s synchronous to the LHC clock.

The Layer-1 calorimeter trigger is organised in 36 calorimeter slices. Each slice consists of a calorimeter section spanning 1/2 of the $\eta$ range and 20° (four trigger towers) in $\phi$. Each CTP7 card covers two calorimeter slices and receives 64 input links in total: 32 from ECAL, 28 from HCAL and four from HF. The regional view of a Layer-1 processing card is extended using intra-board connections via the MicroTCA backplane that allow data sharing between different regions. The Layer-1 trigger aligns and decodes input data, then applies tower-level calibration using lookup tables and calculates sums of ECAL and HCAL transverse energies, finally streaming them to the Layer-2 subsystem for further processing. All calculations are performed at trigger tower precision with the output preserving this granularity. Each of the 18 CTP7 cards provides four 10 Gbit/s links to each of the nine Layer-2 processors which leads to a total of 648 optical links being required to provide the output data to the Layer-2 system. [89, 93]

The Layer-2 calorimeter trigger

The Layer-2 calorimeter trigger consists of an array of nine MP7 processing boards as well as a dedicated MP7 card used for demultiplexing of the output from the processing boards. Each processing node receives energy sums for every calorimeter trigger tower over four optical links operating at 10 Gbit/s from each of the 18 Layer-1 preprocessors (fig. 5.10). The system uses these sums to find $\tau$ lepton, jet, and electron/photon candidates as well as to compute various energy sums (table 5.2). In contrast to the legacy system electron/photon candidates include a tag indicating whether they are isolated or not, but are not split into separate collections anymore. In turn the calorimeter trigger propagates 12 such candidates to the $\mu$GT. Similarly, the distinction between forward and central jets has been dropped in favour of a larger collection of 12 jets that are sent to the $\mu$GT.

<table>
<thead>
<tr>
<th>Physics object</th>
<th>$\eta$ region</th>
<th>number of objects</th>
</tr>
</thead>
<tbody>
<tr>
<td>electron or photon candidates</td>
<td>$</td>
<td>\eta</td>
</tr>
<tr>
<td>jets</td>
<td>$</td>
<td>\eta</td>
</tr>
<tr>
<td>$\tau$ lepton candidates</td>
<td>$</td>
<td>\eta</td>
</tr>
<tr>
<td>total transverse energy</td>
<td>N/A</td>
<td>1</td>
</tr>
<tr>
<td>transverse hadronic energy</td>
<td>N/A</td>
<td>1</td>
</tr>
<tr>
<td>missing energy</td>
<td>N/A</td>
<td>1</td>
</tr>
<tr>
<td>missing hadronic energy</td>
<td>N/A</td>
<td>1</td>
</tr>
</tbody>
</table>

Table 5.2: Physics objects found by the upgraded calorimeter trigger.

As previously mentioned, the Layer-2 trigger processes events in a time-multiplexed fashion similar to the operating principle of the HLT where a given processing node is supplied the full
5.3 The upgraded Level-1 trigger

Figure 5.10: System design of the upgraded calorimeter trigger. An array of 18 CTP7 processing cards form the Layer-1 which acts as preprocessing and multiplexing layer. The entire data for a given event are then sent to one of nine MP7 cards that use the input data to reconstruct physics objects, sort them and forward them to a demultiplexer card that receives the nine data streams and merges them to allow processing by the GT.
Chapter 5  The Level-1 Trigger

event information for a fraction of bunch crossings in a round-robin fashion (fig. 5.10).

The Layer-2 calorimeter trigger is able to apply advanced pile-up mitigation by profiting from the complete event information being available in a single processor: By determining the average energy density of a given event a real-time estimate of the pile-up level can be computed and used to correct energy values of Level-1 object candidates. The availability of full trigger tower granularity greatly improves both the energy and position resolution of \( \tau \) and electron/photon candidates as well as the energy resolution for global energy sums constructed by the upgraded calorimeter trigger. [94]

In a final stage the Layer-2 system sorts each collection of trigger candidates separately and propagates the best objects of each collection to the demultiplexer card. This card receives event information from each of the 9 Layer-2 processing units, concentrates and serialises the data streams and forwards a single stream to the \( \muGT \). [93]

5.3.3 The muon trigger

As described in section 5.2.2, the primary purpose of the Level-1 muon trigger is the identification and measurement of muons with high precision. To improve upon the legacy system the upgrade of the Level-1 muon trigger introduces muon isolation based on pile-up subtracted energy sums as well as an improvement in the muon parameter precision, in particular the transverse momentum measurement.

To improve the measurement of muon track parameters especially in the challenging detector region of the overlap between the DT and CSC detector system, the track finder architecture was re-structured from detector-based to region-based: While there were separate track finding systems for the DT, CSC, and RPC detectors in the legacy trigger, the upgraded system introduces a barrel, overlap, and endcap muon track finder (fig. 5.11). Each upgraded track finder system receives all available muon detector information for the region it covers (fig. 5.12).

In order to accommodate the track finders’ increased resolution, muon track information that is sent to the \( \muGMT \) was doubled in size from 32 bit to 64 bit data structures. This enables the Level-1 trigger to move to a linear scale for the muon track parameters while maintaining flexibility for possible later changes. [96]

Triggering in the barrel region

In the barrel region the Level-1 trigger consists of the TwinMUX concentrator system which merges DT track segments and RPC hits within one muon station to so-called super-primitives and forwards them to the Barrel Muon Track Finder (BMTF). The BMTF constructs muon tracks from these super-primitives. Each BMTF processor finally forwards its best three track candidates to the \( \muGMT \). [96]

The TwinMUX system. The TwinMUX system is implemented in 60 custom processing cards equipped with a Virtex-7 FPGA and replaces the SC system in the DT trigger path. From the DT system it receives hits in the \( \eta \) direction as well as track segments in the r-\( \phi \) plane. It also receives RPC hits which it attempts to merge with the DT track segments in order to construct super-primitives.
5.3 The upgraded Level-1 trigger

Figure 5.11: Side view of a quarter of the CMS detector. For track finding in the Level-1 trigger the detector was segmented into three regions: the barrel (|\eta| < 0.83), the overlap region (0.83 ≤ |\eta| ≤ 1.23), and the endcaps (1.23 < |\eta| ≤ 2.4). Figure adapted from [95].

The TwinMUX receives the DT track segments at 480 Mbit/s from the DT mini-crates (see section 5.2.2) via CuOF boards that convert the galvanic inputs from the mini-crates to optical signals. The RPC hit information is sent by the RPC LB at 1.6 Gbit/s. The TwinMUX system synchronises the incoming data streams to each other, constructs up to four super-primitives per muon station and fans these out to the BMTF processors for a given wedge as well as its neighbour wedges. It also forwards the track segment data along the \eta direction for each muon station.

The duplication of track segment and hit information is necessary in order to improve track reconstruction efficiency in the BMTF: A muon track may cross the boundary between wedges due to magnetic bending. To avoid poor reconstruction due to missing track segments for such tracks, the BMTF processor for a given wedge uses track segments from the neighbouring wedge.

The TwinMUX system also provides unaltered DT track segments from the barrel muon stations in the overlap area to the OMTF system. It furthermore doubles as the read-out system for the DT detector. [97]

The Barrel Muon Track Finder. The Barrel Muon Track Finder (BMTF) is responsible for reconstructing muon tracks in the barrel area (|\eta| < 0.83). Its track finding algorithm builds on the method used in the DTTF, described in section 5.2.2. Following the DTTF design it is logically split into twelve 30° wedges, where each wedge consists of six sectors containing independent DT and RPC detectors. In contrast to the DTTF which receives track segments in the r-\phi plane from the SCs, the BMTF receives super-primitives built by the TwinMUX.
Figure 5.12: Block diagram of the upgraded CMS Level-1 muon trigger. Tracks are reconstructed by the endcap, overlap, and barrel muon track finders (EMTF, OMTF, and BMTF) using partly pre-processed hit or track segment information from the muon detectors. DT track segments are sent with a Copper to optical fibre mezzanine (CuOF) to the TwinMUX system which also receives the RPC hits from the barrel region. It propagates DT track segment information and RPC hits directly to the OMTF system while constructing track segments from DT and RPC information for use by the BMTF. The concentrator and preprocessor fan-out (CPPF) receives and concentrates hits from the RPC in the endcap region via the LBs and provides this information to the EMTF system. Hits from the CSC detectors are sent to the OMTF and EMTF without pre-processing. The OMTF also receives the RPC hits directly from the LB. Fully reconstructed tracks are propagated to the μGMT where information from the Calorimeter Layer-2 system is used to assign isolation values. Finally, the best eight muon tracks are sent to the μGT.
The upgraded Level-1 trigger system from DT and RPC hits. They describe track segments within a detector module and similarly to DTTF track segments consist of the ϕ coordinate, the bending angle, and quality bits indicating the confidence in the measurement. The BMTF also receives hit information along the η direction for each muon station in a given wedge. The RPC hits are used to correct mismeasurement of track segments that were assigned to a bunch-crossing too early or too late by the DT system.

The BMTF uses the super-primitives analogously to the track segments in the r-ϕ plane in the DTTF, extrapolating between the ϕ-segments to construct tracks with a coarse η coordinate while in parallel processing the η information in a pattern matching algorithm. In a later stage the algorithm attempts to match and join the two track types. Due to processing constraints the BMTF currently uses only the two best of the four super-primitives received from the TwinMUX system. The two remaining track segments are discarded.

While the DTTF assigned the transverse momentum using the difference in the ϕ coordinates in two muon stations and utilising the bending angle ϕb only for selection of a high-pT or low-pT LUT, the BMTF algorithm extends this method by using the ϕb value to assign the pT value in some cases.

The rate of the DTTF algorithm suffers from incorrectly promoting tracks with low-pT to high-pT muons. In order to mitigate this effect, the bending angle of track segments used in the reconstruction of a muon track can be used. The bending angle is highly correlated with transverse momentum above about 20 GeV. The BMTF uses this fact to compute a pT value from ϕb and use it instead of the value for high-pT muons computed by the legacy Δϕ algorithm in those cases when its value is lower than the one computed by that algorithm. In this way incorrectly promoted muons are avoided by using the algorithm providing the lower pT (fig. 5.13). [98]

The BMTF is implemented in 12 MP7 processing cards of which each operates on track segments and hit information from one wedge. Additionally, the information from neighbouring wedges is multiplied by the TwinMUX system and shared with the given track finder processor in order to avoid inefficiencies at the edges of a wedge. These data are received over 30 10 Gbit/s optical links from the 15 TwinMUX processors responsible for a given wedge and its neighbours.

The best three muons from each wedge are then sent to the μGMT over one 10 Gbit/s link. [96, 99]

### Triggering in the overlap region

The Overlap Muon Track Finder (OMTF) system reconstructs tracks in the overlap regions between the barrel and endcaps (0.83 ≤ |η| ≤ 1.23) on both sides of the CMS detector. It consists of 12 MTF7 processing boards, each of which processes DT track segments, CSC track stubs, and RPC hit information from a 60° detector sector. After a future upgrade each processor will additionally receive this track information from a 10° overlap with a neighbouring sector to maintain reconstruction efficiency at the boundaries.

A given OMTF processor operates on track information from the DT, CSC, and RPC system. Data from the DT system are provided by the TwinMUX system via 10 Gbit/s links while track stubs from the CSC detectors and hit information from the RPC system are received directly
Chapter 5 The Level-1 Trigger

Figure 5.13: The BMTF $p_T$ assignment algorithm uses the bending angle $\phi_b$ in addition to the difference in azimuthal coordinates between two track segments $\Delta \phi$. If the bending angle indicates a low transverse momentum it always uses the algorithm based on $\Delta \phi$. If $\phi_b$ suggests a high-$p_T$ muon, the lower of either the output of the $\Delta \phi$ algorithm or an algorithm based on the bending angle itself is selected. This method protects against incorrectly identified high-$p_T$ muons.

from the MPCs at 3.2 Gbit/s for the CSC information and the LBs at 1.8 Gbit/s in the case of the RPC hits.

The OMTF system synchronises the incoming data streams to each other and the LHC clock, and applies a so-called Golden Pattern (GP) algorithm in order to find muon tracks. A GP is a pre-computed hit pattern corresponding to a certain transverse momentum range and charge sign. The GP algorithm compares the hit patterns for a given event with the stored set of GPs and selects the track parameters corresponding to the GP with the smallest difference to the actual hit pattern. The $\phi$ coordinate is defined to be measured in muon station 2 of either the barrel or endcap region. Finally, a ghost busting algorithm is applied to eliminate duplicate muon tracks before transmitting the best three tracks per 60° sector to the $\mu$GMT. A detailed description of the OMTF GP algorithm is given in [95].

Triggering in the endcap region

In the forward region ($1.23 < |\eta| \leq 2.4$) the Endcap Muon Track Finder (EMTF) system uses pattern recognition to reconstruct tracks. Similarly to the OMTF system it consists of 12 MTF7 processing boards that each cover a 60° sector as well as an overlap of 20° with the neighbouring sector. Each EMTF board receives CSC track stubs from the MPCs at 3.2 Gbit/s as well as RPC hit information from the endcaps at 10 Gbit/s from the CPPF system. At the time of writing the EMTF however uses only the CSC information to build tracks. An extension of the algorithm to include RPC hits is being studied.

Compared to the CSCTF algorithm, the EMTF pattern recognition algorithm is able to use all 18 available LCTs per station for track finding instead of formerly three LCTs. The EMTF
5.3 The upgraded Level-1 trigger

Algorithm segments a given sector into four zones of constant $\theta$ ranges in the $\phi$-$z$ plane. A segmentation in constant ranges of $\theta$ was chosen because bending and multiple scattering windows in $\theta$ are approximately constant for the entire endcap. A zone is effectively a two-dimensional lattice with 120 discretised values of $\phi$ for each of the four stations, which over the 60° of a sector gives a resolution of roughly half a degree. Each LCT in an event is converted from its raw position variables into ones that are consistent with the discretised $\phi$ values of the zone so that they may be compared with each other.

The pattern recognition process works by comparing all track segments in a zone with predefined patterns: If a sufficient number of segments match, the algorithm considers a potential track to have been found.

In certain cases the same hits can be matched to different patterns, creating duplicate tracks. To mitigate this problem the EMTF applies two ghost busting algorithms already during the pattern recognition process. The first is to choose only the highest quality pattern at a given $\phi$ position in station 2. Additionally, patterns matched at neighbouring $\phi$ positions are compared and the lower quality one subsequently deleted.

An internal quality is calculated by interleaving two other quality-type words, layer and straightness, which are respectively defined by the number of stations present with matching segments in a given pattern and by how close the hits are to a straight line.

Once a quality code is assigned to each pattern detected, the three highest quality patterns per zone are found in a sorting step. This means that a maximal total of 12 track candidates per sector is sent to the next module. A track candidate is comprised of a reference pattern and the LCTs that could be matched to it, as well as the associated quality word.

In a separate step $\Delta \phi$ and $\Delta \eta$ are computed between each matched LCT and the corresponding pattern. These values are used to retrieve the $p_T$ value from a LUT that is filled with values that were computed using Boosted Decision Trees (BDTs). The LCT position values in station 2 are used as the global $\eta$ and $\phi$ values for the track. If this LCT is missing for a given track the position in station 3, followed by station 4 is used.

After the track finding step there are potentially 12 track candidates in each sector. For each sector a final ghost busting step is performed. If two tracks share an LCT the track candidate with the lower quality is discarded. After this step the remaining track candidates are sorted before the three highest quality tracks of each sector are transmitted to the $\mu$GMT. [100]

The upgraded Global Muon Trigger

The upgraded Global Muon Trigger ($\mu$GMT) receives 36 muon track candidates via 12 optical fibres at 10 Gbit/s from each of the three track finding systems as well as 1008 energy sums of calorimeter towers at the $2 \times 2$ trigger tower granularity from the Layer-2 calorimeter trigger via 2810 Gbit/s optical links. It sorts the input muons in two stages while in parallel removing duplicate muon candidates found in the overlap regions between neighbouring track finder processors and extrapolating the tracks to the interaction region. The extrapolated coordinates are used to compute isolation variables in a final step by using the calorimeter tower sums for the corresponding region. The best eight muons are then sent to the $\mu$GT.

The $\mu$GMT is described in greater detail in chapter 6.
Chapter 5  The Level-1 Trigger

5.3.4 The upgraded Global Trigger

The upgraded Global Trigger (μGT) represents the last algorithmic stage of the Level-1 trigger. It receives inputs from both the demultiplexer card of the Layer-2 calorimeter trigger, and the μGMT as well as from so-called external condition boards that receive and convert simple galvanic inputs from various subsystems to optical signals that can be received by the μGT processing boards. As described in sections 5.3.2 and 5.3.3, the calorimeter and muon trigger paths provide physics objects that can include amongst others spatial coordinates, a measure of the quality, as well as transverse energy or momentum values. In contrast, the external condition signals are simple binary pulses signifying either a vetoed or an accepted event. They are received as low-voltage differential signalling (LVDS) signals using 8 position 8 contact (8P8C) (commonly known as RJ45) connectors on a galvanic patch panel that concentrates them into very-high-density cable interconnect (VHDCI) connectors in order to allow the space-efficient reception of 64 external condition signals on a single MicroTCA module, designed to convert the galvanic signals for optical transmission to the μGT processors.

While the GT system could use 128 algorithms to form a FinOR decision, the μGT was upgraded with the possibility to use up to 512 algorithms. The technical trigger and external condition signals were merged to allow them to be used both as veto signals as well as in combination with other inputs to trigger a read-out signal. Finally, the possibility to compute the invariant mass for two input objects was added.

The μGT integrates the formerly separate modules of the GT in one MP7 processing board. Due to trigger algorithms growing both in number as well as in complexity the system was designed in such a way as to support operation of up to six μGT boards in parallel. This allows to split the set of algorithms among multiple boards, thereby reducing the logic use on a single chip. The output of each of these boards is sent to a custom processor card called the FinOR board via two galvanic LEMO cables from each μGT processor to signify either a local FinOR or veto signal. The FinOR board merges these signals in order to form a global FinOR signal that is sent via LEMO connectors to the TCDS. [101]

5.3.5 The Trigger Control and Distribution System

The CMS TCS and TTC systems were upgraded primarily due to the need for an increase in partitions from the original 32 to at least 40. This upgrade also provided the opportunity to revisit the interface and split in responsibilities between the GT and TTC systems. While the legacy GT was deeply integrated with the TCS system, these responsibilities were merged with the TTC system to form the Trigger Control and Distribution System (TCDS).

The TCDS distributes timing, control, and synchronisation data to the detector front-ends while receiving status information from each sub-detector read-out system. Each sub-detector can provide status information indicating readiness to receive triggers initiating a detector read-out via the Trigger Throttling System (TTS) mechanism. Only if all detectors signal readiness via this system can a trigger be issued by the TCDS. The clock signal that is distributed together with fast control information is synchronous to bunch crossing events in the LHC. This allows data taking in tune with the collisions provided by the accelerator and synchronised across the various detector systems that make up CMS. [102]
The TCDS was installed during LS1 and went into operation already with the beginning of LHC Run-2.

5.4 Online control software for the Level-1 trigger

The entire Level-1 trigger electronics are exposed to the CMS central run-control (RCMS) as a single sub-system. However, the different trigger systems are housed in different crates and built by different teams, which makes a distributed control and monitoring system very useful.

Such a system is provided by the Trigger Supervisor C++ software, which uses the CMS-wide XDAQ middleware. A Trigger Supervisor-based system is represented by a cell which offers both a web-interface for direct control by a human operator and a Simple Object Access Protocol (SOAP)-based application programming interface (API). Cells can be highly customised by the sub-system’s maintainer. For central control of the Level-1 trigger electronics by RCMS the so-called Central Cell is provided (fig. 5.14). [104]

![Diagram](image)

Figure 5.14: Conceptual overview of the CMS Trigger Supervisor (TS). The central run control infrastructure (RC) connects to a single point of entry, the central cell (labelled TS central node). The central cell then steers the various sub-system control cells which manage their respective online software infrastructure (OSWI). Communication with central run control and between cells is carried out via SOAP messages. Each cell additionally provides access via a web interface. [104]

Configuration data for the complete CMS detector are stored in the Online Master Database system (OMDS), an Oracle database. It can be easily accessed by Trigger Supervisor cells using XDAQ’s TStore service. The Level-1 trigger stores e.g. its different configurations in this database. [105]

During the upgrade of the Level-1 trigger (see section 5.3) the online software underwent
Chapter 5 The Level-1 Trigger

a significant overhaul. Taking advantage of the homogenisation in the trigger hardware and firmware SoftWare for Automating the conTrol of Common Hardware (SWATCH) provides centralised features such as the setup and monitoring of the AMC13 read-out board and provisions for easy implementation of custom monitoring. It is organised into modules that mirror the hardware in order to allow sub-system developers to easily adapt a SWATCH cell to their needs. [106]

5.5 Offline software for Level-1 trigger emulation

For the analysis of the Level-1 trigger performance as well as studies of anticipated features, a software emulation of the Level-1 trigger is used. This software is organised in modules, similarly to the processing boards and has to provide bit-level emulation of the firmware implementation. The emulators are implemented within the CMS Software (CMSSW) framework [107, 108] and are run in parallel to data taking in order to continuously monitor the hardware for errors, as well as offline for in-depth analysis of the recorded data as well as to predict the impact of changes to the Level-1 trigger firmware before it is deployed.
Chapter 6

THE UPGRADED GLOBAL MUON TRIGGER

In this chapter the development and design of the upgraded Global Muon Trigger (μGMT) is presented. The requirements for such a system are given before discussing the hardware implementation as well as the characteristics of the muon and calorimeter inputs to the μGMT. The bulk of the chapter is represented by a detailed outline of the algorithms used together with their implementation in the chip. Finally, monitoring and control infrastructure is described before giving a short overview of the technological challenges overcome during the development of this system.

