The Search for a Heavy Top-Like Quark

by

Michael M. H. Luk

Sc.M., Brown University, 2010
M.A.St., University of Cambridge, 2008
B.Sc., A.R.C.S., Imperial College London, 2007

Thesis

Submitted in partial fulfillment of the requirements for
the Degree of Doctor of Philosophy
in the Department of Physics at Brown University

Providence, Rhode Island
May 2013

Abstract of “The Search for a Heavy Top-Like Quark” by Michael M. H. Luk, Ph.D., Brown University, May 2013

The Standard Model (SM) has stood as one of the cornerstones of modern physics for the past 40 years; yet it is likely to remain only as an effective field theory. In particular, new beyond the SM theories are required to explain both the baryon-antibaryon asymmetry observed in the Universe and the unnatural lightness of the tentatively observed Higgs boson. Such theories often predict heavy partners to the six observed quarks, which naturally explain these phenomena.

In this thesis, we present two searches for a heavy top-like quark at the LHC using data collected by the CMS detector. In 2011, we searched for a sequential fourth generation t' using $5\,\mathrm{fb}^{-1}$ of data at a $\sqrt{s} = 7$ TeV centre of mass energy [1]. In 2012, we searched for a vector-like T that partners the SM top quark using $19.6\,\mathrm{fb}^{-1}$ of data at $\sqrt{s} = 8$ TeV [2,3]. In both analyses, we used a semileptonic event topology, consisting of events with exactly one lepton. No excess over SM predictions are observed and the heavy top-like quark is ruled out below the mass of $700\,\mathrm{GeV}/c^2$ across the entire phase space. This is the first analysis of its kind and we set the most stringent limit on this entire class of beyond the SM theories to date.

This dissertation by Michael M. H. Luk is accepted in its present form by
the Department of Physics as satisfying the
dissertation requirement for the degree of
Doctor of Philosophy

Date........................ ....

Ulrich Heintz, Advisor

Recommended to the Graduate Council

Date........................ ....

Meenakshi Narain, Reader

Date........................ ....

Gerald Guralnik, Reader

Approved by the Graduate Council

Date........................ ....

Peter M. Weber
Dean of the Graduate School

Acknowledgments

Over the course of my academic career, I have been helped by countless people and whilst it is impossible to thank everyone individually, there are a few I would be remiss not mention.

First and foremost, I would like to thank my advisor Professor Ulrich Heintz for his wisdom, understanding, and unwavering support. To my readers Professors Meenakshi Narain and Gerald Guralnik, thank you for taking the time to enlighten me in both politics and particle physics in equal measure. Gerry, the mass of the pion is... well, actually I'm still not sure.

To the CMS collaboration, thank you for the opportunities and experiences; I hope that you will not break down without me! I would like to thank Dr Alexey Ferapontov for assisting with the Monte-Carlo generation of my analysis, Dr Norraphat Srimanobhas for his frequent help battling CMSSW, Dr Gennadiy Kukartsev for sharing his unlimited supply of statistical knowledge, and Dr Daniel Ferrante for the numerous conversations about physics and all things besides. I would also like to thank all my fellow post-graduates; in particular, Dr Michael Segala, for the countless exceptional discussions and for sharing his enthusiasm of CERN. In addition, I am grateful to both Edith Hanson and Christiana Christodoulou for reading my thesis and for providing much needed feedback.

From Cheltenham College, I would like to acknowledge my physics teachers Roger Lambert, Dominic Faulkner, and Tom Adams. I am indebted to them for their tireless enthusiasm and endless encouragement, without which I would never have discovered the magnificent world of physics. I shall be eternally obliged to my tutor, Zoe Roberts (née Powell), who spent untold hours outside class practising my french and for helping me on my first steps outside Cheltenham. And from Imperial College, I would like to thank Dr Edward Grace for both his infinite guidance and for the many shared pints across the years.

I reserve special thanks to my parents and family; they have given me all the freedom and opportunities that one could ever have. Finally, I would like to express my gratitude to Dr Helen Hanson for her patience and love - you have pushed me to be more than I thought possible.

The masses of the charged pions are 139.57018 ± 0.00035 MeV/c^2 and the mass of the neutral pion π^0 is 134.9766 ± 0.0006 MeV/c^2 [4].

To the discovery of more than just particles.

Preface

Since the 1970s, the Standard Model (SM) has proved one of the most accurate and predictive physical theories to date. However, new beyond the Standard Model (BSM) theories such as Supersymmetry (SUSY) and String Theory are already required to explain certain phenomena at the Planck scale. These are but two of the leading candidates for Grand Unified Theories (GUT) in a hugely rich field of research.

Indeed, we are fast approaching a time where the Standard Model may break down, where observations of inexplicable experimental results will call for additional physics to be present. Dark energy, the cosmological constant, strong CP, and hierarchy problems are but a few of the issues that we have already encountered at our relatively low energy scales, which hint at some grander underlying structure in nature.

Part I of this thesis is dedicated to the foundations of the theory [5, 6]. Part II formulates the Standard Model [7–9] and Part III discusses the Large Hadron Collider (LHC) and the Compact Muon Solenoid (CMS) detector [10–12]. Finally, Parts IV and V will discuss beyond the Standard Model physics and the searches for their signatures at the CMS detector.

An introduction to Group Theory and the classification of particles can be found in Appendix A. The Lorentz transformations are described in Appendix B and finally additional material on the search for a heavy vector-like T quark can be found in Appendix C.

I Notations and Conventions

Throughout this thesis we shall be using the $\mathbf{diag}(+,-,-,-)$ signature.

Greek indices will typically run over all space-time coordinates i.e. $\mu, \nu = 0, 1, 2, 3$ whilst Roman indices $i, j = 1, 2, 3$ the three spatial components. Unless otherwise stated, we shall use the Heaviside-Lorentz units where $\epsilon_0, \mu_0 = 1$ and natural units where $\hbar = c = k_B = 1$. One can easily restore these constants in any equations through simple dimensional analysis.

Indices will follow Einstein's summation convention where a contravariant (superscript) index with a covariant (subscript) one indicates the implicit sum

$$A^i B_i = g_{ij} A^i B^j \equiv \sum_i A_i B_i = A \cdot B. \tag{1}$$

Expressed in terms of matrices, a Hermition operator H is defined by

$$H^\dagger = H \tag{2}$$

where $\dagger \equiv (T\star)$ is known as the Hermition conjugate, equivalent to the operation of complex conjugating the transpose. Additionally, unitary operators U are defined by

$$U^\dagger U = \mathbb{I} \tag{3}$$

II Length Scales

We shall also use the common notation $[\psi]$ to mean the mass dimension of the quantity ψ. In natural units the following relationships between mass M, length L and time T hold:

$$\begin{aligned} [M] &= [L]^{-1} && \text{(4a)} \\ [L] &= [T] && \text{(4b)} \\ [E] &= [M] && \text{(4c)} \end{aligned}$$

We can thus convert all units to one independent dimension.

The fundamental constants of nature are

$$\begin{aligned} [c] &= LT^{-1} \\ [\hbar] &= L^2 M T^{-1} \\ [G] &= L^3 M^{-1} T^{-2} && \text{(5a)} \end{aligned}$$

where the speed of light $c = 3 \times 10^8$ m s^{-1}, Planck's constant $\hbar = 1.0545 \times 10^{-34}$ m^2 kg s^{-1}, and the gravitational constant $G = 6.67398 \times 10^{-11}$ m^3 kg^{-1} s^{-2}.

Then, converting to mass units, we discover that the Gravitational constant is the only one that remains dimensionful. We can express G in mass units as

$$G = \frac{\hbar c}{M_{pl}^2} = \frac{1}{M_{pl}^2} \tag{6}$$

where $M_{pl} \sim 10^{19}$ GeV is known as the Plank mass. Therefore $G \sim 10^{-38} - 10^{-39}$ GeV, which defines the characteristic strength of the gravitational force. Equivalently, it sets the length scale $l_{pl} \sim 10^{-33}$ cm of gravity. Below this length, our conventional theory is expected to break down and a quantum theory of gravity will dominate. On the other extreme of the length scale is the cosmological horizon with $\sim l_{\text{cosmo}} \approxeq\ 10^{80} l_{pl}$.

The characteristic energy (mass) scale is typically set by the bosonic fields of the gauge theories as these naturally arise from the symmetries of nature. Below these energy scales, one would typically observe merely an effective theory. Historically, the original Fermi theory of weak interactions was developed from experimental observations [13]. It was found to be only an effective theory of what is now known as the electroweak force [14], which is mediated by the weak bosons $W^{\pm}, Z^0$. The masses of these particles define the cut-off regime below which the effective theory is still valid.

Since the gluons are massless, the natural scale of quantum chromodynamics is $\mathcal{O}(1)$ GeV set by the pion, which are the pseudo-Goldstone bosons of the theory.

The characteristic strength of the electromagnetic force is determined by the dimensionless fine-structure constant $\alpha = e^2/4\pi\epsilon_0\hbar c \sim 10^{-2}$.

We will see in Part IV that there are already some experimental and theoretical hints that the Standard Model of particle physics is merely an effective one. In Chapter 4, we shall motivate why the Standard Model is expected to be valid only to the TeV scale - the regime where the LHC is currently probing.

We summarise the scales in Fig. 1 on a logarithmic plot where energy increases to the right, and equivalently, length increases to the left.

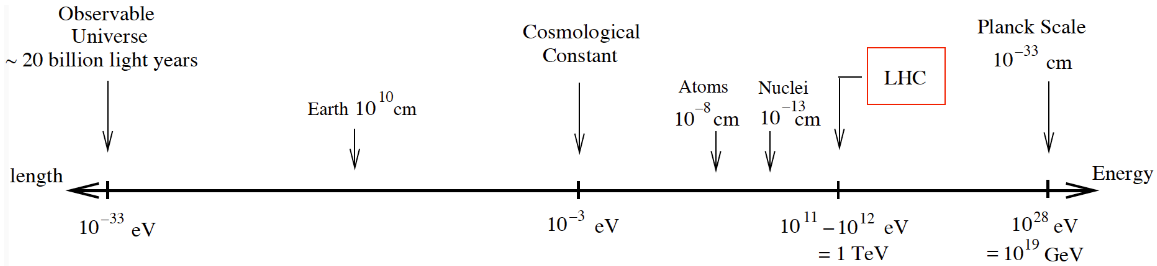


Figure 1: Shows the length scales involved in the Universe [6].

III Group Axioms

Definition A group G is a set of elements $(g_0, g_1, ...)$ with a multiplication law that obeys the following four properties:

1. Identity - There exists a unique element $e \in G$ such that for every element $g \in G$ the equation $g \cdot e = e \cdot g = g$ holds.

2. Inverse - For each element $g \in G$, there exists an element $g' \in G$ such that $g \cdot g' = g' \cdot g = e$.

3. Associativity - The multiplication law is associative such that the equation $(g_1 \cdot g_2) \cdot g_3 = g_1 \cdot (g_2 \cdot g_3)$ holds $\forall g_1, g_2, g_3 \in G$.

4. Closure - The set is closed under multiplication such that $g_1 \cdot g_2 = g \in G$ holds $\forall g_1, g_2 \in G$.

We shall typically be concerned with Lie groups, which are a special type of group, with both an an analytic multiplication map and continuous group elements.

IV The Postulates of Quantum Mechanics

The postulates of quantum mechanics (QM):

- For every dynamical system there exists a wave function ψ that is continuous, square integrable, and single valued function of the parameters of the system, from which all possible predictions about the physical properties of a system can be obtained.

- Measurements of the state of the system are performed by $\hat{H}\psi = E\psi$, where $\hat{H}$ the Hamiltonian and E the energy of the system. The notation $\hat{}$ is used to indicate that the Hamiltonian is an operator (often dropped when unambiguous). This is an eigenvalue equation which gives stored information of the state.

- Every dynamical variable may be represented by a Hermition operator whose eigenvalues represent the possible results of carrying out a measurement of the dynamical variables.

- Immediately after such a measurement the wave-function of the system collapses to a particular state, identical to the eigenfunction corresponding to the eigenvalue obtained as a result of that measurement.

- Between measurements the system evolves according to the Time Dependent Schrödinger Equation (TDSE): $\hat{H}\psi = i\hbar\frac{\partial\psi}{\partial t}$.

Two observables are said to be "compatible" if their operators have a common set of eigenstates. They commute if they do so, otherwise there is an uncertainty relation for these two observables. In terms of observables, the wave-function collapses to one of the basis eigenstates upon measurement and therefore two variables can only be simultaneously measured if they commute, otherwise there is an inherent uncertainty in the measurements.

In particular, the operators that represent the position and the momentum of a particle in one-dimension are $\hat{x}$ and $\hat{p} \equiv -i\hbar\frac{\partial}{\partial x}$ respectively. These act on a state by $\hat{x}\psi = x\psi$ and $\hat{p}\psi = p\psi$, giving the position and the momentum of the measured particle. Since these are incompatible observables, there is an uncertainty relation:

$$[\hat{x}, \hat{p}] \equiv \hat{x}\hat{p} - \hat{p}\hat{x} = i\hbar \tag{7}$$

known as the Heisenberg uncertainty relation.

Contents

List of Tables

List of Figures

Part I

Basics

Chapter 1

Symmetry

Symmetries have been used as an invaluable tool by physicists over the past century. Particle physicists apply Group Theory, the mathematical language of these symmetries, to simplify otherwise intractable calculations.

In this approach, elements of a group transform one state to another whilst leaving the total system invariant. In general, we can construct matrix representations of the group by identifying the abstract group elements with linear matrices and identifying the group operation with the standard matrix multiplication. The finite set of basis eigenstates for these matrices then furnishes the vector space on which these transformations act. The group and the vector space together describe irreducible representations that form the building blocks of all physical theories.

1.1 Space-Time Symmetries

There are two types of symmetries that govern all particles. First, the symmetries of space-time are *global* large scale symmetries that are independent of position. The second are internal *local* symmetries, known as "gauge" symmetries, that describe three of the four fundamental forces that control interactions in particle physics.

We shall examine both classes of symmetries in the absence of the final fundamental force, gravity. We do so for two reasons: first, gravity is very weak, and therefore will not play a major role in dynamics of particle interactions at the subatomic level. Second, it belongs to a class of complicated and poorly understood theories that are what is known as non-renormalisable [8].

We assume that particles and their interactions are always invariant under these space-time symmetries. This will ensure that our equations that govern the laws of Nature shall be relativistic, satisfying Einsteins equations of Special Relativity. These global symmetries are defined by the Poincaré group, which describe the symmetries of translations, rotations and boosts in flat space.

1.1.1 The Lorentz Group

The rotations and boosts of flat-space form the Lorentz group, denoted $SO(1,3)$. This group naturally arises from Einstein's Special Relativity, which requires that all reference frames were equivalent.

A metric is used to parameterise distances and angles in space. The solution for flat space $\mathcal{M}^4$ from Einstein's equations [15] gives the Minkowski metric

$$\eta_{\mu\nu} = \eta^{\mu\nu} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \tag{1.1}$$

which defines the infinitesimal path length

$$ds^2 = \eta_{\mu\nu} dx^\mu d^\nu x \ = dt^2 - dx^2 - dy^2 - dz^2 \tag{1.2}$$

Further, the metric can be used to "raise" and "lower" any space-time index. In flat-space, we use the Minkowksi metric: $x^\mu = \eta^{\mu\nu} x_\nu$ and similarly $x_\nu = \eta_{\mu\nu} x^\mu$.

Notice that the metric η can be redefined by an additional sign change and an appropriate redefinition of the coordinates [16]. However, any physical result must be independent of such a choice and we shall proceed with (1.2).

The set of linear transformations $\Lambda : M^4 \to M^4$ where $\Lambda \in SO(1,3)$ are real matrices that transform the coordinates $x'^{\mu} = \Lambda^{\mu}{}_{\nu} x^{\nu}$, leaving the lengths invariant. Physically, this means that the path length of the space-time to be independent of the reference frame.

Therefore, we require

$$\begin{aligned} \eta_{\mu\nu} x'^{\mu} x'^{\nu} &= \eta_{\mu\nu} \Lambda^{\mu}{}_{\sigma} \Lambda^{\nu}{}_{\rho} x^{\sigma} x^{\rho} \\ &= \eta_{\mu\nu} x^{\mu} x^{\nu} \end{aligned} \tag{1.3}$$

For the last equality to follow, the metric must also be invariant under the Lorentz transformation such that

$$\Lambda^{\mu}{}_{\sigma} \eta_{\mu\nu} \Lambda_{\rho}{}^{\nu} = \eta^{\rho\sigma} \tag{1.4}$$

We will show that the set of Λ form a group under matrix multiplication. Take $\Lambda_1, \Lambda_2 \in SO(1,3)$ as 4×4 matrices that satisfy (1.4). It is straightforward to see that $\Lambda_1 \Lambda_2$ also satisfies (1.4). Additionally, if Λ satisfies (1.4), then $\det(\Lambda) = \pm 1$. The inverse is given by $\Lambda^{-1} = \eta^{-1} \Lambda^T \eta$ and

$$\begin{aligned} (\Lambda^{-1})^T \eta \Lambda^{-1} &= \eta \Lambda \eta^{-1} \eta \eta^{-1} \Lambda^T \eta \\ &= \eta \Lambda \underbrace{\eta^{-1} \Lambda^T \eta}_{\Lambda^{-1}} \\ &= \eta \end{aligned} \tag{1.5}$$

suppressing all indices. Thus Λ^{-1} is also a Lorentz transformation; in particular, $\Lambda_{\lambda}{}^{\nu} \Lambda^{\rho}{}_{\nu} = \delta_{\lambda}{}^{\rho}$. Finally the identity element is simply $\Lambda^{\mu}{}_{\nu} = \delta^{\mu}{}_{\nu}$ and the set of Lorentz transformations indeed form a group.

Now, one can write a generic Lorentz transformation as

$$\Lambda = \begin{pmatrix} \lambda & \underline{\beta}^T \\ \underline{\alpha} & R \end{pmatrix} \tag{1.6}$$

where $\lambda \in \mathbb{R}$, $\underline{\alpha}, \underline{\beta} \in \mathbb{R}^3$, and R is a 3×3 real matrix.

Then, the constraints from (1.4) imply that

$$\lambda^2 = 1 + \underline{\alpha} \cdot \underline{\alpha} \quad , \quad \lambda \underline{\beta} = R^T \underline{\alpha} \quad , \quad R^T R - \underline{\beta}\underline{\beta}^T = \mathbb{I}_3 \tag{1.7}$$

The first constraint can be rewritten as $\lambda = \pm\sqrt{1 + \underline{\alpha} \cdot \underline{\alpha}}$, which means that $\lambda \geq +1$ or $\lambda \leq -1$, with $\underline{\beta} = \frac{R^T \alpha}{\pm\sqrt{1+\underline{\alpha}\cdot\underline{\alpha}}}$.

The Lorentz transformations with ${\Lambda^0}_0 \geq 1$ form a subgroup of the Lorentz group. Specifically,

$$\det \Lambda = +1 \quad \text{and} \quad {\Lambda^0}_0 \geq 1 \tag{1.8}$$

is called the orthochronous Lorentz group, denoted $SO(1,3)^+$.

The proper orthochronous transformations consist only of rotations and boosts that preserve time-like, space-like, and null vectors, which have components that obey $\overline{x}^2 < t^2$, $\overline{x} > t^2$ and $\overline{x} = 0$ respectively.

Note that any general Lorentz transformation can be written as a proper orthochronous one with a discrete (parity and/or time) transformation. We shall therefore restrict our discussions to orthochronous ones and suppress the additional + superscript.

The Lorentz transformation can be written infinitesimally as

$${\Lambda^\mu}_\nu = {\delta^\mu}_\nu + \delta{\omega^\mu}_\nu \tag{1.9}$$

Then the constraint that the metric is invariant yields

$$\eta_{\mu\nu}{\Lambda^\mu}_\rho{\Lambda^\nu}_\sigma = \eta_{\rho\sigma} \implies \delta w_{\rho\sigma} = -\delta\omega_{\sigma\rho} \tag{1.10}$$

The antisymmetric nature of ω implies that there are $4 \times 3/2 = 6$ degrees of freedom. Considering three spatial directions, a generic rotation lives in $SO(3)$ and has three parameters. The remaining three degrees of freedom are are known as the Lorentz "boosts".

We can introduce the six basis matrices ${(\mathcal{M}^{\rho\sigma})^\mu}_\nu$ for these transformations. These are defined by

$${(\mathcal{M}^{\rho\sigma})^\mu}_\nu = \eta^{\rho\mu}{\delta^\sigma}_\nu - \eta^{\sigma\mu}{\delta^\rho}_\nu \tag{1.11}$$

and are antisymmetric in the indices $\rho, \sigma = 0, 1, 2, 3$.

The infinitesimal transformation is now

$$\omega^{\mu}{}_{\nu} = \frac{1}{2}\Omega_{\rho\sigma}\left(\mathcal{M}^{\rho\sigma}\right)^{\mu}{}_{\nu} \tag{1.12}$$

where $\Omega_{\rho\sigma}$ is antisymmetric in $[\rho\sigma]$ and contain six constants, which specify the magnitude under a particular basis transformation.

These basis matrices of the symmetry are known as "generators" of the Lie group and obey the "Lie algebra" defined by the commutation relations

$$[\mathcal{M}^{\rho\sigma}, \mathcal{M}^{\tau\nu}] = \eta^{\sigma\tau}\mathcal{M}^{\rho\nu} - \eta^{\rho\tau}\mathcal{M}^{\sigma\nu} + \eta^{\rho\nu}\mathcal{M}^{\sigma\tau} - \eta^{\sigma\nu}\mathcal{M}^{\rho\tau} \tag{1.13}$$

where we have suppressed the matrix indices.

A finite Lorentz transformation can then be expressed in terms of the exponential

$$\Lambda = \exp\left(\frac{1}{2}\Omega_{\rho\sigma}\mathcal{M}^{\rho\sigma}\right) \tag{1.14}$$

Notice that one can also construct these $\Lambda \in SO^{+}(1,3)$ matrices directly from Einstein's equivalence principle, as shown in Appendix B. Without loss of generality, the Lorentz boost in the z-direction transforms the space-time coordinate (in four-vector notation) as

$$\begin{pmatrix} ct' \\ x' \\ y' \\ z' \end{pmatrix} = \begin{pmatrix} \gamma & 0 & 0 & -\beta\gamma \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ -\beta\gamma & 0 & 0 & \gamma \end{pmatrix} \begin{pmatrix} ct \\ x \\ y \\ z \end{pmatrix} \tag{1.15}$$

where

$$\gamma = \frac{1}{\sqrt{1-\beta}} \quad , \quad \beta = \frac{v}{c} \tag{1.16}$$

In turn, this transformation in special relativity corresponds to the one given by the symmetry generator

$$(\mathcal{M}^{03}) \equiv \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix} \tag{1.17}$$

Further, by defining the rapidity ϕ as

$$e^{\phi} = \gamma(1+\beta) \quad , \quad e^{-\phi} = \gamma(1-\beta) \tag{1.18}$$

we see that the Lorentz transformations are hyperbolic rotations of the coordinates in Minkowski space are given as

$$\begin{pmatrix} ct' \\ x' \\ y' \\ z' \end{pmatrix} = \begin{pmatrix} \cosh\phi & 0 & 0 & -\sinh\phi \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ -\sinh\phi & 0 & 0 & \cosh\phi \end{pmatrix} \begin{pmatrix} ct \\ x \\ y \\ z \end{pmatrix} \tag{1.19}$$

Similarly, this transformation is given by the generator

$$(\mathcal{M}^{12})^{\mu}{}_{\nu} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \tag{1.20}$$

which generates rotations in the (x, y) plane. In general, we define the $\mathcal{M}^{0i}$ as boosts and $\mathcal{M}^{ij}$ for $i, j \neq 0$ as rotations [6]. In this way, the equations of special relativity and the Lorentz transformations are equivalent. Thus, imposing Lorentz invariance will yield a relativistic theory.

Representations of the Lorentz group

In general, all physical particles are described by fields that transform under some representation of the Lorentz group $D[\Lambda]$, where a representation obeys the following axioms:

- $D[\Lambda_1]D[\Lambda_2] = D[\Lambda_1\Lambda_2]$,
- $D[\Lambda^{-1}] = D[\Lambda]^{-1}$,
- and $D[\mathbb{I}] = \mathbb{I}$.

Although the fields can transform non-trivially, we can ensure that our theory is relativistic by simply constraining the permissible interactions between multiple fields to those that are invariant under the Lorentz group. A quantity that is invariant under a Lorentz transformation is known as a Lorentz scalar.

The Lorentz group $S0(3,1)$ is the double cover of $SU(2)$, such that the representations of the Lorentz group can be classified by two $SU(2)$ spin values (j, j') [5]. The smallest irreducible representations are $(0,0)$, $(0,\frac{1}{2})$, $(\frac{1}{2},0)$, and $(\frac{1}{2},\frac{1}{2})$:

- $(0,0)$ is the tensor product of two one-dimensional states; it forms the singlet state of the tensor product representation. The field that transforms trivially under this singlet representation is known as the scalar field.
- $(0,\frac{1}{2})$ is known as the left-handed Weyl representation. It has 2 states and describes the transformation of the spin-$\frac{1}{2}$ field. We call the field that transforms under $(0,\frac{1}{2})$ representation of $SO(1,3)$ the left-handed Weyl spinor field denoted ψ_L.
- Similarly, $(\frac{1}{2},0)$ is the right-handed Weyl spinor representation with two states described by the ψ_R field.
- The spin-1 representation $(\frac{1}{2},\frac{1}{2})$ is four-dimensional and is known as the vectorial or vector-like representation.

We shall now construct Dirac spinors, composed of a left and a right-handed spinor field, which transforms under $(\frac{1}{2},0) \oplus (0,\frac{1}{2})$. All fermions are spin-$\frac{1}{2}$ fields and are described by these Dirac spinors [8].

Dirac spinors transform invariantly under time and parity transformations in addition to the Lorentz symmetry. Parity transformations acts on the weights that define the states as

$$(j, j') \to (j', j) \tag{1.21}$$

Physically, this means that imposing relativity $SO(1,3)$ and the condition that spin-$\frac{1}{2}$ particles transform invariantly under parity requires left and right-handed fields come together (each transforming under opposite-handed irreducible representations of $SU(2)$).

This implies that any theory that is invariant under parity and Lorentz symmetry, such as quantum chromodynamics, can be represented in terms of Dirac spinors. In this case, the parity transformation simply switches the left and right-handed components, leaving the total field invariant.

In theories that are not invariant under parity, for example in the electroweak theory, we use Weyl spinors. The left and right handed fields then transform under independent left-handed $(0, \frac{1}{2})$ and right-handed $(\frac{1}{2}, 0)$ representations respectively. Such theories that treat left and right handed fields differently are known as chiral theories.

We shall return to both the field and particle content of the Standard Model in Chapter 3 but for now, we shall just mention the irreducible representations in which they lie. For a relativistic theory, only certain combinations of fields are permissible in each term of the equations of motion. These are the ones that are in total Lorentz scalars. The specific representation under which each physical field transforms is determined the experimentally observed interactions.

Thus, to couple consistently with all other particles, the Higgs field is chosen to be a scalar field. The fermions exhibiting the effect known as spin, which distinguish them into two equivalent categories, naturally fall into the spin-$\frac{1}{2}$ representation and are therefore spinor-fields. The gauge fields, having four degrees of freedom, are spin-1 vector fields. Finally, the field tensor and the metric tensor are constructed from the other fields and have spin-2. They transform under the higher representations $(1,0) \oplus (0,1)$ and $(1,1)$ respectively.

It turns out that space-time is also translationally invariant. We will now consider the Poincaré group, which is the (semi-direct) product of the group of translations and the Lorentz group.

1.1.2 The Poincaré Group

The Poincaré group is a 10-dimensional Lie group that define the symmetries of flat Minkowski space-time. There are four additional transformations Π_μ ($\mu = 1, 2, 3, 4$), one for each of the directions in space-time.

The Poincaré algebra is governed by the following commutation relations [5]:

$$[\Pi_\mu, \Pi_\nu] = 0 \tag{1.22a}$$

$$\frac{1}{i}[\Lambda_{\mu\nu}, \Pi_\rho] = \eta_{\mu\rho}\Pi_\nu - \eta_{\nu\rho}\Pi_\mu \tag{1.22b}$$

$$\frac{1}{i}[\Lambda_{\mu\nu}, \Lambda_{\rho\sigma}] = \eta_{\mu\rho}\Lambda_{\nu\sigma} - \eta_{\mu\sigma}\Lambda_{\nu\rho} - \eta_{\nu\rho}M_{\mu\sigma} + \eta_{\nu\sigma}\Lambda_{\mu\rho} \tag{1.22c}$$

where (1.22c) by itself defines the Lie algebra for the Lorentz group.

In fact, a theory must be invariant under the Poincaré symmetries (of flat-space) for it to be relativistic. However, since translational symmetries are often straight-forward and typically uninteresting, we shall restrict our discussions to Lorentz invariant theories and assume that the generalisation to the full Poincaré group follows trivially.

1.2 The Principle of Least Action

A "field" $\phi(\bar{x}, t)$ is a function defined at every point of space-time. Typically, they are used to describe particles that will naturally arise upon quantisation, described in the next Chapter 2.

For now, we shall continue by defining the "action"

$$S \equiv \int dt L\left(\phi(t), \dot{\phi}(t)\right) \tag{1.23}$$

where L is the Lagrangian that depends on a field ϕ and its momenta $\dot{\phi}$. The Lagrangian density $\mathcal{L}$ is defined as $L = \int dtd^3x\mathcal{L}$. L and $\mathcal{L}$ are commonly both referred to as the Lagrangian. For now, we shall assume that this parameterises the path along which a particle takes.

Classically, the action is a functional of a particle's coordinates with q and momenta $\dot{q}$ replacing ϕ and $\dot{\phi}$ respectively. This means that the action is dependent on all points along the path $(q, \dot{q})$ of the particle, with only the start $(q(t_i), \dot{q}(t_i))$ and end $(q(t_f), \dot{q}(t_f))$ points fixed. In quantum mechanics, particles are described by fields and the functional integral naturally generalises to an integration over *all* field configurations representing all possible paths.

Note that the form of L is currently arbitrary and just parameterises the space-time that we are considering. However, we shall see that it will describe the path on which particles travel.

Hamilton's principle [17] states that nature extremises the path that a particle takes between points $\phi(t_i)$ and $\phi(t_f)$. In terms of the action, this means that,

$$\begin{aligned} \delta S &= 0 \\ &= \delta \int_{t_i}^{t_f} d^4x\mathcal{L}\left(\phi(x^\mu), \dot{\phi}(x^\mu)\right) \end{aligned} \tag{1.24}$$

where we have written $d^4x = dtd^3x$.

We find,

$$\begin{aligned} S + \delta S &= \int d^4x\mathcal{L}\left(\phi + \varepsilon\delta\phi, \partial_\mu\phi + \varepsilon\partial_\mu\phi\right) \\ \Rightarrow \delta S &= \int d^4x \left(\delta\phi\frac{\partial\phi}{\partial\phi} + \delta(\partial_\mu\phi)\frac{\partial\mathcal{L}}{\partial\phi}\right) \\ &= \int d^4x \left(\delta\phi\frac{\partial\mathcal{L}}{\partial\phi} + \underbrace{\partial_\mu\left(\delta\phi\frac{\partial\mathcal{L}}{\partial\left(\partial_\mu\phi\right)}\right)}_{\star} - \delta\phi\partial_\mu\frac{\partial\mathcal{L}}{\partial\left(\partial_\mu\phi\right)}\right) \end{aligned} \tag{1.25}$$

where $\star$ vanishes when integrated by parts and taking the surface term to drop off sufficiently quickly at infinity.

Taking $\delta S = 0$, which is valid for all values of $\delta\phi$, we arrive at the Euler-Lagrange equations

$$\frac{\partial \mathcal{L}}{\partial \phi_i} = \partial_\mu \frac{\partial \mathcal{L}}{\partial \left(\partial_\mu \phi_i\right)} \tag{1.26}$$

where $\mu = 0, 1, 2, 3$ denotes the space-time index. Notice here that we have generalised to n different fields by appending the label $i = 1, \cdots, n$ and . This gives the "equations of motion" of the fields ϕ_i. It is therefore sufficient to specify the Lagrangian to completely determine a particle's equations of motion.

For consistency with Newton's laws, the form of the Lagrangian is chosen to be

$$L \equiv T - V \tag{1.27}$$

where T and V are the kinetic and potential energy of the system respectively. We shall typically assume this form of the Lagrangian.

Note that Newton's laws were empirical, based on his observations from data. In general, the form of the Lagrangian is somewhat arbitrary. Finding which Lagrangians are representative of nature is one of the crucial components of developing theories in particle physics.

It turns out that, by assuming that nature is symmetric under particular symmetry groups, one can restrict the infinite set of possible Lagrangians to those that provide a very good description of the Universe.

1.2.1 Noether's Theorem

Noether's theorem states that there is a conserved quantity for every continuous symmetry of the action that leaves it invariant.

For a given Lagrangian and symmetry transformation $\phi \to \phi + \varepsilon\delta\phi$, we have

$$\begin{aligned}
\mathcal{L}\left(\phi, \partial_\mu \phi\right) &\to \mathcal{L}\left(\phi + \varepsilon\delta\phi, \partial_\mu\phi + \varepsilon\partial_\mu\phi\right) \\
&= \mathcal{L}\left(\phi, \partial_\mu \phi\right) + \varepsilon\delta\phi\frac{\partial\mathcal{L}}{\partial\phi} + \varepsilon\delta(\partial_\mu\phi)\frac{\partial\mathcal{L}}{\partial\left(\partial_\mu\phi\right)} \\
&= \mathcal{L}\left(\phi, \partial_\mu \phi\right) + \varepsilon\delta\phi\partial_\mu\frac{\partial\mathcal{L}}{\partial\left(\partial_\mu\phi\right)} + \varepsilon\delta(\partial_\mu\phi)\frac{\partial\mathcal{L}}{\partial\left(\partial_\mu\phi\right)} \qquad \text{by Euler-Lagrange} \\
&= \mathcal{L}\left(\phi, \partial_\mu \phi\right) + \varepsilon\partial_\mu j^\mu
\end{aligned} \tag{1.28}$$

If this transformation leaves the action invariant, then $\mathcal{L} \to \mathcal{L}$ and there is a "conserved" current

$$j^\mu = \frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi)}\delta\phi \tag{1.29}$$

where

$$\partial_\mu j^\mu = 0 \tag{1.30}$$

Note that j^0 is identified as the charge density, $\bar{j}$ the current density, and the total charge is defined as

$$Q = \int d^4x j^0 \tag{1.31}$$

We can again generalise the result in (1.29) to n fields, with ϕ_a for $a = 1, \ldots, n$. The continuous symmetry transformation can be written as

$$\delta\phi_a(x) = T_a(\phi) \tag{1.32}$$

where T_a denote the infinitesimal "generators" of the symmetry, which correspond to the basis transformations of the group.

Then (1.29) becomes

$$j^\mu = \frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi_a)}T_a \tag{1.33}$$

The Energy-Momentum Tensor

One of the most important conserved quantities in physics is the energy-momentum tensor that arises from the symmetry of space-time. We consider the infinitesimal translation

$$x^{\nu} \to x^{\nu} - \varepsilon^{\nu} \implies \phi_a(x) + \varepsilon^{\nu}\partial_{\nu}\phi_a(x) \tag{1.34}$$

where the transformation in the field gets an additional sign change, since we are considering "active" transformations [6].

Under this shift, the Lagrangian becomes

$$\mathcal{L}(c) \to \mathcal{L}(x) + \epsilon^{\nu}\partial_{\nu}\mathcal{L}(x) \tag{1.35}$$

Assuming that our equations are symmetric under such a transformation, we can invoke Noether's theorem to give us four conserved currents $(j^{\mu})_{\nu}$, one for each of the translations ε^{ν} with $\nu = 0, 1, 2, 3$ where

$$(j^{\mu})_{\nu} = \frac{\partial \mathcal{L}}{\partial\left(\partial_{\mu}\phi\right)}\partial_{\nu}\phi_a \equiv T^{\mu}_{\ \nu} \tag{1.36}$$

where we have defined $T^{\mu}_{\ \nu}$ the "energy-momentum" tensor (sometimes referred to as the stress-energy tensor). It satisfies,

$$\partial_{\mu}T^{\mu}_{\ \nu} = 0 \tag{1.37}$$

and the four conserved quantities are given by

$$E = \int d^3x T^{00} \quad \text{and} \quad P^i = \int d^3x T^{0i} \tag{1.38}$$

where E and P^i (i=1,2,3) are the total energy and the total momentum of the field configuration respectively.

Chapter 2

Quantisation

2.1 Classical Electrodynamics

In 1865 Maxwell proposed a unified theory for classical electromagnetism [18]. Maxwell formulated his theory in terms of the electric and magnetic fields, denoted $E(x)$ and $\overline{B}(x)$ respectively, with equations

$$\begin{aligned} \nabla \cdot \overline{E} &= \frac{\rho}{\epsilon_0} && (2.1a) \\ \nabla \cdot \overline{B} &= 0 && (2.1b) \\ \nabla \times \overline{E} &= -\frac{\partial \overline{B}}{\partial t} && (2.1c) \\ \nabla \times \overline{B} &= \mu_0 J + \mu_0 \epsilon_0 \frac{\partial \overline{E}}{\partial t} && (2.1d) \end{aligned}$$

where ρ is the charge density, J the current and ϵ_0 and μ_0 are the permittivity and permeability of free space, which we shall set to unity for the remainder of this chapter.

Maxwell's approach was to introduce the "vector potential" A_μ where

$$\overline{B} = \nabla \times \overline{A} \tag{2.2}$$

and

$$\overline{E} = -\nabla \phi - \frac{\partial A}{\partial t} \tag{2.3}$$

Explicitly,

$$A_\mu = (\phi, -\boldsymbol{A}) \tag{2.4}$$

where ϕ is a scalar potential and $\boldsymbol{A}$ is the vector field.

Maxwell embedded this information into the electromagnetic "field strength" defined as

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu \tag{2.5}$$

where

$$F^{\mu\nu} = \begin{pmatrix} 0 & -E_x & -E_y & -E_z \\ E_x & 0 & -B_z & B_y \\ E_y & B_z & 0 & B_x \\ E_z & -B_y & B_x & 0 \end{pmatrix} \quad , \quad F_{\mu\nu} = \begin{pmatrix} 0 & E_x & E_y & E_z \\ -E_x & 0 & -B_z & B_y \\ -E_y & B_z & 0 & B_x \\ -E_z & -B_y & B_x & 0 \end{pmatrix} \tag{2.6}$$

The current is then defined as

$$J^\mu \equiv (\rho, J) \tag{2.7}$$

With these definitions, the Maxwell equations (2.1) are automatically solved using the identities

$$\partial^\lambda F^{\mu\nu} + \partial^\mu F^{\lambda\mu} + \partial^\mu F^{\nu\lambda} = 0 \quad \Longrightarrow \quad \nabla \cdot B = 0 \quad \& \quad \nabla \times E + \frac{\partial B}{\partial t} = 0 \tag{2.8}$$

$$\partial_\mu F^{\mu\nu} = J^\nu \quad \Longrightarrow \quad \nabla \cdot E = \rho \quad \& \quad \nabla \times B - \frac{\partial E}{\partial t} = J^\mu \tag{2.9}$$

One can treat A_μ as a fundamental field with a kinetic term that is quadratic in derivatives. With this in mind, the Lagrangian that gives rise to electromagnetism is

$$\mathcal{L}_{EM} = -\frac{1}{4} F_{\mu\nu} F^{\mu\nu} - J^\mu A_\mu \tag{2.10}$$

with the action

$$S_{EM} = \int d^4x \left(-\frac{1}{4} F_{\mu\nu} F^{\mu\nu} - J^\mu A_\mu \right) \tag{2.11}$$

Extremising the path we find

$$
\begin{aligned}
\delta S &= \int d^4x \left(-\frac{1}{2}F_{\mu\nu}\delta F^{\mu\nu} - J^\mu \delta A_\mu\right) \\
&= \int d^4x \left(-F_{\mu\nu}\partial^\mu \delta A^\nu - J^\mu \delta A_\mu\right) \\
&= \int d^4x \left(-\partial^\mu F_{\mu\nu} - J^\nu\right)\delta A_\mu \quad \text{integrate by parts} \\
&= 0
\end{aligned} \tag{2.12}
$$

where the last line is required by Hamilton's principle. This automatically gives the condition $\partial^\mu F_{\mu\nu} = J^\nu$ for the inhomogeneous Maxwell equations in (2.9). The homogeneous ones (2.8), come from the identity

$$
\partial^\lambda F^{\mu\nu} + \partial^\mu F^{\lambda\mu} + \partial^\mu F^{\nu\lambda} = 0 \tag{2.13}
$$

known as the Jacobi identity.

