Universität Karlsruhe (TH)
Forschungsuniversität • gegründet 1825

Studien zur Suche nach schweren in Top-Quark-Paare zerfallenden Teilchen mit dem CMS-Detektor

Jochen Ott

Diplomarbeit

An der Fakultät für Physik
der Universität Karlsruhe (TH)

Referent: Prof. Dr. Thomas Müller
Korreferent: Prof. Dr. Günter Quast

26. Juni 2009

Zusammenfassung

Die theoretischen Grundlagen des bis heute mit großer Präzision bestätigten *Standard-Modells der Teilchenphysik* wurden bereits in den sechziger und frühen siebziger Jahren des letzten Jahrhunderts gelegt. Ein großer Erfolg war unter anderem die Vorhersage einer dritten Quark-Generation auf Grundlage von beobachteten, CP-verletzenden Zerfällen von Kaonen durch Kobayashi und Maskawa [1].

Während das leichtere Quark dieser dritten Generation, das b-Quark, bereits wenige Jahre später nachgewiesen werden konnte [2], gestaltete sich die Suche nach dem Top-Quark schwieriger. Es wurde erst nach jahrzehntelanger Suche 1995 an den Detektoren CDF und DØ am Tevatron-Beschleuniger entdeckt [3, 4], einem Proton-Antiproton-Beschleuniger mit einer Schwerpunktsenergie von damals 1.8 TeV.

Aufgrund seiner hohen Masse von über $170\,\mathrm{GeV}/c^2$ ist die Rate von Top-Quark-Ereignissen am Tevatron gering, und eine genaue Untersuchung aller Eigenschaften ist schwierig. Während die Messung der Top-Quark-Masse über viele Jahre verbessert wurde und mit einer inzwischen erreichten relativen Unsicherheit von weniger als 1% [5] als Präzisionsmessung gilt, konnte die vom Standard-Modell vorhergesagte Produktion einzelner Top-Quarks erst vor wenigen Monaten mit ausreichend hoher statistischer Signifikanz nachgewiesen werden, um von einer „Entdeckung“ zu sprechen [6].

Die hohe Masse des Top-Quarks hat Anlass zur Spekulation gegeben, ob es womöglich eine besondere Rolle bei der elektroschwachen Symmetriebrechung spielt. Einige derartige Theorien sagen schwere Eichbosonen voraus, die bevorzugt in Top-Quark-Paare zerfallen, etwa das Z'-Teilchen im Topcolor-Modell [7]. Es gibt jedoch auch andere Theorien jenseits des Standard-Modells, die zusätzliche Produktionsmechanismen für $t\bar{t}$-Ereignisse im Bereich hoher invarianter $t\bar{t}$-Massen vorhersagen, etwa Modelle mit Axigluonen [8].

Die Experimente am Large Hadron Collider am CERN eröffnen die Möglichkeit, derartige Theorien für einen großen, bisher nicht zugänglichen Parameterbereich zu testen.

Diese Arbeit beschäftigt sich mit der Suche nach schweren Resonanzen am CMS-Detektor (Compact Muon Solenoid), die in Top-Quark-Paare zerfallen. In solch einem Fall tragen die Top-Quarks einen hohen Impuls und deren Zerfallsprodukte werden im Detektor oftmals nicht als getrennte Objekte rekonstruiert. Die üblicherweise verwendeten Techniken zur Selektion und Rekonstruktion von $t\bar{t}$-Ereignissen setzen

dies jedoch voraus. Daher ist die Entwicklung einer neuen Ereignis-Selektion und -Rekonstruktion für derartige Ereignisse erforderlich.

Zwar gibt es bereits einige sehr interessante Studien zur Selektion und Rekonstruktion solch hochenergetischer Top-Quarks [9, 10], jedoch sind dies fast ausschließlich Studien auf Generator-Ebene ohne realistische Detektorsimulation oder Berücksichtigung von Untergrundprozessen.

Das Ziel der vorliegenden Arbeit ist es, diese Lücke zu schließen. Dazu wird der „μ+Jets"-Zerfallskanal von Top-Quark-Paaren – d. h. der Fall, dass eines der Top-Quarks in ein Myon, ein Neutrino und ein b-Quark und das andere rein hadronisch zerfällt – für hohe invariante Massen des $t\bar{t}$-Systems untersucht.

Das Szenario, das in dieser Arbeit genauer untersucht wird, entspricht mit einer integrierten Luminosität von $200\,\mathrm{pb}^{-1}$ bei einer Schwerpunktsenergie von 10 TeV etwa dem, was man nach dem ersten Jahr Datennahme erwarten kann. Da nicht ganz klar ist, mit welcher Genauigkeit der Detektor und die Simulation verstanden sein werden, wurden möglichst einfache Methoden entwickelt.

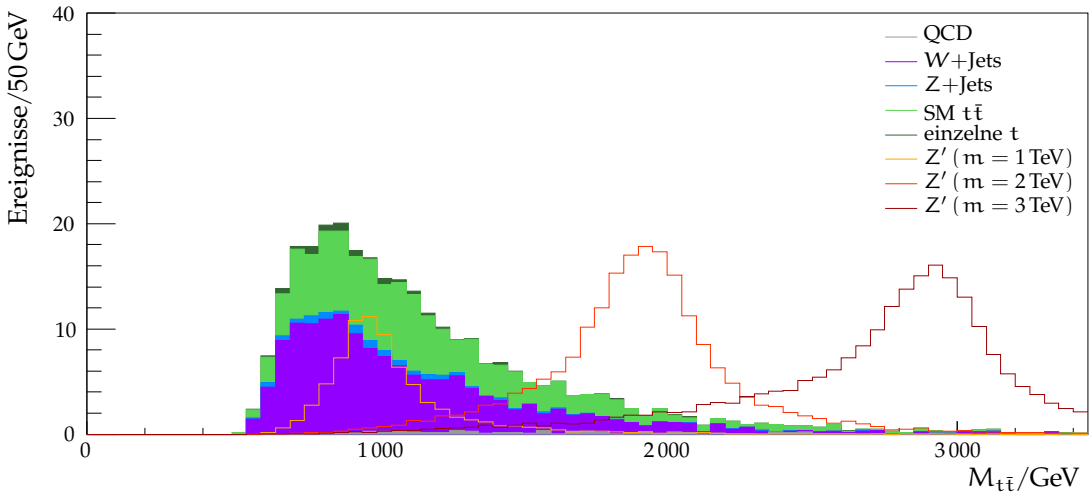

Abbildung 1: Spektrum der rekonstruierten Masse des Top-Quark-Paar-Systems, nach der Ereignis-Selektion. Die eingezeichneten Resonanzen Z' haben Massen von 1, 2 und $3\,\mathrm{TeV}/c^2$ und eine natürliche Breite von jeweils 1 % der Masse. Die angenommenen Wirkungsquerschnitte, $\sigma(pp \to Z' \to t\bar{t}) = 10\,\mathrm{pb}$, sind rein illustrativ, sie entsprechen keinem bestimmten Modell.

Die Ereignis-Selektion basiert auf Schnitten einiger weniger kinematischer Variablen. Dabei wird sowohl die hohe Energie der Top-Quarks als auch die hohe Masse desselben ausgenutzt. Ersteres hat zur Folge, dass die rekonstruierten Jets eine hohe Energie besitzen, Letzteres schlägt sich in einem vergleichsweise hohen

Transversalimpuls des Myons in Bezug auf das b-Quark nieder. Die neu entwickelte Ereignis-Selektion ist dabei für hohe $t\bar{t}$-Massen um ein Vielfaches besser als die Anwendung einer $t\bar{t}$-Ereignis-Selektion wie sie für Standard-Modell-$t\bar{t}$-Studien verwendet wird, die vor allem für relativ kleine $t\bar{t}$-Massen geeignet ist.

Mit *Rekonstruktion* eines $t\bar{t}$-Ereignisses ist in diesem Zusammenhang die Bestimmung der jeweiligen Top-Quark-Viererimpulse und damit auch die Bestimmung der invarianten Masse des $t\bar{t}$-Systems gemeint. Dazu wird die bereits erwähnte Eigenschaft von Zerfällen hochenergetischer Top-Quarks, nämlich dass die Zerfallsprodukte einen kleinen Winkelabstand im Detektor haben, ausgenutzt.

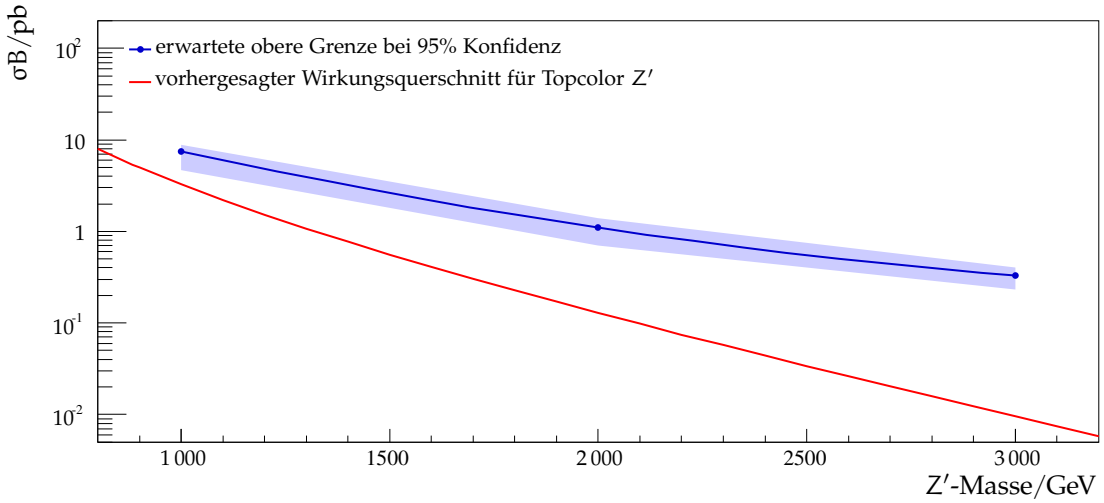

Abbildung 2: Die erwarteten oberen Grenzen für $\sigma(pp \to Z') \times BR(Z' \to t\bar{t})$ für ein 95%-Konfidenzniveau. einer integrierten Luminosität von $L = 200\,pb^{-1}$ und einer Proton-Proton-Schwerpunktsenergie von 10 TeV.

Nach der Rekonstruktion und Ereignis-Selektion erwartet man eine Verteilung der rekonstruierten Masse wie in Abbildung 1 zu sehen. Erwartete Untergründe sind ausgefüllt, während mögliche schmale Resonanzen Z' für verschiedene Massen als Linie eingezeichnet sind. Die verwendeten Wirkungsquerschnitte von 10 pb sind willkürlich gewählt und entsprechen keinem bestimmten Modell.

Um die Ergebnisse der Studien zur Selektion und Rekonstruktion zu quantifizieren, wurden obere Grenzen für den Wirkungsquerschnitt von schmalen Resonanzen Z' bestimmt. Um diese Grenzen zu bestimmen, wird mit einem Daten-Modell und dem Bayestheorem eine Posterior-Verteilung für die Parameter des Modells bestimmt. Diese Parameter enthalten unter anderem den Wirkungsquerschnitt für Z'. Das 95%-Quantil der Marginalverteilung des Posteriors in dem Parameter ist dann die gesuchte obere Grenze. Die Bestimmung dieses Quantils erfordert eine numerische

Integration in einem hochdimensionalen Raum. Dazu wurde ein effizientes Verfahren implementiert, das auf Markov-Ketten beruht.

Das Daten-Modell beruht in weiten Teilen auf Monte-Carlo-Simulation. Um von einer verzerrten Modellierung unabhängiger zu sein, wurde für den besonders anfälligen QCD-Untergrund ein Verfahren entwickelt, das weitgehend datenbasiert ist. Außerdem wurden viele systematische Unsicherheiten, etwa die theoretischen Unsicherheiten der Wirkungsquerschnitte, der Hadronisierung und der Jet-Energie-Rekonstruktion berücksichtigt. Die erwarteten oberen Grenzen, die mit den Markov-Ketten unter Berücksichtigung dieser systematischen Unsicherheiten bestimmt wurden, sind in Abbildung 2 zu sehen. Die erwarteten Grenzen liegen in der Größenordnung einiger Picobarn, für hohe Massen auch darunter. Schon nach einem Jahr Datennahme ist es somit möglich, Modelle auszuschließen, die einen Zusatzbeitrag dieser Größenordnung zum $t\bar{t}$-Massenspektrum vorhersagen. Im Massenbereich von $1\,\mathrm{TeV}/c^2$ liegt die erwartete Ausschlussgrenze auch nur wenig über dem Wirkungsquerschnitt des Topcolor-Z′. Mit weiteren Verbesserungen der Ereignis-Selektion oder durch eine Kombination der Ergebnisse mit verwandten Analysen [11] oder der Einbeziehung des Elektron-Kanals kommt ein Ausschluss des Topcolor-Z′ in diesem Massenbereich in Reichweite.

Axigluonen hingegen besitzen einen höheren Wirkungsquerschnitt, und eine Verbesserung der gegenwärtigen indirekten Massengrenze von $0.9\,\mathrm{TeV}/c^2$ [12] bei 95 % Konfidenz durch direkte Suche erscheint bereits nach dem ersten Jahr Datennahme möglich.

Universität Karlsruhe (TH)
Research University • founded 1825

Studies for the Search of Heavy Particles Decaying in Top Quark Pairs with the CMS Detector

Jochen Ott

Diploma Thesis

At the Department of Physics
of the University of Karlsruhe (TH)

Referee: Prof. Dr. Thomas Müller
Co-Referee: Prof. Dr. Günter Quast

June 26th, 2009

Contents

List of Tables

List of Figures

Introduction

The top quark, discovered 1995 at the Tevatron [3, 4], has the largest mass of all known fundamental particles. With a mass close to the scale of electroweak symmetry breaking (EWSB), it plays a special role in many alternative EWSB theories. Some models predict a heavy resonance which can decay to top quark pairs. If the resonance has an invariant mass of the order of 1 TeV or above, the top quarks from its decay will have large momenta and the decay products of the top quarks will be close in the detector. In this case, the selection and reconstruction applied in standard model $t\bar{t}$ analyses, which assume large angular separation of the decay products, will not yield optimal results.

Therefore, new techniques have to be developed to select and reconstruct $t\bar{t}$ events with high $m_{t\bar{t}}$. While there are interesting generator studies for this problem [9, 10], there are only very few studies assuming a realistic detector and realistic backgrounds.

The aim of this thesis is to fill this gap. Reconstruction and selection techniques are developed for the μ+jets channel, where one top quark decays to a muon, a neutrino and a b quark and the other top quark decays to three quarks.

The studies are based on simulations of the CMS detector (Compact Muon Solenoid) at the Large Hadron Collider (LHC). The assumed integrated luminosity of $L = 200\,\mathrm{pb}^{-1}$ and the center-of-mass energy of 10 TeV correspond approximately (to current knowledge) to the first year of data taking. As it is not clear how well the detector and the Monte-Carlo simulations will be understood at that time, the methods for selection and reconstruction are kept at a simple level.

The event selection is based on cuts on few kinematic variables. The chosen variables reflect the properties of a high-energy $t\bar{t}$ event in which the decay products are not separated in the detector but might overlap. The transverse momenta of the decay products are typically large as well as their relative transverse momenta.

Similar properties are exploited in the event reconstruction which aims to estimate the four-momenta of the two top quarks.

After selection and reconstruction, the spectrum of the invariant mass of the $t\bar{t}$ system is analyzed for narrow resonances Z'. The limits are derived with a frequentist and a purely Bayesian method. Systematic uncertainties are accounted for with a Bayesian approach in both cases. For the Bayesian analysis, an efficient Markov chain Monte-Carlo integration was implemented. In order to be more independent from

Monte-Carlo simulations for the challenging QCD background, a data-based QCD model was developed.

The thesis is structured as follows. The first chapter briefly reviews the theoretical framework and introduces models with heavy resonances decaying in a top quark pair. The second chapter gives an overview of the LHC and the CMS detector. In the third chapter, the reconstruction of physics object and $t\bar{t}$ events is discussed. The fourth chapter shows that new selection methods are required for high mass resonances, possible cut variables and results of a cut optimization are discussed. In the last chapter, it is shown how selected and reconstructed events can be used to determine upper limits on the cross section of a narrow resonance Z' and how the QCD model will be extracted from data.

1 Theory Overview

The standard model of particle physics is the most accurate and complete model describing elementary particles and their interactions. The formal framework of this theory, the particles and interactions are reviewed in this chapter. Also, processes important for the present studies, namely top quark pair production and decay as well as $t\bar{t}$ production beyond standard model processes, are discussed.

The overview presented in this chapter is not a general introduction to quantum field theory or the standard model. Rather, it tries to review some of the important concepts for readers already familiar with the subject. This implies that there is a lack of completeness of the presented arguments. For a more complete introduction, there are many textbooks available, for example [13] and [14].

Before going into details, a short remark on the relevance of the standard model is in order. The necessary formalism and the standard model were already developed in the sixties and seventies of the last century, at a time when the experimentally accessible energies were some orders of magnitude below the ones directly available today. A very remarkable milestone in the development of the standard model was the prediction of a third generation of quarks by Kobayashi and Maskawa [1]. Their prediction was based on the observation of CP violation in rare kaon decays. It can be seen as a major success for the theory to make a prediction of unknown particles based on rare decays because this requires a deep understanding and correct application of the underlying principles of the theory in order to connect such seemingly unrelated topics.

Since then, a huge number of experiments were conducted in order to either measure its parameters to high precision or to show the existence of physics *beyond* the standard model. Until today, there is no convincing measurement which proves the standard model wrong.[1] While this is a great success for the standard model, it is also unsatisfactory because any physics beyond the standard model could point to a more complete model. The standard model itself cannot be the last word because it does not account for gravitation, Dark Matter or CP violation strong enough to explain the matter-antimatter imbalance in the universe.

[1] Apart from neutrino masses. But those can be incorporated into the standard model without challenging the validity of the rest of the theory.

1.1 Theoretical Framework: Quantum Field Theory

The standard model can be seen as an instance of a broad range of theories which are all formulated in the language of a quantized field theory and are based on the same principles:

- A quantum field theory is completely described by its Lagrangian density. Writing the theory as a Lagrangian *density* instead of a total Lagrangian ensures the *locality* of the theory.
- The Langrangian must be invariant under Lorentz transformations and translations (Poincaré invariance).
- The theory has to fulfil *microcausality* (which is closely linked to locality) which states that physical observables at space-like separated points must commute, i. e. physical quantities at space-like separated points must be independent.
- The spectrum of the Hamiltonian (i. e. the energy) has to be bounded from below.

Another principle which is not as basic as the ones just mentioned is the one of *renormalizability*. Demanding renormalizability means that the theory should be valid up to a given scale Λ and corrections for physics at an energy scale E satisfying $E \ll \Lambda$ should be of order E/Λ.

A further principle is the *anything goes* principle: if a term can be added to the Lagrangian without introducing more degrees of freedom and respecting all other principles (including renormalizability), it must be added to the Lagrangian.

1.1.1 Action Principle

As in classical physics, the principle of least action can be used. It states that a system described by the Lagrangian L evolves from its initial to its final state such that the action S, defined as

$$S = \int L(t)dt$$

is minimal. As in classical physics, differential equations for the fields follow from that requirement (the *Euler-Lagrange equations* or "equations of motion").

Therefore, the fundamental entity of any quantum field theory is the Lagrangian density $\mathcal{L}$ with

$$L = \int d^4x \, \mathcal{L}.$$

1.1.2 *Canonical Form* of the Lagrangian

The most general possible Lagrangian for a theory with N real Lorentz-scalar fields which satisfies all principles stated above is

$$\mathcal{L} = -\frac{1}{2}A_{ij}\partial_\mu\phi^i\partial^\mu\phi^j - \frac{1}{2}B_{ij}\phi^i\phi^j - C,$$

where the matrices A and B are hermitian and positive definite and C is not observable and can be set to zero.

A and B can be diagonalized by redefining the fields ϕ_i using a linear transformation, yielding the fields φ_i. In *canonical form*, the most general Lagrangian is

$$\mathcal{L} = -\frac{1}{2}\partial_\mu\varphi^i\partial^\mu\varphi^i - \frac{1}{2}m_i^2\left(\varphi^i\right)^2$$

where the interpretation of m_i as mass of the particle φ^i can be checked by the equations of motion yielding the usual relativistic energy-momentum relation (note that $\hbar = c = 1$ is used throughout).

1.1.3 Fermion Fields

The Lie algebra of the Lorentz group consists of three generators for rotation, J_i, and three generators for boost, K_i, satisfying the commutation relations

$$\begin{aligned}[J_i, J_j] &= i\epsilon_{ijk}J_k\\ [J_i, K_j] &= i\epsilon_{ijk}K_k\\ [K_i, K_j] &= -i\epsilon_{ijk}J_k.\end{aligned}$$

It is locally isomorphic to $SU(2) \times SU(2)$, which can be seen by introducing the operators

$$L_i = \frac{1}{2}(J_i + iK_i), \qquad R_i = \frac{1}{2}(J_i - iK_i) \tag{1.1}$$

satisfying

$$\begin{aligned}[L_i, L_j] &= i\epsilon_{ijk}L_k\\ [R_i, R_j] &= i\epsilon_{ijk}R_k\\ [L_i, R_j] &= 0.\end{aligned}$$

The representation used for transforming fields under the Lorentz group used so far was $(0, 0)$.

Studying other possible irreducible representations of the Lorentz group leads to other types of particles. Particles transforming as $(\frac{1}{2}, \frac{1}{2})$ are vector particles. The only

other two representations needed for theories satisfying the principles stated above are the *spinor representations*.

The most simple non-trivial matrix representation satisfying the commutation relation of $SU(2)$ are the matrices $\sigma_i/2$ where the σ_i are known as Pauli matrices.

Using $\sigma_i/2$ as representation for the L_i and 0 as representation for the R_i (in equation 1.1), those operators act on two fields typically written as two components of a single field called *left-handed Weyl spinor* Ψ_L. Switching representations (i. e. 0 for the L_i and $\sigma_i/2$ for the R_i) yields a *right-handed Weyl spinor* Ψ_R.

As the Lagrangian density has to be a Lorentz-scalar, fields have to be multiplied such that the combination transforms as $(0,0)$.

In the formulation of the standard model, the *Majorana notation* for spinors is used. It is a single, four component field Ψ_M defined as

$$\Psi_M = \begin{pmatrix} \Psi_L \\ i\sigma_2\Psi_L \end{pmatrix}.$$

The two lower components $i\sigma_2\Psi_L$ transform like a right-handed Weyl spinor.

1.1.4 Interactions

New particle fields can be added to the Lagrangian density, as long as the principles above are satisfied. However, it is useful to have a recipe which automatically ensures that they are satisfied. The recipe of choice is using local gauge invariance which is presented here. It has the advantage that relatively little input is required. On the other side, it is not the most general approach.

Using local gauge invariance to introduce interactions, a new local gauge symmetry group is postulated and interacting particles are arranged in multiplets transforming under non-trivial representations of this group. To satisfy the gauge symmetry, one massless field is required for each generator of the group.

Example for Local Gauge Invariance: QED As an example, take the Lagrangian for a spin $\frac{1}{2}$ particle Ψ,

$$\mathcal{L} = \bar{\Psi}(i\gamma^\mu\partial_\mu - m)\Psi \tag{1.2}$$

and postulate local $U(1)$ invariance. For a local $U(1)$ transformation with the local phase $-q\lambda(x)$, Ψ should transform non-trivially as

$$\Psi(x) \to e^{-iq\lambda(x)}\Psi(x).$$

If using that in equation 1.2 $\mathcal{L}$ is not invariant; rather

$$\mathcal{L} \to \mathcal{L} + q(\partial_\mu\lambda)\bar{\Psi}\gamma^\mu\Psi.$$

Because local gauge invariance must hold (as postulated), everything must be done to ensure invariance of $\mathcal{L}$. This can be achieved by introducing a new vector field A_μ transforming as

$$A_\mu \to A_\mu + \partial_\mu \lambda$$

and adding the term

$$-(q\bar{\Psi}\gamma^\mu\Psi)A_\mu$$

to the Lagrangian.

The total Lagrangian

$$\mathcal{L} = \bar{\Psi}(i\gamma^\mu(\partial_\mu - iqA_\mu) - m)\Psi$$

is now invariant under local $U(1)$ transformations.

Because of the *anything goes* principle, a kinematic term for A_μ has to be added,

$$\mathcal{L}_{A_\mu\text{-kin}} = -\frac{1}{4}F_{\mu\nu}F^{\mu\nu} = -\frac{1}{4}(\partial_\mu A_\nu - \partial_\nu A_\mu)(\partial^\mu A^\nu - \partial^\nu A^\mu).$$

Problems with Local Gauge Invariance Note that no mass term for A_μ can be added without violating local gauge invariance. Thus, introducing interactions through local gauge invariance leads to massless gauge fields. Experimentally however, the interaction particles W and Z *have* mass. One has two choices:

1. Give up the local gauge invariance as construction principle for interactions (note that this principle does not appear in the list of principles given above).
2. Find a way to explain the masses of W and Z consistent with local gauge invariance.

The standard model is formulated using the second alternative using the *Higgs mechanism* which provides an explanation for gauge boson masses consistent with local gauge invariance (see section 1.2.1). If predictions made with this assumptions are true (for example, if the Higgs particle exists), that would be a strong indication that local gauge invariance actually is *more* than a convenient way to introduce interactions.

1.1.5 Feynman Graphs

At a hadron collider, the major question to theory are (differential) cross sections of a given initial state α and final state β. This question can be addressed by perturbative calculations.

For an initial state α and final state β, the cross section $d\sigma(\alpha \to \beta)$ is (by virtue of Fermi's Golden Rule):[2]

$$d\sigma(\alpha \to \beta) \propto |\mathcal{M}_{\alpha\beta}|^2 (2\pi)^4 \delta(p_\alpha - p_\beta) d\beta. \tag{1.3}$$

The delta function ensures energy and momentum conservation. The formula holds for a defined point in phase space for both initial and final states. To get the total cross section of a process, one has to integrate the formula over the final state phase space.

$\mathcal{M}_{\alpha\beta}$ is the matrix element of the initial and final state of the process and can be interpreted as the transition probability of a system from state α to state β. Therefore, it can be evaluated using the time evolution of the system. As the initial and final states are non-interacting (they are separated particles), time evolution can be evaluated from $-\infty$ to ∞. As in quantum mechanics, time evolution can be written as the time ordered exponential (or Dyson series):

$$\mathcal{M}_{\alpha\beta} = \langle \beta | \mathrm{T} \exp\left(-i \int d^4x \mathcal{H}(\S) \right) | \alpha \rangle,$$

where $\mathcal{H}$ is the Hamiltonian density and T exp denotes the formal exponential series where the terms $\mathcal{H}(x)$ are time-ordered.

Approximate evaluating this expression is done through expansion of the exponential, truncating the series at some (low) order of the coupling constants. As $\mathcal{H}$ consists of products of creation and annihilation operators, this expansion yields a sum of matrix elements of products of creation and annihilation operators.

Using Wick's theorem, one can rewrite the product as a sum of terms of creation and annihilation operators in normal order. Those individual terms can be visualized by *Feynman graphs*. The *Feynman rules* determine the translation between a Feynman graph and the corresponding term of the matrix element.

1.2 The Standard Model of Particle Physics

The particles of the standard model can be divided into gauge bosons of the interactions and the "matter" part (fermions). The Higgs particle fits neither of those and plays a special role.

The matter particles interact via the electromagnetic, weak and strong interactions. The corresponding gauge groups are $U_Y(1)$, $SU_L(2)$ and $SU_C(3)$ where the Y is the *weak hypercharge*, L indicates that only left-handed fermions carry this charge and C stands for *color*.

[2] The factor neglected depends on the relative velocities of the incident particles.

name	symbol	$SU(2)_L$	$SU_C(3)$	y(for $U_Y(1)$)
left-handed leptons	$P_L L_m$	**1**	**2**	$-\frac{1}{2}$
right-handed charged leptons	$P_R E_m$	**1**	**1**	-1
left-handed quarks	$P_L Q_m$	**3**	**1**	$+\frac{1}{6}$
right-handed up-type quarks	$P_R U_m$	**3**	**1**	$+\frac{2}{3}$
left-handed down-type quarks	$P_L D_m$	**3**	**1**	$-\frac{1}{3}$

Table 1.1: The particles of the standard model and the representation under the gauge groups. The dimension of the representations is given in bold face, for example **1** is the trivial representation, **3** the 3-dimensional representation. For the weak hypercharge y, the fields transform under $\lambda(x)$ like $\Psi(x) \to \exp(i\lambda(x)y)\Psi(x)$. More explanation in the main text.

Those interactions are associated to in total 12 gauge bosons: B_μ, 3 W^a_μ and 8 G^α_μ.

The matter fields of the standard model can be classified according to their interactions, i. e. the transformation representation for the three gauge groups. In table 1.1, the standard model particles are given in this classification.

There, the weak lepton duplet L_m consists of a left-handed neutrino Majorana spinor as first component and the left-handed charged lepton Majorana spinor as second component: $L_m = (\nu_m, \mathcal{E}_m)$. Accordingly, the weak quark duplet Q_m is $Q_m = (\mathcal{U}_m, \mathcal{D}_m)$ The index m denotes the generation running from 1 to 3. U_m, D_m and E_m are the *right-handed* up-type quarks, down-type quarks and charged leptons, respectively.

The most general Lagrangian with this particle contents and with those gauge groups is

$$\begin{aligned}\mathcal{L}_{fg} = &-\frac{1}{4}G^\alpha_{\mu\nu}G^{\alpha\mu\nu} - \frac{1}{4}W^a_{\mu\nu}W^{a\mu\nu} - \frac{1}{4}B_{\mu\nu}B^{\mu\nu} - \frac{g_3^2\Theta_3}{64\pi^2}\epsilon_{\mu\nu\lambda\rho}G^{\alpha\mu\nu}G^{\alpha\lambda\rho}\\ &-\frac{g_2^2\Theta_2}{64\pi^2}\epsilon_{\mu\nu\lambda\rho}W^{a\mu\nu}W^{a\lambda\rho} - \frac{g_1^2\Theta_1}{64\pi^2}\epsilon_{\mu\nu\lambda\rho}B^{\mu\nu}B^{\lambda\rho}\\ &-\frac{1}{2}\bar{L}_m\not{D}L_m - \frac{1}{2}\bar{E}_m\not{D}E_m - \frac{1}{2}\bar{Q}_m\not{D}Q_m\\ &-\frac{1}{2}\bar{U}_m\not{D}U_m - \frac{1}{2}\bar{D}_m\not{D}D_m\end{aligned}$$

in which the gauge field-strengths are given by

$$\begin{aligned}G^\alpha_{\mu\nu} &= \partial_\mu G^\alpha_\nu - \partial_\nu G^\alpha_\mu + g_3 f^\alpha_{\beta\gamma}G^\beta_\mu G^\gamma_\nu\\ W^a_{\mu\nu} &= \partial_\mu W^a_\nu - \partial_\nu W^a_\mu + g_2\epsilon_{abc}W^b_\mu W^c_\nu\\ B_{\mu\nu} &= \partial_\mu B_\nu - \partial_\nu B_\mu.\end{aligned}$$

The $f^{\alpha\beta\gamma}$ and ϵ_{abc} are the structure constants of the gauge groups $SU_C(3)$ and $SU_L(2)$. The terms involving these give rise to Feynman rules with vertices of 3 and 4 gauge bosons.

The coupling constants g_1, g_2 and g_3 are for the $U_Y(1)$, the $SU_L(2)$ and strong interactions.[3] The notation for the gauge-covariant derivatives $\not{D}$ is a bit sloppy: its actual meaning depends on the field they are applied to and follow from the gauge transformation properties listed in the table above. For example, for the weak quark duplet Q_m (which takes part in all three interactions):

$$D_\mu = \frac{1}{2}\partial_\mu + \left(-\frac{i}{2}g_3 G_\mu^\alpha \lambda_\alpha - \frac{i}{2}g_2 W_\mu^a \tau_a - \frac{i}{6}g_1 B_\mu\right) P_L + \text{h. c.}$$

The parameters Θ_1 and Θ_2 are physically not observable. So up to now, there are only four free parameters in the Lagrangian: g_1, g_2, g_3 and Θ_3.

While it is worth noting that this Lagrangian already includes all the matter and interaction particles, it is also important to note that there are *no* mass terms, neither for the gauge bosons nor for the fermions. They cannot be added directly because this would violate gauge invariance. How to deal with this shortcoming is explained in the next section.

1.2.1 Higgs Mechanism

The interaction fields introduced by the gauge principle are massless. However, experiments have shown $m_W \neq 0 \neq m_Z$. Simply adding a mass term to the standard model Lagrangian would not keep local gauge symmetry. So either one has to give up the principle of local gauge symmetry or a way has to be found to include masses without violating this principle. As the local gauge invariance has proven a very useful concept, the latter way is the one pursued in the standard model.

A set of possibilities to achieve that is the existence of a field interacting with the gauge bosons to give mass to. In the ground state of the system, this field shall have a non-zero expectation value.[4] In order to apply perturbation theory, fields have to be rewritten such that the ground state corresponds to the field configuration where all fields are vanishing. The coupling of Z to that field is then a quadratic term of Z in the Lagrangain, a mass term for Z.

[3] Through symmetry breaking, linear combinations of g_1 and g_2 appear as electromagnetic and weak coupling constants.

[4] If written manifestly gauge-invariant, i. e. the gauge group representation for the field is a linear transformation of the field multiplet.

This mechanism is called *spontaneous symmetry breaking*. Note, however, that the Lagrangian still has the full gauge symmetry. The only thing that is "special" about this case is that there is no parameterization of the fields such that both (1) the gauge symmetry transformations are linear transformations of field multiplets and (2) the ground state is the state in which all fields are vanishing.

The minimal way to implement this is to postulate a weak doublet of scalar complex fields, the Higgs doublet ϕ. Introducing the Higgs doublet adds four degrees of freedom, but three of them have already been observed as they correspond to the masses of W^+, W^- and Z.

The Higgs mechanism is minimal in the sense that it only adds one degree of freedom not yet observed. The non-zero vacuum expectation value for the Higgs field is achieved by the Higgs potential

$$\mathcal{L}_{\text{Higgs pot.}} = \lambda(\phi^\dagger\phi - \mu^2/2\lambda)^2.$$

The minimum of the field is not at zero but at $v = \mu^2/\lambda$.

As was already noted at the end of the last section, also fermion masses are not gauge invariant and are now explained by the Higgs mechanism. After rewriting fields about the vacuum expectation value v, only one real field, H remains from the Higgs doublet and the Higgs sector of the standard model is:

$$\begin{aligned}
\mathcal{L}_{\text{Higgs}} = &-\frac{1}{2}\partial_\mu H \partial^\mu H - \lambda v^2 H^2 - \lambda v H^3 - \frac{\lambda}{4}H^4 \\
&- \frac{1}{8}g_2^2(v+H)^2\left(W_\mu^1 - iW_\mu^2\right)^2 \\
&- \frac{1}{8}(v+H)^2(-g_2 W_\mu^3 + g_1 B_\mu)^2 \\
&- \frac{1}{\sqrt{2}}(v+H)(f_{mn}\bar{\mathcal{E}}_m P_R E_n + \text{h. c.}) \\
&- \frac{1}{\sqrt{2}}(v+H)(g_{mn}\bar{\mathcal{U}}_m P_R U_n + \text{h. c.}) \\
&- \frac{1}{\sqrt{2}}(v+H)(h_{mn}\bar{\mathcal{D}}_m P_R D_n + \text{h. c.}).
\end{aligned} \tag{1.4}$$

From this Lagrangian, the Higgs mass parameter can be determined as $m_H = 2\lambda v^2$.

Redefining the fields such that the mass matrices f_{mn}, g_{mn} and h_{mn} are diagonal, the fermion masses are given by

$$m = \frac{v}{\sqrt{2}}c$$

where c is the coupling to the Higgs (the diagonal entries after mass matrix diagonalization).

Diagonalizing the mass matrix for the W fields leads to the fields

$$W_\mu^\pm = \frac{1}{\sqrt{2}}(W_\mu^1 \mp iW_\mu^2)$$

with the mass

$$m_W = \frac{g_2^2 v^2}{4}.$$

As g_2 and m_W are known from experiment, the numerical value of v can be determined; it is $v = 246$ GeV.

It can further be shown that W_μ^3 and B mix to what is known as photon and Z described by the mixing angle Θ_W (Weinberg angle). The electromagnetic coupling constant (that is, the absolute value of the electron charge) is given by $e = g_1 \cos\Theta_W = g_2 \sin\Theta_W$.

The total standard model Lagrangian is

$$\mathcal{L}_{SM} = \mathcal{L}_{fg} + \mathcal{L}_{Higgs}.$$

Given the gauge symmetries and particle contents, $\mathcal{L}_{SM}$ is the most general Lagrangian density compatible with the principles formulated in section 1.1.