6.1 Introduction

The μGMT forms the last stage of the muon processing chain in the Level-1 trigger of the Compact Muon Solenoid (CMS) experiment. It receives and ranks almost fully reconstructed muon objects from the three track finder systems and selects the eight highest-ranked tracks before forwarding them to the upgraded Global Trigger (μGT). In order to determine these eight final muons the μGMT sorts the input muon objects while removing duplicates in parallel. Additionally, it computes an isolation variable that is attached to each of the final muons and allows the μGT to suppress muons originating from hadronic jets. As a byproduct of the isolation functionality the μGMT extrapolates tracks back to the nominal interaction region. This information is propagated to the μGT as well, improving the invariant mass resolution for calculations involving at least one muon as well as correlation conditions that depend on the $\Delta\phi$ value of two particles.

6.2 Requirements

The μGMT receives up to three muon objects each from 36 independently operating track finder processors, as well as energy sums from the calorimeter trigger. It is tasked with selecting eight final muons based on a rank derived from their transverse momentum and reconstruction quality, as well as computing isolation variables using the received calorimetric information. Several requirements emerge from these considerations.

Conversion of azimuthal coordinates During muon track reconstruction each track finder processor assigns a $\phi$ value relative to its internal coordinate system to the reconstructed muon object. The μGMT has to convert these coordinates to the global reference coordinate system.
Chapter 6 The upgraded Global Muon Trigger

Duplicate removal As described in section 5.3.3, both the neighbouring track finder systems as well as the individual processors within a track finder system overlap in coverage in order to eliminate inefficiencies at the boundaries. This feature however leads to the construction of duplicates in these overlap regions. The μGMT is tasked with the removal of such so-called ghosts.

Establishment of the most useful muon tracks in the given event The μGMT receives up to 108 muon objects from the track finding layer, and forwards eight muons to the μGT. It therefore ranks the input muons based on both reconstruction quality and transverse momentum before finding the eight highest ranked muons in a fast sorting module.

Calculation of isolation variables In order to reduce the number of muons originating from hadronic jets, the μGMT is able to compute isolation variables for each muon selected for transmission to the μGT by using energy deposits that are received from the calorimeter system. A future version of the μGMT firmware may alternatively use data received from the calorimeter trigger to confirm a muon measurement.

Extrapolation of the spatial coordinates to the nominal interaction region As mentioned in section 5.3.3 the track finder systems measure a muon’s spatial coordinate within the muon system. In particular, the azimuthal coordinate of muons with low transverse momentum is significantly affected by magnetic bending. As the calorimeter is significantly closer to the interaction point than the muon stations and hadronic jets are less affected by magnetic bending than muons, the μGMT has to extrapolate the muon tracks back to the interaction region in order to improve the assignment of the isolation variables. The extrapolated azimuthal coordinates are additionally forwarded to the μGT in order to allow correct computation of the invariant mass as well as cuts on ∆ϕ between two trigger objects.

Provision of debugging and monitoring data for each triggered event In order to guarantee the correct operation of the trigger chain, the μGMT is required to transmit data to the central data acquisition (DAQ) system. This collection of data contains the set of input and output muons for the bunch crossing that triggered a read-out event, as well as a window of at least one bunch crossing before and after the read-out signal arrived. The μGMT additionally transmits so-called intermediate muons\(^1\) that allow more in-depth diagnostics by giving insight into the internal operation of the μGMT. It furthermore is required to transmit the energy sums received from the calorimeter trigger to the DAQ. This information can be examined almost in real-time and is used by operators to verify the correct functioning of the system. It is also used as input to the emulator software for offline analysis and studies.

Real-time monitoring of link stability and muon rates In addition to the debugging information sent to the DAQ, the μGMT possesses the ability of rudimentary self-monitoring, incrementing error counters as well as recording input, intermediate, and output rates that can be read out by the online software.

\(^1\)The set of muon objects remaining after the first sorting stage.
6.3 Hardware of the upgraded Global Muon Trigger

6.2.1 Operational requirements

Further constraints are imposed on the system due to external limitations.

**Latency**  The legacy Global Muon Trigger (GMT) operated in the time frame of 225 ns which was also the latency budget foreseen for the upgraded system. The migration to serial optical links for communication means that about 150 ns are used on serialisation and deserialisation of link data, leaving only 75 ns to perform the μGMT’s calculations. This tight requirement necessitates the use of highly optimised algorithms that can be completed on a short time scale. Additionally, in order to reduce the number of deserialisation-serialisation steps in the upgraded system it was decided to integrate the formerly separate regional sorter stages into the μGMT, further tightening the demands on algorithm efficiency.

**Pipelined architecture**  Each Large Hadron Collider (LHC) collision has to be inspected within a time span of less than 4 µs due to limitations of the buffer memory on the detector. The Level-1 trigger is expected to arrive at a read-out decision for each event without any deadtime. This goal is achieved by processing multiple events simultaneously within a pipeline operating at the LHC bunch crossing frequency of 40 MHz. The fundamental unit of operation within the Level-1 trigger is therefore a 40 MHz clock period, called bunch crossing interval.

6.3 Hardware of the upgraded Global Muon Trigger

The μGMT is implemented in an Imperial Master Processor 7 (MP7) board, described in section 5.3.1. It is housed in a Micro Telecommunications Computing Architecture (MicroTCA) crate in the underground counting room, in close vicinity of the μGT system. The MP7 card is equipped with a large Xilinx Virtex-7 VX690T field-programmable gate array (FPGA) and provides (72 + 72) 10 Gbit/s optical input and output links. Each track finding system transmits its muon candidates via 12 fibres, requiring half the μGMT input capacity. The calorimeter trigger requires 28 links to provide the energy sum information necessary to compute isolation variables. An optical patch panel installation concentrates the individual fibres from all track finder processors as well as the fibres from the calorimeter trigger into two dedicated trunk cables, each of which is connected to one of the two input Multi-fibre Termination Push-on (MTP) connectors on the μGMT. The transceivers for each of the MTP connectors are placed on opposite sides of the FPGA, allowing a clean separation of muon and calorimeter information at the inputs and allowing optimised placement of the algorithm blocks as described in section 6.7.

6.4 Inputs to the upgraded Global Muon Trigger

As mentioned earlier the μGMT receives input data at 10 Gbit/s. Data are asynchronously transmitted in 32 bit words at 250 MHz using 10b/8b coding. Padding and alignment words are

---

2The fraction of the acquisition time when no events can be recorded.
3Data words dedicated to aligning multiple data lanes to each other.
Chapter 6  The upgraded Global Muon Trigger

inserted at the transmitter stage, leading to an effective data rate of 240 MHz, which translates to 192 bit per link and bunch crossing.

6.4.1 Muon data

Muon information is encoded in 62 bit data structures, transmitted within two 32 bit words over the link. The most significant bits (MSBs) of each link word are used to store control bits with information global to the event under consideration, such as the three least significant bits (LSBs) of the bunch crossing number in the LHC orbit as well as a bit indicating the beginning of the orbit. These control bits are used for online monitoring purposes, as explained in section 6.6. In one bunch crossing data describing three muon tracks can be transmitted per link.

Table 6.1 illustrates the input muon data format for transmission via the optical links. Data fields are arranged in an order that optimises routing of signals in logic blocks pertaining to complex operations and thereby allows faster processing in order to optimise the system’s latency, see sections 6.5 and 6.7.

Input scales

The legacy system was required to encode pseudorapidity $\eta$ and transverse momentum $p_T$ using non-linear conversion functions between the integer values used by the firmware and their physical representation (“scales”) due to bandwidth reasons. In the upgraded muon system this limitation has been mitigated with the introduction of high bandwidth serial links, and thus the scales for pseudorapidity, azimuthal angle $\phi$, and transverse momentum are all linear as well as common to the three track finder systems. In contrast, the so-called H/F bit as well as the track addresses have track finder-specific meaning. Four quality bits received for each muon object are split into two parts where the two MSBs use a common scale among track finders, while the two remaining bits have track finder specific meaning but retain an ordering which allows their use in the cancel-out logic. The azimuthal coordinate $\phi$ is encoded as an 8 bit two’s complement signed integer by the track finder processors. Its value is local to the processor’s internal coordinate system where $\phi = 0^\circ$ lines up with the lower sector or wedge boundary of the individual processor. The $\mu$GMT thus needs to convert this value to the global coordinate system used by the CMS Level-1 trigger (see section 6.5.2). Transmission in this way allows the track finder logic to be designed without encoding explicit knowledge of the detector area that is covered by a given processor. Furthermore, the bandwidth required for transmission of the azimuthal angle to the $\mu$GMT is reduced from 10 bit to 8 bit.

The detailed input scale definitions for muon candidates are listed in table 6.2.

Quality definition  The quality scale is defined at the track finder level. The full 4 bit of the quality code are used by the $\mu$GMT to determine a sort rank, as well as which muon to cancel in case of the occurrence of duplicates. A modified form of the quality code is also forwarded to the $\mu$GT, where the two MSBs of the quality value are used to determine in which trigger algorithms a muon can be used in (e.g. trigger algorithms requiring one or two muons, see table 6.3). Additionally, the $\mu$GMT sends the quality code 0b1111 to the $\mu$GT if the halo bit was set for a muon by the EMTF, indicating a track parallel to the beamline.
Table 6.1: Input muon data sent during one bunch crossing. Three muons are transmitted in a total of six 32 bit words. The most significant bit at the far-left position is not used for muon information, but to transmit control data pertaining to the given bunch crossing.

<table>
<thead>
<tr>
<th></th>
<th>BX0</th>
<th>φ</th>
<th>H/F</th>
<th>η</th>
<th>qual</th>
<th>p_T</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0</td>
<td>31 30</td>
<td>22 21</td>
<td>12 8</td>
<td>0</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>BX0</th>
<th>φ</th>
<th>H/F</th>
<th>η</th>
<th>qual</th>
<th>p_T</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Res.</td>
<td>track addresses</td>
<td>VCH</td>
<td>CH</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>31 30</td>
<td>22 21</td>
<td>12 8</td>
<td>0</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>B0</th>
<th>φ</th>
<th>H/F</th>
<th>η</th>
<th>qual</th>
<th>p_T</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>B1</td>
<td>track addresses</td>
<td>VCH</td>
<td>CH</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>31 30</td>
<td>22 21</td>
<td>12 8</td>
<td>0</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>B2</th>
<th>φ</th>
<th>H/F</th>
<th>η</th>
<th>qual</th>
<th>p_T</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>Res.</td>
<td>track addresses</td>
<td>VCH</td>
<td>CH</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>31 30</td>
<td>22 21</td>
<td>12 8</td>
<td>0</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

62 data bits (for each muon):
- p_T: Transverse momentum; 9 bit word
- qual: 4 bit word indicating reconstruction quality
- η: Pseudorapidity; 9 bit word
- H/F: Halo bit for Endcap Muon Track Finder (EMTF); set if the muon travelled parallel to beamline. (Currently not implemented in EMTF.)
- η-fine bit indicating accurate measurement of η for Barrel Muon Track Finder (BMTF). Ignored for Overlap Muon Track Finder (OMTF)
- φ: Local azimuthal angle in track finder processor’s local coordinate system; encoded as a two’s complement 8 bit word
- CH: Charge bit
- VCH: Valid charge
- track addresses: Information on track segments used to construct given track, up to 29 bit wide

4 control bits:
- BX0: Bunch crossing zero bit (1 = bunch crossing zero of LHC orbit; 0 = else)
- B0/1/2: Three lowest bits of bunch crossing counter
- Res.: Reserved for future use
Table 6.2: Given in this table are the scale definitions of muon candidates transmitted to the \( \mu \)GMT: Name of the parameter, number of bits, the unit (value of the least significant bit), the integer range used and a comment.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Width (bit)</th>
<th>Unit ( u )</th>
<th>Range</th>
<th>Comment</th>
</tr>
</thead>
<tbody>
<tr>
<td>( p_T )</td>
<td>9</td>
<td>0.5 GeV</td>
<td>0 to 511</td>
<td>Value of 0 indicates no valid muon data; (bit value(-1)(x)( u ) ( p_T ) is defined at 90% of the plateau efficiency.</td>
</tr>
<tr>
<td>Quality</td>
<td>4</td>
<td></td>
<td>0 to 15</td>
<td>2’s complement. bit value ( x ) ( u ) gives the centre of the bin.</td>
</tr>
<tr>
<td>( \eta )</td>
<td>9</td>
<td>0.010875</td>
<td>(-230 ) to ( 230 )</td>
<td>2’s complement. bit value ( x ) ( u ) gives the lower edge of the bin.</td>
</tr>
<tr>
<td>H/F</td>
<td>1</td>
<td></td>
<td></td>
<td>Indicates halo-muon when sent by EMTF or ( \eta )-fine bit when sent by BMTF.</td>
</tr>
<tr>
<td>( \phi_l )</td>
<td>8</td>
<td>2( \pi )/576</td>
<td></td>
<td>2’s complement. bit value ( x ) ( u ) gives the lower edge of the bin.</td>
</tr>
<tr>
<td>Muon charge</td>
<td>1</td>
<td></td>
<td></td>
<td>1: negative, 0: positive</td>
</tr>
<tr>
<td>Charge valid</td>
<td>1</td>
<td></td>
<td></td>
<td>1: charge sign is valid, 0: charge sign cannot be determined</td>
</tr>
<tr>
<td>Track address</td>
<td>29</td>
<td></td>
<td></td>
<td>See section 6.5.3</td>
</tr>
</tbody>
</table>

Table 6.3: Muon quality scale as interpreted by the \( \mu \)GT.

<table>
<thead>
<tr>
<th>Quality code at ( \mu )GT</th>
<th>Quality code at ( \mu )GT</th>
<th>Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>0b0000</td>
<td>0b00xx</td>
<td>cosmic muon trigger (EMTF single hit tracks)</td>
</tr>
<tr>
<td>0b0100</td>
<td>0b01xx</td>
<td>accepted in Open triggers</td>
</tr>
<tr>
<td>0b1000</td>
<td>0b10xx</td>
<td>accepted in double muon and Open triggers</td>
</tr>
<tr>
<td>0b1100</td>
<td>0b11xx</td>
<td>accepted in single and double muon as well as Open triggers</td>
</tr>
<tr>
<td>0b1111</td>
<td>N/A</td>
<td>special code sent if halo muon bit is set</td>
</tr>
</tbody>
</table>
6.4.2 Calorimetry data

Each energy sum is encoded as a 5 bit word, allowing six energy sums to be packed into one 32 bit word provided by the link protocol. A total of $6 \times 6$ energy sums can thus be sent during one LHC clock period. This means that one ring in constant $\eta$ can be transmitted per link. The $\mu$GMT is built to accept either 28 or 32 links from the calorimeter trigger depending on implementation choices to be made in the future (see section 6.5.5).

The two MSBs of each incoming 32 bit link word are used as space to store control bits pertaining to the transmitted bunch crossing, as explained in section 6.4.1, while the remaining 30 bit store the six energy sum word, see table 6.4.

Table 6.4: Data format for the six 32 bit data words transmitted during one LHC clock from the calorimeter trigger.

<table>
<thead>
<tr>
<th>Res.</th>
<th>BX0</th>
<th>$E_5$</th>
<th>$E_4$</th>
<th>$E_3$</th>
<th>$E_2$</th>
<th>$E_1$</th>
<th>$E_0$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>Res.</td>
<td>$E_{11}$</td>
<td>$E_{10}$</td>
<td>$E_9$</td>
<td>$E_8$</td>
<td>$E_7$</td>
<td>$E_6$</td>
</tr>
<tr>
<td>2</td>
<td>Res.</td>
<td>$E_{17}$</td>
<td>$E_{13}$</td>
<td>$E_{15}$</td>
<td>$E_{14}$</td>
<td>$E_{13}$</td>
<td>$E_{12}$</td>
</tr>
<tr>
<td>3</td>
<td>Res.</td>
<td>$E_{23}$</td>
<td>$E_{20}$</td>
<td>$E_{21}$</td>
<td>$E_{20}$</td>
<td>$E_{19}$</td>
<td>$E_{18}$</td>
</tr>
<tr>
<td>4</td>
<td>Res.</td>
<td>$E_{29}$</td>
<td>$E_{26}$</td>
<td>$E_{27}$</td>
<td>$E_{26}$</td>
<td>$E_{25}$</td>
<td>$E_{24}$</td>
</tr>
<tr>
<td>5</td>
<td>Res.</td>
<td>$E_{35}$</td>
<td>$E_{32}$</td>
<td>$E_{33}$</td>
<td>$E_{32}$</td>
<td>$E_{31}$</td>
<td>$E_{30}$</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>31</th>
<th>30</th>
<th>29</th>
<th>24</th>
<th>19</th>
<th>14</th>
<th>9</th>
<th>4</th>
<th>0</th>
</tr>
</thead>
</table>

180 data bits:
- $E_i$ Energy values for towers with fixed $\eta$ ($36 \times 5$ bit).
- $E_0$ begins at $\phi = 0^\circ$; $\Delta \phi = 10^\circ$

4 control bits:
- BX0 Bunch crossing zero bit
  - (1 = bunch crossing zero of LHC orbit; 0 = else)
- B0/1/2 Three lowest bits of bunch crossing counter

6.5 Design and algorithms of the upgraded Global Muon Trigger

The $\mu$GMT algorithms are written in the Very High Speed Integrated Circuit Hardware Description Language (VHDL) [109]. They have been implemented in a highly modular fashion (fig. 6.1) in order to remain flexible towards changes due to revised requirements or updated studies indicating the need for different algorithms. This was especially important because only three years were available for the development and implementation of the system. As the track finder systems were designed in parallel to the $\mu$GMT, performance studies for the system converged only shortly before commissioning of the upgraded Level-1 trigger, making multiple designs for certain algorithms necessary.

As described in section 6.4, the MP7 card represents the link data as 32 bit words at 240 MHz internally while the $\mu$GMT fundamentally operates at the LHC clock of 40 MHz. In those cases where the muons have to be compared to each other it is necessary to deserialise the six 32 bit words received over one LHC clock period, corresponding to 25 ns. This deserialisation step however requires one LHC clock period in latency. For this reason a significant portion of the
Figure 6.1: The μGMT algorithm. Muons are received from the three track finding systems and sorted in two stages. In parallel, muon tracks are extrapolated to the interaction point and muon isolation variables are computed by using energy sum information received from the calorimeter trigger. The best eight muons are sent to the μGT.
6.5 Design and algorithms of the upgraded Global Muon Trigger

Table 6.5: Contents of data structure used for muon track data.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Width (bit)</th>
<th>Comment</th>
</tr>
</thead>
<tbody>
<tr>
<td>Muon charge</td>
<td>1</td>
<td>1: negative, 0: positive</td>
</tr>
<tr>
<td>Charge valid</td>
<td>1</td>
<td>1 indicates that the charge was measured accurately.</td>
</tr>
<tr>
<td>$\eta$</td>
<td>9</td>
<td>$\eta &gt; 0$ for $z &gt; 0$.</td>
</tr>
<tr>
<td>$\eta$-fine</td>
<td>1</td>
<td>1 indicates that the $\eta$ coordinate could be measured accurately.</td>
</tr>
<tr>
<td>Quality</td>
<td>4</td>
<td>Reconstruction quality.</td>
</tr>
<tr>
<td>$p_T$</td>
<td>9</td>
<td>Transverse momentum.</td>
</tr>
<tr>
<td>$\phi$</td>
<td>10</td>
<td>$\phi$ coordinate in global coordinate system.</td>
</tr>
<tr>
<td>Halo muon</td>
<td>1</td>
<td>Muon travelling parallel to beamline. Can only be set for EMTF.</td>
</tr>
</tbody>
</table>

$\mu$GMT logic is operated at the native link speed of 240 MHz, with the sorting and cancel-out unit as well as parts of the isolation assignment operating at the LHC bunch crossing frequency.

6.5.1 Deserialisation of input data

As explained earlier, a deserialisation stage allows to examine the complete information for a given event at the same time and is thus required as a preprocessing step for both the sorting and cancel-out units in the case of muon data, as well as for isolation assignment in the case of both muon and energy information. In the $\mu$GMT the data are inserted into a pipeline with six stages, operating at 240 MHz in order to extract the six 32 bit words that constitute the event information for one bunch crossing at each 40 MHz clock tick.

Deserialisation of muon tracks

Muon data are received via 36 input links, equally split between the three track finder systems (BMTF, OMTF, and EMTF). The muon deserialisation unit for a single link receives three muons encoded in two 32 bit words each at every bunch crossing. It contains a pipeline designed to construct both a muon data structure as well as a dedicated track address storage object. The muon data structure stores the main muon track information of each of the three input muons received over the given link, while the track addresses are used for duplicate removal, described in section 6.5.3. The muon track information is essentially the data structure that will be sent to the $\mu$GT eventually and contains track parameters such as transverse momentum ($p_T$), spatial coordinates, and reconstruction quality (see table 6.5). Additionally, for each muon object a so-called empty bit is generated and set if a muon with $p_T = 0$ is received. The empty bit indicates that no valid data are filled in a given muon data structure and has to be taken into account at various places in the algorithm. This is required because the track finder systems can in principle transmit any data instead of a muon if the field for $p_T$ is left empty.

In parallel to the main pipeline the global $\phi$ coordinate is calculated from the local value provided by each track finder processor at 240 MHz as described in section 6.5.2.
Chapter 6  The upgraded Global Muon Trigger

In a third computation unit besides the main pipeline and the computation of the global azimuthal coordinate, a sort rank is computed by the muon deserialisation unit. This rank is used to compare muons during the sorting stage and is dependent on transverse momentum and reconstruction quality. The rank can be flexibly calculated by a programmable look-up table (LUT) for each track finder system independently. It is currently defined as

\[ R = p_T + 4 \cdot \text{quality code} \]  \hspace{1cm} (6.1)

in order to weight the 4 bit reconstruction quality value higher, while maintaining the stronger contribution of the 9 bit transverse momentum.

Sort rank assignment is performed at 240 MHz since the data format was designed so that both transverse momentum and reconstruction quality words arrive in the first of the two muon data words. Thus, the sort rank lookup operation can be effectively hidden with regard to the latency budget as it can be computed in parallel to the deserialisation step.