Furthermore the transformation $A^\mu \to A^{\mu\prime} = A^\mu + \partial^\mu\chi = \left(\phi - \frac{\partial\chi}{\partial t}, A + \nabla\chi\right)$ leaves the action invariant, since

$$
\begin{aligned}
F'^{\;\mu\nu} &= \partial^\mu A'^{\;\nu} - \partial^\nu A'^{\;\mu} = \partial^\mu\left(A^\nu + \partial^\nu\chi\right) - \partial^\nu\left(A^\mu + \partial^\mu\chi\right) \\
&= F^{\mu\nu}
\end{aligned} \tag{2.14a}
$$

$$
\begin{aligned}
J^\mu A'_\mu &= J^\mu A_\mu + J^\mu \partial_\mu \chi \\
&= J^\mu A_\mu
\end{aligned} \tag{2.14b}
$$

where the final equality in (2.14b) follows because $\int d^4x J^\mu \partial_\mu\chi = -\int d^4x \left(\partial_\mu J^\mu\right)\chi$, and $\partial_\mu J^\mu = 0$ (since $\partial_\mu\partial_\nu F^{\mu\nu} = 0$ with the antisymmetric $F^{\mu\nu}$).

Therefore, the entire action is invariant under the transformation $A^\mu \to A^\mu + \partial^\mu\chi$. We will see that this seemingly coincidental property proves fundamental to the development of modern particle physics.

2.2 The Schrödinger Equation

In the 1920s, Austrian physicist Erwin Schrödinger developed an equation that describes the evolution of a quantum particle over time [19]. In his formulation, the wave-function ϕ describes the physical state of the system with the equation itself governing the evolution of this state over time.

This equation is now known as the Schrödinger equation of quantum mechanics:

$$H\phi = i\hbar\frac{\partial\phi}{\partial t} \tag{2.15}$$

In the non-interacting, non-relativistic scenario the Hamiltonian is $H = \frac{p^2}{2m} = -\frac{\hbar}{2m}\nabla^2$, where m denotes the mass of the quantum particle that we are considering. In this case,

$$-\frac{\hbar}{2m}\nabla^2\phi = i\hbar\frac{\partial\phi}{\partial t} \tag{2.16}$$

where ϕ is identified as the trivial (scalar) field that lives in the trivial spin-0 representation of the Lorentz group $SO(1,3)$.

Unfortunately, we cannot simply elevate (2.16) to a relativistic theory as it inherently treats space and time on unequal footings. Furthermore, quantum mechanics treats space as an operator $\hat{x}\phi = x\phi$ and time a parameter [20].

To construct a relativistic theory, one either needs to promote time to an operator or demote space to a parameter and quantise in a completely new manner. It turns out that the former is very complicated [6]. We will therefore follow the latter. This leads us to a situation where both x and p are scalars and thus commuting objects. In this case, the canonical commutation relation used to quantise the fields

$$[\hat{x}, \hat{p}] = i\hbar \tag{2.17}$$

is no longer valid [20].

We will follow the second approach. Instead of promoting parameters to Hermition operators that act on the states in the Hilbert space representing particles, we interpret the particles themselves as the Hermition operators that act on the vacuum. These Hermition operators $\hat{\phi}(x^\mu)$ shall be parameterised by the space-time coordinates x^μ such that $\hat{\phi}(x^\mu)\,|0\rangle$. For conciseness, we shall remove the explicit $\hat{}$ in our notation.

One can interpret the $\phi(x^\mu)$ field as an excitation of the vacuum creating a state in the Hilbert space representing a particle. This concept shall become clearer when we construct quantum fields $\phi(x^\mu)$ out of a linear combination of raising and lowering operators in the canonical way.

We shall now discuss the different types of quantum fields that will be present in our theory before moving to quantise them into something that resembles particles.

2.3 Spin-0 Fields

The Lorentz scalar fields with spin-0 transform under the trivial Lorentz representation, such that the Lorentz transformation $S = \mathbb{I}$ and $\phi' = S\phi = \phi$.

This is the simplest case and it is the field that appears in the relativistic Schrödinger equation. The obvious way to construct such a theory is to use the relativistic Hamiltonian whose eigenvalue is the energy of the relativistic particle of the system.

The Hamiltonian for a free particle (of mass m and momentum p) is thus

$$H \equiv \sqrt{m^2c^4 + p^2c^2} \tag{2.18}$$

In the non-relativistic limit, we return to the canonical Hamiltonian

$$\begin{aligned} H &= mc^2\sqrt{c^2 + \left(\frac{p}{m}\right)^2} \qquad \text{for } p \ll mc \\ &= mc^2\left(c^2 + \frac{1}{2}\left(\frac{p}{m}\right)^2\right) \\ &= mc^2 + \frac{p^2}{2m} \end{aligned} \tag{2.19}$$

where mc^2 is the constant rest energy of the system.

To put the time and space coordinates on equal settings, we *square* the operator on both sides of the Schrödinger equation,

$$-\hbar^2 \frac{\partial^2 \phi}{\partial t^2} = (m^2 c^4 - \hbar^2 c^2 \nabla^2)\phi \tag{2.20}$$

In natural units, with $c = 1$ we find the Klein-Gordon equation

$$\left(\partial^2 - m^2\right) \phi = 0 \tag{2.21}$$

which is simply the operator version of the relativistic relation $E^2 = p^2 c^2 + m^2 c^4$.

We have reached a classical relativistic analogue to the Schrödinger equation for spinless non-interacting fields. This is currently just a wave equation and we still need to quantise the scalar fields ϕ.

2.3.1 Negative Energy States

Note that *squaring* the operator H^2 implies that there are (equally valid) negative energy solutions

$$E = -\sqrt{m^2 c^4 + p^2 c^2} \tag{2.22}$$

This poses the obvious issue that the vacuum would no longer be the lowest energy state. There would therefore be infinite cascades of particles down these states, giving off infinite quantities of energy.

Dirac's solution was to consider that the infinite sea of negative energy states is already entirely filled by Pauli's exclusion principle [21]. This made them them inconsequential since we only ever observe energy differences.

Whenever electrons, described by left-handed spinor field ψ_L are elevated out of these negative energy states, the holes that are left are seen as antiparticles described by the right-handed spinor field ψ_R. This interpretation turned out not to be correct as ψ_L and ψ_R are not in general antiparticles of each other. Additionally, bosons do not exhibit Pauli's exclusion principle; thus these negative energy states would never be filled.

By moving to quantum field theory this issue will naturally be resolved. Fields ϕ, ψ with negative energy states $\phi^\dagger, \psi^\dagger$, are instead elevated to operators and will no longer carry intrinsic energy. Then, the fields will operate on the ground state $|0\rangle$ as $\phi\,|0\rangle$ and $\psi\,|0\rangle$ creating (different) particles and the operators $\overline{\phi}, \overline{\psi}$ will create their "antiparticles".

2.3.2 The Klein-Gordon Lagrangian for Scalar Spin-0 Fields

We wish to embed the Klein-Gordon equation of (real) scalar fields (2.21) into a Lagrangian. It is trivial to see that the Lagrangian that gives rise to this equation of motion is

$$\mathcal{L}_{\mathrm{KG}}(x) = -\frac{1}{2}\partial_\phi\phi\partial^\mu\phi - \frac{1}{2}m^2\phi^\dagger\phi \tag{2.23}$$

This Lagrangian is explicitly shown to be Lorentz invariant in Appendix B.1 for the Lorentz scalar field ϕ. Notice that the (squared) mass is the coefficient of the quadratic term in the fields. This is a general property of field theories and stems ultimately from applying the classical Hamiltonian into the quantum theory described by the field ϕ (2.20). The particles that arise upon quantisation of the field ϕ will thus have mass m.

Note that there is a symmetry

$$\phi \longrightarrow e^{i\alpha}\phi \approx (1 + i\alpha)\phi \tag{2.24}$$

that leaves the Lagrangian invariant. The associated conserved current is

$$j^\mu_{\mathrm{KG}} = i(\phi\partial^\mu\phi^\dagger - \phi^\dagger\partial^\mu\phi) \tag{2.25}$$

2.4 Spin-$\frac{1}{2}$ Fields

In 1922, Otto Stern and Walther Gerlach found that a beam of otherwise indistinguishable electrons were separated into two paths when passed through a magnetic field [22]. An additional non-physical and purely quantum property, known as spin, that separated out these two electron beams was required.

The simplest way to describe electrons, and indeed, all fermions, with two equivalent states was to assign them an additional degree of freedom and define them as spinor fields in the spin-$\frac{1}{2}$ representation of $SU(2)$.

To accomplish this, we will look for the analogue of the Klein-Gordon equation for spin-$\frac{1}{2}$ fields. We wish to embed this information into a spinor field ψ by

$$\psi = \begin{pmatrix} \psi_1 \\ \psi_2 \end{pmatrix} \tag{2.26}$$

with ψ_1 spin 'up' and ψ_2 spin 'down' fields in the spin-$\frac{1}{2}$ representation.

Historically, Dirac attempted to find a 2×2 operator that acted on these spinors, defining the equation of motion for these fermions. To ensure that his theory was relativistic, he wanted to find an operator that would naturally yield the Klein-Gordon equation. His approach was to write the equation of motion

$$\left(i\partial\!\!\!/ - m\right)\psi = 0 \tag{2.27}$$

known as the Dirac equation, with

$$\begin{aligned} \partial\!\!\!/ &= \gamma^\mu \partial_\mu \\ &= \gamma^0\partial_0 + \gamma^1\partial_1 + \gamma^0\partial_2\gamma^0\partial_2 + \gamma^3\partial_3 \end{aligned} \tag{2.28}$$

where γs were 2×2 matrices.

He noted that operating with $\partial\!\!\!/$ gives

$$\begin{aligned} -im\partial\!\!\!/\psi &= \partial\!\!\!/\partial\!\!\!/\psi \\ (-im)^2\psi &= \gamma^\mu\partial_\mu\gamma^\nu\partial_\nu\psi \\ \implies -m^2\psi &= \gamma^\mu\gamma^\nu\partial_\mu\partial_\nu\psi \\ \implies 0 &= (\gamma^\mu\gamma^\nu\partial_\mu\partial_\nu + m^2)\psi \end{aligned} \tag{2.29}$$

gave the Klein-Gordon equation if $\gamma^\mu\gamma^\nu = -\eta^{\mu\nu}\mathbb{I}$ or equivalently

$$\{\gamma^\mu, \gamma^\nu\} = 2\eta^{\mu\nu}\mathbb{I} \tag{2.30}$$

where $\{\cdot, \cdot\}$ denotes the anti-commutator defined as

$$\{\gamma^\mu, \gamma^\nu\} \equiv \frac{1}{2}\left(\gamma^\mu\gamma^\nu + \gamma^\nu\gamma^\mu\right) \tag{2.31}$$

Unfortunately, a solution for (2.30) does not exist for 2×2 matrices. In fact, the lowest dimension solution is 4×4 and this requires us to describe the two spin states with a four-component field

$$\begin{pmatrix} \psi_1 \\ \psi_2 \\ \psi_3 \\ \psi_4 \end{pmatrix} \equiv \begin{pmatrix} \psi_L \\ \psi_R \end{pmatrix} \tag{2.32}$$

with ψ_L, ψ_R two-component left and right handed Weyl spinors respectively. This combination of left and right handed Weyl spinors lives in the $(\frac{1}{2}, 0) \otimes (0, \frac{1}{2})$ representation of $SO(1,3)$ and is known as the two-component Dirac spinor, which by construction satisfies the Dirac equation.

Although Dirac was originally concerned with the need for a four-component field to describe two spin states, we shall later observe when we quantise that Dirac had inadvertently discovered antiparticles.

One solution of (2.30) is

$$\gamma^0 = \begin{pmatrix} 0 & \sigma^0 \\ \sigma^0 & 0 \end{pmatrix} \quad , \quad \gamma^i = \begin{pmatrix} 0 & \sigma^i \\ -\sigma^i & 0 \end{pmatrix} \tag{2.33}$$

where σ^i are the Pauli spin-matrices and $(\gamma^0)^2 = +1$, $(\gamma^i)^2 = -1$. This solution is known as the Weyl (or chiral) representation.

Writing out the Dirac equation explicitly in four-component form, one has

$$\begin{pmatrix} 0 & 0 & \partial_0 - \partial_3 & -\partial_1 + i\partial_2 \\ 0 & 0 & -\partial_1 - i\partial_2 & \partial_0 + \partial_3 \\ \partial_0 + \partial_3 & \partial_1 - i\partial_2 & 0 & 0 \\ \partial_1 + i\partial_2 & \partial_0 - \partial_3 & 0 & 0 \end{pmatrix} \begin{pmatrix} \psi_1 \\ \psi_2 \\ \psi_3 \\ \psi_4 \end{pmatrix} = -im \begin{pmatrix} \psi_1 \\ \psi_2 \\ \psi_3 \\ \psi_4 \end{pmatrix} \tag{2.34}$$

or equivalently

$$i\overline{\sigma}^\mu \partial_\mu \psi_R = +m\psi_L \tag{2.35a}$$

$$i\sigma^\mu \partial_\mu \psi_L = +m\psi_R \tag{2.35b}$$

where $\sigma^\mu \equiv \left(\sigma^0 = \mathbb{I}, \sigma^1, \sigma^2, \sigma^3\right)$ and $\overline{\sigma}^\mu \equiv \left(\sigma^0 = \mathbb{I}, -\sigma^1, -\sigma^2, -\sigma^3\right)$.

For later convenience, we define the fifth γ^5 matrix as

$$\gamma^5 = i\gamma^0\gamma^1\gamma^2\gamma^3 = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix} \tag{2.36}$$

This is useful to construct the projection operator of the Dirac field onto its left-handed and right-handed components with

$$\psi_L = \frac{1}{2}\left(1 - \gamma^5\right)\psi \quad , \quad \psi_R = \frac{1}{2}\left(1 + \gamma^5\right)\psi \tag{2.37}$$

As such, theories that have γ^5 terms are chiral and will affect the left and right-handed fields differently. The handedness of the particle, determined by the Lorentz representation under which it transforms, is known as its "chirality".

Helicity

"Helicity" (h) is defined as the sign of the projection of the spin onto its momentum vector with $h \equiv \sigma \cdot p$. Formally, the projection operators that project onto the positive $(+)$ and negative (-) helicity states are defined by

$$\psi_- = \frac{1}{2}\left(1 - \frac{\sigma \cdot p}{|p|}\right)\psi \quad , \quad \psi_+ = \frac{1}{2}\left(1 + \frac{\sigma \cdot p}{|p|}\right)\psi \tag{2.38}$$

where σ are the three Pauli spin matrices and p the three-momentum of the particle described by ψ. In four-vector notation, we take the definition

$$\frac{\sigma \cdot p}{|p|} \equiv \begin{pmatrix} \frac{\sigma \cdot p}{|p|} & 0 \\ 0 & \frac{\sigma \cdot p}{|p|} \end{pmatrix} \tag{2.39}$$

where σ are themselves 2×2 matrices.

Notice that for massive particles it is always possible to perform a Lorentz transformation to a reference frame where its helicity is reversed. In contrast, massless particles have an intrinsic helicity, with the negative (positive) helicity state coinciding with the definition of left (right)-handed chirality, since $\frac{\sigma \cdot p}{|p|} \to \gamma^5$ for $m = 0$ [16].

Therefore, one can think of helicity as an apparent chirality, where the two definitions are equivalent in the massless limit. Since spin can be determined experimentally, measurements of helicity thus provide a mechanism to deduce the chirality of particles.

Indeed, neutrinos, which are considered massless in the Standard Model, have only been measured with negative helicity and therefore left-handed chirality [23].

2.4.1 The Dirac Lagrangian for Spin-$\frac{1}{2}$ Particles

Notice that the Dirac spinors transform non-trivially under the Lorentz group since

$$\psi' = \exp\begin{pmatrix} \frac{1}{2}\phi_k\sigma^k + \frac{1}{2}i\theta_k\sigma^k & 0 \\ 0 & -\frac{1}{2}\phi_k\sigma^k + \frac{i}{2}\theta_k\sigma^k \end{pmatrix}\psi \tag{2.40}$$

where σ_k are the canonical Pauli spin matrices and θ, ϕ are (constant) arbitrary phases. The transformation rule of the spinor fields ψ and the Lorentz invariance of the Dirac Lagrangian are shown in Appendix B.2.

Using these spinor fields, one constructs the Dirac Lagrangian:

$$\begin{aligned}\mathcal{L}_{\text{Dirac}} &= i\psi_L^\dagger \overline{\sigma}^\mu \partial_\mu \psi_L + i\psi_R^\dagger \sigma^\mu \partial_\mu \psi_R - m(\psi_L^\dagger \psi_R + \psi_R^\dagger \psi_L) \\ &= \overline{\psi}\left(i\gamma^\mu \partial_\mu - m\right)\psi \end{aligned} \tag{2.41}$$

with $\psi \equiv \begin{pmatrix} \psi_L \\ \psi_R \end{pmatrix}$, $\psi_L^\dagger = (\psi_1^\star, \psi_2^\star)$ where $\dagger$ represents the standard Hermition conjugate and $\overline{\psi} \equiv \psi^\dagger \gamma^0$ is known as the "Dirac adjoint".

Notice that the mass m, which will represent the mass of the particles after quantisation, is again the coefficient of the quadratic term in the field. However, in contrast to the Klein-Gordon Lagrangian (2.23), there is only a linear dependence on m. This is consistent with the Lagrangian having a mass dimension of four.

2.5 Spin-1 Fields

The vector fields of the Lorentz group live in the $(\frac{1}{2}, \frac{1}{2})$ spin-1 vectorial representation. These fields are typically indicated with an additional index μ, that runs over the four space-time directions.

We shall see the vector field A_μ arise in an extended Dirac equation as

$$\mathcal{L}_{\text{Dirac+vector}} = \overline{\psi}\left(\gamma^\mu\left(i\partial_\mu - eA_\mu\right) - m\right)\psi \tag{2.42}$$

where e is a constant. To ensure that this Lagrangian is Lorentz invariant, the vector fields A_μ are required to transform under the $(\frac{1}{2}, \frac{1}{2})$ vectorial representation as $A_\mu \to \Lambda^\mu{}_\nu A^\nu$. The transformation rule and the invariance of this Lagrangian are described in Appendix B.3.

In particular, we shall see that the Standard Model (SM) bosons are described by vector fields to give the correct experimentally observed couplings to the fermions. However, the four independent degrees of freedom of a vector field A_μ (where $\mu = 0, 1, 2, 3$) are slightly too general to describe the observed degrees of freedom in bosons.

In fact, there is an infinite class of equivalent ways to represent the bosons with the vector fields, all related by symmetries. From this infinite set, a choice must be made to remove the redundancy. This is done by imposing ad-hoc constraints, known as "gauge fixing" [7]. To match the expected degrees of freedom of the spin-1 vector fields in quantum mechanics, the procedure removes one degree of freedom for massive fields and two for massless ones.

Notice that since the physics described by the theory is symmetric, the final result must be independent of the choice of gauge even though the intermediate calculations may not be.

2.6 Path Integral Quantisation

In 1803, Thomas Young performed his famous double slit experiment where photons, which were thought to be quantised discrete particles, were fired at two slits and appeared to travel through both simultaneously [24]. A wave-like interference pattern was observed on a screen some distance away, with peaks and troughs at points of constructive and destructive superposition respectively.

It was not until 1927, that Niels Bohr, using ideas from Heisenberg's uncertainty principle, proposed the wave-particle duality of all particles [25]. Bohr proposed the notion of complementarity, that photons obey wave-like behaviour described by a wave-function until they are measured as discrete objects.

In the Copenhagen Interpretation of quantum mechanics, quantum particles travel along all paths simultaneously each with some probabilistic weighting; the most probable trajectory often coincides with the classical non-quantum path. Any measurement, collapses the wave-function to one particular point forcing the particle to take only one of the infinite possible paths [20].

In 1948 Feynman completed the development of path integral methods, with the path integral describing the quantum paths that particles take [26]. This procedure is known as the path integral quantisation and naturally describes the wave-particle duality.

We start with the canonical Schrödinger equation, expressed in terms of the Hamiltonian

$$H\phi = i\frac{\partial \phi}{\partial t} \tag{2.43}$$

Recall, that this is not manifestly Lorentz covariant as it treats both space and time differently. We reiterate that this will automatically be addressed when we move to quantum fields, since ϕ will transform as a Lorentz scalar field.

Taking H to be independent of time, the solution of this equation can be written as

$$\begin{aligned}\phi(q,t) &= \langle q|\, e^{-iH(t-t_0)}\, |\phi(t=t_0)\rangle \\ &\equiv |q,t\rangle \end{aligned} \tag{2.44}$$

which describes the state of a single particle at position q and time t. The quantity $U(t_f,t_i) \equiv e^{-iH(t_f-t_i)}$ is a unitary operator known as the time evolution operator.

We proceed by defining the correlation function between two states as

$$\langle q_f,t_f|q_i,t_i\rangle = \langle q_f|\, e^{iH(t_f-t_i)}\, |q_i\rangle \tag{2.45}$$

where we have used the evolution operator to extract the explicit time dependence. One of the postulates of quantum mechanics states that the modulus squared of this quantity defines the probability that a particle will evolve from state $\phi(q_i,t_i)$ to the state $\phi(q_f,t_f)$ [20].

The time interval $T \equiv t_f - t_i$ can now be split into $N+1$ components, each of length $\delta t = T/N+1$ such that

$$\langle q_f,t_f|q_i,t_i\rangle = \int_{-\infty}^{\infty}\prod_{j=1}^{N} dQ_j\, \langle q_f|\, e^{-iH\delta T}\, |Q_N\rangle\, \langle Q_N|\, e^{-iH\delta t}\, |Q_{N-1}\rangle \cdots \langle Q_1|\, e^{-iH\delta t}\, |q_i\rangle \tag{2.46}$$

where we have inserted the complete set of states $\int dQ'\, |Q'\rangle\, \langle Q'| = \mathbb{I}$.

Typically $H = \frac{p^2}{2m} + V(q)$ and it follows that

$$\begin{aligned}\langle Q_{j+1}|\, e^{-iH\delta t}\, |Q_j\rangle &= \langle Q_{j+1}|\, e^{-i\left(\frac{p^2}{2m}+V(q)\right)\delta t}\, |Q_j\rangle \\ &= \int dP'\, \langle Q_{j+1}|\, e^{-i\delta t\frac{p^2}{2m}}\, \left|P'\right\rangle \left\langle P'\right|\, e^{-i\delta t V(q)}\, |Q_j\rangle \qquad \text{by completeness} \\ &= \int dP' e^{-i\delta t\frac{P'^{\,2}}{2m}}\, e^{-i\delta t V(Q_j)}\, \left\langle Q_{j+1}|P'\right\rangle \left\langle P'|Q_j\right\rangle \\ &= \int dP' e^{-i\delta t\frac{P'^{\,2}}{2m}}\, e^{-i\delta t V(Q_j)}\, \frac{e^{iP'Q_{j+1}}}{\sqrt{2\pi}}\frac{e^{-iP'Q_j}}{\sqrt{2\pi}} \end{aligned} \tag{2.47}$$

where the last line follows from the identity that $\langle p|q'\rangle = \frac{1}{\sqrt{2\pi}} e^{-ipq'}$.

Note that the Baker-Cambell-Hausdorff formula states that for general operators A and B

$$e^A e^B = e^{A+B-\frac{1}{2}[A,B]+\cdots} \tag{2.48}$$

where $\cdots$ in the exponent indicates higher order commutators [8]. However, all commutator terms here are $\mathcal{O}(\delta t)^2$ and can therefore be neglected in the limit $\delta t \to 0$.

Then,

$$\langle Q_{j+1}| e^{-iH\delta t} |Q_j\rangle \quad = \quad \int \frac{dP'}{2\pi} e^{-iH(P',Q_i)\delta t} e^{iP'\delta t \frac{(Q_{j+1}-Q_j)}{\delta t}} \tag{2.49}$$

which in the limit that $\delta t \to 0$ and $N \to \infty$ gives $(Q_{j+1} - Q_j)/\delta t \to \dot{Q}$.

Substituting (2.49) into the correlation function (2.46) yields the "functional" integral

$$\langle q_f, t_f | q_i, t_i\rangle = \int_{-\infty}^{\infty} d[q]d[p]e^{iS_H} \tag{2.50}$$

where we have defined

$$d[p] \equiv \prod_{j=1}^{N} \frac{dp_j}{2\pi} \qquad d[q] \equiv \prod_{j=1}^{N} dq_j \tag{2.51}$$

and

$$S_H(p,q) \equiv \int_{t_i}^{t_f} dt\, [p(t)\dot{q}(t) - H\left(p(t), q(t)\right)] \tag{2.52}$$

is the Hamiltonian action.

Assuming $H = \frac{p}{2m} - V(q)$, one can explicitly do the Gaussian integral over the momenta p_j, which gives

$$\langle q_2, t_2 | q_1, t_1\rangle = \int D[q]e^{iS} \tag{2.53}$$

where $D[q] \equiv \prod_{i=1}^{n} \left(\frac{m}{2\pi i}\right) dq_i$ is the infinite product of the coordinates q_i. Finally, the action can be written as

$$S = \int d^4x \mathcal{L} = \int_{t_i}^{t_f} dt \left[\frac{m}{2}\dot{q}^2(t) - V\left(q(t)\right)\right] \tag{2.54}$$

We can interpret the functional integral of (2.53) as the sum over all possible paths $\int D[q]$ between (q_1, t_1) and (q_2, t_2), with each path given the statistical weight e^{iS}.

Varying this by a small quantity $q' = q_0 + \varepsilon\delta_0$, we have the statistical weight $e^{iS[q_0+\varepsilon\delta q_0]} = e^{iS[q_0]+i\varepsilon\delta q_0 \frac{\delta S[q_0]}{\delta q}}$ where $\frac{\delta S}{\delta q}$ is the Euler-Lagrange variation with respect to q such that

$$\frac{\delta S}{\delta q} = \frac{d}{dt}\frac{\partial S}{\partial \dot{q}} - \frac{\partial S}{\partial q} \tag{2.55}$$

To simplify, we perform a Wick rotation [8], where we take $dt \to i\delta t$, $S = \int dt\mathcal{L} \to i\int dt\mathcal{L} = iS$ and therefore $e^{iS} \to e^{-S}$. Looking at a small variation, the path $q' = q_0 + \varepsilon\delta q_0$ gets the weight

$$e^{S[q_0]}e^{-\varepsilon\delta q_0 \frac{\delta[q_0]}{\delta q}} \tag{2.56}$$

It is now clear that when the variation $\frac{\delta S}{\delta q}$ is large, the corresponding weight is will be small.

Therefore, out of the infinite possible paths, those with the largest variations of the action are the least probable and those that minimise the action are the most probable paths.

We can easily generalise this notion to the path integral of fields that, recall, is manifestly Lorentz invariant, by taking the Lagrangian

$$\mathcal{L} = \frac{1}{2}\partial^\mu\phi\partial_\mu\phi - V(\phi) \tag{2.57}$$

and making the following identifications

$$q \longrightarrow \phi \qquad p \longrightarrow \Pi = \dot{\phi} \tag{2.58}$$

Note that since the action is in the exponential, it is required to be dimensionless. This imposes an additional constraint on the permissible terms in the Lagrangian. This is consistent with the Lagrangian density being proportional to energy and thus having units of mass in one dimension. In four dimensions, the Lagrangian density has a mass dimension $+4$ and the integration measure d^4x has mass dimension -4 as required.

In general, to find the mass dimension of a field, one can simply look at its kinetic term; for scalar fields the kinetic term is of the form $\partial\phi\partial\phi$ and for spinor fields $\psi\not{\partial}\psi$. Thus, the mass dimension of the scalar fields is $[\phi] = +1$ and $[\psi] = +3/2$ for spinor fields.

The correlation function is then

$$\langle \phi_f(t_f)|\phi_i(t_i)\rangle = \int d[\phi] e^{iS[\phi]} \tag{2.59}$$

where $\int d[\phi] = \int \prod_n \frac{1}{2\pi i} d\phi_n$ is now the sum over all field configurations and the action $S[\phi]$ is

$$S = \int_{t_i}^{t_f} dt \int d^3x \left[\frac{1}{2}\partial^\mu \phi \partial_\mu \phi - V(\phi)\right] \tag{2.60}$$

By specifying the action, we can now compute the correlation function of these fields. One can also calculate arbitrary expectation values, which define the probabilities of interactions by a similar prescription.

2.6.1 Expectation Values with Path Integral Quantisation

In quantum mechanics, every physical observable is represented by a Hermition operator that acts on the states that define the system. Performing measurements of this observable is equivalent to acting with this operator, with the possible measurements the eigenvalues of the (matrix) operator.

The "expectation value" of an operator Q on the state ϕ is the predicted mean value of this measurement and is defined as

$$\langle Q\rangle_\phi = \langle\phi| Q |\phi\rangle \tag{2.61}$$

In general, we wish to calculate quantities like $\langle q_2, t_2| Q(t') |q_1, t_1\rangle$ and $\langle q_2, t_2| P(t') |q_1, t_1\rangle$. Following analogous arguments for the correlation function above, we find

$$\langle q_2, t_2| Q(t') |q_1, t_1\rangle = \int D[q] Q(t') e^{iS} \tag{2.62}$$

We now define the functional derivative for some function $f(x)$ by

$$\frac{\delta}{\delta f(y)} f(x) \equiv \delta(x-y) \tag{2.63}$$

Next, we add two additional external source functions $f(t)$ and $h(t)$ to the Lagrangian, such that

$$\mathcal{L} \longrightarrow \mathcal{L} + f(t)Q(t) + h(t)P(t) \tag{2.64}$$

Thus,

$$\langle q_2, t_2 | q_1, t_1 \rangle_{f,h} = \int D[q] e^{\int dt (\mathcal{L} + fQ + hP)} \tag{2.65}$$

The expectation value of a generic operator can now be found by repeated operations of the functional derivative. For example,

$$\begin{aligned} \langle q_2, t_2 | \, Q(t') \, | q_1, t_1 \rangle &= \frac{1}{i} \frac{\delta}{\delta f(t')} \, \langle q_2, t_2 | q_1, t_1 \rangle_{f,h} \Big|_{f=0,h=0} \\ &= \int D[q] Q(t') \, e^{iS + i \int dt (fQ + hP)} \Big|_{f,h=0} \\ &= \int D[q] Q(t') e^{iS} \end{aligned} \tag{2.66}$$

and

$$\begin{aligned} \langle q_2, t_2 | \, P(t') \, | q_1, t_1 \rangle &= \frac{1}{i} \frac{\delta}{\delta h(t')} \, \langle q_2, t_2 | q_1, t_1 \rangle |_{f,h=0} \\ &= \int D[q] P(t') e^{iS} \end{aligned} \tag{2.67}$$

We can generalise this method to find the expectation of any higher order operators by simply taking successive functional derivatives [27].

2.7 First Quantisation

We have so far made no attempt to quantise the fields in such a way as to convert the classical Lagrangians into something that resembles particles. We shall now describe the procedure of first quantisation, where the quantum states correspond to single particles.

In 1927, Werner Heisenberg heuristically proposed that there was an inherent uncertainty in the simultaneous measuring of related quantities, such as the position and the momentum of a particle [28]. This uncertainty was empirically postulated to describe a fundamental limit of the accuracy that measurements can be made. It therefore provides a natural measure of quantisation in quantum mechanics.

Later that year, Earle Hesse Kennard [29] provided the exact form for the uncertainty equation,

$$\sigma_x \sigma_p \geq i\hbar \tag{2.68}$$

in terms of σ_x, and σ_p, the standard deviation of position and momentum respectively, with $\hbar$ related to the fundamental quantity known as Planck's constant $h = 2\pi\hbar$.

Measurements of momentum and position are found by acting on the state with momentum and position matrix operators, whose eigenvalues represent the measured quantity. The basis states that describe single particles are then linear combinations of eigenstates of the position and momentum.

In this way, the uncertainty relation can be interpreted as the fundamental commutator

$$[\hat{x}, \hat{p}] \equiv \hat{x}\hat{p} - \hat{p}\hat{x} = i\hbar \tag{2.69}$$

where we have assumed the equality in the relation to signify the lower bound.

The non-vanishing commutator can be understood as the difference between operating on a state with $\hat{p}$ then $\hat{x}$ and vice-versa. The difference between the order of operating with $\hat{x}$ and $\hat{p}$ is quantified by Planck's constant and is equivalent to the uncertainty of the measurements in (2.68). Indeed, the probabilistic evolution of such a state with respect to time will be dictated by the time dependent Schrödinger equation [20].

For example, measurement of the momentum of the wave-function $\hat{p}\,|\phi\rangle$ collapses the wave-function $|\phi\rangle$ to an eigenstate $|p\rangle$ with eigenvalue p; immediate measurement of the position $\hat{x}\,|p\rangle$ yields some uncertainty $i\hbar/2$ since $|p\rangle$ is not in general an eigenstate of $\hat{x}$.

In the next section, we will discuss quantum field theory where $\hat{x}$ and $\hat{p}$ have been reduced to allow Lorentz invariance. The scalar fields ϕ (and also the spinor fields ψ) are themselves elevated to operators; the eigenstates will then become eigenvectors of the Hamiltonian. With this formulation, we shall use the canonical uncertainty relation (2.69) to quantise quantum field theory with analogous commutation relations. For fields, the commutation relations that are imposed in quantisation are now heuristically $[\phi, \Pi] > 0$ for Π the field's conjugate momenta.

2.8 Second Quantisation of Scalar Fields

Second quantisation, also known as canonical quantisation, is the procedure to quantise fields in such a way as to describe multi-particle states.

We start with the Klein-Gordon Lagrangian for real scalar fields:

$$\mathcal{L}_{\mathrm{KG}} = -\frac{1}{2}\partial^{\mu}\phi\partial_{\mu}\phi - \frac{1}{2}m^2\phi^2 + \Lambda \tag{2.70}$$

where we have introduced an arbitrary constant Λ that does not affect the Euler-Lagrange equations of motion.

We now define the Hamiltonian density of this system $\mathcal{H}$ in terms of $\mathcal{L}$

$$\mathcal{H} \equiv \Pi\dot{\phi} - \mathcal{L} \tag{2.71}$$

where

$$\Pi \equiv \frac{\partial \mathcal{L}}{\partial \dot{\phi}} \tag{2.72}$$

is the conjugate momentum of the field. This definition is again necessary for consistency with Newton's laws where the Hamiltonian $H \equiv \int d^3x\mathcal{H} = T + V$ gives the total energy of the system.

For the Klein-Gordon Lagrangian, the field momentum is

$$\Pi = \frac{\partial \mathcal{L}}{\partial \dot{\phi}(x)} = \dot{\phi} \tag{2.73}$$

The Hamiltonian density is then

$$\mathcal{H} = \frac{1}{2}\dot{\phi} + \frac{1}{2}\left(\nabla\phi\right)^2 + \frac{1}{2}m^2\phi^2 - \Lambda \tag{2.74}$$

This has a plane wave solution

$$\phi(\bar{x},t) = \int \frac{d^3k}{(2\pi)^3}\frac{1}{\sqrt{2\omega}}\left(a(k)e^{i\bar{k}\cdot\bar{x}-i\omega t} + b(k)e^{i\bar{k}\cdot\bar{x}+i\omega t}\right) \tag{2.75}$$

with $\omega = +\sqrt{k^2+m^2}$, where k is the standard wave-vector and ω the angular frequency, equivalent to the energy $E(k)$ of the system up to a factor $\hbar$.

The $(2\pi)^3\sqrt{2\omega}$ is an additional function to ensure Lorentz invariance; a and b are currently arbitrary functions of k. One can write the four dimensional measure as

$$d\tilde{k} = \frac{d^3k}{(2\pi)^3\sqrt{2\omega}} \tag{2.76}$$

whichis Lorentz invariant by construction [6].

To quantise, we now impose similar commutation relations to those in canonical quantum mechanics

$$\begin{aligned}
\left[\phi(\overline{x},t),\phi(\overline{x}',t')\right] &= 0 && (2.77a)\\
\left[\Pi(\overline{x},t),\Pi(\overline{x}',t')\right] &= 0 && (2.77b)\\
\left[\phi(\overline{x},t),\Pi(\overline{x'},t')\right] &= i\delta(\overline{x}-\overline{x}')\delta(t-t') && (2.77c)
\end{aligned}$$

where we have used natural units with $\hbar = 1$.

From the requirement that the scalar field ϕ is Hermition, giving necessarily real eigenvalues that can be physically measured, we have $\phi^\dagger = \phi$ and thus

$$\begin{aligned}
\phi^\dagger &= \int d\tilde{k}\left(a^\dagger(k)e^{-i\overline{k}\cdot\overline{x}+i\omega t} + b^\dagger(k)e^{i\overline{k}\cdot x+i\omega t}\right)\\
&= \int d\tilde{k}\left(a^\dagger(-k)e^{i\overline{k}\cdot\overline{x}+i\omega t} + b^\dagger(-k)e^{-i\overline{k}\cdot\overline{x}+i\omega t}\right)\\
&= \phi && (2.78)
\end{aligned}$$

This requires, $a^\dagger(-k) = b(k)$, such that

$$\phi(x) = \int d\tilde{k}\left(a(k)e^{ik\cdot x} + a^\dagger(k)e^{-ik\cdot x}\right) \tag{2.79}$$

where we have used the four-vector notation $k \cdot x = k^\mu x_\mu$ and made use of the integral over all d^3k to flip the negative sign in the second component.

The commutation relations of (2.77) now give the following:

$$\left[a(k), a(k')\right] = 0 \tag{2.80a}$$
$$\left[a^\dagger(k), a^\dagger(k')\right] = 0 \tag{2.80b}$$
$$\left[a(k), a^\dagger(k')\right] = (2\pi)^3\, 2\omega\delta^3\left(\overline{k} - \overline{k}'\right) \tag{2.80c}$$

One can now explicitly expand the fields in the Hamiltonian. Noting that $\Pi = \dot{\phi}$ and that $\int d^3x e^{ix \cdot y} = (2\pi)^3\delta(y)$, we have

$$H = \frac{1}{2}\int \tilde{dk}\omega\left(a^\dagger(k)a(k) + a(k)a^\dagger(k)\right) - V\Lambda \tag{2.81}$$

Using the commutation relations (2.80), this becomes

$$H = \int \tilde{dk}\omega a^\dagger(k)a(k) + \underbrace{\frac{1}{2}\delta^3(0)\int d^3k}_{E_0} - V\Lambda \tag{2.82}$$

where V is the volume over which we integrate. Note that the last two terms are infinite, the first E_0 is interpreted as the infinite zero point energy of the vacuum, and the second is the integral over infinite space.

We thus define
$$\Lambda = \frac{1}{2V}\delta^3(0)\int d^3k \tag{2.83}$$
which yields
$$H = \int \tilde{dk}\omega a^\dagger(k)a(k) \tag{2.84}$$

Note that the Hamiltonian in (2.81) looks very similar to a simple harmonic oscillator [20]. We can identify $a^\dagger$ and a as the quantum mechanical operators respectively that act on the vacuum, where

$$\left[H, a^\dagger\right] = \omega a^\dagger \quad , \quad [H, a] = -\omega a \tag{2.85}$$

In a single particle system with one $a^\dagger$, this implies that

$$H a^\dagger \left|E\right\rangle = (E+\omega)\, a^\dagger \left|E\right\rangle \quad , \quad H a \left|E\right\rangle = (E-\omega)\, a \left|E\right\rangle \tag{2.86}$$

Repeated application of $a^\dagger$ and a furnishes the ladder of excited states with energies, $\ldots, E-\omega, E, E+\omega, E+2\omega, \ldots$ with a ground state having energy $H\left|0\right\rangle = \frac{1}{2}\omega\left|0\right\rangle$. This exactly corresponds to the $SU(2)$ ladder operators (A.5) described in Appendix A.2. Notice that although there are negative frequency modes (corresponding to the term $e^{ik\cdot x}$ in the field expansion), the spectra of energies has a well-defined and positive minimum.

In terms of multi-particle systems, acting repeatedly with $a^\dagger(k)$ on the ground state, gives particles described by ϕ with momentum k and energy ω with normalised states

$$\left|k\right\rangle = \sqrt{2\omega}\, a^\dagger(k)\left|0\right\rangle \tag{2.87}$$

The n-particle state can be constructed as $\left|k_1, k_2, \ldots, k_n\right\rangle = a^\dagger_{k_1} a^\dagger_{k_2} \cdots a^\dagger_{k_n}\left|0\right\rangle$, where the operation $H = E_N \propto a^\dagger a \equiv N$ defines the number operator that counts the number of particles in the system (or equivalently the energy, up to some normalisation factor [20]).

The entire spectrum of states can be generated in this way with repeated application of creation and annihilation operators on the ground state. The probability for transitions between any two states is given by the expectation value $\left\langle k_f | k_i\right\rangle > 0$.

Notice that simply imposing the commutation relations on fields and their conjugate momenta yields the whole spectrum of non-interacting (multi-)particle states. For real scalar fields, we see that there is only one set of creation and annihilation operators and the scalar bosons they describe are necessarily neutral. On the other hand, the complex scalar fields admit two different sets of creation and annihilation operators (related by complex conjugation). This allows the field to describe charged scalar bosons.

2.8.1 The Cosmological Constant

We have just arbitrarily cancelled a seemingly divergent term by another infinite quantity in (2.83). A detailed analysis of this cancellation requires the theory of renormalisation, which cares only for measurable energy differences. Traditionally, we regard these infinite terms as coming from a more fundamental, higher energy theory, which when viewed from our low energy prospective yields an effective theory plagued with these mathematical quirks.

However, since gravity is encapsulated in a theory that is intricately related to space-time, the sum over all zero-point energies will necessarily contribute to the stress-energy tensor $T_{\mu\nu}$ that appears in Einstein's equations. This describes the energy of the particular space-time and for flat-space this is (1.38).