1.2.2 Particle Masses

The matrices appearing in the last three lines of equation 1.4 give rise to fermion masses. They can be diagonalized as mentioned in section 1.1.2. Given the Dirac spinors

$$e_m \equiv P_L \mathcal{E}_m + P_R E_m \tag{1.5}$$
$$u_m \equiv P_L \mathcal{D}_m + P_R D_m \tag{1.6}$$
$$d_m \equiv P_L \mathcal{U}_m + P_R U_m \tag{1.7}$$

the mass terms can be written as

$$\mathcal{L}_{\text{fermion-mass}} = -\frac{1}{\sqrt{2}} v (f_m \bar{e}_m e_m + g_m \bar{u}_m u_m + h_m \bar{d}_m d_m).$$

Note that there are no masses for neutrinos in the standard model. To account for experimental evidence which strongly points to neutrino masses, they can be included if one adds right-handed neutrinos. Without adding neutrino mass terms, the Higgs part has 15 parameters (ten masses, the Higgs self-coupling and four parameters for the Kobayashi-Maskawa matrix).

1.2.3 Kobayashi-Maskawa Matrix

The freedom to redefine fields was exploited in the last section to diagonalize the matrix of fermion-Higgs couplings (mass matrix). There is no reason for the fields in the weak quark duplets to be the same.[5] If redefining the fields such that the mass matrix is diagonal (this can be done redefining down-type quark fields only), the charged-current part coupling to quarks can be written as

$$\mathcal{L}_{cc} = i\frac{g_2}{2\sqrt{2}}\left(W^+_\mu V_{mn}\bar{u}_m\gamma^\mu(1+\gamma_5)d'_n + W^-_\mu (V^\dagger)_{mn}\bar{d}'_m\gamma^\mu(1+\gamma_5)u_n\right) \tag{1.8}$$

where m and n are generation indices and V_{mn} is the *Kobayashi-Maskawa* matrix (also called *Cabbibo-Kobayashi-Maskawa* matrix). This matrix is a unitary 3×3 matrix. Not all its parameters are physically significant because all but one of the complex phases can be absorbed in a redefinition of the quark field phases. This leaves four parameters: three angles and one complex phase.

Via processes where dominant matrix elements contain couplings between a W-boson and a quark, the matrix elements V_{mn} can be measured. Also, it is the complex phase of this matrix which is responsible for the CP violation in the weak sector.[6] The precise measurement of those matrix elements is another major experimental challenge for the standard model (for example, non-unitarity of V would be an indication of a fourth quark generation).

1.3 Full Events at Hadron Colliders

In the setting of a hadron collider, some more specific remarks are in order.

As the LHC is a proton-proton collider, one has to deal with the fact that neither the initial state particles, nor all final state particles are really *elementary particles* in the sense of the theory. Rather, as initial state particles, all quarks and gluons (called *partons* in this context) can appear as initial state particles of a matrix element.

To describe proton contents, one introduces *parton density functions* (pdf) for each parton type in the proton. A pdf for a particular parton is a function of a scale q^2 and the momentum fraction x: $a(q^2, x)$. x is the momentum fraction of the proton the parton carries, measured in the lab frame. At a fixed scale q^2, the pdf $a(q^2, x)$ can be interpreted as the probability density in x that a probe of virtuality q^2 finds a parton of momentum fraction x.

[5] For the leptons, this can be achieved by redefining neutrino fields only. This is possible for vanishing neutrino masses. If including neutrino masses, a corresponding mixing matrix for the lepton sector is needed.

[6] Since CP was discovered not to be an exact symmetry of nature, the smallness of the CP-violating parameter in the *strong* sector, Θ_3, is a puzzle.

The parton density functions are closely connected to the structure functions which naturally arise in the most general formulation of hadronic cross sections. While in principle, it should be possible to derive the pdfs *ab initio* from the standard model Lagrangian, there was so far no success in doing so because this includes intrinsically non-perturbative kinematic regions.

However, one can use perturbation theory to derive equations which connect the pdfs at two different scales: the *DGLAP* (Dokshitzer-Gribov-Lipatov-Altarelli-Parisi [15]) equations. Therefore, it is sufficient to measure the pdf at only one scale, using subsequently the DGLAP equations for all other scales.

Introducing pdfs as separate concept of the matrix element of the hard process is called *factorization ansatz*.

In the factorization approximation, the cross section of a reaction $pp \to X$ can be written as

$$\sigma_{pp\to X} = \int dx_a dx_b f_a(x_a, \mu_F^2) f_b(x_b, \mu_F^2) \left(\hat{\sigma}_0(\mu_F^2) + \alpha_s(\mu_R^2)\hat{\sigma}_1(\mu_F^2) + \dots\right)_{ab\to X}. \quad (1.9)$$

Here, f are the pdfs for a and b, respectively. μ_F is the *factorization scale* and μ_R is the *renormalization scale* for the running coupling constant α_S. $\hat{\sigma}_0$, $\hat{\sigma}_1$,... are the leading order, next-to-leading order cross sections. Formally, the cross section is invariant under change of μ_R and μ_F, if one includes *all* perturbative orders. For calculations, concrete values for μ_R and μ_F have to be chosen. They should be of the same order as the momentum transfer in the hard process, for resonances for example, one often chooses the mass of the resonance for both μ_R and μ_F. Deviations due to different values of μ_R and μ_F express the uncertainty from perturbation theory and have to be included as systematic uncertainties if using those values.

As depicted in figure 1.1, the description of full events invloves several steps:

1. Describing the *parton density function* at the factorization scale μ_F. As described above, this is done by numerically solving the DGLAP equations. The algorithm applied is basically the same as for parton shower evolution (see 3.) and consists of iteratively splitting the partons according to the perturbative splitting functions describing the processes $G \to GG$, $G \to q\bar{q}$ and $q \to qG$. This gives rise to additional final state partons, the "initial state radiation".
2. The *hard interaction* involving two initial state particles from step 1 yielding a small number of final state particles.
3. *Shower evolution* of the final state partons from step 2, yielding more and more partons with decreasing virtuality. This step is treated perturbatively by numerical Monte-Carlo integration the DGLAP equations.
4. *Hadronization*: As the shower evolution is based on perturbative QCD which breaks down at scales of order $\Lambda_{QCD} \approx 200$MeV, it cannot used to describe

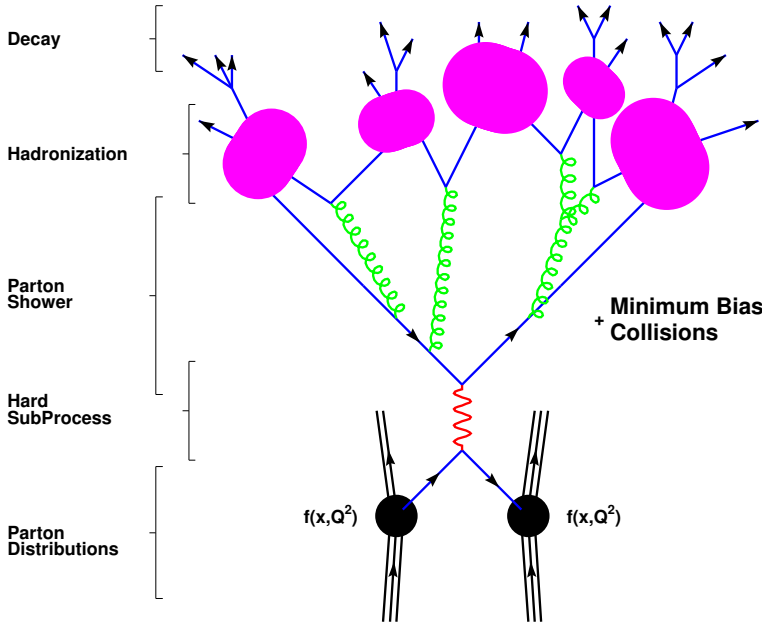

Figure 1.1: Illustration of the steps describing events at hadron colliders [16].

how confinement mechanisms turn partons into hadrons. There are several phenomenological models to describe this process, like the Lund string model used in the Pythia event generator [17] or a cluster fragmentation model in Herwig++ [18]. The parameters of those models are tuned to data.

5. Most hadrons of the previous step are unstable and *decay* via the strong or electromagnetic interaction. Known decay modes and branching ratios and decay models where the hadrons decay according to the available phase space are used. This yields "stable" particles which reach the detector. This includes hadrons like pions, kaons, protons and neutrons and leptons from the hard interaction or decays such as muons, electrons and neutrinos.
6. Interaction of the proton remnants which is called the *underlying event*. This is also described by phenomenological models tuned to available data. It usually adds additional soft hadrons to the final state.
7. For a given proton-proton bunch crossing, more than one inelastic scattering can occur. This is called *multiple interaction*. As the $2 \to 2$-processes with soft final state particles have by far the largest cross section, those additional interactions are referred to as "minimum bias collisions" and typically only add soft hadrons to the final state.

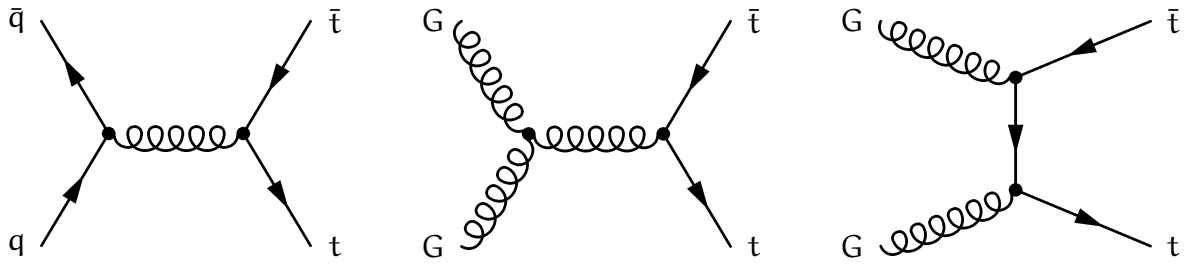

Figure 1.2: The leading order Feynman graphs for $t\bar{t}$ production at a hadron collider. Top quark pairs are produced via $q\bar{q}$ initial state (left) or GG initial state (center and right).

1.4 $t\bar{t}$ Production and Decay in the Standard Model

Top quarks can be produced either as pair of top and anti-top ($t\bar{t}$) via the strong interaction or as single top quarks via the electroweak interaction. The $t\bar{t}$ cross section is about 2.5 times larger than the single top cross section. More importantly, $t\bar{t}$ events can be selected with much higher purity and efficiency than single top events.[7] Therefore, previous studies of top quark properties mostly focus on the $t\bar{t}$ channel which is discussed in this section.

Top-quark pair production at a hadron collider can be divided according to its initial state into the "production channels" GG and $q\bar{q}$ (see figure 1.2). The next-to-leading order cross section for top-quark pair production for $m_t = 171$ GeV at $\sqrt{s} = 10$ TeV using the CTEQ6.5 PDF set is [19]

$$\sigma_{t\bar{t}} = \left(414^{+36}_{-38}(\text{scales})^{+20}_{-18}(\text{pdf})\right)\text{pb}.$$

The central value is determined with renormalization and factorization scales at $\mu_R = \mu_F = m_t$. The scale uncertainty is the largest deviation from the central value if varying μ_R and μ_F independently with the constraints

$$\frac{1}{2}m_t \leqslant \mu_F \leqslant 2m_t \qquad \frac{1}{2}m_t \leqslant \mu_R \leqslant 2m_t \qquad \frac{1}{2}\mu_R \leqslant \mu_F \leqslant 2\mu_R.$$

The PDF uncertainties are evaluated with the scales set to their central values varying the PDF according to the errors provided by the collaboration which provides the PDF. They roughly represent a 90 % confidence level.

The $t\bar{t}$ cross section as a function of the invariant mass of the top pair system, $m_{t\bar{t}}$, is shown in figure 1.3. To illustrate how additional resonances can alter this spectrum, the spectrum is given for different graviton models. The standard model

[7] In fact, a good selection of single top events is so difficult that the 5σ discovery of the single top process has succeeded only very recently [6].

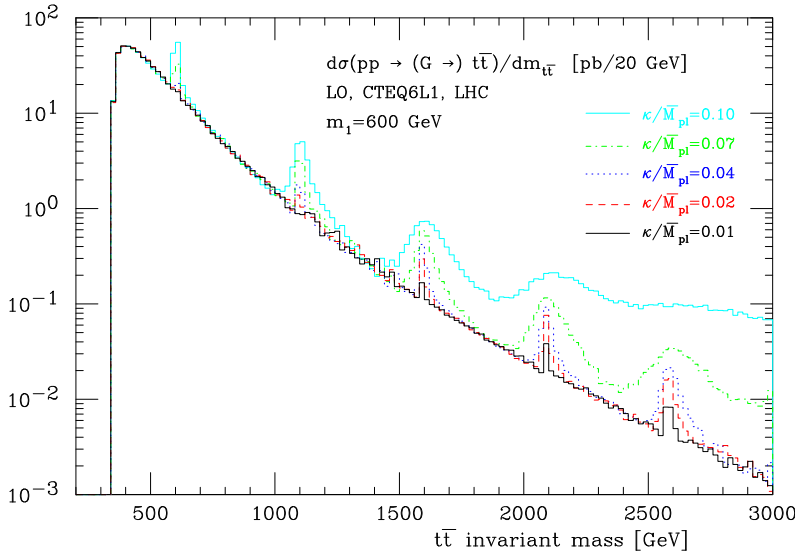

Figure 1.3: Spectrum of the invariant mass of the $t\bar{t}$ system for $pp \to t\bar{t}$ including s-channel gravitons in the RS model. The mass of the first Kaluza-Klein mode is set to $m_1 = 600\,\text{GeV}$, the lines represent different values for $\kappa/\bar{M}_{pl}$ [20].

$t\bar{t}$ production cross section has its largest cross section just above the production threshold at $2m_t$ and is steeply falling.

At the LHC, the GG production channel accounts for approximately 84 % of the overall $t\bar{t}$ cross section at 10 TeV center-of-mass energy [12]. This ratio decreases for high $m_{t\bar{t}}$, as the parton density function for gluons is falling for larger x faster than the pdf for quarks.

Top-quark decay is only possible via flavor change and therefore always includes radiating a *W* boson while the top quark line changes flavor into a down-type quark (see figure 1.4). Because of the large KM matrix element V_{tb}, this will be a b quark in nearly 100 % of the cases.

The *W* boson in turn can decay in either a quark-antiquark pair (*hadronic decay*) or in a charged lepton and neutrino (*leptonic decay*).

Unlike other heavy flavor decays, the *W* boson in this decay is on-shell because of the large top mass. For the same reason, there is a large phase space which implies a short lifetime for the top quark of $\tau_t \approx 5 \cdot 10^{-25}$ s [21]. The time scale for hadronization is given by $1/\Lambda_{QCD} \approx 7 \cdot 10^{-24}$ s (for $\Lambda_{QCD} = 100$ MeV).[8] Therefore, the top quark decays without hadronization allowing measurements of decay properties not possible with other quarks, such as the *W*-boson helicity [22].

[8] Calculating hadronization scale by using the time interval of light propagation through a meson with experimental meson radii of ≈ 1fm yields the same conclusions.

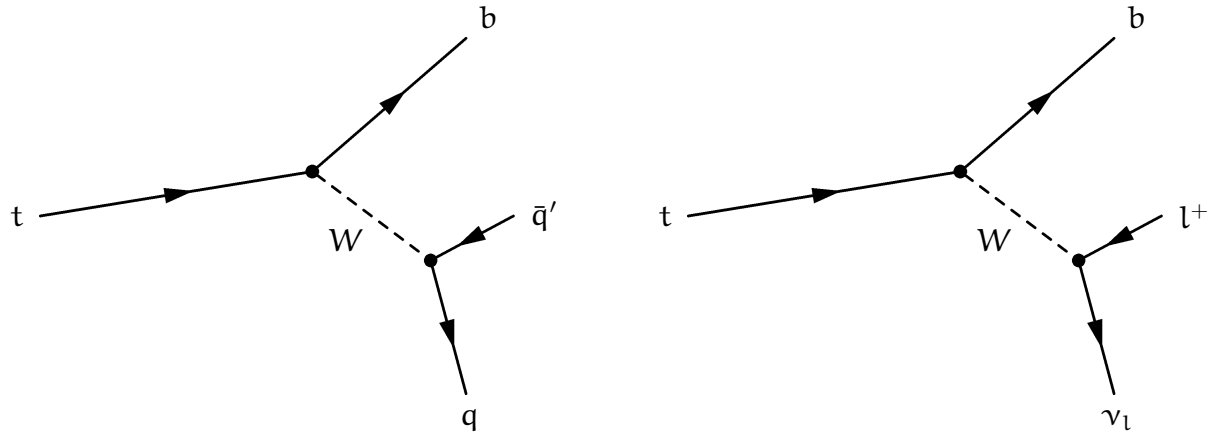

Figure 1.4: Top quark decay to a b quark and a *W* boson. The *W* boson decays either to $q\bar{q}'$ or a charged lepton and the corresponding neutrino ($l\nu_l$).

Top-quark-pair decays can be divided into three classes according to the *W*-boson decays: [21]

- *fully hadronic* decay where both *W* bosons decay to $q\bar{q}'$ (46.2 %),
- *semi-leptonic* decay where one *W* boson decays to $q\bar{q}'$ and the other to a charged lepton and a neutrino (43.5 %),
- *fully leptonic* decay where both *W* bosons decay to a charged lepton and a neutrino (10.3 %).

The lepton channels are further divided by the lepton flavor into the *e*, μ and τ channels. As the detector signatures of *e* and μ are much clearer than for τ (which decays before reaching the detector and often decays hadronically), most results for the "lepton" channels do not use τ.

In the studies presented here, only the semi-leptonic decay to a μ is considered. This channel has a branching ratio of 14.5 % and is called "muon+jets".

1.5 Beyond Standard Model Contributions to $t\bar{t}$ Production

The top quark is unique in its large mass and (equivalently) large coupling to the Higgs particle. Indirect limits can be put on the Higgs mass with precision measurements of m_W and m_t [23]. Furthermore, Higgs production with gluon-gluon fusion via a top quark loop is the dominant production process at the LHC [24].

In the introduction of the Higgs mechanism in section 1.2.1, it was mentioned that the Higgs can be seen as an example of theories for mass generation compatible with local gauge invariance. So even if there is no Higgs mechanism in nature, the large top mass implies a large coupling to the new sector. Not surprisingly, the top

quark plays a special role in some alternative mechanisms of electroweak symmetry breaking.

Model-Independent Search In the present studies, a *model-independent* search for heavy resonances to $t\bar{t}$ is performed, i. e. no specific model which would explain this resonance is assumed. Rather, the $t\bar{t}$ spectrum is searched for structure not explained by the standard model. The origin of the resonance could be an additional broken $U(1)$ symmetry which couples mainly to top quarks, scalar (Higgs-like) bosons, heavy gluons or resonances arising in models with extra dimensions such as Rundall-Sundrum [25] and ADD models [26]. for an overview and specific references, see [20]. If a new resonance decays only into quarks or gluons, the $t\bar{t}$ channel is the most sensitive as the huge irreducible dijet background limits the sensitivity in other final states.

The general case of a new resonance decaying to $t\bar{t}$ can be introduced by adding an additional, minimal effective term for this resonance X to the standard model Lagrangian. By specifying all its couplings to the fermions and gauge bosons, all other properties like cross section and decay modes follow automatically.

This approach was implemented in the event generator MadGraph/MadEvent as the package `topBSM` [20]. With this package, it is possible to specify the spin, whether it couples to gluons and parity (vector and axial-vector coefficients of the coupling). Most properties affect production cross section and angular distributions of the final state $t\bar{t}$ pair but have very little impact on the general search strategy developed here.[9]

1.5.1 Topcolor Z'

Limits can be cited both as mass-dependent upper limit or as limit on the mass of the resonance X. The latter is only possible if one has chosen a specific model which predicts cross sections for the resonance.

As benchmark for searches for resonances with $t\bar{t}$ final state, the topcolor Z' [7] has been chosen for Tevatron (in proton-antiproton collisions with $\sqrt{s} = 1.96$ TeV). For easier comparison, when citing mass limits, we use cross sections predicted by that model [27].

The topcolor Z' arises in an alternative model of electroweak symmetry breaking. In this model the gauge group for QCD ($SU_C(3)$) is embedded into a larger group, for example $SU_1(3) \times SU_2(3)$ with couplings h_1 and h_2 respectively and $h_2 \gg h_1$.

[9] Of course, this will change if and when a resonance is discovered and all its properties are to be determined.

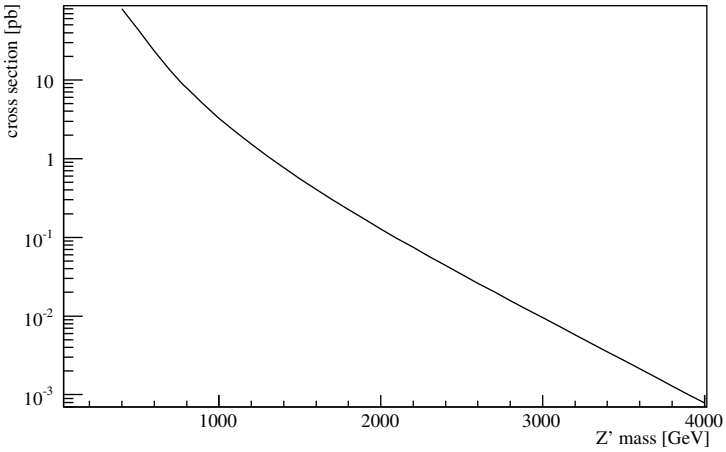

Figure 1.5: Leading order topcolor $pp \to Z' \to t\bar{t}$ cross sections for the LHC at $\sqrt{s} = 10$ TeV [29]. The currently excluded mass range for Z'_t is $m_{Z'} < 760$ GeV at 95 % confidence level [28].

The first gauge group couples only to the first two, the second group only to the third generation of quarks. As no extra gluon-like particles have been observed, symmetry has to be broken to the one observed in the standard model, $SU_C(3)$. Both the breaking of this new symmetry as well as electroweak symmetry breaking is achieved by a $t\bar{t}$ "condensate".

To explain the mass differences and prevent formation of $b\bar{b}$ condensate, an additional mechanism is needed. One possibility is the introduction of the topcolor Z', Z'_t.

The cross section times branching ratio in $t\bar{t}$ predicted for the LHC for different masses of Z'_t can be seen in figure 1.5.

With the higher center-of-mass energy available at the LHC, the current 95 % C. L. lower limit on $m_{Z'_t}$ of $m_{Z'_t} > 760\,\text{GeV}$ [28] mass could be improved with an amount of data equivalent to about the first year of running.

2 The CMS Detector at the LHC

This chapter gives a short introduction to the Large Hadron Collider (LHC) and the Compact Muon Solenoid detector (CMS). For a more complete overview, see [30] for LHC and [31], [32] and [33] for CMS.

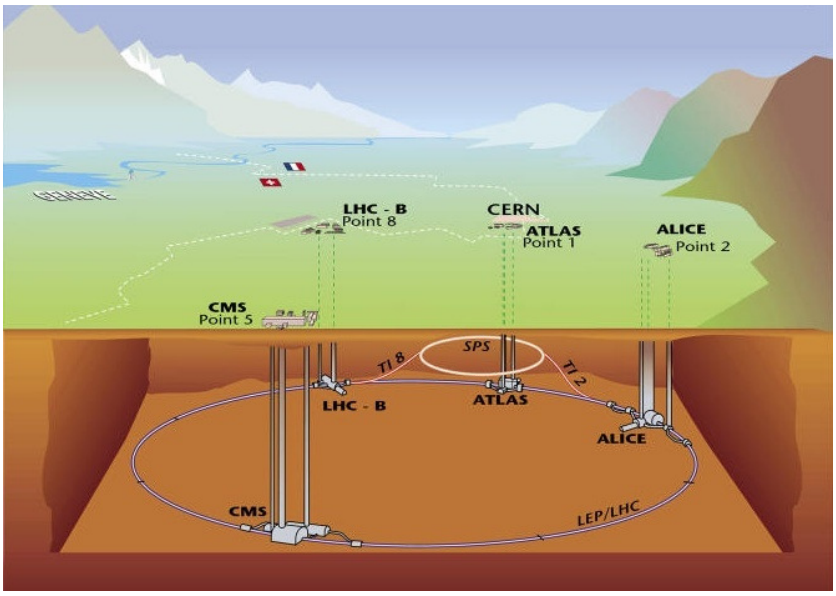

Figure 2.1: Overall view of the LHC with the four major detectors [34].

2.1 The Large Hadron Collider

The Large Hadron Collider is the world's largest particle collider located at CERN[1] near Geneva at the Franco-Swiss border (see figure 2.1). It is designed to collide protons at an unprecedented center-of-mass energy of 14 TeV or lead ions at a center-of-mass energy of 5.5 $^{\text{TeV}}/_{\text{nucleon}}$.

The LHC consists of a ring-shaped tunnel of 27 km circumference about 100 m below the surface where more than one thousand magnets are installed to bend the proton beams around the ring. Four major experiments are installed on the ring:

[1] European Organization for Nuclear Research (french: Conseil Européen pour la Recherche Nucléaire)

the general-purpose detectors ATLAS and CMS, and the special-purpose detectors ALICE[2] and LHCb[3].

One major question addressed with the two general-purpose detectors ATLAS and CMS is the existence of the Higgs particle or, more generally, the nature of the electroweak symmetry breaking. Another major question is the existence of new particles with masses at the TeV scale. LHC might be the first collider to produce such particles. A popular theory which can be either discovered or excluded (for many parameter sets, at least) is supersymmetry (SUSY) which postulates a (broken) symmetry between bosons and fermions. It predicts many new particles with masses at the TeV scale. There are also many other searches for physics beyond the standard model possible which will cover a much larger parameter space than before. For example, searches for a fourth quark generation (called b' and t') or for W and Z like particles with a much higher mass (W' and Z').

Apart from the center-of-mass energy, the luminosity L is the most important quantity to measure the performance of a collider. For a given process with cross section σ, the number of events per second is given by

$$N = L\sigma.$$

L is given by[4]

$$L = \frac{f n N^2}{A}$$

where f is the beam revolution frequency, n is the number of bunches per beam, N is the number of particles per bunch and A is the geometrical cross section of the beams at the interaction point. The design luminosity for the LHC is $10^{34}\,\mathrm{cm}^{-2}\mathrm{s}^{-1}$. The event rates for different processes can be seen in figure 2.2.

The use of protons for both particle beams allows a higher luminosity than for a proton-antiproton design as used by Tevatron because it is very difficult to produce anti-protons in sufficient quantities. The two counter-rotating proton beams require opposite magnetic fields. As the LHC tunnel[5] is too small for two separate rings of magnets, the LHC uses twin magnets consisting of two sets of coils but sharing the mechanical structure and cooling system. For an energy of 7 TeV per beam, a

[2] A Large Ion Collider Experiment, mainly addressing the physics of very dense, strongly interacting matter [35].

[3] It is dedicated to precision measurements of CP violation and and rare b-meson decays [36].

[4] Note that this formula is a simplification: it does not take into account the finite crossing angle of the beams and assumes that the area of the beams is well-defined.

[5] The tunnel was originally built for LEP, the Large Electron-Positron Collider which collided electrons and positrons at a center-of-mass energy between 90 GeV at startup in 1989 and 209 GeV just before shutdown in 2000. A review of its main results can be found in [38].

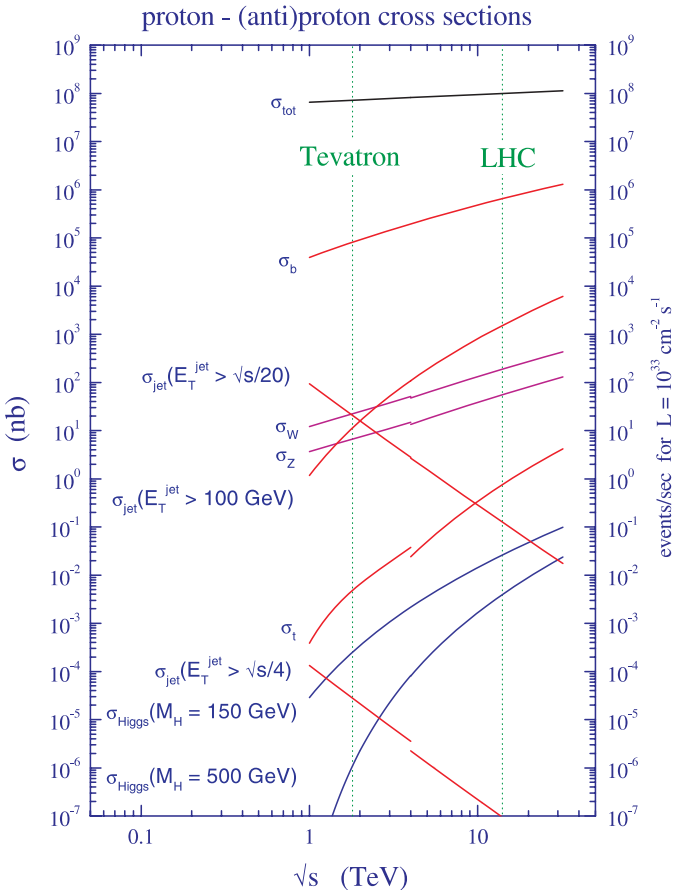

Figure 2.2: Next-to-leading order cross sections for different processes for proton-antiproton (for $\sqrt{s} < 4\,\text{TeV}$) or proton-proton collisions (for $\sqrt{s} > 4\,\text{TeV}$) [37]. The LHC design luminosity is $10^{34}\,\text{cm}^{-2}\text{s}^{-1}$, so it is a factor ten higher than what is given in the figure. From the plot one can see that for example the Higgs boson cross section is several orders of magnitude below the total cross section. Important background processes for $\text{t}\bar{\text{t}}$-production like W-boson and Z-boson production have comparably large cross sections. So from this plot alone one can conclude that developing criteria for selecting the interesting events is extremely important.

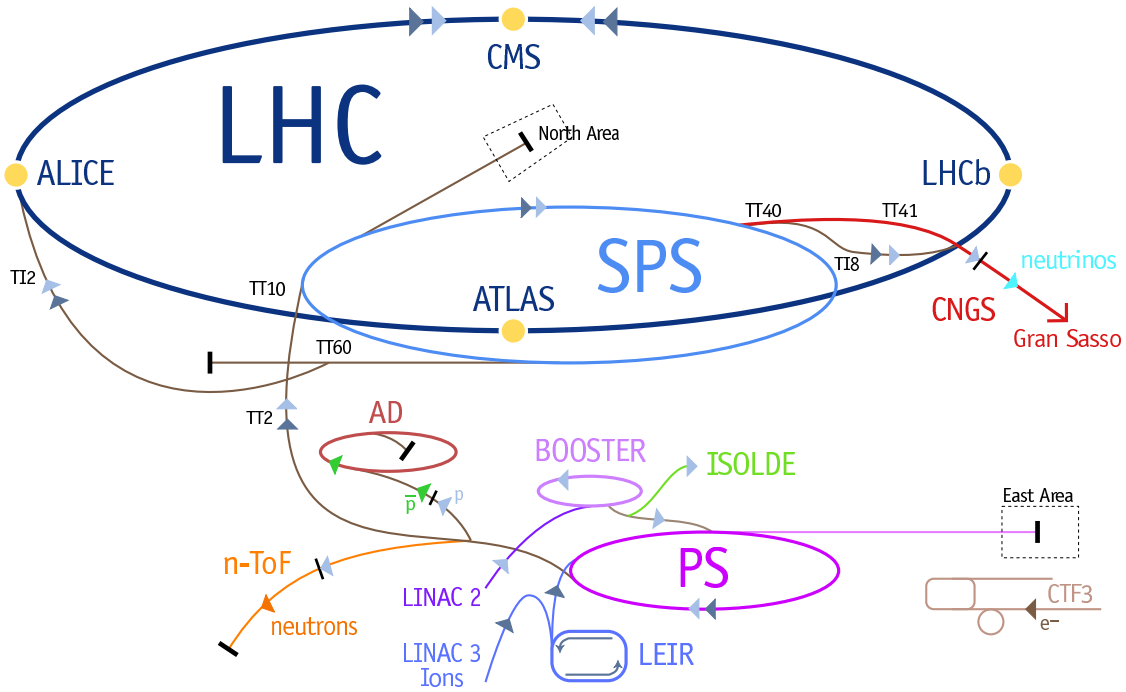

Figure 2.3: Schematic overview of the LHC and adjoint accelerators [39]. The accelerator chain to fill the LHC is Linac 2 – Booster – PS – SPS.

magnetic dipole field of 8.33 T is required. In total, 1 232 dipole magnets are installed. The ability to produce and reliably operate such a strong magnetic field is the current limitation for the beam energy. Protons are arranged into 2 808 bunches per beam; each bunch consists of about $1.1 \cdot 10^{11}$ protons.

In the LHC, the protons are accelerated by radio frequency cavities which increase the proton energy by $0.5\,^{\mathrm{MeV}}/_{\mathrm{turn}}$. Thus, it takes about 20 minutes to ramp up the beam from its injection energy of 450 GeV to the design energy of 7 TeV. The acceleration up to the injection energy is provided by the chain of accelerators (see figure 2.3) Linac 2 – Proton Synchrotron Booster (PBS) – Proton Synchrotron (PS) – Super Proton Synchrotron (SPS). For a full fill of LHC, multiple cycles of those accelerators are needed, yielding a theoretical minimal filling time of around 70 minutes. Experience from other accelerators has shown that actual filling times are about factor 6 higher than this minimum, so an according estimate for the LHC would be 7 hours.

Current Status After promising *first beam* at September 10th 2008, a major incident just 9 days later delayed first collisions by presumably one year.[6]

A superconducting electrical connection between two magnets developed a resistance leading to an electrical arc. This heated the liquid helium, building up more pressure than the installed release valves could handle. The subsequent burst severely damaged several magnets. The affected sectors have to be warmed up, and 53 magnets had to be replaced. At the same time, devices for sensitive electrical measurements are installed in order to detect similar resistances early so that countermeasures can be taken quickly. Also, new helium release valves are installed to prevent the damage in a similar incident.

In superconducting magnets, a cable can undergo a transition from superconducting to the resistive state. If the heat released from the cable cannot be absorbed quickly enough, the whole magnet is affected and becomes resistive. This is known as *quench*. By repeated (intentional) quenching, a magnet can be *trained*. During those quenches, the cables rearrange into a more stable position such that after training, higher magnetic fields can be reached [41].

Training the magnets in the LHC has shown that it takes a long time before reaching a stable operation for the magnetic field strength required at the high design beam energy of 7 TeV. Therefore, for the first run, a beam energy of 5 TeV is used instead. This first run is planned to last about one year and will yield an integrated luminosity of about $200\,\mathrm{pb}^{-1}$. Therefore, the present studies use throughout a scenario with $\sqrt{s} = 10\,\mathrm{TeV}$ and $\mathrm{L} = 200\,\mathrm{pb}^{-1}$.

2.2 The CMS detector

The Compact Muon Solenoid detector is shown in figure 2.4. It is 22 m long, has a diameter of 15 m and weights about 12 500 t. It is "compact" when compared to the other general-purpose detector at LHC (ATLAS) which has about twice the volume and half the weight of CMS [42].

As general-purpose detector, a high angular coverage is important in order to identify and measure a large phase space of final state particles escaping the beam pipe. Such detectors are known as "4π detectors". In order to achieve that, CMS consists of a cylindrical barrel whose symmetry axis is the beam pipe, and two endcaps.

The superconducting solenoid installed between the hadron calorimeter and muon system (see figure 2.4) produces a very homogeneous magnetic field of 3.8 T in the

[6] At time of writing, the most recent schedule available is the one from the Chamonix workshop. It foresees *first beam* at the end of September 2009 and first collisions about four weeks later.[40]

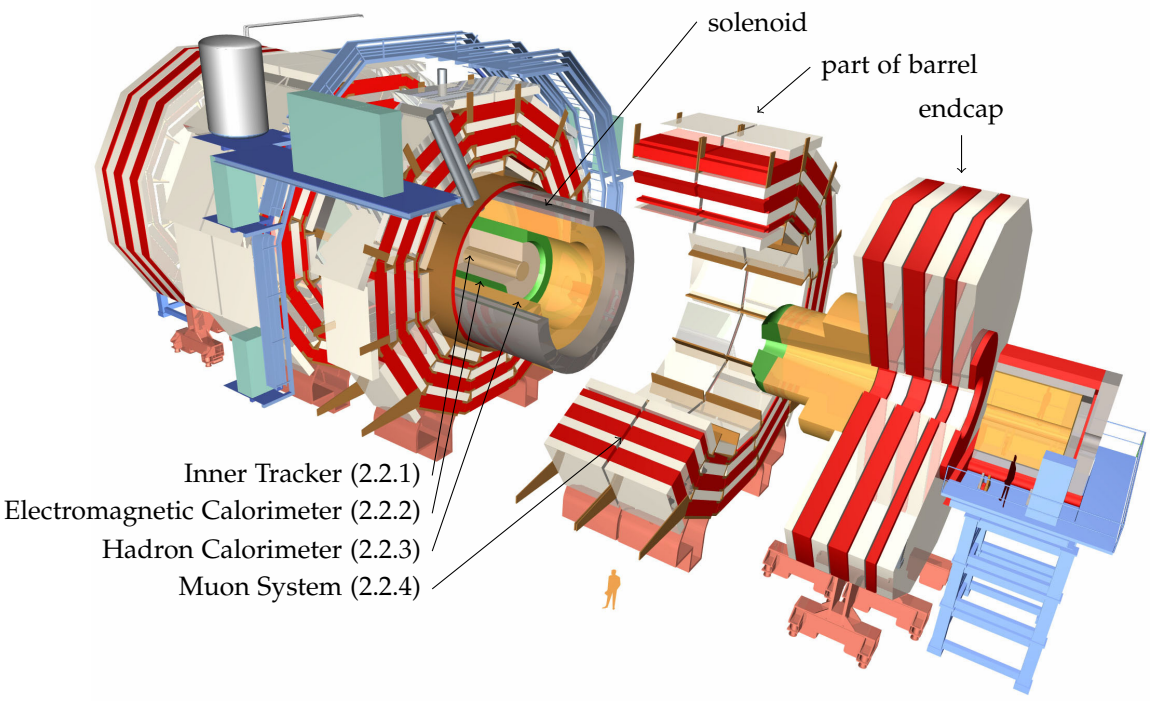

Figure 2.4: The CMS detector at the LHC [33]. It is divided into the central barrel and two endcaps. It is 22 m long and has a diameter of 15 m. The subdetectors are explained in the respective sections given in the legend.

whole tracker volume. This allows a momentum measurement of charged particles by measuring the bending radius of its trajectory.