The MP7 core firmware provides an indication whether the link is currently transmitting alignment words or valid data, by setting a data valid bit. If this bit is cleared the input data words for the given link contain no valid data. The µGMT masks any link transmitting invalid data automatically and then ORs the data valid bits over one bunch crossing for all links before setting its output links to the resulting state. In this way it can be guaranteed that alignment data are not erroneously sent as physics data to the µGT while ensuring that any valid muon data are dealt with correctly. These alignment data are expected to be received during a gap in the LHC orbit that is dedicated for the transmission of such information, however it can influence monitoring counters and therefore must be suppressed. A schematic of the muon deserialiser unit is shown in fig. 6.2.

Additional error checking logic, not illustrated in fig. 6.2 uses the control bits mentioned in section 6.4.1. This logic is described in section 6.6.

Deserialisation of energy sums

Deserialisation of energy sums for a single link proceeds similarly to the deserialisation in the main pipeline of the muon deserialisation unit. The six incoming 32 bit link words are buffered at 240 MHz in a pipeline and extracted at each 40 MHz clock tick. A total of 36 energy sums are received by each link per LHC clock period. The sums received from all 28 or 32 links are subsequently merged and provided to the algorithm logic as a large array of arrays data structure containing \( 36 \times 28 \) energy sums. Each sub-array in the data structure corresponds to the information received over a single link, and thus contains the 5 bit energy values corresponding to calorimeter towers with constant \( \eta \). A schematic of the energy sum deserialiser unit for a single link is shown in fig. 6.3.

Similarly to the muon deserialiser, a dedicated monitoring logic uses the received control bits described in section 6.4.2. It is described in section 6.6.

6.5.2 Conversion of azimuthal coordinates at the inputs

As mentioned in section 6.4.1, the µGMT receives a \( \phi \) coordinate with its origin set relative to a track finder processor’s local coordinate system. It then converts these coordinates to the
6.5 Design and algorithms of the upgraded Global Muon Trigger

Figure 6.2: The muon deserialisation unit for a single input link. A 32 bit data word from a track finder is received at 240 MHz and separated into parallel streams. The entire link words are processed by a deserialisation pipeline that stores the current value at each 240 MHz clock tick. This allows the construction of a large $3 \times 62$ bit data structure that contains the three input muons. In parallel, the data word corresponding to the local $\phi$ value is extracted and used to calculate the global azimuthal coordinate by adding a configurable value depending on which track finder processor is connected to the given unit. Care must be taken close to $\phi = 0$ in order to correctly treat the wrap-around. Additionally, a rank is computed from $p_T$ and the quality value from the track finders using a LUT. Finally, the data valid bit, generated by the link logic is OR-ed and stored for usage in the serialisation logic before transmission to the $\mu$GT.

Figure 6.3: The energy sum deserialiser unit for a single input link. Input data consist of six 32 bit link words per bunch crossing that are buffered in a pipeline before the five energy sum values stored in each word are extracted at each 40 MHz clock tick and stored in an array representing a ring of calorimeter towers with fixed $\eta$ coordinate.
Chapter 6 The upgraded Global Muon Trigger

global reference frame by applying an offset that is configurable for each link independently. The global azimuthal coordinate is stored in an unsigned 10 bit data word with an LSB value of $2\pi/576$. It follows that valid bit values range from 0 to 575. The µGMT protects against an under- or overflow at $\phi = 0^\circ$ in those cases when a negative local $\phi$ value is added to a small offset or a large local $\phi$ value is added to a large offset, applying simple arithmetical operations if such a situation is detected.

Figure 6.4 shows the global CMS trigger coordinate system overlayed with the relative coordinate systems used by the track finders. Two conventions are used as the BMTF processors take one drift tube (DT) wedge of 30° into account, while the OMTF and EMTF use the 60° sectors of the cathode strip chamber (CSC) system as their reference frame. In both cases, the lower boundary of the wedge/sector corresponds to $\phi_l = 0^\circ$.

- The BMTF transmits a signed value encoded in 2’s complement. The range expected is $-8$ to $56$ (corresponding to a range of $-5^\circ$ to $35^\circ$), where the first and last 8 values (corresponding to $5^\circ$ each) are overlapping with the previous and next processor.

- The EMTF and OMTF also transmit a signed value encoded in 2’s complement. The expected range for the EMTF is $-32$ to $100$ (corresponding to $-20^\circ$ to $62.5^\circ$), while the OMTF transmits a range of $-16$ to $100$ (corresponding to $-10^\circ$ to $62.5^\circ$). The values in the negative range as well as the last 4 values (corresponding to a range of $20^\circ$ and $10^\circ$ as well as $2.5^\circ$ respectively) are overlapping with the previous and next processor.

6.5.3 Duplicate removal

As described in section 5.3.3, duplicate muon tracks can be produced in two cases:

1. Hit information is split at the muon station-level, therefore track finder systems necessarily receive data from muon stations that also provide track segments outside the assigned pseudorapidity coverage region of the given track finder, see fig. 6.5. It follows that neighbouring track finder systems independently reconstruct the same muon track twice at their border.

2. A further source of duplicate muon tracks arises because individual track finder processors receive track segments from neighbouring wedges or sectors. This sharing of track segments is necessary in order to mitigate track reconstruction inefficiencies at the wedge and sector boundaries.

The µGMT removes such duplicates using a sophisticated matching and cancelling system. It was designed to allow the use of different cancel-out algorithms for each type of overlap area. Following studies presented in section 7.3, the µGMT is currently configured to use a track address-based cancel-out algorithm between processors of the BMTF and a coordinate-based scheme for all other boundaries.

The µGMT cancel-out system exploits the fact that duplicate tracks can only appear between adjacent track finder processors, thereby reducing the number of required comparisons significantly compared to a scheme in which all tracks have to be compared with each other. On
Figure 6.4: The relative $\phi$-scale for the track finders has its origin at the edge of the wedge or sector under consideration. The range is different for BMTF as it processes 30° wedges (black), compared to OMTF and EMTF which process 60° sectors (red).
Chapter 6 The upgraded Global Muon Trigger

Figure 6.5: Cross-sectional view of a quarter of the CMS detector. The muon system is split in coverage by $\eta$. The BMTF, OMTF, and EMTF systems receive track segments from their respective pseudorapidity region (see section 5.3.3). A complication arises because certain muon stations (drawn hatched in the figure) are shared between neighbouring track finder systems. Track segments from these stations are sent to both track finding systems and lead to possible duplicates. Figure adapted from [95].
Table 6.6: Contents of the data structure used in the cancel-out units. The synthesis tool detects whether a given cancel-out unit uses spatial coordinates or track addresses from either BMTF or EMTF and removes the unemployed signals, avoiding unnecessary routing effort.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Width/bit</th>
<th>Comment</th>
</tr>
</thead>
<tbody>
<tr>
<td>η-fine</td>
<td>1</td>
<td>1 indicates that the η coordinate could be measured accurately.</td>
</tr>
<tr>
<td>η</td>
<td>9</td>
<td>η &gt; 0 for z &gt; 0.</td>
</tr>
<tr>
<td>ϕ</td>
<td>8</td>
<td>ϕ coordinate in coordinate system local to individual track finder.</td>
</tr>
<tr>
<td>BMTF track address</td>
<td>21</td>
<td>Track addresses from BMTF system.</td>
</tr>
<tr>
<td>EMTF track address</td>
<td>16</td>
<td>Track addresses from EMTF system.</td>
</tr>
<tr>
<td>Quality</td>
<td>4</td>
<td>Used to determine which muon shall be removed in case of duplicates.</td>
</tr>
<tr>
<td>Empty</td>
<td>1</td>
<td>Tag to indicate that muon slot was determined to be empty at inputs.</td>
</tr>
</tbody>
</table>

6.5 Design and algorithms of the upgraded Global Muon Trigger

The firmware level, three cancel-out units with similar design are used: a dedicated cancel-out unit for intra-track finder cancel-out, and one per endcap for the borders between BMTF and OMTF as well as between OMTF and EMTF. This is necessary as each OMTF sector needs to be compared to four BMTF wedges, while only three EMTF sectors border a given OMTF sector. When searching for duplicates within a track finder system only comparisons with one neighbouring wedge or sector are required due to the symmetrical relationship in the comparison.

All cancel-out units receive muon track information containing either the spatial coordinates or track addresses, as well as the reconstruction quality and a bit indicating whether the muon data were determined to be invalid at the inputs (see table 6.6). The cancel-out units dedicated to duplicated muons between two different systems contain so-called wedge comparator units which assign the four BMTF wedges or three EMTF sectors to be compared to a given OMTF sector. The cancel-out unit dedicated to intra-track finder duplicates does not require a wedge comparator, as each wedge or sector is only compared to its neighbour in ascending ϕ (the wedge or sector with the highest ϕ is compared to the one with ϕ = 0°). The comparison of two wedges or sectors is performed in the wedge checker unit that ensures that all three muons from the first wedge or sector are compared to the three muons from the second wedge or sector. Within the wedge checker unit nine ghost checker units are instantiated to perform the $3 \times 3$ comparisons. A ghost checker unit uses the reconstruction quality as reported by the track finder systems to determine which muon should be removed once a match has been found. It then returns one ghost bit for each muon to indicate if it was identified as a duplicate. The firmware can be configured at synthesis time to use either track address-based or coordinate-based cancel-out,

---

4In case of equal qualities the muon constructed in the wedge or sector with lower azimuthal angle is used. For inter-track finder comparisons the muon generated in the OMTF system is removed in case of ties.
Chapter 6 The upgraded Global Muon Trigger

leading to a different ghost checker unit to be generated. The ghost bits for a particular track as well as the muon’s empty bit generated at the inputs are merged in two steps in the wedge checker unit and the cancel-out units. The wedge checker unit additionally uses the empty bits (see section 6.5.1) to disable ghost busting for invalid muons. Finally, bit arrays are returned by each cancel-out unit containing tags indicating which of the input muons shall be disregarded in further processing. A graphical representation of the cancel-out unit for the boundary between OMTF and EMTF is shown in fig. 6.6.

Coordinate-based cancel-out

The coordinate-based cancel-out scheme uses the spatial coordinates of the muons under consideration to determine whether they can be matched. It computes a distance

$$\Delta R^2 = (f_1 \Delta \eta)^2 + (f_2 \Delta \phi)^2,$$

where $f_1$ and $f_2$ are programmable parameters that allow to compensate for the difference in scales and measurement resolution between the $\phi$ and $\eta$ coordinates. The value of $\Delta R$ is compared to a threshold in order to determine whether two muons are duplicates of each other.

The local $\phi$ coordinate reported by the track finder systems is 8 bit in width while the $\phi$ coordinate in the global reference frame uses 10 bit. As the cancel-out algorithm is only performed between wedges or sectors that are neighbouring each other, the coordinate-based cancel-out algorithm uses the local azimuthal coordinates in order to minimise the number of signals that need to be routed across the FPGA. The firmware adds a programmable offset to the second $\phi$ coordinate in order to align the two local coordinate systems before the values of $\Delta \phi$ and $\Delta \eta$ are computed and used as inputs to a LUT that returns a single bit indicating whether the two input muons are classified as duplicates. In addition to $\Delta \phi$ and $\Delta \eta$, the $\eta$-fine bit can be used when a BMTF wedge is compared with a OMTF sector in order to stretch the cancel-out window in pseudorapidity when $\eta$ could not be measured precisely by the BMTF. Simultaneously to the ghost identification, the reconstruction quality of both muons is compared, allowing the ghost bit to be only issued for the lower quality muon. The behaviour in case of ties is detailed in the previous section.

In order to minimise the size of the LUT both $\Delta \phi$ and $\Delta \eta$ are reduced in width to 5 bit for $\Delta \eta$ and 3 bit for $\Delta \phi$. The values were chosen as muon separated by more than 0.3 in pseudorapidity or 0.075 in azimuthal angle will never be considered ghosts (section 7.3.1). A parallel logic checks if $\Delta \eta$ or $\Delta \phi$ are larger than 0.3 or 0.075 respectively and overrides the output of the LUT in those cases. A scheme of the cancel-out logic is shown in fig. 6.7.

The latency of the cancel-out module is less than 25 ns as it must fit in a 40 MHz clock cycle together with the count wins step of the sorter unit (section 6.5.4).

Track address-based cancel-out

The track address-based cancel-out algorithm is based on the fact that a limited number of track segments is available in each muon station to construct a muon object in the track finder systems. If the same track segment is used for two tracks, they can be tagged as probable duplicates with great confidence.
6.5 Design and algorithms of the upgraded Global Muon Trigger

Figure 6.6: Schematic of the cancel-out unit for the boundary between OMTF and EMTF. Track information for each muon is received by the unit and spread to six wedge comparator units, each responsible for one OMTF sector. Within a wedge comparator unit, three wedge checker units are tasked with comparing each muon within the given OMTF sector with all muons from one of the three adjacent EMTF sectors. The ghost checker unit receives two muon tracks and returns two bits indicating whether they are duplicates. These ghost bits are merged in both the wedge checker units and the cancel-out unit itself, in the ghost bits mergers in order to return two bit vectors containing the tags indicating cancelled muons. Dotted lines indicate grouping of individual signals for simplification. Not drawn is additional logic using the empty bits to disregard muons with invalid data as well as logic merging the empty bit for a given muon with its ghost bits. Both these circuits are implemented in the wedge checker unit.
Figure 6.7: The μGMT coordinate-based cancel-out logic. Differences in azimuthal coordinates and pseudorapidity ($\Delta \phi$ and $\Delta \eta$) are computed and used as input to a LUT together with the optional $\eta$-fine bit in those cases when a BMTF muon is under consideration. The LUT returns a single bit indicating the detection of a ghost. In order to reduce the memory consumption of the LUT only the lower bits of $\Delta \phi$ and $\Delta \eta$ are used as inputs while in parallel a simple calculation is done to determine if the two muons are far apart. Finally, the reconstruction quality of the two muons is compared and only the lower muon is assigned the ghost bit.
Both the BMTF and EMTF systems are capable of transmitting track addresses to the μGMT, both using a local addressing scheme which means that the track segments are identified relative to the track finder processor’s location.

**Track addresses from the BMTF**  As explained in sections 5.2.2 and 5.3.3, the BMTF extrapolates between track segments found in radially neighbouring stations in order to construct complete muon tracks. In addition to track segments from the wedge under consideration it uses segments from adjacent wedges to avoid inefficiencies at the boundaries due to muons crossing from one wedge into the other as a result of magnetic bending in the azimuthal direction.

The BMTF receives four track segments per muon station from the TwinMUX, however currently only two of these segments are considered by the logic. Its extrapolation algorithm utilises at most one track segment from each of the four radially separated muon stations to build a track, allowing the description of a track by listing its constituent track segments. Additionally, the BMTF employs at most one track segment from neighbouring wedges while constructing a muon track and does not extrapolate from that segment to others, i.e. the track finding algorithm extrapolates muon tracks only a maximum of one step into a neighbouring wedge. The μGMT takes advantage of this fact in the track address-based cancel-out algorithm and tags any two tracks sharing a single track segment as duplicates.

The track address-based cancel-out algorithm is a duplicate removal scheme that was extended from the legacy Drift Tube Track Finder (DTTF) system. To facilitate the implementation at the track finder side it was decided to group the four track segments into two sets of two: track segments (0, 1) and track segments (2, 3).

As the BMTF track address completely describes a given track, it uses significant transmission bandwidth in uncompressed form. To allow this information to be received in compact form, the associated wheel and wedge containing the muon station for each track segment are localised relative to the reference point of the track finder processor. In contrast, the reference wheel needs to be encoded in the track address word. The constituents are then addressed relative to this reference. To uniquely identify the track constituents the following information is thus required:

- Wheel number \( w \)

- Track segment encoding \( t \):
  - \( 4 \times 1 \) bit to select track segment set (0, 1) or (2, 3) for each station
  - \( 2 \) bit for innermost station (the used track segment was found in the reference wedge)
  - \( 3 \times 4 \) bit for stations 2, 3, and 4

- Reference wedge (not explicitly encoded, given by link number)

The wheel number is encoded in two parts. Bit 3 indicates the “side” of the detector (i.e. positive or negative) while bits 1 and 2 encode the number. For bit 3 the code ‘1’ means negative and ‘0’ means positive (see table 6.7).
Table 6.7: Wheel encoding: The first bit serves as a dedicated sign bit.

<table>
<thead>
<tr>
<th>Wheel Label</th>
<th>Binary encoding</th>
<th>Hexadecimal encoding</th>
</tr>
</thead>
<tbody>
<tr>
<td>-2</td>
<td>0b110</td>
<td>0x6</td>
</tr>
<tr>
<td>-1</td>
<td>0b101</td>
<td>0x5</td>
</tr>
<tr>
<td>-0</td>
<td>0b100</td>
<td>0x4</td>
</tr>
<tr>
<td>+0</td>
<td>0b000</td>
<td>0x0</td>
</tr>
<tr>
<td>+1</td>
<td>0b001</td>
<td>0x1</td>
</tr>
<tr>
<td>+2</td>
<td>0b010</td>
<td>0x2</td>
</tr>
<tr>
<td>None</td>
<td>0b111</td>
<td>0x7</td>
</tr>
</tbody>
</table>

The track segments are encoded in three distinct fields. The first field contains one bit for each station indicating whether the track segment from that station is a member of the track segment set (0, 1) or (2, 3). The remaining fields indicate whether the first or second track segment within such a set was used for each station separately. The track segment encoding is described in more detail in the following.

For station 1 the encoding 0x2 and 0x1 for first and second track segment of a given set and 0x3 for “no segment used” was agreed on. (For the first station the reference point — i.e. the central wedge and wheel with lower $|\eta|$ — is used.)

Track segments from stations 2 to 4, are identified by the following addressing scheme: For each station a 4 bit number is used that identifies wheel and wedge relative to the reference point as shown in fig. 6.8. In the direction of the beam, higher numbers identify the wheel closer to the central wheel (e.g. 0x2 vs. 0xA). For the azimuthal direction, higher numbers correspond to larger $\phi$ values (e.g. 0x1 vs. 0x5). Within a station the two possible numbers encode whether the first or second track segment of a given set was used for a particular track (e.g. 0x0 vs. 0x1). The code 0xF is used to signify “no segment used” for a particular station.

The encoding within the input muon words (see section 6.4.1) is illustrated in table 6.8 with the detailed coding of the selector field described in table 6.9.

Table 6.8: Encoding of track addresses for BMTF within the 32 bit of the second muon word described in table 6.1.

<table>
<thead>
<tr>
<th>s</th>
<th>w</th>
<th>t</th>
</tr>
</thead>
<tbody>
<tr>
<td>26</td>
<td>22</td>
<td>19</td>
</tr>
</tbody>
</table>

The track-address based cancel-out logic in the $\mu$GMT is implemented as a series of hard-coded

Table 6.9: Coding of the selector field shown in table 6.8.

<table>
<thead>
<tr>
<th>Station 1 selector</th>
<th>Station 2 selector</th>
<th>Station 3 selector</th>
<th>Station 4 selector</th>
</tr>
</thead>
<tbody>
<tr>
<td>26</td>
<td>25</td>
<td>24</td>
<td>23</td>
</tr>
</tbody>
</table>
Figure 6.8: Track segment encoding for wedge $N$ in one muon station. Each square corresponds to a muon station in a specific wheel and wedge. The four pairs of numbers identify the four track segments in the given muon station. The first number determines whether track segment set (0, 1) or (2, 3) was used, while the second number selects which of the two track segments within the given set was selected. This means that e.g. $(0x0, 0x8)$ is the first track segment in the reference station, while $(0x1, 0x5)$ is the fourth track segment from the muon station located in the wheel further from the detector’s centre and wedge $N + 1$. In this way the $\mu$GMT can determine which of the four track segments from a muon station was used to build a given track.
checks for all possible combinations of wheel and track segment encodings that indicate the presence of duplicates. The general cancel-out logic ensures that only muons from neighbouring wedges are compared.

In order to demonstrate the track address-based cancel-out algorithm an example (illustrated in fig. 6.9) is given in the following. In this example the \( \mu \text{GMT} \) is comparing two tracks received from neighbouring BMTF wedges.

1. **Encoding of track found by BMTF processor for wedge \( N \):** \( w = 0x1, s = 0x2, \) and \( t = 0x29CF \)

   This encoding describes a track originating from wheel +1 \( (w = 0x1) \) that was constructed from the following segments:
   - First track segment \( (0x2) \) of the first set \( (0x2=0b0010) \) in station 1
   - Second segment in wheel 1 and wedge \( N \) \( (0x9) \) of the first set \( (0x2=0b0010) \) in station 2
   - First segment in wheel 1 and wedge \( N + 1 \) \( (0xC) \) of the second set \( (0x2=0b0010) \) in station 3
   - No track segment from station 4 was used \( (0xF) \)

2. **Track as found by track finder for neighbour wedge \( N + 1 \):** \( w = 0x1, s = 0x3, \) and \( t = 0x3F88 \)

   The second track was found in wheel +1 \( (w = 0x1) \) and built from the following segments:
   - No track segment from stations 1 \( (0x3) \) and 2 \( (0xF) \)
   - First segment in wheel 1 and wedge \( N + 1 \) \( (0x8) \) of the second set \( (0x3=0b0011) \) in station 3
   - First segment in wheel 1 and wedge \( N + 1 \) \( (0x8) \) of the first set \( (0x3=0b0011) \) in station 4

The track segment used from station 3 is common for both tracks (it is the first segment of the second set in wheel 1 and wedge \( N + 1 \)). The tracks are therefore flagged as duplicates.

**Track addresses from EMTF** Each EMTF processor receives up to two track segments per CSC. A track segment is purely constructed from CSC hit information. Currently resistive plate chamber (RPC) hits are not included in the track address-based cancel-out scheme.

Similar to track addresses for BMTF tracks, the information sent by the EMTF indicates for each muon station which segment from which station was used to build a given track. If a segment is used by two tracks, these should be tagged as duplicates.

The EMTF sends two fields for each station, indicating the segment ID (either 0 or 1) and the chamber ID (identifying the CSC chamber that constructed the track segment), illustrated in table 6.10. If no segment from a station is used, then the associated chamber ID is set to zero.