Einstein's General Relativity [15,30] would thus require the cosmological constant $\Lambda = E_0/V$, such that the equations of motion are

$$R_{\mu\nu} - \frac{1}{2} g_{\mu\nu} = -8\pi G T_{\mu\nu} + \Lambda g_{\mu\nu} \tag{2.88}$$

Current measurements find $\Lambda \approxeq 10^{-35} s^{-2}$ [31,32]. This value, equivalent to 10^{-47} GeV^4 is much smaller than all other physical scales and indicates a seemingly unnatural cancellation of zero-point contributions to the cosmological constant.

Using dimensional analysis, the mass scale is determined by Newton's gravitational constant $G = 1/M_{pl}^2$, where $M_{pl} \approxeq 10^{19}$ GeV is the Planck mass.

In the absence of any other scale, one would expect a cosmological constant for the SM of order M_{pl}^4, since the zero point energy of the vacuum in $T^{\mu\nu} \approxeq 1/M_{pl}^2$ also has mass dimension -2. This is approximately a factor 10^{120} too large. This discrepancy is known as the cosmological constant problem and is one of the unresolved issues of the SM.

2.9 The Spin-Statistics Theorem

One can construct two particle states by applying two creation operators on the vacuum.

$$\left|\overline{k}, \overline{k}'\right\rangle = 2\sqrt{\omega\omega'} a^\dagger(\overline{k}) a^\dagger(\overline{k}') \left|0\right\rangle \tag{2.89}$$

Additionally, one can also construct the state by applying the operators in the reverse manner

$$\left|\overline{k}', \overline{k}\right\rangle = 2\sqrt{\omega'\omega} a^\dagger(\overline{k}') a^\dagger(\overline{k}) \left|0\right\rangle \tag{2.90}$$

As we have discussed, for the scalar field ϕ, with spin $j = 0$, the creation operators of both scalar fields commute from (2.80). This implies that both two particle systems share common eigenstates and can exist simultaneously.

However, for particles with half-odd integer spins, the Pauli exclusion principle states that two particles cannot exist in exactly the same state. Therefore, we will need a different procedure to quantise the spin-$\frac{1}{2}$ spinor fields that describe the fermions.

To quantise fermions, such that they obey the correct statistics, we use anti-commutation relations

$$\left\{a_1^\dagger, a_2^\dagger\right\} = 0 \implies a_1^\dagger a_2^\dagger = -a_2^\dagger a_1^\dagger \tag{2.91}$$

This implies that

$$a_1^\dagger a_1^\dagger \left|0\right\rangle = -a_1^\dagger a_1^\dagger \left|0\right\rangle = 0 \tag{2.92}$$

and therefore the two identical particles cannot occupy the same state simultaneously.

The "spin statistics theorem" says that we quantise integer spin particles using commutation relations and anti-commutation relations for $\frac{1}{2}$ odd-integer spin particles [8]. This procedure naturally gives the correct spin statistics for Bosonic and Fermi-Dirac particles.

2.10 Second Quantisation of Spinor Fields

From the Dirac equation (2.41), we note that the conjugate momentum of the Dirac spinor field is

$$\Pi = \frac{\partial \mathcal{L}}{\partial \dot{\psi}} = i\overline{\psi}\gamma^0 = i\psi^\dagger \tag{2.93}$$

We continue quantising the spin-$\frac{1}{2}$ Dirac fields ψ, with the anti-commutation relations:

$$\{\psi_\alpha, \psi_\beta\} = 0 = \left\{\psi_\alpha^\dagger, \psi_\beta^\dagger\right\} \tag{2.94a}$$

$$\{\psi_\alpha, \psi_\beta\} = \delta_{\alpha\beta}\delta^{(3)}\left(\overline{x} - \overline{y}\right) \tag{2.94b}$$

where $\alpha, \beta = 1, 2$ label the two possible spin states.

Since the classical solution for the Dirac equation is a sum of plane waves, we can write the quantum operators $\psi, \psi^\dagger$ in terms of their expansion in a similar way to the scalar field case,

$$\psi(x) = \int \frac{d^3p}{(2\pi)^3} \frac{1}{\sqrt{2E_p}} \sum_s \left[b_p^s u^s(p) e^{ip\cdot x} + c_p^{s\dagger} v^s(p) e^{-ip\cdot x} \right] \tag{2.95a}$$

$$\psi(x)^\dagger = \int \frac{d^3p}{(2\pi)^3} \frac{1}{\sqrt{2E_p}} \sum_s \left[b_p^{s\dagger} u^s(p)^\dagger e^{-ip\cdot x} + c_p^s v^s(p)^\dagger e^{ip\cdot x} \right] \tag{2.95b}$$

where $p \cdot x = p^\mu x_\mu$ runs over four dimensions and s labels the spin. For fermions with spin $\frac{1}{2}$, $s = +\frac{1}{2}$ or $s = -\frac{1}{2}$.

As the fields are elevated to operators, the Fourier modes $b_p^{s\dagger}$ are considered operators that create particles associated to the spinors $u^s(p)$ and $c_p^{\alpha\dagger}$ can be thought of as an operator creating particles associated to $v^s(p)$.

These spinors correspond to the positive ($\psi \sim e^{-iEt}$) and negative ($\psi \sim e^{+iEt}$) energy solutions to the Dirac equation. Taking the simple ansatz that $\psi^+ = u(p)e^{-ip\cdot x}$ for the positive energy solutions, we see

$$(i\gamma^\mu p_\mu - m)\, u(p) = 0 \tag{2.96}$$

This has the solution

$$u(p) = \begin{pmatrix} \sqrt{p\cdot\sigma}\zeta \\ \sqrt{p\cdot\bar{\sigma}}\zeta \end{pmatrix} \tag{2.97}$$

with $\sigma' = (1, \sigma^i)$, $\bar{\sigma}^\mu = (1 - \sigma^i)$ and for any two-component spinor ζ which we normalise to $\zeta^\dagger\zeta = 1$ [6].

Similarly, the negative energy solution is $\psi^- = v(p)e^{ip\cdot x}$ such that

$$(i\gamma^\mu p_\mu + m)\, v(p) = 0 \tag{2.98}$$

with solution

$$v(p) = \begin{pmatrix} \sqrt{p\cdot\sigma}\eta \\ -\sqrt{p\cdot\bar{\sigma}}\eta \end{pmatrix} \tag{2.99}$$

for some two-component spinor η, normalised to $\eta^\dagger\eta = 1$.

With these solutions, the anti-commutation relations (2.94) can now be re-expressed as

$$\begin{aligned} \left\{b_p^r, b_q^{s\dagger}\right\} &= (2\pi)^3\,\delta^{rs}\delta^{(3)}(p-q) && (2.100a) \\ \left\{c_p^r, c_q^{s\dagger}\right\} &= (2\pi)^3\,\delta^{rs}\delta^{(3)}(p-q) && (2.100b) \end{aligned}$$

with all other anti-commutators vanishing [6].

We now look at the Hamiltonian $H = \int d^3x\,\left(i\overline{\psi}\gamma\gamma^0\partial_0\psi - \mathcal{L}\right)$. Expanding the fields in terms of their expansions (2.95) and rearranging the terms, we find

$$\begin{aligned} H &= \int \frac{d^3p}{(2\pi)^3}\psi\left(-j\gamma^j\partial_i + m\right)\psi \\ &= \int \frac{d^3p}{(2\pi)^3}E_p\left[b_p^{s\dagger}b_p^s + c_p^{s\dagger}c_p^s - (2\pi)^3\,\delta^{(3)}\,(0)\right] \end{aligned} \quad (2.101)$$

with $j = 1, 2, 3$.

The negative infinite term $-\,(2\pi)^3\,\delta^{(3)}\,(0)$ could in principle partially cancel the positive infinite term from quantising the bosonic fields but again this requires a seemingly miraculous and unnatural cancellation of divergent quantities.

Although we have quantised with the anti-commutation relations, there are still relevant commutation relations

$$\begin{aligned} \left[H, b_p^r\right] = -E_p b_p^r \quad &\text{and} \quad \left[H, b_p^{r\dagger}\right] = -E_p b_p^{r\dagger}b_p^r && (2.102a) \\ \left[H, c_p^r\right] = -E_p c_p^r \quad &\text{and} \quad \left[H, c_p^{r\dagger}\right] = -E_p c_p^{r\dagger} && (2.102b) \end{aligned}$$

Therefore, we can construct infinite towers of particle states using the creation operators in a similar manner to the scalar case (2.85). The $b^\dagger$ and $c^\dagger$ creation operators create distinct particles, indicating that the fermionic and anti-fermionic particles are different.

Further, the Hamiltonian now contains two number operators $N_b \propto b^\dagger b$ and $N_c \propto c^\dagger c$. Since the Hamiltonian is conserved (since energy is always conserved), the total number of particles

$$N \equiv N_b + N_c \tag{2.103}$$

is constant.

We have now introduced all possible states for both the bosonic spin-0 and fermionic spin-$\frac{1}{2}$ fields through the commutation and anti-commutation relations respectively.

2.10.1 Expectation Values with Canonical Quantisation

In Section 2.6.1, we used the path integral to construct the correlation function in a manifestly covariant formulation.

However, it is often simpler to compute directly the expectation value using the commutation relations that canonically quantise fields described in Sections 2.8 and 2.10. In canonical quantisation, multi-particle initial and final states are explicitly constructed using the creation and annihilation operators, which can then be used to derive more complicated expectation values as follows

The ground state for an interacting theory $|\Omega\rangle$ is in general not the same as the ground state of the free theory $|0\rangle$. However, one can always normalise the correlation functions appropriately. For example, the two point correlation function is expressed as

$$\langle\Omega|\,\phi\left(x\right)\phi\left(y\right)|\Omega\rangle = \lim_{t\to\infty}\frac{\langle 0|\,\phi\left(x\right)\phi\left(y\right)\exp\left[-i\int_{-t}^{t}dtH_I\right]|0\rangle}{\langle 0|\exp\left[-i\int_{-t}^{t}dtH_I\right]|0\rangle} \tag{2.104}$$

where H_I is the interacting portion of the Hamiltonian.

We typically use the perturbative expansion of the exponential and then Fourier expand the fields in terms of their creation and annihilation operators with (2.75) and (2.95). We then rearrange terms in the expansion such that all annihilation operators are on the right using the commutation relations (2.80) and (2.100). This procedure is known as "normal ordering" and simplifies the correlation functions to physically measurable quantities [8].

Furthermore, one can see that, to leading order of the exponential, the expectation value of any interaction is directly proportional to the terms in the Hamilton that contain exactly the fields involved in the process. There is therefore a direct correspondence between the terms that couple specific fields in Lagrangian and the computed probability of their interactions. In particular, the coupling strength is to leading order merely the coefficient of that specific Lagrangian term.

The higher order corrections from the exponential describe quantum processes and correspond to situations where unobserved particles are spontaneously created and annihilated from the vacuum. These unobserved interactions occur between the measured initial and final states, with the unphysical particles forming closed "loops". Notice that these terms always require additional vertices where these loops start and end. In the computation, this manifests itself as the additional prefactors of the leading-order coupling strength g, where for the typical weak coupling scenario $g < 1$ the perturbative corrections become progressively smaller.

In particle physics experiments, we often collide together two beams of particles with well defined momentum. The scattering of these particles is measured through the out-going particles that travel through detectors. The probability of an initial state $|i\rangle$ evolving to a final state $|f\rangle$ is expressed in terms of the cross section σ as

$$P_{fi} = \sigma \times \text{number density} \tag{2.105}$$

It is this probability, defined by the cross section, that is often the subject of interest and the measured quantity at particle colliders such as the Large Hadron Collider.

On the other hand, theoretical predictions for P_{fi} are made through the scattering (S) matrix defined as

$$S = \frac{\exp\left[-i\int dt H_I\right]}{\langle 0|\exp\left[-i\int_{-t}^{t} dt H_I\right]|0\rangle} \tag{2.106}$$

where the probability is given by

$$P_{fi} = |\langle f| S |i\rangle|^2 \tag{2.107}$$

In almost all physical cases, a valid perturbative expansion exists for the S-matrix. This allows us to truncate the series after a few terms and greatly simplifies calculations. From here, we can perturbatively compute probabilities of any particle interactions merely by specifying Lagrangian or its corresponding Hamiltonian (2.71).

Two Particle Scattering

For two particle scattering with initial particles A and B (mass m_A, m_B and momenta p_A and p_B) and final particles C and D (mass m_C, m_D and momenta p_C and p_D) , the cross section can be computed directly [9]. We shall restrict ourselves to the simple situation for non-polarised states without spin as these considerations are highly Lagrangian dependent and often present as non-trivial prefactors in the derivation Ref. [8].

In the limit that $m_A, m_B \ll s$ relevant in the majority of experimental situations, where $s = (p_A + p_B)^2$ is the centre of mass energy of the system, we find

$$\begin{aligned} d\sigma &= (2\pi)^4 \delta^{(4)}(P_f - p_A - p_B) \frac{\overline{|\mathcal{M}_{fi}|}^2}{2s} \frac{d^3 p_C}{2E_C (2\pi)^3} \frac{d^3 p_D}{2E_D (2\pi)^3} \\ &= \overline{|\mathcal{M}_{fi}|}^2 \frac{p}{32\pi^2 s^{3/2}} d\Omega \end{aligned} \tag{2.108}$$

where P_f is the final state momentum, $p \equiv |p_C|$, $d\Omega$ the infinitesimal solid angle around particle C and E_i are the energies of particle i. Notice that the factor $\delta^{(4)}(P_f - p_A - p_B)$ ensures momentum conservation. $\mathcal{M}_{fi}$ are the probability of interactions between initial and final states and can be set by experiment.

Lifetimes and Decay Widths

Restricting (2.108) to the decay of an unstable particle $R \to C + D$, the probability of decay is given by

$$d\Gamma = \frac{(2\pi)^4 \delta^{(4)}(P_f - p_R)}{2E_R} \overline{|\mathcal{M}_{fi}|}^2 \prod_{j=f} \frac{d^3 p_j}{2E_j (2\pi)^3} \tag{2.109}$$

with Γ known as the decay width.

In the rest frame of the final particle R, we have $E_R = m_R = \sqrt{s}$. It follows that we have the transition probability

$$d\Gamma = \overline{|\mathcal{M}_{fi}|}^2 \frac{p}{32\pi^2 m_R^2} d\Omega \tag{2.110}$$

which for an isotropic decay, with unpolarised state R,

$$\Gamma = \overline{|\mathcal{M}_{fi}|}^2 \frac{p}{8\pi m_R^2} \tag{2.111}$$

Notice that this in general implies the decay of a heavy particle R is inversely proportional to its mass-squared. Further, the momentum of one of the final state particles $p = |p_C|$ (or equivalently $p = |p_D|$) that appears in this equation, can be written as

$$p = \frac{1}{2m_R} \left\{ \left[m_R^2 - (m_C + m_D)^2 \right] \left[m_R^2 - (m_C + m_D)^2 \right] \right\}^{\frac{1}{2}} \tag{2.112}$$

If the initial state particle is sufficiently heavy in comparison to the final state particles, this reduces to $p = \frac{1}{2} m_R$.

Explicitly, the matrix elements $\mathcal{M}_{fi}$ correspond to the transition probabilities between initial and final states and are related to the theoretical probabilities of (2.107) by

$$P_{fi} = \delta_{fi} + (2\pi)^4 \, \delta^{(4)}(P_f - P_i)(-i\mathcal{M}_{fi}) \prod_{j=f,i} \frac{1}{\sqrt{2E_j}} \tag{2.113}$$

Furthermore, Feynman proposed a diagrammatical representation for each perturbative term in the S-matrix of (2.106). Along with these “Feynman diagrams”, he associated a set of “Feynman rules” to quickly write down the contribution of each term [8, 16]. With the relation in (2.113), this procedure can thus be used to directly compute the theoretical value for $\mathcal{M}_{fi}$, allowing a comparison to the probabilities observed in experiments.

We are therefore left with specifying the Lagrangian that describes these particles. As we have seen, this completely determines all particle dynamics: the probabilities of all particle interactions through the S-matrix (2.106) and all equations of motion through the Euler-Lagrange equation (1.26).

Historically, the Lagrangian that governs all particle was constructed piecewise to describe various experimental observations and theoretical insights; it is known as the Lagrangian of the “Standard Model” [9].

Part II

The Standard Model and Beyond

Chapter 3

The Standard Model

The Standard Model (SM) provides both a mathematically consistent framework and an accurate description of almost all experimental results to date. In this chapter, we shall review the key theoretical concepts involved and the equations that govern nature.

To construct the SM, we identify the symmetries of nature and the group representations of the fermion and scalar fields. From here, we write down the most general renormalisable Lagrangian permissible by the symmetries. Renormalisability shall be discussed in Section 3.3.1. Historically, the group symmetries that were required were found to be too constrictive to give the exact particle spectrum observed in experiments and a symmetry breaking mechanism was introduced [9].

3.1 The Standard Model

The symmetries that govern the SM are described by the "Yang-Mills" non-abelian product group

$$SU(3)_C \times SU(2)_L \times U(1)_Y \tag{3.1}$$

where C denotes the colour symmetry of quantum chromodynamics (QCD), L the left-handed fields of the weak theory (the right-handed fields will be singlets of $SU(2)$) and $Y = 2Q - T_3$ the weak hypercharge of the unified electroweak theory where Q is the electromagnetic charge and $T_3 = \frac{\sigma_3}{2}$ is the third component of isospin.

We shall see that this product group will define the Lagrangian and therefore the equations of motion that govern all observed particle dynamics.

3.1.1 Particle Content

In 1964, Murray Gell-Man and George Zweig independently used data collected in e^- and p^+ collisions to suggest that the neutron and the proton were composite particles, which were made up of three point-like spin-$\frac{1}{2}$ particles of $-\frac{1}{3}$ or $+\frac{2}{3}$ charge [33,34]. These fundamental particles were named "quarks". Over the course of the 50 years since, the "zoo" of all composite particles has slowly been organised by their characteristic properties. It was discovered that there was an underlying, finite and indeed small set of particles, which grouped together, make up all known matter.

The fundamental particles are grouped into the fermions, which are spin-$\frac{1}{2}$ particles obeying Fermi-Dirac statistics, and the gauge bosons, one set for each symmetry group, which are spin-1 particles that obey Bose-Einstein statistics. The fundamental fermions are further subdivided into leptons and quarks, each with distinctive properties. These elementary particles are summarised in Fig. 3.1

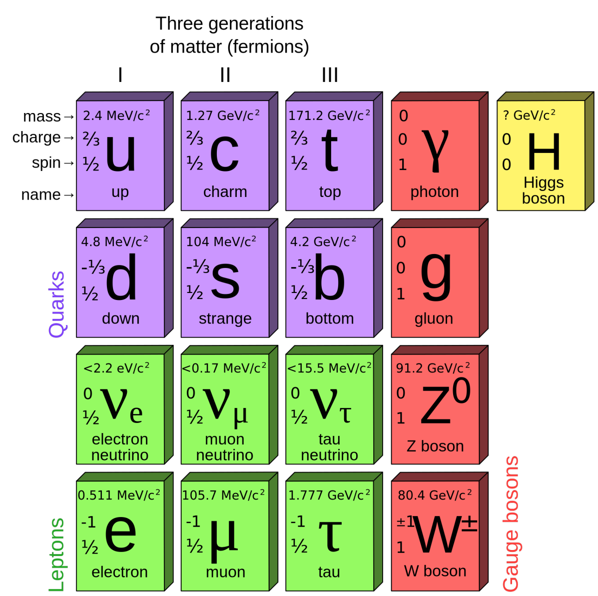


Figure 3.1: Fundamental particles of the Standard Model. [35]

Along with these particles, there is also an associated set of right-handed particles for all fermions except the neutrinos.

The fermions naturally fall into three families, known as "generations", each a heavier copy of the preceding one with all quantum numbers the same [4]. The classification of the composite particles known as baryons, which are made up of quarks, is described in Appendix A.

3.1.2 Fundamental Forces

There are four fundamental forces that govern all known interactions in nature summarised in Table 3.1; the three that make up the SM are governed by a special symmetry group, known as a "gauge" group. The fourth, gravity, remains one of the largest missing components of the theory.

- The electromagnetic force is responsible for both electromagnetic attraction and repulsion and is governed by the theory known as quantum electrodynamics (QED).
- The weak force, which governs how protons decay to neutrons is described by the unified electroweak theory.
- The strong force is the strongest of the four fundamental forces and is governed by quantum chromodynamics (QCD). It is responsible for binding atomic nuclei together and its strong coupling requires it be a non-perturbative theory. Fortunately the notion of asymptotic freedom will allow us to do meaningful calculations on measurable quantities.
- The theory of gravity known as General Relativity was derived by Einstein [36, 37], and is not currently part of the SM. It is inherently difficult to quantise as the theory describes smooth and continuous space-time. Furthermore, it is by far the weakest of the four forces and is therefore very difficult to probe experimentally.

Gauge Group	Force	Gauge Bosons [spin]	Acts on	Strength	Range
$U(1)_Y$	Electromagnetism	γ [1]	electric charges	10^{-2}	∞
$SU(2)_L$	Weak	3 $W^{\pm}, Z^0$ [1]	fermions	10^{-5}	10^{-18} m
$SU(3)_C$	Strong	8 gluons [1]	quarks	1	10^{-15} m
-	Gravity	graviton* [2]	mass	10^{-39}	∞

Table 3.1: The ingredients of the SM: the listed strengths are normalised to the strong force and are derived from the characteristic constants of the theory [4]. * the graviton is the hypothetical boson for the quantised general relativistic theory of gravity.

During the early 1960s Sheldon Lee Glashow, Abdus Salam, and Steven Weinberg independently unified the $SU(2)_L$ and $U(1)_Y$ groups to describe the electroweak force [14,38,39].

However, this product group was too restrictive to incorporate particles with mass. In 1964, three independent groups of theorists Brout & Englert [40]; Guralnik, Hagen, & Kibble [41]; and Higgs [42] proposed a mechanism for spontaneous symmetry breaking (SSB). This mechanism broke just enough of the imposed symmetries with

$$SU(2)_L \times U(1)_Y \longrightarrow U(1)_{\mathrm{EM}} \tag{3.2}$$

to allow mass into the theory, whilst leaving the photon of electromagnetism massless. Furthermore, SSB predicted an additional particle known as the Higgs boson, evidence for which has been recently seen [43].

The interactions between the elementary particles are summarised in Fig. 3.2; each line indicates the possible interactions between particles.

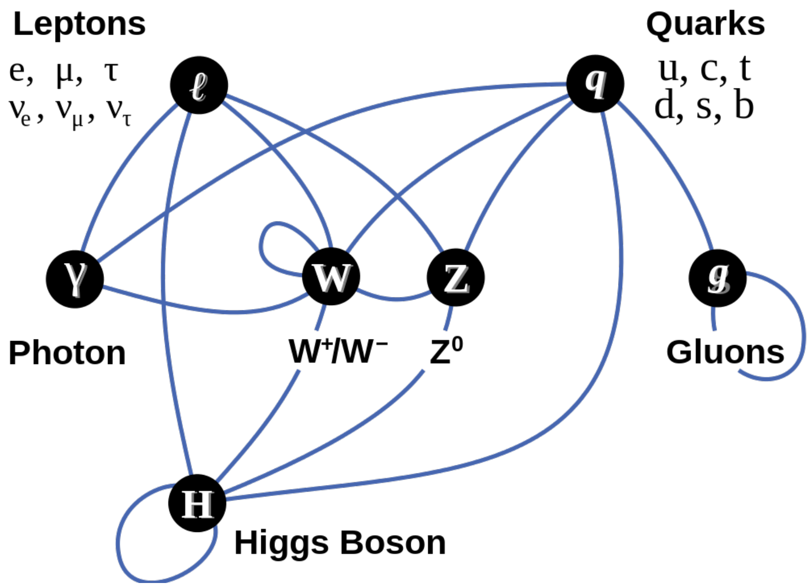


Figure 3.2: Interactions between the fundamental particles. [35]

Physically observed fermions and bosons are embedded into Lorentz fields that fall into particular irreducible representations of the gauge groups. The representations of the gauge groups are described by the quantum numbers

$$(C, L)_Y \tag{3.3}$$

where C, L, and Y label the representations of the $SU(3)_C$, $SU(2)_L$ and $U(1)_Y$ gauge groups respectively. Since the hypercharge is one-dimensional, the quantum number Y is just a constant.

Using this notation, the representations of the fermion spin-$\frac{1}{2}$ fields are shown in Table 3.2 below.

Field	Representation	Particles
Q_L	$(\mathbf{3},\mathbf{2})_{1/3}$	left-handed quarks
L_L	$(\mathbf{1},\mathbf{2})_{-1}$	left-handed leptons
U_R	$(\overline{\mathbf{3}},\mathbf{1})_{-4/3}$	right-handed up-type quarks
D_R	$(\overline{\mathbf{3}},\mathbf{1})_{2/3}$	right-handed down type quarks
E_R	$(\mathbf{1},\mathbf{1})_{2}$	right-handed electron

Table 3.2: Fermion fields of the Standard Model.

All fermion fields describe both fermions and anti-fermions since they arise together in the spinor field in(2.95a). As such, all fermions (and similarly all scalar bosons) have antiparticle counterparts.

There are three sets of known gauge fields, one for each of the SM gauge groups. These fields will arise from imposing the gauge symmetries on our theory. These are spin-1 vector fields and fall into the representations given in Table 3.3.

Gauge Group	Field	Representation
$U(1)_Y$	B_μ	$(\mathbf{1},\mathbf{1})_0$
$SU(2)_L$	A_μ	$(\mathbf{1},\mathbf{3})_0$
$SU(3)_C$	G_μ	$(\mathbf{8},\mathbf{1})_0$

Table 3.3: Gauge fields of the Standard Model.

The quantisation of these gauge fields are physically manifested as the gauge bosons, particles that mediate the interactions of their corresponding gauge groups.

The additional scalar field ϕ, known as the Higgs field, is introduced to break the electroweak $SU(2)_L \times U(1)_Y$ symmetry and transforms under the representation given in Table 3.4.

Field	Representation	Description
ϕ	$(\mathbf{1},\mathbf{2})_{-1}$	complex scalar doublet

Table 3.4: Scalar field of the Standard Model.

After electroweak spontaneous symmetry breaking (EWSSB), this field gives rise to the (tentatively) observed Higgs boson. Additionally, the gauge bosons that remain are the eight gluons from $SU(3)_C$, and the electroweak gauge bosons, which are linear combinations of the electroweak gauge fields A and B of $SU(2)_L \times U(1)_Y$. These include the photon (γ) of QED, which remains massless, and the massive $W^{\pm}$ and Z^0 of the weak theory.

The neutral photon is the quantisation of light and was first proposed by Einstein in 1905 to explain the photoelectric effect and [44]. The $W^{\pm}$, Z^0 were directly observed in 1983 in pp collisions at CERN [45–48]. The eight gluons of the strong force were experimentally discovered in 1979 in e^-p^+ collisions at DESY in Hamburg [49]. The gluons exhibited confinement and asymptotic freedom properties similar to quarks. In total, there are twelve known gauge bosons [4].

The SM also exhibits a few global symmetries, such as the conservation of baryon and lepton numbers separately with the symmetry group $U(1)_B \times U(1)_e \times U(1)_\mu \times U(1)_\tau$ (where e, μ and τ indicate the three lepton families). The global symmetries are not imposed but can be viewed as accidental symmetries that arise coincidentally from certain parts of the gauge theory.

There is however, an assumed global charge-parity-time (CPT) symmetry of the SM [8, 9, 21]. This is taken to be an exact symmetry of nature since all particle interactions have been observed to obey this. Therefore, the equations of the motion of the left-handed particles must also be satisfied by the right-handed antiparticles. Indeed, the Feynman-Stueckelberg interpretation states that antiparticles can be viewed as the charge conjugate particles, travelling backwards in time [50].

Finally, there are nineteen free parameters in the theory, all of which shall be discussed in this chapter. These are measurable quantities that are not defined by the SM and must be experimentally measured. It is possible that this rather large set of free parameters stems from a deeper underlying theory [50].

3.2 Gauge Theories

We have already described the quantisation procedure that gives rise to particles from bosonic (spin-1) and fermionic (spin-$\frac{1}{2}$) fields in Section 2.4 and Section 2.5. These were Lorentz vector and spinor fields that transform under a particular representation of the Lorentz group.

To construct a relativistic theory, one merely needs to write down a Lagrangian that transforms trivially under the Lorentz group. Each term, constructed from a tensor product of these Lorentz fields, is therefore required to form a singlet of the Lorentz group. In general, a symmetry can be imposed in this way, by requiring that the total Lagrangian transforms trivially under that symmetry group.

However the most general Lorentz invariant Lagrangian is too broad to describe the universe that we observe. Additional local symmetries, in the form of Lie groups and Lie algebras, were required to restrict the permissible class of equations to those that describe the observed dynamical theories of particles [51]. The resultant Lagrangians describe local theories that depend on space-time position and are known as "gauge theories".

These local symmetries represent a redundancy of information in the theory and understanding them is central to tackling the many infinities that arise from integrating quantum fields over all of space-time [27].

In all gauge theories, the quantisation of the Lagrangian gives rise to an over-counting of possible field configurations, all related by gauge symmetry [8]. There is no canonical procedure for selecting one solution that describes a quantum state from the range of physically equivalent ones. However, it is possible to break the gauge symmetry by choosing a particular gauge fixing condition. Doing so requires the addition of Faddeev-Popov ghost fields to preserve unitarity [52, 53]. These fields correspond to unphysical virtual particles that never appear as external states and whose exact formulation is dependent on gauge choice. In general, we shall ignore these complications.

The simplest case of a gauge symmetry that we can consider is of the abelian $U(1)$ group. We shall see that imposing a local $U(1)$ symmetry on the Dirac Lagrangian (2.41) will result in the quantum theory of electromagnetism (EM), known as quantum electrodynamics

(QED). This theory partly governs the dynamics and interactions between photons and electrons.

3.3 Quantum Electrodynamics $U(1)_{\mathbf{EM}}$

As mentioned, the starting point of QED is the Dirac Lagrangian:

$$\mathcal{L}_{\text{Dirac}}(x) = \overline{\psi}\left(i\partial\!\!\!/ - m\right)\psi \tag{3.4}$$

There is an obvious one-dimensional global $U(1)$ symmetry. Transforming the field ψ by the phase $e^{i\alpha}$ with α constant, leaves the Lagrangian invariant.

To *gauge* a symmetry, we attempt to make the factor $\alpha = \alpha(x)$ depend on the space-time coordinate. For this one-dimensional case we will suppress the "generator" Y of the symmetry, which form a basis of the gauge group, such that $\alpha(x) \equiv \alpha(x) \cdot Y$. In this way, we will elevate the global $U(1)$ symmetry of the classical Lagrangian to a local one, making it *gauge* invariant.

We do this by imposing a local $U(1)$ symmetry on the Dirac field $\psi(x)$ such that the field transforms as

$$\psi(x) \longrightarrow e^{i\alpha(x)}\psi(x) \tag{3.5}$$

A complication arises from the kinetic energy term in the Lagrangian of such a field. This is typically of the form $(\partial\phi)^{\dagger}\partial\phi$ for scalars and $\overline{\psi}i\partial\!\!\!/\psi$ for spinors where $\overline{\psi} = \psi^{\dagger}\gamma^{0}$. In general, this is not in invariant under the gauge transformation such that the Dirac equation is also not invariant:

$$\begin{aligned} \mathcal{L}_{\text{Dirac}}(x) &\rightarrow \overline{\psi}e^{-i\alpha(x)}\left(i\partial\!\!\!/ - m\right)e^{i\alpha(x)}\psi \\ &= \overline{\psi}\left(i\partial\!\!\!/ - m\right)\psi - \overline{\psi}\gamma^{\mu}\frac{\partial\alpha}{\partial x^{\mu}}\psi(x) \end{aligned} \tag{3.6}$$

We need an additional term to cancel the extra factor in (3.6). We therefore construct the "covariant derivative" D_{μ} that will have the correct transformation properties to explicitly cancel this contribution.

We introduce the gauge field A^μ that transforms under this local $U(1)$ symmetry as

$$A_\mu \longrightarrow A'_\mu = A_\mu + \frac{1}{e}\partial_\mu \lambda(x) \tag{3.7}$$

where e is the "charge" of the field A_μ, which is experimentally identified as the electromagnetic charge of the electron. Note that we have again suppressed the explicit dependence on the basis generators such that $A_\mu \equiv A_\mu \cdot Y$.

Since we are imposing a one-dimensional symmetry, we expect $\lambda(x)$ to also be one-dimensional and to belong to the $U(1)$ gauge group. Explicitly, we start with the generic Maxwell equations (2.9), and set the sources equal to zero such that

$$0 = \partial^\mu F_{\mu\nu} = \partial^\mu(\partial_\nu A_\mu - \partial^\mu A_\nu) \tag{3.8}$$

Under the transformation of (3.7), $\partial^2\lambda = 0$; this is just a wave-equation with solution $\lambda = e^{i\theta}$. Since θ is one-dimensional, and λ is periodic with periodicity 2π, λ is therefore a member of $U(1)$.

We can now define the covariant derivative

$$D_\mu \psi = \partial_\mu \psi + iq\zeta_\mu \psi \tag{3.9}$$

for some currently arbitrary constant q and vector field ζ_μ. This transforms as

$$D_\mu \psi \longrightarrow i(\partial_\mu \alpha)(e^{i\alpha}\psi) + e^{i\alpha}\partial_\mu \psi + iq\zeta'_\mu e^{i\alpha}\psi \tag{3.10}$$

Since $\overline{\psi} \to \overline{\psi}e^{-i\alpha}$ (where $\overline{\psi} = \psi^\dagger\gamma^0$), we only have to require $[D_\mu, e^{i\alpha}]\psi = 0$ to make the term $\overline{\psi}D_\mu\psi$ gauge invariant. Thus, by taking $\zeta_\mu \to \zeta'_\mu = \zeta_\mu - \frac{1}{q}(\partial_\mu\alpha)$, we find $D_\mu(e^{i\alpha}\psi) = e^{i\alpha}D_\mu\psi$ as required.

Note that this transformation rule of ζ is exactly the same as that for the electromagnetic $U(1)$ symmetry in (3.7), with the identification that $q = -e$, $\alpha = \lambda$ and $\zeta = A_\mu$. Indeed, the experimental coupling between the electron and the photon is precisely described by identifying the gauge field A_μ with the photon and spinor field ψ with the electron.

We therefore find the fundamental covariant derivative D_μ for QED to be given by

$$D_\mu \psi = \partial_\mu \psi - ieA_\mu \psi \tag{3.11}$$

where the field A_μ transforms as

$$A_\mu \longrightarrow A_\mu + \frac{1}{e}\partial_\mu \alpha(x) \tag{3.12}$$

with $\alpha(x)$ corresponding to the local symmetry transformation and where we have again suppressed the generator where $A_\mu \equiv A_\mu \cdot Y$.

Note that it is this covariant derivative that introduces interaction terms between the electrons ψ and photons A_μ. The charge e acts as the coupling constant and determines the strength of interaction. For a chargeless field with $e = 0$, the covariant derivative is equal to the conventional partial derivative and the theory becomes a non-interacting one. The field transformation of A_μ is no longer well-defined as the A_μ in this case decouples from the theory.

Thus, the Dirac Lagrangian with this additional $U(1)$ symmetry is

$$\mathcal{L}_{\mathrm{Dirac},U(1)}(x) = \overline{\psi}(\not{D} - m)\psi \tag{3.13}$$

where $\overline{\psi} \equiv \psi^\dagger \gamma^0$.

Now, the transformation $\psi \to e^{i\alpha(x)}\psi$ gives

$$\begin{aligned}
\mathcal{L}_{\mathrm{Dirac},U(1)}(x) \quad &\to \quad \overline{\psi} e^{-i\alpha(x)} \left(i\gamma^\mu \partial_\mu - m\right) e^{i\alpha(x)}\psi + \overline{\psi} e^{-i\alpha(x)} \left(i\gamma^\mu \left(iqA_\mu - \partial_\mu \alpha(x)\right)\right) e^{i\alpha(x)}\psi \\
&= \quad \overline{\psi}\left(i\gamma^\mu \left(\partial_\mu + iqA_\mu\right) - m\right)\psi \\
&= \quad \mathcal{L}_{\mathrm{Dirac},U(1)}(x)
\end{aligned} \tag{3.14}$$

as desired.

There is one more complication that we need to consider. Since there are currently no derivative terms of A_μ in (3.13), A_μ could be treated as a constant background field that can be ignored. The theory would then reduce back to the original Dirac Lagrangian with the regular partial derivatives that is not invariant under the local $U(1)$. To avoid this, we need to introduce dynamics for the field A_μ. The appropriate gauge invariant term is

$$\mathcal{L}_A^{kin}(x) = -\frac{1}{4} F_{\mu\nu} F^{\mu\nu} \tag{3.15}$$

where $F^{\mu\nu} = \frac{i}{q}\left[D^\mu, D^\nu\right]$ is the field strength.

Explicitly,

$$\begin{aligned} F^{\mu\nu}\psi(x) &= \frac{i}{q}\left[D^\mu, D^\nu\right]\phi \\ &= \left(\partial^\mu A^\nu - \partial^\nu A^\mu - q\left[A^\mu, A^\nu\right]\right)\phi(x) \end{aligned} \tag{3.16}$$

where ϕ is an arbitrary scalar field. The commutator in this equation stems from the commutator of the group generators; for this 1-dimensional case, we have simply suppressed the explicit dependence on the generator Y by the relabelling $A_\mu \equiv A_\mu Y$. In particular, the commutator of these generators defines the so-called "Lie Algebra" [54]. This shall become clearer when we consider more complicated group theories.

For the abelian $U(1)$ theory that we are considering, we have commuting symmetry generators (and thus commuting fields A_μ) and so the commutator here vanishes and the field strength $F^{\mu\nu}$ reduces to the one seen in classical electromagnetism (2.5), which is gauge invariant by (2.14).

Finally, the full QED $U(1)$ gauge invariant Lagrangian that describes the interaction of electrons ψ with photons A_μ is

$$\mathcal{L}_{\mathrm{QED}}(x) = \overline{\psi}(\not{D} - m)\psi - \frac{1}{4} F^{\mu\nu} F_{\mu\nu} \tag{3.17}$$

If we expand the covariant derivative in (3.17), we find

$$\begin{aligned} \mathcal{L}_{\mathrm{QED}}(x) &= \overline{\psi}(\not{\partial} - m)\psi - \frac{1}{4} F^{\mu\nu} F_{\mu\nu} - J^\mu A_\mu \\ &= \mathcal{L}_{\mathrm{Dirac}} + \mathcal{L}_{\mathrm{EM}} \end{aligned} \tag{3.18}$$

where we have identified the Noether current associated to the $U(1)$ symmetry of the field A_μ by

$$J^\mu = ie\overline{\psi}\gamma^\mu\psi \tag{3.19}$$

Therefore, electromagnetism naturally arises from imposing a local $U(1)$ symmetry on the Dirac Lagrangian. We will use a similar approach for the weak and strong fundamental forces. Specifying Lie groups and imposing gauge invariance on the Lagrangian, we shall write down all terms that invariant under the symmetry transformations. In turn, this Lagrangian shall give rise to equations of motion that represent nature.

To ensure gauge invariance for a general group, a vector field is introduced for every group generator. In the case of electromagnetism, the generator $\alpha(x) \in U(1)$ required the existence of the vector field A_μ to construct the covariant derivative in (3.11).

3.3.1 Renormalisabilitiy

By embedding a local $U(1)$ symmetry on the Dirac Lagrangian, we have somewhat miraculously ended up with the equations of motion that govern the physical theory of QED.

Instead of starting at the Dirac equation, we could have naïvely written down the most generic $U(1)$ gauge invariant Lagrangian. This Lagrangian would contain QED but also additional terms that would appear in the invariant potential, for example

$$V\left(\overline{\psi}\psi\right) \supset \overline{\psi}\gamma^5\gamma^\mu\psi + \overline{\psi}\gamma^5\psi + \overline{\psi}\gamma^\mu\psi + \overline{\psi}\gamma^\mu\gamma^\nu\psi \tag{3.20}$$

The first two terms are "chiral", distinguishing left and right-handed Weyl spinors, whilst the remaining two are vector-like terms that do not. In general, these introduce additional interactions into the theory [8].

Furthermore, there could also be higher order terms, such as $|\psi|^4$. However, almost all of these terms fall into the category of "non-renormalisable" theories, which yield uncontrolled divergent quantities. Specifically, terms that have positive mass-dimension coupling constant $[g] > 0$ are non-renormalisable.

This is simple to see directly. The Lagrangian density must have mass dimension $[\mathcal{L}] = 4$ for the action S to be dimensionless. Then, any term $\mathcal{L} \supset gf(\psi)$ with coupling $[g] > 0$, would require the (operator) $f(\psi)$ to have dimension $[f(\psi)] < 4$ (equivalently a length dimension of > -4). Integrating this over the whole of four space gives

$$\begin{aligned} S &= \int d^4x\mathcal{L} \\ &= g\int d^4x f(\psi) \;\propto\; \infty \end{aligned} \tag{3.21}$$

where $\left[\int d^4x\right] = -4$.