The main detector requirements were defined as follows:

- good muon identification over a large range of momenta and angles;
- good reconstruction efficiency and momentum resolution for charged particles in the inner tracker;
- good electromagnetic energy resolution, good diphoton and dielectron mass resolution with high angular coverage and efficient photon and lepton isolation at high luminosities;
- good missing-transverse-energy and dijet-mass resolution.

These requirements are implemented by different specialized subdetectors which are arranged concentrically around the beam line. Before going into the details of the individual subdetectors, the interactions of various high-energy particles passing through matter are reviewed. Figure 2.5 shows a cut through the CMS detector and some particles passing the sub-detectors:

- *Electrons* leave a track in the inner tracker and interact mainly by radiating

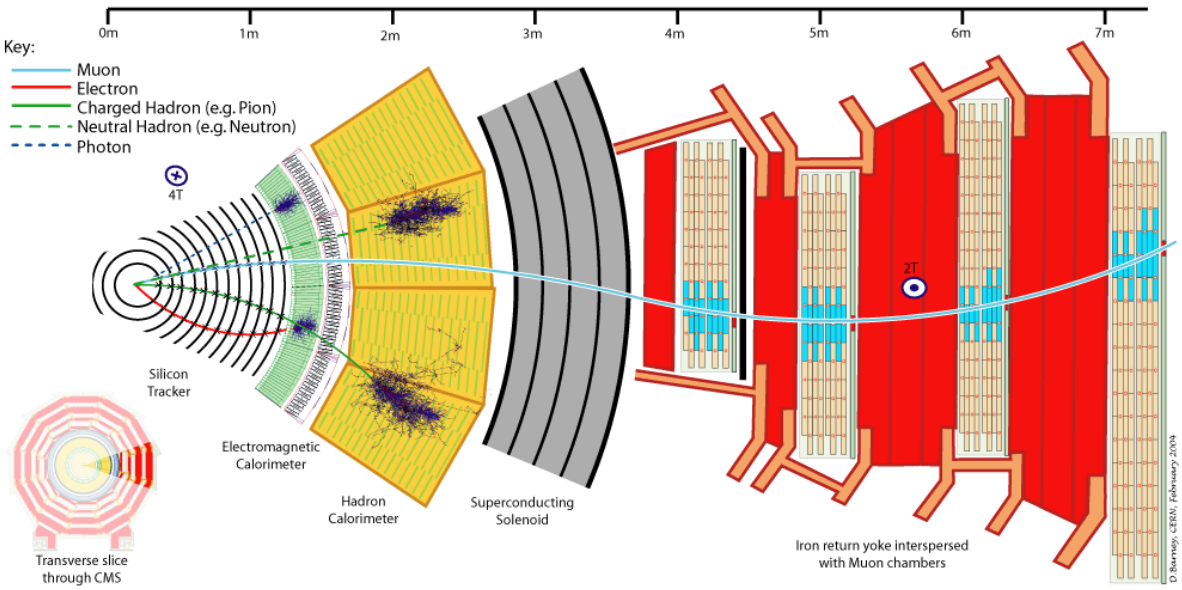

Figure 2.5: Slice of CMS showing the sub-detectors and interacting particles [33]: Muons are the only type of detectable particle reaching the muon system installed as outmost system; electrons and photons shower mainly in the electromagnetic calorimeter and hadrons in the hadron calorimeter. Charged particles leave a track in the inner silicon tracker.

high-energy photons (bremsstrahlung) which in turn mainly produce electron-positron pairs. This leads to a cascade of secondary particles of decreasing energy ("electromagnetic shower").

- *Photons* do not have a track. In the calorimeter, they behave similar to electrons.
- *Muons* leave a track in the inner tracker. Energy loss due to bremsstrahlung is much lower as for electrons, so they loose only very little energy in the calorimeters. While all other types of particles are absorbed in the calorimeters, muons reach the outer parts of the detector which are therefore used to identify muons.
- *Charged hadrons* (mainly pions, kaons and protons; usually in jets) leave a track in the tracker and produce through strong interaction with nuclei in the material a shower in the electromagnetic and hadron calorimeters. Hadron showers are typically much longer than electromagnetic showers originating from electrons and photons. Therefore, most of the energy is deposited in the hadron calorimeter.
- *Neutral hadrons* (mainly neutrons, kaons; usually in jets) leave no track in the

tracker. Like charged hadrons, they produce a hadron shower mainly in the hadron calorimeter.

- *Neutrinos* do not interact with any detector component. They can be identified indirectly by using momentum conservation in the transverse (x-y) plane: the sum of momenta of all final state particles must be zero. Any deviation is interpreted as the sum of the transverse momenta of neutrinos (or other non-interacting particles).

With pattern recognition algorithms, it is possible to analyze the low-level detector response yielding high-level physics objects such as electrons, muons and jets (see chapter 3).

Reconstructing τ-leptons and b mesons (in b-jets) is more difficult because they decay before reaching the detector. Analyzing patterns produced by their decay products, identification is still possible but performance is not as good as for electrons, muons or jets. As those objects are not used in this analysis, their reconstruction is not discussed.

The high event rate at the LHC is another major factor in the design of CMS. The total inelastic cross section is expected to be about 100 mb. The collision frequency of the proton bunches is 40 MHz. At the LHC design luminosity, there will be on average around 1000 particles per proton-proton interaction and about 20 interactions per bunch crossing. Processing the event data for each bunch crossing would require handling an enormous amount of data of about 20 TB per second[7]. This is neither technically possible nor is it required from a physics point of view. Therefore, interesting events are filtered on-line with a *trigger*.

The coordinate system of CMS is defined as follows: the origin is at the nominal collision point, the x-axis points to the center of the LHC ring, the y-axis points vertically upwards. Thus, the z-axis points along the beam direction such that the (x, y, z)-axes are a right-handed coordinate system. The polar angle Θ is measured from the z-axis, the pseudorapidity is defined as $\eta = -\ln \tan \frac{\Theta}{2}$.

2.2.1 CMS Tracker

Coming from the interaction point, the first layer a particle passes is the tracker [43]. Its task is to measure the tracks of the charged particles. From the curvature of the track and the magnetic field of 3.8 T produced by the solenoid, one can derive the sign of the charge and the momentum of the particle. The hits in the tracker can also

[7] The actual value of this figure depends on the assumptions. The 20 TB/s is calculated given the bunch crossing rate of 40 MHz and the size of a (already zero-suppressed) RAW event data record of 0.5 MB.

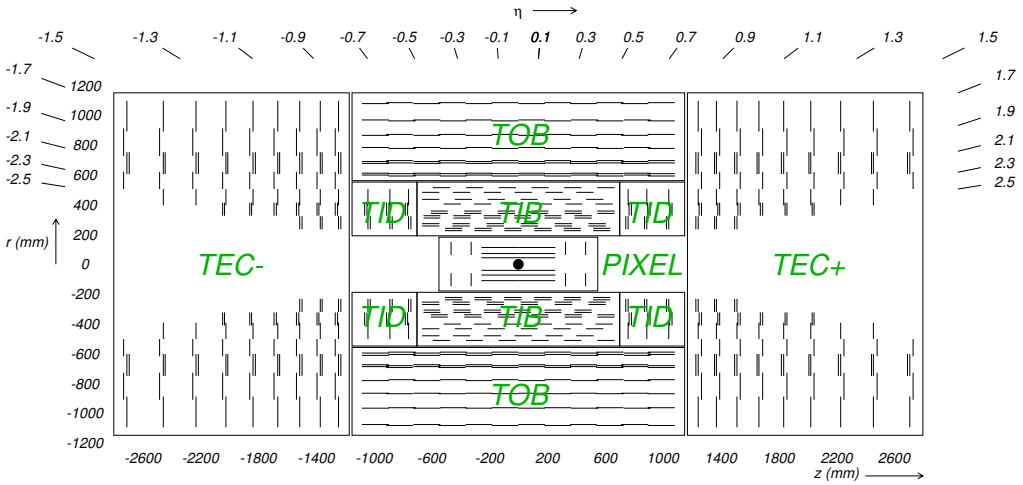

Figure 2.6: Schematic cross section through the CMS tracker [31]. Each line represents a detector module, double lines indicate back-to-back modules which deliver stereo hits.

be used to reconstruct primary and secondary vertices of an event (see section 3.1 and section 3.2).

The high particle flux will cause high radiation damage. The tracker has to be designed such that it is operational nonetheless for about 10 years. The high collision rate demands for a high timing resolution and fast response. This requires a high power density of about 60 kW for the whole tracker volume and adequate cooling to the working temperature of the detectors of $-10°$C. However, this is in conflict with the aim of keeping the material in the tracker to a minimum in order to avoid any particle interaction beyond the electronic excitation in the silicon detectors.

Tracker elements are installed in a large tracker support tube and aligned to ± 0.5 mm. Using event tracks for alignment, the true position of the tracker elements can be determined to high precision allowing measurements of the impact parameter with a precision of up to 10 µm for high-p_T particles. Resolution is limited by multiple scattering for low p_T and by pixel-size and alignment for high p_T.

A schematic drawing of the tracker is shown in figure 2.6. It consists of 1 440 pixel and over 15 000 strip detector modules with an active area of about 200 m^2. It covers the pseudorapidity range $|\eta| < 2.5$.

Pixel Tracker In order to keep the occupancy[8] at the few percent level, pixel detectors have to be used at the innermost layers. There are three cylindrical pixel

[8]The occupancy is the fraction of responding channels for an event.

layers (BPix) at radii of 4.4, 7.3 and 10.2 cm and two endcaps at $|z| = 34.5$ cm and $|z| = 46.5$ cm (FPix). The pixel size is $100 \times 150\,\mu m^2$, the total number of pixels being about 66 million with an area of about $1\,m^2$. The pixels cover the full pseudorapidity range of the tracker.

For each charged particle, the pixel detector delivers three high precision space points.

One read-out chip controls a group of 52×80 pixels. Read-out is carried out only if a Level-1 trigger has fired. This implements the mentioned requirement for data reduction at a very early stage.

Silicon Strip Tracker The region between $r = 20$ cm and $r = 116$ cm is instrumented by silicon strips. The silicon strip tracker consists of an inner and an outer part.

The Tracker Inner Barrel (TIB) consists of four layers, complemented by three disk layers at each end (Tracker Inner Disks — TID). The strips in the barrel are aligned along the beam axis and radial in the disks. TIB/TID delivers up to four $r - \phi$ measurements per track. With a strip pitch of about 100 µm (different for inner and outer layers), it reaches a single point resolution of about 30 µm.

The TIB is surrounded by the Tracker Outer Barrel (TOB) which consists of 6 layers and extends to $r = 116$ cm. It provides additional six $r - \phi$ measurements with point resolutions of about 50 µm (a higher precision of 35 µm is achieved by smaller strip pitches in the two outer layers).

Beyond the z range of the TOB ($|z| \leqslant 118$ cm), the Tracker EndCaps (TEC+ and TEC−, where the sign denotes the location along the z axis) cover the region $124\,cm \leqslant |z| \leqslant 282$ cm. Each TEC consists of nine disks providing up to nine ϕ measurements per trajectory.

As can be seen in figure 2.6, there are stereo modules in each part of the strip tracker. Those stereo modules carry a second module which is mounted back-to-back with a stereo angle of 100 mrad. This provides measurement of the second coordinate (z in the barrel and r in the disks). The measurement of this additional coordinate has a lower resolution raging from 230 µm for the TIB to 530 µm for the TOB.

2.2.2 Electromagnetic Calorimeter

The electromagnetic calorimeter (ECAL) [44] is a homogeneous calorimeter made of almost 70 000 lead tungstate ($PbWO_4$) crystals. Lead tungstate is a high density material ($8.3\,g/cm^3$) transparent for visible light. It has a short scintillation decay time: 80% of the light is emitted in the 25 ns LHC bunch crossing interval. The high density

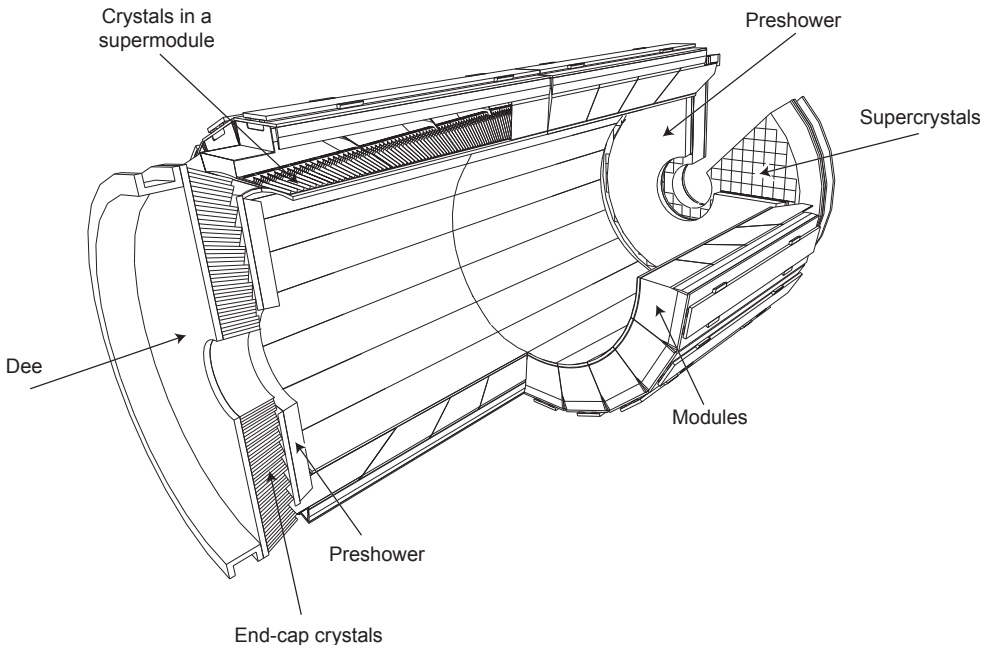

Figure 2.7: Schematic cross section through the CMS electromagnetic calorimeter (ECAL).[31]

is correlated to a short radiation length[9] of 0.89 cm and a small Molière radius[10] of 2.2 cm. The blue-green scintillation light (4.5 photoelectrons per MeV) are collected by photodetectors attached to each of the crystals.

The ECAL layout can be seen in figure 2.7. In the barrel region ($|\eta| < 1.479$), there are 360 crystals along ϕ and (2×85) crystals along η. The crystal cross section corresponds to about 0.0174×0.0174 in $\eta \times \phi$ or $22 \times 22\,\mathrm{mm}^2$ at the inner face of the crystal and $26 \times 26\,\mathrm{mm}^2$ at the outer crystal face. Crystal length is 230 mm, corresponding to about 26 radiation lengths. The inner faces of the crystals are at $r = 129$ cm. The crystals are separated by a structure of thin walls made of aluminum facing the crystals and two layers of glass fiber-oxy resin.

The endcaps (EE) cover the pseudorapidity range $1.479 < |\eta| < 3.0$. They start at around $|z| = 315$ cm. Crystals are grouped into mechanical units of 5×5 crystals (supercrystals). The crystals and supercrystals are arranged in a rectangular $x - y$

[9] The radiation length is both (a) the length a high-energy electron looses all but $1/e$ by bremsstrahlung and (b) $\frac{7}{9}$ of the mean free path for e^+e^- pair production for high-energy photons. Therefore, it is the appropriate scale to describe high-energy electromagnetic cascades [21].

[10] R_M being the Molière radius, a cylinder with radius R_M contains 90% of the shower energy; about 99% of the energy is contained in a cylinder with radius $3.5R_M$ [21]. Thus, the Molière radius is the transverse scale of electromagnetic showers.

grid. The inner faces have a cross section of $28.6 \times 28.6\,\mathrm{mm}^2$, the cross section of the outer faces is $30 \times 30\,\mathrm{mm}^2$. Each endcap is divided into two halves, or *Dees*.

Preshower Detector The preshower detector is a detector to identify neutral pions in the endcaps within $1.653 < |\eta| < 2.6$. It also supports the correct identification of electrons and improves position resolution for photons and electrons. The preshower is a sampling calorimeter with two alternating types of layers: lead layers initiate showers from incoming high-energy electrons and photons while layers of silicon strips measure the deposited energy. The preshower is 20 cm thick consisting of a lead layer which is 2 radiation lengths thick ($2X_0$) followed by the first sensor layer, and another lead layer of thickness $1X_0$ before reaching the second sensor plane. Thus, about 90% of incident photons produce a e^+e^- pair before reaching the second sensor layer.

The energy resolution of the ECAL is given by different contributions: a stochastic term S which scales like $\sqrt{E}$, a noise term N with a scaling of E and a constant term C. The total resolution σ for an energy deposit E is thus given by

$$\left(\frac{\sigma}{E}\right)^2 = \left(\frac{S}{\sqrt{E}}\right)^2 + \left(\frac{N}{E}\right)^2 + C^2.$$

In test beam experiments conducted in 2004 with electrons, the resolution effects were found as expected and the parameters for the formula above explicitly measured (for an energy measurement with 3×3 crystals):

$$\left(\frac{\sigma}{E}\right)^2 = \left(\frac{0.028}{\sqrt{E}}\right)^2 + \left(\frac{0.12}{E}\right)^2 + (0.003)^2$$

where E is the unitless numerical value of the energy if measured in GeV. For example, an electron with $E = 120\,\mathrm{GeV}$ can be measured with a relative energy resolution better than 0.45%.

2.2.3 Hadron Calorimeter

The CMS hadron calorimeter (HCAL) [45] is used to measure the energy of hadron jets. It is also important for measuring the missing transverse energy, $\not{E}_T$ (which is calculated using information from the ECAL and HCAL). As can be seen in figure 2.8, it consists of the hadron barrel (HB), endcap (HE), outer (HO) and forward (HF) calorimeters. All parts of the HCAL except the HO are restricted by the outer extent of the ECAL ($r = 1.77\,\mathrm{m}$) and the inner extent of the magnetic coil ($r = 2.95\,\mathrm{m}$).

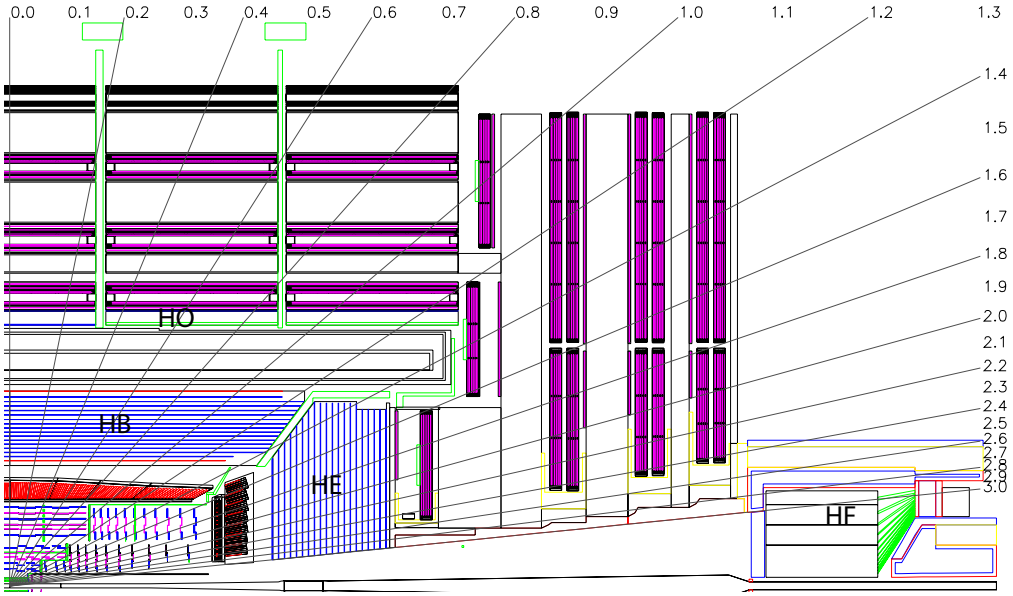

Figure 2.8: Scheme of the CMS hadron calorimeter (HCAL) consisting of the barrel (HB), endcap (HE), outer (HO) and forward (HF) calorimeters. The outer (red-magenta) parts belong to the muon system. The lines indicate the pseudorapidity η [31].

The hadron calorimeter barrel (HB) covers the pseudorapidity range $|\eta| < 1.3$. The barrel is divided at $z = 0$ into two half-barrels. Each half-barrel consists of 18 identical azimuthal wedges, each of them segmented into four sectors yielding a segmentation along ϕ of $\Delta\phi = 0.087$. Along η, the scintillator is divided in 16 sectors resulting in the same segmentation as ϕ of $\Delta\eta = 0.087$. The wedges are composed of alternating layers of absorber and scintillator. The absorber layers are about 50 mm thick and made of a composition of 70% Cu and 30% Zn (brass) with a density of 8.53 g/cm^3. The radiation length X_0 is about 1.5 cm and the nuclear interaction length $\lambda_I = 16.4$ cm. The total absorber thickness at $\eta = 0$ is $5.8 \cdot \lambda_I$ increasing to $10.6 \cdot \lambda_I$ at $|\eta| = 1.3$; the ECAL adds about $1.1 \cdot \lambda_I$ of material.

The hadron calorimeter endcaps (HE) cover the pseudorapidity range $1.3 < |\eta| < 3$. They were designed to minimize the dead region between HB and HE. In the HE region, the total absorber thickness is about $10 \cdot \lambda_I$, including the ECAL.

The forward calorimeter (HF) covers the region $2.9 < |\eta| < 5.2$. It has mainly be designed to withstand the large particle fluxes in this region. The signal is generated by Cherenkov light, so it is mostly sensitive to the charged component of the hadron showers. It has a coarser segmentation than the other parts of 0.175×0.175 in $\eta - \phi$. Information of the HF can be used to measure real-time luminosity.

In the central pseudorapidity region, the combined stopping power of ECAL and HCAL is not sufficient. Therefore, in the range $|\eta| < 1.3$, the outer calorimeter (HO) is installed to catch tails of hadron showers, using the solenoid coil as absorber material. It consists of five rings along the z axis at $r = 4.07\,\text{m}$. At $\eta = 0$, HB has the minimal absorber depth. Therefore, an additional absorber in the form of a 19.5 cm thick iron ring is installed. In this way, the minimal absorber thickness is extended to $11.8 \cdot \lambda_I$, except at the barrel-endcap boundary region.

2.2.4 Muon System

The muon system plays a central role in CMS. It is used for triggering, muon identification and measurement. As can be seen in figure 2.9, it consists of three detector types: drift tubes cover the central region, cathode strip chambers cover the outer regions. To control uncertainties, an independent and fast-responding system of resistive plate chambers is installed.

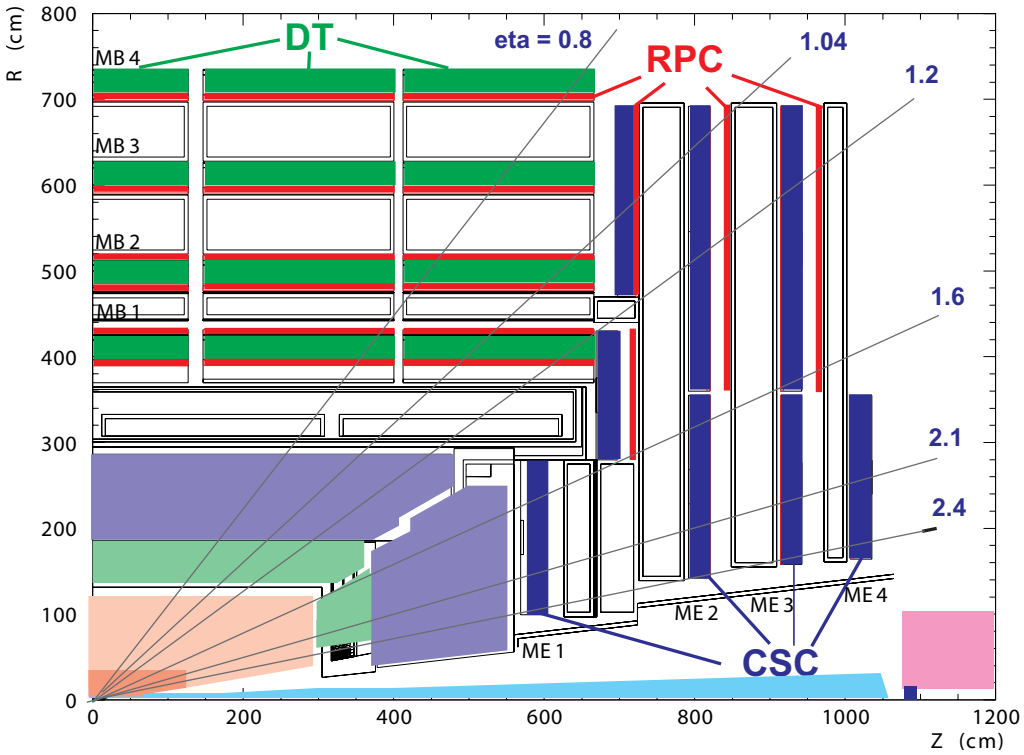

Figure 2.9: Scheme of the CMS muon system.[32] It consists of three types of detectors: drift tubes (DT), resistive plate chambers (RPC), and cathode strip chambers (CSC).

The barrel muon detector consists of four layers of drift tube chambers (DT)

arranged cylindrically around the beam line. Muons with $|\eta| < 0.8$ cross all four layers, muons with $|\eta| < 1.2$ at least one. The space between the drift tube chambers is occupied by the yoke-iron which gives rise to a 2 T magnetic field (see figure 2.5). Each chamber is made of three superlayers and each superlayer consists of four layers of drift tubes. The wires in the outer superlayers are parallel to the beam line and provide an $r - \phi$ measurement for each track. In the inner superlayer, wires are perpendicular to the beam line and thus provide a measurement of z (this third superlayer is not present at the outermost chambers which only consist of two superlayers). The resolution of a single wire is about 300 µm in $r - \phi$. Combination of the eight ϕ measurements leads to a precision of about 100 µm, a figure less than or comparable to the effects of multiple scattering up to $p_T \approx 200\,\text{GeV}$.

The endcap muon system consists of four layers of cathode strip chambers (CSC). Muons with $1.2 < |\eta| < 2.4$ cross three or four CSC layers. In the barrel-endcap overlap range of $0.9 < |\eta| < 1.2$ muons are detected by both DTs and CSCs. The CSCs are multiwire proportional chambers where 6 anode wire planes are interleaved with 7 cathode panels. The panels are segmented into strips along ϕ while wires run approximately perpendicular to the strips. The radial position of the wires provide a measurement for r while the induced charges on the strips provide a measurement for ϕ.

The resistive plate chambers (RPCs) have a time resolution much better than the LHC bunch crossing time of 25 ns but a worse space resolution compared to DTs and CSCs. The signal from RPCs can be used to unambiguously identify the relevant bunch crossing to which a track is to be associated and provide some position information in the overlap region where there are fewer measurements from the other systems. The CSCs cover a pseudorapidity region of $|\eta| < 1.6$.

If measuring the muon momentum with the muon system only, the relative resolution is about 9% for the central η region and $p_T < 200\,\text{GeV}$. It is limited by multiple scattering. At $p_T = 1\,\text{TeV}$, the momentum resolution varies between 15% and 40%, depending on $|\eta|$. Combining information from the muon system with information from the inner tracker improves the momentum resolution substantially and yields resolutions of around 1% or better for central muons with $p_T < 100\,\text{GeV}$ and resolutions between 5% and 10% for $p_T = 1\,\text{TeV}$, depending on $|\eta|$.

2.2.5 Trigger

The task of the trigger system is to reduce the event rate to a manageable level. Because of the high bunch crossing frequency of 40 MHz, event filtering has to be done at a very early stage in order to reduce the data rate from the detector. Therefore, the first trigger level (Level-1 or L1 Trigger) is implemented in hardware.

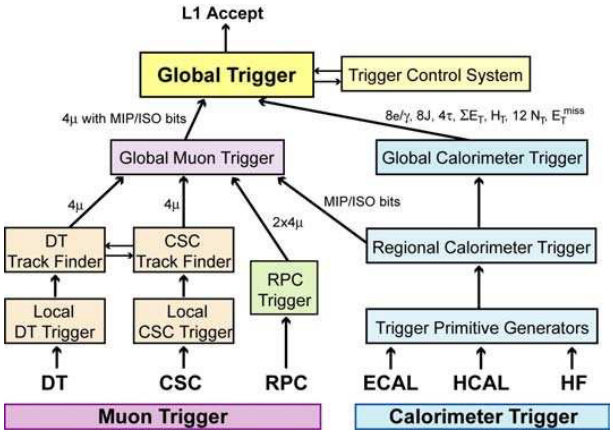

Figure 2.10: Architecture of the L1 Trigger of CMS [31].

However, it has only very limited capabilities of combining information of different detector elements. Therefore, a second trigger step, the High-Level Trigger (HLT) is introduced which is implemented in software and run on usual computer hardware and uses similar or same algorithms as are used for full event reconstruction.

The L1 Trigger architecture is shown in figure 2.10. At the bottom, the Local Triggers based on energy deposits in calorimeter trigger towers and hit patterns in muon chambers are produced. These *trigger primitives* are propagated to the regional triggers which combine the information of several trigger primitives to calorimeter and muon objects. They are passed to the global muon and calorimeter triggers which rank the muon and calorimeter objects and pass the highest-rated objects to the Global Trigger, the top entity of the L1 Trigger hierarchy. It decides whether to read out and pass the event to the High-Level Trigger (HLT) or to reject the event.

The L1 Trigger output rate will be about 50 kHz, while HLT output rate will be about 150 Hz.

2.2.6 Computing and Software

More Information about computing in CMS can be found in [46].

The computing centers used to reconstruct and analyze the data are distributed around the world. They are organized in three layers (Tier-0, Tier-1 and Tier-2) which provide different resources and services.

To better understand the roles of the different Tier layers, one has to know the different data format in which an event is represented:

1. The RAW data format saves the “raw” detector information (e.g. individual pixel hits).

2. The RECO format contains all reconstructed objects (see chapter 3). It is too large for frequent transfers and contains much more information than most analyses will need.
3. AOD ("Analysis Object Data") contains a subset of RECO sufficient for most analyses.

More recently, there are plans that all high-level analyses should be based on yet another format provided by *Physics Analysis Toolkit* (PAT) which uses data from AOD as input. It does some data processing common to most analyses and conversion in the own format.

Those data formats and the reconstruction algorithms are implemented in the CMS software framework (CMSSW). It has a very modular approach and is based on the event data model (EDM) which treats every event independently.

The Tier-0 center is located at only one site, CERN. It accepts RAW data from the detector and does a first calibration and reconstruction. Based on immutable trigger information, events are divided into *primary datasets*, archived at the Tier-0 and distributed to at least one Tier-1 center such that each event is saved at least twice.

There are currently eight Tier-1 centers. They receive data from the Tier-0 and provide archiving of the RAW data. They also provide extensive CPU power for re-reconstruction, calibration, AOD extraction and other common processing steps. The Tier-1 centers redistribute the events to the smaller Tier-2 centers.

The numerous Tier-2 centers provide CPU and storage resources for the individual analyses, such as the one presented here. From the perspective of a user doing an analysis, the part of the workflow which involves those computing resources can be summarized as follows:[11]

1. Identify the datasets needed for the study. That comprises all signal and background processes for Monte-Carlo, and the primary dataset/trigger for data. Most Monte-Carlo datasets are produced centrally and reside on one or more Tier-2 centers. Information about the available datasets and the Tier-2 centers at which they reside can be found in the multiple sources. For example, the Dataset Bookkeeping System (DBS) can be used. Detailed data on parameters for Monte-Carlo generated data is often not available in DBS. In this case, information is collected on a CMS TWIKI page (for example [47], [48]), or on a website recommended by the Physics Analysis Group, such as [49].
2. Write a CMSSW module which runs on AOD (or PAT) format and already do a first selection of events ("pre-selection"). This modules writes the pre-selected

[11] There are many different possibilities of how to organize an analysis. The one discussed here is the one used for the present studies.

events to an output file.

3. Send the program which runs the module to the grid ("grid job"). As events are treated independently, parallelization is trivial: one only has to send multiple grid jobs such that each job processes a subset of events and merge the results afterwards. Splitting data processing in more than one job also increases robustness: one failed job does not affect others. The splitting and job submission is done with a user frontend such as CRAB, the official CMS program for this task. Due to major issues with earlier versions of CARB, an alternative tool has been developed in Karlsruhe, *grid-control*. Job debugging is difficult as job submission is a black-box operation: the grid job is sent and the job output can only be retrieved only after some hours, sometimes days. Managing data samples and running grid jobs is the single most time consuming overhead for a CMSSW physics analysis which is based on data from the grid.
4. Copy the output data to a local machine.
5. Analyze the data. Typically, the first step is to do some cross checks and applying a tighter selection. The rest is very analysis specific.

Of course, those steps are iterated many times.

As CMSSW changes frequently, it has proven useful to use an own data format, beginning as output of step 2. This way, the analysis code in step 5 can be used unchanged even for major changes in CMSSW. Only the "data retrieval module" of step 2 has to be adapted to the new software release. There are more advantages such as a smaller event data format and therefore faster copying and processing.[12]

However, there are some disadvantages using this approach as well: If the same data is to be shared among different users at the analysis level, they either have to agree on one format or resort to one of the official formats RECO/AOD/PAT.

[12] If processing time is bound by IO operations and not by CPU power. This is the case for many analysis steps.

3 Reconstruction

In the previous chapter, the CMS detector and its sub-detectors were discussed. On the raw detector response of a muon for example, algorithms are applied to combine individual hits in the muon chambers to track segments and further to tracks which are then linked to information from the calorimeters and the inner tracker. This yields an estimate for the momentum of the muon and parameters for the muon's track. Such *reconstruction* is done for other physics-objects like electrons, jets, and missing transverse energy, $\not{E}_T$.

In this chapter, track and vertex reconstruction is discussed, followed by reconstruction algorithms for muons, electrons, jets and $\not{E}_T$. Finally, reconstruction techniques specific for top quark physics are discussed, namely top jet reconstruction and $t\bar{t}$ event reconstruction.

3.1 Tracks

Charged particles from the interaction region leave hits in the tracker. To reconstruct the trajectories or *tracks* of those particles is the purpose of the reconstruction algorithm presented here [32]. It is divided in 5 steps which will be discussed in more detail below:

1. Hit reconstruction from strip and pixel information
2. Seed generation
3. Track building
4. Ambiguity resolution
5. Final track fit

Hit reconstruction The hit reconstruction for strips starts with gain-calibrated and zero-suppressed strip data. It uses strips with a signal-to-noise ratio S/N with $S/N > 3$ as seed and clusters adjacent strips if they fulfil $S/N > 2$. A cluster is only kept, if its total S/N exceeds 5. For each cluster, its position, including the error, is determined.

For pixel hit reconstruction, pixels with $S/N > 5$ are used and clustered with adjacent pixels if they have $S/N > 5$. It is kept if the total S/N exceeds 10.1. Pixel cluster position and error are determined.

The reconstruction efficiency for hits can be determined using Monte-Carlo and is above 99.5 %. The fraction of ghost hits (hits reconstructed that were not simulated) is less than 0.01 %. Those values can also be checked with data-driven methods and studies are already ongoing to use data from cosmic muons for tracker and Monte-Carlo validation [50, 51].

Seed generation Seeds are used as starting point for the trajectory building. They are first estimates for the five track parameters, including errors. Without further constraints, three hits are needed for each estimate. Using the reconstructed vertices as constraint on the track parameters, two hits are sufficient.

Track building Starting with the seeds, tracks are built iteratively by propagating the track to the next layer, including effects of multiple scattering and energy loss in the material. In the region predicted by the extrapolation, hits are searched and the track parameters are re-estimated using the new hit information with a Kalman filter [52]. If more than one hit can be found, a trajectory candidate is created for each compatible hit. If no hit is found in the current layer, an "invalid hit" is inserted in the track and the extrapolation continues to the next layer.

In order to avoid biases and exponential growth of track candidates, the tracks are grown in parallel and the list of tracks is truncated after each layer using their normalized χ^2 and number of invalid hits.

The algorithm can be tuned to limit CPU time consumption. For example, in the default configuration, no track candidates with invalid hits in two consecutive layers are considered. Also, during HLT, pattern recognition can stop after five or six hits, even if there are much more in order to save CPU time.

Ambiguity resolution In the algorithm described above, more than one candidate can be found for only one seed or the same track can be reconstructed starting at different seeds. Therefore, a track is removed if there is a "better" track which shares at least 50% of the hits. A track is "better" if it has more hits; if the tracks have the same number of hits, only the track with smaller χ^2 is kept.