The chamber IDs identify those chambers in a muon station that are part of an overlap region. For station 1 (the station closest to the interaction region) three chambers overlap with
6.5 Design and algorithms of the upgraded Global Muon Trigger

Figure 6.9: Illustration for example of track address-based cancel-out algorithm. The two tracks are duplicates because of the shared track segment in station 3.
Table 6.10: Encoding of track addresses for EMTF within the 32 bit of the second muon word described in table 6.1.

<table>
<thead>
<tr>
<th>station 4</th>
<th>station 3</th>
<th>station 2</th>
<th>station 1</th>
</tr>
</thead>
<tbody>
<tr>
<td>chamber</td>
<td>segment</td>
<td>chamber</td>
<td>segment</td>
</tr>
<tr>
<td>17</td>
<td>15</td>
<td>14</td>
<td>13</td>
</tr>
<tr>
<td>13</td>
<td>11</td>
<td>10</td>
<td>9</td>
</tr>
<tr>
<td>9</td>
<td>7</td>
<td>6</td>
<td>5</td>
</tr>
<tr>
<td>5</td>
<td>3</td>
<td>2</td>
<td></td>
</tr>
</tbody>
</table>

Figure 6.10: Illustration of the chamber ID assignment for CSCs in an EMTF processor for the station closest to the interaction region, station 1 (a), as well as station 2/3/4 (b). These chamber IDs are used to address track segments in the μGMT cancel-out system. Shaded chambers are associated with the neighbouring sector, but used for track finding to retain high efficiency at the overlaps.
neighbouring stations on each side while for station 2/3/4 only two chambers overlap on each side (see fig. 6.10).

Analogously to the BMTF cancel-out algorithm, two EMTF tracks are tagged as duplicates if they share a track segment.

The cancel-out algorithm finds shared track segments using the following criteria:

- The segment ID is identical for two track segments from the same station.
- In station 1 the relationship chamberID(sector N)+3 == chamberID(sector N+1) holds.
- In stations 2, 3, and 4 the following correspondence applies:
  \[ \text{chamberID(sector N)} + 2 == \text{chamberID(sector N+1)} \]

At the time of writing the track address-based cancel-out algorithm has been implemented in the \( \mu\)GMT firmware and emulator, but has not been studied due to the missing implementation of track addresses in the EMTF emulator.

### 6.5.4 Muon sorting

Muon sorting in the \( \mu\)GMT proceeds in two stages. The final sorters for the track finding systems, implemented in dedicated processing boards in the legacy system (see sections 5.2.2 to 5.2.2) are contained in the \( \mu\)GMT as regional sorting stage in order to recover the latency otherwise required for the serialisation and deserialisation process introduced by a separate sorter stage (see section 6.2.1). The outputs of the regional sorting stages are fed into a final sorter that identifies the eight highest ranked muons for transmission to the \( \mu\)GT.

The regional stage consists of five independent sorter units that each cover a contiguous range in pseudorapidity: A large sorter unit selects the best eight muons from 36 input tracks from the BMTF while four smaller sorting units receive 18 muons each from the positive and negative regions in pseudorapidity of the OMTF and EMTF systems respectively and find the best four amongst them. The regional sorting stage works in parallel to the cancel-out units and uses so-called disable bits in order to exclude duplicate or invalid muons from being selected. The disable bits are expected to arrive one 40 MHz clock tick late, during which time the sorting unit can perform preprocessing in order to optimise the system’s latency. A final sorting stage receives the 24 muons from the regional stages and selects the best eight of these.

All sorter units utilise the same basic architecture, developed previously for the legacy GMT [79]: Each input muon’s sort rank \( (R_{\mu_i}) \) is compared to the sort ranks of all other muons and a so-called greater equal matrix is created (see eq. (6.3)). The greater equal matrix contains a bit for each muon-muon combination indicating whether the first muon was ranked higher or equal to the second one.

\[
GEM = \begin{pmatrix}
0 & R_{\mu_1} \geq R_{\mu_2} & R_{\mu_1} \geq R_{\mu_3} & \ldots \\
R_{\mu_2} \geq R_{\mu_1} & 0 & R_{\mu_2} \geq R_{\mu_3} & \ldots \\
R_{\mu_3} \geq R_{\mu_1} & R_{\mu_3} \geq R_{\mu_2} & 0 & \ldots \\
\ldots & \ldots & \ldots & 0
\end{pmatrix}
\]  

(6.3)
The lower half of the greater equal matrix is not explicitly computed, but $GEM_{ij} = \sim GEM_{ji}$ is posited as $R_{\mu_i} \geq R_{\mu_j} = \text{true}$ automatically means that $R_{\mu_j} > R_{\mu_i}$ is false. Ties are in this way broken in favour of the muon with lower index because track finder systems in general send better muons earlier within a given bunch crossing, which in turn leads to a lower index being assigned in the $\mu$GMT.

A subsequent module counts the number of times a given muon was ranked higher than its partner, taking into account the received disable bits in order to disregard cancelled muons. From this information so-called select bits can be calculated by positing that the highest ranked muon would be ranked higher $N - 1$ times (as it was compared to all muons except itself) and similarly for the remaining muon slots. These select bits are then used by multiplexers to pick the highest ranked muons which are then sent to the next stage. Apart from muon objects, the sorting units also provide the associated empty and index bits that can be used to associate a muon object at the output of the sorter to a given input. The regional sorters furthermore output the associated muon sort ranks. A schematic of a regional sorter unit is shown in fig. 6.11.

Figure 6.11: Example of a regional sorter unit. Muon sort ranks from the deserialisation unit are used to construct a greater equal matrix. This matrix stores a bit array for each muon that represents if it was ranked higher or equal to the other muons processed by the unit. The matrix is used by a module that first counts the number of “wins” for each muon, taking into account the disable bits received from the cancel-out unit. It then calculates select bits that uniquely identify one of the sorter’s input muons for each output object. Apart from muon objects the regional sorter unit propagates the muon’s sort rank and index bits, as well as the empty and disable bit vectors.
6.5 Design and algorithms of the upgraded Global Muon Trigger

6.5.5 Muon isolation

As most analyses using a muon trigger for data taking are largely interested in muons from direct decays rather than originating from jets, the \( \mu \)GMT is equipped with functionality to examine energy deposits in the calorimeter around a muon’s track, potentially indicating the presence of a jet. The isolation unit determines whether a muon track originates from an area with large energy deposits and if so clears the isolation bit for the given muon. A jet is electrically neutral and therefore does not bend in the magnetic field while muons are affected by magnetic bending. As the muon’s coordinates are measured at station 2 by the track finding systems an extrapolation unit (see section 6.5.6) is required to determine the coordinates at the nominal interaction region.

The \( \mu \)GMT muon isolation algorithm receives sums of calorimeter trigger tower energies in a cone around every second trigger tower. Energy sums at full trigger tower granularity cannot be received due to bandwidth constraints. It then determines which trigger tower sum corresponds to each muon destined for transmission to the \( \mu \)GT and calculates both an absolute isolation value based on the pure energy value as well as a relative isolation value that takes into account both the calorimeter energy sum as well as the muon’s transverse momentum.

The acceptance range of the muon system is \( |\eta_{\text{acc}}| < 2.4 \). Within this region, the calorimeter operates with a segmentation of 52 \( \times \) 72 trigger towers in \( \eta \times \phi \). The towers have a constant size of 0.087 in \( \phi \) while the size in \( \eta \) varies. For \( |\eta| < 1.74 \) trigger towers have a constant size of 0.087 in \( \eta \) which increases for higher \( \eta \) values, as shown in table 6.11a. The towers are identified by indices that start at 1 (−1) for the first central tower in positive (negative) \( \eta \) and are incremented (decremented) towards higher \( |\eta| \).

The \( \mu \)GMT uses a reduced granularity of 28 \( \times \) 36 so-called tower sums, which each consist of two calorimeter trigger towers in \( \phi \) and have varying size in \( \eta \) as shown in table 6.11b. These tower sums can be calculated by the calorimeter trigger and transmitted to the \( \mu \)GMT. The \( \mu \)GMT calculates sums of 5 \( \times \) 5 tower sums that are used for the isolation assignment. To construct the isolation variable for muons that are found at the edge of the muon system, two additional rings of constant \( \eta \) are required on both sides of the detector. This means a total of 32 rings of constant \( \eta \) and thus 32 links are required.

An alternative to the above described algorithm arises because the calorimeter trigger calculates energy sums of cones around each tower with a diameter of eight trigger towers for its jet finding algorithm. These pre-summed cones are suitable for muon isolation. If these cones are used in the \( \mu \)GMT, the cone energy around every other trigger tower would be sent (except for high \( \eta \)), corresponding to the tower sum granularity also used in the scheme described above. In this case 28 \( \times \) 36 energy sums of these cones would have to be sent to cover the full muon acceptance. Note that the number of received objects is reduced with regard to the tower sums because no additional information is required at the edges of coverage in pseudorapidity as for the tower sums. A total of 28 links is required in this scheme.

Figure 6.12 shows a sketch of the different regions and cones, as they would be used by the \( \mu \)GMT.

The isolation unit (fig. 6.13) receives a set of so-called calorimeter index bits for each input muon, which indicate the calorimeter trigger tower corresponding to the muon’s originating point at the nominal interaction region. It additionally receives the index bits for the final eight
Chapter 6 The upgraded Global Muon Trigger

Table 6.11: Sizes and positions of trigger towers in positive $\eta$ (a), as well as sizes and positions of tower sums in positive $\eta$ together with the corresponding tower indices $i_\eta$ (b). The towers and tower sums are symmetric for negative $\eta$.

(a)  

<table>
<thead>
<tr>
<th>$i_\eta$</th>
<th>$\eta_{\text{low}}$</th>
<th>$\Delta \eta$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.0</td>
<td>0.087</td>
</tr>
<tr>
<td>2</td>
<td>0.087</td>
<td>0.087</td>
</tr>
<tr>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>19</td>
<td>1.566</td>
<td>0.087</td>
</tr>
<tr>
<td>20</td>
<td>1.653</td>
<td>0.087</td>
</tr>
<tr>
<td>21</td>
<td>1.74</td>
<td>0.09</td>
</tr>
<tr>
<td>22</td>
<td>1.83</td>
<td>0.1</td>
</tr>
<tr>
<td>23</td>
<td>1.93</td>
<td>0.113</td>
</tr>
<tr>
<td>24</td>
<td>2.043</td>
<td>0.129</td>
</tr>
<tr>
<td>25</td>
<td>2.172</td>
<td>0.15</td>
</tr>
<tr>
<td>26</td>
<td>2.322</td>
<td>0.178</td>
</tr>
<tr>
<td>27</td>
<td>2.5</td>
<td>0.15</td>
</tr>
<tr>
<td>28</td>
<td>2.65</td>
<td>0.35</td>
</tr>
</tbody>
</table>

(b)  

<table>
<thead>
<tr>
<th>$i^\text{sum}_\eta$</th>
<th>$\eta_{\text{low}}$</th>
<th>$\Delta \eta$</th>
<th>trigger towers ($i_\eta$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.0</td>
<td>0.174</td>
<td>1, 2</td>
</tr>
<tr>
<td>2</td>
<td>0.174</td>
<td>0.174</td>
<td>3, 4</td>
</tr>
<tr>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>10</td>
<td>1.566</td>
<td>0.174</td>
<td>19, 20</td>
</tr>
<tr>
<td>11</td>
<td>1.74</td>
<td>0.19</td>
<td>21, 22</td>
</tr>
<tr>
<td>12</td>
<td>1.93</td>
<td>0.242</td>
<td>23, 24</td>
</tr>
<tr>
<td>13</td>
<td>2.172</td>
<td>0.15</td>
<td>25</td>
</tr>
<tr>
<td>14</td>
<td>2.322</td>
<td>0.178</td>
<td>26</td>
</tr>
<tr>
<td>15</td>
<td>2.5</td>
<td>0.15</td>
<td>27</td>
</tr>
<tr>
<td>16</td>
<td>2.65</td>
<td>0.35</td>
<td>28</td>
</tr>
</tbody>
</table>

5x5 sums  

<table>
<thead>
<tr>
<th>trigger tower</th>
<th>tower sum</th>
<th>muon pass-through</th>
<th>calorimeter cone</th>
</tr>
</thead>
</table>

Figure 6.12: Illustration of a calorimeter cone calculated at trigger tower granularity for low $\eta$ regions. At higher $\eta$ a tower sum consists of only one trigger tower in $\eta$.  

96
6.5 Design and algorithms of the upgraded Global Muon Trigger

muons to be sent to the $\mu$GT which it uses to select the correct calorimeter index bits. The calorimeter index bits then select the eight tower sums corresponding to the final muon tracks out of the 1008 values received from the calorimeter trigger. Two LUTs finally compute absolute isolation from the chosen tower sums, and relative isolation from both the tower sums and the final muon’s transverse momentum values.

![Diagram](image)

Figure 6.13: Schematic of the isolation assignment unit. Index bits for the final eight muons as calculated by the global sorter stage are used to select a set of calorimeter tower sum index bits. These index bits point to the calorimeter regions which correspond to the position of each of the 108 input muons at the nominal interaction region. Finally, two LUTs are used to compute both absolute and relative isolation.

A detailed study of the $\mu$GMT isolation algorithm and a possible replacement by a muon confirmation algorithm utilising calorimeter information is given in [110].

6.5.6 Muon track extrapolation

As explained in section 6.5.5, the isolation algorithm uses calorimeter region indices corresponding to the spatial coordinates of the input muons at the interaction region. As the track finder systems provide a muon’s coordinates at muon station 2, the $\mu$GMT extrapolates all input muon tracks back to the nominal interaction region. The extrapolated azimuthal coordinates can furthermore be propagated to the $\mu$GT in addition to the muon coordinates at muon station 2, which improves the efficiency of trigger paths that include cuts on either $\Delta \phi$ or invariant mass calculations involving a muon (see section 7.4).

The extrapolation unit receives the 32 bit link words as well as the global $\phi$ value from the deserialisation unit with a delay of one and two 240 MHz clock ticks, respectively. From the input link words it extracts the values for $p_T$, $\eta$, and the muon’s charge. The lower six bits of the transverse momentum as well as the six MSBs of the pseudorapidity are used as inputs to two LUTs that return $\Delta \phi$ and $\Delta \eta$ values, respectively. The value of $\Delta \eta$ is encoded in a signed integer.
Chapter 6  The upgraded Global Muon Trigger

of 4 bit width which can be added or subtracted to the received pseudorapidity depending on the \( \eta \) value’s sign. The \( \Delta \phi \) value is a 3 bit unsigned integer that requires further processing. In order to extrapolate the azimuthal coordinate, the \( \mu \)GMT in parallel adds and subtracts the \( \Delta \phi \) value from the global \( \phi \) value. It then selects one of the two values based on the muon’s charge. The \( \mu \)GMT finally selects the extrapolated coordinates only if the transverse momentum of the muon is below a threshold value. Above this value the muon’s track is assumed to be bent less than the size of a calorimeter trigger tower making the extrapolation superfluous. This conditional extrapolation allows to use only the lower bits of the transverse momentum, reducing the size of the extrapolation LUTs. As the LUTs currently use the six LSBs of the transverse momentum value, the threshold value above which no extrapolation is performed is set to 32 GeV. After the extrapolation step the azimuthal coordinate is bounds-checked and if necessary adjusted to the correct range. Finally, both coordinates are converted to a calorimeter trigger tower index for consumption by the isolation unit. The extrapolated \( \phi \) values are separately propagated to the \( \mu \)GMT sorting logic where index bits generated by the final sorter are used to select those values corresponding to the final eight muons. These values are then sent to the \( \mu \)GT together with the associated muon objects.

While the information required for the muon track extrapolation is spread across two successive 32 bit words, it is possible to operate the extrapolation unit at 240 MHz in order to reduce resource usage and minimise routing effort (see section 6.7.2). Every odd result is constructed from the combination of the second 32 bit word of the \( N \)th muon with the first 32 bit word of the \( N + 1 \)th muon instead of the 32 bit words associated with the same muon, however these results are disregarded in a final deserialisation stage which provides the outputs of the extrapolation units at 40 MHz to the following units. A schematic of the extrapolation unit is shown in fig. 6.14.

The muon input data format was optimised in order to allow calculation of the extrapolated coordinates largely in parallel to the deserialisation unit, meaning that no latency is added to the \( \mu \)GMT processing pipeline by the extrapolation unit as the remaining deserialisation step is performed in the shadow of the sorting stages. In order to allow the immediate computation of the extrapolation residuals, the values for \( p_T \) and pseudorapidity are contained in the first of the two link words constituting a muon. Conversely, the charge bit is included in the second muon input word as it is required one 240 MHz clock tick later in order to select the correct \( \phi \)-extrapolation. The extrapolation unit begins processing with a delay of one 240 MHz clock tick due to routing constraints in the FPGA.

6.5.7 Serialisation of muon output data

Data destined for transmission to the \( \mu \)GT are serialised by a dedicated unit in the \( \mu \)GMT. This unit receives muon data at 40 MHz and serialises these to 32 bit words to be passed to the MP7 transceiver logic at 240 MHz. The \( \mu \)GMT transmits two muons per link to the \( \mu \)GT and fills the remaining two link words with zero. The serialiser is tasked with creating the appropriate data structure for transmission. It is furthermore responsible for the serialisation of the intermediate muon information that is solely used for monitoring purposes (see section 6.6.1).

The serialiser accepts the eight final muons from the final sorter as well as their associated extrapolated azimuthal coordinates, index bits, and isolation bits. It splits this information first
Figure 6.14: Schematic of the μGMT extrapolation unit for one input link. Values for $\Delta \phi$ and $\Delta \eta$ are computed by LUTs from the six higher bits of the muon’s pseudorapidity and the six LSBs of the transverse momentum, and used to calculate the extrapolated coordinates. The so-computed values are only used if the muon’s transverse momentum is below a threshold value, above which bending is considered to be negligible. Finally, the muon’s coordinates are mapped to calorimeter region indices by LUTs. Additionally, the extrapolated azimuthal coordinate is provided to the sorter unit for transmission to the μGT. The majority of the extrapolation pipeline is clocked at 240 MHz, with only the results being deserialised to the LHC bunch crossing frequency of 40 MHz.
Chapter 6 The upgraded Global Muon Trigger

per group of four links (quad) and then by individual link before merging all data for the two muons transmitted via a given link into a single output data structure, represented by an array of six 32 bit words (see table 6.12). As only two output muon objects are transmitted per link, two frames remain empty. These two empty frames are placed at the beginning of the bunch crossing in order to allow the receipt of the output muon objects two 240 MHz clock ticks later than otherwise possible.

Table 6.12: Data format for the six 32 bit data words transmitted during one LHC clock to the μGT.

<table>
<thead>
<tr>
<th>0</th>
<th>No data</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>No data</td>
</tr>
<tr>
<td>2</td>
<td>η</td>
</tr>
<tr>
<td></td>
<td>31 23 22 19 18 10 9 0</td>
</tr>
<tr>
<td>3</td>
<td>Res.</td>
</tr>
<tr>
<td></td>
<td>31 21 20 11 10 4 3 2 1 0</td>
</tr>
<tr>
<td>4</td>
<td>η</td>
</tr>
<tr>
<td></td>
<td>31 23 22 19 18 10 9 0</td>
</tr>
<tr>
<td>5</td>
<td>Res.</td>
</tr>
<tr>
<td></td>
<td>31 21 20 11 10 4 3 2 1 0</td>
</tr>
</tbody>
</table>

53 data bits (for each muon):

- \( \phi_{\text{extrap.}} \): Azimuthal angle at the interaction point; 10 bit word
- \( p_T \): Transverse momentum; 9 bit word where \( p_T = 0 \) indicates an empty muon (no valid data)
- qual: 4 bit word indicating reconstruction quality
- η: Pseudorapidity; 9 bit word
- aISO: Absolute isolation (1 = isolated)
- rISO: Relative isolation (1 = isolated)
- CH: Charge bit (1 = negative, 0 = positive)
- VCH: Valid charge (1 = valid, 0 = not valid)
- idx bits: Index bits to tag corresponding input muon, 7 bit word
- \( \phi \): Azimuthal angle in global coordinate system; 10 bit word

A multiplexer selects each word in the array successively at a frequency of 240 MHz. A schematic of the serialiser is shown in fig. 6.15.

The unit additionally receives valid bits for all activated muon and energy inputs separately, indicating whether the respective input links were sending valid data or alignment words as well as the outputs of the μGMT algorithms for transmission to the μGT. If either valid bit is set, the serialiser unit propagates this status to its output links, otherwise all output links are instructed to send alignment words.
6.5 Design and algorithms of the upgraded Global Muon Trigger

Figure 6.15: Schematic of the \( \mu \text{GMT} \) serialiser unit for one output link. Two muons are registered with the 40 MHz clock and split into two 32 bit segments. The four 32 bit words are then inserted into an array together with two words filled with zeroes. A multiplexer selects these words at a frequency of 240 MHz for transmission to the \( \mu \text{GT} \). Not drawn is the logic to merge the extrapolated azimuthal coordinate, index bits, and the isolation bits into the muon data structure as well as circuitry determining whether input link data were declared valid and setting the output link status accordingly.
6.6 Monitoring and control of the upgraded Global Muon Trigger

The μGMT provides multiple monitoring facilities using both data recorded by the DAQ after receipt of a Level-1 Accept (L1A) signal as well as separate monitoring information calculated on-chip for all received data and read out by the online control and monitoring software in aggregate.

6.6.1 Monitoring with triggered events

As mentioned in section 5.3.1 all Level-1 trigger systems utilise the AMC13 module as interface to the Frontend Readout Optical Link (FEROL) [111] board of the central DAQ system. The connection between the MP7 and the AMC13 is a 4 Gbit/s serial link on the backplane of the MicroTCA crate, leading to a hard limit for the size of the so-called read-out record — the set of data transmitted to the DAQ upon receipt of the L1A signal.

The MP7 firmware provides facilities to send the contents of the input and output link buffers to the AMC13 module upon receipt of a L1A signal. This functionality is highly configurable, allowing to set a so-called read-out window that determines the number of link words that are transmitted to the DAQ, as well as providing facilities that allow sending additional data for a subset of triggered bunch crossings.