Note that it is still possible to admit some non-renormalisable terms into the theory by introducing a momentum cutoff p^4/Λ^2, with Λ^2 large enough to suppress these interactions at the relatively low energies experiments have hitherto probed. The infinities that plague the low energy theory will naturally be cancelled by contributions from beyond the SM (BSM) terms, typically introduced through additional particles. Above this cutoff, new physics enters into the theory and renders the grander high energy theory perfectly renormalisable.

3.4 Electroweak Unification $SU(2)_L \times U(1)_Y$

Experiments of weak processes were seen to involve electromagnetic charged currents [45–48, 55, 56]. These results implied that there was some unified gauge theoretic description of the electroweak force.

In 1961, Glashow proposed the symmetry structure of what is now known as the Weinberg-Salam electroweak theory. The symmetry structure he proposed unified electromagnetism with the weak force through a gauge theory with the product group

$$G_{\text{EWK}} = SU(2)_L \times U(1)_Y \tag{3.22}$$

This theory required a set of four gauge fields corresponding to the four group generators of the product group. These fields describe massless messenger particles that mediate the unified electroweak interaction. Physically, however, the weak force was found to be short range implying that these particles are massive. The underlying symmetry must be broken by the Higgs field H in such a way as to give rise to masses to the particles exchanged in weak interactions but not to the massless photon, which is exchanged in electromagnetism interactions.

Taking the generic fermion field ψ the transformation under the local $SU(2)_T \times U(1)_Y$ group is given by

$$\psi(x) \longrightarrow e^{i\alpha(x)\cdot T + i\beta(x)Y}\psi(x) \tag{3.23}$$

where T is the generator for $SU(2)$ and Y is the hypercharge of $U(1)_Y$.

The electroweak force was found to be a "chiral" theory that distinguishes left and right handed fields [57, 58]. Furthermore, to match the experimental results all left-handed fermion fields are required to belong to the $T = 1/2$ doublet representation of $SU(2)_L$, with all right-handed fields belong to the singlet $T = 0$ representation [8].

Then, the covariant derivative in the fundamental representation is constructed as

$$D_\mu\psi(x) = \left(\partial_\mu - igA^a_\mu(x)T_a - ig'B_\mu(x)Y\right)\psi(x) \tag{3.24}$$

where the gauge couplings g and g' of $SU(2)$ and $U(1)$ respectively. The three fields $A^a_\mu(x)$ are the $SU(2)_L$ gauge fields associated to their respective weak generators $T_i = \sigma_i/2$ and the field B_μ is the gauge field of $U(1)_Y$ associated to the hypercharge generator $Y \propto \mathbb{I}_{2\times 2}$.

Explicitly, one can write

$$\begin{aligned} gA^a_\mu T_a + g'B_\mu Y &= \frac{g}{2}\left(A^1_\mu\sigma^1 + A^2_\mu\sigma^2 + A^3_\mu\sigma^3\right) - \frac{g'}{2}B_\mu\mathbb{I}^{2\times 2} \\ &= \frac{1}{2}\begin{pmatrix} gA^3_\mu - g'B_\mu & g(A^1_\mu - iA^2_\mu) \\ g(A^1_\mu + iA^2_\mu) & -gA^3_\mu - g'B_\mu \end{pmatrix} \end{aligned} \tag{3.25}$$

The transformation properties of the gauge field under $SU(2)_L \times U(1)_Y$ are given by

$$A_\mu(x)\cdot T \longrightarrow e^{i\alpha(x)\cdot T}A_\mu(x)\cdot Te^{-i\alpha(x)\cdot T} + \frac{1}{g}e^{i\alpha(x)\cdot T}i\partial_\mu e^{-i\alpha(x)\cdot T} \tag{3.26a}$$

$$B_\mu(x)Y \longrightarrow B_\mu(x)Y + \frac{1}{g'}\partial_\mu\beta(x)Y \tag{3.26b}$$

Thus,

$$D_\mu \psi(x) \longrightarrow e^{i\alpha(x)\cdot T + i\beta(x)Y} D_\mu \psi(x) \tag{3.27}$$

transforms exactly as (3.23), ensuring gauge invariance.

Further, writing the field in its left $L(x)$ and right-handed $R(x)$ components, we have

$$D_\mu L(x) = \left(\partial_\mu - ig\frac{1}{2}A_\mu(x)\cdot T - ig'\frac{1}{2}B_\mu(x)Y\right) L(x) \tag{3.28a}$$

$$D_\mu R(x) = \left(\partial_\mu - ig' B\,(x)\,Y\right) R(x) \tag{3.28b}$$

The kinetic term for the fermions is then

$$\mathcal{L}_{\text{fermions}}(x) = \overline{L}(x) i\gamma^\mu D_\mu L(x) + \overline{R}(x) i\gamma^\mu D_\mu R(x) \tag{3.29}$$

The field strength for the $SU(2)_L$ gauge fields is

$$F_{\mu\nu}(x)\cdot T = \partial_\mu A_\nu(x)\cdot T - \partial_\nu A_\mu(x)\cdot T - gA^i_\mu(x)A^j_\nu(x)\,[T_i, T_j] \tag{3.30}$$

where $i, j = 1, 2, 3$ and labels the generators of the $SU(2)_L$ group. This obeys the transformation rule

$$F_{\mu\nu}\cdot T \longrightarrow e^{i\alpha(x)\cdot T} F_{\mu\nu}(x)\cdot T e^{-i\alpha(x)\cdot T} \tag{3.31}$$

The analogous field strength for the $U(1)_Y$ gauge fields is

$$B_{\mu\nu}(x)Y = \partial_\mu B_\nu(x)Y - \partial_\nu B_\mu(x)Y \tag{3.32}$$

which transforms similarly to the canonical Maxwell field strength under gauge transformation (2.14).

The electroweak Lagrangian for the gauge fields is described by a generalisation of the equivalent terms in the EM Lagrangian,

$$\mathcal{L}_{\text{gauge}}(x) = -\frac{1}{4}F_{\mu\nu}F^{\mu\nu}(x) - \frac{1}{4}B^{\mu\nu}(x)B_{\mu\nu}(x) \tag{3.33}$$

where we have suppressed all generator dependence. $\mathcal{L}_{\text{gauge}}$ now represents a non-trivial interacting theory with non-linear self coupling terms dependent on the coupling constant g from $F_{\mu\nu}$. These non-abelian gauge fields are charged under the gauge group and will couple to all fields with any such charge, including themselves. In a similar manner, gravity is also inherently a non-linear theory, since the gravitational field couples to everything with energy density including itself.

Importantly, since left-handed and right-handed fields transform under different representations of $SU(2)_T \times U(1)_Y$ there is no way to add an ad-hoc mass term for fermions $\propto m\overline{\psi}\psi$ as there was in QED. Specifically, the term

$$m\overline{\psi}\psi = m\overline{\psi}_L\psi_R + m\overline{\psi}_R\psi_L \tag{3.34}$$

cannot form an (invariant) singlet under $SU(2)_L$ [54]. It is therefore forbidden by the gauge symmetry. The mass terms for the gauge fields are also forbidden for an analogous reason. In this way, all particles in the electroweak theory are currently massless due to the gauge symmetry. We are therefore required to introduce a symmetry breaking mechanism to match the observed masses of particles [4].

3.5 Spontaneous Symmetry Breakdown

Adding mass terms in by hand breaks the gauge invariance and in fact ruins the renormalisability of the theory, yielding ill-defined infinites [8].

Experimentally however, we observe only one massless vector boson, the photon, with all other particles gaining mass [4]. To reconcile the theory with the observation that the weak force is indeed short-ranged, we must generate masses in such a way as to preserve the renormalisibility of the theory.

The notion of spontaneous symmetry breaking (SSB), first introduced in 1964, provided a mechanism for particles to gain mass. However, it was not immediately apparent that such a theory was inherently renormalisible and it was not until 1971 that 't Hooft and Veltman showed the explicit renormalisibility of spontaneously broken gauge theories [59, 60]. The mechanism quickly became main-stream and shall be described in this section.

SSB stems from the assumption that we live in an imperfectly symmetric universe, with nature having symmetries that are not manifest to us. To formulate such a theory, we keep the Lagrangian symmetric but introduce at least one field (H) with an asymmetric particle vacuum state, characterised by a non-zero vacuum expectation value

$$\langle 0| H |0\rangle \neq 0 \tag{3.35}$$

In the SM the field H (known as the Higgs field) has to have spin-0 since the vacuum would otherwise have a non-zero angular momentum breaking the rotational invariance observed in nature. Furthermore, the Universe's translational invariance implies that the vacuum expectation must be independent of space-time. Thus, this spinless field uniformly permeates all space-time and by assumption must couple to all other massless particles to give them their observed mass. The asymmetric particle vacuum of the Higgs field will be defined by the vacuum expectation value v, which is a free parameter of the SM.

Picking a particular ground state lifts the degeneracy of the minimum energy states, spontaneously breaking the manifest symmetry. Note that the underlying symmetry of the theory, essential for renormalisability, is preserved and is merely hidden in the equations of motion.

In particular, the QM vacuum is defined as a complex state in which pairs of particles and antiparticles are spontaneously created and annihilated (and indeed it is the ground state on which the creation and annihilation operators of Chapter 2 act). If these virtual particles interact strongly enough amongst themselves, the vacuum forms a permanent high density state known as a "condensate". For such a vacuum state to have a condensate, we must cross a thermodynamical phase transition. This requires an interaction between particles that has sufficiently strong attraction at low density in addition to a sufficiently strong repulsion at high density (to prevent a runaway situation) [7].

In the SM, the condensate known as the Higgs vacuum condensate is assumed to exist. The otherwise massless fermions and bosons that populate the vacuum interact with condensate giving them their masses.

3.5.1 Global Symmetries

We first look at the SSB of a global symmetry following Ref. [7]. We start with an underlying global $U(1)$ symmetric Lagrangian with complex scalar field φ.

The field transforms as

$$\varphi \longrightarrow e^{i\alpha}\varphi \quad , \quad \varphi^\dagger \longrightarrow e^{-i\alpha}\varphi^\dagger \tag{3.36}$$

where for $U(1)$ scalar (singlet) fields $\varphi^\dagger = \varphi^\star$.

The Lagrangian is

$$\mathcal{L} = \partial_\mu\varphi\partial^\mu\varphi - V(\varphi\varphi^\dagger) \tag{3.37}$$

where the potential has of the form

$$V(\varphi^\dagger\varphi) = \frac{1}{2}\mu^2\varphi^\dagger\varphi + \frac{1}{4}\lambda\left(\varphi^\dagger\varphi\right)^2 \tag{3.38}$$

with constants $\lambda, \mu \in \mathbb{R}$.

There are three possible situations for the configuration of the potential:

- $\lambda < 0$
- $\lambda \geq 0,\ \mu^2 > 0$
- $\lambda > 0,\ \mu^2 < 0$

The first case, $\lambda < 0$, is unphysical and corresponds to a potential with no stable minima.

The second case, $\lambda > 0$, $\mu^2 > 0$, has a unique stable minimum at $|\varphi| = 0$ and is shown in Fig. 3.3.

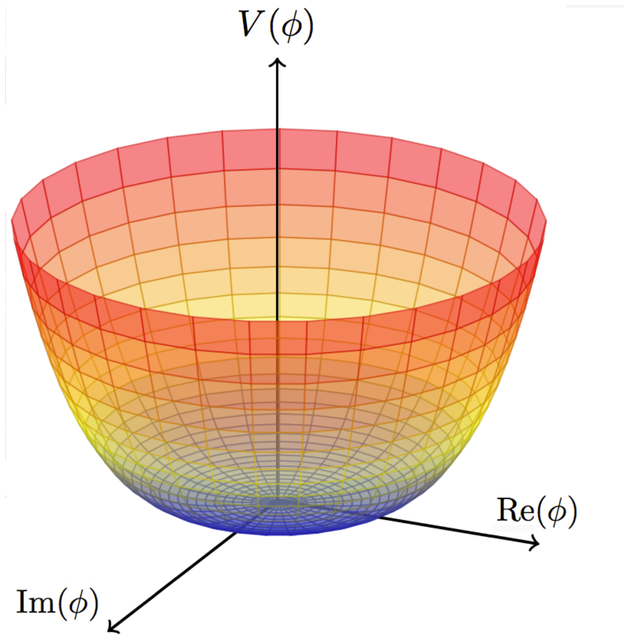


Figure 3.3: Vacuum potential with $\lambda \geq 0$, $\mu^2 > 0$.

This vacuum has a vanishing expectation value, where

$$\langle 0|\, \varphi \,|0\rangle = 0 \tag{3.39}$$

at $|\varphi| = 0$. This corresponds to the situation without a condensate where the manifest symmetry preserved.

Now the final case is $\lambda > 0, \mu^2 < 0$; the potential in this case is called a "Mexican hat" and is shown in Fig. 3.4.

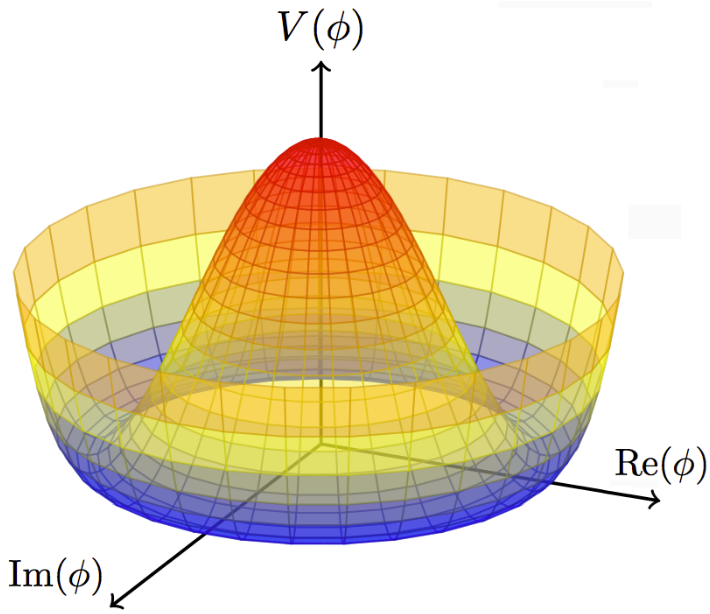


Figure 3.4: Vacuum potential with $\lambda > 0$, $\mu^2 < 0$.

The potential $V(\varphi)$ is then attractive at small values and repulsive at large values of φ. These features are exactly the behaviour necessary for a symmetry breaking condensate.

The vacuum expectation value is now non-vanishing, where

$$\langle 0| \varphi |0\rangle = \sqrt{\frac{-2\mu^2}{\lambda}} \tag{3.40}$$

with minimum at $\varphi_0 = e^{iq\theta}\sqrt{-2\mu^2/\lambda} \equiv v/\sqrt{2}$, where θ is the symmetry group generator of $U(1)$ and $q \in \mathbb{R}$ an arbitrary constant. The symmetry represents a redundancy in the theory and there is an infinite choice of equivalent ground states that lie on the circle of radius $v/\sqrt{2}$. Nature breaks this degeneracy by choosing one particular value of θ. Without loss of generality, we can choose the ground state with $q = 0$.

We now re-express the field, shifted around this minimum such that

$$\varphi = \frac{1}{\sqrt{2}}\left(\chi_1 + i\chi_2 + v\right) \tag{3.41}$$

where χ_1, χ_2 are real scalar fields that have zero vacuum expectation values. The two fields correspond to excitations in the radial and tangential directions to the circle of degenerate minima respectively.

Substituting this into the original Lagrangian (3.37), we find

$$\begin{aligned}\mathcal{L}(x) &= \frac{1}{2}\left[\partial_\mu\chi_1\partial^\mu\chi_1 + 2\mu^2\chi_1^2\right] + \frac{1}{2}\partial_\mu\chi_2\partial^\mu\chi_2 \\ &\quad -\frac{\lambda}{16}\left(\chi_1^2+\chi_2^2\right)\left(4v\chi_1+\chi_1^2+\chi_2^2\right) - \frac{1}{4}\mu^2 v^2\end{aligned} \tag{3.42}$$

We now have a theory of real massive scalar fields χ_1, a massless real scalar field χ_2 and six interaction terms.

Notice that the field χ_1 corresponding to oscillations in the asymmetric radial direction to the potential has become massive. On the other hand, the field χ_2, representing the oscillations in the symmetric tangential direction, remains massless. In terms of the potential, rotating along the circle of degenerate minima (that define equivalent vacuum states) costs zero energy.

This effect happens in the general case where a (continuous) symmetry is broken and is known as Goldstone's theorem. The massless fields describe the so-called Nambu-Goldstone bosons [61, 62].

Above the peak of the Mexican hat energy at $|\phi| = 0$ the potential looks symmetric. At lower energies below this condensation threshold, the particles will condense in the vacuum and roll down the potential, spontaneously breaking the symmetry of the theory as it chooses a particular vacuum.

In principle, the form of the symmetry breaking potential can be dynamically created by radiative corrections generated by quantum loops of gauge bosons and fermions [63]. However, the exact origin of the electroweak potential in the SM is still poorly understood and is a subject of ongoing research [8].

3.5.2 Gauge Symmetries

We shall extend the discussion of symmetry breaking to general non-abelian gauge theories. The n dimensional representation of the gauge group G is spanned by the n-component scalar fields H. There are N generators T_i of the symmetry, represented by $n \times n$ matrices, with one vector field $A_{i\mu}$ associated to each.

Here the Lagrangian is

$$\mathcal{L}(x) = (D_\mu H)^\dagger D^\mu H - V(H^\dagger H) - \frac{1}{4} F_{\mu\nu} F^{\mu\nu} \tag{3.43}$$

where the covariant derivative in a general representation is

$$D_\mu H = \left(\partial_\mu - igA_{j\mu} T^j\right) H \tag{3.44}$$

and the field strength is given by

$$F_i^{\mu\nu} = \partial^\mu A_i^\nu - \partial^\nu A_i^\mu - g f^{ijk} A_j^\mu A_k^\nu \tag{3.45}$$

for the "structure constants" $f^{ijk} = f_{ijk}$ (where $[T_i, T_j] = f^{ijk} T_k$) and some coupling constant g of the group.

$\mu, \nu = 0, 1, 2, 3$ are the space-time indices and $i, j, k = 1, \cdots, N$ are the group generator indices of $SU(N)$. The field strength term $F_{\mu\nu} F^{\mu\nu}$ is responsible for interactions of the gauge fields amongst themselves and is not concerned with the form of the scalar field's potential.

Abelian Symmetry Breaking

We shall first restrict our discussions to the abelian group $G = U(1)$, where $f^{ijk} = 0$, with only one generator. In this case, the scalar field transforms as $\phi \to e^{-ig\theta}\phi$, and the vector field as $A_\mu \to A_\mu + \partial_\mu \theta$, with $\theta = \theta(x)$, where we have again defined $A_\mu \equiv A_{1\mu} T^1$.

As in the case for the global symmetry breaking, the interesting situation has the potential where $\lambda > 0$, $\mu^2 < 0$ (corresponding to Fig. 3.4). The scalar fields can again be rewritten as (3.41) but the the gauge invariant kinetic term is now

$$\begin{aligned} \mathcal{L}(x) &= D_\mu H D^\mu H^\dagger \\ &= \frac{1}{2} \left(\partial_\mu \chi_1 - g A_\mu \chi_2\right)^2 + \frac{1}{2} (gv)^2 \left(A_\mu \frac{1}{gv} \partial_\mu \chi_2\right) \left(A^\mu + \frac{1}{gv} \partial^\mu \chi_2\right) \end{aligned} \tag{3.46}$$

The coupling terms on the right hand side indicate that χ_2 may be regarded as some shift of the field A_μ rather than an independent physical field.

In fact, we can use the gauge freedom to redefine the vector field as

$$A_\mu \longrightarrow A'_\mu = A_\mu + \frac{1}{gv}\partial_\mu \chi_2 \tag{3.47}$$

where the A'_μ is just the gauge transformation of A^μ with $\theta = \chi_2/gv$.

We can then redefine the scalar fields with the following transformations

$$\begin{aligned}
\chi_1 \longrightarrow \chi'_1 &= -v + \cos(g\theta)(v+\chi_1) + \sin(g\theta)\chi_2 \\
&\approxeq \chi_1 + g\theta\chi_2 \qquad (3.48a)\\
\chi_2 \longrightarrow \chi'_2 &= \cos(g\theta)\chi_2 - \sin(g\theta)(v+\chi_1) \\
&\approxeq \chi_2 - g\theta\chi_1 - g\theta v \qquad (3.48b)
\end{aligned}$$

where the second line in each transformation is valid for small g.

Finally, we use the gauge invariance to make the choice

$$\theta(x) = \frac{1}{g}\tan^{-1}\left(\frac{\chi_2}{v+\chi_1}\right) \tag{3.49}$$

In this gauge, there are only two remaining fields, χ'_1, which we shall relabel as h, and $A'_\mu = A_\mu + \partial_\mu\theta \approxeq A_\mu + 1/gv\partial_\mu\chi_2$, which we shall relabel A_μ.

The full Lagrangian then becomes

$$\begin{aligned}
\mathcal{L}(x) &= \frac{1}{2}\left[\partial_\mu h\partial^\mu h + 2\mu^2 h^2\right] - \frac{1}{4}F_{\mu\nu}F^{\mu\nu} + \frac{(gv)^2}{2}A_\mu A^\mu \\
&\quad + \frac{g^2}{2}A_\mu A^\mu h^2 + g^2 A_\mu A^\mu hv - \frac{\lambda}{16}h^4 - \frac{\lambda}{4}h^3 v - \frac{1}{4}\mu^2 v^2
\end{aligned} \tag{3.50}$$

We can now see that the scalar field h has a mass $m_h = \sqrt{-2\mu^2}$ and A_μ acquires the mass $m_A = gv$. Similar to the global symmetry breaking, the h field is massive, but now, due to the gauge freedom representing the local symmetry, the would-be Goldstone boson have been "eaten" by the originally massless gauge field A_μ that thus gain mass.

In the case of a 1-dimensional representation, the two transverse components of the gauge field and the two real parameters for the complex scalar field has become a single real spin-0 field h with one degree of freedom and the massive spin-1 vector field A_μ with three degrees of freedom, where A_μ acquires an additional (longitudinal) polarisation state seen in (3.47).

If this were a physical theory and the field A_μ was observed to be massless, then we could instead choose the gauge $\partial_\mu \theta = 0$, which removes both the longitudinal polarisation and the mass term.

Specifically for the local $U(1)$ case, SSB broke the symmetry of the only generator in the gauge group. In general, SSB need only break part of the group and can leave a subgroup of the symmetry generators unbroken. This is useful since the photon needs to be protected by a symmetry from gaining a mass.

Non-Abelian Symmetry Breaking

We shall now revert back to the full non-abelian Lagrangian, which by definition has non-vanishing commutation relations:

$$[T_i, T_j] = if^{ijk}T_k \tag{3.51}$$

with f^{ijk} the structure constants, which are defined by the particular group considered.

For the gauge group $G = SU(N)$ the arbitrary group element can be written as

$$U = \exp\left(-\frac{i}{v}\sum_{k=1}^{N}\theta_a(x)T^a\right) \in G \tag{3.52}$$

where θ_a are $N^2 - 1$ gauge fields associated to the T^a matrix generators of the group. If the symmetry breaking leaves the vacuum invariant under a subgroup S, such that the M generators of S are labelled $T_j \in S$ for $j = 1, \ldots, M$ and $M < N$, we have

$$T_j v \;=\; 0, \qquad j = 1, \ldots, M; \tag{3.53a}$$

$$T_k v \;\neq\; 0, \qquad k = M+1, \ldots, N \tag{3.53b}$$

where the generators T_k are the orthogonal complement of $T_j \in (G)$ (where $j = 1, \ldots, m$) and assuming the potential has a minimum at $|H|^2 = v^2/2$. Strictly speaking the generators belong to the Lie algebra ($\mathcal{L}(G)$) associated to the Lie group.

In this case, we can pick a gauge for U, such that the first M components vanish and we can write

$$U = \exp\left(-\frac{i}{v}\sum_{k=M+1}^{N} \theta_k T^k\right) \tag{3.54}$$

The the field H may now be parameterised as

$$H = \exp\left(\frac{i}{v}\sum_{i=M+1}^{N} \theta_a T^a\right)(v + h(x)) \tag{3.55}$$

such that the transformation rule is $H \to H' = v + h(x)$. By assumption, this general non-abelian group defines a symmetry of the Lagrangian. We therefore require the fields to transform as

$$H' = UH = v + h \tag{3.56a}$$

$$A'_\mu = UA_\mu U^\dagger + \frac{i}{g}(\partial_\mu U)U^\dagger \tag{3.56b}$$

$$F'_{\mu\nu} = UF_{\mu\nu}U^\dagger \tag{3.56c}$$

which leaves the total Lagrangian (3.43) invariant.

Similar to the abelian case (3.50), the gauge choice for U has left the Lagrangian independent of the gauge fields θ_k. We are left with

$$\begin{aligned}\mathcal{L}(x) \;=\;& \frac{1}{2}\left(\partial_\mu h^T \partial^\mu h - h^T \mathcal{M}_h^2 h\right) \\ & -\frac{1}{2}\partial_\mu A_{i\nu}\left(\partial^\mu A_i^\nu - \partial^\nu A_i^\mu\right) + \frac{1}{4}g^2 v^T A_\mu A^\mu v \\ & + ig\partial_\mu h^T A^\mu h + \frac{g^2}{\sqrt{2}}v^T A_\mu A^\mu h + \frac{1}{2}g^2 h^T A_\mu A^\mu h \\ & + \text{higher order terms}\end{aligned} \tag{3.57}$$

where the surviving field h has squared-mass determined by the matrix:

$$\left(\mathcal{M}_h^2\right)_{ab} = \left.\frac{\partial^2 V}{\partial H_a \partial H_b}\right|_{H=v} \tag{3.58}$$

The vector gauge bosons eat the would-be Goldstone bosons and have a squared-mass matrix

$$\left(\mathcal{M}_A^2\right)_{lk} = g^2\left(v^T T_l T_k v\right) \tag{3.59}$$

which is real, symmetric, and positive definite. It follows immediately from (3.53) that the gauge fields corresponding to oscillations in the directions of symmetry A_k^μ for $k = 1, \ldots, M$ are massless whereas the fields $k = M+1, \ldots, N$ are massive. Notice again that the massive components also acquire the additional longitudinal component from (3.56b).

3.6 Electroweak Spontaneous Symmetry Breaking

In nature the fermions are observed to be massive particles [4]. Therefore, we need to introduce the explicit mechanism to break the underlying electroweak symmetry $SU(2)_L \times U(1)_Y$. For the quantisation of light, the photon, to remain massless, the unbroken $U(1)_{\text{EM}}$ symmetry of electromagnetism (with the symmetry generator Q, representing the charge) is needed.

The simplest way to break the electroweak Lagrangian is to introduce a complex scalar doublet, canonically denoted H, known as the Higgs field. Since the (Yukawa) terms that describe the coupling of the Higgs field with the leptons must be singlets of both the Lorentz and gauge groups, the scalar is chosen to live in the **2** representation of $SU(2)_L$, with hypercharge $Y = 1$ of $U(1)_Y$.

We shall prepend a superscript g to denote the Lagrangian before electroweak spontaneous symmetry breaking (EWSSB) and a superscript m for the Lagrangian afterwards, describing the physical mass basis that we observe.

Before EWSSB, we start with the canonical Lagrangian for the scalar field

$$^g\mathcal{L}_{\text{Higgs}}(x) = -\frac{1}{2}\left(D_\mu H_i\right)^\dagger D_\mu H_i - V(H^\dagger H) \tag{3.60}$$

where $V(H^\dagger H) = \frac{1}{2}\mu^2 H^\dagger H + \frac{1}{4}\lambda\left(H^\dagger H\right)^2$ and $\lambda > 0, \mu^2 < 0$. The covariant derivative for the electroweak gauge group is given by

$$D_\mu H = \left(\partial_\mu - igA_{j\mu}T^j - ig' B_\mu Y\right) H \tag{3.61}$$

where T_j are the generators of the $SU(2)_L$ gauge group with coupling strength g and Y are the generators of $U(1)_Y$ with coupling g'.

We choose this vacuum to have the expectation value

$$\langle 0| H |0\rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v \end{pmatrix} \tag{3.62}$$

with the minimum at

$$H_0 = \frac{1}{\sqrt{2}} e^{i\left(A_1 T^1 + A_2 T^2 + (A_3 T^3 - \frac{1}{2}BY)\right)} \begin{pmatrix} 0 \\ v \end{pmatrix} \tag{3.63}$$

which still appears symmetric under the full electroweak gauge group $SU(2)_L \times U(1)_Y$.

To match experimental observation, we require that the vacuum is annihilated by three of the generators, leaving one unbroken symmetry direction. Explicitly, we can express the generators as

$$T^{i\prime} \begin{pmatrix} 0 \\ v \end{pmatrix} \neq 0 \quad , \quad Q \begin{pmatrix} 0 \\ v \end{pmatrix} = 0 \tag{3.64}$$

where we have defined

$$T^{1\prime} \equiv T^1 \tag{3.65a}$$

$$T^{2\prime} \equiv T^2 \tag{3.65b}$$

$$T^{3\prime} \equiv T^3 - \frac{1}{2}Y \tag{3.65c}$$

corresponding to the three broken directions. The final unbroken generator is

$$Q \equiv T_3 + \frac{1}{2}Y \tag{3.66}$$

and corresponds to the remaining EM charge symmetry. Our vacuum is therefore chargeless in terms of the electromagnetic charge Q. This is the simplest choice we can make to give the correct observed particle spectra [64]. Further (3.66) implies that the Higgs field, written as an electroweak doublet, has both charged and neutral components:

$$H \equiv \begin{pmatrix} \phi^+ \\ \phi^0 \end{pmatrix} \tag{3.67}$$

We can now write the field in terms of its oscillations around the minimum:

$$H(x) = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v + h(x) \end{pmatrix} \tag{3.68}$$

where $v \in \mathbb{R}$ and $h(x)$ is a real scalar field with one degree of freedom. Experimentally, this vacuum expectation value has been set at $v = 246$ GeV through indirect measurements of the muon life-time [65].

Notice that this choice (3.68) has "gauged away" three Goldstone bosons. Explicitly, we can write the full invariant $SU(2)_L \times U(1)_Y$ Higgs field as

$$H(x) = \frac{1}{\sqrt{2}} e^{i\theta\cdot\tau} \begin{pmatrix} 0 \\ v + h(x) \end{pmatrix} \tag{3.69}$$

where $\tau_i = \frac{\sigma_i}{2}$ are the generators of the $SU(2)$ group, which correspond to the broken directions of symmetry. The gauge fields θ_i associated to these generators are gauged away in (3.68). For small variations, we see

$$H(x) \approxeq \frac{1}{\sqrt{2}} \begin{pmatrix} \theta_2 + i\theta_1 \\ v + h(x) - i\theta_3 \end{pmatrix} \tag{3.70}$$

and so our four fields θ_i (i=1,2,3,) and h are indeed independent and fully parameterise the deviations from the vacuum H_0. Since the Lagrangian is locally $SU(2)_L$ invariant, we can gauge away the three gauge fields $\theta(x)$ to get the result claimed in (3.68). The three fields $\theta(x)$ would have given rise to three Goldstone bosons, each with one degree of freedom, but in this gauge they are unphysical and will be eaten by the electroweak bosons of the theory, which thus gain mass. Specifically, the two charged components of the Higgs gives will give rise to the $W^{\pm}$ bosons and the neutral degree of freedom in the Higgs field will be eaten by the Z^0.

To simplify our equations, we now define the Weinberg angle θ_W that relates the coupling g' and g:

$$\theta_W \equiv \tan^{-1}\left(\frac{g'}{g}\right) \tag{3.71}$$

This relationship between the weak and electromagnetic forces unifies the two theories. There is two things to notice: first, since this angle is determined from the two coupling constants, it is not a free parameter of the SM and second, since $g \neq g'$ the two theories are not (strictly speaking) fully unified.

We can then re-express the neutral gauge fields in convenient linear combinations called the mass eigenbasis where

$$\begin{pmatrix} Z_\mu \\ A_\mu \end{pmatrix} = \underbrace{\begin{pmatrix} \cos\theta_W & -\sin\theta_W \\ \sin\theta_W & \cos\theta_W \end{pmatrix}}_{R(\theta_W)} \begin{pmatrix} A^3_\mu \\ B_\mu \end{pmatrix} \tag{3.72}$$

where $R(\theta_W)$ denotes a 2-dimensional rotation with an angle θ_W. Similarly for the charged gauge fields as

$$\begin{pmatrix} W^+ \\ W^- \end{pmatrix} = \begin{pmatrix} \cos\theta_W & -\sin\theta_W \\ \sin\theta_W & \cos\theta_W \end{pmatrix} \begin{pmatrix} A^1 \\ A^2 \end{pmatrix} \tag{3.73}$$

This basis diagonalises the mass matrix and defines the physical eigenstates that are observed in experiments.

Explicitly, these fields can be written as

$$W^{\pm}_\mu = \frac{1}{\sqrt{2}}\left(A^1_\mu \mp i A^2_\mu\right) \tag{3.74a}$$

$$Z_\mu = \cos\theta_W A^3_\mu - \sin\theta_W \theta B_\mu \tag{3.74b}$$

$$A_\mu = \sin\theta_W A^3_\mu + \cos\theta_W B_\mu \tag{3.74c}$$

with inverses

$$
\begin{aligned}
A^1_\mu &= \frac{1}{\sqrt{2}}\left(W^+_\mu + W^-_\mu\right) && \text{(3.75a)}\\
A^2_\mu &= \frac{i}{\sqrt{2}}\left(W^+_\mu - W^-_\mu\right) && \text{(3.75b)}\\
A^3_\mu &= \cos\theta_W Z_\mu + \sin\theta_W A_\mu && \text{(3.75c)}\\
B_\mu &= -\sin\theta_W Z + \cos\theta_W A_\mu && \text{(3.75d)}
\end{aligned}
$$

Notice that $W^\pm$ has the form of raising and lowering operators and are just the linear combinations of generators in $SU(2)$. They are transformations between states in an irreducible representation and will change the charge of the system as well as transferring energy and momentum. On the other hand, the Z_μ and A_μ are "Cartan" generators, whose eigenvalues define the weights of the algebra [5]. These fields transfer energy and momentum but do not change the charge of the system. In terms of particle interactions $W^\pm$ carry ± 1 EM charge whereas Z^0 and the photon γ are electromagnetically neutral.

The covariant derivative for the $SU(2)_L \times U(1)_Y$ can now be written

$$
D_\mu H = \frac{1}{\sqrt{2}}\left(\partial_\mu h + i\frac{g}{2\cos\theta_W}(v+h)Z_\mu\right)\begin{pmatrix}0\\1\end{pmatrix} - i\frac{1}{2}g(v+h)W_\mu\begin{pmatrix}1\\0\end{pmatrix} \quad (3.76)
$$

where we have used $T_i = \frac{\sigma_i}{2}$ and $Y = \frac{1}{2}\mathbb{I}_{2\times 2}$ seen in (3.25).

Upon expansion around the minima, the quadratic term in the covariant derivative ($D_\mu H^\dagger D^\mu H$) that couples H to the gauge fields becomes

$$
{}^m\mathcal{L}_{H,\text{gauge}}(x) = \frac{1}{2}\partial^\mu h\partial_\mu h + \frac{1}{4}g^2\left(v+h\right)^2\left(\frac{1}{2\cos\theta_W^2}Z^\mu Z_\mu + W^{\mu\dagger}W_\mu\right) - \frac{1}{8}\lambda\left(h^2+2\nu h\right)^2 \quad (3.77)
$$

We can now read off the masses of the fields as the coefficients of the quadratic terms:

$$\begin{aligned} m_W^2 &= \frac{1}{4} g'^2 v^2 && (3.78a) \\ m_Z^2 &= \frac{1}{4}\left(g'^2 + g^2\right) v^2 = \frac{m_W^2}{\cos^2 \theta_W} && (3.78b) \\ m_A^2 &= 0 && (3.78c) \end{aligned}$$

The masses of the weak gauge bosons provide a non-trivial prediction for our formulation where

$$\rho \equiv \frac{m_W^2}{m_Z^2 \cos^2 \theta_W} \tag{3.79}$$

is expected to be equal to unity for the SM. This identity highly constrains the form that any BSM theory can have.

The $W^{\pm}$ and Z^0 bosons mediate the electroweak interaction and have a short range, as they have relatively large masses. These masses are not predicted by theory as the couplings g' and g are not constrained. In 1983, the weak bosons were discovered by UA1 and UA2 collaborations [47, 48], with mass:

$$\begin{aligned} m_W &= 80.385 \pm 0.015 \ \text{ GeV} && (3.80a) \\ m_Z &= 91.876 \pm 0.0021 \text{ GeV} && (3.80b) \end{aligned}$$

The final field A corresponds to the single unbroken generator Q of the electromagnetic symmetry $U(1)_{\text{EM}}$. The photon of QED represents the quantisation of this field, which is massless by construction $m_A = 0$ to be consistent with experimental measurements [4].

Currently, the theory predicts that A_μ ($\mu = 1, 2, 3, 4$) has four components but experimentally we only measure 2 degrees of freedom, one for longitudinal polarisation and another for transverse polarisation. There is a leftover gauge symmetry which we can remove by again fixing a gauge. Common choices include the Lorentz gauge ($\partial^\mu A_\mu = 0$ with $\Box A^\mu = 0$) and the Coulomb gauge ($\nabla \cdot \boldsymbol{A} = 0$ with $A^0 = 0$) [7]. Similar constraints are imposed to remove one degree of freedom from the massive gauge bosons that are spin-1 fields with three degrees of freedom [64].

After symmetry breaking, the potential can be rewritten as

$$V(H^\dagger H) = \frac{1}{2}m_h^2 h^2 + \frac{1}{4}\lambda v h^3 + \frac{1}{16}h^4 \tag{3.81}$$

where the last two terms on the right-hand-side are the cubic and quartic self-couplings of the Higgs boson.

The Higgs boson thus has squared-mass

$$m_h^2 = \frac{1}{2}\lambda v^2 \tag{3.82}$$

for $v = \sqrt{2\mu^2}/\lambda$.

The kinetic terms of the gauge fields before EWSSB are

$$^g\mathcal{L}_{\text{gauge}}(x) = F^{\mu\nu,i}F_{i,\mu\nu} + B^{\mu\nu}B_{\mu\nu} \tag{3.83}$$

where $i = 1, 2, 3$ labels the generators of $SU(2)_L$. We can rewrite this in the mass basis by defining

$$F_{\mu\nu}^W \equiv \frac{1}{\sqrt{2}}\left(F_{\mu\nu 1} - iF_{\mu\nu 2}\right) = d_\mu W_\nu - d_\nu W_\mu \tag{3.84a}$$

$$F_{\mu\nu}^Z \equiv \partial_\mu Z_\nu - \partial_\nu Z_\mu \tag{3.84b}$$

$$F_{\mu\nu}^A \equiv \partial_\mu A_\nu - \partial_\nu A_\mu \tag{3.84c}$$

where $d_\mu \equiv \partial_\mu - igA_{\mu 3} = \partial_\mu - ieA_\mu - ig\cos\theta_W Z_\mu$, and $e \equiv g\sin\theta_W$.

With these definitions, the kinetic terms of the gauge fields after EWSSB become

$$\begin{aligned} ^m\mathcal{L}_{\text{gauge}}(x) &= -\frac{1}{2}F^{W,\mu\nu\dagger}F_{\mu\nu}^W - \frac{1}{4}F^{Z\mu\nu}F_{\mu\nu}^Z - \frac{1}{4}F^{A\mu\nu}F_{\mu\nu}^A \\ &\quad + iW^\mu W^{\nu\dagger}\left(eF_{\mu\nu}^A + g\cos\theta_W F_{\mu\nu}^Z\right) \\ &\quad + ig^2\left(W^2 W^{\dagger 2} - \left(W\cdot W^\dagger\right)^2\right) \end{aligned} \tag{3.85}$$

In summary, we started with the Higgs field and four identical massless gauge fields $(A_\mu^1, A_\mu^2, A_\mu^3, B_\mu)$ and an exact unbroken electroweak symmetry $SU(2)_L \times U(1)_Y$ at high energies. At low energies, the non-symmetric Higgs vacuum breaks the manifest symmetry leaving a residual $U(1)_{\text{EM}}$ unbroken. We are left with the three massive gauge bosons $W^{\pm}$ and Z that mediate the weak force and the one massless boson γ corresponding to the photon of electromagnetism.

Counting the degrees of freedom: before EWSSB there are four for the Higgs complex scalar doublet, and two each for the four massless gauge fields for a total of twelve; after EWSSB, there is one for the Higgs boson, three each for the three massive vector bosons, which acquire an additional polarisation as in (3.47), and two for the photon that remains massless (again giving a total of twelve).

The gauge and Higgs bosons of electroweak theory have all been observed and are tabulated in Table 3.5 below.

Boson	EM Charge	Mass	Date of Discovery
photon (γ)	0	0	1900 (Paul Villard [66])
$W^{\pm}$	± 1	80.385 ± 0.015 GeV	1983 (UA1 and UA2 [47, 48])
Z^0	0	91.1876 ± 0.0021 GeV	1983 (UA1 and UA2 [45, 46])
Higgs (h)	0	125.3 ± 0.6 GeV	2012* (CMS and ATLAS [43, 67])

Table 3.5: The experimentally measured bosons and the dates of their discovery. The electromagnetic charge is normalised to that of the electron, the photon, the Z^0, and the Higgs boson have charge 0, whilst $W^{\pm}$ has charge ± 1. * the Higgs boson has only been tentatively discovered and needs verification.