Final track fit The Kalman filter used in the track building step updates track estimation using only the current estimation and the new hit; in this sense, it is "local". A global fit which is performed in this step gives a better estimate for the track parameters.

3.2 Vertices

Vertex reconstruction [32] is done in two steps: grouping tracks into vertex candidates, *vertex finding*, and finding the best estimate for vertex parameters, given the tracks associated to the vertex, *vertex fitting*. Which vertex finding algorithm is used, depends on the physics use case and the requirements.

For HLT, a fast vertex finding algorithm is needed. Similarly to track seed finding, pixel hits are clustered into groups of three, *pixel triplets*, and treated as track candidates, *tracklets*. After estimation of the tracklets' parameters, they are clustered based on their z coordinate at the point of closest approach to the beam line (z_0). The clustered tracks are then used to perform the vertex fit. The vertex with the highest sum p_T^2 of associated tracklets is tagged as the *primary vertex*.

For higher accuracy, fully reconstructed tracks can be used to find the primary vertex. Tracks are preselected to have a transverse impact parameter compatible with the beam line, and $p_T > 1.5\,\text{GeV}$. They are then clustered according to z_0, rejecting incompatible tracks. The clusters are fitted to get the vertex parameters.

Primary vertex resolution depends on the tracker alignment and physics process. For a perfectly aligned tracker and $t\bar{t}$ events, the resolution is 13 µm for the x coordinate and 18 µm for the z coordinate.

3.3 Muons

For details of the muon reconstruction, please refer to [53].

As discussed in section 2.2 and figure 2.5, muons leave hits in the tracker, almost no energy in the calorimeters and are the only particles typically passing the muon system. There are several strategies to identify and reconstruct muons:

1. Using information from the muon system only yields so-called *standalone muons*.
2. Using tracker information for reconstruction and hits in the muon system only for identification yields *tracker muons*.
3. Combining the information from both tracker and muon system to perform the reconstruction and identification yields *global muons*.

Before reconstructing those objects, a so-called "local reconstruction" is performed which produces a *track segment* for each DT and CSC chamber.

Standalone muon reconstruction The standalone reconstruction uses the information of all three muon systems, DT, CSC and RPC (see section 2.2.4).

Tracks are built using different steps, similar to those used for track reconstruction in the inner tracker:

1. Seed generation
2. Track building
3. Track cleaning (or ambiguity resolution)
4. Final track fit

The track segments from local reconstruction are used as track seeds. For track building, the current track state is propagated to the next layer and updated with information from the new track segment. Muon propagation takes into account muon energy loss, multiple scattering and the (non-uniform) magnetic field. If no track segment is found, it is further extrapolated to the next layer. The lower quality of track pairs sharing a large fraction of hits is removed in the track cleaning step. As final step, the Kalman filter is applied backwards and the track is extrapolated to the interaction region.

Global muon reconstruction The global muon reconstruction uses the standalone muons and the tracks from the inner tracker as input.

For each standalone muon, a "region of interest" for tracks is defined in order to reduce the set of track candidates. It is a rectangular region in η-ϕ-space. The parameters of the region of interest are derived from the primary vertex and the parameters from standalone muon reconstruction.

In order to match the standalone track to one inner track, the two tracks are propagated to a common surface where the track parameters are compared. Which surface to choose depends on the dominant errors, for both, the track parameters for the tracks and the errors made by track propagation. It is chosen differently for low and high p_T muons. Possible choices are the outer surface of the tracker, the inner surface of the muon system or somewhere in between.

After propagation to a common surface, track parameters of the two tracks are compared. For the best matching tracker tracks, a global fit of the hits in the tracker and the muon system is performed. If more than one tracker track was chosen, only the fit with the best global χ^2 is retained such that for each reconstructed standalone muon, there is at most one reconstructed global muon.

Tracker muons The algorithms presented so far all start at the muon system and therefore rely on reconstructed track segments in the muon chambers. To identify muons with a low number of hits in the muon system (for example, low p_T muons), a complementary approach has been developed, starting with tracker tracks and searching for compatible hits or individual track segments in the muon chambers. In this case, no global track fit is performed but the momentum is taken from the tracker track.

The muon candidate is propagated to the muon chambers and the crossed chambers are recorded. This list can be compared with the list of standalone muon track segments and based on that, some criteria can be defined for muon identification.

There are some additional steps performed for all reconstructed muons which are useful for better muon identification. For example, the muon track is propagated through the ECAL and HCAL and the energy deposited in the ECAL crystals and HCAL towers is recorded.

In many analyses, including this one, it is important to correctly identify *prompt* muons, i. e. muons from the hard interaction, in contrast to muons from the decay of b-mesons or kaons and pions. As those non-prompt muons are accompanied by nearby hadronic activity, the energy deposited in a cone around the muon track is usually much larger for this type of muons. This property is called *isolation*.

The isolation algorithm sums up all energy deposited in the ECAL and HCAL and all tracks' momenta within a cone of $\Delta R \leqslant 0.3$ in the $\eta - \phi$ plane around the muon track. A small inner cone called "veto cone" is excluded to prevent using energy deposited by the muon itself. The radius of this vecto cone is $\Delta R \leqslant 0.07$ for the ECAL and $\Delta R \leqslant 0.1$ for the HCAL.

The V+jets group of CMS proposes a muon selection based on a study of muon identification [54] for prompt muons as follows:[55]

1. Use global muons where the global track fit has $\chi^2/\text{ndof} < 10$.
2. The transverse impact parameter d_0 with respect to the beamspot has to fulfil $|d_0| < 0.2\,\text{mm}$.
3. The number of hits in the tracker has to be $\geqslant 11$.
4. The energy deposited in the veto cone has to be $< 6\,\text{GeV}$ in the HCAL and $< 4\,\text{GeV}$ in the ECAL.
5. For the isolation energies determined as described above in the cone $\Delta R \leqslant 0.3$ (iso_{hcal} and iso_{ecal}, for the sum track momenta $\text{iso}_{\text{track}}$) compute the "combined relative isolation" as

$$\text{CombRelIso} := \frac{\text{iso}_{\text{hcal}} + \text{iso}_{\text{ecal}} + \text{iso}_{\text{ecal}}}{p_T}. \tag{3.1}$$

The cut on CombRelIso is analysis dependent, a typical cut is $\text{CombRelIso} \geqslant 0.1$. For this analysis, a modified muon identification is used, as discussed in section 4.2.

3.4 Electrons

As electrons do not play a central role in the present studies, electron reconstruction [32, 56] is reviewed only briefly.

Electron reconstruction uses information from the inner tracker and the ECAL. Direct energy measurement of an electron with the ECAL is complicated by the fact that an electron traversing the inner tracking system radiates bremsstrahlung. The amount of tracker material varies from 0.4 to over 1.6 radiation lengths, depending on $|\eta|$ (see figure 3.1). About 35% of the electrons radiate more than 70% of their initial energy in the inner tracker. The bremsstrahlung photons deposit their energy in the ECAL. However, their direction in ϕ does not coincide with the electron ϕ direction because the electron trajectory is bent by the magnetic field.

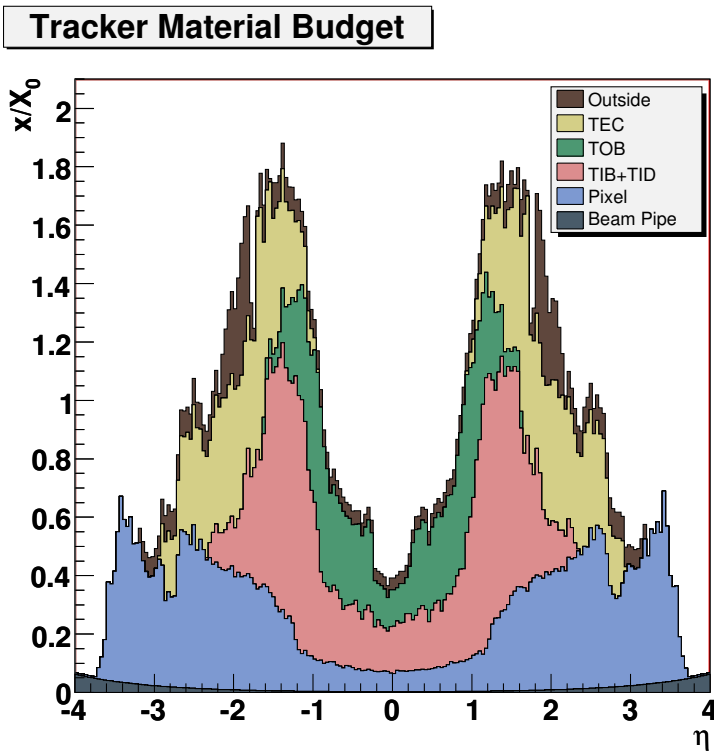

Figure 3.1: Tracker material budget in units of the radiation length X_0 [31].

Electron Candidate Generation To collect both, the main electron shower and the bremsstrahlung photons spread in ϕ, dedicated clustering algorithms are applied to hits in ECAL crystals, yielding so-called *superclusters*.

Using the reconstructed superclusters, a dedicated track finding algorithm, involving a *Gaussian sum filter* (GSF) is used [57]. The Kalman filter usually used for track reconstruction works well for Gaussian noise contribution to the track state. The current track state is given by the track parameters and their errors which are assumed to be Gaussian. As bremsstrahlung leads to highly non-Gaussian tails in

the position prediction for the electron for the next layer, a better approach in this case is to use sums of Gaussians as the error model for the track state.

The track finding algorithm also proceeds in the five steps outlined in section 3.1 but uses different algorithms for seed generation, track building and final track fit. Seeds consisting of two pixel hits are generated by using the supercluster energy and direction information which is propagated backwards to the innermost layers where two matching pixel hits are searched. Track building and final backward fitting is similar as for general tracks described in section 3.1 but with a different energy loss model and using the GSF.

Electron Identification Once electron candidates have been found as described, additional cuts are applied to reject pions or other fakes. At start-up, simple cut-based identification criteria are used as they are relatively simple to understand and check in data. Depending on the analysis, an identification with high purity and low efficiency or with high efficiency and low purity might be preferred.

An identification widely used for start-up scenarios for the high purity use case applies the following cuts:

1. $H/E < 0.015$ (for the barrel region) or $H/E < 0.018$ (in the endcap region) where H is the energy in the HCAL and E the energy in the ECAL;
2. $\Delta\eta < 0.0025$ (barrel) or $\Delta\eta < 0.0040$ (endcap), and
3. $\Delta\phi < 0.02$ (both barrel and endcap), where $\Delta\eta$ and $\Delta\phi$ are the distance in η and ϕ between the electron track extrapolated to the ECAL and the position of the supercluster;
4. $\sigma_{\eta\eta} < 0.0092$ (barrel) or $\sigma_{\eta\eta} < 0.025$ (endcap), where $\sigma_{\eta\eta}$ is the η-width of the 5×5 crystal matrix [58].

This identification is sometimes called "tight Fixed Threshold Identification". The cut values are different for the barrel ($|\eta| \leqslant 1.479$) and endcap region as the ECAL layout changes significantly.

Other methods for electron identification exist, such as classifying the electron candidates into different categories and applying category-dependent cuts [32]. Also, more complex approaches exist, which are based on multivariate techniques such as likelihood [59] and artificial neural networks [60].

3.5 Jets

Jet reconstruction aims to identify the narrow cone of hadrons from quark or gluon hadronization (see section 1.3). Ultimately, jet reconstruction aims to deliver an estimate of the four-vectors of the final state quarks and gluons of the hard interaction.

In this sense, they try to "reverse" the hadronization and detector response to get information of the hard matrix element as depicted in figure 3.2.

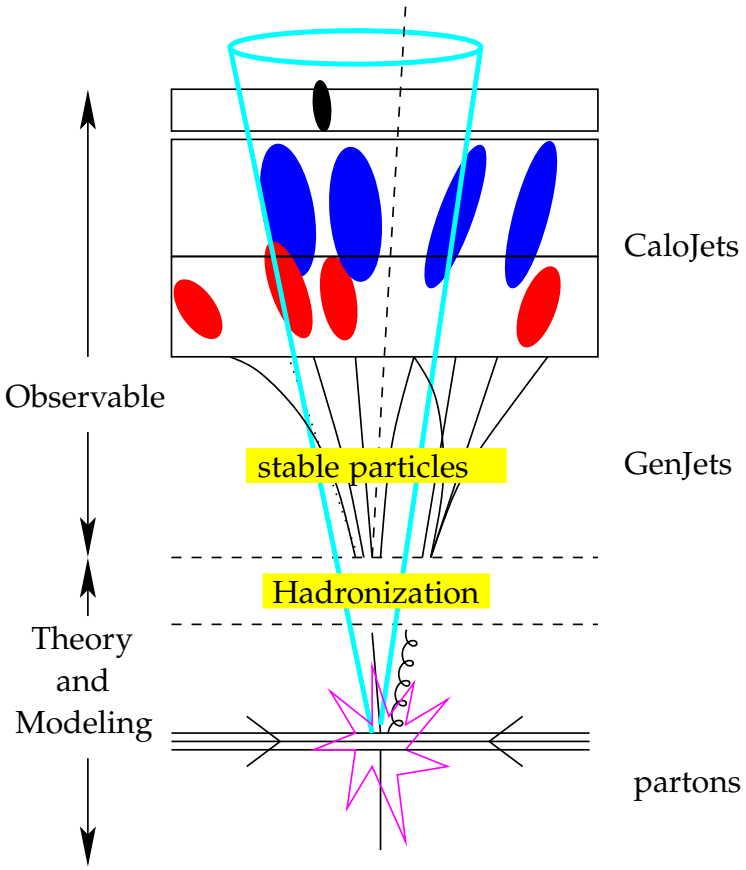

Figure 3.2: The goal of the jet reconstruction is to "reverse" the hadronization in order to measure the momenta of quarks and gluons [61]. This depicts the reverse of figure 1.1.

Jet algorithms are formulated such that they can be applied to any set of four-vectors, *pseudo-particles*, as input. The advantage of this abstract definition is that the same algorithm can be run on simulated partons, stable generated particles after hadronization and decay, as well as detector objects like calorimeter hits. This allows for a comparison between jets without introducing artefacts by using different algorithms. The output of a jet algorithm are the jets, i. e. a set of disjoint subsets of the input set.

An obvious desirable property of a jet algorithm is a good performance in the original goal of estimating the quark and gluon momenta for a large kinematic region. Two non-trivial desired properties of a jet algorithm are:

- It should be *infrared-save*: given a fixed input collection, an additional soft four-vector in the input should not change the number of jets. This requirement is motivated by the large probability of additional soft "infrared" gluon emissions in a hadron shower, and the possibility of additional soft contributions from the underlying event.
- It should be *collinear save*: splitting a four-vector in the input into two adjacent four-vectors (which add up to the original one) should yield the same jet clustering. This ensures that an additional collinear splitting in the shower ($G \to GG$ or $G \to q\bar{q}$) does not alter the jet clustering result.

3.5.1 Cone Algorithms

The iterative cone algorithm [62] is an instance of broad range of cone algorithms. Jets produced by such algorithms contain all pseudo-particles in a cone of fixed size in the η-ϕ plane.[1]

The configuration parameter for this cone algorithm is the cone radius in the η-ϕ plane, R_0 and a seed threshold $E_{T,0}$.

The algorithm finds stable cones in the event. "Stable" means that the jet axis (a four-vector) conincides with the sum of the four-vectors of its constituents. It starts with the pseudo-particle which has the largest E_T above the threshold $E_{T,0}$ as an initial estimate ("seed") for a cone axis $\vec{a}_0$. Then, a new estimate for the cone axis, $\vec{a}_{n+1}$ is calculated by taking the sum of the momenta of all pseudo-particles in the cone centered around $\vec{a}_n$ with radius R_0. This step is repeated until the axis converges. The pseudo-particles contained in this cone are removed from the input collection and the procedure is repeated until no seeds are left.

Jets constructed this way consist of all particles in the cone around its central axis. However, it might happen that after removing the pseudo-particles from the first jet, a nearby jet is found whose cone overlaps with the cone from the first jet. In this case, the jet shape of the second jet in $\eta - \phi$ is not a disk but the difference of two disks.

The algorithm has the advantage that it is very simple and fast. Therefore, it is used in the High Level Trigger in CMS [62]. However, if seeds are defined via a threshold, it is neither collinear-save nor infrared-save.

To see the non-infrared-safety, consider an event with two pseudo-particles. One can find a configuration in which there are three stable cones in this event: the two cones containing only one particle each and a cone containing both particles. Furthermore, there are configurations for which a cone algorithm only finds the two jet consisting of one particle. Now, if there is an additional (soft) particle between the

[1] As two jets must not contain the same pseudo-particle, this is not exactly true as will be discussed below in more detail.

two hard particles, this additional particle can serve as seed and the third jet is found. Therefore, algorithms for which such an example is found, are not infrared-save.

One approach to fix this behavior is to seach for *all* stable cones in the event (for the event discussed, that would include all three jets). Once all stable cones are found, a split-merge procedure is run on those jets which ensures that the jets do not share pseudo-particles. If choosing the split-merge procedure carefully, such an algorithm is infrared-save and collienar-save. [63]

The main problem remaining is finding provably all stable cones in the event. This is non-trivial as the algorithms have to run on large input sets and straight-forward solutions have exponential runtime behavior.

One idea is to insert a seed between each pair of particles. While this works for the example discussed, this algorithm known as "midpoint cone algorithm" [64] is not infrared-save either. Attempts to extend this idea to more than two points to restore infrared-savety have failed [63].

In the present studies, the *SIScone algorithm* (seedless infrared-save cone algorithm) is used. It provably finds all stable cones in the events with a runtime behavior in $\mathcal{O}(N^2)$ for N input particles and applies an appropriate split-merge procedure. [63]

3.5.2 Cambride-Aachen Algorithm

Besides cone algorithms, there is another class of algorithms, *sequential clustering algorithms*. They all follow the same scheme:

1. Find two pseudo-particles a, b with minimal distance $d(a, b)$. If the minimal distance is larger than a cutoff d_0, all remaining pseudo-particles are the jets and the algorithm has terminated.
2. Otherwise, replace the pseudo-particles a and b by one with momentum $p = p_a + p_b$.
3. Go to step 1.

The only configuration parameter for such jet algorithms is the cutoff distance d_0. The definition of the distance d leads to different jet algorithms.

The distance used in the jet algorithm should be invariant under boosts in z-direction and rotation about the z-axis. Otherwise, two processes connected by a Lorentz transformation on the parton-level would be reconstructed differently.[2] The Euclidean distance in the $\eta - \phi$ plane does not fulfil this requirement. Rather, the quantity *rapidity* y is introduced which is defined as

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

[2]For boosts along or rotations around the x and y-axes, an invariance is not necessary because they do not transform a process into another possible process.

Differences in y are invariant under boosts in z-direction. The rapidity y coincides with the pseudorapidity η for particles with zero mass.

The simplest choice is to use the distance in the $y - \phi$ plane which fulfils the requirements just mentioned:

$$d_{y\phi}(a, b) = \sqrt{(\phi_a - \phi_b)^2 + (y_a - y_b)^2}.$$

This is called the Cambridge-Aachen jet algorithm.

Other well-known distance definitions are

$$d(a, b) = \min(p_{T,a}^{2p}, p_{T,b}^{2p}) d_{y\phi}(a, b)$$

which yield the k_T algorithm for $p = 1$ and the anti-k_T algorithm for $p = -1$.

In each iteration, the Cambridge-Aachen algorithm combines two four-vectors to one. Reversing this combination can be used to find substructure of the jet. As will be discussed in section 3.6, this has interesting applications.

3.5.3 Input Sets for Jet Algorithms

Jets are determined by the jet algorithm and the type of the input pseudo-particles. At CMS, calorimeter towers, particle flow objects, or generated particles are used as input objects.

Calorimeter Towers [32] are constructed from hits in the electromagnetic calorimeters and hadron calorimeters. As the hadron calorimeter has a much coarser segmentation than the electromagnetic calorimeter, corresponding ECAL cells are added to match the HCAL geometry. This yields 4 176 calorimeter towers which are used as input for the jet algorithm. The resulting jets are called *CaloJets*.

Particle Flow [65] builds a single list of reconstructed particles including photons, neutral hadrons and charged hadrons. To achieve optimal resolution, the reconstruction combines information from the tracker and the calorimeters. Jets produced with particle flow objects as input are called *PFJets*. They have better energy resolution than jets built from calorimeter towers, especially for low jet energies [66].

Generated Particles (usually all generated stable particles) can be used in Monte-Carlo studies and for comparison with reconstructed jets. Those jets are called *GenJets*.

3.5.4 Jet Energy Corrections

For a fixed true jet energy and direction, the energy recorded in the calorimeters depends on the electromagnetic fraction of the jet, because the calorimeter *response*[3] R is different for the electromagnetic and hadronic part of the shower. Such calorimeter behavior is known as "non-compensating". Moreover, even if considering only one type of particle, the calorimeter response depends on the energy of the particle.

Therefore, jets built from calorimeter towers have to be corrected for the calorimeter response R. They are corrected by scaling the four-momentum with a single correction factor. The correction is factorized into sequentially applied steps to account for different physics effects and dependences of the response function:[67]

1. The *level 1* or *offset* corrections [68] to account for "pile-up" (energy from additional pp-interactions) and electronic noise.
2. After the *level 2* or *relative* jet energy correction, the response R is flat in η. This correction factor is derived by comparing jet response with arbitrary η with jets in the central region. This correction can be derived from data using *dijet balance*, i. e. the fact that in a $2 \rightarrow 2$ QCD process the two jets should have equal p_T [69].
3. The goal of the *level 3* or *absolute* correction is to make jet response R flat versus the true jet p_T. The true jet p_T is defined using a matching GenJet for Monte-Carlo studies and via Z+jet p_T-balance [70, 71] or γ+jet p_T-balance [72] in data.
4. *Level 4,5, 6*, and *7* corrections correct for effects of the electromagnetic fraction (4), the flavor dependence of the response (5), the underlying event (6), and a correction to estimate not the GenJet p_T as was done so far, but rather the parton p_T for the originating parton of the hard process (7). Those corrections are not used in the present studies.

Instead of applying jet energy corrections based on calorimeter information only, it is also possible to use information from the tracker. This possibility of improving the jet energy resolution by using tracker information for jet energy correction is studied for CMS and is known as the *jet plus tracks* algorithm [73].

Jets built from particle-flow objects do need no or only little correction because the input objects for the jet algorithm have already been calibrated [65].

3.5.5 Missing Transverse Energy

For details of $\vec{\not{E}}_T$ reconstruction, refer to [74].

[3] The response R is defined as the ratio of the measured and the true energy. The "jet response" is the corresponding ratio for p_T.

As the initial particles in a hard interaction have no transverse momenta, the sum of all final state particles' $\vec{p}_T$ must vanish. The calorimeter has a high angular coverage and captures almost all energy of the final state particles. Neutrinos leave the detector undetected. Therefore, the residual to zero transverse momentum is the sum of the neutrinos' $\vec{p}_T$ and other weakly interacting particles. This residual is called missing E_T, or $\vec{\not{E}}_T$.

Using calorimeter information, $\vec{\not{E}}_T$ can be reconstructed as

$$\vec{\not{E}}_T = -\sum_{i=1}^{N} E_i \sin\Theta_i(\cos\phi_i\hat{x} + \sin\phi_i\hat{y})$$

where the sum runs over all calorimeter towers. Θ_i and ϕ_i are the coordinates of the calorimeter towers, and $\hat{x}$ and $\hat{y}$ are the unit vectors in x and y direction.

As jets are corrected for the calorimeter response, this must also be taken into account for reconstruction of missing energy. Missing energy is corrected for jet energy scale by

$$\vec{\not{E}}_T^{\,\mathrm{corr}} = \vec{\not{E}}_T - \sum_{i=1}^{N_{\mathrm{jets}}} \vec{p}_{T,i}^{\,\mathrm{corr}} - \vec{p}_{T,i}^{\,\mathrm{raw}}. \tag{3.2}$$

$p_{T,i}$ denotes the jet p_T before or after applying the jet correction. As jets with $p_T < 10\,\mathrm{GeV}$ have a large correction uncertaintya and are not used for $\vec{\not{E}}_T$ correction. Energy deposition in the electromagnetic calorimeter by electrons is well measured and does not have to be corrected. However, jet clustering can also cluster energy depositions from electrons. This would lead to spurious corrections of $\vec{\not{E}}_T$. Therefore, jets above a given electromagnetic fraction are excluded in equation 3.2.

Muons escape the detector leaving only very little energy in the calorimeters. Therefore, $\vec{\not{E}}_T$ also has to be corrected for the muons in the event by subtracting all muons' $\vec{p}_T$ from $\vec{\not{E}}_T$ and adding the energy deposited in calorimeter towers.

The corrections for jet energy scale and muons are called *Type-1* corrections. So-called *Type-2* corrections account for pile-up and out-of-cone effects. It is also possible to include tracker information for $\not{E}_T$ calculation. As for jets, algorithms exist to make use of the tracker information to improve $\not{E}_T$ resolution [75]. Another approach is to use particle flow objects instead of calorimeter towers to reconstruct $\not{E}_T$ which has shown to improve the resolution compared to calorimeter based $\not{E}_T$ for low energies [65].

In the current studies, calorimeter-based $\not{E}_T$ with Type-1 corrections is used.

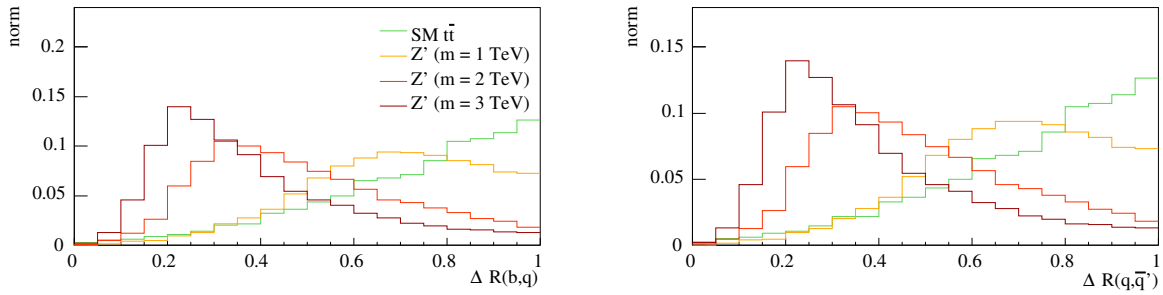

Figure 3.3: The ΔR distance between the b-quark and a light quark (left) and between the two light quarks from the W-boson decay (right) for the hadronically decaying top quark.

3.6 Boosted Top Jet Reconstruction

If a particle with high energy decays into hadrons, those decay products still carry a large component of the original direction of the decaying particle. While all decay products might end up in a single reconstructed jet, calorimeter resolution might still be sufficient to resolve a substructure of the jet originating from the individual decay products. This is only possible if the decay products are separated enough to be resolvable by the calorimeter. The ΔR distance between the decay products of a particle with mass m and transverse momentum p_T is of the order m/p_T. For CMS calorimeter resolution of $\Delta\eta = \Delta\phi = 0.087$ (see section 2.2.3), and a top quark as decaying particle, this means that p_T up to about 2 TeV could be resolved. Of course, this is only a rough estimate as it assumes a constant ΔR and requires to identify large energy depositions in two adjacent calorimeter towers as substructure. For a more complete picture, see figure 3.3 which shows the ΔR distribution between the decay products of a hadronically decaying top quark.

The search for this substructure can be achieved by reversing the sequential clustering of the Cambridge-Aachen jet algorithm (section 3.5.2) yielding "subjets" of the original jet.

This idea can be applied for different analysis. For example, it can be used to search for high-p_T Higgs bosons decaying to a b-quark pair [76].

Here, an algorithm is presented which aims to reconstruct hadronically decaying top quarks with a high boost [10]. This algorithm was implemented in the CMSSW framework where it is called `CATopTag` Algorithm [77]. This is the algorithm described here.

The general idea is to use the cluster sequence of a sequential clustering algorithm to find "subjets" from the decay products of the top quark. In the case of a top quark,

three subjets are expected: a b-subjet and two subjets from the *W*-boson decay.

The algorithm proceeds as follows:

1. Cluster the input particles with the Cambridge-Aachen algorithm with the distance cutoff $d_0 = 0.8$. Those are called "hard jets".
2. Require that the hard jets are hard and central: $p_T > 250\,\text{GeV}$ and $|y| < 2.5$.
3. Reverse the clustering steps of the Cambridge-Aachen algorithm as follows:
   a) Decompose the hard jet in two constituents which both fulfil $p_T^{\text{cluster}} > \delta_p \times p_T$, with $\delta_p = 0.05$. This p_T requirement rejects soft subjets which are not expected in the signal. If reversing the first clustering step does not yield two subclusters fulfilling this criterion, the de-clustering is repeated for the harder of the two subclusters. This is iterated until a valid decomposition is found or proven to be impossible.
   b) Decompose the two clusters from the previous steps again, by using the same algorithm applied to the two clusters as the "hard jets".

   This declustering yields between 0 and 4 "subjets".
4. To allow for an interpretation of the subjets as the decay products of the top quarks, the jet is required to have at least three subjets. In case of four subjets, the subjet with lowest p_T is interpreted as additional soft radiation. The following kinematic cuts are applied:
   a) The invariant mass of the jet of the hard jet must be consistent with the top mass: $100\,\text{GeV} < m_{\text{hard jet}} < 250\,\text{GeV}$.
   b) The invariant dijet masses built with the three subjets with largest p_T must fulfil $m_{\text{dijet}} > 50\,\text{GeV}$ for all three possible combinations.

The original algorithm [10] uses the scalar sum of the transverse energy in the event, E_T, as first estimate of the top quark p_T. For higher p_T, the decay products are expected to be closer in ΔR. Therefore, the parameters d_0 and δ_p of the top jet algorithm are chosen dependent on this E_T.

In the implementation of this algorithm in CMSSW, this E_T-dependent choice of parameters is the default configuration [78]. The event's E_T is calculated by summing all calorimeter towers. While this is adequate for the all-hadronic decay channel, it does not account for the muon and the neutrino ($\not{E}_T$) in the muon+jets channel and makes a direct comparison for different channels more difficult.

Therefore, it was suggested to change this dependence for an easier application of this algorithm on the muon+jets channel. As shown by the authors of the implementation of this algorithm [77], choosing the fixed values $d_=0.8$ and $\delta_p = 0.05$ still yield a high efficiency and therefore, those values are now generally used. As can be seen in figure 3.4, the efficiency is about 40 % for high p_T of the top quark.

This algorithm can be applied to other sequential cluster algorithms such as k_T and anti-k_T. However, decomposition of those jets yields subjets which are closer to

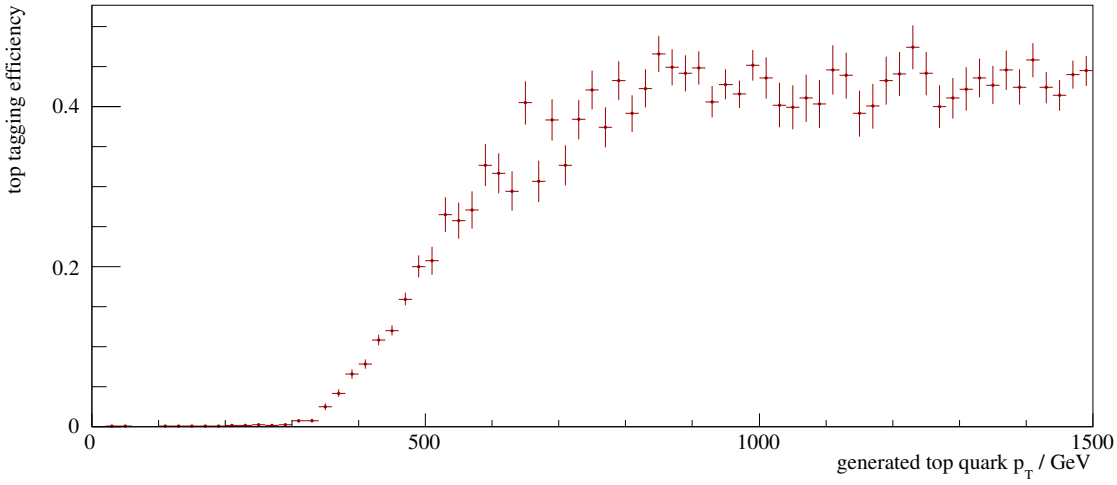

Figure 3.4: The top jet reconstruction efficiency as function of the generated top quark p_T. For this figure, all Z' and standard model $t\bar{t}$ in the semileptonic μ decay channel were used, after applying the loose pre-selection. The samples and the selection are discussed in more detail in section 4.1 and section 4.2, respectively. The efficiency plotted here is defined as the number of events with one tagged top jet divided by all events, as function of the generated top-quark p_T of the hadronically decaying top quark.

the axis of the hard jet in ΔR and leads to worse discrimination for the dijet mass criterion [77].

For the present studies, the algorithm is applied as described to Cambridge-Aachen jets built from calorimeter towers. It is used only for the event selection discussed in section 4.5. For the rest of the analysis, SIScone jets built from calorimeter towers are used.

One problem that arises both for the top jets but also for the SIScone jets in the context of boosted top-quark jets are jet energy corrections. The relative and absolute corrections are derived from QCD dijet samples which mainly contain jets from light quarks or gluons. Therefore, they are not applicable to jets originating from top quarks. One might be tempted to correct the found "subjets" because they originate either from a b-quark or from a light quark from the *W*-boson decay. However, as those subjets are closely merged, it is conceptually not possible to assign the constituents of a subjet to a single parton from the top-quark decay. But this is necessary for the mere concept of a jet correction whose aim is to give an estimate for the original parton four-momentum. Therefore, one should not attempt to correct

the subjets.

Rather, the goal is to estimate the original top-quark momentum only. A correction of this type is conceptually well-defined.Such jet corrections are not officially available in CMS.

To see the effect of jet energy correction for the case that the decay products of the top quark end up in a single jet, events are studied where all decay products of a hadronically decaying top quark have $\Delta R < 0.3$ to one reconstructed SIScone jet. The resulting jet response (uncorrected and with standard corrections) can be seen in figure 3.5. To create the plot, bins in top quark p_T are created such that there are at least 400 entries in each bin. For each bin, the median jet p_T is divided by the top quark p_T to get the response.

As can be seen, applying standard corrections is better than using uncorrected jets. But even with jet energy corrections, the true top quark p_T is underestimated. This is due to (a) energy deposits of the b and light quarks outside the cone and (b) the fact that the applied jet corrections were defined to correct a reconstructed jet p_T to the p_T of the matching GenJet, not a correction on parton level what would be needed here and (c) the neutrino involved in the b meson decay. The latter is connected to a well-known feature of jet energy corrections: as they are derived on a QCD sample which mainly consists of jets originating from gluons, they do not correctly estimate the p_T for other jets. So this problem always arises if the flavor composition under study does not represent the flavor composition of the sample the jet energy corrections were derived for. This difference is accounted for in higher-level jet energy scale corrections (see section 3.5.4).

3.7 $t\bar{t}$ Event Reconstruction

The goal of the $t\bar{t}$ event reconstruction is to estimate the four-vectors of the two top quarks. This allows to determine $m_{t\bar{t}}$ but also the measurement of other quantities such as $|\eta_t| - |\eta_{\bar{t}}|$. This quantity has a symmetrical distribution around zero in the standard model. Any asymmetry would be a pointer to new physics as discussed in [79].

In the following, $m_{t\bar{t}}$ will denote the *true* mass of the system, while $M_{t\bar{t}}$ refers to the *reconstructed* mass of the $t\bar{t}$ system.

The algorithm for $t\bar{t}$ reconstruction presented here consist of two steps: For each event

1. build a list of hypothesis and
2. select a hypothesis from the list

where "hypothesis" refers to a mapping of the top quark decay products to re-

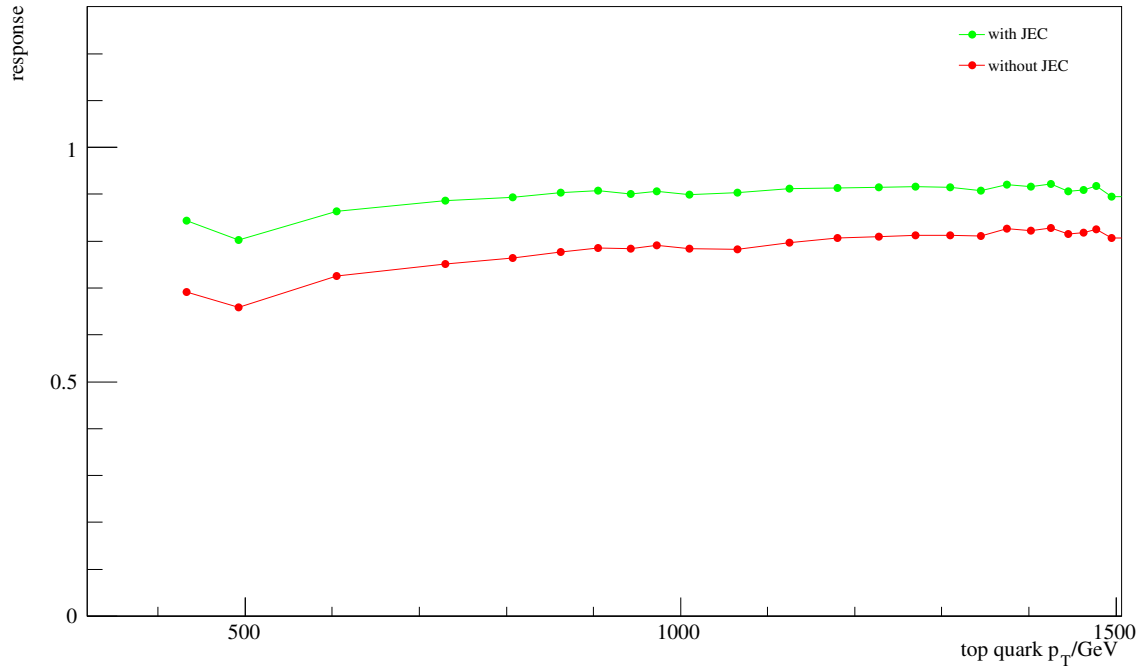

Figure 3.5: Response for high-p_T, hadronically decaying top quarks. The reconstructed jet's p_T versus the true top quark p_T is shown. Jets are SIScone jets with cone radius $\Delta R = 0.5$ which match the decay products of a top quark. Jet energy corrections considered are level-2 and level-3 corrections.

constructed objects.[4] Each hypothesis thus also gives values for the top quark four-momenta and other kinematic variables.