The read-out record of the μGMT thus consists of the input data record, shown in tables 6.1 and 6.4, as well as the output data record (see table 6.12), and an intermediate muon data record that is only used for monitoring purposes and corresponds to the output data record, except for the omission of the isolation and index bits as well as the extrapolated $\phi$-coordinate.

In normal operation the μGMT sends a standard event record data structure at most read-out events. For every $10^7$th read-out event a larger event record is transmitted, allowing detailed debugging of synchronisation problems. The standard event record consists of input data associated to the triggering bunch crossing, as well as intermediate and output data from a window of five bunch crossings around the triggering bunch crossing. The larger validation event in contrast provides a window of five bunch crossings around the triggering bunch crossing for the complete set of input, output, and intermediate data.

The maximal size of the read-out record was computed by assuming an upper limit for the read-out rate of 100 kHz, leading to a maximal average event size of 5 kB. This value can be exceeded for validation events, but standard events are required to be smaller in order to compensate. In several areas within both the MP7 and AMC13 firmware there are large buffers that guarantee all events to be transmitted with no deadtime to the DAQ even under high rates of more than 100 kHz. Furthermore, the system can signal to the Trigger Control and Distribution System (TCDS) (described in section 5.3.5) when buffers approach capacity, thereby throttling the trigger.
During the 2016 data taking year the \mu GMT read-out system was improved by the addition of zero suppression logic. In the \mu GMT this feature monitors the value of the muon’s transverse momentum. If all muons of a given link have a value corresponding to zero, these data are disregarded and not sent to the DAQ system. After activation of this system the average \mu GMT read-out record could be reduced from 8504 B to 558 B in total size. The zero suppression logic is designed in a flexible way, allowing the application of zero suppression to arbitrary input data structures.

6.6.2 Monitoring via online software

The \mu GMT calculates several aggregate quantities for every luminosity section\(^5\) which can be read out by the online software together with the counter for the luminosity section. These quantities are computed for all data received by the \mu GMT during a given luminosity section and therefore complement the monitoring information from section 6.6.1.

Rate monitoring

The \mu GMT allows the computation of muon rates at various stages in the system’s logic. Muon counting logic is instantiated for each input link as well as at the outputs of all sorting units. The counting logic consists of two 32 bit registers where the first register is incremented if a muon data word with transverse momentum greater than zero is received. At each luminosity section boundary the \mu GMT firmware resets the first register after its value has been copied to the second register. The online monitoring software reads both the luminosity section counter and the values of the muon counters via a custom network protocol based on TCP/IP — IPbus \[^{112}\] — that provides software access to the \mu GMT registers. It then computes the muon rate by dividing the provided counts by 23.3 s (the approximate length of a luminosity section). The counters are continuously monitored by the online software and warnings are issued in case of lower or higher rates than expected for nominal operation.

As the muon rates for the barrel region ($|\eta| < 0.8$) can be used to estimate the instantaneous luminosity, the \mu GMT rate monitoring system is additionally exploited by the CMS Collaboration as a complementary offline luminometer. In fig. 6.16 the rates of muons in the barrel as recorded at the output of the \mu GMT barrel sorter show the effect of a so-called Van der Meer scan \[^{113}\] in the CMS interaction point. During such a scan the LHC moves the two beams through each other in the transverse plane, allowing experiments to calibrate their luminosity measurements per colliding bunch pair from the LHC machine parameters.

As can be seen in fig. 6.17, the rates at the output of the barrel muon sorter stage are strongly correlated to the values obtained from the Pixel Cluster Counting (PCC) algorithm, described in \[^{114}\].

Luminosity section counter

In order to both measure the number of muons received during a given luminosity section, as well as correlate the value of the muon counters to a given luminosity section, a dedicated unit was introduced that uses signals received from the TCDS

---

\(^{5}\)A luminosity section is defined as $2^{18}$ LHC orbit periods and corresponds to about 23.3 s. Its begin and end are synchronised centrally for the entire CMS experiment by fast control signals sent by the TCDS.
Figure 6.16: Behaviour of the barrel muon rates as measured at the output of the μGMT barrel sorter during a Van der Meer scan.

Figure 6.17: Ratio of the luminosities per 10 luminosity sections measured from the barrel muon (DT) and PCC algorithms over the full 2016 data taking year. [114]
to detect the start of a CMS data taking run. The logic waits for the reception of an orbit counter reset signal, followed by a bunch counter reset signal as this marks the begin of the first luminosity section. It then counts the number of bunch counter reset signals, each of which indicating the completion of one LHC orbit. After $2^{18}$ orbits the luminosity section counter is incremented and a reset signal is sent to all muon counters. As a debugging feature the online software can override this periodic muon counter reset by setting a special register. In this mode a reset signal can be initiated by the control software. The luminosity section counter is exposed for read-out by the online monitoring software and is used to correlate a given monitoring value with the associated luminosity section.

**Link monitoring**

As described in sections 6.4.1 and 6.4.2, the μGMT expects the presence of markers indicating the bunch crossing number corresponding to the received event. At each bunch crossing, the μGMT checks both the tag indicating the first bunch crossing of the LHC orbit, as well as the three LSBs of the downstream systems’ bunch crossing counters and compares them to its internal bunch crossing counter. For each link two separate error counters are incremented for disagreements with either the bunch crossing zero marker or the LSBs of the bunch crossing counter. A read operation via the IPbus protocol automatically resets them to zero.

Additionally, the MP7 board provides link integrity monitoring, using both the alignment words mentioned in section 6.4 as well as cyclic redundancy check (CRC) sums computed by the link logic. Error counters for these monitorables are exposed by the SoftWare for Automating the conTrol of Common Hardware (SWATCH) framework.

**6.7 Technological challenges**

**6.7.1 Distributed firmware development**

Due to the complexity of designing a processing board equipped with a modern high-performance FPGA and high-speed optical links the upgraded Level-1 trigger was foreseen to use largely common hardware (described in section 5.3.1), designed by several universities. The μGMT is implemented in the MP7 board, which includes infrastructure firmware that is amongst others responsible for the read-out system, optical links, and IPbus controls.

Due to the stringent schedule, development of the infrastructure firmware progressed in parallel to the algorithm design, introducing logistical and technological challenges during the development.

**Physical separation of algorithm from infrastructure logic**

In order to allow the parallel development of infrastructure and algorithm logic it is useful to separate the circuitry for each of these domains on the chip-level.

This can be accomplished by so-called floorplanning. Floorplanning means the restriction of a firmware block to a certain region on the FPGA using area constraints. The synthesis tool is
supplied with a so-called constraints file that contains statements limiting placement of a given firmware block to a certain area of the chip.

The Virtex-7 FPGA used in the MP7 card is subtended into 20 clock regions. The logic elements in each clock region can be supplied by 12 different clock signals with almost no skew. Every clock region is equipped with four high-speed transceivers. The transceivers of 18 clock regions are connected to the optical links of the MP7 card while the remaining transceivers provide links to the backplane.

In the MP7, the infrastructure logic has been constrained to the edges of the chips, encompassing the chip resources for the high-speed optical links as well as two clock regions reserved for the read-out and IPbus protocol implementation. The two clock regions correspond to those with transceivers connected to the backplane for communication with the AMC13 read-out board. The remaining area of the FPGA is dedicated to the algorithm logic for the μGMT.

By maintaining this separation, parallel development of adjacent firmware blocks could be greatly simplified as the occurrence of resource conflicts was minimised.

**Build scripts**

The MP7 infrastructure firmware is maintained in an Apache Subversion (SVN) source control repository. Stable versions are published as tags within this system and can be retrieved from the central repository by scripts that include the option to automatically build a minimal version of the firmware. In contrast, the main development branch of the μGMT firmware is stored in a Git repository due to advantages during the initial development.

In order to automatically integrate the algorithms for the μGMT into the retrieved infrastructure firmware, specially customised software was created. This allowed fast turnaround after the release of new infrastructure firmware and the ability to flag any problems caused for the build process promptly.

**6.7.2 Routing on large FPGAs**

Modern FPGAs, such as those used in the upgrade of the CMS Level-1 trigger provide a large amount of logic and memory resources, but are mostly constrained by the difficulty to route signal paths between different components due to high component and signal densities.

The optimal design for such a high-density chip utilises comparatively high clock frequencies with short paths between registers and processing of streams of data, contained as much as possible within a single clock region in order to maintain low levels of clock skew.

For the μGMT however, it is necessary to maintain a comparatively low clock frequency of 40 MHz, in particular in the cancel-out and sorting units (see sections 6.5.3 and 6.5.4) due to the necessity of comparing all muons to each other in parallel. This constraint leads to difficulty during routing. In the following, strategies are described to combat this challenge.

---

6Clock skew describes a phenomenon in which a clock signal arrives at different components at different times.
6.7 Technological challenges

Restriction of firmware modules to contiguous areas

For the development of the MP7 algorithms care was taken to process most data at the link frequency of 240 MHz as described in section 6.5. For the firmware blocks at the LHC frequency of 40 MHz and other modules with challenging routing environments special area constraints were introduced to improve routing performance.

In particular the deserialising and extrapolating logic was constrained to the clock regions associated to their assigned optical links. Similarly, the routing intensive cancel-out units were each constrained to the clock regions opposite of the ones receiving associated data (fig. 6.18). In contrast the sorting units were left unconstrained as they receive data from several clock regions and experience showed routing algorithms in the synthesis software to be superior to manual floorplanning.

Finally, the serialiser firmware was constrained to clock regions housing associated output links, analogously to the deserialisation units.

Build farms for use with Xilinx ISE

The current μGMT firmware is being synthesised with the Xilinx Vivado Design Suite that was created for the synthesis and implementation of large logic designs. However during initial development the MP7 firmware was only compatible with the Xilinx ISE Design Suite, an older software suite that was intended for less routing-heavy chip designs.

In order to accomplish routing with Xilinx ISE, a feature called SmartXplorer was used. This tool allows to attempt the parallel routing of a given chip design using different strategies on multiple hosts. In addition it allows the usage of multiple so-called cost tables. The cost tables contain information for the initial placement of various components on the FPGA and are required by the routing algorithms applied by Xilinx ISE. Using an alternative cost table can significantly improve routing results as it may be more suitable to a given design.

The μGMT made use of a build farm composed of spare read-out unit (RU) servers of the DAQ system, helping to reduce synthesis time from several days to less than 10 hours.

6.7.3 Ensuring stability of the firmware

During development it is imperative to continuously ensure the correctness of the firmware. Once the μGMT firmware code reached a stable state, so-called test benches were created. These test benches use known input and output data to validate the firmware. They can be used in a build system that is executed each night in order to verify the correctness of the firmware as well as the ability to build it without timing errors.

Test benches

Test benches are small wrapper programs written in VHDL whose purpose is to test a specific firmware module. The test bench instantiates the module under test and reads a file containing

---

7 A timing error is an error condition appearing if the placement and routing steps can not establish a chip layout where all paths have signal propagation times under the requested limits.

---
Figure 6.18: Illustration of area constraints for the $\mu$GMT firmware. The Virtex-7 FPGA is split into 20 clock regions with 10 on each side of the chip. Muon data arrive at the optical links connected to the clock regions on the left while calorimeter data arrive on the right-hand side. Deserialisation and extrapolation logic is constrained to those clock regions associated with their assigned links, while cancel-out units are constrained to clock regions opposite the connected deserialisation units. Not shown are the area constraints for the transceiver logic as well as the serialiser logic. The transceiver logic in each clock region is placed close to the FPGA’s edges, while the serialiser logic is placed within the area foreseen for the algorithm logic. The serialiser logic also shares the clock regions used by the cancel-out units because the output transceivers used are placed on the right-hand side of the FPGA.
6.7 Technological challenges

test data as well as the expected output. The input is piped into the unit under test and the output is synchronised to the expected values from the external file. An error counter is incremented in case of mismatches and returned after execution completes. More detailed information can be output in case of errors, allowing efficient debugging.

In order to test the $\mu$GMT firmware, several testbenches were written, testing individual modules as well as the complete $\mu$GMT algorithm firmware. Additionally, several different test files were generated, each optimised to find certain classes of errors. The current $\mu$GMT logic is routinely tested with test files generated from samples containing $t\bar{t}$ pairs as well as $J/\psi$ particles that decay to two muons.

Continuous integration

In order to routinely test the firmware’s functional correctness as well as ensure the ability to complete routing without timing errors, a continuous integration system was set up using the Jenkins continuous integration server [119]. This system was developed for software projects where it is used to run unit and integration tests as well as build the software project.

In the $\mu$GMT project the Jenkins server was set up to run all test benches after each commit to the central version control system. A small python script was written that examines the output of the test benches and from that determines success or failure. This python script is used by the continuous integration server as interface to the result of the test benches. Furthermore, the Jenkins server was set up to perform a complete firmware synthesis every night in order to ensure that routing succeeds. Similarly to the script used for the test benches, a small python script was written to parse the synthesis output and determine success or failure for usage by the Jenkins server. In case of failure an e-mail is automatically sent to a predefined group of developers and a dashboard indicates the current build and test status (fig. 6.19).

Figure 6.19: Screenshot of the dashboard for the continuous integration server Jenkins. Shown are the various tests and builds that can be programmed to be run at a certain time or triggered by specific actions. A green circle with a tick mark indicates a test or build process that has recently completed successfully, a red circle with an exclamation mark indicates failure. The stability over the last five runs is illustrated by “weather”, ranging from sunny to thunderstorms.
In this chapter several studies conducted on the performance of the muon trigger are shown in the context of analyses involving J/ψ mesons that decay to two muons (see also chapter 2). In the first section, a new double muon trigger algorithm using an opposite charge requirement on two muons is presented and its performance measured. Such an algorithm was already feasible in the legacy trigger system and could potentially reduce the trigger rate by almost a factor two compared to an identical double muon trigger without such a requirement. For this reason it was introduced during the 2015 data taking period with the legacy trigger in order to cope with the increased luminosity. In the following section the pure double muon detection efficiency for the upgraded Global Muon Trigger (μGMT) is optimised by investigating various matching windows using Monte Carlo (MC) generated events. Both the expected rate and efficiency are measured on data taken in 2016. Subsequently, a study is presented on a novel trigger algorithm utilising a restriction on the invariant mass of two muons. To calculate the invariant mass, the spatial coordinates at the vertex are computed by the μGMT (see section 6.5.6) and propagated to the upgraded Global Trigger (μGT). The study shows the effect of using the extrapolated coordinates as well as improvements gained through the use of a more elaborate extrapolation algorithm. Finally, the expected efficiency and rate of trigger algorithms utilising this novel feature are shown.

7.1 Monte Carlo event samples

For all analyses presented in the following, four simulated event samples were used. The efficiency of the opposite charge requirement in the legacy trigger system is computed with a sample of 500 000 J/ψ → μμ events with a total centre-of-mass energy of \( \sqrt{s} = 13 \text{ TeV} \). The sample was generated centrally with the Pythia [120] event generator. Detector simulation was performed using the Geant4 [121, 122, 123] simulation toolkit, while reconstruction was done using the CMS Software (CMSSW) framework [107, 108]. The transverse momentum of the generated J/ψ mesons was required to be above 7.5 GeV at the generator level. Additional minimum bias events1 were mixed with the desired event to emulate the environment of an average pile-up\(^2\) of 40, the value expected to be produced by the Large Hadron Collider (LHC) during Run-2. For technical reasons the efficiency of the opposite charge requirement

---

1 All inelastic scattering processes occurring during proton-proton collisions.
2 The mean number of proton-proton collisions during one bunch crossing.
in the upgraded trigger system was computed with a similar sample of almost one million \( \Upsilon \rightarrow \mu\mu \) events at \( \sqrt{s} = 13 \text{ TeV} \). Again, the events were generated centrally with \textsc{Pythia} and reconstructed using \textsc{CMSSW}. The transverse momentum of the simulated \( \Upsilon \) mesons is generated larger than \( 8 \text{ GeV} \) and the added pile-up is in the range of 28 to 62, drawn from a flat distribution.

For the estimation of the effect on the double muon efficiency by the cancel-out unit, a sample of \( \Upsilon \rightarrow \mu\mu \) MC events with a total centre-of-mass energy of \( \sqrt{s} = 13 \text{ TeV} \) is used. A data set of 15 million events was centrally generated with \textsc{Pythia} and simulated using \textsc{Geant4} before being passed to \textsc{CMSSW} for reconstruction. The generation of \( \Upsilon \) mesons was restricted at the generator-level to a range from 0 to 2.4 in pseudorapidity and 0.17 to 1.57 in azimuthal angle in order to cover all types of overlaps between track finder processors due to the symmetries of the \textit{Compact Muon Solenoid (CMS)} detector while simultaneously reducing the number of particles required for high statistical significance. Additionally, the generator was restricted to creating \( \Upsilon \) mesons in the transverse momentum range of 20 GeV to 120 GeV with a flat distribution in order to create a significant amount of close-by muon pairs from decaying high transverse momentum \( \Upsilon \) particles.

The ghosting probability within the muon system is derived with a generated sample of negatively charged single muons. A data set of 15 million events was generated centrally using a specialised particle gun included in the \textsc{CMSSW} framework as generator. The spatial restrictions used for the third \( \Upsilon \rightarrow \mu\mu \) sample were applied, while the generated transverse momentum was restricted to a range of 3 GeV to 70 GeV. In order to provide a constant number of both low-transverse momentum and therefore strongly bending muons as well as almost straight high-\( p_T \) particles a flat transverse momentum distribution was required at the generation step.

For the data samples used to study the cancel-out unit performance in terms of efficiency and ghosting probability no pile-up was added.

The performance of the extrapolation unit was estimated by determining the deviation of the spatial coordinates reported by the Level-1 trigger from the MC generated values. For this study the \( \Upsilon \rightarrow \mu\mu \) event sample without pile-up used for the evaluation of the cancel-out performance was employed. Additionally, the invariant mass resolution of the Level-1 trigger was studied using the \( \Upsilon \rightarrow \mu\mu \) MC event sample with transverse momentum greater than 8 GeV and pile-up in the range of 28 to 62, also used for the measurement of the opposite charge tagging performance.

### 7.2 Charge tagging performance

As detailed in chapter 3, the instantaneous luminosity delivered by the \textit{LHC} during its Run-2 significantly exceeds the maximal value recorded in Run-1. Because the Phase-1 upgrade of the Level-1 trigger was scheduled to enter operation in 2016, the second year of Run-2 (section 5.3), the \textit{CMS} experiment was still taking data with the legacy Level-1 trigger system in 2015. This required the development of novel trigger algorithms for the legacy trigger system to better cope with the increased luminosity delivered by the upgraded \textit{LHC}.

As can be seen in fig. 7.1a the muon rate rises linearly with luminosity. Furthermore, fig. 7.1b shows that most of the single muon rate originates from low-transverse momentum muons. A
7.2 Charge tagging performance

(a) The background muon rates in the resistive plate chamber (RPC) system can be seen to grow linearly as a function of instantaneous luminosity. [124]

(b) The rate for a single muon trigger path as a function of the transverse momentum threshold. [125]

Figure 7.1: The muon trigger rate follows a linear dependence on the instantaneous luminosity while the single muon trigger rate is inversely proportional to the transverse momentum threshold.

The vast majority of these soft muons constitute background that can be safely ignored for analyses of quarkonia. Thus, the most straightforward rate suppression strategy is to simply increase the muon transverse momentum threshold in the Global Trigger (GT). For quarkonium studies described in chapter 2 it is however important to collect a significant sample of low-$p_T$ $J/\psi$ mesons that in turn decay to low-$p_T$ muons. For this reason it becomes vital to efficiently distinguish muons that resulted from $J/\psi$ decays with low $p_T$ from muons originating from background processes.

Typically, quarkonia that decay to muons are collected using a double muon trigger which by itself aids in the rejection of a large amount of background processes. However, with the increased luminosities provided by the Run-2 LHC it was expected that also double muon triggers would have to be either raised in threshold beyond usefulness or prescaled significantly, similarly reducing their utility. One of the instruments that can be used to control the rate from background processes is an opposite charge requirement on a lower-$p_T$ double muon trigger algorithm. If charge tagging in the Level-1 trigger is efficient, this can reduce the rate by almost a factor two (see fig. 7.2).

In the 2015 Level-1 trigger the charge is measured in each track finder independently and then merged in the Global Muon Trigger (GMT) according to complex rules when two complementary measurements of a track are available (see section 5.2.2 for a description of the GMT). For high-$p_T$ muons the charge cannot be determined reliably, as their tracks are indistinguishable from straight lines at the Level-1 trigger: In the absence of magnetic bending it is impossible to determine the charge. However, muons with high transverse momenta can be selected by using
Chapter 7 Using the upgraded Level-1 Trigger for physics

Double Muon, $q>7$ rate
Double Muon opposite charge, $q>7$ rate

Runs 279588-279991, $\sqrt{s} = 13$ TeV
ZeroBias

Figure 7.2: The gain in rate rejection when adding an opposite charge requirement to a double muon trigger algorithm as a function of transverse momentum.
a higher threshold double muon trigger algorithm without the opposite charge requirement because the rate at higher thresholds is significantly lower.

The following studies show the expected and obtained performance of trigger algorithms utilising the opposite charge requirement both for the year 2015 as well as the expected performance using the upgraded trigger system.

### 7.2.1 Study of opposite charge requirement in the Level-1 trigger for LHC Run-2

To establish the feasibility of a trigger algorithm including an opposite charge requirement, studies were performed both on data collected during a run period in 2012 as well as on MC generated events that simulated the accelerator and detector configuration for 2015 (see section 7.1). For the 2012 data set the Level-1 trigger results were recreated with the Level-1 trigger emulator in the configuration used in 2012, but with a modification done to remedy a problem found in the charge assignment logic of the Cathode Strip Chamber Track Finder (CSCTF). The Level-1 trigger decision for the simulated events was emulated using the configuration foreseen for 2015. For both the data taken in 2012 and the simulated data set only events containing two muons within an invariant mass window of \( M(\mu\mu) \in [3, 3.2] \) and with opposite charge as determined by the full offline reconstruction algorithms were considered.

A further study comparing the 2015 detector and trigger configuration to the Phase-1 upgrade trigger system was performed using a separately generated MC data sample to determine the \( \mu \)GMT performance. The same invariant mass window and requirements as before were applied on the two muons.