3.7 The Fermion Sector

3.7.1 The Leptons

We shall now describe the leptons in the electroweak sector with gauge group $SU(2)_L \times U(1)_Y$ and do not interact via the strong force. The leptons are spin-1/2 fermions (represented by spinor fields) and are singlets of $SU(3)_C$.

It was experimentally observed in 1957 by Chien-Shiun Wu and collaborators that parity is not a symmetry of nature [68]. In particular, the weak force is non-chiral and distinguishes between left and right-handed particles.

There are currently six known flavours of left-handed leptons that can be arranged into three generations organised into $SU(2)_L$ doublets **2**. The three generations behave identically, except that the particles in the higher generations are more massive. The left-handed neutral neutrinos have $T_3 = +\frac{1}{2}$ and the left-handed charged leptons have $T_3 = -\frac{1}{2}$, both with hypercharge $Y = -1$.

The three left-handed families of leptons can be written as,

$$L_1 = \begin{pmatrix} \nu_e \\ e_L \end{pmatrix}_L \quad , \quad L_2 = \begin{pmatrix} \nu_\mu \\ \mu_L \end{pmatrix}_L \quad , \quad L_3 = \begin{pmatrix} \nu_\tau \\ \tau_L \end{pmatrix}_L \tag{3.86}$$

where ν_L and e_L, μ_L, τ_L are the left-handed neutrino and left-handed charged leptons respectively.

The right-handed leptonic fields consist *only* of the three charged leptons, each placed into a singlet **1** of $SU(2)_L$; these have $T_3 = 0$ with weak hypercharge $Y = 2$. There are no right-handed neutrinos in the SM, which requires their (Dirac) mass to be identically zero.

The three right-handed familes of leptons are

$$R_1 = e_R \quad , \quad R_2 = \mu_R \quad , \quad R_3 = \tau_R \tag{3.87}$$

with e_R, μ_R, τ_R the right-handed charged leptons.

This framework is consistent with definition of Q in (3.66), where the left-handed electron has charge -1, the neutrino 0 and the right-handed electron has charge $+1$.

Leptonic Currents

Since the leptons are fermionic fields, the $SU(2)_L \times U(1)_Y$ gauge invariant kinetic part of the leptonic Lagrangian can be written as

$$^{g}\mathcal{L}_{\text{lepton,kin}}(x) = Li\not{D}_L L + \overline{R}i\not{D}_R R \tag{3.88}$$

Explicitly, the covariant derivative can be written as:

$$
\begin{aligned}
(D_{\mu,L}L)_i &= \partial_\mu L_i - igA^a_\mu (T^a)_{ij} L_j - ig' B_\mu Y_1 L_i \qquad \text{LH lepton fields} && (3.89a)\\
(D_{\mu,R}\overline{R}) &= \partial_\mu \overline{R} - ig' B_\mu Y_{\overline{R}} \overline{R} \qquad \text{RH lepton fields} && (3.89b)
\end{aligned}
$$

where the field A^a_μ is associated to the generators T^a of $SU(2)_L$ and B_μ to the generator Y of $U(1)_Y$. In this way, the covariant derivative couples the leptons to the gauge fields that describe the bosons of the theory.

By Noether's theorem, one can compute the currents related to the electroweak symmetries of the Lagrangian. Since the symmetry transformations are given by the generators of the group (1.33), the associated gauge fields will naturally couple to the Noether currents.

We shall first define

$$T = T_L \frac{1}{2}(1-\gamma_5) \tag{3.90}$$

where recall from (2.37), $\frac{1}{2}(1-\gamma_5)$ is the projection operator onto the left-handed component and T_L represents only the left-handed component of the $SU(2)_L$ generators. In the fundamental representation, the left-handed generators of $SU(2)_L$ are $T_{Li} = \frac{1}{2}\sigma_i$, where σ_i (i=1,2,3) are the usual Pauli matrices and the right-handed generators are $T_{Ri} = 0$.

Similarly,

$$
\begin{aligned}
Y &= 2Q - T_3 = Y_L \frac{1}{2}(1-\gamma_5) + Y_R \frac{1}{2}(1+\gamma_5) && (3.91a)\\
Y_L &= 2Q - T_{L3} && (3.91b)\\
Y_R &= 2Q && (3.91c)
\end{aligned}
$$

where Q is the generator for the residual $U(1)_{\text{EM}}$ symmetry. This symmetry vector-like (independent of handedness), indicating that the electromagnetic force, unlike the weak force, acts on left and right-handed particles in the same way.

The electromagnetic current can be found by Noether's theory from the symmetry of the A_μ field to be:

$$\begin{aligned} j^\mu_{EM} &= \overline{L}\gamma^\mu \left(T_{L3} - \frac{1}{2}\right) L - \overline{R}\gamma^\mu R \\ &= -\overline{e}\gamma^\mu e \end{aligned} \tag{3.92}$$

where we have defined the field e to be

$$e \equiv \begin{pmatrix} e_L \\ e_R \end{pmatrix} \tag{3.93}$$

and $\overline{e} = e^\dagger\gamma^0$ the Dirac conjugate field.

In the electroweak theory, the electromagnetic current couples the photon field (A_μ) to the leptons with coupling strength $g\sin\theta_W$. This coupling arises as a prefactor of the term j^μ_{EM} in the derivation of the current and is simply pulled out.

The neutral current that couples to Z_μ is similarly

$$\begin{aligned} J^\mu_n &= \overline{L}\gamma^\mu \left(\cos^2\theta_W\sigma_3 + \sin^2\theta_W\right) L + 2\sin^2\theta_W \overline{R}\gamma^\mu R \\ &= \frac{1}{2}\left(\overline{\nu}_e\gamma^\mu\left(1-\gamma_5\right)\nu_e - \overline{e}\gamma^\mu\left(1-\gamma_5 - 4\sin^2\theta_W\right)e\right) \end{aligned} \tag{3.94}$$

This introduces the interaction between Z_μ and both the neutrinos ($\overline{\nu}_e\nu_e$) and charged leptons (e^+e^-) of all three generations. The coupling strength of these processes is $\frac{g}{2}\cos\theta_W$.

Finally, the charged current from the symmetries of the $W^\pm$ field is

$$J^\mu = \overline{L}\gamma^\mu T_{L+} L = \overline{e}\gamma^\mu\left(1-\gamma_5\right)\nu_e \tag{3.95}$$

where we have defined

$$T_{L+} = \frac{1}{2}\left(T_{L1} + iT_{L2}\right) = \frac{1}{2}\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \tag{3.96}$$

This J^μ current couples to the W_μ fields and has charge $\Delta Q = +1$ with effective coupling $\frac{g}{2\sqrt{2}}$.

The kinetic term (3.88) introduces interactions between the leptons and the the observed gauge bosons and can now be rewritten with these identifications:

$$ {}^{m}\mathcal{L}_{\text{lepton,kin}}(x) = i\overline{L}\partial\!\!\!/ L + i\overline{R}\partial\!\!\!/ R + \frac{g}{2\sqrt{2}}\left(J^\mu W_\mu + J^{\mu\dagger} W_\mu^\dagger\right) + e j^\mu_{EM} A_\mu + \frac{g}{2\cos\theta_W} J^\mu_n Z_\mu \tag{3.97} $$

where we have defined the coupling of EM coupling $e \equiv g\sin\theta_W$.

Notice that had we included the right-handed neutrino field in $\nu \equiv \begin{pmatrix} \nu_L \\ \nu_R \end{pmatrix}$, it would have been projected out by $\frac{1}{2}\left(1-\gamma^5\right)$ from (2.37), in both the neutral (3.94) and charged (3.95) weak currents. Since the neutrino has only ever been observed to interact via the weak force in the SM, there is no reason for the right-handed neutrino to exist at all. Notice that since, Lorentz boosts can transform left-handed particles to right-handed particles only if they have mass (and vice-versa), these particles are assumed to be massless [64].

Lepton Masses

Since, the Higgs field H lives in the $T = \frac{1}{2}$ doublet with hypercharge $Y = 1$, we can construct a invariant singlet of $SU(2)_T \times U(1)_Y$ that couples the leptons to the Higgs boson H:

$$ {}^{g}\mathcal{L}_{H,\text{lepton}}(x) = -\sqrt{2}G_e\left(\overline{L}(x)H(x)R(x) + \overline{R}(x)H(x)^\dagger L(x)\right) \tag{3.98} $$

with coupling G_e, where there is a copy of these terms for each generation. We shall again use the notation of superscripts (g) and (m) to indicate the gauge and mass eigenbases corresponding to before and after EWSSB.

By expanding around the Higgs minimum, we find

$$ \begin{aligned} {}^{m}\mathcal{L}_{H,\text{lepton}}(x) &= -G_e\left(v+h(x)\right)\left(\overline{e}_L e_R + \overline{e}_R e_L\right) \\ &= -G_e\left(v+h(x)\right)\overline{e}e \end{aligned} \tag{3.99} $$

The mass of the charged leptons is thus

$$m_e = 2G_e v \tag{3.100}$$

with G_e distinct for each generation. This set of three coupling constants that define the leptons masses are free parameters in the SM and need to be experimentally determined.

Explicitly for three observed lepton generations, we can generalise the Lagrangian in (3.98) to

$$^g\mathcal{L}_{H,\text{lepton}}(x) = -\sqrt{2}\left(\overline{L}_N f_{NM} H R_{NM} + \overline{R}_M f^{MN} H L_N\right) \tag{3.101}$$

where the indices $M, N = 1, 2, 3$ run over the three fermion generations and f_{NM} is an unspecified 3×3 matrix that can be diagonalised with $f \to V^\dagger f U$, for some 3-dimensional unitary transformations $U, V \in U(3)$.

In fact, since the right and left-handed fields transform under different representations, we have enough freedom to choose U and V separately such that $R \to U^\dagger R$ and $L \to V^\dagger L$. This rotates the fields into the mass basis without requiring any off-diagonal couplings that would mix the lepton generations (although quantum affects mean that there are in fact highly suppressed lepton generation violating effects).

The dates that the lepton were discovered and their masses are tabulated in Table 3.6 below.

Lepton	EM Charge	Mass	Date of Discovery
electron (e^-)	-1	0.511 MeV	1897 (J. J. Thompson [69])
electron neutrino (ν_e)	0	~ 0	1956 (C. Cowan and F. Reines [70])
muon (μ^-)	-1	105.658 MeV	1936 (C. Anderson [71])
muon neutrino (ν_μ)	0	~ 0	1962 (L. Lederman et al [72])
tau (τ^-)	-1	1776.82 ± 0.16 MeV	1975 (M. L. Perl et al [73])
tau neutrino (ν_τ)	0	~ 0	2000 (DONUT collaboration [74])

Table 3.6: The experimentally measured leptons and the dates of their discovery [4]. The charged leptons e^-, μ^-, τ^- are all negatively charged under the unbroken electromagnetic symmetry with charge -1 normalised to the magnitude of the electron's charge. The uncertainty in the electron and muon masses are insignificant. All these leptons all have oppositely charged antiparticles.

From the original electroweak Lagrangian defined by (3.88) and (3.98), we have

$$^{g}\mathcal{L}_{\text{EWK,lepton}} = {}^{g}\mathcal{L}_{H,\text{lepton}} + {}^{g}\mathcal{L}_{\text{lepton,kin}} \tag{3.102}$$

After EWSSB, the final electroweak Lagrangian for the leptonic sector is given

$$^{m}\mathcal{L}_{\text{EWK,lepton}} = {}^{m}\mathcal{L}_{H,\text{lepton}} + {}^{m}\mathcal{L}_{\text{lepton,kin}} \tag{3.103}$$

with definitions found in (3.97) and (3.99).

3.7.2 The Quarks

The electroweak theory of the quarks will be very similar to that of the leptons. There are also three known families of quarks. The left-handed quark fields form doublets of $SU(2)_L$, with up-type quarks having $T_3 = +\frac{1}{2}$, and down-type quarks $T_3 = -\frac{1}{2}$, both carrying hypercharge $Y = 1/3$.

In contrast to the leptons, all right-handed quark states exist. The right-handed fields live in a singlet of $SU(2)_L$ and have $Y = -4/3$ and $Y = 2/3$ for the up-type and down-type quarks respectively.

Explicitly, the left-handed fields can be written as

$$\psi_{L,1} = \begin{pmatrix} u^g \\ d^g \end{pmatrix}_L \quad , \quad \psi_{L,2} = \begin{pmatrix} c^g \\ s^g \end{pmatrix}_L \quad , \quad \psi_{L,3} = \begin{pmatrix} t^g \\ b^g \end{pmatrix}_L \tag{3.104}$$

where the superscript g shall be used to indicate the gauge eigenbasis.

The right-handed fields are

$$\begin{aligned} \psi_{R,1+} &= u^g_R \quad , \quad \psi_{R,2+} = c^g_R \quad , \quad \psi_{R,3+} = t^g_R \\ \psi_{R,1-} &= d^g_R \quad , \quad \psi_{R,2-} = s^g_R \quad , \quad \psi_{R,3-} = b^g_R \end{aligned} \tag{3.105}$$

To obtain experimentally consistent electric charges of the quarks, we require $Q(u^g_L) = 2/3$ for the left-handed up-type, $Q(d^g_L) = -1/3$ for the left-handed down-type quarks. The right-handed quarks are oppositely charged with $Q(u_R^g) = -2/3$, and $Q(d_R^g) = 1/3$ for the up and down-types respectively.

Quark Currents

The gauge kinetic term for the quark fields is

$$^g\mathcal{L}_{\text{quark,kin}}(x) = \overline{\psi}_{L,N} i \not{D} \psi_{L,N} + \overline{\psi}_{R,N+} i \not{D} \psi_{R,N+} + \overline{\psi}_{R,N-} i \not{D} \psi_{R,N-} \tag{3.106}$$

with D_μ the $SU(2)_L \times U(1)_Y$ fundamental covariant derivative from (3.24), where $\not{D} \equiv \gamma^\mu D_\mu$.

As with the leptonic case, the kinetic part can be re-expressed in terms of the conserved currents, explicitly introducing the measured interactions of the gauge bosons with the quarks.

For conciseness, we shall consider the fermion field ψ, representing both left-handed L and right-handed R quarks, such that (3.106) can be rewritten as

$$^g\mathcal{L}_{\text{quark,kin}} = \overline{\psi} i \gamma^\mu D_\mu \psi \tag{3.107}$$

As in the leptonic case, we can rotate the gauge fields to the mass basis to find

$$^m\mathcal{L}_{\text{quark,kin}} = \overline{\psi} i \gamma^\mu \partial_\mu \psi + \frac{g}{2\sqrt{2}} \left(J^\mu W_\mu + J^{\mu\dagger} W_\mu^\dagger \right) + e j_{EM}^\mu A_\mu + \frac{g}{2\cos\theta_W} J_n^\mu Z_\mu \tag{3.108}$$

where the currents are defined as

$$\text{EM} \quad j_{EM}^\mu = 2\overline{\psi}\gamma^\mu Q \psi \tag{3.109a}$$

$$\text{neutral} \quad J_n^\mu = 2J_3^\mu - 2\sin^2\theta_W j_{EM}^\mu = \overline{\psi}\gamma^\mu \left((1-\gamma_5) T_{L3} - 4\sin^2\theta_W Q \right) \psi \tag{3.109b}$$

$$\text{charged} \quad J^\mu = \overline{\psi}\gamma^\mu (1-\gamma_5) T_{L+} \psi \tag{3.109c}$$

with $T_{L+} = T_{L_1} + iT_{L_2}$. As in the case of the leptons, the generator Q does not depend on chirality and the electromagnetic current j_{EM}^μ is vector-like. The neutral and charged currents on the other hand, which couple to the weak gauge bosons ($W^{\pm}$ and Z^0), are seen here to be chiral.

Quark Masses

The generations of the up-type quarks of charge 2/3 quarks u, c, t, and down-type quarks of charge $-1/3$ quarks d, s, b are distinguished by mass. If all families had the same mass, the quarks would be physically indistinguishable from each other and the mixing angle between the generations would be irrelevant.

This mixing angle arises from the desire to write the masses in a similar manner to the leptons with

$$\mathcal{L}_{\text{quark,mass}} = -\sum_q m_q \bar{q} q \tag{3.110}$$

where m_q correspond to the different quark masses. However, just as for leptons, there are no $SU(2)_L \times U(1)_Y$ invariant mass terms; we therefore need to introduce them via the Higgs mechanism.

Historically, the u, c, d quarks were observed and the s quark was predicted in order to explain the lack of flavour changing neutral currents (FCNC) in the SM. As we shall see, this ensures that the mixing angle between the generations can be diagonalised when coupled to the neutral Z boson in the mass basis.

As with leptons, the quark masses arise from the coupling to the $T = \frac{1}{2}, Y = 1$ scalar H Higgs field. Recall that, we can break the symmetry of the theory by choosing a vacuum for the Higgs to be

$$H_0 = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v \end{pmatrix} \tag{3.111}$$

One can in general also have the conjugate field H^c that can be defined as follows

$$H^c \equiv i\sigma_2 H^\star = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \begin{pmatrix} H_1^\dagger \\ H_2^\dagger \end{pmatrix} = \begin{pmatrix} H_2^\dagger \\ -H_1^\dagger \end{pmatrix} \tag{3.112}$$

Then, $YH^c = +H^c$, where H^c also lives in an equivalent $T = \frac{1}{2}$ representation of $SU(2)_L$. This follows since

$$i\sigma_2\sigma^\star = \sigma \quad , \quad i\sigma_2\sigma^\star = -\sigma i\sigma_2 \tag{3.113}$$

where $\sigma^\star$ is the complex conjugate of any of the Pauli matrices. This in turn implies that

$$i\sigma_2 \left(e^{\frac{i}{2}\alpha\cdot\sigma}\right)^\star = e^{\frac{i}{2}\alpha\cdot\sigma} i\sigma_2 \tag{3.114}$$

Thus, H^c transforms exactly as H under both $SU(2)_L$ and $U(1)_Y$ and has the vacuum expectation value

$$\langle 0| H^c |0\rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} v \\ 0 \end{pmatrix} \tag{3.115}$$

We shall denote the quark fields q_- to be down-type quarks with charge $-\frac{1}{3}$ and q_+ up-type quarks with charge $+\frac{2}{3}$. In this notation, the masses of the quarks arise from the coupling to the Higgs field in

$${}^g\mathcal{L}_{H,\text{quark}} = -\sqrt{2}\left(\overline{\psi}_L \Gamma_- q_{-R} H + \overline{\psi}_L \Gamma_+ q_{+R} H^C + h.c.\right) \tag{3.116}$$

where $\Gamma_\pm$ are some arbitrary complex matrices that act on $q_\pm$ to diagonalise the mass matrix. Notice that without explicit mass terms, the basis we choose would have been arbitrary.

Keeping just the mass terms after EWSSB, we find

$${}^m\mathcal{L}_{\text{quark,mass}} = -\left(\overline{\psi}_L m_- q_{-R} + \overline{\psi}_L m_+ q_{+R} + h.c.\right) \tag{3.117}$$

where we have taken $m_\mp = \Gamma_\mp v$.

The other terms from (3.116) couple the Higgs boson h with the quarks in

$${}^m\mathcal{L}_{\text{quark,int}} = -\left(\overline{\psi}_L \Gamma_- q_{-R} h + \overline{\psi}_L \Gamma_+ q_{+R} h + h.c.\right) \tag{3.118}$$

where ${}^{m}\mathcal{L}_{H,\text{quark}} \equiv {}^{m}\mathcal{L}_{\text{quark,mass}} + {}^{m}\mathcal{L}_{\text{quark,int}}$. These terms shall be dropped for this discussion.

The physical mass matrix arises by diagonalising the Lagrangian, such that the "mass matrix" $m_-^\dagger m_-$ gives the masses for d, s, b and $m_+^\dagger m_+$ the masses for u, c, t. Recall the basic property that any Hermition matrix $m^\dagger m$ can always be diagonalised by a unitary transformation R, such that

$$Rm^\dagger m R^{-1} = D^2 \tag{3.119}$$

where D is a diagonal matrix with real positive eigenvalues. Then, we define $L^\dagger = mR^{-1}D^{-1}$, where $L^\dagger L = \mathbb{I}$ by construction. Thus, $LmR^\dagger = D$.

We define the unitary transformations $L_\pm$ and $R_\pm$ such that

$$L_- m_- R_-^{-1} = \begin{pmatrix} m_d & 0 & 0 \\ 0 & m_s & 0 \\ 0 & 0 & m_b \end{pmatrix} \quad , \quad L_+ m_+ R_+^{-1} = \begin{pmatrix} m_u & 0 & 0 \\ 0 & m_c & 0 \\ 0 & 0 & m_t \end{pmatrix} \tag{3.120}$$

for the quarks $q = d, s, b, u, c, t$ where $m_q > 0$.

The change of basis of the q_- and q_+ gauge fields into the mass basis is then given by

$$\begin{pmatrix} d \\ s \\ b \end{pmatrix}_L = L_- q_{-L} \quad , \quad \begin{pmatrix} d \\ s \\ b \end{pmatrix}_R = R_- q_{-R} \quad , \quad \begin{pmatrix} u \\ c \\ t \end{pmatrix}_L = L_+ q_{+L} \quad , \quad \begin{pmatrix} u \\ c \\ t \end{pmatrix}_R = R_+ q_{+R} \tag{3.121}$$

As desired, the Lagrangian (3.117) can now be rewritten as

$$\mathcal{L}_{\text{quark,mass}} = -\sum_q m_q \overline{q} q \tag{3.122}$$

where q runs over the six quarks and m_q are the physically observed masses. As in the leptonic case, there are no constraints on these masses in the SM and they remain free parameters that need to be experimentally determined. There are currently six observed quarks (with six corresponding antiquarks) with masses and discovery dates tabulated in Table 3.7.

Quark	EM Charge	Mass	Date of Discovery
up (u)	$+\frac{2}{3}$	1.7-3.1 MeV	1968 (SLAC [75,76])
down (d)	$-\frac{1}{3}$	4.1-5.7 MeV	1968 (SLAC [75,76])
charm (c)	$+\frac{2}{3}$	$1.290^{+0.050}_{-0.110}$ GeV	1974 (SLAC [77] and BNL [78])
strange (s)	$-\frac{1}{3}$	0.100^{+30}_{-20} GeV	1968 (SLAC [75,76])
top (t)	$+\frac{2}{3}$	173.5 ± 1.2 GeV	1995 (CDF [79] and D0 [80])
bottom (b)	$-\frac{1}{3}$	4.190^{180}_{-60} GeV	1977 (Fermilab E288 [79])

Table 3.7: The experimentally measured quark masses and the dates of their discovery. The up-type quarks u, c, t have electromagnetic charge $+2/3$ and the down-type quarks d, s, b have charge $-1/3$. All listed particles have anti-quark counterparts with opposite EM charge.

From the original electroweak Lagrangian defined by (3.106) and (3.116), we have

$$ {}^{g}\mathcal{L}_{\text{EWK,quarks}} = {}^{g}\mathcal{L}_{H,\text{quarks}} + {}^{g}\mathcal{L}_{\text{quarks,kin}} \tag{3.123} $$

After EWSSB, the final electroweak Lagrangian for the quark sector is given

$$ {}^{m}\mathcal{L}_{\text{EWK,quarks}} = {}^{m}\mathcal{L}_{H,\text{quarks}} + {}^{m}\mathcal{L}_{\text{quarks,kin}} \tag{3.124} $$

with definitions found in (3.108), (3.117), and (3.118).

The GIM Mechanism

Since the photon A^μ is massless, the mass and gauge eigenbasis of the electromagnetic current j^μ_{EM} must be identical. This can be seen directly from (3.109) as the charge Q is one-dimensional and therefore trivially commutes with the unitarity transformations L and R used to diagonalise the quark fields ψ, where the term in (3.108) is therefore invariant.

The weak currents are more complicated. The charged weak current of (3.109b), which couples to the W-boson in (3.108), now has the additional matrix $V \equiv L_+ L_-^{-1}$ in the physical mass basis, known as the Cabibbo-Kobayashi-Maskawa (CKM) matrix. Using the transformations on the quark fields defined in (3.121), we find the charged current is to be

$$J^{\mu} = \overline{q}_{+}\gamma^{\mu}\left(1-\gamma_{5}\right)q_{-} = \left(\begin{array}{ccc} \overline{u} & \overline{c} & \overline{t} \end{array}\right)\gamma^{\mu}\left(1-\gamma_{5}\right)V\left(\begin{array}{c} d \\ s \\ b \end{array}\right) \tag{3.125}$$

Similarly, the neutral current that couples to the Z-boson in (3.108) is

$$\begin{aligned} J_{n}^{\mu} &= \overline{\psi}_{L}\gamma^{\mu}\sigma_{3}\psi_{L} - 2\sin^{2}\theta_{W}\overline{\psi}\gamma^{\mu}Q\psi \\ &= \overline{q}_{+}\gamma^{\mu}\frac{1}{2}\left(1-\frac{8}{3}\sin^{2}\theta_{W}-\gamma_{5}\right)q_{+} - \overline{q}_{-}\gamma^{\mu}\frac{1}{2}\left(1-\frac{4}{3}\sin^{2}\theta_{W}-\gamma_{5}\right)q \\ &= \sum_{q=u,c,t}\overline{q}\gamma^{\mu}\frac{1}{2}\left(1-\frac{8}{3}\sin^{2}\theta_{W}-\gamma_{5}\right)q - \sum_{q=d,s,b}\overline{q}\gamma^{\mu}\frac{1}{2}\left(1-\frac{4}{3}\sin^{2}\theta_{W}-\gamma_{5}\right)q \end{aligned} \tag{3.126}$$

Since this is diagonal in the quark fields, there are no generation mixing terms that couple to the Z boson in the physical mass basis. This suppression of these flavour changing neutral currents (FCNCs) at leading order is known as the GIM mechanism after Sheldon Lee Glashow, John Iliopolus, Luciano Maiani, who showed this in 1970.

Due to quantum effects, there is in fact some generation mixing but since these loop corrections are higher order in the coupling g, their contributions shall be negligible for small coupling.

The CKM Matrix

There are many parameterisations of $V \equiv L_+ L_-^{-1}$. We give a canonical one that uses the three Euler angles $\theta_{12}, \theta_{23}, \theta_{13}$ and CP-violating complex phase δ_{13} where

$$\begin{aligned} V &\equiv \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} \\ &= \begin{pmatrix} c_{12}c_{13} & s_{12}c_{13} & s_{13}e^{-i\delta_{13}} \\ -s_{12}c_{23} - c_{12}s_{23}s_{13}e^{i\delta_{13}} & c_{12}c_{23} - s_{12}s_{23}s_{13}e^{i\delta_{13}} & s_{23}c_{13} \\ s_{12}s_{23} - c_{12}c_{23}s_{13}e^{i\delta_{13}} & -c_{12}s_{23} - s_{12}c_{23}s_{13}e^{i\delta_{13}} & c_{23}c_{13} \end{pmatrix} \end{aligned} \tag{3.127}$$

characterises the probability for transitions between the quarks.

We have again used the notation $c_{ij} \equiv \cos\theta_{ij}$ and $s_{ij} \equiv \sin\theta_{ij}$. Notice that the couplings between generations i and j vanish if $\theta_{ij} = 0$.

The four constants θ_{12}, θ_{23}, θ_{13} and δ_{13} are free parameters of the SM. Their experimentally measured values are listed in Table 3.8 below.

θ_{12}	$=$	$13.04 \pm 0.05°$
θ_{13}	$=$	$0.201 \pm 0.011°$
θ_{23}	$=$	$2.38 \pm 0.06°$
δ_{13}	$=$	$1.20 \pm 0.08°$

Table 3.8: Experimentally observed CKM Euler angles. [4]

Since $V^\dagger V = \mathbb{I}$, we have the relation $\sum_i |V_{ij}| = \sum_j |V_{ij}| = 1$. This is known as "weak universality". The remaining constraints of unitarity, imply $\sum_j V_{ij} V_{jk}^\star = 0$ where $i \neq k$. In particular,

$$V_{ub}^\star V_{ud} + V_{cb}^\star V_{cd} + V_{tb}^\star V_{td} = 0 \tag{3.128}$$

Note that this equation defines a closed triangle known as the "unitary triangle" and has area $A = \frac{1}{2}\mathrm{Im}\,(V_{cb} V_{cd}^\star V_{ub}^\star V_{ud})$. Measuring this triangle thus quantifies the complex phase and provides an important test for CP violation in the SM [7]. Notably, the area vanishes if the theory has no complex phase.

We cannot follow the same logic for leptons as we assume that the right-handed neutrino field ν_R does not exist. There is therefore no Yukawa term of the form $\psi_L \phi \nu_R$. The different flavour neutrinos are all degenerate with identically zero mass, with their weak gauge eigenstates equivalent to their mass eigenstates. Specifically, the unitary transformation L_ν, which transforms the state ν_L^g to the mass basis $\nu_L^m \equiv L_\nu \nu_L^g$, is arbitrary as it is not constrained by any mass matrix. Thus we can always choose the leptonic equivalent to the CKM matrix to be diagonal $V_{\text{lepton}} \equiv L_\nu L_e^\dagger = 1$.

Matter-Antimatter Asymmetry

To explain the observed matter-antimatter asymmetry in the Universe [81], Andrei Sakharov proposed three conditions required for baryon-generating interactions to produce matter and antimatter at different rates; such processes are known as "baryogenesis" [82]. Equivalent conditions for the leptons which give rise to "leptogenesis" are discussed elsewhere [83].

The Sakharov conditions are:

- baryon number violation
- interactions out of thermal equilibrium
- C and CP symmetry violations.

The first is a trivial condition required to produce an excess of baryons. The interaction must occur outside thermal equilibrium as otherwise the the CPT theorem, which states that CPT is always conserved [8], will assure compensation between processes increasing and decreasing the baryon number. Finally, we shall consider C and CP-violation.

Since CPT is always conserved, CP-violation requires a violation of T-symmetry. Then, since T is an anti-linear operator that complex conjugates terms in Lagrangian [5], any term with a residual complex phase will necessarily violate T and thus CP-symmetry.

Recall the Feynman-Stueckelberg interpretation [84] where antimatter is thought of as matter travelling backwards in time. A violation of T-symmetry would mean thus mean that interactions and their equations of motion are asymmetric. It follows that a CP-violating complex phase gives rise to a baryon-antibaryon asymmetry and thus an observed excess of matter.

In a similar manner, C-symmetry must also be violated as otherwise the creation of left-handed baryons and right-handed anti-baryons (and vice-versa) would be produced in equal parts. We shall restrict our discussion to CP-violation as it provides natural motivation for the BSM physics that we will probe in Chapter 8. Discussions of C-violating interactions, which occur only in the weak interaction, can be found elsewhere [8].

Whilst most terms in the SM Lagragnian are CP-invariant, the charged weak current in (3.109) coupling to the W boson is not, as it contains the CKM matrix. The relevant part of the Lagrangian is

$$\mathcal{L}_{W,\text{quark}} = \sum_i \sum_j \frac{-g}{2\sqrt{2}} \left(\overline{u}_i \gamma^\mu \left(1-\gamma^5\right) V_{ij} d_j W_\mu + \overline{d}_j \gamma^\mu \left(1-\gamma^5\right) V^\star_{ij} u_i W^\dagger_\mu \right) \tag{3.129}$$

where under CP conjugation, we find [9]:

$$\overline{u}_i \gamma^\mu \left(1-\gamma^5\right) d_j W_\mu \quad \longrightarrow \quad \overline{d}_j \gamma^\mu \left(1-\gamma^5\right) u_i W^\dagger_\mu \tag{3.130a}$$

$$\overline{d}_j \gamma^\mu \left(1-\gamma^5\right) u_i W^\dagger_\mu \quad \longrightarrow \quad \overline{u}_i \gamma^\mu \left(1-\gamma^5\right) d_j W_\mu \tag{3.130b}$$

Then,

$$\mathcal{L}_{q,W} \longrightarrow \sum_i \sum_j -\frac{g}{2\sqrt{2}} \left(\overline{u}_i \gamma^\mu \left(1-\gamma^5\right) V^\star_{ij} d_j W_\mu + d_j \gamma^\mu \left(1-\gamma^5\right) V_{ij} u_i W^\dagger_\mu \right) \tag{3.131}$$

which is only invariant if the CKM matrix $V_{ij} = V^\star_{ij}$ is pure real.

For N families of quarks, the CKM matrix V is a unitary $N \times N$ matrix with N^2 real parameters, such that $V^\dagger V = \mathbb{I}$. There are $2N$ quark fields that can each absorb one phase, however there is one global $U(1)$ symmetry of J^μ common to all quark fields. This implies that there are $N^2 - (2N-1)$ free parameters in general. In N-dimensional space, there is global $SO(N)$ symmetry between the generations, which carries $N(N-1)/2$ real Euler angles. This leaves $1 + N(N-3)/2$ complex phases.

For two families is one real mixing angle and no other free parameters. Indeed, Cabibo's original formulation in 1963 used a 2×2 matrix to account for the four observed quarks u, d, c, s at the time [85]. In 1964 however, CP violation was measured in neutral kaon decays by James Cronin and Val Fitch [86]. This led to Kobayashi and Maskawa in 1973 to theorise the existence of a third generation, which allows a complex phase [87]. For three generations, there are four real parameters, three of which are placed in the pure real $SO(3)$ rotation matrix, leaving one leftover complex phase that violates CP symmetry.

Although all three Sakharov conditions hold in the SM, experimental measurements indicate that the SM contributions are orders of magnitude too small to account for the Universe's asymmetry [4, 8]. A fourth generation of quarks would introduce three complex phases, giving us more freedom for which CP violation to occur. This possibility and a search for the fourth generation top-like quark is discussed in Chapter 8.

The Electroweak Lagrangian

In summary, the full electroweak Lagrangian before EWSSB can be written in the gauge basis as

$$ {}^{g}\mathcal{L}_{\text{EWK}} = {}^{g}\mathcal{L}_{\text{Higgs}} + {}^{g}\mathcal{L}_{\text{gauge}} + {}^{g}\mathcal{L}_{\text{EWK,lepton}} + {}^{g}\mathcal{L}_{\text{EWK,quark}} \tag{3.132} $$

with definitions from (3.60), (3.83), (3.102), and (3.123).

Similarly, the full electroweak Lagrangian after EWSSB can be written in the mass basis as

$$ {}^{m}\mathcal{L}_{\text{EWK}} = {}^{m}\mathcal{L}_{H,\text{gauge}} + {}^{m}\mathcal{L}_{\text{EWK,gauge}} + {}^{m}\mathcal{L}_{\text{EWK,lepton}} + {}^{m}\mathcal{L}_{\text{EWK,quark}} \tag{3.133} $$

with definitions from (3.77), (3.85), (3.103), and (3.124).

3.8 Quantum Chromodynamics $SU(3)_C$

The final ingredient to the SM is Quantum Chromodynamics (QCD) that describes the strong force. It is governed by the non-abelian gauge group $SU(3)_C$ representing the colour symmetry.

Historically, the hadrons were discovered to be composite, made up of hypothetical particles called "quarks". The quarks and therefore the hadrons were considered to live in representations of the approximate $SU(3)_F$ flavour symmetry.

QCD evolved from these ideas to describe the quarks as an interacting field theory with an exact $SU(3)_C$ symmetry. The eight gauge bosons of this group associated to the 8 generators of $SU(3)_C$ (in the adjoint representation) are physically manifested as the gluons. These act as messenger particles between quarks transferring a *strong* binding force to form hadrons.

The QCD Langrangian is

$$\mathcal{L}_{\text{QCD}} = -\frac{1}{4}\sum_{i=1}^{8} F^a_{\mu\nu}F^{a,\mu\nu} + i\sum_{f=1}^{6} \overline{\Psi}^A_f \gamma^\mu (\delta_A^{\ B}\partial_\mu + eA^a_\mu (T_a)_A^{\ B})\Psi_{B,f} - \sum_{f=1}^{6} m_f \overline{\Psi}^A_f \Psi_{A,f} \quad (3.134)$$

where $F_{\mu\nu} \equiv F^a_{\mu\nu}T^a$, $A_\mu \equiv A^a_\mu T^a$ with $a = 1,\ldots,8$, labelling the adjoint generators of $SU(3)_C$, $f = 1,\cdots,6$ are the flavour indices of the six quarks and $A, B = 1,2,3$ are the three colour indices.

The QCD field strength is defined as

$$F^a_{\mu\nu} = \partial_\mu A^a_\nu - \partial_\nu A^a_\mu + g f^a_{\ bc} A^b_\mu A^c_\nu \quad (3.135)$$

where again $a, b, c = 1,\cdots,8$ runs over the 8 adjoint generators, g_s is the strong coupling constant and $f^a_{\ bc}$ is the totally antisymmetric structure constants.

By construction this Lagrangian is gauge invariant under the $SU(3)_C$ transformations. The gauge fields A^a_μ transform under the adjoint representation of $SU(3)_C$ and the quarks transform under the fundamental representation of $SU(3)_C$.

Explicitly,

$$A_\mu \longrightarrow g(x)A_\mu g(x)^{-1} - e^{-1}\partial_\mu g(x) g(x)^{-1} \quad (3.136a)$$

$$\Psi_{Af} \longrightarrow g_A^{\ B}(x)\Psi_{Bf} \quad (3.136b)$$

with $g(x) \in SU(3)_C$ and $A, B = 1,2,3$ the colour indices.

Furthermore, the particles that are governed by the QCD interaction have an additional non-perturbative vacuum state that needs to be taken into account. The vacuum state contains both a gluon and quark chiral condensates and spontaneously breaks the approximate global $SU(N_f) \times SU(N_f)$ chiral flavour symmetry between the quarks. This process leaves only the diagonal $SU(N_f)$ part of the chiral group unbroken [8].

The non-vanishing chiral condensate therefore is charachterised by

$$\langle \overline{\psi}_i \psi^i \rangle \sim \exp\left(i\theta_{\mathrm{QCD}}\right) \tag{3.137}$$

where θ_{QCD} is the QCD vacuum angle and ψ_i are the different flavours of quark fields.

Notice that a CP violating phase is introduced into the theory for non-vanishing θ_{QCD}, however experimental evidence suggests that QCD is always invariant under CP symmetry, requiring $|\theta_{\mathrm{QCD}}| \lesssim 10^{-9}$ [88]. There is no mechanism within the SM that naturally gives rise to the conservation of CP-symmetry and forces the vacuum angle to be so (unnaturally) small. This is known as the strong CP problem of the SM [50].

We shall consider the massless limits in comparison to the QCD scale. If we take the case $N_f = 2$, where the u, d quarks are considered to be exactly massless (with an exact $SU(2)_F$ flavour symmetry), the condensate gives rise to the three massless pions, $\pi^{\pm}$ and π, as the Nambu-Goldstone bosons [8]. The three pions are the mesons (composite particles of a quark-antiquark pair) composed of linear combinations of the u and d quarks, with each pion associated to a state in the adjoint representation of the $SU(2)_F$ symmetry [5]. Notice that these Nambu-Goldstone bosons simply represent the fluctuations along the direction of symmetry of the QCD vacuum, which in this case is $SU(2)_F$ symmetric.

Similarly, if we now include the s quark to be massless in our considerations and extend to $N_f = 3$, we find the full octet of pseudoscalar mesons as the Nambu-Goldstone bosons [5].

Since the flavour symmetry is not exact, due to the non-degenerate masses of the quarks, the pseudoscalar mesons are in fact pseudo-Nambu-Goldstone bosons each with some small mass. The mass difference between the charged pseudoscalar mesons $\pi^{\pm}$, $K^{\pm}$ and their corresponding neutral partners π^0, K is due to electromagnetic corrections. The electromagnetic mass difference has been determined in the chiral limit using current algebra for the pions in Ref. [89] and the kaons Ref. [90]. Interestingly, Dashen's theorem [91] states that the squared mass differences for the pions and kaons are equal in the chiral limit, i.e. $\Delta M_K^2 - \Delta_\pi^2 = 0$, where $\Delta_P^2 = M_{P\pm}^2 - M_{P0}^2$. However, more recent research indicates that the electromagnetic corrections, due to the non-vanishing quark masses, may violate this relation [92–95]. These particles have all been experimentally measured [4] and they are classified in Appendix A.4.1.

Arbitrarily including the heavier flavours does not actually give rise to any new physically observed particles. This is because the approximate flavour symmetry breaks down and the composite mesons are too heavy to form bound states. However, physics of course will be independent of θ_{QCD} in the massless limit of QCD.

Ongoing research in lattice QCD [96] will hopefully yield a full theory of the colour force: including a determination of the vacuum structure, a description of the the observed hadron spectrum and provide a natural explanation for confinement.

3.8.1 Confinement and Asymptotic Freedom

The gluons along with the quarks are never observed free. They obey the law of "confinement" [8], which asserts that the dynamics of QCD allow only singlets of $SU(3)_C$ to be physically observable.

In fact, the coupling strength between both quarks and gluons, collectively known as partons, becomes increasingly large as the distances between them increase. Trying to pull apart these partons from each other, eventually leads to enough potential energy to spontaneously create new hadrons in the vacuum by drawing out quarks from the quark sea. This process is known as "hadronisation" and it is these hadrons that are observed in our detector experiments.