The criteria to select one hypothesis from the list are derived on Monte-Carlo simulations where the *correct* hypothesis can be defined by matching the reconstructed objects to the generated objects. The properties of this correct hypothesis can then be used for hypothesis selection.

Other $t\bar{t}$ reconstruction methods use the known masses for the top quark and the *W*-boson to correct mismeasured jets via a *kinematic fit* where the four-vectors of the reconstructed objects are scaled to fulfil the mass constraints.

The cross section for $t\bar{t}$ production in the standard model is high for low $m_{t\bar{t}}$. In this region, the top quarks can be assumed to be in rest and their decay products will be well separated in the detector. To select events in the muon+jets channel, one usually demands four jets with large p_T, an energetic, isolated muon, and missing

[4] While this map as defined refers to all six $t\bar{t}$ decay products, this is not necessary for the reconstruction of the top quark four-momenta: in this case, association on the level of the two top quarks is sufficient.

transverse energy from the neutrino.

3.7.1 Neutrino Reconstruction

$\vec{\not{E}}_T$ is interpreted as the x and y components of the neutrino momentum. The z component of the neutrino cannot be determined because the initial boost of the parton system is not known. As the neutrino originates from the decay $W \to \mu\nu$, the following equations hold (assuming massless ν, μ and that the W was exactly "on-shell"):

$$\begin{aligned} p_\nu^2 &= p_\mu^2 = 0 \\ m_W^2 &= (p_\nu + p_\mu)^2. \end{aligned}$$

Re-arranging these equations yields a quadratic equation for $p_{\nu,z}$. This equation has either zero, one or two real solutions. The case of exactly one solution has almost zero probability but is straight-forward. For the other cases, several strategies exist.

In the case of no real solution, one could take the real part of the complex solutions. Another method is to interpret this as a mismeasurement of $\not{E}_T$ and rescale $\not{E}_T$ by a factor as close to unity as possible such that there is a real solution. Both methods have been used during the studies presented here. However, the difference in reconstruction performance turned out to be negligible. Therefore, the simpler of the two methods — taking the real part — was chosen.

In case of two solutions, there are also several strategies:

1. take the $p_{\nu,z}$ closer to 0. That reflects the observation that the decay products of $t\bar{t}$ events are central, i. e. $\eta \approx 0$.
2. take the $p_{\nu,z}$ which is closer to the lepton. This is plausible if one assumes that the decaying W-boson has a boost.
3. Defer the decision and make hypotheses with both solutions.
4. There are other possibilities: The procedure for selecting the neutrino solution could depend on some additional information about the two solutions, for example $|p_{\nu,z,1} - p_{\nu,z,2}|$. In such scenarios, Monte-Carlo information is used to tune the selection procedure.

In the present studies, option 3. is chosen, i. e. both neutrino solutions enter the list of hypotheses.

3.7.2 $t\bar{t}$ Reconstruction

The list of hypotheses consists of all assignments of jets to the decaying top quarks. That is, each jet is either assigned to the leptonically decaying top quark, to the hadronically decaying top quark or none of them. The top quark four-momentum

is then simply the sum of the four-momenta of the assigned decay products in the particular hypothesis. For N jets, this yields 3^N different hypotheses. This number is doubled if there are two solutions for the neutrino z component. Hypotheses where no jet is assigned to one of the top quarks are thrown away.

The highest-p_T jet assigned to each top quark is labelled "b-jet". However, it does not matter much whether it really *is* the b-jet or not.

As shown in figure 3.3, the decay products of the top quarks are close together in the detector. This can be used for hypothesis selection by defining the quantity C for each hypothesis which is defined as the sum of ΔR of the top quarks and their decay products:

$$C := \Delta R(p_{t,\mu}, p_\mu) + \Delta R(p_{t,\mu}, p_\nu) + \Delta R(p_{t,\mu}, p_{b,\mu})$$

where $p_{t,\mu}$ is the four-vector of the muonically decaying top quark, $p_{b,\mu}$ is momentum of the b-jet of the muonically decaying top quark.

The value of C is the same for all hypotheses which only differ on the side of the hadronically decaying top quarks. Therefore, an additional rule has to be stated. For now, the hypothesis with the lowest C is selected where all other jets are assigned to the hadronically decaying top quark.

The result of the event reconstruction can be seen in figure 3.6.

Optimizing Hypothesis Selection Since the definition of C does not distinguish between hypotheses which differ only in the jet assignment of the hadronically decaying top quark, the modified quantity C′ is introduced, defined as

$$C' := C - f_1 \Delta R(p_{t,\mu}, p_{t,h}) - f_2 M_{t\bar{t}}$$

where f_1 and f_2 are positive constants to be determined. Taking the hypothesis with minimal C′ will favor hypotheses with larger separation of $t\bar{t}$ and higher $M_{t\bar{t}}$.

The constants f_1 and f_2 can be chosen such that the peak of the reconstructed $M_{t\bar{t}}$ distribution has smaller width and is shifted towards the correct value. For example, studied values are $f_1 = 0.5$ and $f_2 = (0, 0.001, 0.005)$ (for 2, 3, or more jets respectively).

In order to choose optimal values for f_1 and f_2, a measure of the performance for the reconstruction is needed. The influence of background is neglected in the following discussion, as it contributes a continuum in the $M_{t\bar{t}}$ spectrum which is not changed very much by the reconstruction technique.

A possible criterion for optimizing the reconstruction would be minimizing the variance of the reconstructed $M_{t\bar{t}}$ distribution. This seems reasonable, because the generated mass distribution has a very small width, so the width of the reconstructed $M_{t\bar{t}}$ distribution gives an estimate of the reconstruction resolution.

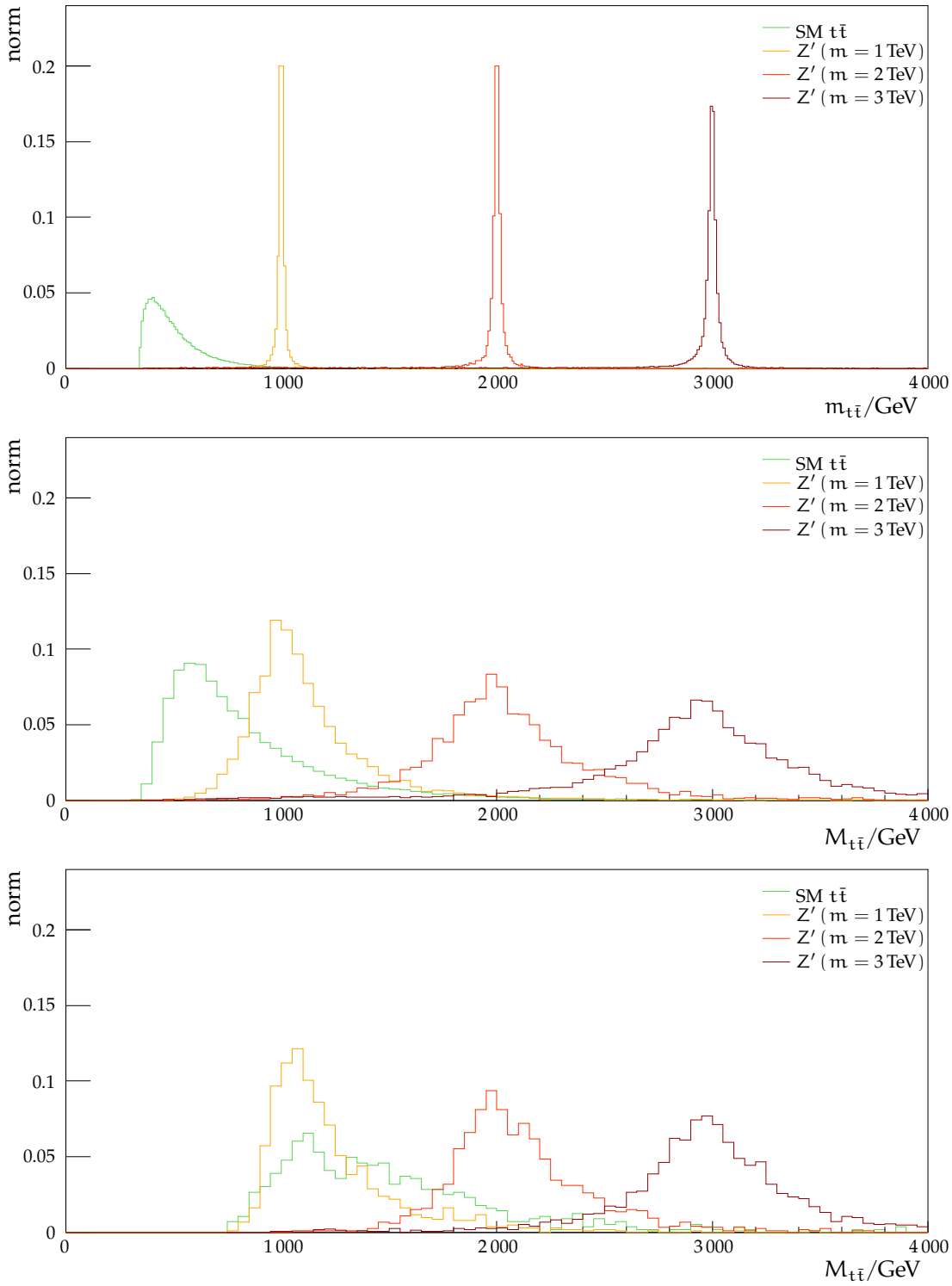

Figure 3.6: Distribution for the generated mass (top) and reconstructed mass (central and bottom) of the $t\bar{t}$ system. The reconstruction uses the hypothesis with smallest C. The central plot was created on a sample with a loose pre-selection (section 4.2), the bottom plot with the final selection (section 4.4.3). As can be seen, the selection has a major impact on the resolution of the mass reconstruction.

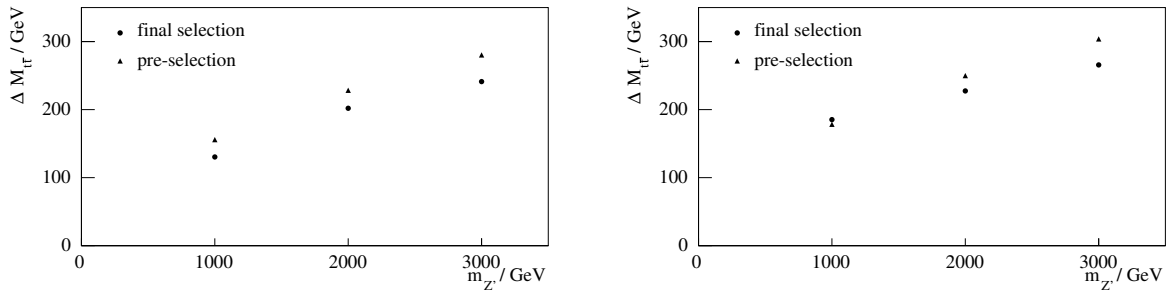

Figure 3.7: Resolution of the reconstruction. The width of a Gaussian fitted to the reconstructed $M_{t\bar{t}}$ distribution is shown versus the true Z' mass. The left plot uses a hypothesis selection with minimal χ, the right one with χ'. The pre-selection and final selection are defined in section 4.2 and section 4.4.3, respectively.

However, using the hypothesis selection for which the $M_{t\bar{t}}$ distribution variance is smaller, does not yield optimal limits on the Z' cross section (for the method of determining limits, see chapter 5). This can be better understood, if one realizes that the limit determination is similar to a fit which is driven mainly by the number of events in the "peak region" while minimizing the variance of the whole $M_{t\bar{t}}$ distribution primarily affects the tails of this distribution as they contribute quadratically to the variance. However, for events in the tails, there is often a better, but rarely a good hypothesis which would contribute to the peak region. Optimizing the $M_{t\bar{t}}$ variance therefore leads to smaller tails in the $M_{t\bar{t}}$ distribution, but the peak region remains largely unaffected.

To circumvent this problem, one could optimize not the width of the whole $M_{t\bar{t}}$ distribution but rather the width of the peak. The peak region can be defined by as $\hat{\mu} \pm 1.5\hat{\sigma}$ where $\hat{\mu}$ and $\hat{\sigma}$ are the values from a Gaussian which is fitted in the same region (of course, this has to be done iteratively). This was done for the resolution plots in figure 3.7. As can be seen, hypothesis selection with χ instead of χ' yields a better resolution. As the definition of χ is simpler, using χ is the preferred method for reconstruction.

Note, however, that the performance of hypothesis selection depends on the jet definition: for the plots shown here, a jet definition was used which requires jet $p_T \geqslant 50\,\text{GeV}$. If relaxing this requirement to lower p_T values, the picture changes. Similarly, the selection has major impact on the resolution, as was already shown in figure 3.6.

However, the impact of the event selection on the final result is much larger than differences between χ and χ' for hypothesis selection. Therefore, more work was

done for optimizing the event selection which is discussed in the next chapter. In the following, the hypothesis selection is always done with C.

4 Event Selection

In this chapter, Monte-Carlo event generation for signal and background processes is discussed in section 4.1. In section 4.2, a loose pre-selection as reference point is defined and the definitions for the physics objects used in this study are reviewed. In section 4.3, a typical selection for a $t\bar{t}$ event topology as expected for standard model $t\bar{t}$ production is applied. It will be shown that the selection efficiency degrades substantially for high $m_{t\bar{t}}$. In section 4.4, a selection for high $m_{t\bar{t}}$ and its optimization is discussed. Finally, section 4.5 briefly discusses the possibility of a selection based on the boosted top jet reconstruction introduced in section 3.6.

4.1 Event Generation and Simulation

As discussed in section 1.3, events at hadron colliders are modelled in separate steps. The integration in equation 1.9 is carried out by choosing values for x_a and x_b according to their parton distribution functions which take the role of a density for the integral. A random final state is chosen according to the final state phase space ($d\beta$ in equation 1.3). The matrix element for this process is evaluated yielding one generated event with a concrete set of final state particles with defined four-momenta. This process is called "event generation".

The final state partons then undergo showering, hadronization, and decay as depicted in figure 1.1. As generation of hard events and simulation of the shower are relatively independent, they can be carried out by different tools:

- MadEvent [80] is an event generator which uses MadGraph [81] to calculate all tree-level Feynman graphs for given initial and final state particles. MadGraph calculates the Feynman graphs using Feynman rules which can be specified by the user. Therefore, physics beyond the standard model can be included easily. For example, the Z' processes used in the present studies have been generated using the MadGraph package `topBSM` [20].
- Pythia [17] is both a tool for event generation and for simulation of shower, hadronization and decay. The shower simulation of *all* Monte-Carlo samples used for the present studies have been made with Pythia.

So far, event generation, hadronization and decay is covered. For each event, the output consists of a set of leptons and stable hadrons. Those particles are passed to the CMS detector simulation which is based on Geant4 [82], a toolkit for the detailed

simulation of the passage of particles through matter, accounting for a large number of different interactions and effects. This type of detector simulation is called full simulation.

For some processes, several million events have to be generated in order to reliably estimate the number of events passing a selection. As processing a single event with GEANT is a very time-consuming process, an alternative tool is implemented known as fast simulation [32]. Instead of detailed simulation of the individual particles interacting with the detector, the detector response is parameterized for all particles using the full simulation. This parameterization can then be applied efficiently to other simulated events.

The result of both the full and the fast simulation are low-level detector objects such as energy deposits in calorimeter cells which can be used as input for the reconstruction algorithms discussed in chapter 3. In some cases like tracks or muons, the fast simulation directly produces reconstructed higher-level objects.

As the different samples, each containing a certain process, have very different cross sections and are produced in different quantities, they have to be re-weighted to a common integrated luminosity L. For a process with cross section σ where a filter is applied with filter efficiency f, the correspoding weight w is

$$w = \frac{L\sigma f}{N}$$

where N is the number of events passing the filter. The combination σf is called *effective cross section*. The weight is a direct measure of the available Monte-Carlo statistics: a unity event weight means that as many events are available as Monte-Carlo sample as are expected in data for the assumed luminosity. For most analysis, an event weight much smaller than one is desirable. Otherwise, the Monte-Carlo sample itself is affected by a statistical error as large as or even larger than the statistical fluctuations expected in data.

The cross sections for the processes discussed in the following sections are leading-order results obtained from the generators, unless otherwise noted. For a next-to-leading cross-section σ_{NLO}, the so-called K-factor is defined as

$$K = \frac{\sigma_{NLO}}{\sigma_{LO}}.$$

In general, the K-factor is a function of the momenta of the final state particles. However, for many cases, taking a constant K-factor is a satisfactory approximation and all generated events are scaled by the K-factor to obtain approximate NLO results.

As discussed in section 1.4, the present studies concentrate on the muon+jets decay channel of $t\bar{t}$ events from high mass resonances. The expected signal signature is a high-p_T muon and at least two jets. In the following sections, background processes with a similar signature are discussed.

4.1.1 QCD Background

At a hadron collider, "QCD processes" dominated by Feynman graphs with strong interaction vertices only, have the largest cross section. In those events, there is no lepton in the final state of the hard scattering. However, even in QCD processes, reconstructed muons can appear:

1. By *heavy flavor* decay: $b\bar{b}$ or $c\bar{c}$ production leads to mesons which can decay to $\mu + X$.
2. By *decays in flight*: $\pi^{\pm}$ and kaons in any hadron shower can decay in the detector volume to $\mu + X$. The μ produces a track in the muon system which is matched to some nearby inner track (for example, the $\pi^{\pm}$ track).
3. By *punch-through*: a high-energy shower can reach the muon system. Those hits might be combined to tracks and matched to an inner track, produced by charged hadrons from the punched-through jet.

The third possibility is extremely rare for global muon reconstruction. Therefore, simulating those is difficult and would require a very good understanding of the detector, including unusual detector effects.

The first two cases have been simulated in the QCD Monte-Carlo sample used for the present studies. The generation starts with the Pythia event generator producing "minimum bias" events, consisting of all $2 \to 2$ processes with q, $\bar{q}$ and G as initial and final state particles.

Pythia generates the events, simulates the shower, hadronization, and decay of unstable particles. For a usual Monte-Carlo production, only very short-lived particles which cannot reach the detector are decayed. For this sample, Pythia is configured to simulate also the decay of long-living kaons and pions within a cylinder around the beam axis with a radius of 1.5 m and length of 6 m. Only events with at least one muon within $|\eta| < 2.5$ and $p_T > 5\,\text{GeV}$ are passed to the detector simulation.

As the cross section for the simulated $2 \to 2$ processes dramatically increases for low $\hat{s}$, a generator level filter is applied. Otherwise, a huge number of Monte-Carlo events would be necessary to have enough statistics after the event selection which does not select events with low $\hat{s}$. Therefore, QCD events are produced for different

sub-sample	cross section (pb)	generated events	weight
$\hat{p}_T > 50\,\text{GeV}$	280 000	7 382 989	7.59
$\hat{p}_T > 150\,\text{GeV}$	5 000	1 300 477	0.77
$\hat{p}_T > 250\,\text{GeV}$	700	1 094 551	0.128
$\hat{p}_T > 350\,\text{GeV}$	112	1 143 532	0.020

Table 4.1: Number of produced μ-enriched QCD events, i. e. event with at least one generated μ with $p_T > 5\,\text{GeV}$ and $|\eta| < 2.5$, for the different sub-samples and their respective effective leading-order cross section. The event weight is given for a luminosity of $L = 200\,\text{pb}^{-1}$.

ranges of $\hat{p}_T$ defined as

$$\hat{p}_T^2 = \frac{\hat{t}\hat{u}}{\hat{s}}$$

with the Mandelstam variables in the parton system $\hat{s}$, $\hat{t}$ and $\hat{u}$. The Mandelstam variables are defined for $2 \to 2$ processes as

$$\begin{aligned} s &= (p_1 + p_2)^2 = (p_3 + p_4)^2 \\ t &= (p_1 - p_3)^2 = (p_2 - p_4)^2 \\ u &= (p_1 - p_4)^2 = (p_2 - p_3)^2, \end{aligned}$$

where p_1 and p_2 are the momenta of the initial state particles, and the p_3 and p_4 the momenta of the final state particles.

In case of filtering on generator level, the event selection has to ensure that events failing those filters on generator level would not have been selected.

The generated QCD samples and event numbers are summarized in table 4.1. The samples have overlapping $\hat{p}_T$ ranges. To avoid double counting, this overlap is removed by applying a filter using the Monte-Carlo information about $\hat{p}_T$.

4.1.2 Electroweak Backgrounds

Apart from QCD, the most important backgrounds are processes with a muon in the final state of the hard scattering. This is only possible for events with a Z, γ, or W boson as part of the hard process. Examples of contributing tree-level Feynman graphs are shown in figure 4.1.

The "W+jets" and "Z+jets" background samples were generated with MadGraph and MadEvent. As discussed in section 4.1.2, there are processes with no or only one hard parton in the final state. Such events will not pass the event selection.

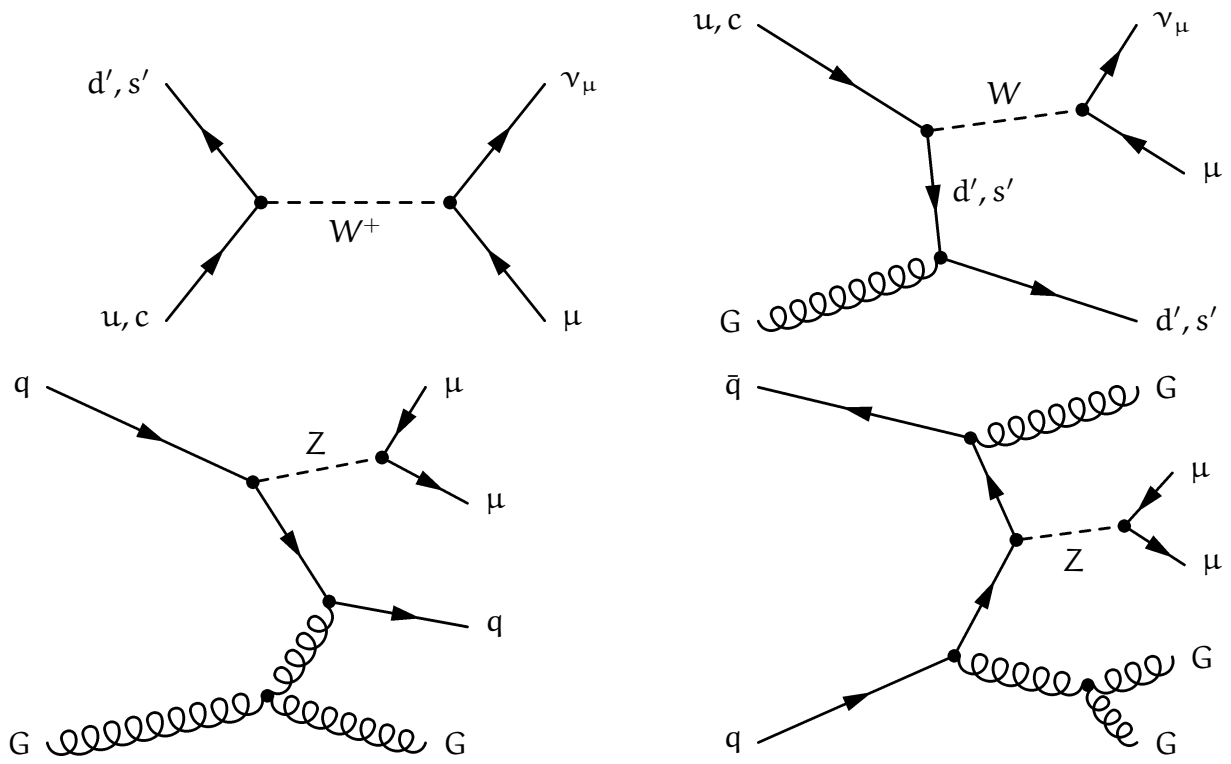

Figure 4.1: Examples of processes with a W or Z boson decaying to one or two muons. The process on the upper left contains no hard parton as final state particle. Therefore, there will be no reconstructed jet and this event will not pass the event selection which cuts on at least two hard jets. The other processes contain more and more hard partons in the final state.

However, they have the largest cross section. Therefore, most of events in a "W+jets" Monte-Carlo sample do not pass the event selection and the number of remaining events is very low even for large samples containing 10 million simulated W+jets events, the size of the centrally produced W+jets sample.

Therefore, a sample with the tenfold number of events has been used. This sample has been processed by the fast detector simulation.

The number of generated events and cross sections are summarized in table 4.2. For W+jets, the cross section was calculated by scaling the leading-order cross section from MadGraph by the K-factor of $K = 1.14$ given in [83].

Only events where the Z or W boson decays to leptons were generated. For processes with a Z-boson, the matrix elements containing a virtual photon have to be included in the calculation. As the cross section for this Z/γ^*+jets process is very large for small $m_{l^+l^-}$, a generator-level cut of $m_{l^+l^-} > 50\,\text{GeV}$ is applied.

sample	cross section (pb)		generated events	weight
W+jets, $W \to l\nu$	40 000	(NLO)	101 467 739	0.079
Z+jets, $Z \to l^+l^-$	3 700	(LO)	9 964 055	0.074

Table 4.2: The cross sections of the W+jets and Z+jets background samples, processed with the fast simulation. The weight is given for $L = 200\,pb^{-1}$. The cross section for W+jets includes the K-factor of $K = 1.14$, the Z+jets cross section is leading order. In the generated events, the Z or W boson decays always to leptons. For Z+jets, a cut on the dilepton-mass of $m(l^+l^-) > 50\,GeV$ has been applied.

sample	NLO cross section (pb)	generated events	weight	K-factor
t-channel, $t \to l\nu b$	41.6	281 756	0.081	1.12
tW-channel, all modes	29.0	169 048	0.030	1.06
$t\bar{t}$, all modes	414.0	1 028 304	0.034	1.29

Table 4.3: Generated $t\bar{t}$ and single top events. For the calculation of the K-factors, refer to [84, 85, 19]. The t-channel sample only contains leptonically decaying W-bosons, while the other samples include all decay modes. The event weight assumes $L = 200\,pb^{-1}$.

4.1.3 Top-Quark Backgrounds

The last category of background processes considered for this studies are processes with top quarks in the final state. $t\bar{t}$ production in the standard model has already been discussed in section 1.4. Feynman graphs for single top production are shown in figure 4.2.

Apart from the processes shown in figure 4.2, single top quarks can be produced in the s-channel, $q\bar{q}' \to W^+ \to t\bar{b}$, and the charge conjugated process for $\bar{t}$ production. This process has a cross section times branching ratio in leptons of about 1.6 pb. (This cross section was obtained by scaling the LO generator result with the K-factor derived in [84].) Compared to the other single top backgrounds, it is negligible and was therefore not included in this studies.

4.1.4 Signal

The signal process used in the present studies is the production of a massive gauge-boson like particle, Z'. The leading order Feynman graph of the production and

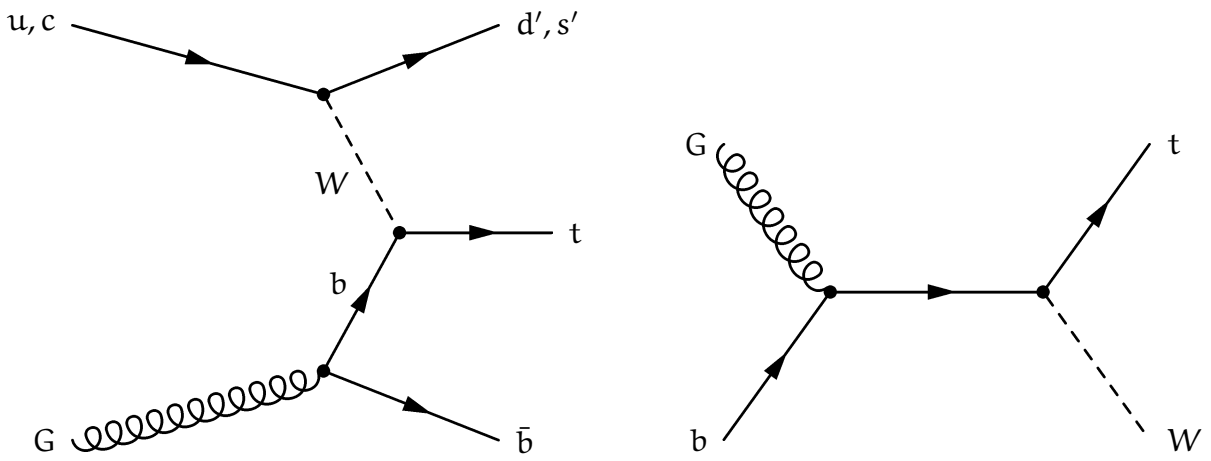

Figure 4.2: Feynman graphs for single top production: (a) t-channel, (b) in association with a W-boson, "tW-channel".

decay of Z' is shown in figure 4.3.

In order to study the experimental resolution of the $t\bar{t}$ reconstruction, Monte-Carlo samples for Z' resonances decaying in $t\bar{t}$ have been produced for $m_{Z'} = 1$, 2 and 3 TeV. The width of the Z' was set to 1 % of the mass. As shown in figure 3.7, this is well below the detector resolution. Therefore, it is easy to compare different reconstruction methods for $t\bar{t}$ as the natural width and detector effects do not have to be unfolded.

Details on the sample sizes are given in table 4.4. As the search concentrates on the $t\bar{t}$ decay channel, all cross sections given for Z' are denoted σB to indicate that the value includes the branching ratio in $t\bar{t}$, i. e. $\sigma B = \sigma(Z') \times BR(Z' \to t\bar{t})$.

4.2 Object Definition and Pre-Selection

As generator-level filters have been applied on some of the generated Monte-Carlo samples, a first selection has to be done to ensure a consistent sample which can be used as reference point for comparisons of further cuts. Also, this pre-selection step ensures that the event reconstruction as discussed in section 3.7 is possible for all events.

The objects definition used for this and the following selections is:

- *Jets* are SIScone jets built from calorimeter towers with relative (level-2) and

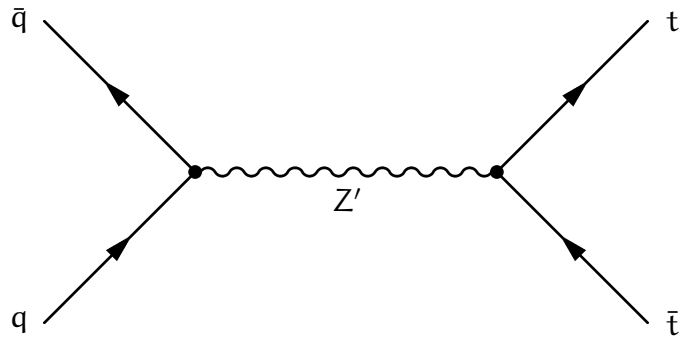

Figure 4.3: The leading order Feynman graph for Z′ production and decay in $t\bar{t}$.

Z′ mass in TeV	decay mode	generated events
1	fully leptonic	49 000
1	semi-leptonic	65 000
1	fully hadronic	49 000
2	fully leptonic	25 000
2	semi-leptonic	25 000
2	fuly hadronic	25 000
3	fully leptonic	20 000
3	semi-leptonic	50 000
3	fully hadronic	8 000

Table 4.4: Number of generated events for the Z′ samples which are used as signal for this study. The width is set to 1 % of the mass for all samples. Unlike the notation used otherwise, "lepton" includes the τ lepton in this case.

absolute (level-3) jet energy corrections applied. A jet is required to have $p_T \geqslant 50\,\text{GeV}$and $|\eta| < 2.4$. The p_T requirement is motivated by the hard p_T spectrum of the signal. In the chosen η range, the tracker is available and could be used for b-tagging in future studies. Moreover, for high $|\eta|$, the background contribution rises much more than the signal contribution.

- *Muons* are defined with a $p_T \geqslant 10\,\text{GeV}$ and $|\eta| < 2.1$. Additionally, some cuts on identification are required which loosely follow the V+jets recommendation discussed at the end of section 3.3: all global muons with
  - a global track fit with $\chi^2/\text{ndof} < 10$,
  - a transverse impact parameter $d_0 < 0.2\,\text{mm}$, and
  - the number of valid hits in the tracker $n \geqslant 11$

process	expected events	efficiency
QCD	$4.98 \cdot 10^6$	0.087
W+jets	$18.5 \cdot 10^3$	0.0023
Z+jets	3 122	0.0017
single top	1 264	0.089
$t\bar{t}$, semileptonic μ	8 214	0.68
$t\bar{t}$, other	8 113	0.11
Z', $m_{Z'} = 1\,\text{TeV}$, semileptonic μ	238	0.80
Z', $m_{Z'} = 2\,\text{TeV}$, semileptonic μ	250	0.84
Z', $m_{Z'} = 3\,\text{TeV}$, semileptonic μ	253	0.85
Z', $m_{Z'} = 1\,\text{TeV}$, other	323	0.19
Z', $m_{Z'} = 2\,\text{TeV}$, other	412	0.24
Z', $m_{Z'} = 3\,\text{TeV}$, other	460	0.27

Table 4.5: Number of expected events after the pre-selection for $L = 200\,\text{pb}^{-1}$. The efficiency is calculated with respect to the generated number of events. For Z', an arbitrary cross section of $\sigma B = 10\,\text{pb}$ was used. For $t\bar{t}$ final state, the events are split into semileptonic μ decay channel and all other decay channels. For single top, the numbers for t-channel and tW were added.

are used. Note that neither a cut on the deposited energy in the veto cone, nor the isolation requirement is applied here because for signal, the muon is often close to a jet.

- *Electrons* are defined by the "tight" cuts as discussed in section 3.4 with $|\eta| < 2.5$ and $p_T > 10\,\text{GeV}$.
- *Leptons* in the sense of this analysis are only electrons and muons.
- $\not{E}_T$ is based on calorimeter towers and is corrected for jet energy scale and muons as discussed in section 3.5.5.

The pre-selection requires at least two jets and that the lepton with largest p_T is a muon.

After this pre-selection, the expected number of events for each process is listed in table 4.5. Selected events with $t\bar{t}$ in the final state not decaying to muon+jets come mainly from the dilepton channel. The cut efficiency in the table is calculated with respect to the unselected sample, so some care has to be taken if interpreting those numbers as generator level cuts have been applied, as discussed above.

4.3 Selection for Low $m_{t\bar{t}}$

In the standard model, the cross section for $t\bar{t}$ production as a function of $m_{t\bar{t}}$ has its largest value just above its turn-on at $2m_t$ (see figure 3.6). Therefore, the top quarks can be assumed to be in rest in the lab frame. As discussed in 1.4, each of the top quarks decays to three fermions. As the top quarks are at rest, they decay isotropically and the decay products will typically have a large angular separation and will be reconstructed as separate objects.

An example of a typical selection for standard model $t\bar{t}$ in the muon+jets decay channel is

- require at least four jets. The p_T cut in the jet definition for the third and fourth jet is relaxed to $p_T \geqslant 30\,\text{GeV}$, while for the two leading jets, $p_T \geqslant 50\,\text{GeV}$ is required.
- require exactly one isolated muon and no isolated electron with $p_T \geqslant 20\,\text{GeV}$. Lepton isolation is defined as CombRelIso < 0.1 with the definition of the combined relative isolation CombRelIso from equation 3.1.

The event yields after applying those additional cuts are given in table 4.6. As can be seen, the selection yields a good signal-to-background ratio for standard model $t\bar{t}$ of about 2/1. However, the selection efficiency for Z′ degrades for large $m_{Z'}$.

A stack plot showing the expected number of events after applying this selection versus the reconstructed invariant $t\bar{t}$ mass, $M_{t\bar{t}}$, is shown in figure 4.4. As expected, the background is quite low for high values of $m_{Z'}$. As it is hard to judge or compare the sensitivity from this plot, the significances are given in table 4.7.

4.4 Selection for High $m_{t\bar{t}}$

As shown in the previous section, the selection efficiency decreases for high $m_{t\bar{t}}$. This is due to the changing event topology and jets begin to merge for high $m_{t\bar{t}}$, which was shown in figure 3.3.