In the presented studies the efficiency is computed by considering any muon pair that was measured by the Level-1 trigger system and satisfies the double muon trigger algorithm’s quality requirements in the denominator and adding the opposite charge requirement for muon pairs to be included in the numerator (eq. (7.1))

\[
\epsilon = \frac{\text{number of Level-1 muon pairs with opposite charge}}{\text{number of Level-1 muon pairs}}
\]  

(7.1)

The study does not require the explicit correct measurement of both muons’ charges, but any muon pair with opposite charges is considered to contribute to the trigger algorithm’s efficiency. In this way events with mismeasured charge of both muons are counted as efficient. The efficiency, \( \epsilon_{\text{OS}} \), of such an opposite charge requirement could thus be written as

\[
\epsilon_{\text{OS}}(\mu\mu) = \epsilon_{\text{CS}}(\mu) \cdot \epsilon_{\text{CS}}(\mu) + (1 - \epsilon_{\text{CS}}(\mu)) \cdot (1 - \epsilon_{\text{CS}}(\mu))
\]

(7.2)

where \( \epsilon_{\text{CS}} \) describes the efficiency of correctly measuring the charge of a single muon.

**Results**

Figure 7.3 shows the results for studies on the opposite charge requirement in the legacy Level-1 trigger. The performance is near-optimal in the barrel region of \(|\eta| < 0.8\) and remains above 85% up to \(|\eta| = 1.6\) (see fig. 7.3a). For larger pseudorapidity values the data samples used in the study did not provide a sufficient number of events. However, this has no influence on
the validity of the study, as the B physics trigger algorithms utilising such a requirement were foreseen to be restricted to a pseudorapidity range of $|\eta| < 1.6$.

Figure 7.3b shows the dependence of the opposite charge requirement on the J/$\psi$ transverse momentum. While charge tagging for low-transverse momentum muons is assumed to be very efficient for the barrel region, the behaviour of the opposite charge requirement at high $p_T$ was a concern at the outset of the study. The plot shows that high efficiency can be maintained also for J/$\psi$ mesons with considerable transverse momenta. The significantly higher efficiency of the trigger at low transverse momentum for 2015 MC generated events compared to 2012 data can be attributed to upgrades in the endcap muon system, as can be seen in fig. 7.3a. The upgrade of the endcap muon system is detailed in sections 4.6.2 and 4.6.3.

Figure 7.3: Plots of the Level-1 trigger efficiency for the opposite charge requirement in the invariant mass window 3 GeV to 3.2 GeV.

A comparison of the legacy Level-1 trigger in the 2015 configuration with the Phase-1 upgrade deployed for 2016 is shown in fig. 7.4. The efficiency at high pseudorapidities could be significantly improved while a large drop around $|\eta| = 1$ is being investigated by the Overlap Muon Track Finder (OMTF) group. In fig. 7.4b a significant decrease in efficiency for low-$p_T$ J/$\psi$ mesons compared to the legacy system is shown, most likely due to the aforementioned drop at $|\eta| = 1$. For high transverse momenta a slight increase of the upgraded trigger’s efficiency compared to the legacy system can be seen, largely due to the improvements at high pseudorapidity. The high efficiency at large transverse momenta of the J/$\psi$ meson is maintained.

Figure 7.5 shows a further study of the efficiency drop in the OMTF system. The efficiency is computed in the same way as for the above study, by requiring both muons to have been detected by the Level-1 trigger and adding the opposite charge requirement in the numerator. When compared to the complete pseudorapidity range of the muon system (fig. 7.5a), the OMTF
7.2 Charge tagging performance

(a) The efficiency of the opposite charge requirement as a function of $\eta$. The upgraded system shows higher efficiency at high pseudorapidity than the legacy system, but suffers from a significant drop at $|\eta| = 1$.

(b) Plot of the opposite charge requirement efficiency vs. transverse momentum. The efficiency shows a drop for low $p_T$ J/$\psi$ mesons, but recovers and remains stable for moderately high transverse momentum J/$\psi$ mesons.

Figure 7.4: Plots of the Level-1 trigger opposite charge requirement efficiency for the upgraded Level-1 trigger.

system (fig. 7.5b) displays a significant reduction in efficiency for muon pairs where the muon with lower transverse momentum ("trailing leg") has a $p_T$ of less than 4 GeV.

7.2.2 Analysis of 2015 data

The efficiency of the opposite charge requirement in the legacy Level-1 trigger was validated using data taken with special High-Level trigger (HLT) paths during the beginning of the run in 2015. Two paths were introduced:

- **HLT_Dimuon0er16_Jpsi_NoVertexing** — a pass-through path for the Level-1 double muon algorithm including the opposite charge requirement under study. No threshold is applied on the transverse momentum, but a restriction in pseudorapidity of $-1.6 \leq \eta \leq 1.6$ is implemented.

- **HLT_Dimuon0er16_Jpsi_NoOS_NoVertexing** — a pass-through path on a Level-1 double muon trigger algorithm identical to the above but without the opposite sign requirement.

Both paths are prescaled by a factor 50, i.e. only every 50th event that matches the particular trigger is actually stored on disk. The study considers all events with two offline reconstructed

---

3Level-1 pass-through triggers transfer events accepted by Level-1 triggers to storage without any additional processing or selection.
Chapter 7 Using the upgraded Level-1 Trigger for physics

Figure 7.5: Plots of the Level-1 trigger opposite charge requirement efficiency in the upgrade system. Efficiency displayed as a function of both the leading (higher $p_T$) and trailing (lower $p_T$) muons’ transverse momenta. Plots drawn for the full coverage of the muon system (a) and a restricted range of $0.8 \leq |\eta| \leq 1.3$ (b).

muons that were measured in both the inner tracking and the muon system, have opposite charge, and are in the invariant mass window $M(\mu\mu) \in [3, 3.2]$ GeV.

While suffering from the small number of events recorded, the plots in fig. 7.6 show a satisfactory efficiency for the proposed trigger path.

The total efficiency was determined to be $\epsilon_{OS}(\mu\mu) = 0.952 \pm 0.019$.

As a result of the above studies a dedicated trigger algorithm exploiting the opposite charge requirement in order to select $J/\psi$ mesons was introduced in 2015 and has contributed to the data set used by the B physics analysis groups since then.

7.3 Removal of duplicate muon candidates

As explained in section 5.3.3, in the upgraded Level-1 trigger system muon primitive information is shared between all neighbouring track finder processors. This is done in order to improve the muon reconstruction efficiency at the track finder processor coverage boundaries and is implemented both between processors of the same track finder system as well as between those of adjacent systems. While the reconstruction efficiency is improved in this way, the rate of fake double muon events is increased as well because the same track can be reconstructed in two track finder processors separately. As single muon events are more common than double muon events by up to two orders of magnitude at common transverse momentum thresholds for double muon trigger algorithms (see fig. 7.7), it is imperative to suppress such so-called ghosts to avoid the double muon trigger rate becoming dominated by duplicates of single muons.

The $\mu$GMT is equipped with a cancel-out system that can use either spatial coordinates or track addresses to detect duplicates (see section 6.5.3). For each boundary type a different choice can be made, however track addresses are only provided by the Barrel Muon Track Finder (BMTF) system at the time of writing. While the Endcap Muon Track Finder (EMTF) firmware is
7.3 Removal of duplicate muon candidates

Figure 7.6: Plots of J/ψ trigger efficiency for the opposite sign (OS) double muon trigger path in the invariant mass window 3 GeV to 3.2 GeV plotted against the J/ψ pseudorapidity (a) and transverse momentum (b).

Note: Efficiencies can be larger than 1 because the two trigger paths are prescaled and thus contain different events.

Figure 7.7: Comparison of the Level-1 trigger single muon and double muon rates as a function of transverse momentum thresholds. Only muons passing the double muon quality threshold of $q > 7$ are considered. For the double muon trigger algorithm the same threshold in transverse momentum is applied for both muons.
equipped to provide track addresses, this has not yet been implemented in the EMTF emulator and has thus not been studied.

7.3.1 Estimating the spatial separation of duplicates in the Level-1 trigger

The spatial coordinate-based cancel-out system considers two muons to be duplicates if they are located in an area $\Delta R^2 \leq (f_1 \cdot \Delta \phi)^2 + (f_2 \cdot \Delta \eta)^2$ around each other (see section 6.5.3). It is thus useful to investigate the expected distance between a muon and its ghost in order to find optimal values for $f_1$ and $f_2$.

To determine this value, a simulated single muon data set without pile-up was used. The analysis considered only events in which the Level-1 trigger detected two muons, indicating that a duplicate muon had been generated. The distance between the two Level-1 muon objects is shown in figs. 7.8 to 7.12.

Figure 7.8 shows that the vast majority of duplicates occurring with the BMTF system are contained within a pseudorapidity range of $|\Delta \eta| < 0.125$ when the $\eta$-fine bit is set (fig. 7.8a), while this range increases to $|\Delta \eta| < 0.3$ when the $\eta$-fine bit was cleared because the pseudorapidity of the muon or its duplicate could not be measured precisely (fig. 7.8b). The azimuthal range of duplicates is shown to be $|\Delta \phi| < 0.2$ in fig. 7.8c.

In the OMTF system the maximal distance between a muon and its ghost in pseudorapidity is about $|\Delta \eta| < 0.2$, and in azimuthal angle approximately $|\Delta \phi| < 0.1$, although the majority of duplicates are contained in $|\Delta \phi| < 0.025$ as shown in fig. 7.9.

Duplicates in the EMTF system are usually separated by less than 0.1 in both pseudorapidity and azimuthal angle, see fig. 7.10.

Events with duplicates in neighbouring track finder systems were also examined. In figs. 7.11a to 7.11c the expected difference in pseudorapidity and azimuthal angle between two muons in the BMTF and OMTF systems is shown to be smaller than 0.15 and 0.025, respectively. The difference in pseudorapidity is not influenced by the accuracy of the measurement in the BMTF system, indicating that the imprecision originates mostly from the OMTF system.

In figs. 7.12a and 7.12b $|\Delta \eta| < 0.1$ and $|\Delta \phi| < 0.075$ are shown for the majority of events with duplicate muons in the EMTF and OMTF systems.

In the above data the distance of a BMTF muon track from its duplicate was shown to be large compared to the other systems. As track addresses are available for this region, the cancel-out method based on track addresses was chosen for muons originating from the BMTF system. For the remaining cancel-out units two sets of values for the cancel-out windows were evaluated against a baseline of cancel-out windows with $|\Delta \eta| < 0.1$ and $|\Delta \phi|$ maximally large at about 0.076 — the cancel-out window is restricted in azimuthal angle to slightly more than 0.076 due to firmware limitations outlined in section 6.5.3. The two sets are labelled aggressive and conservative, where the aggressive tuning attempts to maximise the potential efficiency (thereby accepting higher rate due to ghosts), while the conservative set of parameters is tuned to reduce the ghost rate (table 7.1).
7.3 Removal of duplicate muon candidates

(a) Distance in pseudorapidity between two duplicates in the BMTF system where both muons have a well measured pseudorapidity.

(b) Distance in pseudorapidity between two duplicates in the BMTF system where one or both muons have coarsely measured pseudorapidity.

(c) Azimuthal distance between two duplicates in the BMTF system.

Figure 7.8: Distance between two muons reconstructed by the BMTF in events with a single generated muon. A threshold of $q > 7$ on the output quality was applied.
Chapter 7 Using the upgraded Level-1 Trigger for physics

Figure 7.9: Distance between two muons reconstructed by the OMTF in events with a single generated muon. A threshold of \( q > 7 \) on the output quality was applied.

Figure 7.10: Distance between two muons reconstructed by the EMTF in events with a single generated muon. A threshold of \( q > 7 \) on the output quality was applied.
7.3 Removal of duplicate muon candidates

(a) Distance in pseudorapidity between two duplicates in the BMTF and OMTF system where both muons have a well measured pseudorapidity.

(b) Distance in pseudorapidity between two duplicates in the BMTF and OMTF systems where the BMTF muon has coarsely measured pseudorapidity.

(c) Azimuthal distance between two duplicates in the BMTF and OMTF systems.

Figure 7.11: Distance between two muons reconstructed by the BMTF and OMTF systems in events with a single generated muon. A threshold of q > 7 on the output quality was applied.
Table 7.1: Parameters considered for usage in the cancel-out look-up tables (LUTs) for the coordinate-based matching algorithm in the µGMT. For the duplicate removal between wedges of the BMTF system the track address-based algorithm was chosen.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Baseline</th>
<th>Conservative tuning</th>
<th>Aggressive tuning</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>OMTF</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\Delta R$</td>
<td>0.1</td>
<td>0.2</td>
<td>0.05</td>
</tr>
<tr>
<td>$f_\eta$</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>$\Delta \eta_{\text{max}}$</td>
<td>0.1</td>
<td>0.2</td>
<td>0.05</td>
</tr>
<tr>
<td>$f_\phi$</td>
<td>1</td>
<td>8</td>
<td>2</td>
</tr>
<tr>
<td>$\Delta \phi_{\text{max}}$</td>
<td>$\approx 0.076$</td>
<td>0.025</td>
<td>0.025</td>
</tr>
<tr>
<td><strong>EMTF</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\Delta R$</td>
<td>0.1</td>
<td>0.1</td>
<td>0.05</td>
</tr>
<tr>
<td>$f_\eta$</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>$\Delta \eta_{\text{max}}$</td>
<td>0.1</td>
<td>0.1</td>
<td>0.05</td>
</tr>
<tr>
<td>$f_\phi$</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>$\Delta \phi_{\text{max}}$</td>
<td>$\approx 0.076$</td>
<td>$\approx 0.076$</td>
<td>0.025</td>
</tr>
<tr>
<td><strong>BMTF &amp; OMTF</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\Delta R$</td>
<td>0.1</td>
<td>0.3</td>
<td>0.15</td>
</tr>
<tr>
<td>$f_{\eta, \text{fine}}$</td>
<td>1</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>$\Delta \eta_{\text{fine, max}}$</td>
<td>0.1</td>
<td>0.15</td>
<td>0.15</td>
</tr>
<tr>
<td>$f_{\eta, \text{coarse}}$</td>
<td>1</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>$\Delta \eta_{\text{coarse, max}}$</td>
<td>0.1</td>
<td>0.15</td>
<td>0.15</td>
</tr>
<tr>
<td>$f_\phi$</td>
<td>1</td>
<td>3</td>
<td>6</td>
</tr>
<tr>
<td>$\Delta \phi_{\text{max}}$</td>
<td>$\approx 0.076$</td>
<td>$\approx 0.076$</td>
<td>0.025</td>
</tr>
<tr>
<td><strong>EMTF &amp; OMTF</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\Delta R$</td>
<td>0.1</td>
<td>0.2</td>
<td>0.075</td>
</tr>
<tr>
<td>$f_\eta$</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>$\Delta \eta_{\text{max}}$</td>
<td>0.1</td>
<td>0.2</td>
<td>0.075</td>
</tr>
<tr>
<td>$f_\phi$</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>$\Delta \phi_{\text{max}}$</td>
<td>$\approx 0.076$</td>
<td>$\approx 0.076$</td>
<td>0.025</td>
</tr>
</tbody>
</table>
7.3 Removal of duplicate muon candidates

Figure 7.12: Distance between two muons reconstructed by the EMTF and OMTF systems in events with a single generated muon. A threshold of $q > 7$ on the output quality was applied.

7.3.2 Double muon trigger rates

Ghosting probability

To estimate the rate of duplicates for each cancel-out region a simulated sample of single muon events (described in section 7.1) was used. In order to approximate the transverse momentum distribution produced by proton-proton collisions the sample was re-weighted with a $1/p_T^4$ distribution as it was found to be proportional to the measured transverse momentum spectrum at CMS in [126]. In the following, the percentage of single muon events that generated two muons in the Level-1 trigger is shown both at the $\mu$GMT inputs as well as at the outputs, demonstrating the impact of the $\mu$GMT cancel-out unit. Muons passing the quality thresholds for double muon algorithms\(^4\) are considered.

In fig. 7.13 the ghosting probability as a function of pseudorapidity of the generated muon is shown. As can be seen in fig. 7.13a, the probability for ghosts is almost 20% before duplicates are removed by the $\mu$GMT. Figure 7.13b shows that the $\mu$GMT cancel-out unit performs better than or equally well as the legacy system for close to the entire range of pseudorapidity, except for the region covered by the OMTF system as well as the overlap region between the BMTF and OMTF systems when using the baseline and aggressive tuning. The conservative cancel-out tuning improves in the aforementioned region, leading to comparable or better results than the legacy Level-1 trigger in almost the entire pseudorapidity region covered by the Level-1 trigger.

Figure 7.14 shows the probability for ghosts as a function of azimuthal angle. The spikes in the $\mu$GMT input (fig. 7.14a) can be attributed to the overlaps between track finder processors. As the event sample was generated for a slice of 80° in azimuthal angle of the detector, a large

\(^4\)For the legacy system $q > 2$ is required, while the upgraded system applies a threshold of $q > 7$. 

125
peak can be seen in the centre of the considered region. This peak corresponds to the overlap between the processors of all track finding systems ranging from $\phi = 0.96$ to $\phi = 1.3$ as well as the overlap region of two BMTF wedges at $\phi = 0.785 \pm 0.087$ (see fig. 6.4 for an illustration of the track finder processor coverage). In fig. 7.14b it can be seen that the $\mu$GMT effectively removes the structures at the overlaps between track finder processors. Furthermore, it can be seen that the baseline tuning has a larger ghosting probability than both the aggressive and conservative tuning. The reduction in ghost probability for the conservative and aggressive tuning compared to the baseline tuning can be attributed to the significant enlargement of the cancel-out windows in pseudorapidity for the duplicate removal between BMTF and OMTF.

In fig. 7.15 the probability for ghosts is drawn as a function of transverse momentum. As can be seen, an excess of ghosts appears for low-$p_T$ muons at the input of the $\mu$GMT. This is problematic because as already mentioned the transverse momenta are distributed following a $1/p_T^4$ curve during proton-proton collisions. Figure 7.15b shows that the upgraded system maintains lower ghosting probability than the legacy system both for low-$p_T$ muons as well as for higher transverse momenta. While the ghosting probability for the aggressive tuning increases, it remains below both the baseline tuning and the legacy system.

### Double muon trigger rate

Using the Level-1 trigger emulator a range of proton-proton collision runs from late August 2016 to early September 2016 was employed to estimate the double muon trigger rates as a function of

---

3CMS runs 279 588 to 279 991.

126
7.3 Removal of duplicate muon candidates

Figure 7.14: Probability for the occurrence of duplicates in the Level-1 trigger for different $\mu$GMT configurations drawn versus azimuthal angle. Plots were drawn both with (a) and without (b) the input ghost rate. Events with a single generated muon were used for this study.

Figure 7.15: Probability for the occurrence of duplicates in the Level-1 trigger for different $\mu$GMT configurations drawn versus transverse impulse. Plots were drawn both with (a) and without (b) the input ghost rate. Events with a single generated muon were used for this study.
Chapter 7 Using the upgraded Level-1 Trigger for physics

Symmetrical\textsuperscript{6} transverse momentum thresholds and for trigger paths with thresholds of 18 GeV and 5 GeV for the leading and trailing muon respectively, as well as for no threshold on the muons’ transverse momentum as a function of the leading muon’s pseudorapidity. Furthermore, an estimate of the proportion of ghosts contributing to the measured rate was established. The runs used had a peak instantaneous luminosity from $1 \times 10^{34} \text{cm}^{-2} \text{s}^{-1}$ to $1.3 \times 10^{34} \text{cm}^{-2} \text{s}^{-1}$ and an initial average pile-up value between $\langle n \rangle = 35$ and $\langle n \rangle = 40$.

The double muon rates were computed using a zero bias data set\textsuperscript{7} and counting the number of events with two muons satisfying both the transverse momentum thresholds as well as the requirements on the muon quality for the double muon trigger paths, $N_{\text{pass}}$. This value is multiplied by a normalisation factor, consisting of the number of colliding bunches $n_{\text{coll}}$, the LHC revolution frequency $f_{\text{LHC}} = 11 \text{kHz}$, as well as the total number of events considered $N_{\text{tot}}$ to arrive at the event rate $R$:

$$ R = N_{\text{pass}} \cdot \frac{n_{\text{coll}} \cdot f_{\text{LHC}}}{N_{\text{tot}}} $$

(7.3)

To compute the rate of ghost events contributing to the double muon rate an additional requirement for $N_{\text{pass}}$ was introduced, demanding the presence of only one offline reconstructed muon. In particular the offline reconstructed muon was required to

- have passed loose reconstruction criteria: (see also [127])
  - the muon was reconstructed by the particle flow algorithm
  - the muon was found in the tracker by the reconstruction algorithm
- have been detected in at least the innermost muon station.

In this way only events with one reconstructed but two triggering muons are counted, thus indicating the presence of a duplicate.

Figure 7.16 shows the expected double muon trigger rate as well as the contribution of ghosts to the double muon trigger rate as a function of symmetrical transverse momentum thresholds.

In fig. 7.17 the rates of double muon triggers for the open\textsuperscript{8} as well as 18 GeV and 5 GeV transverse momentum thresholds as a function of pseudorapidity are shown. An imbalance between the positive and negative detector sides for high pseudorapidities can be attributed to two faulty muon stations in the negative endcap during the considered run period.

Figure 7.18 shows the contribution of ghosts to the double muon trigger rates as a function of pseudorapidity. As can be seen, the contribution from ghosts is about one fifth of the corresponding total double muon rate.

The rates for the examined double muon trigger paths as well as the contributions from ghosts are listed in table 7.2. While the rate of the double muon trigger path for the aggressive tuning is larger than for the baseline cancel-out parameters, the ghost rate is almost equal for the two tunings. This indicates a slightly higher efficiency for the aggressive tuning.

\textsuperscript{6}The same threshold for the leading and trailing muon.
\textsuperscript{7}Data taken with an unbiased trigger algorithm that reacts merely to the beam presence monitors indicating two colliding bunches.
\textsuperscript{8}No transverse momentum threshold.
7.3 Removal of duplicate muon candidates

Figure 7.16: Rates of double muon events (a) and ghosts (b) drawn as a function of symmetrical $p_T$ thresholds.