The large couplings of QCD mean that perturbative calculations naïvely breaks down. However, David Gross, Frank Wilczek and David Politzer [97, 98] calculated the "running" coupling constant of the strong interaction due to quantum corrections to be given by the β-function:

$$\beta(\alpha) \equiv \frac{\partial \alpha}{\partial \ln \mu} = \frac{\alpha^2}{\pi}\left(\frac{n_f}{3} - \frac{11}{6}\right) \tag{3.138}$$

where $\alpha = g_s^2/4\pi$, n is the number of quark generations, n_f is the number of quark flavours and μ is the energy scale of the given physical process.

Thus, if this quantity is negative, the coupling strength decreases with increasing energy. This effect is known as the "principle of asymptotic freedom". Since we observe three generations $n = 3$ and $n_f = 6$ quarks, we find indeed that $\beta < 0$. We are therefore able to probe regimes of weak QCD coupling using our high energy collider experiments.

3.9 The Parameters of the Standard Model

In summary, there are 19 free parameters of the SM:

- the 9 fermion masses in Tables 3.6 & 3.7
- the 4 CKM angles in Table 3.8
- the vacuum expectation value $v = 246$ GeV [65]
- the Higgs mass $m_H = 125.3 \pm 0.6$ GeV [43]
- the QCD vacuum angle $\theta_{\text{QCD}} \sim 0$ [88]
- and the 3 couplings g', g and g_s for the SM gauge groups $U(1)_Y$, $SU(2)_L$ and $SU(3)_C$ respectively [99].

All three gauge couplings can be expressed in terms of the dimensionless parameter $\alpha_i = g_i^2/4\pi hc$. Similar to the QCD case, quantum corrections cause all gauge couplings to run with respect to the energy scale we consider [8]. Experimental measurements that are extrapolated to higher energies give the coupling strengths in Fig. 3.5 [99].

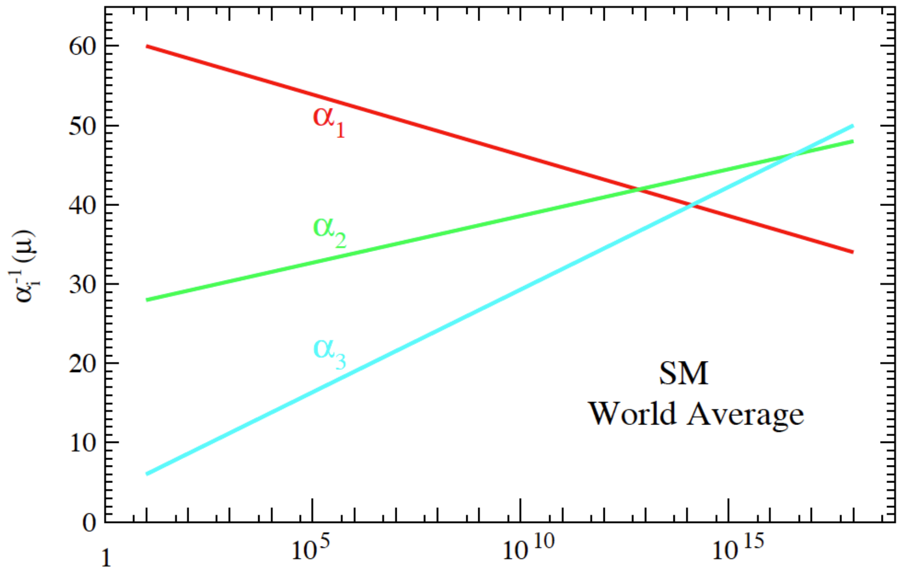


Figure 3.5: The running couplings α_i of the three SM gauge groups as a function of the energy μ, where $i = 1, 2, 3$ represent the gauge groups $U(1)_Y$, $SU(2)_L$ and $SU(3)_C$ respectively. In some BSM theories, such as supersymmetry, the gauge couplings are predicted to coincide at $\sim 10^{16}$ GeV [99].

3.10 The Fundamental Forces

The complete SM Lagrangian, which governs the strong, weak and electromagnetic forces is

$$\mathcal{L}_{\text{SM}} = \mathcal{L}_{\text{EWK}} + \mathcal{L}_{\text{QCD}} \tag{3.139}$$

where $\mathcal{L}_{\text{EWK}}$ is the combined electroweak Lagrangian defined in (3.132) and $\mathcal{L}_{\text{QCD}}$ is defined in (3.134).

The final fundamental force, gravity, is described by Einstein's equations (2.88). These are the equations of motion from the Einstein-Hilbert Lagrangian for General Relativity:

$$\mathcal{L}_{\text{EH}} = \frac{1}{28\pi G} R\sqrt{-g} \tag{3.140}$$

where R is the Ricci scalar, G, Newton's constant and $g = \det(g_{\mu\nu})$ is the determinant of the metric tensor. Since the mass dimension of the coupling $\frac{1}{G}$ is positive since $[G] =$

-2, the theory is inherently non-renormalisable and as consequence has so far not been quantised [15].

These Lagrangians, together with their Euler-Lagrange equations of motion (1.26) and the S-matrix that determine the probabilities of interactions (2.106), completely define all known physical interactions.

Chapter 4

Beyond The Standard Model

We are motivated to look for a more fundamental theory of nature as there are already theoretical and experimental hints that the SM is incomplete. The main unresolved issues within the framework of the SM include:

- Quantum gravity
- Neutrino masses
- Dark energy and dark matter
- Baryon-antibaryon asymmetry
- The hierarchy problem and naturalness.

4.1 Quantum Gravity

A quantum theory of gravity is missing entirely from the SM. Many attempts have been made to describe gravity in a consistent quantised manner, including Supersymmetry, String theory and Loop Quantum Gravity [100–103] but so far none have been widely accepted.

4.2 Neutrino Masses

There is an unresolved complication arising from neutrino measurements. Neutrinos are observed to oscillate between flavours [104], indicating that their flavour and mass eigenstates are inequivalent. In turn, this implies that their masses are non-vanishing.

This is fundamentally inconsistent with the SM, which has only left-handed neutrinos and thus no possibility of a Dirac mass term. Furthermore, they are very difficult to study experimentally as they are both electromagnetically uncharged and colourless under the $SU(3)_C$ gauge symmetry. Therefore, they interact only via the weak and gravitational forces.

Neutrino physics is a very active field of research and the exact way in which they fit into the SM is not very well understood. The seesaw mechanism or the neutrino being a new type of particle known as a "Majorana" fermion are but two of the leading theories to resolve this issue. However, this remains an active area of research and is discussed elsewhere [9, 64].

4.3 Dark Energy and Dark Matter

Only 4.9% of the universe is composed of regular baryonic matter [105]. The SM provides no natural explanation or candidate for either dark energy and dark matter that are hypothesised to make up the remaining 95.1%.

Measurements of the cosmic microwave background indicate that approximately 68.3% of the universe is composed of dark energy [105, 106]. This is a hypothetical form of energy that permeates all space, responsible for the accelerating expansion of the universe [107]. The cosmological constant, introduced in Einstein's equations of General Relativity [15] or quintessence [108] provide potential explanations for dark energy.

Dark matter makes up the remaining 26.8% of the universe [105]. Difficulty to detect such matter except by gravitational rotation curves in galaxies [109, 110] indicates that its constituents interact only via gravitational effects. No such weakly interacting particles exist within the SM framework. Supersymmetry, which predicts a lightest stable particle (LSP), is one model that provides a natural dark matter candidate [100].

4.4 Baryon-Antibaryon Asymmetry

Baryon-antibaryon asymmetry observed in the universe indicates that the equations of motion violate charge-parity (CP) symmetry. Although the SM provides one complex CP violating phase in the CKM matrix (3.127), recent measurements [4] are too restrictive to describe observations.

A sequential fourth generation of quarks is one of the simplest extensions to the SM that may introduce enough freedom in the CKM matrix to explain this observation. This is discussed further in Chapter 8 and a search for the fourth generation heavy top-like quark (t') is described in Chapter 9.

4.5 The Hierarchy Problem

There is a large hierarchy between the different scales in particle physics Fig. 1. Most notably, the difference between the electroweak $\mathcal{O}(10^2)$ GeV and Planck scales $\mathcal{O}(10^{19})$ GeV is dramatic and introduces unnatural artefacts in the breaking of electroweak symmetry.

Spontaneous symmetry breaking of the electroweak gauge symmetry by the Higgs field, gives rise to the observed masses of all particles including itself. To second order the Higgs boson has mass

$$m_h^2 = \frac{1}{2}\lambda v^2 + \Delta m_{h,i}^2 \tag{4.1}$$

where $\Delta m_{h,i}$ indicates quantum corrections from particle i.

Since the coupling of the Higgs field to the top-quark is the largest, indicated by the top having the largest mass of the fundamental particles, its quantum contribution to the Higgs mass will be the most significant. At one-loop level, the coupling gives rise to a contribution with a Feynman diagram shown in Fig. 4.1-(a).

The scattering amplitude can be directly computed using Feynman rules [8], giving

$$\Delta m_{h,t}^2 \;\propto\; \frac{\lambda_t^2}{(4\pi)^2}\int d^4k\frac{1}{k^2} \;\sim\; \frac{\lambda_t^2}{(4\pi)^2}\Lambda^2 \tag{4.2}$$

where λ_t is the coupling between the top-quark and the Higgs field.

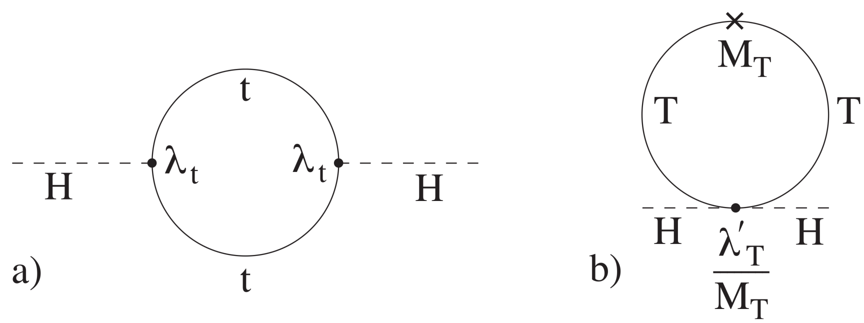


Figure 4.1: Loop contributions to the Higgs mass. (a) the SM top-quark quadratic coupling and (b) an added heavy partner T quartic coupling to the Higgs field.

Due to the integration over four-momenta, this quantity is a quadratically divergent and will flow up to the momentum cut-off Λ^2 where our model is no longer applicable. If there is no new physics beyond the SM, the natural cut-off scale is the Planck scale, where quantum gravity effects become large and the notion of space-time loses its meaning.

However, tentative measurements of the Higgs boson indicate that its mass falls around 125 GeV [43,67]. An incredible fine tuning of the SM parameters to an accuracy of $\mathcal{O}(10^{34})$ is required to unnaturally cancel this loop correction.

These additional fields typically stem from a deeper, underlying theory. Looking at the contributions to the Higgs mass, we have

$$m_h^2 \sim 125^2\,\mathrm{GeV}^2 \;=\; m_{h,0}^2 + \Delta m_{h,t}^2 + \cdots \tag{4.3}$$

where $\cdots$ indicate other quantum contributions.

Taking $m_{h,0}^2 \ll \Delta m_{h,t}^2$, we see that

$$(125\ \mathrm{GeV})^2 \sim \frac{\lambda_t^2}{(4\pi)^2}\Lambda^2 \tag{4.4}$$

Thus, new physics is likely to enter at

$$\Lambda \;\sim\; \mathcal{O}(1)\,\mathrm{TeV} \tag{4.5}$$

where we have assumed, for perturbation theory to hold, the coupling is at most $\lambda_t \sim \mathcal{O}(1)$.

Many beyond the SM theories have been proposed that introduce heavy partners to the SM fields. These naturally cancel the divergence from the top-loop as well as similar contributions from other SM fields [8]. For example. Supersymmetry introduces a heavy parter to the top-quark known as the 'stop' quark T, shown in Fig. 4.1-(b). In this case, the two contributions cancel:

$$\Delta m_{h,t}^2 + \Delta m_{h,T}^2 \ \sim \ 0 \tag{4.6}$$

In general there can be more than one heavy-quark contribution that cancel the SM quadratic divergence. This is indeed the case in the class of Little Higgs Models. The phenomenology of such theories and a direct search for the predicted heavy top-like quark (T) are described in Chapter 10 and Chapter 11 respectively.

Both the fourth generation t' and the Little Higgs T searches are carried out at the Large Hadron Collider with data collected by the Compact Muon Solenoid Experiment.

Part III

The LHC and the CMS Detector

Chapter 5

The Large Hadron Collider

The Large Hadron Collider (LHC) at the Conseil Européen pour la Recherche Nucléaire (CERN) is currently the most powerful particle accelerator in the world. Located in the Geneva-region, it is a proton-proton (pp) collider measuring 26.7 km circumference and placed underground at a depth between 45 to 170 meters [111, 112].

5.1 The Injection Complex

Protons are stripped from a plasma made of hydrogen H_2 molecules at the Duoplasmatron by a large electric field. These are accelerated by electric fields, focused by radiofrequency quadrapole magnets and emitted in bunches at 90 KeV in 25 ns intervals into the linear accelerator LINAC2.

A 200 MHz, 50 MeV alternating electric field is applied in the LINAC2, accelerating the protons in periods. When the protons would otherwise have been subjected to the reverse (decelerating) electric field, they are shielded inside drift tubes. These drift tubes and the spaces between them become progressively larger at increasing energies (speeds). The beam is focused with a combination of Coulomb repulsion between same-sign charges and the magnetic attraction between parallel currents.

The protons then exit the LINAC2 and pass through the Proton Synchrotron Booster at 1.4 GeV, being bent around by a 4 T magnetic field. They are then injected into the Proton Synchrotron (PS) and accelerated to 26 GeV, before being ejected into the Super Proton

Synchrotron (SPS), where they are further accelerated to 450 GeV. They are finally injected into the LHC in both directions, where they are accelerated to their peak centre-of-mass energy of 14 TeV at design luminosity.

The LHC injection complex is shown in Fig. 5.1.

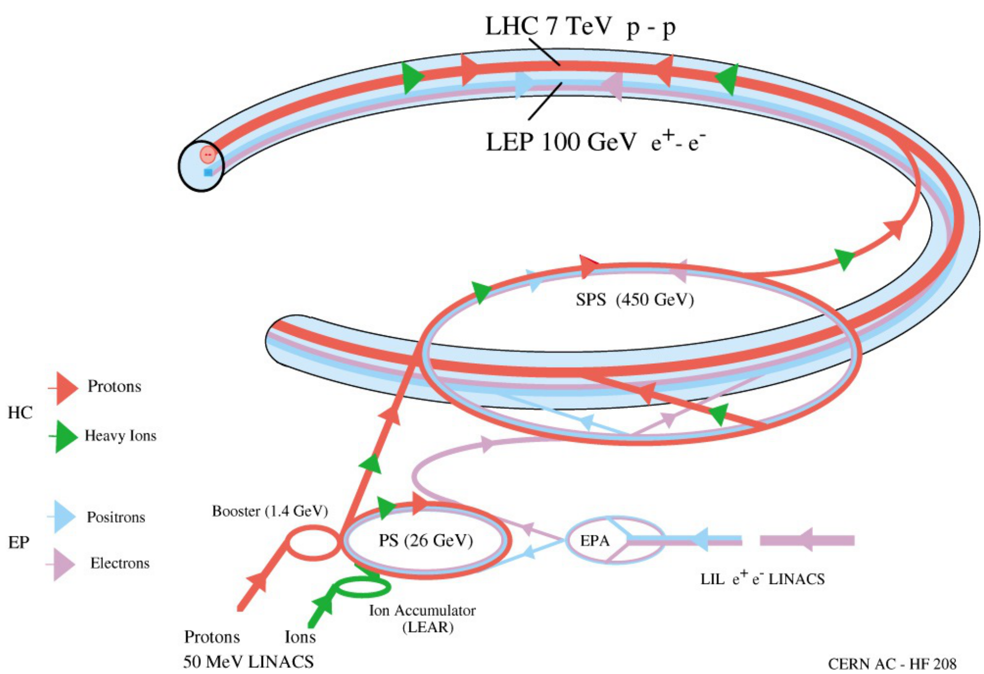


Figure 5.1: The LHC injection complex. At design luminosity, each beam is accelerated to 7 TeV. The LEP injection complex that accelerates electrons and positrons is also shown. [113]

5.2 The Large Hadron Collider

The ring that transports the beam houses two vacuum pipes, each containing radio-frequency cavities, magnets, and one of the two counter-rotating beams of protons. A cross section of the LHC beam is shown in Fig. 5.2.

The superconducting dipole magnets, cooled by liquid helium at 1.9 K, provide a magnetic field of 10 T used to keep the protons orbiting around the ring in a roughly circular trajectory.

The 1,232 dipole magnets are created by passing a large current down coils of copper-clad niobium-titanium cables; the resulting magnetic field created is shown in Fig. 5.3.

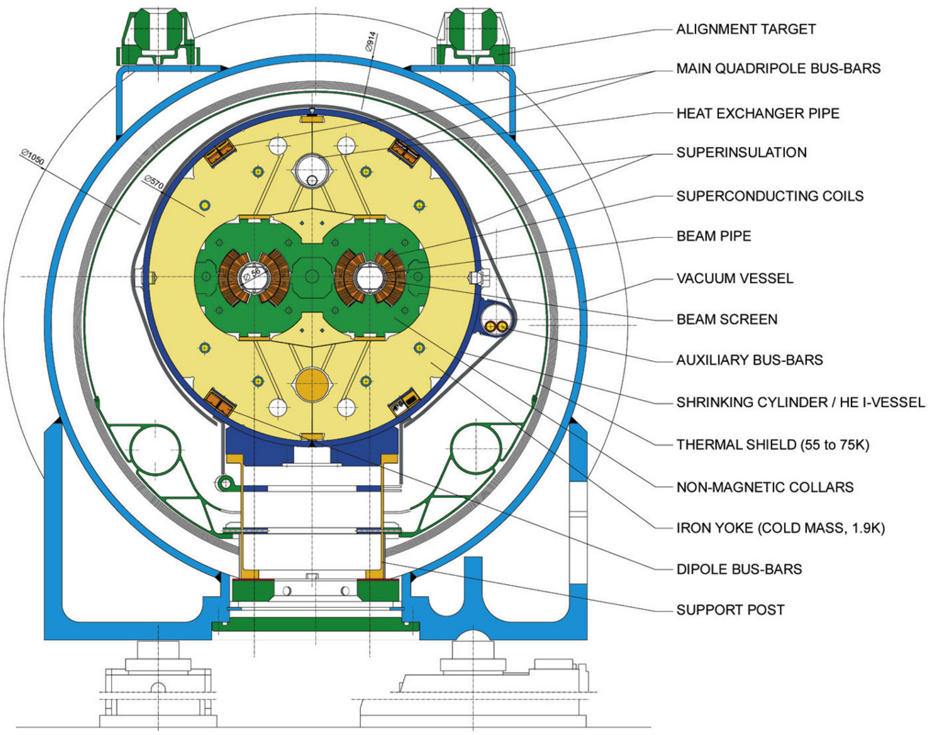


Figure 5.2: The LHC cross section. [113]

Finally, a further 392 quadrupole magnets are placed around the LHC ring to focus the two beams.

Radio-frequency (RF) cavities are resonating structures used to transfer electrical energy for particle acceleration. Each RF cavity generates a longitudinal oscillating voltage delivering 2 MV (an accelerating field of 5 MV/m), which is applied across an isolated gap in the vacuum chamber such that particles see an accelerating voltage at the gap on each revolution. There are 16 such cavities, with a frequency of 400 MHz, organised in four cylindrical refrigerators called "cryomodules" (two per beam); these are installed at a long straight section of the LHC known as Point 4.

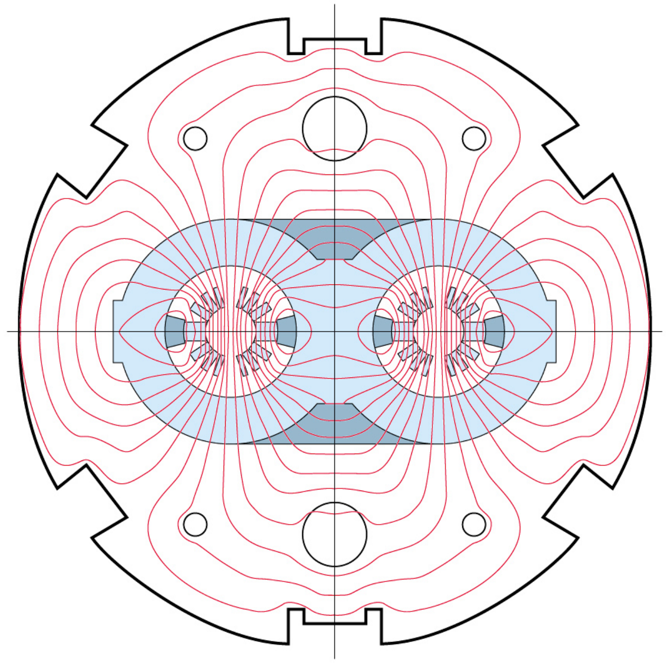


Figure 5.3: A Cross-Section of the Computed Magnetic Flux Map at $B = 10\,\mathrm{T}$ generated by the LHC dipole. [113]

5.3 The Experiments

Five experiments lie at four designated interaction points along the beam-line where the two counter-rotating beams cross. ATLAS (A Toroidial LHC Apparatus) and CMS (Central Muon Solenoid) are the two large general purpose detectors. The remaining detectors are LHC-B, designed to study Bottom b quark physics, ALICE (A Large Ion Collider Experiment) studying heavy ions, and TOTEM (Total Cross Section, Elastic Scattering and Diffraction Dissociation), designed to study physics in the very forward regions close to the beam-line. The entire accelerator complex is shown in Fig. 5.4.

The interactions between the constituents of the colliding protons, known as *partons*, produce new particles. These new particles are measured at each of these detectors, designed specifically to probe the Standard Model and beyond.

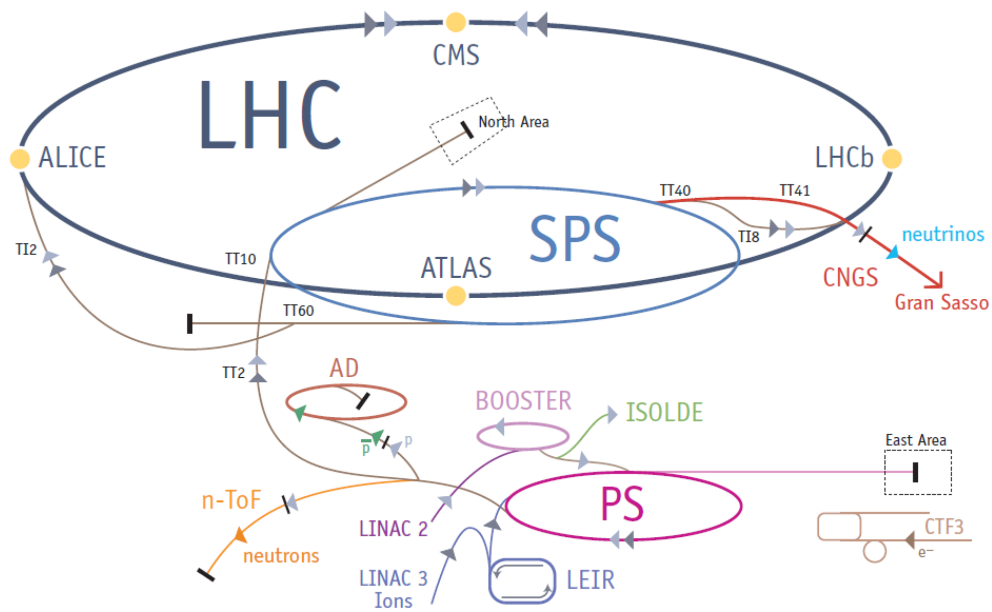


Figure 5.4: The CERN accelerator complex. [113]

5.4 Conventions

At hadron colliders, the momentum of colliding protons is almost entirely along the direction of the beam, where the beam-line is chosen to lie along the z-axis by convention. The longitudinal and transverse directions are defined as the directions perpendicular and parallel to the beam axis, respectively.

Furthermore, since the energy of the colliding hadrons is not evenly distributed amongst their constituent hadrons, interactions may give rise to events with significant total momentum along the z-axis. This is especially difficult to detect, most notably due to the lack of detecting material along the beam. More importance will therefore be placed on particles travelling transverse to the beam, where the total momentum and energy are well known.

We introduce the transverse variables p_T and E_T, which shall denote the transverse momentum and the transverse energy. We also denote the missing energy as $\not{E}_T$ and the scalar sum of transverse momenta of reconstructed particles as H_T .

Since, the momenta along the beam-line is difficult to reconstruct, we define the rapidity y that is independent of Lorentz boosts along the z-axis

$$y = \frac{1}{2} \ln \frac{E + p_z c}{E - p_z c} \tag{5.1}$$

For the typical case at hadron colliders, the mass is negligible compared to the momentum and we can convert this quantity to the pseudo-rapidity

$$\eta = -\ln \left[\tan \left(\frac{\theta}{2} \right) \right]. \tag{5.2}$$

which is dependent on the conventional cylindrical polar angle θ relative to the anticlockwise proton beam direction

Chapter 6

The Compact Muon Solenoid

The Compact Muon Solenoid (CMS) detector is a general purpose detector, situated approximately 100 m underground at the LHC interaction point 5 (P5), located near the village of Cessy in France. The detector, shown in Fig. 6.1, is a large cylindrical solid-angle magnetic spectrometer with a superconducting solenoid that applies an axial magnetic field of 4 T.

CMS is designed to accurately reconstruct trajectories, momenta, and energies of all SM particles. Particles created in *pp* collisions travel radially from the interaction point through the detector, which deposits energy in the various subsystems. A cross section of the CMS detector is shown in Fig. 6.2. Different subsystems are situated in a near-hermetic layered structure around the nominal interaction vertex at the centre of the detector and are designed to achieve both high precision and accuracy.

The innermost cylindrical layer of the detector closest to the interaction point is the silicon tracker. This is enclosed by a lead tungstate crystal electromagnetic calorimeter (ECAL), with a lead-silicon preshower detector in the endcaps. In turn, the ECAL is surrounded by a brass-scintillator hadronic calorimeter (HCAL). Finally, muons are identified by gas-ionisation muon chambers that are embedded in the steel return yoke outside the solenoid. These form the outermost layer of the detector. The very forward regions are also equipped with various calorimeters. A two-tier hardware/software trigger system is then implemented to select interesting proton-proton collision events.

Each of the subsystems that compose the CMS detector shall be discussed in the following sections. A full description of the CMS detector can be found elsewhere [10–12,114,115].

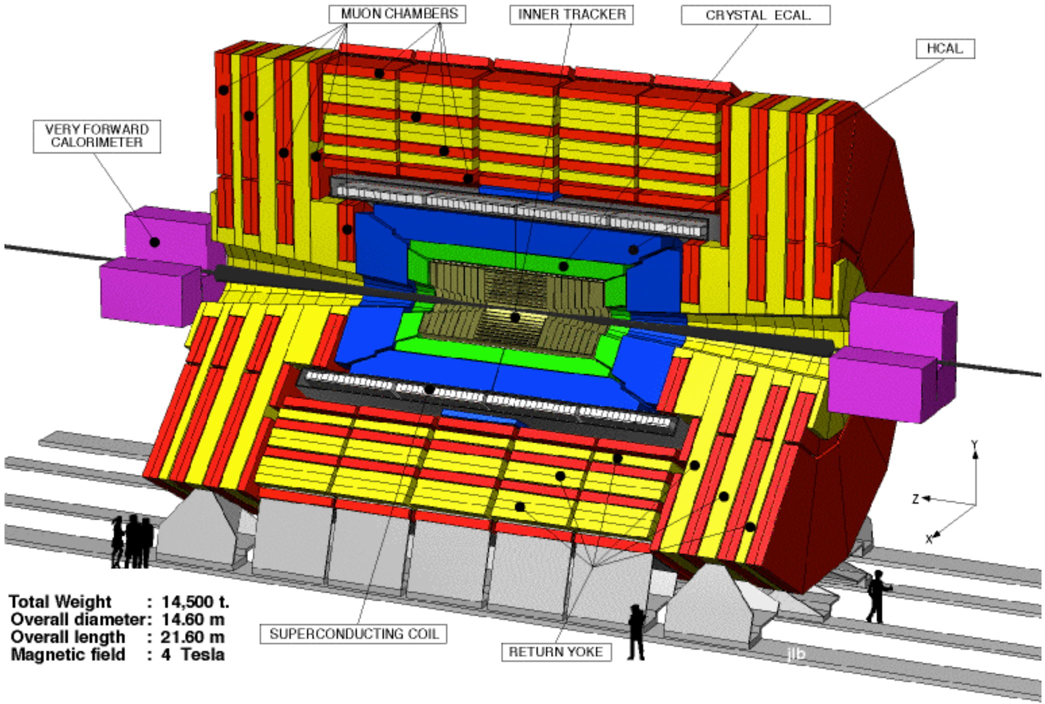


Figure 6.1: The CMS detector. [12]

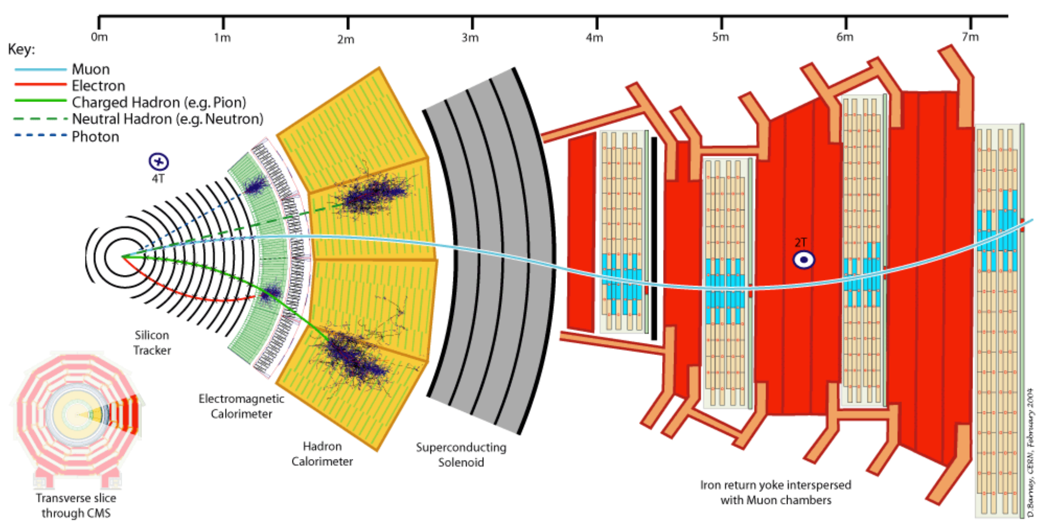


Figure 6.2: A cross section of the CMS detector. [12]

The SM particles that can be identified at CMS are the visible leptons, photons, mesons, hadrons and neutrinos. Aside from the neutrinos, all other SM particles that pass through the tracker leave distinct energy signatures. In particular, electrons and photons deposit the majority of their energy in the electromagnetic calorimeters and the hadrons deposit energy in the hadronic calorimeter. Muons deposit energy throughout the detector, and since they are minimally ionising particles, they are likely to be the only particles, with the exception of neutrinos, to travel through the muon system. Neutrinos are very weakly interacting and are only reconstructed by computing the total missing transverse energy of the system.

The reconstruction algorithms of all physics objects are discussed in Section 7.

6.1 Coordinate System

CMS uses a right-handed coordinate system with the origin at the nominal collision point. The x-axis points towards the centre of the LHC ring; the y-axis points upwards, and the z-axis runs parallel to the beam-line. In spherical coordinates, the polar angle θ is measured from the z-axis and the azimuthal angle ϕ from the x-axis in the plane transverse to the beam.

6.2 The Superconducting Solenoid and Magnetic Coil

The entire CMS detector is designed around the superconducting solenoid and the desire to create a large homogeneous magnetic field. Tracks from charged particles that emerge from pp collisions, are identified by their circular path in the magnetic field. The curvature can be used to compute a particle's velocity and momentum. Therefore, a strong field is necessary to provide high resolution measurements.

Creating a 4 T homogeneous magnetic field was the single largest engineering challenge when designing the CMS detector; it is currently the largest superconducting magnet ever built.

Using Maxwell's equations (2.1), the magnetic field at CMS is created by passing the largest possible current down the conducting wire of the solenoid. To minimise the resistance, the wire must be superconducting, which requires that the temperature of the solenoid be lowered (using a cold mass inside a cryostat). The solenoid's conductor consists of a 32 strand Rutherford cable embedded in very high purity aluminium, with an operating current of 19.5 kA and a precision of 10^{-4}%. The conductor is mechanically reinforced with an aluminium alloy.

Due to the number of Ampére turns required to generate a 4 T magnetic field (41.7 MA-turn) the winding is composed of four layers. Further, the radial extent of the coil is very small in relation to the distance from the beam-line creating a well understood and uniform magnetic field. The number of coils and thus the strength of the magnetic field was limited by the transportation and building restrictions of the detector.

Within the magnetic coil, this free bore magnet houses the tracker, the ECAL and the HCAL. It measures 6.3 m in diameter, 12.5 m in length, has a weight of 22 tonnes and stores 2.6 GJ of energy at full current.

The field runs parallel to the beam-line and is returned through the iron yoke, which has a saturation of -1.8 T, shown in Fig. 6.3. The return yoke through which the flux is guided, weighs $10,000$ tonnes and is composed of five barrel wheels and two endcaps, each individually consisting of three discs going out to 14 m in diameter. The twelve sided return yoke is interspaced by the gas chambers of the muon detector and also acts as additional absorbing layers typically allowing only muons and neutrinos through.

The magnet also provides the majority of the structural support and is designed to withstand the high magnetic field that it produces.

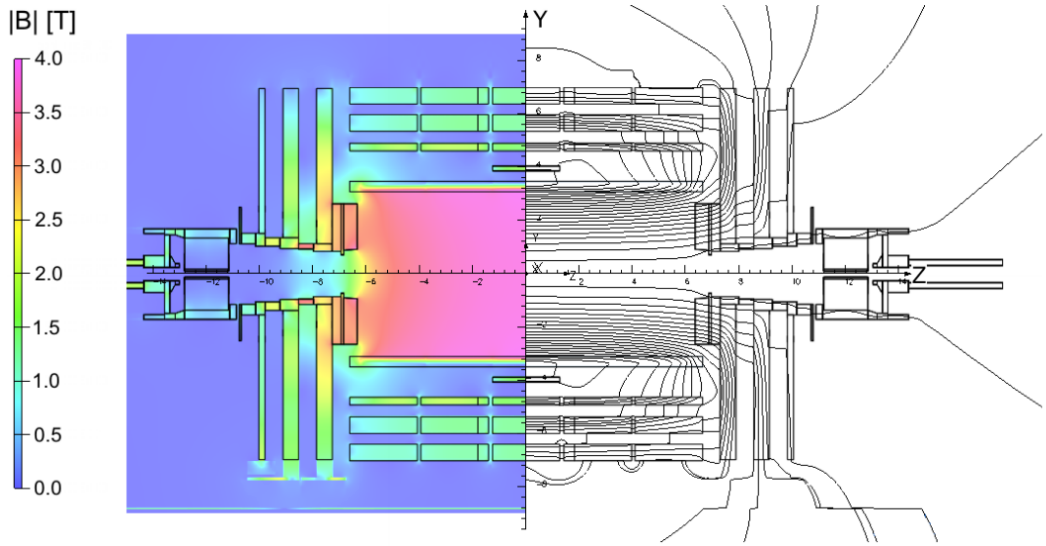


Figure 6.3: The magnetic field of CMS mapped in the CMS barrel yoke using cosmic rays. Value of (left) $|B|$ and (right) field lines in the longitudinal direction. Neighbouring magnetic field lines are separated by 6 Wb. [116]

6.3 The Tracking System

The tracker is an ionisation detector and forms the innermost layer relative to the interaction point; it is therefore the first system through which particles, created in pp collisions, pass. High granularity, precision, fast output, and the necessity to survive in a high radiation environment motivates the choice of silicon based technology.

A 4 T homogeneous magnetic field is applied over the entire tracker subdetector. The tracker consists of the silicon pixel detectors in the innermost region and the silicon strip detectors in the inner and outer barrel layers, shown in Fig. 6.4.

Since the average flux decreases with increasing radius, larger silicon panels can be used further out to reduce cost. The flux is approximately 1 MHz/mm^2 at $r = 4$ cm, 60 kHz/mm^2 at $r = 22$ cm down to 3 kHz at $r = 115$ cm. The fall is not exactly proportional to the naïve $1/r^2$ due to the magnetic field.

The silicon tracker covers the azimuthal range $0 < \phi \leq 2\pi$ and the pseudorapidity range $|\eta| \leq 2.5$.

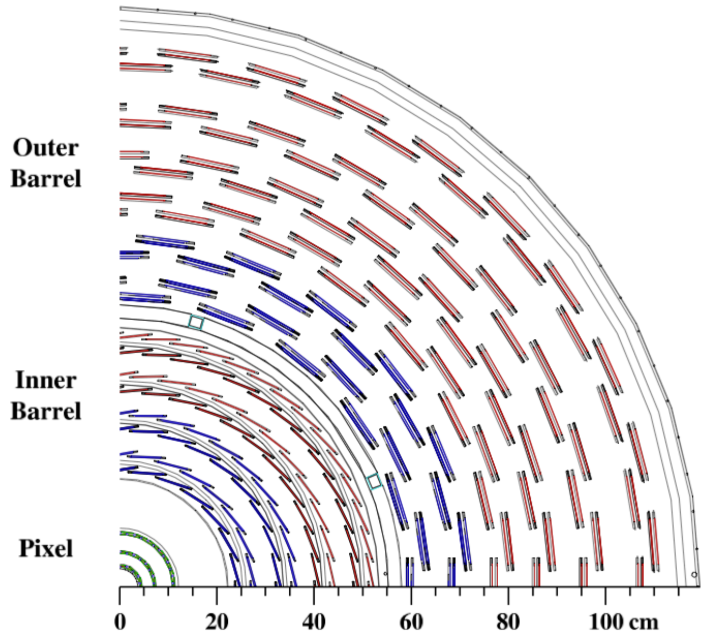


Figure 6.4: Cross section of the CMS tracker. [12]

The entire tracker is housed in a support tube suspended on the HCAL barrel and discs. The support tube measures 5.30 m long and has an inner diameter of 2.38 m, with 33 mm thick cylindrical wall. The tracking information is primarily used to reconstruct tracks but it is also used in the triggering systems (discussed Section 6.7).

Ionisation Detectors

Ionising radiation from particles that traverse the detector liberate valence electrons (leaving "holes") in the silicon; these flow around a circuit and are measured as current. Since the energy required to ionise one valence electron is constant and dependent only on detector material, the number of charged carriers is directly proportional to the amount of ionising energy that is deposited, which in turn is proportional to the energy of the incident particles.

The speed of the electrons is determined by the size of the applied electric field; thus one can, after calibration, determine the amount of energy of the particles from current measurements. These electron and holes travel very quickly, providing the excellent energy and time resolution desired.

6.3.1 Silicon Pixel Tracker

The tracker is designed to keep the occupancy below 1%. Silicon pixels that measure $100 \times 15\,\mu\text{m}^2$ (in $r-\phi$ and z directions respectively) are used in the inner tracker. This requirement is driven by the desire to have excellent impact parameter resolution $\mathcal{O}\left(10^{-4}\right)$ cm defined as the distance between primary and secondary interaction points, where secondary vertices typically arise from decays of heavy particles close to the interaction points.

The inner CMS pixel tracker has three cylindrical barrel layers at radii 4.4, 7.0, 10.2 cm around the beam-line, with these layers extending roughly 53 cm along the z axis. Additionally, the pixel detector has endcaps, two silicon discs on either side of the pixel barrel with a hole for the beam-line ($6\,\text{cm} < r < 15\,\text{cm}$). These are placed at $z = \pm 34, \pm 46$ cm along the beam-line (with $z = 0$ cm the nominal interaction point), shown in Fig. 6.5.

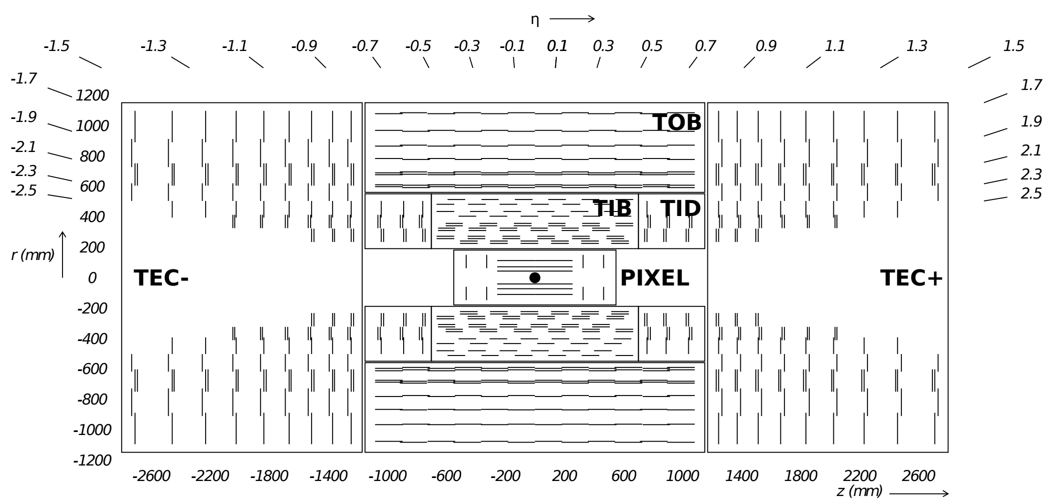


Figure 6.5: Layout of the CMS tracker. [12]

In total there are 66 million pixels, read out by 16,000 readout chips, giving a spatial resolution of between 10 μm and 20 μm.