The decay products of a high-mass resonance will have large transverse momenta. On the other hand, backgrounds have the largest cross section for relatively small energy and p_T scales. This feature of the signal can be seen in different variables which are shown in figure 4.5:

- The large transverse momentum of the leading jet and the second jet (jets are sorted by p_T).
- Large missing transverse energy, $\not{E}_T$.
- The H_T variable, which is defined as the scalar sum of the transverse momenta of all jets, leptons, and $\not{E}_T$, according to the object definition.

process	expected events	efficiency
QCD	245	$4.8 \cdot 10^{-5}$
W+jets	1 373	0.074
Z+jets	156	0.050
single top	147	0.12
$t\bar{t}$, semileptonic μ	3 236	0.39
$t\bar{t}$, other	617	0.08
Z', $m_{Z'} = 1\,\text{TeV}$, semileptonic μ	96	0.40
Z', $m_{Z'} = 2\,\text{TeV}$, semileptonic μ	77	0.31
Z', $m_{Z'} = 3\,\text{TeV}$, semileptonic μ	55	0.22
Z', $m_{Z'} = 1\,\text{TeV}$, other	26	0.08
Z', $m_{Z'} = 2\,\text{TeV}$, other	27	0.07
Z', $m_{Z'} = 3\,\text{TeV}$, other	22	0.05

Table 4.6: Number of expected events for $L = 200\,\text{pb}^{-1}$ after applying the event selection as described in section 4.3. The cut efficiency is calculated with respect to the pre-selection. The cross section for Z' is $\sigma B = 10\,\text{pb}$.

$m_{Z'}$ in TeV	S	B	$S/\sqrt{B}$
1	86 ± 1.1	1269 ± 10	2.3
2	59 ± 1.4	76 ± 3	6.5
3	32 ± 0.8	8.4 ± 1.3	10.1

Table 4.7: The significances $S/\sqrt{B}$ for the events with a $M_{t\bar{t}}$ reconstructed mass in the window $m_{Z'} \pm 300\,\text{GeV}$ after applying the selection for $\sigma B = 10\,\text{pb}$. The calculation of the significance and the error is discussed in section 4.4.1.

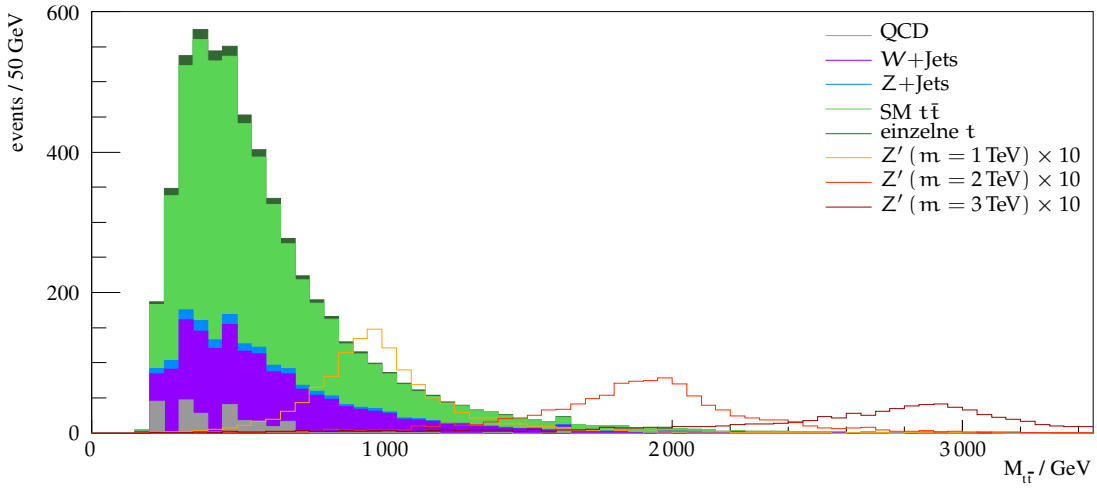

Figure 4.4: Reconstructed invariant mass of the $t\bar{t}$ system, after applying the event selection for a standard model topology as described in section 4.3. As cross section for Z', $\sigma B = 100\,\text{pb}$ was used in this plot. Note that the other plots always use $\sigma B = 10\,\text{pb}$, therefore Z' is indicated with "$\times 10$". The reconstruction is based on the hypothesis with minimal C as discussed in section 3.7.2.

- $H_{T,lep}$, the "leptonic H_T" is a variation of H_T and defined as the scalar sum of $\not{E}_T$ and the leading μ.

Another important class of event variables for selection are variables associated to the muon. As discussed in section 3.3 and section 4.1.1, muon properties are quite different for QCD events and signal events. Some variables are shown in figure 4.6:

- The combined relative isolation, as used in the standard model selection.
- The ΔR of the muon to the nearest jet, where the "nearest jet" is defined via ΔR.
- The transverse momentum of the muon with respect to the axis of the nearest jet, $p_{T,rel}$.
- The invariant mass of the muon and the nearest jet, $M_{\mu,j}$.
- The variable z_μ was studied on generator level by [9] which suggest to use it for identifying boosted tops. It is defined as

$$z_\mu = \frac{E_\mu}{E_\mu + E_b},$$

where E_b is the hadronic energy in the vincinity of the muon (here, the nearest jet is used). This variable is shown as an example of a variable which has shown very good separation power on generator level but is not very useful as cut variable after reconstruction.[1]

[1] However, "hadronic energy" in the cited paper is defined differently, by running the k_T jet algorithm

For the last three observables, the nearest jet is interpreted as the hadron decay product of a heavy flavor decay, or the other hadrons of a jet in case of a decay-in-flight. The ΔR, $p_{T,rel}$ and $M_{\mu,j}$ should be small for QCD and large for signal. In the definition of "nearest jet" the jet p_T cut is relaxed to $p_T \geqslant 30\,\text{GeV}$. Otherwise, the nearest jet is in many cases not the hadronic energy associated with muon production for QCD.

While these muon-associated variables are a good handle against QCD, *W*+jets and *Z*+jets events contain muons from the hard process. Those muons are not correlated with jet activity. Therefore, those variables do not separate those backgrounds from the signal.

For this analysis, all the observables above and many more have been studied. The final selection makes use of the following cut variables:

- p_T of the leading jet,
- A combined cut on $\Delta R(\mu, j)$ and $p_{T,rel}(\mu, j)$, and
- $H_{T,lep}$.

The choice of these variables is not only based on their separation power of signal and background, but also the ability to use this selection cuts for modelling QCD from data (see section 5.2) and the possibility to combine the obtained results with other groups using a similar selection.

Once a set of variables has been chosen, the actual cut values have to be optimized. A full optimization which is based on the final analysis (see chapter 5) is too time consuming to be feasible. Instead, the significance of a counting experiment is optimized.

4.4.1 Cut Optimization

The significance of a counting experiment is given by $s = S/\sqrt{B}$, where S is the expected number of signal events and B the expected number of background events. If B is affected by a systematical error ΔB, the significance becomes

$$s = \frac{S}{\sqrt{(\Delta B)^2 + B}}. \tag{4.1}$$

The cuts are optimized for this significance using the number of signal and background events in a specified region of $M_{t\bar{t}}$, not the whole $M_{t\bar{t}}$ range. This allows to optimize the cut for different $m_{Z'}$ separately and to compare the obtained value with the final selection which aims to have a good significance for a large range of $m_{t\bar{t}}$.

with the muon as jet constituent.

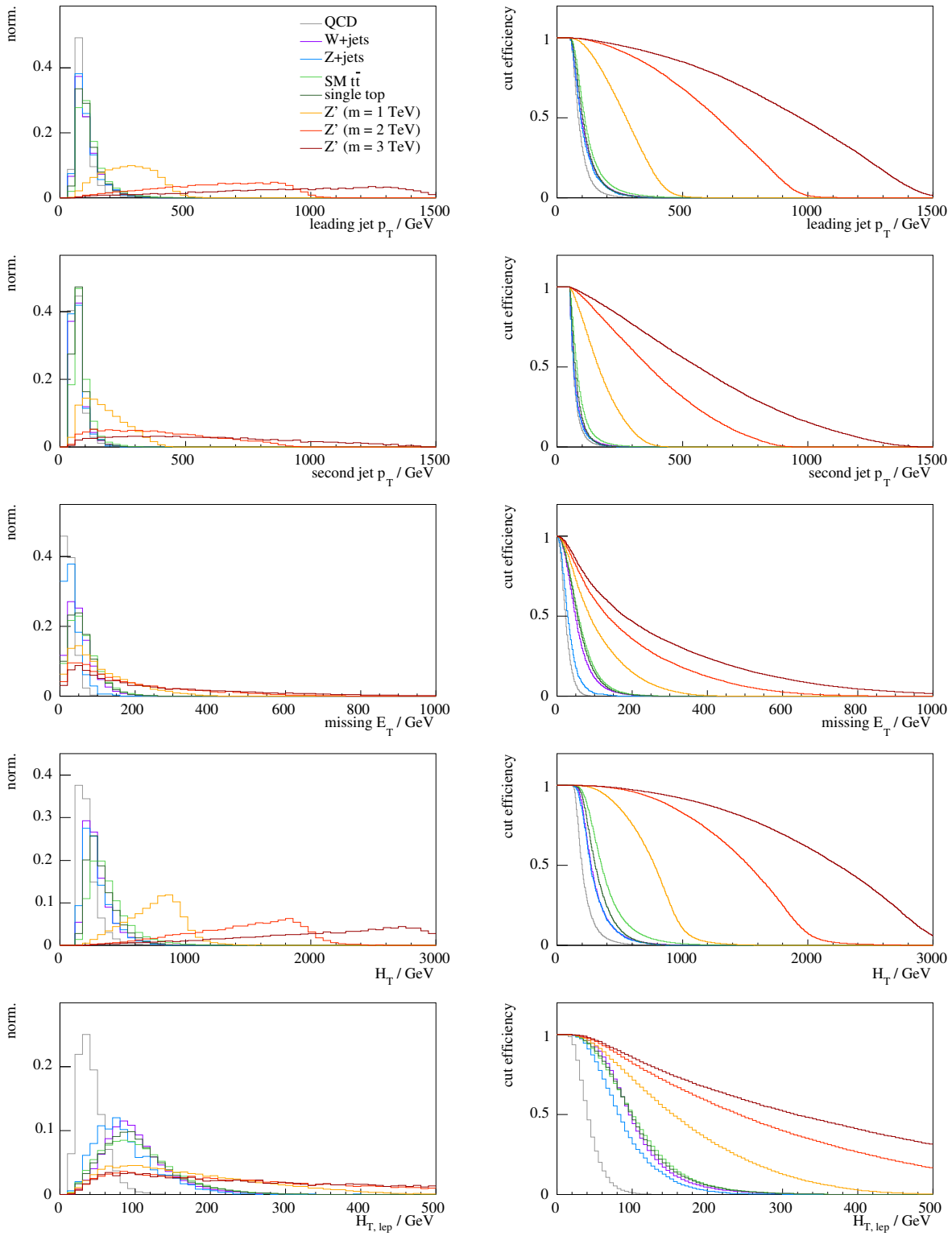

Figure 4.5: Different observables for signal and backgrounds which could be used for signal selection. On the left, the shapes of the variable is shown. On the right, the cut efficiency as function of the cut value is shown. For the definition of the variables, see main text.

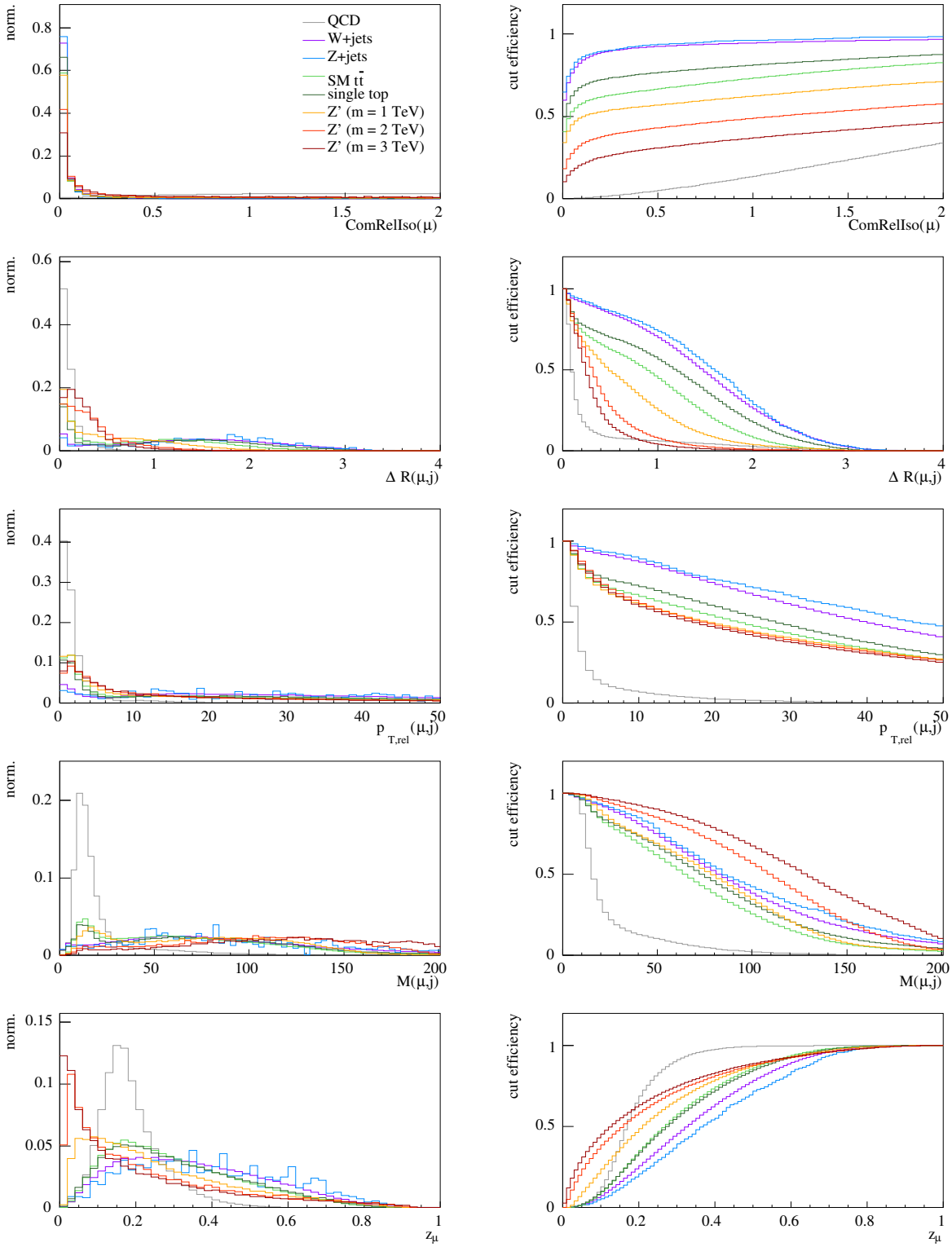

Figure 4.6: Distributions (left) and cut efficiencies (right) for different observables for the muon.

Directly maximizing the significance as a function of the cut values is a problem for many minimization algorithms, as they use numerical derivatives of the objective function, for example the widely used MIGRAD routine in the MINUIT minimization system [86]. Numerical derivatives are computed by evaluating the function in nearby values. However, the number of signal and background events as function of the cut values are non-continuous functions and the numerical derivatives will either evaluate to zero, if the step did not modify the set of selected events, or change the number of selected events even for infinitesimal changes of a cut value. Even for minimization algorithms not using numerical derivatives, an objective function constant in the neighborhood of some points can be a problem.

The solution can be seen easily if the cut procedure is formulated more generally as a re-weighting procedure where events are re-weighted by a factor between 0 and 1. For cut values $c_1, \dots c_n$ which cut on variables $\vec{v} = (v_1, \dots v_n)$, the weight factor for a cut is given by[2]

$$w_{\vec{c}}(\vec{v}) = \prod_{i=1}^{n} \Theta(v_i - c_i) \tag{4.2}$$

where Θ is the Heaviside step function. Now, Θ can be replaced by a smooth function f, provided f behaves similarly, i. e. f is monotonically increasing, $f(-\infty) = 0$ and $f(+\infty) = 1$. Then, the weight is re-defined as

$$w_{\vec{c}}(\vec{v}) = \prod_{i=1}^{n} f(v_i - c_i). \tag{4.3}$$

A function fulfilling those requirements is

$$f(x) = \frac{1}{2}(1 + \tanh(x/b)) \tag{4.4}$$

with $b > 0$. For $b \to 0$, the Θ function is reproduced. Therefore, one has to choose values for b which are not too small in order to get the desired numerical advantages. For each dimension, b can have a different value, so equation 4.2 becomes

$$w_{\vec{c}}(\vec{v}) = \prod_{i=1}^{n} \frac{1}{2} \left(1 + \tanh\left(\frac{v_i - c_i}{b_i}\right)\right). \tag{4.5}$$

In order to obtain a smooth behavior for $S/\sqrt{B}$ as a function of $\vec{c}$, it is enough to choose values for the b_i which are in the order of the distance of two events in this variable.

[2] Without loss of generality, only cuts which cut away the low region for each variable v_i are considered.

However, if choosing b_i appropriately, w can also be interpreted as the *probability* that an event with measured variable values $\hat{\vec{v}}$ passes the selection. This would allow calculating efficiency errors for a selection. To see how this can be done, the measurement $\hat{\vec{v}}$ of the variable $\vec{v}$ is interpreted as the most probable value of a posterior probability density in the variable space, $p_{\hat{\vec{v}}}(\vec{v})$. The factorized approach in equation 4.3 is equivalent with the assumption that this density factorizes in $\vec{v}$. Moreover, it is assumed that

$$p_{\hat{\vec{v}}}(\vec{v}) = p_0(\vec{v} - \hat{\vec{v}}) =: p(\vec{v} - \hat{\vec{v}}),$$

i. e. the posterior for measurements $\hat{\vec{v}}$ and $\hat{\vec{v}}'$ are connected by a translation. This is equivalent to choosing values for b_i which are independent of $\hat{\vec{v}}$.

To determine a sensible interpretation for b_i, consider the one-dimensional case. If the measurement of the variable $\hat{v}$ implies a posterior probability $p(v - \hat{v})$, the probability that the event passes the cut with a cut value c is

$$w_c(\hat{v}) = \int_c^\infty p(v - \hat{v})\, dv. \tag{4.6}$$

And it follows

$$\frac{dw_c}{d\hat{v}}(\hat{v}) = p(c - \hat{v}).$$

On the other hand (from equation 4.3)

$$\frac{dw_c}{d\hat{v}}(\hat{v}) = f'(\hat{v} - c).$$

Using f introduced in equation 4.4, the corresponding posterior probability density is

$$p(c - \hat{v}) = f'(\hat{v} - c) = \frac{1}{2b}\frac{d}{dx}\tanh(x)\big|_{x=\frac{\hat{v}-c}{b}} = \frac{1}{2b}\operatorname{sech}^2\frac{\hat{v} - c}{b}.$$

The standard deviation σ of this distribution is[3]

$$\sigma = \frac{\pi}{2\sqrt{3}} \cdot b.$$

This can be used to determine the value for the b_i given the error on this variable as standard deviation σ_i.

The errors modelled by this approach is the systematic uncertainty. However, those uncertainties will usually not factorize for different variables. This dependence could

[3]Of course, one could also take a Gaussian distribution for the error model. However, this would require a large number of numerical evaluations of the error function. The evaluation of tanh is much faster, and was therefore chosen as f.

be accounted for by generalizing this approach. However, as the results of a cut optimization typically do not change very much for different values of the b_i, and the significance is a rough estimate anyway, such a generalization is not worthwhile.

As σ_i is the systematic uncertainty for the variable v_i, the significance s from equation 4.1 can be used as objective function for the optimization. The number of selected background events B is given by the sum of the cut probabilities,

$$B = \sum_{k=1}^{N_{bkg}} w(\vec{v}^{(k)}) = \sum_{k=1}^{N_{bkg}} \prod_{i=1}^{n} f(v_i^{(k)} - c_i),$$

where the sum runs over all background events, and the variables of event k are $\vec{v}^{(k)}$. Using the usual Gaussian error propagation in linear approximation, the error on B due to the uncertainty in $\vec{v}_i^{(k)}$ is

$$(\Delta B)^2 = \sum_{k=1}^{N_{bkg}} \sum_{i=1}^{n} \left(\frac{w(\vec{v}^{(k)})}{f(v_i^{(k)} - c_i)} f'(v_i^{(k)} - c_i) \right)^2 \sigma_i^2. \tag{4.7}$$

For the particular choice of f as in equation 4.4, straight-forward numerical evaluation of $f'(x)/f(x)$ is unstable for large negative values of x as both numerator and denominator go to zero. This can be handled by appropriate cut-offs.

Another important point to address is the lack of Monte-Carlo statistics. This can lead to cuts where no events are selected at all of a certain Monte-Carlo sample. As discussed in section 4.4.3, a possible solution is to assume that one Monte-Carlo event passed and interpret the background event yield as an upper limit.

To take this into account, $(\Delta B)^2$ from equation 4.7 is modified by adding $(\Delta B_{MC})^2$, which accounts for the error from limited Monte-Carlo statistics:

$$(\Delta B_{MC})^2 = n_{MC} + \frac{1}{n_{MC} + 1}, \tag{4.8}$$

where n_{MC} is the number of events passing the selection. In case of large n_{MC}, this is the usual poisson error. In case no event passes the selection, the error is one Monte-Carlo event. This is done for each Monte-Carlo sample seperately; of course, B and ΔB have to be weighted with the corresponding event weight for this sample.

Multiple Signal Scenarios In the present studies, the selection should be as independent as possible of the invariant $t\bar{t}$ mass, $m_{t\bar{t}}$. However, this goal is hardly achievable with a common cut-based selection as the event topology changes with $m_{t\bar{t}}$. Of course, it would be possible to optimize selections independently for different $m_{t\bar{t}}$, maybe even based on different observables. However, it would then be necessary

to have a working data-based QCD model (see section 5.2) for each mass point that was optimized. While this might be achievable in principle, it seems better to wait until it can be shown which methods work with real data before adding so much complexity to the analysis.

So the question is how to optimize the selection for a large $m_{t\bar{t}}$ interval of $m_{t\bar{t}}$, say, the range of available Z′ samples 1 TeV $\lesssim m_{t\bar{t}} \lesssim$ 3 TeV.

Taking simply a mixture of all three Z′ samples with equal weight and optimizing $S/\sqrt{B}$ does not lead to satisfactory results, as the significance would be dominated by the sample with $m_{Z'} = 3\,\text{TeV}$. Giving the samples with higher $m_{Z'}$ lower weights still might yield undesirable results, as it is not guaranteed that the optimization simply cuts away a whole sample.

Of course, the sensitivity is always better if optimizing for a specific $m_{t\bar{t}}$ instead of a large $m_{t\bar{t}}$ interval. So after optimizing the selection for the $m_{t\bar{t}}$ interval, a natural question is to quantify the loss of sensitivity of this optimization, compared to an optimization for a specific $m_{t\bar{t}}$.

In the following, $s(m_{Z'}, c)$ denotes a measure of the sensitivity for discovering Z′ with mass $m_{Z'}$, given the selection cuts c.

To address both, the optimization for a $m_{t\bar{t}}$ interval and quantifying the loss of sensitivity, following algorithm seems suitable:

1. Optimize the cuts for all $m_{Z'}$ individually. This yields a set of optimal cut values $\hat{c}(m_{Z'})$ and the maximum significance for each Z′ mass, $\hat{s}(m_{Z'}) := s(m_{Z'}, \hat{c}(m_{Z'}))$.
2. Define the relative significance $r(c)$ of a cut c as

$$r(c, m_{Z'}) = \frac{s(m_{Z'}, c)}{\hat{s}(m_{Z'})}. \tag{4.9}$$

   It compares the sensitivity of the cut c to the optimal cut optimized for the individual sample. Now, search for the cut c which maximizes the *minimum* relative significance, i. e. the cut which maximizes $l(c)$, defined as

$$l(c) = \min_{Z'} r(c, m_{Z'}), \tag{4.10}$$

   where Z′ runs over all three Z′ samples.

The second step yields both, a cut that can be used for an $m_{t\bar{t}}$ interval and a quantification of how much significance degrades by applying a global cut instead of optimizing individually for each $m_{Z'}$.

4.4.2 Cut Optimization Results

The optimization is run on the same variables as the final selection discussed in section 4.4.3. For the optimization, uncertainties for the variables have to be chosen.

The uncertainties for jet p_T, $\not{E}_T$ and $H_{T,lep}$ are dominated by jet energy scale uncertainty. This is assumed to be 10 % (see section 5.4). Therefore, values corresponding to about 10 % for the resulting cuts are chosen. An error corresponding to about 10 % was also chosen for the other cut variables. However, the total error is dominated by Monte-Carlo statistics and the assumed uncertainty has only insignificant impact on the optimization result. Therefore, no attempt to determine more accurate errors was made.

The variables are (with the assumed systematic uncertainty in paranthesis):

1. the primary lepton's transverse momentum, $p_{T,\mu}$ (3 GeV)
2. the leading jet's transverse momentum, $p_{T,lead}$ (30 GeV)
3. $H_{T,lep}$ (30 GeV)
4. a "2D-cut" on $\Delta R(\mu, j)$ (0.05) and $p_{T,rel}(\mu, j)$ (3 GeV) which only retains events where the muon has a value above some threshold in at least one of those variables, and
5. a Z veto: a cut on the invariant mass of two leptons of same flavor and opposite charge where the mass is in a window around the Z mass of 91.2 GeV.

The Z veto cut was fixed to a window width of 10 GeV in order to reduce the number of free variables in the fit. An optimization of this cut shows that this is always close to the optimal value, independent of other cuts. The cut of the lepton p_T was fixed to 20 GeV to have a save margin from the used muon trigger (see section 4.4.3). Optimizing it simultaneously with the other cut variables always yielded the lowest allowed value.

The cut values of the optimization are shown in table 4.8. As expected, the selection imposes higher cuts on p_T and $H_{T,lep}$ for large $m_{Z'}$. The resulting expected signal and background events are given in table 4.9.

After optimization of the individual samples, an optimization for all $m_{Z'}$ is carried out by maximizing the minimum relative significance as discussed in the previous section, using the minimum relative significance defined in equation 4.10 as optimization criterion. This "global" optimization yields a relative significance of 0.74 for all three $m_{Z'}$.

For comparison, the significance of the final selection is given in table 4.10.

4.4.3 Final Selection

The variables which are used in the selection are listed in the previous section. The final selection applies following cuts in addition to the pre-selection:

- $p_{T,\mu} > 25$ GeV,
- $p_{T,lead} > 260$ GeV,
- $H_{T,lep} > 200$ GeV,

$m_{Z'}$ in TeV	$H_{T,lep}$ in GeV	$p_{T,lead}$ in GeV	2D-cut (ΔR, $p_{T,rel}$ in GeV)
1	166	282	$(0.40, 35.0)$
2	185	679	$(0.40, 35.2)$
3	319	716	$(0.003, 35.0)$
global	190	374	$(0.35, 34.9)$

Table 4.8: Results of the cut optimization if optimizing the Z' samples individually and for a global optimization maximizing the minimum relative significance as discussed in section 4.4.1.

$m_{Z'}$ in TeV	S	B	$S/\sqrt{B}$
1	58±0.95	200±4.9	3.8
2	69±1.6	2.6±1.1	36
3	100±1.4	0.84±0.98	75

Table 4.9: The expected number of signal events S and background events B for the optimized cuts from table 4.8. S, B and $S/\sqrt{B}$ are determined counting the number of events with a reconstructed mass $M_{t\bar{t}}$ within $m_{Z'} \pm 300\,\text{GeV}$, including the background error according to equations 4.7 and 4.8. For Z', $\sigma B = 10\,\text{pb}$ is used. The significance is calculated the same way as in table 4.7.

				rel. significance compared to selection	
$m_{Z'}$ in TeV	S	B	$S/\sqrt{B}$	optimized	standard model
1	53±0.91	193±4.3	3.6	0.95	1.7
2	114±2.0	27±1.8	21	0.58	3.2
3	102±1.4	3.9±1.1	45	0.60	4.5

Table 4.10: The significance $S/\sqrt{B}$ for the final selection as decribed in section 4.4.3, calculated as in table 4.8. The relative significance is defined as the ratio of the significance obtained here divided by the significance from another selection. Selections considered for comparison are the cuts optimized individually for the different Z' samples (see table 4.8) and the standard model selection discussed in section 4.3 (see table 4.7).

- for the 2D-cut $p_{T,rel}(\mu, j) > 35\,\text{GeV}$ or $\Delta R(\mu, j) > 0.4$,
- the Z veto with a window of 10 GeV, and
- the muon trigger *HLT_Mu15*.

The final selection has been developed before a detailed study of the optimization and was done using a standard model $t\bar{t}$ sample with $m_{t\bar{t}} \geqslant 1\,\text{TeV}$. As the cross section is steeply falling in $m_{t\bar{t}}$, this sample roughly corresponds to $m_{Z'} = 1\,\text{TeV}$. Accordingly, the relative significance of the final selection is close to the optimized selection for $m_{Z'} = 1\,\text{TeV}$ which can be seen in table 4.10. This is also the reason of values for the 2D-cut values and the cut on $p_{T,lead}$ are close to the optimized values for $m_{Z'} = 1\,\text{TeV}$. The cut value for $H_{T,lep}$ was chosen higher than optimal, as a sample with an *inverted* $H_{T,lep}$ cut is used for building the likelihood function for the analysis (see section 5.2) and a larger region for this inverted cut in $H_{T,lep}$ is more robust.

The cut on $p_{T,\mu} > 25\,\text{GeV}$ is purely historical and will probably relaxed in the future.

The performance of the selection compared to a standard model selection is very good: as shown in the last column of table 4.10, $S/\sqrt{B}$ raises by a factor of over 3 for $m_{Z'} \geqslant 2\,\text{TeV}$. This is because fewer events have four or more jets (as required by the standard model selection) for large $m_{Z'}$. But also for $m_{Z'} = 1\,\text{TeV}$, where the event topology is more standra model like, $S/\sqrt{B}$ is much larger.

The trigger efficiency for different triggers if all other cuts of the final selection are applied is shown in figure 4.7. In this study, the high level trigger *HLT_Mu15* is used which requires a muon with $p_T > 15\,\text{GeV}$. This trigger has no requirement on muon isolation and has a high efficiency for a large range of $m_{t\bar{t}}$. As jets with large p_T are expected, it would also be possible to use jet triggers. For example, *HLT_DiJetAve70* triggers on the presence of two jets with an average corrected $p_T \geqslant 70\,\text{GeV}$. It can also be seen that the efficiency of a trigger for isolated muons, *HLT_IsoMu11*, is decreasing in $m_{t\bar{t}}$, as expected.

The number of expected events and the selection efficiency after applying the selection is given in table 4.11. For QCD, enough statistics is available for the two high $\hat{p}_T$ bins ($\hat{p}_T > 250\,\text{GeV}$) but for the lower $\hat{p}_T$ bins, no event passed the selection and given the high weight, the question is how to interpret zero passing events. The limit given in the table assumes that the selection efficiency for the lowest $\hat{p}_T$ sample is actually zero and that an upper limit for the selection efficiency of the second lowest $\hat{p}_T$ sample is one passing Monte-Carlo event. Both assumptions are now discussed.

An upper limit on the number of expected events for the lowest $\hat{p}_T$ bin is derived assuming that for QCD, $H_{T,lep}$ and the 2D-cut are independent if applying the final selection except for those two cuts (see also QCD estimation in section 5.2, where this assumption is checked). In this case, the selection efficiency of applying the last

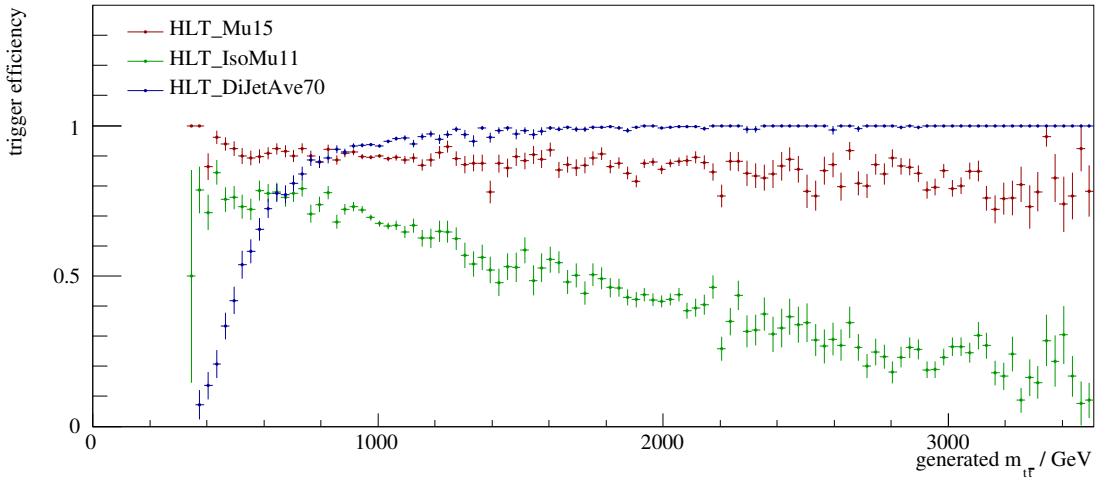

Figure 4.7: Trigger efficiency as function of the generated $t\bar{t}$ mass. The efficiency is calculated with respect to the final selection.

two cuts is given by the product of the selection efficiencies of the two cuts applied individually. This yields an upper limit of 0.1 expected events for the final selection and can be neglected.

The estimation of upper limits on the selection efficiency if no Monte-Carlo event passed the selection can be derived using Bayesian statistics using a flat prior for the selection efficiency [87]. This yields an upper limit of $-\ln(1-\alpha)$ passed Monte-Carlo events at confidence level α in the limit of large Monte-Carlo samples ($N \gtrsim 100$). Assuming one passed Monte-Carlo events cooresponds to $\alpha = 0.63$ which seems reasonable.

A stack plot for the processes after the final selection is shown in figure 4.8. The expected numbers of events for signal and background in a mass window in $M_{t\bar{t}}$ were already given in the last section in table 4.10.

4.5 Selection with the Top Jet Algorithm

As discussed in section 3.6, the top jet algorithm can be used to reconstruct hadronically decaying top quarks with high momentum.

The top jet algorithm can be used for event reconstruction, especially for $t\bar{t}$ events in the full hadronic channel [88]. In principle, this is possible for this study as well. However, a comparison with the other methods developed here is much more difficult. Furthermore, the leptonically decaying top quark, including $\not{E}_T$, should then be reconstructed using the same jet algorithm and jet corrections, for reasons of

process	expected events	efficiency
QCD	< 7.4	$< 1.5 \cdot 10^{-6}$
W+jets	144	$7.7 \cdot 10^{-3}$
Z+jets	8.9	$2.9 \cdot 10^{-3}$
single top	6.3	$5.0 \cdot 10^{-3}$
$t\bar{t}$, semileptonic μ	85	0.010
$t\bar{t}$, other	40	0.005
Z', $m_{Z'} = 1\,\text{TeV}$, semileptonic μ	57	0.24
Z', $m_{Z'} = 2\,\text{TeV}$, semileptonic μ	137	0.55
Z', $m_{Z'} = 3\,\text{TeV}$, semileptonic μ	150	0.60
Z', $m_{Z'} = 1\,\text{TeV}$, other	18	0.06
Z', $m_{Z'} = 2\,\text{TeV}$, other	66	0.16
Z', $m_{Z'} = 3\,\text{TeV}$, other	77	0.17

Table 4.11: Expected events and selection efficiency after applying the final selection. The efficiency is calculated with respect to the pre-selection defined in section 4.2. The assumed cross section for Z' is $\sigma B = 10\,\text{pb}$.

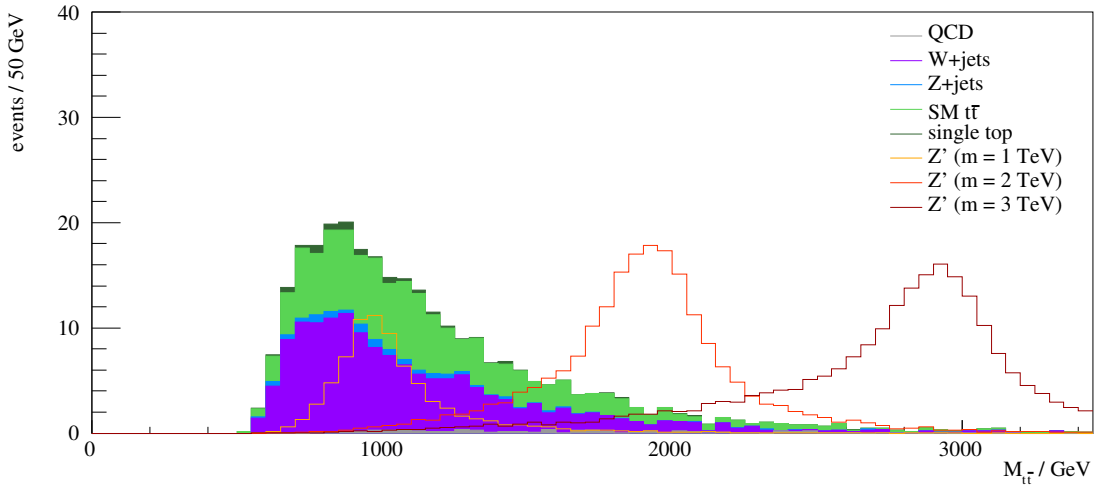

Figure 4.8: The expected number of events for different $M_{t\bar{t}}$ after applying the final selection for $L = 200\,\text{pb}^{-1}$. The cross section for Z' is $\sigma B = 10\,\text{pb}$. Unlike figure 4.4, no scaling of the signal is performed.

consistency. Therefore, in the present study, the top jet algorithm is used for event selection only.