Figure 7.17: Double muon trigger rates drawn as a function of pseudorapidity for no transverse momentum threshold (a) and a threshold of 18 GeV and 5 GeV for the leading and trailing muon, respectively (b).
Chapter 7 Using the upgraded Level-1 Trigger for physics

Figure 7.18: Double muon trigger rates for no transverse momentum threshold (a) and a threshold of 18 GeV and 5 GeV for the leading and trailing muon, respectively (b) when counting only events with one reconstructed muon. This corresponds to the rate contributed to double muon trigger algorithms by single muon events with a duplicated Level-1 muon.

Table 7.2: Rates for double muon trigger paths with and without transverse momentum thresholds.

<table>
<thead>
<tr>
<th>Tuning</th>
<th>DoubleMu0 Rate (kHz)</th>
<th>DoubleMu0 Ghost rate (kHz)</th>
<th>DoubleMu18-5 Rate (kHz)</th>
<th>DoubleMu18-5 Ghost rate (kHz)</th>
</tr>
</thead>
<tbody>
<tr>
<td>No cancel-out</td>
<td>295.7 ± 0.8</td>
<td>211.9 ± 0.7</td>
<td>4.7 ± 0.1</td>
<td>3.42 ± 0.09</td>
</tr>
<tr>
<td>Baseline</td>
<td>61.2 ± 0.4</td>
<td>12.8 ± 0.2</td>
<td>0.52 ± 0.04</td>
<td>0.14 ± 0.02</td>
</tr>
<tr>
<td>Conservative</td>
<td>59.9 ± 0.4</td>
<td>11.6 ± 0.2</td>
<td>0.45 ± 0.04</td>
<td>0.10 ± 0.02</td>
</tr>
<tr>
<td>Aggressive</td>
<td>63.7 ± 0.4</td>
<td>13.8 ± 0.2</td>
<td>0.56 ± 0.03</td>
<td>0.15 ± 0.02</td>
</tr>
</tbody>
</table>
7.3 Removal of duplicate muon candidates

7.3.3 Double muon efficiency

Study using simulated events

The efficiency of the double muon trigger was estimated using a simulated data set containing $J/\psi \rightarrow \mu\mu$ events, described in section 7.1. The study was done assuming transverse momentum thresholds of 18 GeV and 5 GeV for the leading and trailing muon, respectively, corresponding to expected double muon algorithm thresholds in the Level-1 trigger menu for the year 2017.

In order to compute the efficiency, all events containing a generated muon pair with transverse momenta of more than 22 GeV and 9 GeV for the leading and trailing muon respectively are considered in the denominator. These values are chosen to be well above the Level-1 trigger threshold of the considered double muon trigger algorithm and therefore such a muon pair is expected to be detected by the trigger. The numerator is filled with events containing two Level-1 trigger muon objects that satisfy the quality requirements of the double muon trigger path and have transverse momentum above 18 GeV and 5 GeV for the leading and trailing muon, respectively:

$$\epsilon = \frac{\text{number of Level-1 muon pairs with } q > 7 \text{ and } p_T > 18 \text{ GeV and } 5 \text{ GeV respectively}}{\text{number of generated muon pairs with } p_T > 22 \text{ GeV and } 9 \text{ GeV respectively}}$$ (7.4)

For the numerator, the Level-1 muon pairs were additionally matched to generated muon pairs and the same transverse momentum threshold as for the denominator was applied.

The computation is done slightly differently for so-called turn-on curves that show the efficiency as a function of transverse momentum. The transverse momentum requirement on the generated muons is dropped in both the denominator and numerator in order to show the efficiency for the full range of generated transverse momenta.

The plots in fig. 7.19 show the double muon trigger efficiency as a function of different $J/\psi$ observables for a Level-1 trigger transverse momentum threshold of 18 GeV and 5 GeV. The efficiency curves for the legacy system, the input of the $\mu$GMT and the output of the $\mu$GMT for the three cancel-out parameter tunings described in section 7.3.1 are drawn.

For B physics analyses, in particular those probing the $J/\psi$ meson, the cancel-out unit can introduce a challenging problem when close-by muons from the decay of high-$p_T$ $J/\psi$ particles are misidentified as duplicates. Figure 7.19a shows the efficiency as a function of the spatial separation between the two decay products. As can be seen, the $\mu$GMT provides significantly higher efficiency for close-by muons compared to the legacy system and remains at comparable values to the input efficiency.

Examining the efficiency as a function of meson-pseudorapidity in fig. 7.19b shows that significant improvements with respect to the legacy system could be made in particular in the forward area of the detector, above $|\eta| = 1.2$. The drop in efficiency at $|\eta| = 0.2$ can be attributed to a gap in the detector wheels at this pseudorapidity, while the drop in the region of $1 < |\eta| < 1.2$ can be explained by the missing implementation of hit information sharing in the OMTF system (see section 5.3.3).

Figure 7.19c illustrates the decrease in efficiency caused by the $\mu$GMT cancel-out units when plotted as a function of azimuthal angle. At the overlaps between the areas covered by the track
Figure 7.19: Double muon trigger efficiency for different μGMT configurations drawn as a function of spatial distance between the two muons (a), as well as of pseudorapidity (b), azimuthal angle (c), and transverse impulse (d) of the generated J/ψ meson. Events with a generated J/ψ meson decaying to two muons were used for this study.
7.3 Removal of duplicate muon candidates

finder processors peaks in efficiency can be noticed, most likely due to ghosts producing “fake efficiency” when a high transverse muon is duplicated and thereby erroneously contributes to the detection efficiency.

Plotting the efficiency as a function of transverse momentum (fig. 7.19d) indicates that the legacy system mostly tracks the efficiency at the μGMT output for low-$p_T$ J/ψ mesons but degrades for high transverse momenta. This largely mirrors the effect seen in fig. 7.19a as high-$p_T$ mesons predominantly decay to two close-by muons (fig. 7.20a). The drop in efficiency at low transverse momenta can be explained by the trailing muon not passing the threshold of 5 GeV in a significant fraction of events (fig. 7.20b).

**Study on data taken in 2016**

To validate the efficiency estimate conducted in the preceding section on MC generated events, a measurement was performed on data taken from proton-proton collisions during early September 2016 with a peak instantaneous luminosity of about $1.2 \times 10^{34} \text{cm}^{-2} \text{s}^{-1}$ and an initial average pile-up value of $<n> = 40$.

A direct measurement of the double muon efficiency is not feasible due to the low number of double muons in an unbiased data set, however a hybrid approach can be taken. It combines a naive double muon trigger efficiency, computed from the square of the single muon trigger efficiency, with a correction factor obtained from MC generated event samples. This approach follows the one presented in [128].

The double muon efficiency can be written as

$$
\epsilon_{\mu\mu, \text{data}} = \epsilon_{\mu, \text{data}} \cdot \epsilon_{\mu, \text{data}} \cdot Q_{\mu\mu}
$$

(7.5)

where $\epsilon_{\mu}$ is the trigger efficiency for single muons, computable using the so-called Tag&Probe method (described later), and $Q_{\mu\mu}$ is a correction factor to account for any inefficiency specific to

---

9 CMS runs 280 017 to 280 385.
Chapter 7 Using the upgraded Level-1 Trigger for physics

double muon trigger paths. Notably, this correction factor accounts for cases when the \( \mu \)GMT cancel-out unit erroneously removes a muon that was misidentified as a duplicate, therefore reducing the double muon trigger efficiency. The \( \varrho \)-factor can be established by re-writing eq. (7.5) as

\[
\varrho_{\mu \mu} = \frac{\epsilon_{\mu \mu, \text{data}}}{\epsilon_{\mu, \text{data}} \cdot \epsilon_{\mu, \text{data}}} \tag{7.6}
\]

which would require the unfeasible computation of the double muon efficiency from data. However, in first approximation it can be assumed that

\[
\frac{\epsilon_{\mu \mu, \text{data}}}{\epsilon_{\mu, \text{data}} \cdot \epsilon_{\mu, \text{data}}} \approx \frac{\epsilon_{\mu \mu, \text{MC}}}{\epsilon_{\mu, \text{MC}} \cdot \epsilon_{\mu, \text{MC}}} \tag{7.7}
\]

Therefore, eq. (7.6) can be written as

\[
\varrho_{\mu \mu} \approx \frac{\epsilon_{\mu \mu, \text{MC}}}{\epsilon_{\mu, \text{MC}} \cdot \epsilon_{\mu, \text{MC}}} \tag{7.8}
\]

which can be computed purely from MC generated event samples.

As mentioned before, the efficiency of the single muon trigger can be measured from physics data by using the Tag&Probe method. This method allows to compute a single muon trigger efficiency while avoiding the usage of unbiased data such as the above mentioned zero bias data set. Unbiased data sets typically contain very low numbers of events with muons and are therefore not suitable for the computation of trigger efficiencies with the required precision. The Tag&Probe method exploits double muon resonances such as the Z boson or the J/ψ meson. The resonance is reconstructed from a pair of muons, with one leg passing a tight selection criterion (labelled \textit{tag}) and one passing a loose selection criterion (labelled \textit{probe}). Each event is examined for the presence of a tag muon and, if this is found, the trigger efficiency is measured using the probe leg. A muon can be used both as tag and as probe muon if it passes both requirements, increasing the statistical significance of the study.

For the presented study the requirements for the tag leg are

- it is the muon that fired the trigger
- \( p_T,\text{reco} > 27 \text{ GeV} \)
- it passes tight reconstruction criteria that are designed to suppress cosmic muons, as well as muons from in-flight decays and pile-up: (see also [127])
  - the muon was found in both the tracker and at the muon system by the reconstruction algorithm; in particular:
    - at least one hit in the Pixel detector,
    - at least five Silicon Strip Tracker layers, and
    - at least two segments from muon chambers
  - \( \chi^2/\text{d.o.f.} < 10 \) for the muon track fit
7.3 Removal of duplicate muon candidates

- the muon was reconstructed successfully by the Particle Flow reconstruction algorithm [129]
- the track has transverse impact parameter \( d_{xy} < 2 \text{ mm} \) and longitudinal distance from the primary vertex of \( d_z < 5 \text{ mm} \)

- meets the isolation criteria applied by the HLT
- is isolated according to the particle flow reconstruction algorithm

and for the probe to

- pass tight reconstruction criteria
- be separated by \( \Delta R > 0.5 \) from the tag muon.

The muon trigger efficiency for a single leg is then computed on the probe muon by requiring the presence of a muon in the Level-1 trigger close to the reconstructed probe muon and applying the quality threshold for the double muon trigger and the transverse momentum thresholds for the leading or trailing muon in the double muon trigger.

![Figure 7.21](image_url)

Figure 7.21: Single muon efficiencies for the Level-1 trigger as a function of muon transverse momentum as determined by the offline reconstruction algorithm. Curves drawn for complete coverage (a) as well as different pseudorapidity regions corresponding to track finder coverage and overlap regions (b). The efficiency is derived from a single muon sample of a data set taken in 2016.

Figure 7.21 shows the single muon trigger efficiency for the 18 GeV threshold with quality requirements consistent with the double muon trigger applied. These efficiency values can be multiplied with those for the single muon trigger algorithm at the 5 GeV threshold to compute
Chapter 7 Using the upgraded Level-1 Trigger for physics

Figure 7.22: Double muon efficiency of the Level-1 trigger as a function of the leading muon transverse momentum. Curves derived from data are drawn for the complete coverage (a) as well as pseudorapidity regions corresponding to track finder coverage and overlap regions (b). Additionally, the efficiency obtained purely from MC is shown for the complete coverage. The curve from data is derived from the single muon efficiency shown in fig. 7.21 and MC results for the baseline cancel-out unit tuning.
7.4 Invariant mass computation

As explained in section 5.3.4, the μGT is equipped with algorithms to compute the invariant mass of two trigger objects by applying the formula

\[ M^2 = 2 \cdot p_{T,1} \cdot p_{T,2} \cdot (\cosh(\eta_1 - \eta_2) - \cos(\phi_1 - \phi_2)), \]

(7.9)

where \( p_{T,i}, \eta_i, \) and \( \phi_i \) are the suspected decay products’ transverse momentum, pseudorapidity, and azimuthal angle measured at the decay vertex. This formula applies in the case of highly relativistic particles, which is true in the LHC environment. In the μGT firmware it is implemented as a combination of simple differences as well as LUTs in order to compute the cosine and hyperbolic cosine.

The accuracy of eq. (7.9) relies on the precise measurement of the transverse momentum, pseudorapidity, and azimuthal coordinate of both trigger objects. For the muon trigger this leads to a particular challenge, as the \( \phi \) value reported by the track finders is measured in the muon system at considerable distance to the vertex and is therefore significantly influenced by magnetic bending.

Figure 7.23 shows the difference between the azimuthal angle reported by the muon track finders and the generated coordinate as a function of the generated transverse momentum for a MC generated sample of J/ψ mesons with a transverse momentum range of 20 GeV to 120 GeV that decay to two muons (section 7.1). As can be seen, the MC truth values diverge significantly from the quantity measured by the Level-1 trigger at low transverse momenta where magnetic bending most strongly affects the muon track.

7.4.1 Corrections on azimuthal coordinates

The μGMT is equipped with a unit to extrapolate muon tracks back to the vertex by using the measured transverse momentum and pseudorapidity (section 6.5.2). This facility was developed to improve the muon isolation performance, however, the extrapolated coordinates can be additionally propagated to the μGT in order to be used in the computation of the invariant mass of pairs of particles. The LUTs currently used in the μGMT firmware assume the lowest 6 bit of the transverse momentum and the highest 6 bit of the absolute value of the pseudorapidity as
input and return a 3 bit wide value, used to compute the difference between the measured and the assumed actual azimuthal angle. This 3 bit value is left-shifted by 3 bit to create a final 6 bit correction. This means that the current firmware of the μGMT can extrapolate muons up to a transverse momentum of 32 GeV\(^{10}\) and apply a correction from 0.087 rad to 0.7 rad\(^{11}\).

The effect of the μGMT extrapolation unit on the azimuthal coordinate reported by the μGMT is shown in figs. 7.24c and 7.24d, created from the same simulated sample of J/ψ mesons used for fig. 7.23. The measurement is improved for low transverse momenta, but some deviation from the value at the vertex remains, in particular at high \(p_T\). Figures 7.24a and 7.24b show the corrections, \(\Delta \phi(\mu_{\text{L1}}, \mu_{\text{L1, extrap.}}) = |\phi(\mu_{\text{L1}}) - \phi(\mu_{\text{L1, extrap.}})|\), applied by the μGMT as a function of transverse momentum.

An alternative implementation of the extrapolation logic was investigated, where the available bit width for the transverse momentum is increased to 7 bit, while the pseudorapidity is reduced to 5 bit in resolution. Simultaneously, the LUT output width is increased to 4 bit, increasing the range of possible values to a range from 0.044 rad to 0.7 rad. The modifications were implemented in the μGMT emulator software, which was used to study the improved extrapolation logic on the sample described earlier. The changes outlined above lead to further improvements in the extrapolated coordinates at the μGMT outputs as can be seen in figs. 7.25c and 7.25d. The increased granularity and extended range of the extrapolation in transverse momentum is shown in figs. 7.25a and 7.25b.

A firmware implementation of the modified extrapolation logic was deployed in time for the 2017 data taking year.

---

\(^{10}\)The transverse momentum scale uses a step-size of \(\Delta p_T = 0.5\) GeV, which means that the 64 values that are covered by 6 bit can be used to encode a scale up to 32 GeV.

\(^{11}\)The step-size for the azimuthal angle is about 0.0109 rad which is multiplied by 8 due to the bit shift operation. The largest value that can thus be represented by this scale is about 0.7 rad.
Figure 7.24: Corrections applied by the μGMT in the current configuration (a, b) to the azimuthal coordinates received from the track finding systems as well as the resulting deviations of the φ coordinates at the output of the μGMT when compared to MC truth (c, d). Separate plots are drawn for leading (a, c) and trailing (b, d) muons.
Figure 7.25: Corrections applied by a possible modification of the $\mu$GMT firmware to the azimuthal coordinates received from the track finding layer of the muon trigger (a, b) as well as the resultant $\phi$ coordinates at the vertex computed by the $\mu$GMT (c, d). Separate plots drawn for leading (a, c) and trailing (b, d) muons.
7.4.2 Invariant mass resolution for the J/ψ meson after extrapolation of the azimuthal coordinate

The possible invariant mass resolution achievable with the Level-1 trigger quantities is shown in fig. 7.26 for the values received from the track finder systems as well as for the basic and elaborate tunings presented in the previous section. The study is performed on a MC generated sample of events containing J/ψ mesons with transverse momentum larger than 8 GeV that decay to two muons, as well as generated pile-up between 28 to 62 (see section 7.1). As can be seen, the invariant mass resolution is significantly improved already by the baseline extrapolation unit and can be further enhanced by changes to the μGMT firmware.

![Figure 7.26: Invariant mass resolution for a J/ψ → µµ data sample at the Level-1 trigger-level.](image)

7.4.3 Rates and efficiency of an invariant mass J/ψ trigger

In order to establish the feasibility of augmenting existing trigger algorithms dedicated to the collection of J/ψ mesons with an additional threshold for the invariant mass, both the rate and efficiency for such an algorithm were estimated. For this study a double muon trigger algorithm with a \( p_T \) threshold of 4 GeV for both muons as well as a second double muon trigger algorithm with no threshold on the transverse momentum were chosen. If the invariant mass computation is sufficiently reliable, adding a suitable threshold on it would allow the transverse momentum threshold to be lowered while maintaining control over the trigger rates. A lower transverse momentum threshold is particularly important for B physics analyses concerning the study of J/ψ mesons due to their low mass and corresponding low transverse momentum of their decay products.
The efficiency for a double muon trigger algorithm that incorporates an invariant mass threshold was computed from events containing two muons within an invariant mass window of $M(\mu\mu) \in [3, 3.2]$ GeV and with opposite charge as determined by the full reconstruction algorithms. Furthermore, both muons must have been detected by the Level-1 trigger system using a double muon trigger with no transverse momentum threshold. The efficiency was then calculated by adding the invariant mass threshold as constraint:

$$\epsilon = \frac{\text{number of } J/\psi \text{ events detected by DoubleMu0 with invariant mass-threshold}}{\text{number of } J/\psi \text{ events detected by DoubleMu0}}$$  \hspace{1cm} (7.10)

As can be seen in fig. 7.27, it would be necessary to set the invariant mass threshold to 12 GeV in order to retain more than 95% efficiency for a $p_T$-unconstrained trigger algorithm if no extrapolation were used. In contrast, the threshold can be lowered to 7 GeV without suffering significant loss of efficiency when using the elaborate extrapolation. A similar argument holds for a trigger algorithm with a transverse momentum threshold of 4 GeV.

Figure 7.27: Estimated Level-1 trigger efficiency for different invariant mass thresholds used with double muon trigger algorithms that have no transverse momentum threshold (a) and a symmetrical threshold of 4 GeV for both muons (b). The efficiency is computed from a sample of $J/\psi$ mesons that have been detected by a double muon trigger algorithm with no transverse momentum threshold.

Trigger rates as a function of invariant mass thresholds were computed using the same zero bias data set employed in section 7.3.2. Analogous to section 7.3.2, the number of events with two muons satisfying both the transverse momentum thresholds as well as the requirements on the muon quality for the double muon trigger path are counted. The rates are then drawn as a function of the applied invariant mass threshold. As can be seen in fig. 7.28, the rate drops significantly with smaller thresholds.

As explained earlier, an invariant mass threshold of 12 GeV were required to retain high efficiency if no extrapolation was applied at the $\mu$GMT. At this threshold the trigger rate for
7.4 Invariant mass computation

the $p_T$-unrestricted algorithm were 90 kHz, while the rate for the algorithm with a transverse momentum threshold of 4 GeV were about 34 kHz. In contrast, using the elaborate extrapolation logic the invariant mass threshold can be lowered to 7 GeV, reducing the rate of the $p_T$-unrestricted algorithm by more than 10% to about 78 kHz and the rate of the algorithm with a 4 GeV transverse momentum threshold to less than 32 kHz.
The design of the upgraded Global Muon Trigger (μGMT) as well as its performance in the context of quarkonium analyses have been presented.

The theoretical description of quantum chromodynamics (QCD) poses computational challenges due to the gluon self-interactions. While short range interactions can be accurately modelled using perturbative QCD, long range interactions that take place at lower energies cannot be computed in this way because the strong coupling constant grows with decreasing energy. Quarkonia — mesons composed of a heavy quark and its antiquark — are excellent tools to probe the interaction of short and long range effects in QCD both during their formation as well as in their polarisation: Due to the high masses of their constituent particles the timescales for the short range and long range effects during their formation are sufficiently different to allow a separate treatment of the two phenomena. However, in particular the decay of the light J/ψ meson to two muons poses challenges at the trigger level. These problems arise at low meson transverse momenta because of a large background of soft muons from hadronic jets, as well as for J/ψ mesons with high transverse momentum because the close proximity of the two muon decay products can cause their mis-identification as duplicates. A well-tuned muon trigger that can identify these decay products with high accuracy is therefore of utmost importance for quarkonium analyses.

The event selection system for the Compact Muon Solenoid (CMS) experiment — called the trigger — is split into two processing stages: A fast hardware-based Level-1 trigger which operates on a subset of the event data for every bunch crossing is responsible for issuing a read-out signal with a rate of up to 100 kHz. The data read-out from the detector are then further reduced by the High-Level trigger (HLT), a software-based event filter system operating in a large computing centre.

Due to the expectation of a significant increase in the instantaneous luminosity during the second running period of the Large Hadron Collider (LHC), an upgrade of the entire Level-1 trigger system became necessary. While the calorimeter trigger path exploited a time-multiplexed design where a single trigger processor board receives calorimetric information for an entire event, the muon trigger chain was split into three parallel track finding systems which each reconstruct muon tracks in a specific geometric region of the detector.