6.3.2 Silicon Strip Tracker

The silicon strip detector consists of the tracker inner barrel and discs (TIB/TID), supported on the tracker outer barrel (TOB) and both endcaps (TEC$\pm$ where $\pm$ defines the position relative to the nominal interaction point $z = 0$). The entire strip tracker rests on rails in the tracker support tube.

There are ten barrel layers extending to 10.2 cm in z direction and $20\,\text{cm} < r < 116\,\text{cm}$. Further, there are three strip detector endcap discs, placed on either side of the barrel. The reduced particle flux means that the strip detectors can be used. These are placed parallel to the beam axis in the barrel, and in the radial directions in the endcap discs, providing measurements on the trajectory up to $|\eta| < 2.5$.

In the inner barrel the silicon microstrips measure $10\,\text{cm} \times 80\,\mu\text{m}$ with a thickness of $320\,\mu\text{m}$ and in the outer barrel, $25\,\text{cm} \times 180\,\mu\text{m}$ with a thickness of $500\,\mu\text{m}$.

The TIB consists of four cylindrical layers at radii between $20\,\text{cm} < r < 55\,\text{cm}$ centred around the beam-line. Three identical discs that make up the TID are placed at each end of the TIB. These discs span the range $0.2\,\text{m} < r < 0.5\,\text{m}$ and is located between 0.8 m and 0.9 m away from nominal the interaction point in the z-direction.

The TIB/TID is surrounded by the TOB, with an outer radius of 116 cm. The TOB has 6 barrel layers that extend to $z = \pm 118\,\text{cm}$.

The outer barrel has nine endcap discs (TEC) on either side of the TOB. The TEC covers the radii $22.5\,\text{cm} < r < 113.5\,\text{cm}$ with a hole for the beam-line and the range $1.24\,\text{m} < |z| < 2.80\,\text{m}$ along the beam axis.

The silicon strip detector consists of a total of 9.3 million strips and has a $198\,\text{m}^2$ active silicon area. There are 15,148 detector modules with each module containing a set of sensors, a mechanical support structure, and a readout hybrid bonded directly to the sensors. Finally, 50,000 optical fibres are used to extract data for the readout of these modules.

6.4 The Electromagnetic Calorimeters (ECAL)

The CMS ECAL is a hermetic homogeneous calorimeter composed of lead tungstate ($PbWO_4$) designed to detect photons and electrons. In particular, the ECAL is optimised to detect the diphoton decay of the Standard Model Higgs boson. Further, many interesting physical processes produce photons and electrons, which deposit most of their energy in the ECAL.

When a high energy electron ($E \gg 10\,\mathrm{MeV}$) travels through material of high atomic number, the primary mechanism of energy loss is braking energy, Bremsstrahlung. This occurs when electrons are deflected off of other charged particles, with the electron's kinetic energy being converted to pairs of photons with a continuous energy spectrum (higher frequency photons being released for higher energy incident electrons). High energy photons, on the other hand, interact mainly by pair production $\gamma \to e^+e^-$ in the vicinity of a nucleus. Together, these two processes convert incident high energy photons and electrons into EM showers, with secondary interactions continuing to happen until there is insufficient energy, at which point other energy loss mechanisms, such as ionisation, become important.

Energy loss dE is a constant of the material; it is defined in terms of the radiation length X_0, the energy of the incident particle E, and the depth it travels through the material dx, where

$$\frac{dE}{E} = -\frac{dx}{X_0} \tag{6.1}$$

Lead tungstate is used both as the absorbing material, initiating showers, and as the active scintillating material, giving fast response times and fine granularity.

Lead tungstate has a small Molière radius, equivalent to the radiation length in the transverse direction, of 2.10 cm. This allows the ECAL to be compact and to fit, together with both the tracker and the HCAL, within the magnetic coil. Furthermore, the radiation length of $PbOW_4$ is only $X_0 = 0.89\,\mathrm{cm}$, such that the ECAL thickness is larger than $25\,\mathrm{X}_0$ at any point. Thus the majority of the EM showers are contained. The layout of the ECAL is shown in Fig. 6.6.

There are $612,000$ crystals mounted in the central barrel and 7324 crystals in each of the two endcaps. The preshower detector sits in front of each of these two endcap crystals and provides additional detector material. Photons from the EM shower travel down the ECAL crystals, which have both highly polished surfaces and a high refractive index to allow for optimal light collection. The optical transmission is approximately 80% in 25 ns dependent on the wavelength and the crystals emit blue-green scintillation light with maximum at 425 nm.

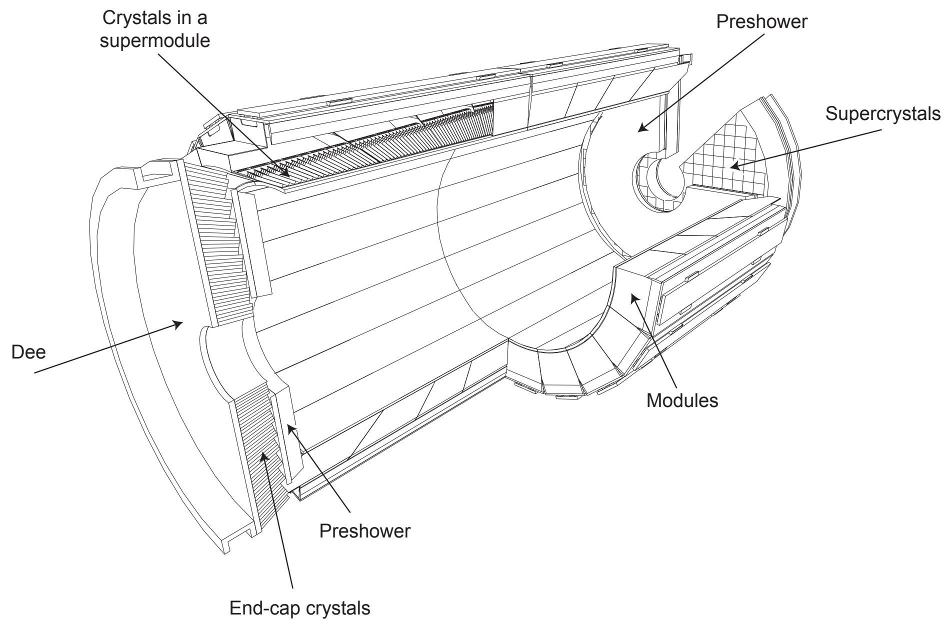


Figure 6.6: Layout of the CMS electromagnetic calorimeter showing the arrangement of crystal modules, supermodules and endcaps, with the preshower in front. [12]

The ECAL provides excellent resolution by using avalanche photodiodes (APDs) and vacuum phototriodes (VPTs) in the barrel and endcap regions respectively to measure the energy deposited by the photons given off in the shower. These are photodetectors made of silicon, with a large applied electric field, that amplify and measure the signal. The current increases exponentially, quickly producing measurable currents.

6.4.1 ECAL Barrel

The $612,000$ barrel crystals of the ECAL (EB) are organised as follows: there are 360 crystals in ϕ, organised in 36 supermodules (18 per half barrel) and 2×85 crystals in η direction, mounted in a quasi-projective geometry to avoid cracks aligned with possible particle trajectories.

The barrel crystals measure 0.0174×0.0174 in $\Delta\eta \times \Delta\phi$. The front faces of the crystals are at $1.29\,\mathrm{m}$ radius and cover $|\eta| < 1.479$, shown in Fig. 6.7.

In the barrel, the light from the crystals is collected and amplified by avalanche photodiodes (APDs). There are two $5 \times 5\,\mathrm{mm}^2$ APDs on the back of each crystal. These devices are

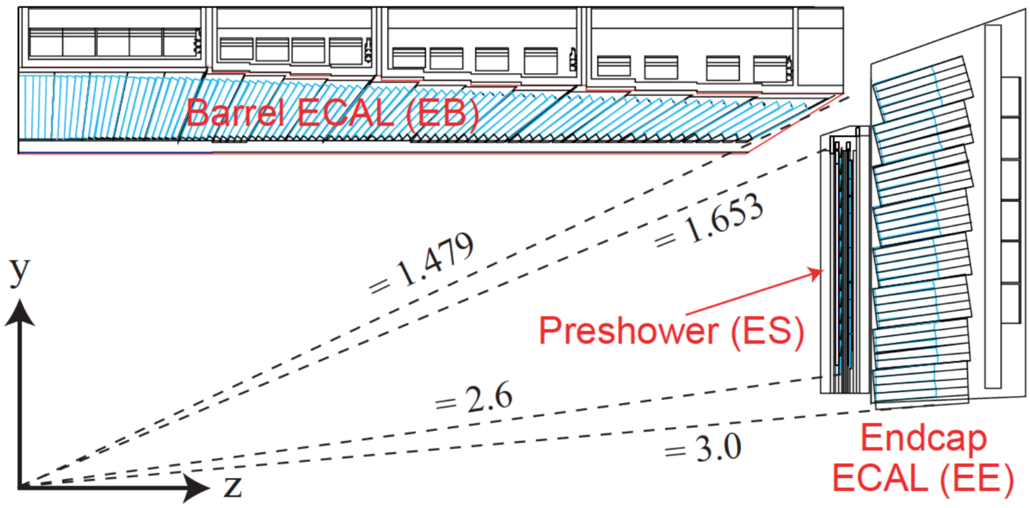


Figure 6.7: Cross section of the CMS ECAL showing the barrel, endcaps and preshower. [12]

photomultipliers, typically composed of doped silicon semi-conductors with a large reverse-bias. Photoelectrons that arrive at the APDs undergo avalanche multiplication with a gain of 50.

6.4.2 ECAL Endcaps

The two ECAL endcaps (EE) extend coverage to $1.479 < |\eta| < 3$ and are located on either side of the barrel, 3.15 m away from the nominal interaction point in the z-direction (shown in Fig. 6.7). The 7,324 crystals in each endcap are grouped into units of 5×5 crystals called superclusters.

The endcap crystals measure 0.01752×0.052 in $\Delta\phi \times \Delta\phi$ and are slightly larger than their barrel counterparts.

The crystals in the EE use VPTs, which are more radiation resistant than silicon diodes. There is one VPT attached to the back of each crystal. The light enters through UV glass with an active area of $280\,\mathrm{mm}^2$ per crystal, and a mean gain factor of 10.2 at $B = 4\,\mathrm{T}$ in the forward regions.

A large electric field is applied across the VPTs. In contrast to the usual two in photodiodes, these contain three electrodes within a vacuum chamber. This generates large current from the very small quantities of scintillation light generated by the lead tungstate crystal. In turn, this signal is converted into a digital signal and sent down the optical fibres.

6.4.3 ECAL Preshower

The ECAL preshower (ES) is a sampling calorimeter with active material consisting of $4,300$ radiation resistant silicon strips organised in two orthogonal planes (similar to the silicon strip panels in the tracker). The ES covers the range $1.653 < |\eta| < 2.6$ in front of each of the endcaps, shown in Fig. 6.7.

Two planes of lead strip panels are used as the absorbing material. These strips sit behind each silicon layer and have a pitch width of 1.9 mm and length of 63 mm. For each of the two layers of lead there is a layer of sensors including read-out electronics.

There is a total $16.5\,\mathrm{m}^2$ of effective area designed to withstand the high radiation levels present in the very forward regions. The silicon strips have dimensions $6.3\,\mathrm{cm} \times 6.3\,\mathrm{cm} \times 0.3\,\mathrm{mm}$ in 32 strips arranged in a circular grid covering most of the endcap. The complete preshower system forms discs approximately 2.5 cm wide, with an outer diameter of 50 cm and an inner diameter of 20 cm to allow for the beam-line.

6.5 The Hadron Calorimeter (HCAL)

In contrast to EM showers, incident hadrons lose the majority of their energy through inelastic collisions with atomic nuclei via the strong force. These interactions typically knock free other hadrons producing hadronic showers. This showering process converts the high energy incident hadron to many low energy particles (all with approximately the same energy). The shower stops when all energy is transferred to the detector through ionisation loss or absorbed through nuclear processes.

The nuclear interaction length λ_{I}, defined in a similar way to the EM radiation length (6.1), is again a material dependent constant. It is typically much longer than the EM radiation length as these inelastic collisions are less probable. Since hadronic showers develop later and have larger longitudinal and lateral dimensions than EM showers, the HCAL is necessarily thicker than the ECAL it surrounds. The HCAL dimensions are restricted by the ECAL to have $r > 1.77\,\mathrm{m}$, and by the inner extent of the magnetic coil at $2.95\,\mathrm{m}$.

In contrast to the CMS ECAL, the HCAL is required to be a sampling calorimeter as there is no material that is both an active scintillator and an absorbing material for hadronic showers. It is composed of alternating active plastic scintillating layers placed interleaved with dense steel and brass layers of absorbing materials. The scintillating layers emit blue-violet light, which is carried down tiny optical fibres of diameter 1 mm that absorb the light from each tile, shifting it to the blue-green spectrum. The optical fibres are spliced to wave-guides composed of high-attention-length clear fibres that carry the light to the photodetecting readout boxes.

Here, the light enters multi-channel hybrid photodiodes (HPDs), consisting of a photocathode held at 8 kV and a standard silicon photodiode. Signals arriving at the HPD receive an amplification of approximately 2000 due to the large electric field. An electronic readout samples each collision, digitises it using special QIE (charge integrated and encode) chips and sends the information to the Data Acquisition System (DAS) for purposes of event triggering.

The HCAL consists of four subdetectors and is organised into the inner and outer barrel (HB and HO), endcaps (HE) and forward (HF) sections. The entire HCAL detector is shown in Fig. 6.8. There are 36 wedges in both the barrel and each of the endcaps, which cover $|\eta| < 3$. These form the last layer of the detector within the magnetic coil. The outer barrel sits outside the coil ensuring no energy that leaks out of the HB goes undetected. The HF is placed at $\pm 11.5\,\mathrm{m}$ in z and covers $3 < |\eta| < 5.2$. Finally, the entire HCAL is mounted on a rigid table.

Since e and γ are also produced in hadronic showers, we want the response ratio for both the EM particles and the hadrons to be roughly equivalent. This motivates the plastic scintillating material used in the HCAL, as it has a well understood response to both the hadronic and EM components of the shower.

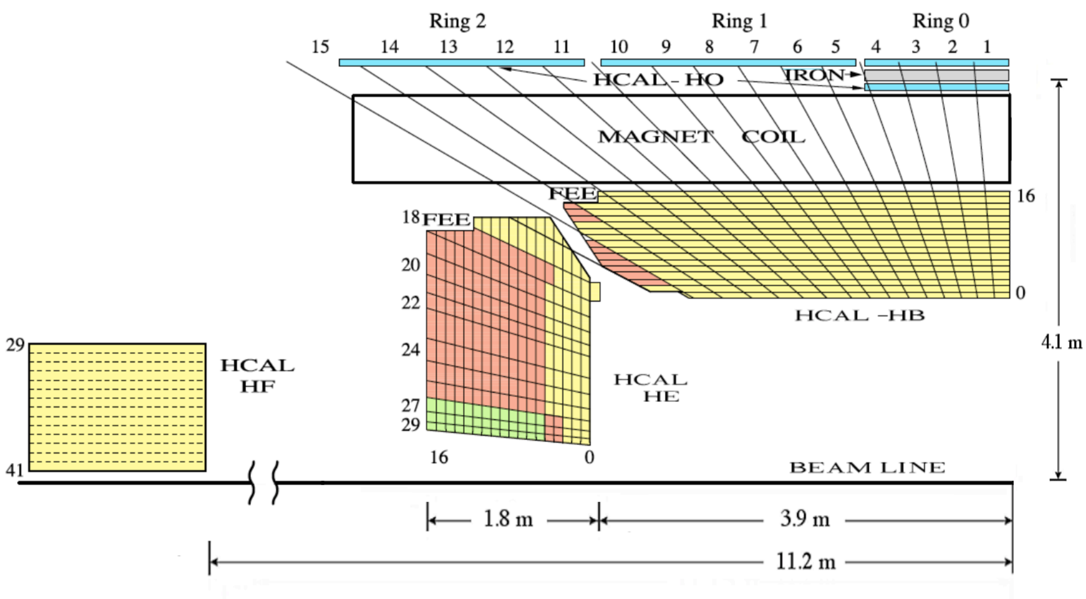


Figure 6.8: Schematic of the Hadronic Calorimeter showing its various components. [117]

6.5.1 HCAL Barrel

The HCAL barrel (HB) is a self-supporting hermetic structure that covers the region $|\eta| < 1.4$. It is composed of passive brass plates interleaved with active plastic scintillator material. The absorber consists of 15 layers of brass scintillator plates with thickness of approximately 5 cm, structurally supported by two external stainless steel plates. The first scintillator plate in the radial direction has a thickness of 9 mm, with all others having a thickness of 3.7 mm.

The total absorbing thickness is $5.82\,\lambda_{\mathrm{I}}$ interaction lengths at $\eta = 0$, increasing to a maximum of $10.6\,\lambda_{\mathrm{I}}$ at $|\eta| < 1.3$. This is extended by about $1.1\,\lambda_{\mathrm{I}}$ at all η when also considering the ECAL.

The HB is segmented into seventeen longitudinal layers measuring 0.087×0.087 in $\Delta\eta \times \Delta\phi$. In the azimuthal direction, there are 36 wedges staggered and segmented in four sectors such that there are no projective dead zones in the full radial extent of the wedge. Layers of tiles are organised into “megatiles” such that the total flux through each is the same.

Optical fibres fit into grooves cut into individual tiles, which carry away the light signal (generated in the plastic scintillators) to the readout boxes.

6.5.2 HCAL Outer Calorimeter

The HCAL outer (HO) detector is located outside the solenoid coil to measure late starting showers that deposit energy after the HB in the region $0 < |\eta| < 1.26$. It uses the solenoid as additional absorbing material and consists of five 2.56 m wide rings placed along the beam-line at $z = 0$, $\pm 2.686\,\mathrm{m}$, $\pm 5.342\,\mathrm{m}$.

The rings are grouped in $= 30^\circ$ ϕ-sectors. The central ring 0, at $\eta = 0$ where the HB has minimal depth, consists of two 10 mm thick scintillating layers on either side of a "tail catcher" iron absorber with thickness 18 cm. These two scintillating layers are placed at radial distances 3.850 m and 4.097 m. The remaining four rings, denoted $\pm 1, \pm 2$, consist of one scintillator layer at 4.097 m, each covering 2.5 m in z. In the radial direction, each HO layer has a total of 40 mm allocated, of which 16 mm is available for the scintillating detector layer. The rest of the space is used for the aluminium honeycomb support structures, itself supported by steel beams.

The total depth of the calorimeter is extended to at least 11 λ_I, with the HO constrained by the geometry of the the barrel structure of the muon system.

6.5.3 HCAL Endcaps

The HCAL endcaps (HE) are attached to the muon endcap yoke. Each HE consists of 14 η towers with 5° ϕ segmentation, covering the region $1.3 < |\eta| < 3.0$. At smaller η, the five outermost layers have finer segmentation of 5° in ϕ and an η-segmentation of 0.087. The remaining eight innermost towers are segmented into 10° in ϕ with η-segmentation between 0.09 to 0.35 at highest η. There are a total of 2304 HE towers and the geometry of the HE is shown in Fig. 6.9.

6.5.4 The Hadronic Forward Calorimeter

The HCAL forward (HF) detector consists of two calorimeters positioned on either side of the interaction point at $z = \pm 11.2\,\mathrm{m}$ covering the regions $2.9 < |\eta| < 5.2$. Each calorimeter is composed of 18 wedges, spanning $\Delta\phi = 20^\circ$, with outer radius 1.30 m and inner radius 12.5 cm to allow for the beam-line.

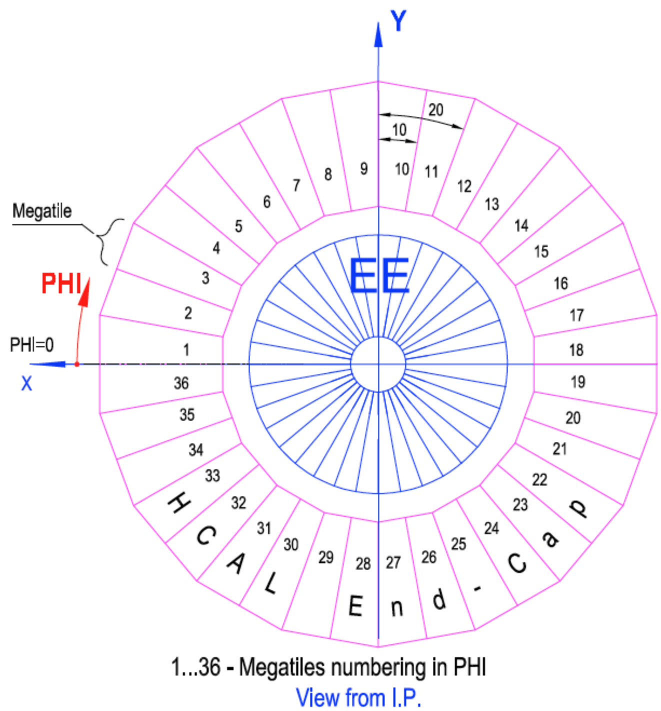


Figure 6.9: Schematic of the Hadronic Calorimeter endcap viewed from the interaction point. [118]

The particle flux is not uniformly distributed and will peak at high η. On average 760 GeV per pp interaction is deposited into the two forward calorimeters, compared to 100 GeV in the rest of the detector. The design of the HF is therefore guided by the necessity for radiation tolerance of the active material.

Radiation resistant quartz crystal fibres are the active scintillators used in this sub-detector. These are placed into grooves in the absorbing steel plates. The fibres provide fast collection of Cherenkov light from charged shower particles and run parallel to the beam-line, grouped in 0.175×0.175 in $\Delta\eta \times \Delta\phi$ towers. Two lengths of fibres are used: longer cables, running the entire HF of length 1.65 m, and shorter fibres that are only 22 cm. These are read out separately and can be used to distinguish electromagnetic showers from hadronic ones, as the former deposits a much larger fraction of its energy in the first 22 cm.

The light from polythene scintillators travels along the crystal into an air-core light guide, which consists of a hollow tube lined with highly reflective custom-made sheets, coated in metal reflectors. Light reflects approximately five times with 50% light loss, before it reaches the readout boxes.

6.6 The Muon System

The muon subdetector has three major functions: it provides muon identification, momentum measurement and triggering information. All three are important and give powerful signatures to many different analyses, including the Standard Model Higgs $H \to 4\mu$ decay.

The design of the muon detector is motivated by the desire to provide excellent resolution for high momentum muons ($p_T \gtrsim 100\,\mathrm{GeV}$), whilst keeping resolution to approximately 5 % up to 1 TeV (below this threshold the resolution is dominated by tracker information).

Further, since the reconstruction of tracking information is not fast enough at the Level 1 trigger, this subdetector provides the only mechanism to trigger on muons.

Since muons are minimally ionising particles, they are expected to traverse all inner subdetectors. The muon systems, consisting of a barrel section and two planar endcap regions, are therefore placed at the edge of the experiment. These cover the range $|\eta| < 2.4$ and form concentric cylinders around the solenoid. The solenoid's return yoke also serves as a hadron absorber for muon identification.

CMS uses three types of gaseous particle detectors for muon identification; drift tubes (DT), cathode strip chambers (CSC) and resistive plate chambers (RPC). The locations of these chambers are shown in Fig. 6.10. There are a total of 1400 muon chambers, of these 250 are DTs, 540 CSCs, and 610 RPCs. The combined barrel and endcaps weigh approximately $12,500$ tonnes and make up approximately $25,000\,\mathrm{m}^2$ of detection plane. The muon system measures approximately 15 m in length, width, and height.

6.6.1 Drift Tubes

The drift tubes (DT) chambers are used for measurements and triggering in the central regions $|\eta| < 1.2$ and form four concentric cylinders around the beam-line shown in Fig. 6.11.

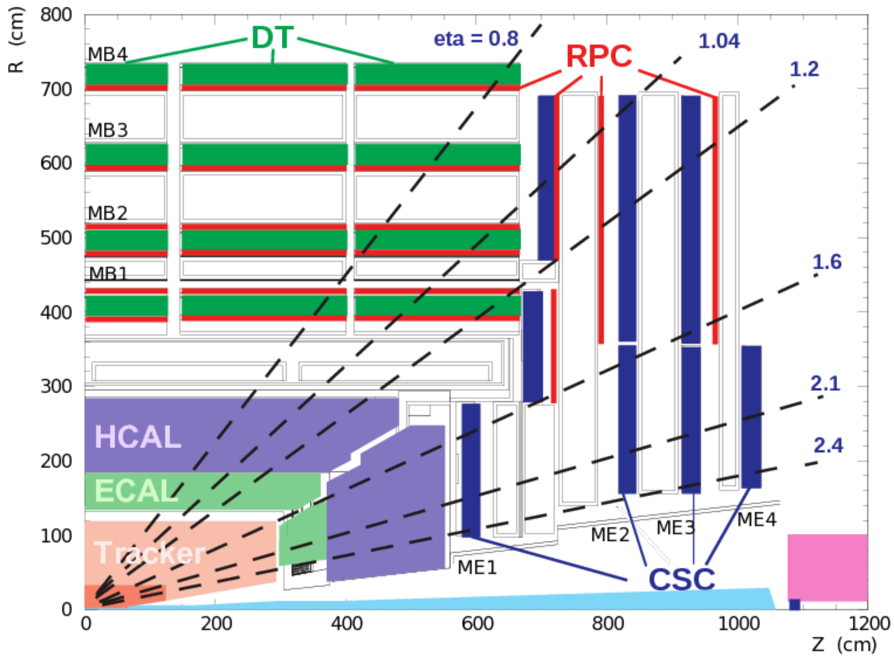


Figure 6.10: Schematic showing the muon systems showing the locations of the drift tubes (DT), cathode strip chambers (CSC) and resistive plate chambers (RPC). [12]

Each drift chamber is on average $2\,\mathrm{m} \times 2.5\,\mathrm{m}$, consisting of eight or twelve drift cell layers grouped into two or three "superlayers", each 4 cm wide.

Individual drift tube systems are composed of a mixture of 85% Ar + 15% CO_2 and a wire held under high voltage. The gas in the DTs is ionised by incident muons when they pass through. These ions *drift* toward the central wire, giving rise to a measurable current used to reconstruct the paths of muons. The length of the wire is 2.4 m and the tube has a transverse dimension of 42 mm, shown in Fig. 6.12. The DT system uses a total of 172, 000 wires.

6.6.2 Cathode Strip Chambers

The Cathode Strip Chambers (CSC) operates well in the large inhomogeneous magnetic field and provide both precise muon measurements and muon triggering information.

There are seven trapezoidal chambers that measure up to 3.4 m long and 1.5 m wide. Each covers either 10° or 20° in ϕ. Almost all the chambers overlap providing continguous ϕ coverage in the muon systems up to $|\eta| < 2.4$. The layout of the CSCs is shown in Fig. 6.13.

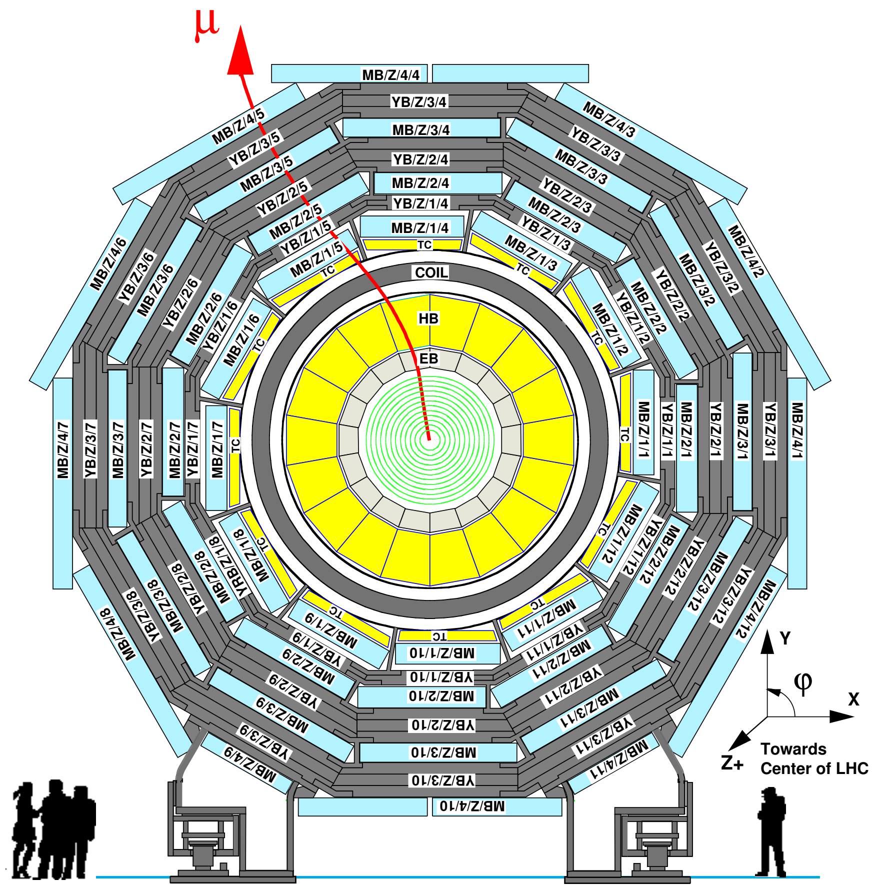


Figure 6.11: A transverse view of the layout of the muon chambers. The DT chambers are shaded in aqua, with the interleaved iron return yolk layers shaded in grey. [12]

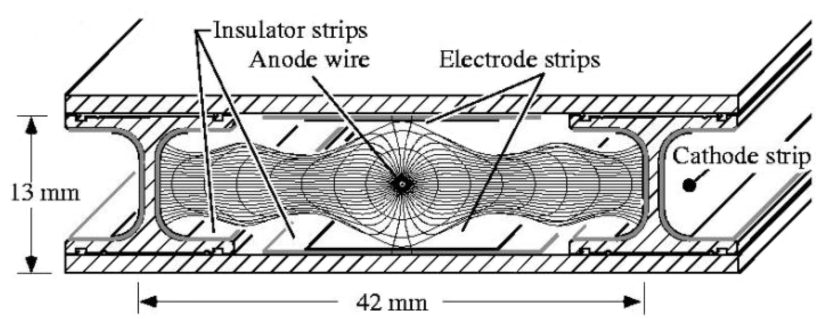


Figure 6.12: A transverse view of a muon DT. [119]

chr

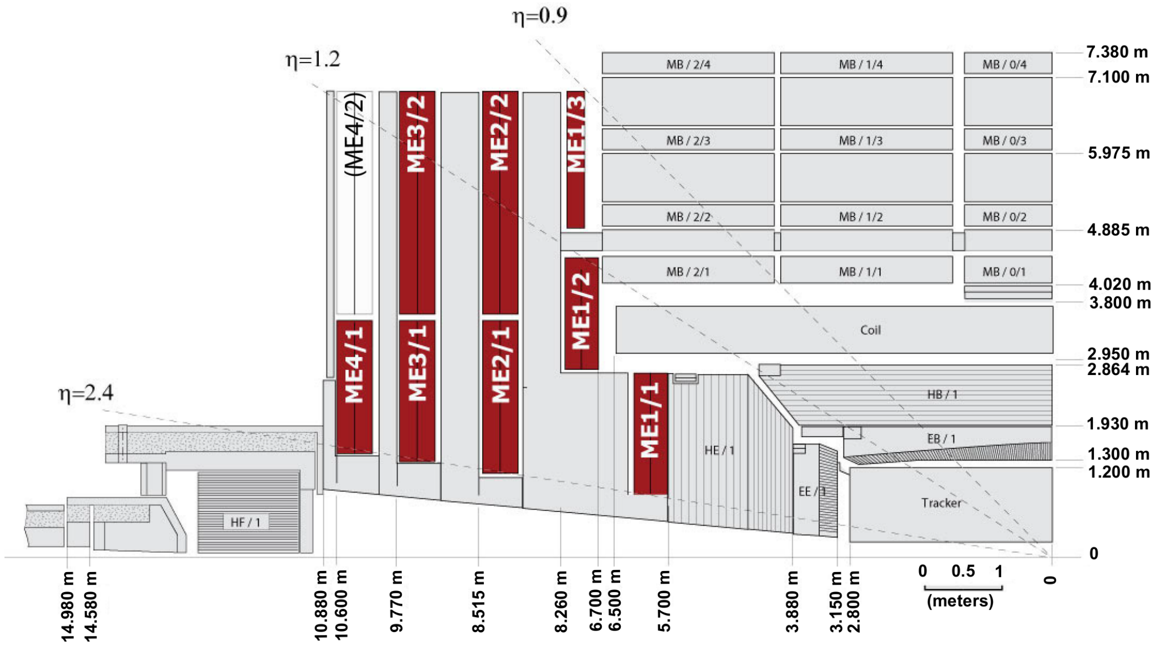


Figure 6.13: The layout of the CSCs, which are highlighted in red. [12]

Each CSC is a multiwire proportional chambers, composed of an array of six positively charged wires interleaved with seven perpendicular running copper strips cathodes separated by a gas volume. These gas gaps are filled with a similar mixture to that in the DTs.

The wire in the chamber runs azimuthally and is used to measure the radial coordinates of muons. In a similar way to the DTs, muons ionise atoms producing electrons that accelerate toward the anode. The positive ions in the gas travel in the opposite direction toward the cathodes, which in turn induces a current in the strips perpendicular to the wires, which can be measured. This procedure gives two position coordinates for each traversing muon. A transverse view of the CSCs is shown in Fig. 6.14.

In total, there are 468 CSCs, with over 2 million CSC wires. A muon in the region $1.2 < |\eta| < 1.4$ crosses either three or four CSCs. In the endcap-barrel overlap $(0.9 < \eta < 1.2)$ muons are detected by both DTs and endcap CSCs.

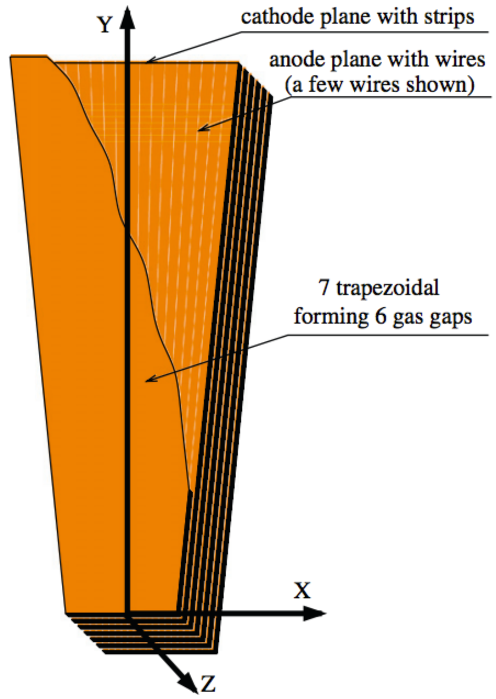


Figure 6.14: A transverse view of the cathode strip chambers. [12]

6.6.3 Resistive Plate Chambers

There are six cylindrical layers of RPCs in the barrel muon system. There is one layer of RPCs in each of the outer two barrel stations and two in each of the inner two stations. In the endcaps, there is one plane of RPCs, which forms a ring in each of the first three stations (providing coverage up to $|\eta| < 1.6$).

The RPCs have good spatial resolution and a very small time resolution of just one nanosecond. They form a redundancy of the muon triggering system, operating in parallel with the DTs and CSCs to allow trigger algorithms to distinguish the proton bunch crossings when triggering on muon candidates.

There are 480 rectangular RPCs in the barrel and a total of 132 trapezoidal RPCs in the endcaps. Each RPC is composed of two parallel oppositely charged electrodes separated by a 2 mm gap filled with $C_2H_2F_4$ gas. The electrodes are planes of plastic with extremely high resistivity, which allows a large potential difference to be used.

The passing muons, which freely traverse the resistive plates, ionise the gas atoms. The liberated electrons are accelerated by the high voltage causing an avalanche of electrons, which in turn pass through the electrodes to be measured as current by the external detecting aluminium strips. A transverse view of the barrel RPCs is shown in Figure 6.15.

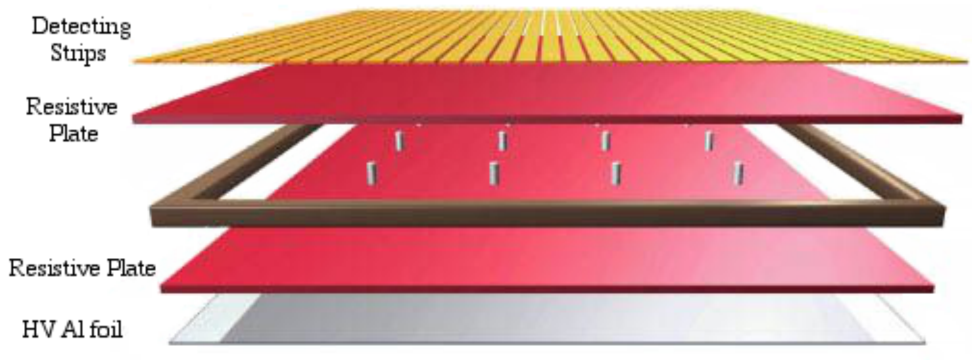


Figure 6.15: Schematic of an RPC. [12]

The pattern of hit strips gives a quick measure of the muon momentum, which is then used by the trigger to make immediate decisions about whether the data should be kept.

6.7 Triggers

The LHC beam consists of protons bunched together into 2,808 bunches with 115 billion protons per bunch. "Bunch crossings" between two beams occur a minimum of 25 ns apart, where the set of interactions that occur at each bunch crossing is known as an "event".

At nominal design luminosity, approximately 20 pp interactions occur superimposed at each bunch crossing of the beam, an effect known as "pileup". The interaction rate is approximately 10^9 Hz, generating 1 Mb per event for zero-suppressed data. This quantity needs to be reduced by $\mathcal{O}(10^6)$ to be feasible for readout and storage.

This is the start of the event selection process and consists of two steps: a hardware Level 1 (L1) trigger and a software High Level Trigger (HLT). Ideally, the trigger systems are designed to accept physically interesting events. To store a subset of potential events that pass a given trigger, a concept known as "prescaling" is used; this randomly selects a specified fraction of passing events, allowing a specific trigger to be kept with reduced bandwidth.

The maximum possible bandwidth out of L1 is approximately 10^5 Hz, corresponding to the input capacity of the High Level Trigger (HLT) processor farm. The output of the HLT farm is on the order of 100 Hz corresponding to the limit that can be processed by the online computer farm, where complicated software algorithms can be run.

Due to the very small bunch crossing intervals, the next event often occurs before the particles from the previous one have left the detector. The data therefore needs to be stored in pipelines that concurrently retain information for many interactions. Further, all subdetectors must also have excellent temporal resolution in order to distinguish between bunch crossings.

6.7.1 Level 1 Trigger

The Level 1 (L1) trigger is a custom designed programmable hardware electronic system. It has local, regional, and global components that consists of the calorimeter trigger, the muon trigger, and the global trigger. The global trigger takes information from the calorimeters and muon systems and ultimately decides whether or not to pass the event to the HLT. It is extremely fast and wholly automatic.

Most of the L1 electronics are installed 90 m away in the counting room adjacent to the experimental cavern that houses the detector. However, some parts of the L1 trigger are mounted directly onto the subdetectors themselves, making them inaccessible during LHC running.

The L1 analyses every bunch crossing with an allowed "latency" of $3.2\,\mu$s. This is the permissible time between a bunch crossing and the distribution of the trigger decision to the detector front-end electronics. Due to this relatively large latency, the data is pipe-lined by storing it in the front end modules in order to remove as much dead-time as possible.

If the L1 returns positive, the information for the event is read out through long optical links and buffered in a Builder Unit (BU). The BU assigns the information of the event through a filter farm to a single processor of a filter unit (FU) where HLT algorithms are performed.

6.7.2 High Level Trigger

The High Level Trigger (HLT) uses highly flexible software algorithms to further reduce the rate of information output. HLT has a complete readout of data; therefore it can perform complicated calculations similar to those done at offline software level. The average HLT processing time for each event is 50 ms but can be up to a few seconds for complicated events.

There is typically a list of HLT triggers that are concurrently deployed. If an event passes any such HLT, then the event is sorted into a distinct dataset for further processing. The trigger menu allows analysers to select analysis specific events.

The cumulative rate for the entire menu needs to be maintained at $\mathcal{O}(100)$ Hz, which is the limit of the computational capabilities. Information that passes the HLT is then transferred and saved in databases around the world in the computing grid. Offline, this information is reconstructed by specifically designed software, building particle candidates for analyses.