The selection requires one top jet, the other variables are optimized for the individual Z′ samples as outlined in section 4.4.2. The result can be seen in tables 4.12 and 4.13. The performance for $m_{Z'} = 1\,\text{TeV}$ is worse. This is expected since the top jet reconstruction efficiency at top quark $p_T \approx 500\,\text{GeV}$ is low which can be seen in figure 3.4. For all $m_{Z'}$, the statistics for the background Monte-Carlo sample is clearly becoming the limiting factor for an estimate of the significance. Therefore, no global optimization has been attempted and no relative significances are given.

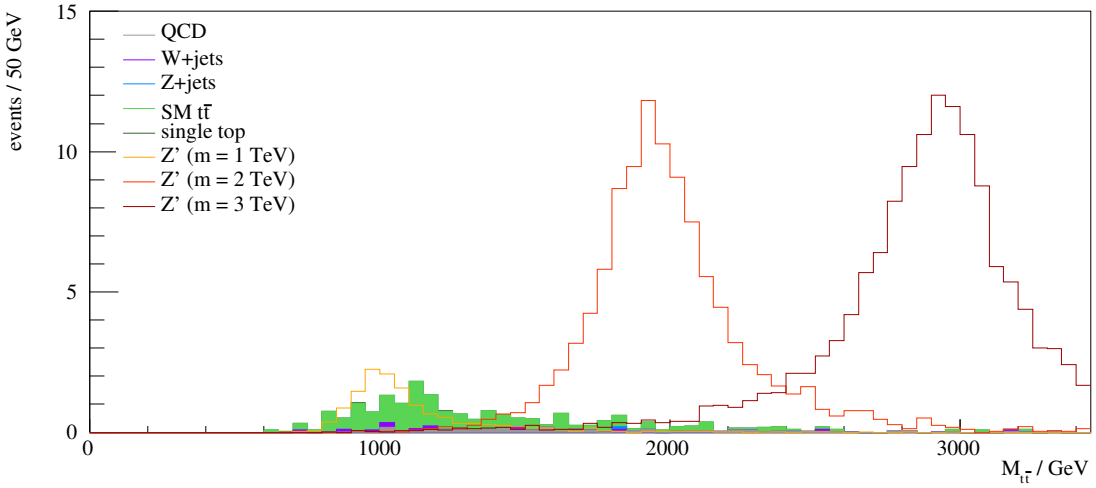

Figure 4.9: The expected number of events after applying a selection requiring one top tagged jet, $H_{T,lep} > 200\,\text{GeV}$ and $p_{T,rel} > 25\,\text{GeV}$. For Z′, $\sigma B = 10\,\text{pb}$ was used.

$m_{Z'}$ in TeV	$H_{T,lep}$ in GeV	$p_{T,lead}$ in GeV	2D-cut (ΔR, $p_{T,rel}$ in GeV)
1	118	235	$(0.36, 22.5)$
2	237	458	$(0.30, 33.2)$
3	262	50	$(0, 0.4)$

Table 4.12: Results of the cut optimization of the selection using the top jet algorithm for different Z′ samples.

As an example, the resulting $M_{t\bar{t}}$ distribution is shown for a relatively loose selection in figure 4.9.

$m_{Z'}$ in TeV	S	B	$S/\sqrt{B}$
1	9.1±0.4	7.6±1.2	3.0
2	43±1.2	0.66±0.98	34
3	52±1.0	0.14±0.96	50

Table 4.13: The expected numbers of signal and background after the optimized top jet selection. S, B and $S/\sqrt{B}$ are calculated as in table 4.8, for $L = 200\,\mathrm{pb}^{-1}$ and $\sigma B = 10\,\mathrm{pb}$. The relative significance is not given here, as a meaningful comparison is hardly possible due to limited Monte-Carlo statistics.

To circumvent the problem of limited Monte-Carlo statistics, one can assume that the "mistag rate" ϵ, i. e. the probability of identifying a jet errornously as top jet, depends only on the transverse momenum of the jet p_T and is independent of all other observables of the event. Plans exist of how to measure this mistag rate using data-driven methods [88]. Once this rate is measured, a pre-selected sample of data (or Monte-Carlo) which would usually undergo the selection of "exactly one top tag" would instead be weighted by $1 - 2\epsilon_1\epsilon_2$ where ϵ_1 and ϵ_2 are the mistag probability of the two leading jets, assuming that there are not more than two hard jets in the event.

However, at the time the present study was conducted, this method was not yet available and was therefore not considered here. Besides, it would be necessary to verify that the mistag rates determined in [88] for the all hadronic channel are valid for the muon+jets channel. This is not self-evident, as the flavour composition of jets changes for a selection with a reconstructed muon and the mistag rate in turn could depend on jet flavour.

5 Analysis

The final goal of the present study is to establish upper limits on the cross section of Z' times branching ratio in $t\bar{t}$ for different $m_{Z'}$. How to determine those limits is the topic of this chapter.

First, the model used to describe the data and the likelihood function is introduced in section 5.1.1, followed by a discussion about how to construct confidence intervals using the likelihood function. There are several ways to do that. In section 5.1.2, an approximate frequentist approach is presented and section 5.1.3 introduces the Bayesian method. For the actual calculation of Bayesian confidence intervals, an efficient algorithm for numerical integration in high dimensions is required. The algorithm used here, Markov chain Monte-Carlo, is discussed in section 5.1.4.

The construction of the likelihood requires estimating probability densities for a variety of processes. While this can be done with Monte-Carlo in many cases, a method is presented to estimate the required quantities for QCD using only few assumptions from Monte-Carlo in section 5.2. The limits with this methods are given in section 5.3.

Finally, sources for systematic uncertainties are discussed in section 5.4.1, how they are incorporated into the statistical methods in sections 5.4.2 and 5.4.3, and finally the results in section 5.5.

5.1 Statistical Method

5.1.1 Model and Likelihood Function

A model means specifying the probability to observe a given number of events with a certain distribution, as function of the parameters of the model. The parameters of the model include the signal cross section for which an upper limit shall be derived.

In this case, the model is given by

$$p(\vec{x}, N|\vec{\beta}) = \frac{\nu(\vec{\beta})^N e^{-\nu(\vec{\beta})}}{N!} \prod_{j=1}^{N} p(x_j|\vec{\beta}), \tag{5.1}$$

where $\vec{x}$ contains the observables of all events, N is the number of observed events, $\vec{\beta}$ are the parameters of the model. The first term on the right hand side is the Poisson

probability to observe N events with a mean of $\nu(\vec{\beta})$ and $p(x|\vec{\beta})$ is the probability density of observing a single event at x.

The model is fully specified if $\nu(\vec{\beta})$ and $p(x|\vec{\beta})$ are given.

The expected number of events $\nu(\vec{\beta})$ can be written as the linear combination of the expected number of events of all contributing processes:

$$\nu(\vec{\beta}) = \sum_{i=1}^{n} \beta_i \nu_i, \tag{5.2}$$

where the sum runs over all n considered processes and ν_i is the theoretically predicted number of events for process i. With this definition, β_i is the ratio of the measured and the theoretically predicted cross section,

$$\beta_i = \frac{\sigma_i}{\sigma_{i,\text{pred}}}.$$

Accordingly, $p(x|\vec{\beta})$ is given by

$$p(x|\vec{\beta}) = \frac{1}{\sum_j \beta_j} \sum_{i=1}^{n} \beta_i p_i(x) \tag{5.3}$$

where $p_i(x)$ is the probability density for variable x for process i.

The predictions for the number of events ν_i and densities $p_i(x)$ for a process i are derived from Monte-Carlo: ν_i is the number of remaining events after applying the selection for the considered luminosity and an estimate for $p_i(x)$ is given by a normalized histogram of the variable.

In the following, process $i = 1$ is the process to determine the cross section for, i. e. the Z' process, and ν_1 is the number of expected events for $\sigma B = 1\,\text{pb}$. Given data, a statement about the true value of σB, or equivalently β_1, shall be derived. In order to do that, first define the Likelihood function of $\vec{\beta}$, given data $\vec{x}$:

$$L(\vec{\beta}|\vec{x}) := p(\vec{x}|\vec{\beta}). \tag{5.4}$$

The likelihood function is a function of the parameters $\vec{\beta}$ for fixed $\vec{x}$, while $p(\vec{x}|\vec{\beta})$ is a density in $\vec{x}$ for fixed $\vec{\beta}$.

How the likelihood function can be used to determine upper limits is subject to the next two sections.

5.1.2 Profile Likelihood and Classical Confidence Intervals

Consider a *hypothesis test* at significance level[1] α where the null hypothesis $\mathcal{H}_0$ is that $\vec{\beta}$ lies in some subset $\mathcal{A}$ and the alternative hypothesis $\mathcal{H}_1$ is $\vec{\beta} \notin \mathcal{A}$. For example,

[1] The significanc level for a hypothesis test is the probability of rejecting the null hypothesis when the null hypothesis is true.

$\vec{\beta} \in \mathcal{A}$ represents the model without signal, $\mathcal{A} = \{\vec{\beta}|\beta_1 = 0\}$.

Then, define the *likelihood ratio* as

$$\lambda(\vec{x}) = \frac{\sup_{\vec{\beta}\in\mathcal{A}}\{L(\vec{\beta}|\vec{x})\}}{\sup_{\vec{\beta}}\{L(\vec{\beta}|\vec{x})\}}. \tag{5.5}$$

If $\vec{x}$ is distributed according to a model with $\vec{\beta} \in \mathcal{A}$, then, in the limit of a large number of events N, $-2\ln\lambda(\vec{x})$ converges towards a χ^2 distribution with the number of degrees of freedom is given by $n - \dim(\mathcal{A})$, the difference of dimensionality of the sets the likelihood ratio is calculated for. In the case of a signal search, $\mathcal{A}$ represents the hypothesis with a fixed cross section β_1 and therefore $k = 1$.

This asymptotic property can be used for the mentioned hypothesis test. For given data $\vec{x}$, $-2\ln\lambda(\vec{x})$ is calculated. A larger value means less compatibility with $\mathcal{H}_0$. Now, the null hypothesis is rejected if the probability of obtaining a value larger than the measured one assuming $\mathcal{H}_0$, is less than the significance level α.

The $1-\alpha$ quantile of the χ^2_k cumulative distribution, y_0, defines the critical region: if $-2\ln\lambda \geqslant y_0$, the null hypothesis is rejected. The probability of rejecting $\mathcal{H}_0$, if it is true, is α, as desired. Note that the value of y_0 does not depend on the likelihood function but only on k and α.

To construct confidence intervals for the signal cross section β_1, choose a fixed α, say $\alpha = 0.05$, and construct the so-called *profile likelihood function* in β_1,

$$\hat{L}(\beta_1|\vec{x}) = \max_{\{\beta_2,\ldots\beta_n\}} L(\vec{\beta}|\vec{x}). \tag{5.6}$$

Now, for each fixed value of β_1, the hypothesis test is carried out as described above, where the null hypothesis $\mathcal{H}_0$ is the hypothesis that the signal cross section is given by β_1. Given data $\vec{x}$, the set of values of β_1 for which the null hypothesis is *not* rejected is the *confidence interval* for β_1 at the confidence level $1-\alpha$.

More generally, a confidence interval for the parameter β_1 at confidence level $1-\alpha$ is defined as an interval $(\theta_1(\vec{x}), \theta_2(\vec{x}))$ with the property that the true value of β_1 is contained in the interval with probability $1-\alpha$ for all true values of β_1.

5.1.3 Bayesian Statistics

Here, another method is also used to construct confidence intervals,[2] the so-called *Bayesian* method. Before applying this method to the model, a brief review of how the Bayesian view differs from the frequentist is in order.

[2]Note that the intervals are not strictly confidence intervals in the sense defined above. Rather, they express a degree of belief, as will be discussed.

In traditional or *frequentist* statistics, probability of an event A is defined as the ratio $p(A) = N_A/N$ in the limit of an ensemble consisting of an infinite number of experiments N where event A is realized N_A times. If estimating a true parameter β_t, it does not make sense to assign a probability to different values of β_t because there is only *one* true value, and the ensemble definition is not applicable. A procedure for deriving a confidence interval $[\beta_1, \beta_2]$ for the true value β_t at a given confidence level α must ensure that the coverage probability[3] is α for *all* possible values of β_t, particularly for the one true (yet unkown) value.

In Bayesian statistics, on the other hand, the definition of probability is extended and refers to the *degree of belief*. With this definition, it is reasonable to specify a probability for the true value β_t. As will be discussed below, one has to make a-priori assumptions for the value of β_t in order to determine the *posterior density* $p(\beta_t)$ for the true value β_t with data. This a-priori assumption is a probability density π for β_t. It expresses the degree of belief about the possible values of β_t *before* analyzing the data. The main criticism on the Bayesian approach is that the choice of π is not "objective". A Bayesian confidence interval with confidence level α for β_t is an interval with posterior probability α. Frequentist coverage properties of this interval are not guaranteed.

To apply the Bayesian definition, the posterior probability density for the true value of $\vec{\beta}$, $\vec{\beta}_t$, has to be derived, given some data $\vec{x}$. This is done using Bayes' Theorem,

$$p(\vec{\beta}_t|\vec{x}) = \frac{p(\vec{x}|\vec{\beta}_t)\pi(\vec{\beta}_t)}{p(\vec{x})}, \tag{5.7}$$

where $p(\vec{x})$ is merely a factor of normalization which ensures that the right hand side is a proper probability density. $p(\vec{x})$ can be formally written as

$$p(\vec{x}) = \int p(\vec{x}|\vec{\beta}_t)\pi(\vec{\beta}_t)\,d\vec{\beta}_t. \tag{5.8}$$

$p(\vec{x}|\vec{\beta}_t)$ is the probability to observe $\vec{x}$, given $\vec{\beta}_t$, so it is just the model as introduced in equation 5.1. $\pi(\vec{\beta}_t)$ is the prior for $\vec{\beta}_t$.

It is not necessary for the prior π to be an actual probability density, as π only appears together with the model probability $p(\vec{x}|\vec{\beta}_t)$. If the model probability is falling fast enough for large $\vec{\beta}_t$, the value of π becomes unimportant. It is therefore often possible to choose a flat prior in the parameter of interest. The signal cross

[3]The coverage probability is the probability of the true value being contained in the given interval, $p(\beta_1 \leqslant \beta_t \leqslant \beta_2)$.

section is determined by β_1. For this parameter, define its prior

$$\pi_1 := 1,$$

assuming that the total prior factorizes,

$$\pi(\vec{\beta}_t) = \prod_{i=1}^{n} \pi_i(\beta_{i,t}).$$

As the denominator of the right hand side of equation 5.7 is merely a constant of normalization, the posterior density in case of flat priors can be written as

$$p(\vec{\beta}_t|\vec{x}) \propto p(\vec{x}|\vec{\beta}_t) = L(\vec{\beta}_t|\vec{x}), \tag{5.9}$$

where L is the likelihood function introduced in equation 5.4.

Given some data $\vec{x}$, the upper limit on β_1 at confidence level α, $\hat{\beta}_1^{\alpha}$, is defined via the marginal probability of β_1,

$$p(\beta_1|\vec{x}) = \int p(\vec{\beta}|\vec{x})\, d(\beta_2, \ldots \beta_n),$$

such that

$$\int_{-\infty}^{\hat{\beta}_1^{\alpha}} p(\beta_1|\vec{x})\, d\beta_1 = \alpha.$$

This integration can only be done numerically. A straight-forward numerical integration algorithm of a function $f(x)$ over a set $A \subseteq R^n$ is to select N evenly distributed random points x_i from A and evaluate

$$I_N := \frac{\lambda(A)}{N} \sum_i f(x_i), \tag{5.10}$$

where $\lambda(A)$ is the Lesbegue measure of A. However, if $|f(x)|$ is large for a small subset $B \subset A$ and small otherwise, the error on I_N will be large.

The solution is to select x_i non-uniformly, as discussed in the next section.

5.1.4 Metropolis-Hastings Markov Chain Monte-Carlo

The original Metropolis algorithm was applied to calculate numerical integrals over the configuration space of a statistical ensemble in order to evaluate thermodynamical quantities. If done using the microstate, the observables have to be weighted with the Boltzmann factor $\exp(-E/kT)$ for each microstate and the convergence of an integral evaluated using the ansatz in equation 5.10 is not good. The idea was, “instead

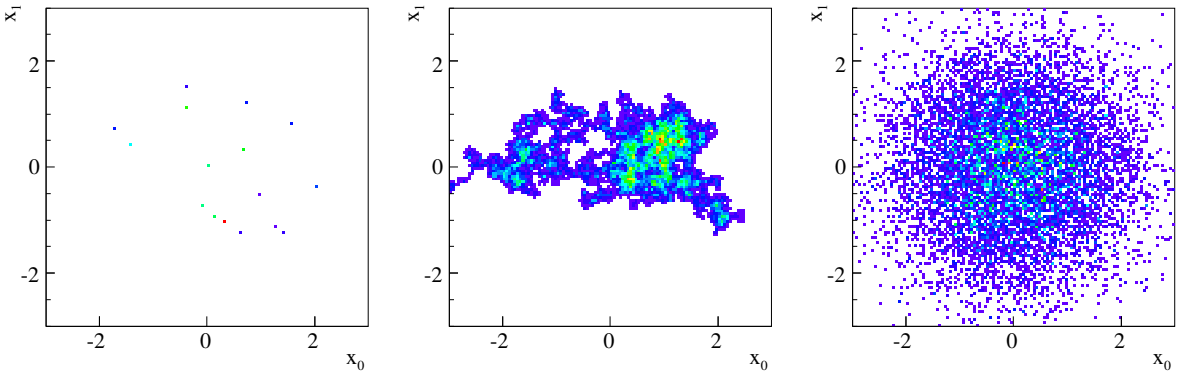

Figure 5.1: Markov chains running on a 2D normal distribution with width one centered at zero. The jump kernel q is Gaussian for all three cases, with different widths σ_q: $\sigma_q = 50$ for the left plot, $\sigma_q = 0.01$ for the central plot, and the optimal width $\sigma_q = 2.38/\sqrt{2}$ (according to equation 5.11) for the right plot. The chain length was $n = 50\,000$ for all three chains.

of choosing configurations randomly, then weighting them with $\exp(-E/kT)$, we choose configurations with a probability $\exp(-E/kT)$ and weight them evenly." [89]

This method was generalized with Markov chains by Hastings [90].

The algorithm produces a sequence of points $x_1, \ldots x_N$ which are distributed according to a probability density $f(x)$. The procedure is as follows:

1. Choose x_1 randomly.
2. Determine a proposed $x' = x_i + y$ where y is chosen according to a symmetric probability density function $q(y)$, called "jump kernel".
3. If $f(x') > f(x_i)$, accpect x' and set $x_{i+1} = x'$.
4. Otherwise, accept x' only with probability $\frac{f(x')}{f(x_i)}$.
5. If x' is not accepted, set $x_{i+1} = x_i$.
6. Go to 2.

The resulting sequence $x_1, \ldots, x_N$ is a Markov sequence as the conditional distribution $p(x_i|x_1, \ldots x_{i-1})$ depends only on x_{i-1}. With standard results from the theory of Markov chains, it can be shown that if $q(y)$ is chosen such that the probability to reach any point in a finite number of steps is non-vanishing, then the sequence $x_1, \ldots x_N$ is distributed according to $f(x)$ [91].

The remaining part of this section addresses the question of how to choose an appropriate jump kernel q and chain length N.

Choosing q Define the *acceptance rate* of a chain as proportion of i for which $x_i \neq x_{i+1}$. A low acceptance rate means that f is small for the proposed x'_{i+1} for these

cases. Clearly, low acceptance rates are to be avoided because only very few points with high multiplicity are in the chain. This case is shown as first plot in figure 5.1. The acceptance rate for this chain was $3.4 \cdot 10^{-4}$.

On the other hand, if the acceptance rate is very high, it means that f is essentially flat in the sampled region. This should be avoided as well as the chain is effectively a random walk covering only a volume $\propto \sqrt{N}$. In this case, the chain length N has to be very high to sample enough points from the tails of f. This case is shown in the middle of figure 5.1 where the chain had an acceptance rate of 99.3%.

In [92], the optimal width for a Gaussian kernel q is derived by maximizing the diffusion speed of the chain in the limit of infinitely many dimensions. It yields a simple result for the optimal standard deviation for q, σ_q:

$$\sigma_q = \frac{2.38 \cdot \sigma_f}{\sqrt{d}}, \tag{5.11}$$

where σ_f is the standard deviation of f and d is the number of dimensions. As one would expect intuitively, the scale of q is in the same order of magnitude as the scale of f. The asymptotic acceptance rate in this limit is 23.4%.

This value for σ_q was used for the right plot in figure 5.1. The acceptance rate in this case was 18.2%.

Usually, σ_f is not known a-priori. However, it can be determined iteratively by using a rough estimate for σ_f for a first chain and subsequently using the variance of the previous chain to estimate σ_f and running a new chain with a value of σ_q chosen according to equation 5.11 for the next chain. This is repeated until the estimate for σ_f converges.

It would also be possible to make "online-updates" for σ_q in the sense that σ_q for the jump attempt $x_i \to x_{i+1}$ is chosen based on the historical chain $x_1, \ldots, x_i$. While this might work well in practice, the resulting chain is not a Markov chain and thus has not automatically the proven convergence property. Therefore, no online update mechanism is used here.

In more than one dimension, the squared standard deviation of f is replaced by the covariance matrix of f. In this case, a multivariate normal distribution can be used for q with the covariance matrix from f, which is estimated iteratively as in the one-dimensional case discussed above.

Choosing a value $\vec{y} = (y_1, \ldots y_d)$ according to the multivariate normal distribution with covariance matrix V is done by computing the Cholesky decomposition of V,

$$V = LL^t,$$

where the matrix L is lower-triangular, and setting

$$\vec{y} = L\vec{n}$$

where the components of $\vec{n}$ are distributed independently according to a normal distribution with mean zero and standard deviation one.

Choosing independent random variables distributed according to a normal distribution is a standard problem for pseudo-number generators. Here, the efficient Ziggurat method [93] which uses uniformly distributed random numbers as input. For that, a fast tausworthe random number generator [94] with seeding modifications described in [95] is used, as implemented in [96].

Choosing N For the present studies, the Markov chain is used to estimate the 95% upper quantile of the marginal distributions for the component representing the signal cross section.

The number of iterations N should be large enough to estimate the quantile q to satisfactory approximation. Unfortunately, the typical "$1/\sqrt{N}$" error cannot be used for Markov chains, as the sequence is correlated by construction.

In [97], an algorithm is presented to determine the number of iterations necessary for estimating the quantile q for a given value x_0, at a desired level of precision r. To determine the 95% upper confidence limit, first a long Markov chain is created and the 95% quantile in the parameter of interest is chosen as x_0. Then, the algorithm is used to estimate the number of iterations necessary to achieve the desired precision for q.

The main idea is to construct a new sequence z_i by setting

$$z_i = \begin{cases} 1 & \text{if } x_i \leqslant x_0 \\ 0 & \text{otherwise.} \end{cases}$$

Note that z_i is not a Markov chain. However, the "thinned" chain $z_i^{(k)}$ constructed by taking every k-th element of the original chain z_i, will behave like a Markov chain, if k is large enough. For such a k, the sequence $z_i^{(k)}$ is described well by a Markov chain consisting of only two states, zero and one. For this chain, it is possible to estimate and extrapolate the error on the quantile q, which has a $1/\sqrt{N_k}$ behavior (where N_k is the length of the chain $z_i^{(k)}$, N/k).

The only task remaining is to find the smallest k such that $z_i^{(k)}$ behaves like a Markov chain. In order to do that, the Bayesian Information Criterion (BIC) [98] is applied. This criterion can be applied more generally to choose between different models which have a different number of free parameters. In this case, the BIC is applied to compare two models describing the behavior of the chain $z_i^{(k)}$: a first-order Monte-Carlo model and a second-order Monte-Carlo model. The smallest k is selected for which the first-order model is preferred by the BIC.

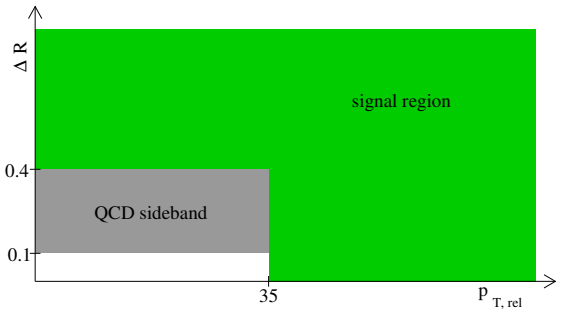

Figure 5.2: Definition of the QCD sideband region. The QCD shape for $M_{t\bar{t}}$ in the final selection and the QCD shape for $H_{T,lep}$ for the low $H_{T,lep}$ region are modelled with events from the shown QCD sideband region.

Now, the Markov chain algorithm can be used to determine upper limits on the signal cross section using the posterior density. However, before doing so, consider an alternative method for estimating the densities $p(x, N|\vec{\beta})$ in equation 5.1 for QCD.

5.2 Data-driven QCD Modelling

As discussed in section 5.1.1, density estimates for the probability densities in $M_{t\bar{t}}$ and the number of expected events after the final selection are required for all processes to calculate the likelihood function. These predictions can be made by applying the analysis on Monte-Carlo data and filling the remaining events in histograms, which are used as the density estimates.

However, QCD models based on Monte-Carlo information only for an event selection requiring reconstructed leptons are usually not very reliable. Therefore, a data-driven model is desirable which predicts the number of events and the distribution of $M_{t\bar{t}}$ for the QCD process in the region of the final selection.

Estimating the $M_{t\bar{t}}$ distribution To estimate the $M_{t\bar{t}}$ distribution for QCD in the signal region, the 2D-cut in the final selection is replaced by the following cut (see left plot in figure 5.2):

1. Require $\Delta R(\mu, j) < 0.4$ and $p_{T,rel}(\mu, j) < 35\,\text{GeV}$. This ensures that there is no overlap between the final selection and the selection for this QCD sideband.
2. Require $\Delta R > 0.1$.

Then, $M_{t\bar{t}}$ is calculated for each event in this QCD sideband. The resulting distribution is used as an estimate of the $M_{t\bar{t}}$ distribution in the signal region.

This method assumes that

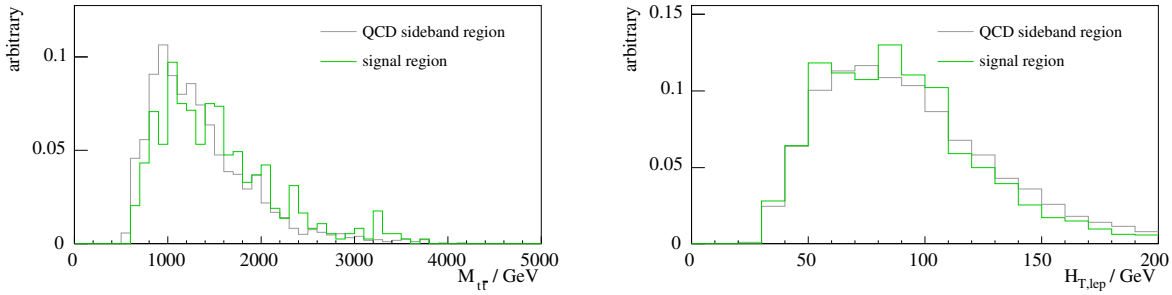

Figure 5.3: $H_{T,lep}$ and $M_{t\bar{t}}$ distributions for the QCD sideband and the final selection. The $H_{T,lep}$ distribution in the QCD region is obtained by inverting the $H_{T,lep}$ cut in the QCD region. For the QCD sideband, all processes have been added, for the final selection, only QCD events are shown.

1. the sideband is highly QCD-enriched and
2. the $M_{t\bar{t}}$ distributions in the sideband region and the final selection are very similar, i. e. $M_{t\bar{t}}$ and the 2D-cut are independent.

The first assumption can be checked by subtracting the other background processes which are presumably well described by Monte-Carlo. For Monte-Carlo, the QCD purity in the sideband is larger than 97%.

The second assumption is valid for Monte-Carlo, as shown in figure 5.3. Of course, this plot cannot be made with data. As a cross-check on data, different variations of the 2D-cut values can be applied to verify that $M_{t\bar{t}}$ is independent of the 2D-cut. However, this can only be done in the QCD sideband where the other backgrounds are negligible. So the independence of $M_{t\bar{t}}$ and the 2D-cut is an assumption that cannot be easily checked with data and so far only relies on Monte-Carlo.

Estimating the number of QCD events To estimate the number of QCD events after the final selection, N_{QCD}, the $H_{T,lep}$ cut is inverted to $H_{T,lep} < 200\,\text{GeV}$. This sample does not only contain QCD events, but also events from the the other samples *W*+jets, Z+jets, $t\bar{t}$, and single top. The distributions for *W*+jets, Z+jets, $t\bar{t}$, and single top are taken from Monte-Carlo, the distribution for QCD is taken from the QCD sideband region with an inverted $H_{T,lep}$ cut.

As one is not interested in deriving the number of events for all processes individually, similar Monte-Carlo distributions are added to caclulate the likelihood function. This is done for *W*+jets, Z+jets, and single top which are combined to a "V+jets" distribution.

The $H_{T,lep}$ distibution of the events in the low $H_{T,lep}$ region is used to construct a likelihood term with the distributions as described. This term mainly constraints the

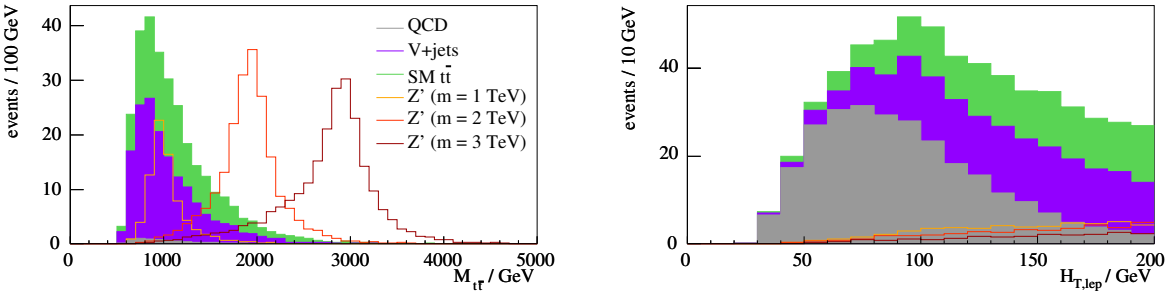

Figure 5.4: The $M_{t\bar{t}}$ distribution after the final selection and the $H_{T,lep}$ distribution in the low $H_{T,lep}$ region, used for the final likelihood function.

number and ratio of QCD and V+jets events.

The final likelihood function for N observed events $\vec{x}$ is

$$
\begin{aligned}
L(\vec{\beta}|\vec{x}, N) &= \frac{\nu(\vec{\beta})^N e^{-\nu(\vec{\beta})}}{N!} G(\beta_{t\bar{t}}, \Delta_{t\bar{t}}) \prod_{\nu=1}^{2} \prod_{j=1}^{N_\nu} p_\nu(x_{\nu,j}|\vec{\beta}) \\
\nu(\vec{\beta}) &= \sum_{i=1}^{n} \beta_i \nu_i \\
p_\nu(x_\nu|\vec{\beta}) &= \frac{1}{\sum_{i=1}^{n} \beta_i} \sum_{i=1}^{n} \beta_i p_{\nu,i}(x_\nu)
\end{aligned}
\tag{5.12}
$$

where n is the number of processes, ν runs from 1 to 2 and denotes the variable type ($M_{t\bar{t}}$ and $H_{T,lep}$), and N_ν is the number of events for the respective variable.

$G(\beta_{t\bar{t}}, \Delta_{t\bar{t}})$ is a Gaussian in $\beta_{t\bar{t}}$ centered at one with width $\Delta_{t\bar{t}}$. It is a constraint on the $t\bar{t}$ cross section. Here, $\Delta_{t\bar{t}} = 0.2$ was chosen. For the actual analysis, a $t\bar{t}$ cross section measurement can be used for $\Delta_{t\bar{t}}$.

Note that such a "constraint" is also valid in frequentist statistics. For this case, consider an analysis which is the same as the one presented here but determines at the same time the $t\bar{t}$ cross section in an appropriate $t\bar{t}$ sideband, in analogy with the QCD sideband. This new likelihood function as function of the $t\bar{t}$ cross section alone is now approximated by the Gaussian in equation 5.12. This interpretation of the constraint will also be useful later, when incorporating the systematic uncertainties in the analysis and the interpretation of the uncertainty $\Delta_{t\bar{t}}$ has to be addressed.

In summary, a method was developed to estimate the number of events for all background processes and their distribution in $M_{t\bar{t}}$ for the signal region. The assumptions entering this method which still rely on Monte-Carlo are:

1. For standard model $t\bar{t}$, the distribution for $H_{T,lep}$ region and for $M_{t\bar{t}}$ is known, as well as the number of events.
2. *W*+jets, Z+jets and single top have the same known distribution for $M_{t\bar{t}}$ and for $H_{T,lep}$. No assumption is made about the cross section.
3. For QCD, the 2D-cut and $M_{t\bar{t}}$ as well as the 2D-cut and $H_{T,lep}$ are independent. No further assumptions about the shape or the cross section is made.
4. For the Z' signals, the distribution in $H_{T,lep}$ and $M_{t\bar{t}}$ is known.

5.3 Results

Given data, it is now possible to calculate the likelihood function as described in the previous section. From the likelihood function, a limit can be derived by the profile likelihood method discussed in section 5.1.2 or by posterior marginalization as motivated in section 5.1.3 with the Markov chain algorithm presented in section 5.1.4.

As real data is not yet available, the method is tested with pseudo-data. This data is generated according to the model in equation 5.12 at the expected rate at the scenario of $L = 200\,pb^{-1}$. No signal events are generated in this case.

With a set of pseudo-data, the limits can be determined as discussed in sections 5.1.2 and 5.1.3. In order to have a reasonable estimate for an *average* case of data, many pseudo-experiments are conducted, yielding a distribution of upper limits. From this distribution, the median is cited as expected limit and the 16% and 84% quantiles as the one-sigma errors.

The results for the profile likelihood method and the Bayesian method are shown in table 5.1. The pseudo-data is drawn at nominal expectation for all processes, except $t\bar{t}$ for which the expectation is shifted randomly according to the constraint, i. e. according to a Gaussian with relative width of 20%. For the likelihood method, 300 pseudo-experiments were conducted and 1 000 for the Bayesian method.

5.4 Systematic Uncertainties

So far, it has been assumed that the input of the model can be determined with high accuracy. This requires in particular that the Monte-Carlo generation of the processes and the detector behave like implemented for the used samples. However, both the theoretical description of the physical processes and the event reconstruction are affected by systematic uncertainties. How those uncertainties can be accounted for in the analysis is the topic of this section.

First, possible sources of systematic uncertainties are discussed. Then, their incorporation is discussed for both presented methods and the resulting expected

	upper limit on σB in pb	
Z' mass in TeV	likelihood method	Bayesian method
1	$6.4^{+3.4}_{-2.3}$	$6.4^{+1.2}_{-2.0}$
2	$0.73^{+0.46}_{-0.32}$	$0.76^{+0.14}_{-0.22}$
3	$0.24^{+0.18}_{-0.08}$	$0.31^{+0.05}_{-0.08}$

Table 5.1: Expected upper limits and 1σ errors for narrow Z' resonances assuming $L = 200\,pb^{-1}$ and no signal. For an explanation of the statistical methods used for derivation of the limits, see sections 5.1.2 and 5.1.3.

upper limits.

5.4.1 Sources of Systematic Uncertainties

Several sources of systematic uncertainties are considered in this study. For most systematic uncertainties, not all Monte-Carlo samples are available. The studied systematic uncertainties and the included samples are

- jet energy scale for all samples
- scale of α_s for W+jets and $t\bar{t}$
- modelling of extra partons for W+jets and $t\bar{t}$
- initial and final state radiation for $t\bar{t}$

The "strength" of a systematic uncertainty i can be parametrized by a parameter δ_i which is zero if the systematic does not apply and ± 1 for the 68% confidence level interval of the assumed uncertainty. To study the effect of systematic uncertainties, Monte-Carlo samples are created for which exactly one δ_i is set to either $+1$ or -1 and the other δ_j are set to zero. How the samples were produced and what $\delta_i = \pm 1$ means physically is explained for each systematic.

Note that the exact choice of parameters which are assumed to cover 68% is conventional and arbitrary to a large extent. The given numbers have been used for the officially produced Monte-Carlo samples for CMS-wide use. The only exception is the jet energy scale uncertainty as this does not require a new Monte-Carlo production. Over the last months, an (implicit) agreement on the strength of this uncertainty has evolved within the top working group which is taken for this study as well.