The μGMT forms the final stage of the upgraded muon trigger chain, receiving up to 108 muon objects from these three track finder systems. It finds duplicates and removes them, while in parallel performing a highly parallel sorting operation in order to select eight final muons to be transmitted to the upgraded Global Trigger (μGT). Furthermore, it is capable of receiving energy sums from the calorimeter trigger that can be used to compute an isolation variable for each.
of the final muons. The coordinates of the muon track received by the \( \mu \text{GMT} \) from the muon track finders are measured in the second station of the muon system, at several meters distance from the interaction region and are therefore significantly affected by magnetic bending. To perform an accurate computation of the isolation variable, the muon track is extrapolated from the position reported by the track finders to the interaction region. This extrapolated coordinate is then also provided to the \( \mu \text{GT} \), allowing a more precise computation of a di-object invariant mass. Finally, the \( \mu \text{GMT} \) is equipped with significant self-monitoring capabilities in order to guarantee the correct operation during data taking. Bunch crossing error counters monitor the link integrity by comparing the internally generated bunch crossing counter to the ones transmitted on the input links and muon counters provide a real-time measurement of the muon rate, allowing to detect defects in the muon systems that could lead to high or low rate from a given region. The muon counters for the barrel region are additionally used as a complementary offline luminometer by the CMS Collaboration and rival the baseline Pixel Cluster Counting (PCC) method in accuracy.

The upgraded Level-1 trigger could only be put into operation in the second year of the LHC Run-2. As the instantaneous luminosity for the first year of Run-2 was already expected to rise above the design values, a novel algorithm to improve the suppression of incorrectly identified \( J/\psi \) mesons at the Level-1 trigger was introduced. A study of the effect on the efficiency by the requirement of opposite charge for the two muons in a double muon trigger algorithm was presented. The performance of this algorithm within the legacy system in both Run-1 and the first year of Run-2 was studied and then compared to the upgraded muon trigger’s performance using the same requirement on the two muons.

As previously mentioned, the identification of \( J/\psi \) mesons with high transverse momenta is particularly challenging due to possible removal of one of the two muon decay products by the \( \mu \text{GMT} \) duplicate detection system. A study examining various tuning parameters for this system was presented. The probability for the production of a duplicate from a given muon track was derived from a Monte Carlo (MC) generated single muon event sample and determined to be about 20% at the inputs when drawn as a function of transverse momentum. This can be reduced effectively by the \( \mu \text{GMT} \) to a value of under 0.2%, depending on the tuning parameters chosen. This study was validated by computing double muon rates as well as the contribution to the rate by single muons that produced a duplicate particle for various transverse momentum thresholds. While the ghost contribution for the input rate is about 70% of the total double muon rate, this drops to 20% after application of the cancel-out algorithm. Comparing the various tuning parameters to each other, a slight advantage of the aggressive tuning is seen as it shows similar ghost rate to the baseline tuning while producing a moderately larger double muon rate, indicating higher efficiency.

Finally, the efficiency computed both from MC generated events as well as from a data sample taken in 2016 for a double muon trigger at the 18 GeV and 5 GeV threshold was presented. The efficiency is first computed from a MC generated sample of \( J/\psi \rightarrow \mu \mu \) events and then from a combination of data taken during 2016 and the MC generated sample. The double muon trigger efficiency is determined to be close to 80% in the plateau. Comparing the various cancel-out parameters in the MC-based study, a small advantage can again be seen for the aggressive tuning, particularly for close-by muons. The double muon efficiency determined partly from
data is shown to be compatible with the value derived exclusively from MC generated events.

The \( \mu \text{GMT} \) is furthermore shown to be superior to the legacy system both in duplicate suppression and efficiency, except for the pseudorapidity region of \( 1 < |\eta| < 1.2 \) where the legacy system’s efficiency is higher because of missing hit information sharing in the Overlap Muon Track Finder (OMTF), as well as the pseudorapidity region of \( 0.7 < |\eta| < 0.95 \) where only the conservative tuning shows comparable ghost rejection power to the legacy Global Muon Trigger (GMT). Sharing of hit information between OMTF processors is foreseen to be added in the future.

Due to the advantages shown using the aggressive tuning, it has been proposed as default tuning for the 2017 data taking run.

A novel feature introduced for the upgraded Level-1 trigger is the computation of di-object invariant mass at the \( \mu \text{GT} \). The effect of the \( \mu \text{GMT} \) extrapolation unit on the invariant mass requirement for double muon trigger algorithms was investigated.

It has been shown that a requirement on the invariant mass can significantly reduce the trigger rate for double muon trigger algorithms with and without transverse momentum thresholds while retaining high detection efficiency. Furthermore, an upgrade in the \( \mu \text{GMT} \) extrapolation unit strongly improves the invariant mass resolution for the J/\( \psi \) meson and allows the invariant mass threshold to be reduced by almost a factor two without sacrificing efficiency. This translates into an additional reduction in rate of more than 10%.

Double muon trigger algorithms utilising the invariant mass restriction were included in the data taking trigger menu for the 2017 run and have enabled the use of lower transverse momentum thresholds while still controlling the trigger rate.

The presented system was commissioned in the winter shutdown of 2015/2016 and has been integrated into the data taking operation of the CMS experiment since May 2016. The transmission of extrapolated azimuthal coordinates was introduced for the 2017 data taking year together with the additional transmission of the extrapolated pseudorapidity. The isolation algorithm is currently under study and may be added at a later date. Likewise, the use of track addresses provided by the Endcap Muon Track Finder (EMTF) for duplicate removal is subject to the outcome of ongoing studies. Finally, the cancel-out system will require re-tuning after improvements and corrections by the track finder groups.

At the time of writing the CMS experiment has used the upgraded Level-1 trigger to successfully record more than 70 fb\(^{-1}\) of proton-proton collisions. The \( \mu \text{GMT} \) has been involved in the readout decision for all events that included muons. This represents about half of the total Level-1 trigger accept rate and corresponds to 90 of 225 activated trigger algorithms for a recent data taking configuration. Comparisons of the recorded data to results obtained from software emulation as well as detailed studies of the system’s behaviour showed excellent performance consistent with expectations.
Every work of the scale of a thesis owes its existence to a multitude of people. I would like to first thank my supervisors Prof. Claudia Wulz and Prof. Eberhard Widmann who helped this thesis come to fruition with their good advice and support. I would like to additionally thank Claudia for providing me the opportunity to visit CERN and contribute to the Level-1 trigger very early in my scientific career, piquing my interest for this fascinating topic.

Furthermore I would like to thank Hannes Sakulin, my supervisor at CERN, for all his help. Hannes was the developer of the original Global Muon Trigger and as such patiently and knowledgeably explained design choices and helped me avoid common pitfalls during the μGMT development. He also took care to support me in my professional development by sending me to training courses and conferences.

Manfred Jeitler deserves great thanks for giving me the initial impetus to come to CERN and for the many lunches with stimulating discussions, in particular when I had just arrived. Thanks go also to the entire HEPHY Vienna Trigger group for help understanding VHDL and support with the Global Trigger during my first on-call shifts. Ivan Mikulec helped me to choose the topic of this thesis and encouraged me along the way for which I am very grateful.

Work on the μGMT was shared initially with Joschka Lingemann who took on the software-side of things and made long days at the office seem shorter with his good humour and quick wit. I also owe him significantly for his help understanding the CMS Software framework and all the useful advice on how to best tackle the double muon efficiency study presented in this work. Thomas Reis took over from Joschka and didn’t miss a beat in continuing the high quality work. His eye for detail and strong work ethic make it a joy to share the maintenance duties for the μGMT with him.

The μGMT was built using the Imperial Master Processor 7 processing board. Alessandro Thea, Dave Newbold, and Greg Iles stood ready to answer any questions I might have and gave great support in case of problems. Alessandro and Dave also took interest in my career and gave me valuable tips navigating the scientific politics for which I am very thankful.

Both the CMS Level-1 trigger and Operations communities were a joy to work with and made coming to work each day a pleasure.

Thank you also to my family for the support and encouragement on my scientific journey so far. My mother for showing me by example that some things are worth working long nights and weekends for and both my father and brother for lighting my scientific curiosity early on in my life.

Finally, I would like to thank Erica for her companionship over the past three years. Our many discussions about various problems I encountered or thoughts I had led to neat solutions or at least helped me gain some distance for reflection. Thank you.
Acknowledgements

My stay at CERN was in part funded by the CERN Austrian Doctoral Student Program, as well as by HEPHY Vienna and the CERN CMS DAQ group for which I am grateful.
Bibliography


Bibliography


Bibliography


Bibliography


Bibliography


Bibliography


Glossary

μGMT

μGT
upgraded Global Trigger. i, iii, 2, 54, 56, 58, 60, 62, 65, 66, 69–72, 74, 76–79, 93, 95, 97–101, 111, 137, 145–147

8P8C
8 position 8 contact. 66

AD
Antiproton Decelerator. 11

ALICE
A Large Ion Collider Experiment. 15, 16

API
application programming interface. 67

ASIC
application-specific integrated circuit. 41, 42, 46

ATLAS
A Toroidal LHC Apparatus. 15–17

AWAKE
Proton Driven Plasma Wakefield Acceleration Experiment. 11

BDT
Boosted Decision Tree. 65

BMTF
Barrel Muon Track Finder. 60–64, 73, 74, 77, 80–84, 86–88, 90, 93, 118, 120, 121, 123–126
Glossary

BPTX
Beam Pickup Timing for the eXperiment. 38

BU
builder unit. 38, 39

Castor
Centauro and Strange Object Research. 28, 29

CERN
European Organisation for Nuclear Research. 9, 12

CMS
Compact Muon Solenoid. i, iii, 1, 2, 4, 7, 14–18, 21–29, 31, 33, 35, 36, 38, 39, 41, 51–54, 56, 61–63, 66, 67, 69, 72, 80, 82, 103, 105, 106, 112, 125, 126, 133, 145–147, 149, 150

CMSSW
CMS Software. 68, 111, 112, 149

CPPF
concentrator and preprocessor fan-out. 62, 64

CRC
cyclic redundancy check. 105

CSC
cathode strip chamber. 18, 22, 31, 33–35, 46–48, 51, 52, 60, 62–64, 80, 90, 92

CSCTF
Cathode Strip Chamber Track Finder. 48, 50, 51, 64, 115

CTP7
Calorimeter Trigger Processor 7. 56, 58, 59

CT-PPS
CMS-TOTEM Precision Proton Spectrometer. 18

CuOF
Copper to optical fibre mezzanine. 61, 62

DAQ
data acquisition. 17, 18, 21, 38, 39, 53, 56, 70, 102, 103, 107, 150
**Glossary**

**DQM**
- data-quality monitoring. 39

**DT**
- drift tube. 22, 31–33, 35, 42, 46–48, 51, 52, 60–63, 80

**DTLT**
- local trigger. 46, 48, 51

**DTTF**
- Drift Tube Track Finder. 46, 48, 49, 51, 61, 63, 87

**ECAL**
- electromagnetic calorimeter. 22, 26, 27, 44, 57, 58

**EMTF**
- Endcap Muon Track Finder. 56, 62, 64, 65, 72–74, 77, 80–85, 87, 90, 92, 93, 118, 120, 122, 124, 125, 147

**EVM**
- event manager. 38

**FEROL**
- Frontend Readout Optical Link. 102

**FinOR**
- final OR. 53, 66

**FPGA**
- field-programmable gate array. 41, 42, 46, 56, 60, 71, 84, 98, 105–108

**FU**
- filter unit. 38, 39

**GB**
- RPC ghost buster. 52

**GbE**
- Gigabit Ethernet. 38, 39

**GCT**
- Global Calorimeter Trigger. 42, 44, 45, 53
Glossary

GEM
Gas Electron Multiplier. 17, 18

GMT
Global Muon Trigger. i, iii, 42, 46–49, 51–53, 71, 93, 113, 147, 149

GP
Golden Pattern. 64

GT
Global Trigger. 41, 42, 44–47, 52, 53, 56, 59, 66, 113, 149

GTL
Global Trigger Logic. 53

HCAL
hadronic calorimeter. 17, 22, 26–28, 44, 56–58

HF
forward hadron calorimeter. 22, 28, 29, 44, 45, 58

HGC
High Granularity Calorimeter. 17

HL-LHC
High Luminosity Large Hadron Collider. 13, 15, 16

HLT
High-Level trigger. 1, 19, 38, 39, 53, 58, 117, 135, 145

HO
outer barrel part. 27

ISOLDE
Isotope mass Separator On-Line facility. 11

L1A
Level-1 Accept. 53, 102

LB
link board. 51, 61, 62, 64
Glossary

LCT
local charged-track. 46, 48, 50, 64, 65

LEIR
Low Energy Ion Ring. 11, 12

LEP
Large Electron–Positron Collider. 9, 10

LHC
Large Hadron Collider. i, iii, 1, 2, 9, 11–19, 21–23, 25, 26, 29, 33–39, 41, 42, 53, 54, 58, 64, 66, 67, 71–73, 75, 77, 78, 99, 100, 103, 105, 107, 111–113, 128, 137, 145, 146

LHCb
LHC beauty. 15, 18, 19

LHCf
LHC forward. 16

LINAC2
Linear accelerator 2. 11, 12

LINAC3
Linear accelerator 3. 11, 12

LS1
Long Shutdown 1. 13, 16–18, 25, 26, 31, 34, 35, 39, 54, 67

LS2
Long Shutdown 2. 13, 15–17, 19, 25

LS3
Long Shutdown 3. 13, 16, 17, 25

LSB
least significant bit. 72, 80, 98, 99, 105

LUT
look-up table. 51, 56, 63, 65, 78, 79, 84, 86, 97–99, 124, 137, 138

LVDS
low-voltage differential signalling. 66
Glossary

MC
Monte Carlo. 111, 112, 115, 116, 133, 134, 136–139, 141, 146, 147

MicroTCA
Micro Telecommunications Computing Architecture. 56, 58, 66, 71, 102

MIP
minimum-ionising particle. 45, 52

MMT
Magnetic Monopole Trapper. 19

MoEDAL
Monopole and Exotics Detector At the LHC. 19

MP7
Imperial Master Processor 7. 56, 58, 59, 63, 66, 71, 75, 78, 98, 102, 105–107, 149

MPC
muon port card. 48, 50, 51, 64

MSB
most significant bit. 72, 75, 97

MTF7
Modular Track Finder Processor 7. 56, 63, 64

MTP
Multi-fibre Termination Push-on. 56, 71

NRQCD
non-relativistic QCD. 5

NTD
Nuclear Track Detector. 19

OMDS
Online Master Database system. 67

OMTF
Overlap Muon Track Finder. 56, 61–64, 73, 77, 80–85, 93, 116, 120, 122–126, 131, 147
Glossary

PACT
Pattern Comparator Trigger. 47, 51

PCC
Pixel Cluster Counting. 103, 104, 146

PS
Proton Synchrotron. 11, 12

PSB
Proton Synchrotron Booster. 11, 12

QCD
quantum chromodynamics. 3–5, 145

RAM
random-access memory. 56

RCMS
central run-control. 39, 67

RCT
Regional Calorimeter Trigger. 44, 45

RPC
resistive plate chamber. 17, 22, 31–33, 35, 36, 46, 47, 51, 52, 60–64, 90, 113

RU
read-out unit. 38, 39, 107

SC
sector collector. 46, 48, 49, 60, 61

SL
superlayer. 32, 33

SOAP
Simple Object Access Protocol. 67

SP
Sector Processor. 48–51
Glossary

**SPS**
Super Proton Synchrotron. 11, 12

**SR**
Sector Receiver. 50, 51

**SRAM**
static random-access memory. 51

**SVN**
Apache Subversion. 106

**SWATCH**
SoftWare for Automating the conTrol of Common Hardware. 68, 105

**TCDS**
Trigger Control and Distribution System. 56, 66, 67, 102, 103

**TCS**
Trigger Control System. 52, 53, 66

**TMB**
trigger motherboard. 48

**TMT**
time-multiplexed trigger. 57

**TOTEM**
TOTal cross section, Elastic scattering and diffraction dissociation Measurement at the LHC. 18, 28

**TPC**
Time Projection Chamber. 15

**TPG**
Trigger Primitives Generator. 42, 44, 45

**TRT**
Transition Radiation Tracker. 16

**TTC**
Timing, Trigger, and Control. 53, 54, 66, 159
Glossary

**TTCci**
Timing, Trigger, and Control (TTC) Versa Module Europa (VME) interface. 54

**TTCmi**
TTC machine interface. 54

**TTS**
Trigger Throttling System. 66

**VELO**
Vertex Locator. 18, 19

**VHDCI**
very-high-density cable interconnect. 66

**VHDL**
Very High Speed Integrated Circuit Hardware Description Language. 75, 107, 149

**VME**
Versa Module Europa. 42, 52, 54, 56, 159

**ZDC**
zero degree calorimeter. 29, 30
Dinyar Rabady

Personal data
Date of birth  November 30, 1986.
Place of birth  Waidhofen an der Thaya, Austria.
Citizenship  Austrian.

Education
2012 – 2018  University of Vienna, Doctoral study of physics.
(expected) PhD thesis: Upgrade of the CMS Global Muon Trigger
Thesis advisors: Claudia-Elisabeth Wulz, Eberhard Widmann, Hannes Sakulin
2006 – 2012  University of Vienna, Study of physics, graduated with distinction.
Master’s thesis: A hardware abstraction layer for the MicroTCA-based Global Trigger for the CMS experiment at CERN
Thesis advisors: Claudia-Elisabeth Wulz, Ivan Mikulec

Work experience
- Implemented system in CMS Run Control to warn operator in case of non-standard configuration choices.
- Work on the Level-1 trigger Phase-II upgrade. Investigated use of Vivado HLS and worked on build system to make collaboration developing firmware easier.
- Round-the-clock on-call expert for the Global Trigger and Global Muon Trigger systems. On-call shifts for the CMS Data Acquisition system.
Design and implementation of the upgrade for the Global Muon Trigger for the CMS experiment.
- Conducted part of performance studies leading to final design of the upgrade by simulating the proposed system in a C++ model.
- Implemented firmware for system in VHDL, taking into account required modifications due to oft-changing requirements. Global Muon Trigger was first to be completed system within the muon trigger upgrade.
- Coordinated joint-tests with multiple subsystems and took leading role ensuring the system’s readiness for data taking during first collisions of 2016.
- Round-the-clock on-call expert for the Global Trigger and Global Muon Trigger systems. On-call shifts for the complete Level-1 trigger.

Work on a fraud detection and protection mail relay server.
- Implemented basic functionality such as VERP as well as specialized features such as providing means to stop and restart conversations between customers if automatic identification of fraud failed.
- Took responsibility for the system’s code quality, in particular working on coding guidelines and ensuring high test coverage.
- Deployment of mail relay server led to reduction in detected phishing attempts by 90%.


- Performed an analysis of the Global Muon Trigger Efficiency of the CMS Experiment using the Root framework. Tuned system to achieve increase of muon detection efficiency of 10% for certain detector regions while confining increase of false positives to <2%.
- Developed a serialized communications protocol in order to implement a hardware abstraction layer providing access to hardware registers via Ethernet. Implemented using a C++ server to expose the hardware registers on the embedded hardware and a python client which communicated with the server via TCP/IP. Serialization was done using Protocol Buffers.

July 5 – Sept. 6, 2010 **“HEPHY@CERN” Summer intern.**

Work on the software of the Level-1 trigger control system of the CMS experiment, shifts for the Level-1 trigger system. Attended the lectures offered by CERN.

Oct. 2008 – Feb. 2011 **Student tutor in the group Aerosol Physics and Environmental Physics, University of Vienna**, Vienna, Austria.

Student tutor in the physics laboratory exercises for biologists.


Help with administrative duties, support for refugees and different problems with IT.

July 2004 **Intern in IT department, Pollmann Austria OHG**, 3822 Karlstein.

- Wrote small application in C# to notify a group of employees if the content of some directory changed.
- Wrote a web interface using ASP.Net and C# to access the employee database via the intranet.

**Awards and appointments**


“For his significant contributions to the upgrade Global Muon Trigger and support to the upgrade Global Trigger”

** Talks and posters**

Aug. 3 – Aug. 10, 2016 **Upgrade of the CMS muon trigger in the barrel region**, *ICHEP 2016*, Chicago, USA.

Poster describing the upgrade of the Level-1 trigger in the barrel region as well as performance studies using data taken with the upgraded trigger.

Feb. 15 – Feb. 19, 2016 **Upgrade of the CMS muon trigger system in the barrel region**, *VCI 16*, Vienna, Austria.

Poster on the Level-1 muon trigger upgrade in the barrel, in particular the Barrel Muon Track Finder and Global Muon Trigger.

Sept. 23 – Sept. 27, 2013 **The CMS Global Muon Trigger upgrade**, *TWEPP ‘13*, Perugia, Italy.

Poster on the status of the upgrade of the Global Muon Trigger for the CMS experiment at the LHC.
June 17, 2011  The upgrade of the CMS Trigger Electronics for SLHC, Joint Annual Meeting of the Swiss Physical Society and Austrian Physical Society, Lausanne, Switzerland.
Contributed Talk on behalf of the CMS collaboration.
Overview of the then-current plans for the upgrade of the Level-1 trigger of the CMS experiment with a focus on the Global Trigger upgrade.

Publications (selection)


J. Lingemann, D. Rabady, H. Sakulin, M. Jeitler, A. Stahl, Upgrade of the CMS Global Muon Trigger, Real Time Conference (RT), 2014 19th IEEE-NPSS.

M. Jeitler, J. Lingemann, D. Rabady, B. Rahbaran, H. Sakulin, Upgrade of the Global Muon Trigger for the CMS experiment, JINST 8 C12017.


Full list of publications: http://inspirehep.net/author/profile/D. Rabady.

Schools
CERN School of Computing 2014, Braga, Portugal,
Lessons and exercises on data storage and management, concurrency, and data analysis methods.

International School of Trigger and Data Acquisition 2011, Rome, Italy,
Lectures and exercises on both software and hardware used in the trigger and data acquisition systems of particle physics experiments.

Languages
German  Mother tongue
English  Mother tongue
French  Basics

Computing skills
Programming  Python, C++, Java  Programs  Xilinx Vivado, Mentor Modelsim, Root, ...
Data storage and exchange  SQL, XML, Protocol Buffers  Misc.  µTCA, VHDL