Jet energy scale (JES) As discussed in section 3.5.4, the jet energy corrections will be derived from data. To reach high accuracy (especially for high p_T), this requires a

large integrated luminosity. A jet energy scale uncertainty of ± 10 % is assumed. To derive the shifted distributions, the four-momenta of all jets are scaled by 1.1 for $\delta = 1$ or 0.9 for $\delta = 0.9$. Through equation 3.2, $\not{E}_T$ depends on the jet energy corrections. This is considered by first adding the transverse momenta of the jets to $\vec{\not{E}}_T$, shifting the jets and $\vec{\not{E}}_T$ with the chosen factor and subtracting them from $\vec{\not{E}}_T$ again. The intermediate shifting of $\vec{\not{E}}_T$ accounts for the error on $\vec{\not{E}}_T$ which is uncorrelated to jet energy scale such as the energy of jets below 10 GeV which are not used for $\vec{\not{E}}_T$ correction in equation 3.2.

Scale uncertainty PYTHIA parameters controlling the shower evolution have been varied: the scale at which the running strong coupling constant α_s is evaluated and the value for Λ_{QCD} which is used to calculate the running of α_s. Those parameters have been changed by a factor of two in both directions.

Extra parton modelling An additional parton in the final state of an interaction can either be simulated as part of the hard matrix element or as part of the parton shower. To avoid possible double counting, a so-called *matching procedure* has to be applied. For this systematic uncertainty, a parameter controlling the matching procedure has been changed.

Initial and final state radiation (ISR/FSR) To simulate more or less initial and final state radiation, a number of shower parameter in PYTHIA have been changed. For *less* ISR/FSR ($\delta = -1$), the default maximum parton virtuality for initial state radiation has been changed from the center-of-mass scale s to the Q^2 of the hard scattering. For more ISR/FSR ($\delta = +1$), Λ_{QCD} for the running strong coupling constant has been set to a different value. Also, the scale at which α_s is evaluated for the initial state radiation is raised.

The systematic uncertainties will change the estimated densities in $M_{t\bar{t}}$ and $H_{T,lep}$ as well as the number of expected events. For most systematic uncertainties, not enough events are available to reliably estimate densities for the sample affected by the uncertainty. However, the densities are often compatible with the original sample, which can be shown using a Kolmogorov test. Therefore, only for systematic uncertainties with samples containing enough statistics, jet energy scale and extra parton modelling, the densities are re-estimated. Otherwise, only the change in acceptance is considered.

The change of acceptance for the different uncertainties is summarized in table 5.2. It refers to the change of acceptance if counting the total number of events in both, (i)

	acceptance change in %							
	JES		scale		extra parton		ISR/FSR	
process	+	−	+	−	+	−	+	−
$t\bar{t}$	43	−28	12	−6	13	9	8	−57
V+jets	49	−37	−23	57	40	−4	—	—
QCD	28	−28	—	—	—	—	—	—
Z' (1 TeV)	17	−17	—	—	—	—	—	—
Z' (2 TeV)	4	0	—	—	—	—	—	—
Z' (3 TeV)	3	−2	—	—	—	—	—	—

Table 5.2: Systematic uncertainties on the selection acceptance, in percent. The + and − columns specify the changes for $\delta = +1$ and $\delta = -1$, respectively. If the systematic uncertainty was not studied for the process, it is denoted — in the table.

the region of the final selection and (ii) the low $H_{T,lep}$ region. These are the events which are used in the likelihood function.

As example of the change a systematic uncertainty has, figure 5.5 shows the expected number with and without jet energy scale.

5.4.2 Profile Likelihood Method

The treatment of systematic uncertainties via likelihood convolution has been used in Z' searches at the Tevatron [99], for the modifications described below, see [100].

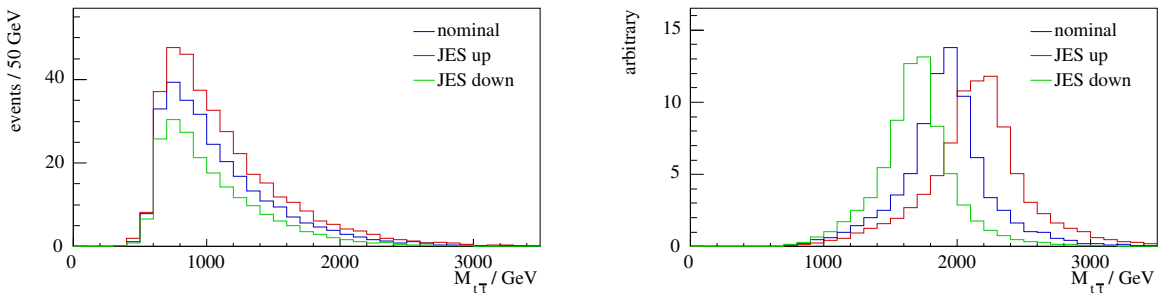

Figure 5.5: Impact of the jet energy scale uncertainty of ±10 % on the shape and expected number of events for QCD (left) and Z' with $m_{Z'} = 2$ TeV (right).

General idea As the systematic uncertainties are independent, their effect on the result is studied by enabling only one of them at a time. For an arbitrary but fixed systematic uncertainty, pseudo-data is generated with signal content β_1. This pseudo-data is used as input to the likelihood function as discussed in section 5.1.2. Particularly, the likelihood is calculated only using density estimates from the Monte-Carlo samples *without* systematic uncertainties.

Maximizing the likelihood function yields an estimate $\hat{\beta}_1$ for the true value of β_1. From the curvature of the likelihood function at the maximum, an error $\sigma_{\hat{\beta}_1}$ on $\hat{\beta}_1$ is estimated. As the pseudo-data is affected by systematic uncertainties, $\hat{\beta}_1$ will systematically deviate from the value of β_1 used to create the pseudo-data by some value Δ.[4] To determine the effect of a systematic uncertainty on limit determination, it is enough to understand the effect of the uncertainty on the *profiled* likelihood $L(\beta_1|\vec{x})$ (section 5.1.2). It is assumed that this effect can be described by this error Δ.

Once Δ has been determined, the profiled likelihood function is convoluted with a Gaussian with width Δ. If β_1 is estimated for the pseudo-data with the *convoluted* likelihood, the estimated value for β_1 will remain unchanged, but the error is calculated with the curvature of the convoluted likelihood and will be $\sqrt{\Delta^2 + \sigma^2_{\hat{\beta}_1}}$. Thus, estimates using the convoluted likelihood will be compatible with the true value, even for data affected by systematic uncertainties.

Modifications So far, it was assumed that Δ is independent of β_1. This is generally not true. Therefore, the convolution of the likelihood function is replaced by the integral

$$L'(\beta_1|\vec{x}) = \int_0^\infty L(\beta_1'|\vec{x}) \frac{1}{N(\beta_1')\sqrt{2\pi}\Delta(\beta_1')} \exp\left(-\frac{1}{2}\left(\frac{\beta_1 - \beta_1'}{\Delta(\beta_1')}\right)^2\right) d\beta_1'. \tag{5.13}$$

Furthermore, the integration is now restricted to the physical region where $\beta_1' \geqslant 0$. For fixed β_1', the Gaussian term in the integrand represents the assumed distribution of the true value of β_1 if the measured value is β_1'. As all β_1-variables are always positive, it is a truncated Gaussian which has to be properly normalized. This is done by $N(\beta_1')$ with

$$N(\beta_1') = \int_0^\infty \frac{1}{\sqrt{2\pi}\Delta(\beta_1')} \exp\left(-\frac{1}{2}\left(\frac{\beta_1' - \beta_1}{\Delta(\beta_1')}\right)^2\right) d\beta_1 = \frac{1}{2}\left(1 + \operatorname{erf}\left(\frac{\beta_1'}{\Delta(\beta_1')}\right)\right).$$

For a fixed systematic uncertainty i, $\Delta_i(\beta_1)$ is determined by scanning over a range of values for β_1 and conducting many pseudo experiments for each value, estimating

[4] Where *systematically* means that averaging over an ensemble of pseudo-experiments yields $\langle \beta_1 - \hat{\beta}_1 \rangle = \Delta \neq 0$.

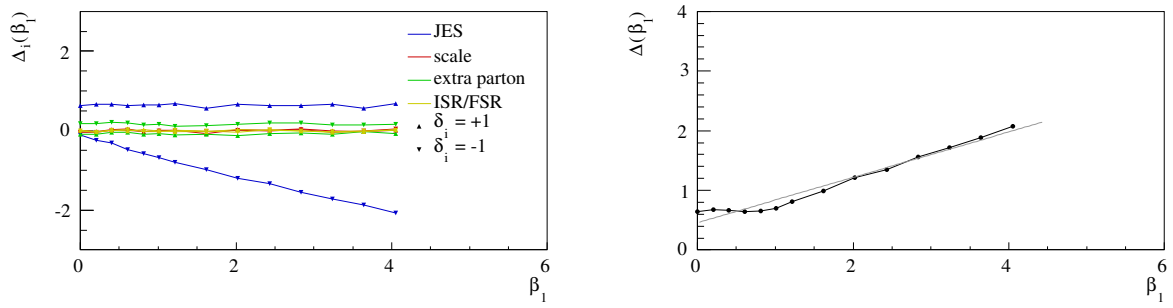

Figure 5.6: The systematic shifts of the signal cross section β_1. For a single systematic source i, the shift $\Delta_i(\beta_1)$ for $\delta_i = \pm 1$ is shown on the left. The total shift $\Delta(\beta_1)$ obtained by symmetrizing and quadratically adding the individual uncertainties. The result is shown in the right plot, with a fitted linear function.

the true value with the likelihood without systematics. The median of the deviation of the estimated and the true value is $\Delta(\beta_1)$. An example of a resulting $\Delta(\beta_1)$ is shown on the left hand side of figure 5.6.

As the systematic uncertainties are two-sided ($\delta = \pm 1$), this is done for $\delta = +1$ and $\delta = -1$ separately. The error for one source of uncertainty is symmetrized by taking the maximum of the two values for $|\Delta_i(\beta_1)|$. The different $\Delta_i(\beta_1)$ thus obtained for each uncertainty are added quadratically which yields the total systematic uncertainty $\Delta(\beta_1)$ which is used in equation 5.13. The total $\Delta(\beta_1)$ and a fitted linear function are shown in the right plot of figure 5.6.

Finally, the interpretation of width for the Gaussian constraint of $t\bar{t}$, $\Delta_{t\bar{t}}$, has to be addressed, which was discussed on page 99. As outlined there, $\Delta_{t\bar{t}}$ can be seen as an approximate likelihood term for a $t\bar{t}$ sideband. Therefore, $\Delta_{t\bar{t}}$ is the *statistical* error of this imaginary sideband measurement and has to be considered in addition to the systematic uncertainties.

5.4.3 Bayesian Method

For the Bayesian treatment, the model from equation 5.12 is modified to include the systematic uncertainties. To do that, the probability density functions for process i, $p_i(\vec{x})$ are altered such that they depend on the parameters $\vec{\delta}$ for the systematic uncertainties. Also, the expected number of events, ν_i depends on $\vec{\delta}$. Those new parameters $\vec{\delta}$ are treated as additional parameters of the mode, i. e. the same as the parameters $\vec{\beta}$, with an appropriate prior.

The dependence chosen here is

$$\nu_i(\vec{\delta})p_i(\vec{x}|\vec{\delta}) = \nu_i p_i(\vec{x}) \prod_{j=1}^{m} \left(\frac{\nu_{i,j}(\mathrm{sgn}(\delta_j))p_{i,j}(\vec{x}, \mathrm{sgn}(\delta_j))}{\nu_i p_i(\vec{x})} \right)^{|\delta_j|}. \tag{5.14}$$

Note that only the product of expectation and density appears. This allows an easy implementation where all $\nu_i p_i$ are histograms normalized to ν_i. The product over j runs over all m systematics. $p_{i,j}$ are density estimates for the process i where systematic j has been enabled. It depends on the sign of δ_j to account for the asymmetry of the systematics:

$$p_{i,j}(\vec{x}, \mathrm{sgn}(\delta_j)) = \mathrm{sgn}(\delta_j)p_{i,j,+}(\vec{x}) - \mathrm{sgn}(\delta_j)p_{i,j,-}(\vec{x}) \tag{5.15}$$

where $p_{i,j,\pm}$ are density estimates for process i with systematic j determined for the samples $\delta_j = \pm 1$.

Equation 5.14 is not the only plausible possibility to parametrize the systematic uncertinaties in the model. However, it has several desirable properties:

- for $\vec{\delta} = 0$, the original model is reproduced,
- if exactly one of the β_j is ± 1 and the others vanish, the density estimate for systematic j (with $\delta_j = \pm 1$) is reproduced. Those two properties ensure that if choosing $\vec{\delta}$ according to a normal distribution around zero with mean one, the model behaves as expected (i. e. consistent with the earlier definition of δ).
- If for all processes i and $\vec{x}$, $p_i(\vec{x}) \neq 0$, then the densities are well-defined for all values of δ_j.

Instead of using the histograms of the variables from Monte-Carlo as density estimates, one could assume a common functional form for the nominal and systematically shifted densities p_i and $p_{i,j,\pm}$ and fit a function to them. The fit parameters can then be written as function of δ_j. However, the process of function selection and fitting has usually to be done manually for each change of the event selection and is thus very time-consuming and hardly automatable. Furthermore, it might be difficult to find a suitable function which could introduce an additional systematic uncertainty.

As priors for the parameters $\vec{\delta}$, uncorrelated Gaussians with mean zero and width one are used. Now, the posterior density is fully specified and can be processed with the Markov chain Monte-Carlo marginalization without further changes.

To generate pseudo-data, values for δ are drawn randomly according to the priors. The scaling parameters $\vec{\beta}$ are set to one, with the exception of $\beta_{t\bar{t}}$ which is chosen randomly according to its constraint in equation 5.12, i. e. according to a Gaussian with mean one and width 0.2.

	upper limit on σB for Z′ in pb		
Z′ mass in TeV	likelihood method	Bayesian method	σB for topcolor Z′ in pb
1	$7.6^{+4.2}_{-2.3}$	$7.5^{+1.4}_{-2.8}$	3.28
2	$2.48^{+0.68}_{-0.37}$	$1.1^{+0.3}_{-0.4}$	0.13
3	$0.75^{+0.13}_{-0.06}$	$0.33^{+0.07}_{-0.10}$	0.0095

Table 5.3: Expected upper limits and 1σ errors for narrow Z′ resonances assuming $L = 200\,pb^{-1}$ and no signal. All systematic uncertainties discussed in section 5.4.1 are accounted for as discussed in sections 5.4.2 and 5.4.3. For the likelihood method, a Gaussian constraint for $t\bar{t}$ has been used and a flat prior in the Bayesian analysis.

5.5 Results Including Systematic Uncertainties

The expected upper limits including all systematic uncertainties are listed in table 5.3 and shown in figure 5.7. Especially for higher masses, the Bayesian method yields lower, less conservative results. This is because this method effectively measures the values of δ for the systematic uncertainties simultaneously and thus limits their impact on the final result.

This robustness of the Bayesian method can be further exploited by replacing the Gaussian prior used for the $t\bar{t}$ cross section by a flat one. This changes the expected upper limit only slightly. The numbers given in table 5.3 have been determined using a flat prior for $t\bar{t}$. In the table, the cross sections for the topcolor Z′ (see section 1.5.1) are also given.

Apart from topcolor Z′, a candidate model for exclusion is the axigluon model [8]. The cross section times branching ratio in $t\bar{t}$ is 45 pb, 1.3 pb and 0.11 pb for axigluon masses of 1, 2 and 3 TeV, respectively. Axigluons of mass M_A have a width of about $\alpha_s M_A$ [101], which cannot be neglected compared to the detector resolution which is of the same order of magnitude (see figure 3.7). Hence, the reconstructed mass distribution will have a width which is about a factor $\sqrt{2}$ larger and upper limits for axigluon cross sections will be higher by about this factor as well. (More detailed studies will have to be carried out to confirm this statement.) This estimate suggests that an exclusion of a 2 TeV axigluon can be done with only small improvements in the analysis of order of 20 % for the limit. Even if this estimate is too optimistic, an improvement of the current axigluon indirect limit of 0.9 TeV [12] should still be possible.

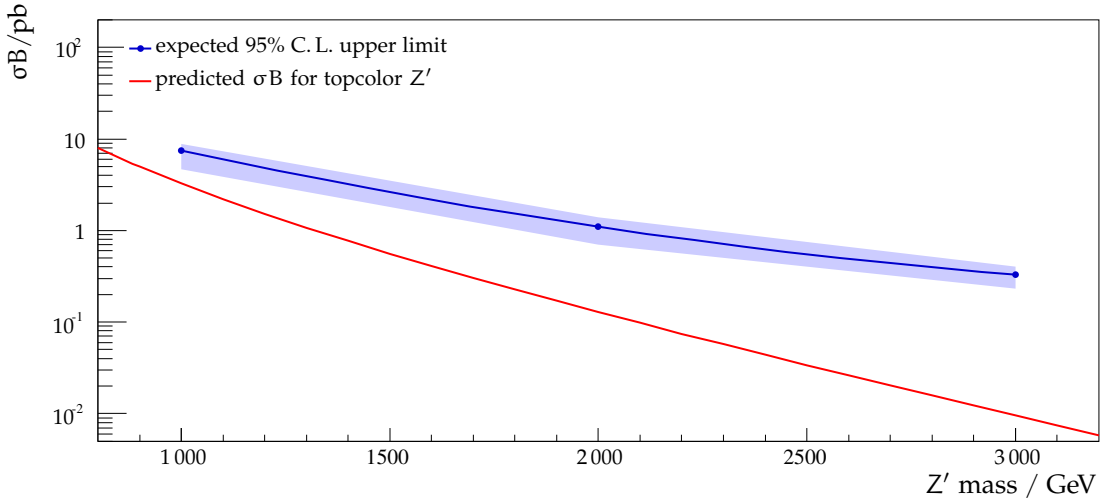

Figure 5.7: The expected upper limits (95% C. L.) on σB for narrow Z' resonances with $L = 200\,pb^{-1}$, including systematic uncertainties.

6 Conclusion and Outlook

In this study, reconstruction techniques for $t\bar{t}$ events with $m_{t\bar{t}}$ at the TeV range at the CMS detector have been developed and studied for a scenario of an integrated luminosity $L = 200\,pb^{-1}$ and 10 TeV center-of-mass energy.

It was shown that the selection and reconstruction used for standard model $t\bar{t}$ studies does not yield optimal results for high $m_{t\bar{t}}$ as the decay products of the top quarks have small angular separation and overlap in the detector.

Possible new cut variables were presented, and a method how to optimize the cut values. It was shown that the newly developed selection method greatly improves the significance compared to a standard model selection.

It was also shown how the top jet reconstruction algorithm could be used for event selection in the μ+jets channel.

Finally, expected upper limits for narrow resonances with masses of 1, 2 and 3 TeV were determined, both with an approximate frequentist method with a Bayesian treatment for systematics and with a purely Bayesian method. For the Bayesian method, a Markov chain Monte-Carlo algorithm was implemented which requires only little configuration and has very good run time behaviour.

The expected upper limits are in the order of a few picobarn or less and an improvement of the current exclusion limits for the topcolor Z' model is well in reach. This analysis can also be extended for broader resonances such as axigluons, for which new mass limits could be set already in the studied luminosity scenario.

This analysis was the first comprehensive study for resonances decaying in $t\bar{t}$ in this mass range and there are still many possible areas for improvements, for example

- the reconstruction technique can be re-optimized for the new selection,
- other jet reconstruction techniques such as particle flow can be studied for possible improvement on resolution,
- for lower masses in the order of $m_{t\bar{t}} = 1\,TeV$, b-tagging algorithms might improve selection and reconstruction,
- the selection can be optimized and applied for each mass point individually,
- the top jet algorithm can be used for reconstruction as well,
- the parameters of the top jet algorithm can be optimized,
- the mistag measurement of the top jet algorithm [88] can be studied for the semileptonic μ channel,

- neural nets can be used for event selection instead of simple cuts (a neural net would also be more suitable for optimizing the selection for a large $m_{t\bar{t}}$ range).

Another area of improvement is the combination of results with other methods, for example a combination with results for low $m_{t\bar{t}}$ using standard model like selection and reconstruction techniques [11] or with a corresponding analysis in the electron channel.

Bibliography

[1] Makoto Kobayashi and Toshihide Maskawa. CP Violation in the Renormalizable Theory of Weak Interaction. *Prog. Theor. Phys.*, 49:652–657, 1973.

[2] S. W. Herb, D. C. Hom, L. M. Lederman, et al. Observation of a Dimuon Resonance at 9.5 GeV in 400-GeV Proton-Nucleus Collisions. *Phys. Rev. Lett.*, 39(5):252–255, Aug 1977.

[3] F. Abe et al. (CDF Collaboration). Observation of top quark production in $\bar{\mathrm{p}}\mathrm{p}$ collisions. *Phys. Rev. Lett.*, 74:2626–2631, 1995.

[4] S. Abachi et al. (DØCollaboration). Observation of the top quark. *Phys. Rev. Lett.*, 74:2632–2637, 1995.

[5] Tevatron Electroweak Working Group. Combination of CDF and DØ Results on the Mass of the Top Quark. 2009.

[6] Dag Gillberg (CDF and DØCollaborations). Discovery of Single Top Quark Production. 2009.

[7] Christopher T. Hill. Topcolor: Top quark condensation in a gauge extension of the standard model. *Phys. Lett.*, B266:419–424, 1991.

[8] Paul H. Frampton and Sheldon L. Glashow. Chiral Color: An Alternative to the Standard Model. *Phys. Lett.*, B190:157, 1987.

[9] Jesse Thaler and Lian-Tao Wang. Strategies to Identify Boosted Tops. *JHEP*, 07:092, 2008.

[10] David E. Kaplan et al. Top Tagging: A Method for Identifying Boosted Hadronically Decaying Top Quarks. *Phys. Rev. Lett.*, 101:142001, 2008.

[11] The CMS Collaboration. Search for narrow resonances in top-pair production close to threshold in the semileptonic muon channel at $\sqrt{s} = 10\,\mathrm{TeV}$. *CMS PAS*, (JME-09-009), 2009. top be published.

[12] German Rodrigo. Axigluon signatures at hadron colliders. *PoS*, RADCOR2007:010, 2007.

[13] Guy Moore and C. Burgess. *The Standard Model – A Primer*. Cambridge University Press, 2007.

[14] David Griffiths. *Introduction to Elementary Particles*. Wiley-VCH, 1987.

[15] Lipatov L N. *Sov. J. Nucl. Phys.* 20:95, 1975.
Gribov V N and Lipatov L N. *Sov. J. Nucl. Phys.*, 15:438, 1972.
Altarelli G and Parisi G. *Nucl. Phys.* B 126:298, 1977.
Dokshitzer Yu L. *Sov. Phys.–JETP* 46:641, 1977.

[16] M. A. Dobbs et al. Les Houches guidebook to Monte Carlo generators for hadron collider physics. 2004.

[17] Torbjorn Sjostrand, Stephen Mrenna, and Peter Skands. PYTHIA 6.4 Physics and Manual. *JHEP*, 05:026, 2006.

[18] M. Bahr et al. Herwig++ Physics and Manual. *Eur. Phys. J.*, C58:639–707, 2008.

[19] Matteo Cacciari, Stefano Frixione, Michelangelo L. Mangano, Paolo Nason, and Giovanni Ridolfi. Updated predictions for the total production cross sections of top and of heavier quark pairs at the Tevatron and at the LHC. *JHEP*, 09:127, 2008.

[20] Rikkert Frederix and Fabio Maltoni. Top pair invariant mass distribution: a window on new physics. *JHEP*, 01:047, 2009.

[21] C Amsler et al. The Review of Particle Physics. *Physics Letters*, B667:1, 2008.

[22] T. Aaltonen et al. (CDF Collaboration). Measurement of W-Boson Helicity Fractions in Top-Quark Decays Using costheta*. *Phys. Lett.*, B674:160–167, 2009.

[23] (ALEPH and CDF and DELPHI and L3 and OPAL and SLD Collaborations). Precision Electroweak Measurements and Constraints on the Standard Model. November 2008.

[24] Michael Spira. QCD effects in Higgs physics. *Fortsch. Phys.*, 46:203–284, 1998.

[25] Lisa Randall and Raman Sundrum. A large mass hierarchy from a small extra dimension. *Phys. Rev. Lett.*, 83:3370–3373, 1999.

[26] Nima Arkani-Hamed, Savas Dimopoulos, and G. R. Dvali. The hierarchy problem and new dimensions at a millimeter. *Phys. Lett.*, B429:263–272, 1998.

[27] Robert M. Harris, Christopher T. Hill, and Stephen J. Parke. Cross Section for Topcolor Z'_t decaying to $t\bar{t}$. 1999.

[28] The DØCollaboration. Search for $t\bar{t}$ Resonances in the Lepton+Jets Final State in $p\bar{p}$ Collisions at $\sqrt{s} = 1.96\,\text{TeV}$. 2008. Preliminary Result for Winter 2008 Conferences.

[29] Robert M. Harris. LHC Cross Sections for Topcolor $Z' \to t\bar{t}$. Via e-mail, August 2008.

[30] Lyndon Evans et al. LHC Machine. *J. Instrum.*, 3:S08001, 2008.

[31] R Adolphi et al. (CMS Collaboration). The CMS experiment at the CERN LHC. *J. Instrum.*, 3:S08004, 2008.

[32] G. L. Bayatian et al. (CMS Collaboration). CMS physics: Technical design report. 1, 2006. CERN-LHCC-2006-001.

[33] CMS Outreach Group. CMS Outreach web page. `http://cms.cern.ch/`, April 2009.

[34] CERN-Education and Technolgy Transfer Division. LIVEfromCERN: Antimatter. `http://livefromcern.web.cern.ch/livefromcern/antimatter/index.html`.

[35] P. Crochet et al. (ALICE Collaboration). The ALICE experiment at the LHC. *Phys. Part. Nucl.*, 39:1074–1081, 2008.

[36] A. Augusto Alves et al. (LHCb Collaboration). The LHCb Detector at the LHC. *JINST*, 3:S08005, 2008.

[37] J M Campbell, J W Huston, and W J Stirling. Hard interactions of quarks and gluons: a primer for LHC physics. *Reports on Progress in Physics*, 70(1):89–193, 2007.

[38] J. Drees. Review of final LEP results or a tribute to LEP. *Int. J. Mod. Phys.*, A17:3259–3283, 2002.

[39] CERN Web Communications. CERN FAQ – LHC the guide. `http://cdsmedia.cern.ch/img/CERN-Brochure-2008-001-Eng.pdf`.

[40] Rolf Heuer. E-mail to CERN personnel, February 2009.

[41] Werner Buckel and Reinhold Kleiner. *Superconductivity*. Wiley-VCH, second edition, 2004.

[42] G Aad et al. (ATLAS Collaboration). The ATLAS Experiment at the CERN Large Hadron Collider. *Journal of Instrumentation*, 3(08):S08003, 2008.

[43] The CMS Collaboration. CMS, tracker technical design report. 1998. CERN-LHCC-98-06.

[44] The CMS Collaboration. CMS: The electromagnetic calorimeter. Technical design report. 1997. CERN-LHCC-97-33.

[45] The CMS Collaboration. CMS: The hadron calorimeter technical design report. 1997. CERN-LHCC-97-31.

[46] The CMS Collaboration. CMS: The computing project. Technical design report. 2005. CERN-LHCC-2005-023.

[47] ProductionSummer2008. `https://twiki.cern.ch/twiki/bin/view/CMS/ProductionSummer2008`, April 2009. Rev. 314.

[48] Paolo Bartalini et al. The 1G Winter09 pp@10TeV Fast Sim Production for Physics with CMSSW 22x. `https://twiki.cern.ch/twiki/bin/view/CMS/ProductionWinter2009`, April 2009. Rev. 95.

[49] Silvano Tosi. Reference Page for Top-Quark Analyses: Samples and Cross-Sections. `http://www.ge.infn.it/~tosi/cms/topMC.html`, 2009.

[50] The CMS Collaboration. Pixel performance in CRAFT. *CMS PAS*, (JME-09-001), 2009. to be published.

[51] The CMS Collaboration. Strip Tracker and tracking performance in CRAFT. *CMS PAS*, (CFT-09-002), 2009. to be published.

[52] R.E. Kalman. A New Approach to Linear Filtering and Prediction Problems. *Trans ASME J Basic Eng D*, 82:35–45, 1960.

[53] G. Abbiendi et al. Muon Reconstruction in the CMS Detector. 2008. Internal note CMS AN-2008/097.

[54] N. Amapane et al. Muon Identification in CMS. 2008. Internal note CMS AN-2008/098.

[55] V+Jets Cross-PAG TWIKI page. `https://twiki.cern.ch/twiki/bin/view/CMS/VplusJets`, April 2009. Rev. 18.

[56] S. Baffioni et al. Electron reconstruction in CMS. *The European Physical Journal C*, 49(4):1099–1116, 2007.

[57] R. Frühwirth and T. Speer. A Gaussian-sum filter for vertex reconstruction. *Nucl. Instrum. and Methods*, A534:217–221, 2004.

[58] CMSSW source code. Version 2_2_3.

[59] Emanuele Di Marco et al. Electron Identification using Likelihood. `https://twiki.cern.ch/twiki/bin/view/CMS/SWGuideLikelihoodElectronID`, 2009. Rev. 15.

[60] Federico Ferri, Leonardno Sala, Roberto Salerno, et al. Electron Identification using Neural Network. `https://twiki.cern.ch/twiki/bin/view/CMS/SWGuideNeuralNetElectronID`, 2009. Rev. 6.

[61] A. Bhatti et al. Performance of the SISCone Jet Clustering Algorithm. *CMS Analysis Note*, 2008. AN-2008/002.

[62] The CMS Collaboration. Performance of Jet Algorithms in CMS. *CMS PAS*, (JME-07-003), 2007.

[63] Gavin P. Salam and Gregory Soyez. A Practical Seedless Infrared-Safe Cone jet algorithm. *JHEP*, 05:086, 2007.

[64] Gerald C. Blazey et al. Run II jet physics. 2000. hep-ex/0005012.

[65] The CMS Collaboration. Particle-Flow Event Reconstruction in CMS and Performance for Jets, Taus and E_T^{miss}. *CMS PAS*, (PFT-09-001), 2009.

[66] The CMS Collaboration. Jet Reconstruction Performance at CMS. *CMS PAS*, (JME-09-007), to be published.

[67] The CMS Collaboration. Plans for Jet Energy Corrections at CMS. *CMS PAS*, (JME-07-002), 2008.

[68] The CMS Collaboration. Offset Energy Correction for Cone Jets. *CMS PAS*, (JME-09-003), to be published.

[69] The CMS Collaboration. Determination of the Relative Jet Energy Scale at CMS from Dijet Balance. *CMS PAS*, (JME-08-003), to be published.

[70] The CMS Collaboration. Jet energy correction using $Z \to ee$ + Jet p_T balance. *CMS PAS*, (JME-09-005), to be published.

[71] The CMS Collaboration. Jet energy correction using $Z \to \mu\mu$ + Jet p_T balance. *CMS PAS*, (JME-09-009), to be published.

[72] The CMS Collaboration. Jet energy calibration with photon+jet events. *CMS PAS*, (JME-09-004), to be published.

[73] The CMS Collaboration. Jet Plus Tracks Algorithm for Calorimeter Jet Energy Corrections in CMS. *CMS PAS*, (JME-09-002), to be published.

[74] The CMS Collaboration. $\not{E}_T$ performance in CMS. *CMS PAS*, (JME-07-001), 2007.

[75] The CMS Collaboration. Missing E_T performance in CMS. *CMS PAS*, (JME-09-006), to be published.

[76] Jonathan M. Butterworth et al. Jet substructure as a new Higgs search channel at the LHC. *Phys. Rev. Lett.*, 100:242001, 2008.

[77] The CMS Collaboration. CATopTag: A Cambridge-Aachen (C-A) based Jet Algorithm for boosted top-jet tagging. *CMS PAS*, (JME-09-001), to be published.

[78] CMSSW source code. Version 3_1_0pre7.

[79] Paola Ferrario and German Rodrigo. Massive color-octet bosons and the charge asymmetries of top quarks at hadron colliders. *Phys. Rev.*, D78:094018, 2008.

[80] Fabio Maltoni and Tim Stelzer. MadEvent: Automatic event generation with MadGraph. *JHEP*, 02:027, 2003.

[81] T. Stelzer and W. F. Long. Automatic generation of tree level helicity amplitudes. *Comput. Phys. Commun.*, 81:357–371, 1994.

[82] S. Agostinelli et al. (GEANT 4 Collaboration). GEANT 4 – a simulation toolkit. *Nucl. Instrum. and Methods*, A506:250–303, 2006.

[83] S. Frixione and M. L. Mangano. How accurately can we measure the *W* cross section? *JHEP*, 05:056, 2004.

[84] John M. Campbell, R. Keith Ellis, and Francesco Tramontano. Single top production and decay at next-to-leading order. *Phys. Rev.*, D70:094012, 2004.

[85] John M. Campbell and Francesco Tramontano. Next-to-leading order corrections to *W*t production and decay. *Nucl. Phys.*, B726:109–130, 2005.

[86] F. James and M. Roos. Minuit: A System for Function Minimization and Analysis of the Parameter Errors and Correlations. *Comput. Phys. Commun.*, 10:343–367, 1975.

[87] Marc Paterno. Calculating efficiencies and their uncertainties. `http://home.fnal.gov/~paterno/images/effic.pdf`, 2003.

[88] The CMS Collaboration. Analysis of boosted tops and backgrounds with data driven methods. *CMS PAS*, (EXO-09-002), to be published.

[89] Nicholas Metropolis, Arianna W. Rosenbluth, Marshall N. Rosenbluth, Augusta H. Teller, and Edward Teller. Equation of state calculations by fast computing machines. *The Journal of Chemical Physics*, 21(6):1087–1092, 1953.

[90] W. K. Hastings. Monte Carlo sampling methods using Markov chains and their applications. *Biometrika*, 57(1):97–109, 1970.

[91] Persi Diaconis. The Markov chain Monte Carlo revolution. *Bull. Amer. Math. Soc.*, 46:179–205, 2009.

[92] A. Gelman, G. O. Roberts, and W. R. Gilks. Efficient metropolis jumping rules. *Bayesian Statistics*, 5:599–607, 1996.

[93] George Marsaglia and Wai Wan Tsang. The Ziggurat Method for Generating Random Variables. *Journal of Statistical Software*, 5(8):1–7, 10 2000.

[94] Pierre L'Ecuyer. Maximally Equidistributed Combined Tausworthe Generators. *Math. Comp.*, (65):203–213, 1996.

[95] Pierre L'Ecuyer. Tables of Maximally Equidistributed Combined LFSR Generators. *Math. Comp.*, (68):261–269, 1999.

[96] M. Galassi et al. GNU Scientific Library Reference Manual (3rd Ed.). `http://gnu.org/software/gsl/`.

[97] Adrian E. Raftery and Steven M. Lewis. The number of iterations, convergence diagnostics and generic metropolis algorithms. In *Practical Markov Chain Monte Carlo (W.R. Gilks, D.J. Spiegelhalter et al.)*, pages 115–130. Chapman and Hall, 1995.

[98] Gideon Schwarz. Estimating the Dimension of a Model. *Ann. Statist.*, 6(2), 1978.

[99] Anthony Allen Affolder et al. (CDF Collaboration). Search for new particles decaying to $t\bar{t}$ in $p\bar{p}$ collisions at $\sqrt{s} = 1.8\,\text{TeV}$. *Phys. Rev. Lett.*, 85:2062–2067, 2000.

[100] T. Aaltonen et al. (CDF Collaboration). Limits on the production of narrow $t\bar{t}$ resonances in $p\bar{p}$ collisions at $\sqrt{s}$ = 1.96 TeV. *Phys. Rev.*, D77:051102, 2008.

[101] Debajyoti Choudhury, Rohini M. Godbole, Ritesh K. Singh, and Kshitij Wagh. Top production at the Tevatron/LHC and nonstandard, strongly interacting spin one particles. *Phys. Lett.*, B657:69–76, 2007.

Danksagungen

Diese Arbeit wäre nicht möglich gewesen ohne die Unterstützung einiger Personen, denen ich an dieser Stelle meinen Dank aussprechen möchte.

An erster Stelle möchte ich mich bei Prof. Dr. Thomas Müller für die Möglichkeit, diese Arbeit in seiner Arbeitsgruppe anzufertigen und die Gelegenheit, an Workshops und Konferenzen teilzunehmen, bedanken.

Desweiteren möchte ich mich bei Prof. Dr. Günter Quast für die Übernahme des Korreferats bedanken.

Ein besonderer Dank gilt Dr. Jeannine Wagner-Kuhr für die vielen hilfreichen Diskussionen und Anregungen, für die sie sich stets viel Zeit nahm sowie für das ausführliche Korrekturlesen dieser Arbeit. Allen Mitgliedern der Top-Gruppe danke ich für die angenehme Arbeitsatmosphäre. Insbesondere möchte ich mich bei Julia und Thorsten für das Korrekturlesen dieser Arbeit und für die gute Zusammenarbeit im Büro bedanken, bei der unzählige Fragen schnell geklärt werden konnten.

Abschließend möchte ich mich ganz besonders bei Thomas bedanken. Die Zusammenarbeit mit ihm war von Anfang an stets sehr angenehm und konstruktiv.

Hiermit versichere ich, dass ich diese Arbeit selbständig verfasst und keine anderen als die angegebenen Hilfsmittel verwendet habe.

Jochen Ott, Karlsruhe, den 26.06.2009