6.8 The Computing Grid

After the L1 and HLT decisions return a positive result, the information of the event is passed onto the LHC computing grid. If we assume an average 1 Mb per event, then, for one effective LHC year of 10^7 seconds, approximately 5 Petabytes of data are produced and needs to be stored at design luminosity. This huge amount of data is sent through a novel distributive computing system and data storage infrastructure called the Worldwide LHC Computing Grid (WLCG). This is composed of tens of thousands of computers around the world organised for data processing and storage, giving worldwide access to scientists.

The LCG is organised as follows. The CERN Tier 0 centre reconstructs events fully. Once stored and once a backup is made, it is sent to 7 Tier 1 locations situated around the world in France, Germany, Italy, Spain, Taiwain, UK and the US. Events are reconstructed again using additional information provided by the experiment to allow finer calibration.

The Tier 1 then sends the most complicated and most interesting physics data to 40 Tier 2 sites. Data is analysed at each step along the way, branching out in this way till physicists can access important data relevant to their specific analysis on a local level.

6.9 The CMS Cavern

The CMS cavern at point 5 of the LHC near Cessy was excavated from an old Large Electron Positron (LEP) Collider access point. The experimental cavern houses the CMS detector and is located 100 m underground, measuring 53 m long, 27 m wide and 24 m high. The onion-like cylindrical CMS detector was originally divided into 15 wedges, which were each tested before being separately lowered into the cavern.

Chapter 7

Physics Objects

The CMS detector is designed for full event reconstruction. In this section we shall describe the reconstruction of all physics objects.

The Kálmán Filter (KF) algorithm is a least squares fit that forms the basis of most CMS reconstruction algorithms; it is described in detail in [120]. Further details of the different physics objects may be found in Ref. [121].

7.1 Vertices

Precise measurement of the primary event vertices (PV) is the first step of event reconstruction. It allows us to assign reconstructed "tracks" (of particle trajectories) to collisions, precisely determining their kinematics, whilst at the same time allowing the identification of long-lived particles that decay at secondary vertices. The distance between primary and secondary vertices is known as the "impact parameter".

Online track reconstruction starts from track candidates, which are a triplet of "hits" where particles interact with the pixel detector. This allows fast determination of the z-coordinate of the primary vertex with a resolution of 50 μm. This information is used in the HLT for triggering.

Offline track reconstruction first requires the selection of tracks that are compatible with the beam-line. These tracks are then clustered so that a vertex can be fitted.

A modified least squares fit known as the Trimmed Kálmán Vertex Filter (TKF) gives the best possible vertex estimate. This algorithm provides the locations of the fitted vertices, as well as a χ^2 probability that indicates the compatibility of the reconstructed tracks with the fitted vertex. Outlier tracks with the highest contribution to the vertex fit χ^2 are removed iteratively if the corresponding χ^2 probability is less than 5%. In this way, we can determine both the position of the vertex and the tracks that come from it.

Although there are a large number of expected vertices in each bunch crossing, there is typically only one vertex that contains the signal event of interest. In general, the vertex with the highest sum of the transverse momenta of all jets will be the most interesting one [122].

7.2 Pileup

During each bunch crossing at the LHC, multiple *pp* collisions can occur, resulting in numerous reconstructed PVs. This effect is known as "pileup" (PU). The expected pileup at the LHC is approximately 20 PVs dependent on the instantaneous luminosity [11]. PU affects the reconstruction of all physics objects and it is corrected in two different ways:

- Charged Hadron Subtraction (`CHS`): also known as `PFnoPU`, attempts to filter all charged hadrons that do not originate from the primary interaction vertex. This technique is very effective but can only be implemented in the pseudorapidity region covered by the tracker.

- `FASTJET`: is a standalone software package, use to calculate the momentum density per unit area ρ due to PU for each event. Reconstruction of jets and lepton isolation cones can then be "ρ-corrected", by subtracting contamination based on their respective areas.

7.3 Tracks

The trajectory of particles through the detector is reconstructed as "tracks". The reconstruction makes use of the combinatorial KF algorithm that integrates track fitting and pattern recognition.

Energy signatures in individual tracker layers form seeds that are extrapolated from layer to layer, accounting for multiple scattering, energy loss and the magnetic field. Starting from the inner pixel layers, a combinatorial exploration to build all possible tracks is done in parallel, limited by a χ^2 and a maximum number of missing hits. Hits that are unambiguously assigned to tracks are removed iteratively. The final collection of tracks is cleaned to remove duplicates.

7.4 Particle Flow

The Particle Flow (`PF`) algorithm aims to reconstruct all stable particles in an event by combining information from all subdetectors, this includes: charged tracks in the tracker and energy deposits in the electromagnetic and hadronic calorimeters, as well as signals in the preshower detector and the muon system [123].

The `PF` algorithm is designed to determine the direction, energy, and type of the stable particles in the event; this includes muons, electrons, photons, charged and neutral hadrons. Once all objects have been reconstructed, the list of particles are clustered to build jets, which are then used to determine more complicated quantities such as the missing transverse E_T, from neutrinos or other exotic invisible particles, and to identify taus from their decay products. The energy calibration is performed separately for each type of particle.

The basic ingredients of the `PF` algorithm are the tracks, formed in the tracker using the combinatorial KF algorithm, the calorimeter clusters, and a link algorithm.

The calorimeter clustering algorithm is performed in both the ECAL and HCAL barrel and endcap separately. Local energy maxima in calorimeter cells are identified as "cluster seeds", which are grown iteratively by aggregating cells with neighbouring ones that have energy deposits. These form the calorimeter seeds of `PF` candidates.

A given particle will likely give rise to several `PF` elements in the various CMS sub-detectors. These elements are connected by the link algorithm. A tentative link is made for each pair of elements with an extrapolation between the sub-detectors. If the extrapolation of the track is within a certain distance to the cluster boundaries the extrapolation continues. The clustering algorithm finds the best track candidates, cleans the spurious seeds, and removes all double counting of track elements.

From the list of track candidates the `PF` algorithm proceeds as follows. The "global muons", which are identified as tracks in the tracker that have extrapolated hits in the muon chambers, are removed from the list of considered particles if the muon's momentum is within three standard deviations of the momentum reconstructed solely from the tracker. This removed collection is known as the `PF muons`.

Electrons are pre-identified by their relatively short tracks and their energy loss by Bremmstrahlung in the tracker layers on their way to the calorimeter. A final identification is made using measurements from both the tracker and ECAL. Each identified electron gives rise to a `PF electron` and its track with associated seeds are removed from the collection.

The remaining tracks may come from neutral hadrons, charged hadrons, photons or in rare instances, additional muons. The neutral hadrons and photons, which have straight trajectories, can be identified from their tracks by comparing the relative momentum and energy deposited in the ECAL and HCAL using a multivariate analysis. These form the collections of `PF photons` and `PF neutral hadrons`. The remaining tracks are checked against a looser muon identification. These are removed, with the last tracks giving rise to `PF charged hadrons`, whose momenta and energy can be taken from track momentum measurements.

The final list of `PF` particles can then be passed to reconstruct jets, the missing transverse energy E_T, and to identify and reconstruct taus.

7.5 Electrons

Electron reconstruction proceeds through a combination of track reconstruction in the ECAL and the tracker [124]. The main algorithm used is the ECAL-driven method. Electron seeds are made from "superclusters" of energy deposited by electron candidates and their Bremsstrahlung radiation.

Due to the magnetic field, the electrons travel along a curved path, radiating photons along the way. These photons reach the ECAL spreading out in ϕ (but not in η). Superclusters of the showering particles in the ECAL are then made by summing energy deposits that are separated only in ϕ.

The superclusters form seeds of the electron tracks, which are extrapolated backwards through the magnetic field towards the pixel detector. Since electrons only deposit a small portion of energy there, the extrapolation is matched to loosely identified electron tracks in the tracker. The extrapolation uses a generalised KF algorithm known as the Gaussian-Sum Filter (GSF) [125]; this accounts for the energy loss from Bremsstrahlung in the silicon tracker layers.

To distinguish electrons from hadronic activity, several selection criteria are imposed. The width of the supercluster in the η direction is required to be small. Further, the electron candidate is required to deposit substantial energy in the supercluster ($E_T > 15$ GeV), with the total ratio of the energy deposited in the ECAL much larger than that in the HCAL.

Finally, since an electron can be mimicked by a combination of a muon (which provides the track) and hadrons in the ECAL, a $\Delta R > 0.1$ cut is imposed between electron and muon candidates. The exact recommendations for the selection is analysis dependent and is discussed in Ref. [126].

7.6 Muons

Muon reconstruction at CMS is done using a combination of the information collected by the tracker, and the dedicated muon systems, which provide a very clean signature for reconstruction.

Muon candidates in CMS are identified by multiple reconstruction algorithms using hits in the silicon tracking system and signals in the muon system.

Since muons are typically the only particles reaching the muon chambers, the "standalone" muon reconstruction algorithm starts from track segments detected in the innermost layers of the muon chambers. From track segments, additional neighbouring hits are added by the KF technique taking into account the inhomogeneous magnetic field using a recursive Runge-Kutta procedure [127]. Finally, muon tracks are propagated back to the interaction point.

A second algorithm starts from the standalone muon tracks, which are extrapolated towards the nominal vertex from the innermost muon chamber hit. The extrapolations are then matched up to tracks, reconstructed by the tracker, forming the "global muon" collection. This track matching is performed by defining an $\eta - \phi$ rectangular region around standalone muon candidates and selecting the most compatible track in the tracker. This procedure works well only if high quality muon tracks are available in the muon detector.

However, in certain instances hit information in the muon system is minimal and standalone muon reconstruction fails. For example, muons with low momenta $6 - 7\,\mathrm{GeV}$ will typically have too few hits to be reconstructed as standalone muons. In such situations, the final reconstruction algorithm, which gives rise to the "tracker muons", is used. This starts from tracks found in the tracking system and associates them to compatible signals in both the calorimeters and the muon chambers. These require only one short segment of hits in the muon system.

The specific algorithm used is analysis dependent.

7.7 Taus

Due to their large mass, the taus often play important roles in SM and BSM analyses. Approximately two-thirds of taus decay hadronically into one or three charged hadrons in about 75% and 25% of the time respectively. The remaining decay width is to the lighter leptons, that are detected as such. The branching ratio into neutral hadrons is on the order of 1%.

`PF` algorithms provide excellent reconstruction of energy and direction of all visible tau decay products. The reconstruction takes in the track collection from the `PF` algorithm and proceeds in one-prong and three-prong topologies. The four-momenta of all particles in a small cone $\Delta R < 0.15$ around the direction of the highest p_T particle is summed to create tau candidates. A boosted decision tree is then used to determine “loose”, “medium” and “tight’ working points to discriminate against muons [128].

7.8 Missing Transverse Energy

During collisions at CMS, the longitudinal momentum along the beam line is not well defined since protons are composite particles and their total energy is unevenly distributed amongst their constituents. Further, many final state particles close to the beam-line are not reconstructed as they do not pass through any detector material, thus carrying away significant momentum.

However, the total transverse momentum of pp constituents is negligible. By measuring all final state visible particles, we can construct the missing transverse energy $\not{E}_T$ from momentum conservation. Within the SM, such missing energy is indicative of neutrinos. In BSM theories, $\not{E}_T$ often characterises a potential signature for weakly interacting particles that do not interact with the detector.

$\not{E}_T$ is reconstructed by the PF algorithm from the visible tracks in the event. The PF algorithm forms the vector sum of the transverse momentum of all reconstructed particles. $\not{E}_T$ is then defined as the modulus of this vector in the opposite direction:

$$\not{E}_T = -\sum_i p_T \qquad \text{i} = \texttt{PF} \text{ objects} \tag{7.1}$$

7.9 Jets

Due to confinement, quarks and gluons are never directly detectable. They spontaneously hadronise and form a collimated spray of hadrons known as “jets”. These are detected as energy signatures in multiple layers of the detector.

All identified hadrons (and their tracks) found by the PF algorithm are clustered into PF jets by using each particle’s direction at the interaction vertex, forming a cone of predefined size. An example of the energy deposited in the detector by jet candidates is shown in Fig. 7.1.

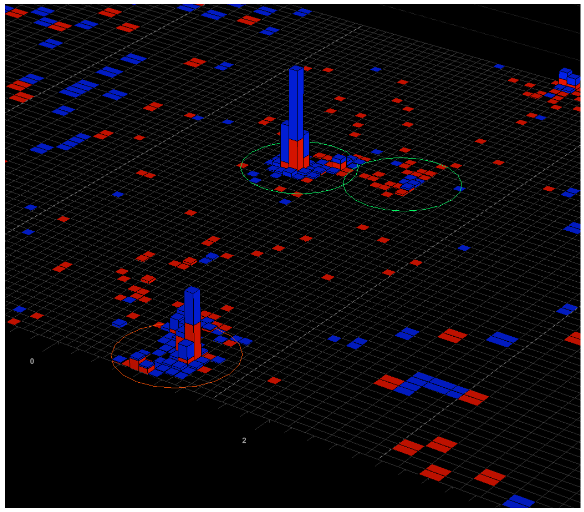


Figure 7.1: Energy deposits of jet candidates in the (red column) ECAL and the (blue column) HCAL are clustered using (green circle) anti-k_t and (orange circle) CA8 algorithms. [129]

The anti-k_t [130] and Cambridge/Aachen [131] inclusive jet finding algorithms are the two main sequential clustering algorithms used in CMS to group neighbouring tracks found in jets. The former is used to cluster `PF` particles into `PFJets`, whilst the latter has additional applications relevant to boosted, high-energy, topologies [132].

Both anti-k_t and Cambridge/Aachen algorithms use a sequential jet clustering algorithm that groups particles by minimising distances between neighbouring tracks. The algorithm looks for the two neighbouring particles and merges them if they are closer together than each individually is to the transverse plane of the beam-line.

The two momenta from the two tracks are added together and treated as one single particle for the rest of the procedure. This process continues iteratively, summing all particles into groups. If the nearest neighbour particles are farther apart than the distance to the beam-line, the particle closer to the beam-line is removed and designated a jet. This continues until all hadronic tracks have been grouped into jets.

The differences between the anti-k_t and Cambridge/Aachen algorithms is the weighting given to neighbouring particles.

On the one hand, the anti-k_t algorithm preferentially merges tracks with higher transverse momentum with respect to its nearest neighbour. This algorithm gives roughly conical jets, that are circular in the $\eta \times \phi$ directions. A typical maximum cone-size of $\Delta R = 0.5$ is used, where $\Delta R = \sqrt{\eta^2 + \phi^2}$, as implemented in FASTJET version 3.0 [133]. These jets are denoted ak5-jets and correspond to the conventional jets used in the majority of physics analyses. Unless otherwise stated, all jets shall be assumed to be ak5-jets.

On the other hand, the Cambridge/Aachen algorithm calculates which tracks to merge based solely on the distance information. This allows larger jets to be constructed with roughly equal momenta subjets. For highly boosted initial objects, tracks from different decay products may be collimated such that they are reconstructed as the same jet and identified as coming from one particle. A larger maximum cone-size of $\Delta R = 0.8$ is used to identify such jet candidates, and information of its constituent subjets is stored to allow easy identification of the various daughter particles. These jets are denoted CA8 jets.

A jet energy calibration (JEC) [134] is typically applied to account for the non-linear, non-uniform response of the CMS calorimeters in all jets. Since most jet constituents are identified and reconstructed with nearly the correct energy by the PF algorithm, only small jet energy corrections need to be applied to each jet. These corrections are between 5% and 10% of the jet energy and were obtained as a function of p_T and η from the GEANT4-based CMS Monte Carlo simulation [135] and early collision data.

7.10 Identification of b-jets

b-quarks are produced in many interesting signal events and can be used to discriminate signal from background. They differ significantly from the light flavour quarks: b quarks have mass $m_b = 4.2$ GeV, lifetime $\tau \approxeq 1.5$ ps, equivalent to 1.8 mm at 20 GeV, and decay predominantly into c-quarks and leptons.

The resulting energy signatures are clustered into anti-k_t jets which are identified by a "b-tagging" algorithm, which can distinguish b from non-b jets. The Combined Secondary Vertex (CSV) algorithm is the main algorithm used by CMS. CSV does a multivariate analysis on the information from tracks, impact parameters, and secondary vertices, giving a b-tagging discriminant value for each ak5-jet. Standard working points "loose", "medium" and "tight" thresholds are then defined by the acceptance rates of light flavour jets using simulation (10%, 1% and 0.1% respectively) [136].

7.11 Identification of W-jets

To reconstruct particles that come from a decay W-boson ("W-jets"), the Cambridge-Aachen algorithm is used with a $\Delta R = 0.8$ cone. These jets are required to have transverse momenta $p_T > 200\,\mathrm{GeV}$ and pseudorapidity $|\eta| < 2.4$.

A jet pruning algorithm [137,138] is used to identify the hadronic decay of the W-boson. This uses a reverse clustering algorithm to construct subjets within the cone of the W-jet candidate. W-jets are required to have exactly two such daughter jets, with a total invariant mass of $60 < m_{\mathrm{jet}} < 130\,\mathrm{GeV}$.

One can also consider imposing an added "mass-drop" condition, which requires each jet to represent a significant portion of the W-jet invariant mass. If applied, the exact value of this cut needs to be studied on a case by case bases.

7.12 Identification of Top-jets

To reconstruct jets that come from top-quarks, "top-jets", the Cambridge-Aachen algorithm is used with a $\Delta R = 0.8$ cone [137,138] . These jets are required to have transverse momenta $p_T > 250\,\mathrm{GeV}$ and $|\eta| < 2.4$. Top-tagged jets are also required to have a total invariant mass $140 < m_{\mathrm{jet}} < 250\,\mathrm{GeV}$ and at least 3 daughter jets (found by a reverse clustering algorithm similar to that in W-jets), with a minimum pairwise invariant mass of the daughters of $50\,\mathrm{GeV}$.

7.13 Event Displays

After full reconstruction with the `PF` algorithm, an event can be diagrammatically displayed. Fig. 7.2 shows events 25787738 and 275039437 recorded during 2012. These events correspond to one with an electron and a muon, and one with two muons.

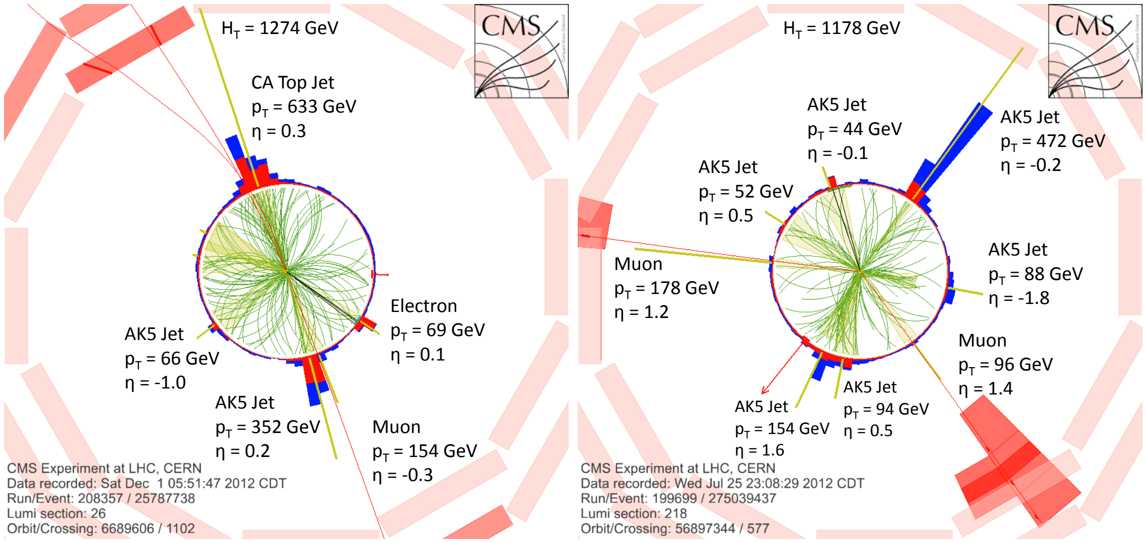


Figure 7.2: Display of the reconstructed objects in event 25787738 and 275039437 recorded by CMS in 2012. The green lines correspond to reconstructed charged tracks in the silicon tracker. These are clustered using both ak5 and CA8 jet algorithms, with cones shown in yellow. The red and blue bars indicate the energy deposited in the ECAL and the HCAL respectively. A red line indicates the reconstructed trajectory of a muon candidate. [129]

7.14 Luminosity

The design instantaneous luminosity is $\mathcal{L} = 10^{34}\,\mathrm{cm}^{-2}\mathrm{s}^{-1}$. At this luminosity, bunches of protons are accelerated to centre-of-mass energies of $\sqrt{s} = 14\,\mathrm{TeV}$, with 25 ns spacings, giving a collision rate of 40 MHz. The total pp cross section at design intensity is expected to be $\sim 100\,\mathrm{mb}$.

The expected rate of events for any process at the LHC is given by the relation:

$$N = \epsilon\sigma\mathcal{L} \tag{7.2}$$

where ϵ is the efficiency of data being recorded, σ[pb] is the cross section of the process and $\mathcal{L}[\mathrm{fb}^{-1}]$ is the instantaneous luminosity. Cross-section is process dependent but instantaneous luminosity can be determined purely in terms of the beam parameters, shown in Table 7.1 for design luminosity.

Parameter	Meaning	Value
n_b	Number of proton bunches per beam	2808
N_b	Number of protons per bunch	1.15×10^{11}
f_{rev}	Frequency of revolution	11.245 kHz
γ_r	Relativistic gamma of the protons	7461
ϵ_n	Normalised transverse beam emittance	3.75 μm rad
β^*	Beta function at the interaction point	55 cm
θ_c	Beam crossing angle	285 μ rad
σ_z	RMS longitudinal bunch length	7.55 cm
σ^*	RMS transverse beam size	16.7 μm

Table 7.1: LHC beam parameters and their design values. [10]

Luminosity Determination

The recorded luminosity at CMS provides the overall normalisation for the expected number of events in data through (7.2). It also allows the real-time monitoring of CMS's performance.

The instantaneous luminosity is determined from signal in the Forward Hadronic calorimeter. CMS uses two methods for online luminosity determination [139]: tower occupancy and total E_T. Both methods are used to determine the mean number of interactions per bunch crossing μ, which is related to luminosity $\mathcal{L}$ by

$$\mu = \frac{\sigma\mathcal{L}}{f} \tag{7.3}$$

where f is the bunch crossing frequency, σ the total (measured) pp cross section and the number of interactions is Poisson with mean μ.

The tower occupancy method is based on a "zero counting" procedure; the average fraction of empty towers is used to infer the mean number of interactions per bunch crossing. For a noiseless calorimeter, the average fraction of empty towers $\langle f_0 \rangle$ is

$$\langle f_0 \rangle = e^{-(1-p)\mu} \tag{7.4}$$

where p is the probability that a tower is not hit in a single bunch crossing.

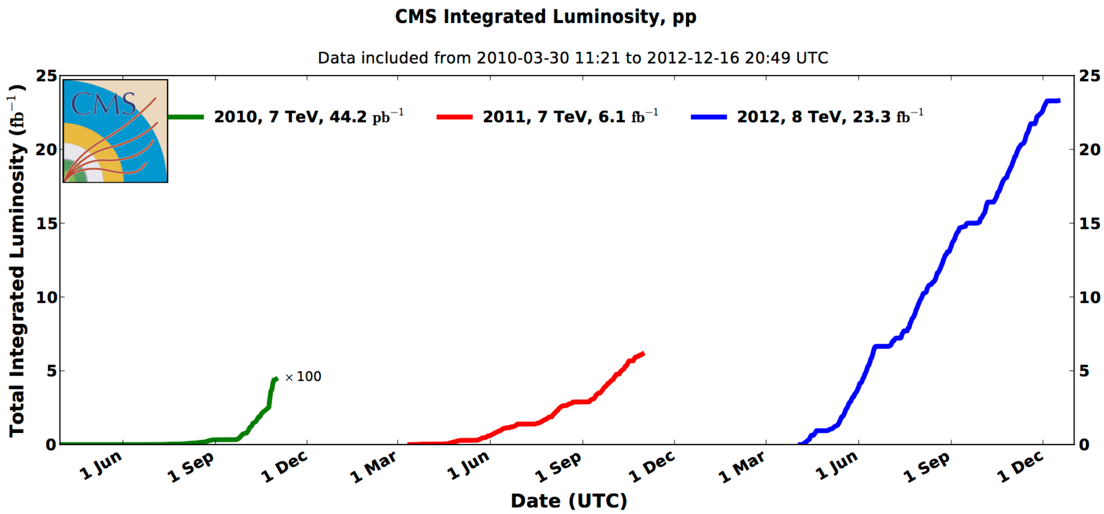


Figure 7.3: The delivered integrated luminosity at the LHC in the 2010-2012 data taking period. [140]

Accounting for noise, this equation becomes

$$-\ln\langle f_0\rangle = (1-p)(1-\epsilon)\mu + N \tag{7.5}$$

where ϵ is a non-linear correction and N the noise offset, both determined through Van der Meer calibration.

Similarly, the E_T measurement method exploits the relationship between the average transverse energy $\langle E_T\rangle$ per tower and the mean number of interactions per bunch crossing:

$$\langle E_T\rangle = v_s\,(1-p)\,\mu + N \tag{7.6}$$

where v_s is the average energy of a single bunch crossing. Both methods agree within uncertainties.

In the first year of data-taking $32\,\mathrm{pb}^{-1}$ of data was recorded at 7 TeV centre-of-mass energy by the CMS detector. In 2011, an additional $5\,\mathrm{fb}^{-1}$ of data was recorded at 7 TeV and finally in 2012, $19.6\,\mathrm{fb}^{-1}$ of data was recorded at 8 TeV. The cumulative amount of "delivered" luminosity by the LHC is shown in Fig. 7.3. The difference between the delivered and the recorded luminosity is due to detector dead-time when CMS was unable to log data.

Part IV

The Fourth Generation

Chapter 8

The Fourth Quark Generation

The Standard Model (SM) of particle physics imposes no constraints on the number of quark generations. Indeed, adding a sequential fourth generation provides a minimal extension to the SM and would allow the resolution of the matter-antimatter asymmetry issue.

We consider a framework where we trivially enlarge the SM by a complete sequential fourth generation of chiral matter $(Q_4, u_4, d_4, L_4, e_4)$. The Yukawa couplings are given by

$$\mathcal{L}_{\text{4th Gen}}(x) = y^u_{pq}\overline{Q}_p H u_q + y^d_{pq}\overline{Q}_p H^\dagger d_q + y^e_{pq}\overline{L}_p H^\dagger e_q + y^\nu_{pq}\overline{L}_p H \nu_q + h.c. \tag{8.1}$$

where $p, q = 1, 2, 3, 4$ are the generation indices and $SU(2)$ contractions are implicit. As in the SM case, the diagonalisation of the quarks into their mass basis give the four generation CKM matrix $V^{4\times4}_{CKM}$.

This 4×4 CKM matrix has three complex phases, which provides a larger phase space in which CP violation can occur [4, 141–144]. One can parameterise the four generation

CKM matrix in terms of the 3×3 matrix of (3.127):

$$
\begin{aligned}
V_{CKM}^{4\times4} &= \begin{pmatrix} & & & 0 \\ & V_{CKM}^{3\times3} & & 0 \\ & & & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \times \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & c_{34} & s_{34} \\ 0 & 0 & -s_{34} & c_{34} \end{pmatrix} \\
&\times \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & c_{24} & 0 & s_{24}e^{-i\delta_2} \\ 0 & 0 & 1 & 0 \\ 0 & -s_{24}e^{i\delta_2} & 0 & c_{24} \end{pmatrix} \times \begin{pmatrix} c_{14} & 0 & 0 & s_{14}e^{-i\delta_3} \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ -s_{14}e^{i\delta_3} & 0 & 0 & c_{14} \end{pmatrix}
\end{aligned} \tag{8.2}
$$

The new Euler angles $(\theta_{14}, \theta_{24}, \theta_{34})$ denote the mixing between the fourth generation and the usual three generations. The complex phases δ_1, δ_2 and δ_3 are also introduced.

The unitarity requirement on the 4×4 CKM matrix allows us to place constraints on the CKM matrix elements. Assuming a fourth generation, the central values and their 1σ deviations can be computed from unitarity constraints on experimental measurements [145]:

$$
\left|V_{CKM}^{4\times4}\right| = \begin{pmatrix} 0.97414^{+0.00032}_{-0.00023} & 0.2245^{+0.0012}_{-0.0012} & 0.04200^{+0.000090}_{-0.000910} & 0.025^{+0.011}_{-0.025} \\ 0.2256^{+0.0011}_{-0.0059} & 0.9717^{+0.0024}_{-0.0105} & 0.04109^{+0.00045}_{-0.00145} & 0.057^{+0.097}_{-0.057} \\ 0.001^{+0.035}_{-0.001} & 0.062^{+0.044}_{-0.062} & 0.910^{+0.079}_{-0.080} & 0.41^{+0.15}_{-0.27} \\ 0.013^{+0.039}_{-0.013} & 0.04^{+0.12}_{-0.04} & 0.41^{+0.14}_{-0.27} & 0.910^{+0.078}_{-0.083} \end{pmatrix} \tag{8.3}
$$

Furthermore, in order not to contradict precision electroweak measurements, the mass splitting of the fourth generation b' and t' quarks is expected to be much smaller than the mass of the W-boson [146–148]. Thus, if kinematically possible and if the fourth generation is sufficiently heavy compared to the SM's third generation, then the decay modes that dominate are $b' \to tW$ and $t' \to bW$, governed by $|V_{tb'}|$ and $|V_{t'b}|$ respectively. Notice that the flavour changing neutral currents are suppressed by the usual GIM mechanism (3.126).

Chapter 9

The Search for a Fourth-Generation Top-Like Quark

In this chapter, we present a direct search for the fourth generation top-like quark with a *semileptonic* final state, consisting of exactly one lepton and at least four jets. We assume that the t' quarks are strongly pair-produced with 100% branching ratio in the $t' \to bW$ channel. The theoretical NNLO cross section for a strongly pair-produced heavy quark, computed by HATHOR [149], is tabulated in Table 9.1 and plotted in Fig. 9.1.

Our search is not limited to fourth generation chiral quarks and is relevant to many beyond the SM theories [150,151]. No excess over the expected SM background is observed in this analysis and a lower mass limit is set on the pair-produced t' quarks at 565 GeV, at 95% confidence level (C.L.), using the full 5 fb^{-1} of data collected during 2011 at a $\sqrt{s} = 7\,\text{TeV}$ centre of mass energy [1].

$m_{t'}$/GeV	$\sqrt{s} = 7$ TeV	$\sqrt{s} = 8$ TeV	$\sqrt{s} = 10$ TeV	$\sqrt{s} = 14$ TeV
300	7.987	12.32	24.41	62.53
325	4.988	7.827	15.78	41.49
350	3.200	5.086	10.47	28.24
375	2.100	3.385	7.111	19.64
400	1.406	2.301	4.922	13.93
425	0.957	1.582	3.468	10.04
450	0.662	1.112	2.481	7.354
475	0.464	0.791	1.800	5.458
500	0.330	0.571	1.322	4.101
525	0.237	0.414	0.982	3.115
550	0.171	0.305	0.737	2.391
575	0.125	0.226	0.558	1.852
600	0.0923	0.170	0.426	1.446
625	0.0685	0.128	0.328	1.138
650	0.0511	0.0970	0.254	0.903
675	0.0384	0.0742	0.199	0.721
700	0.0290	0.0569	0.156	0.579
725	0.0220	0.0440	0.123	0.468
750	0.0168	0.0341	0.0976	0.380
775	0.0129	0.0266	0.0778	0.310
800	0.00988	0.0208	0.0623	0.254
825	0.00761	0.0163	0.0501	0.209
850	0.00589	0.0129	0.0404	0.172
875	0.00457	0.0102	0.0327	0.144
900	0.00355	0.00809	0.0266	0.120
925	0.00277	0.00641	0.0217	0.100
950	0.00216	0.00513	0.0177	0.0838
975	0.00169	0.00409	0.0145	0.0705
1000	0.00133	0.00327	0.0120	0.0595
1025	0.00104	0.00262	0.00985	0.0503
1050	0.000819	0.00211	0.00811	0.0428
1075	0.000643	0.00169	0.00673	0.0364
1100	0.000508	0.00137	0.00557	0.0308
1125	0.000399	0.00111	0.00463	0.0264
1150	0.000819	0.000895	0.00385	0.0226
1175	0.000643	0.000724	0.00322	0.0193
1200	0.000197	0.000583	0.00268	0.0166
1225	0.000155	0.000475	0.00225	0.0143
1250	0.000122	0.000384	0.00187	0.0123
1275	0.0000963	0.000312	0.00157	0.0106
1300	0.0000760	0.000253	0.00132	0.00921
1325	0.0000698	0.000206	0.00111	0.00797
1350	0.0000472	0.000167	0.000933	0.00692
1375	0.0000373	0.000136	0.000784	0.00601
1400	0.0000292	0.000111	0.000661	0.00523
1425	0.0000230	0.000898	0.000557	0.00455
1450	0.0000181	0.000729	0.000470	0.00345
1475	0.0000142	0.000594	0.000397	0.00302
1500	0.0000114	0.000482	0.000335	0.00263

Table 9.1: Cross section σ in pb for $t'\bar{t}'$ pair production computed for different pp center of mass energies $\sqrt{s}$ at NNLO [149].

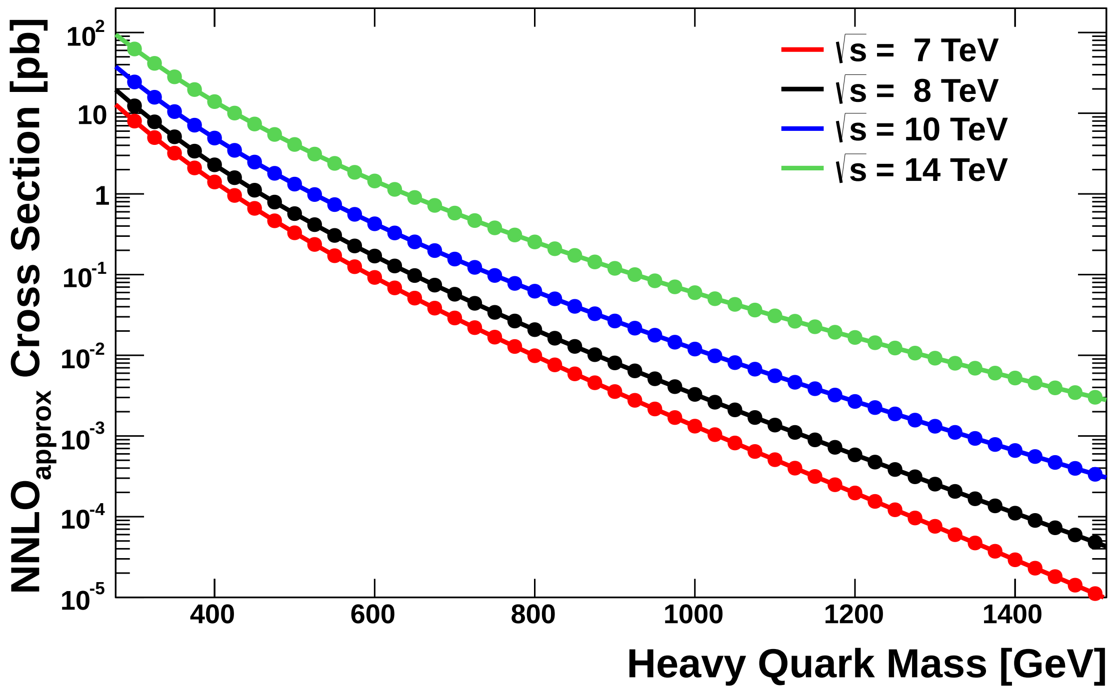


Figure 9.1: Cross section σ in pb for $t'\bar{t}'$ pair production computed for different pp center of mass energies $\sqrt{s}$ at NNLO [149].

9.1 Event Selection

We select events with exactly one isolated electron (e) or muon (μ) and at least four jets. The full decay chain for events considered is

$$pp \rightarrow t'\bar{t}' \rightarrow bW^+\bar{b}W^- \rightarrow bl\nu\bar{b}q\bar{q} \tag{9.1}$$

Events with one electron (e+jets) and one muon (μ+jets) are analysed separately, with the event selection optimised in each channel to maximise signal sensitivity. Statistical results from both channels are then combined to give a final limit. The differences in event requirements are minimal and are mainly motivated by detector and trigger limitations.

There are several SM processes that are expected to give rise to a lepton-plus-jets signature. Indeed, the SM $t\bar{t}$ production process is an irreducible background whose final state particles exactly match that of the signal. Therefore, two discriminating variables are used to distinguish the signal from the background: the reconstructed t' mass denoted $\mathrm{M_{fit}}$ and the scalar sum of transverse energies of the lepton, jets and missing p_T, denoted H_T .

All events are required to have at least four jets with $p_T > 120, 90, 35, 35\,\mathrm{GeV}$ and $|\eta| < 2.4$, missing energy $p_T > 20\,\mathrm{GeV}$ and at least one b-tagged jet.

Additionally, both channels require events to have a primary vertex with $|z| < 24\,\mathrm{cm}$ ($|r| < 2\,\mathrm{cm}$), where z (r) is the distance from the nominal interaction parallel (perpendicular) to the beam line.

In the μ+jets channel, the muon is required to have $p_T > 35\,\mathrm{GeV}$, with $|\eta| < 2.1$. In the e+jets channel, the electron is required to have $p_T > 30, 35, 45\,\mathrm{GeV}$ (to match the trigger thresholds that were updated during data collection) with $|\eta| < 2.5$ excluding the barrel-endcap transition region in the detector between $1.44 < |\eta| < 1.56$.

In both cases the lepton is required to have a transverse impact parameter $|\mathrm{d_{xy}}| < 0.02\,\mathrm{cm}$ and a longitudinal impact parameter $|\mathrm{d_z}| < 1\,\mathrm{cm}$, such that they are identified as arising from the primary vertex.

9.2 Event Yield

Each SM background is computationally generated and scaled to the process's respective cross section. This value is either measured in data or computed directly, dependent on whether such a measurement exists. The event yields after selection of data and the SM background can be found in Table 9.2.

Signal samples of a pair-produced t' are also generated computationally using Monte Carlo (MC) techniques [152]. These are generated in 25 GeV intervals between 400-625 GeV. The signal selection efficiencies $\epsilon = n_s/N_s$, where n_s is the number of selected events and N_s is the total number of generated events, can be found in Table 9.3. The signal cross section is the parameter of interest and shall be computed at each mass point with a statistical limit calculator discussed in Section 9.4.

Process	Cross section	e+jets	μ+jets
$t\bar{t}$	154 pb	3950 ± 490	5460 ± 670
W+jets	31 nb	462 ± 55	750 ± 110
Single-t production	85 pb	208 ± 24	336 ± 45
Z+jets, WW, WZ, ZZ	3.1 nb	49 ± 8	69 ± 11
QCD	-	78 ± 9	5 ± 5
Total Background	-	4750 ± 560	6620 ± 800
Data	-	4734	6448

Table 9.2: Number of MC and data events after selection in the e+jets channel in 4.98 fb^{-1} and μ+jets channel in 4.90 fb^{-1} of data analysed at $\sqrt{s} = 7\,\text{TeV}$. The listed uncertainty accounts for both statistical and cross section uncertainty. The QCD uncertainty is calculated using data driven methods described in [1].

Mass [GeV]	400	425	450	475	500
Electron Efficiency	$4.5 \pm 0.4\%$	$4.7 \pm 0.4\%$	$5.0 \pm 0.4\%$	$5.0 \pm 0.4\%$	$5.1 \pm 0.4\%$
Muon Efficiency	$5.4 \pm 0.4\%$	$5.6 \pm 0.4\%$	$6.0 \pm 0.4\%$	$6.1 \pm 0.4\%$	$6.2 \pm 0.4\%$
Mass [GeV]	525	550	575	600	625
Electron Efficiency	$5.0 \pm 0.4\%$	$5.2 \pm 0.4\%$	$5.1 \pm 0.4\%$	$5.0 \pm 0.4\%$	$5.1 \pm 0.4\%$
Muon Efficiency	$6.4 \pm 0.4\%$	$6.5 \pm 0.4\%$	$6.6 \pm 0.4\%$	$6.6 \pm 0.4\%$	$6.5 \pm 0.4\%$

Table 9.3: Signal selection efficiency in the μ+jets and e+jets channels for different t' masses. The listed error is purely statistical.

9.3 Event Reconstruction

Since the SM $t\bar{t}$ background is irreducible the reconstructed mass is used as a discriminating variable. We perform a kinematic fit of each event to the $t'\bar{t}' \to bW\bar{b}W \to \nu b q_1 \bar{q}_2 \bar{b}$ process. There are two steps in the reconstruction of the t'-quark mass: the assignment of reconstructed objects to the quarks and a kinematic fit to improve the resolution of the reconstructed mass of the t'-quark candidates [153]. The four-momenta resulting from the kinematic fit of the particles in the final state must satisfy three constraints, where m is the invariant mass of the corresponding particles in parentheses, m_W is the W-boson mass, M_{fit} is a free parameter in the fit (representing the reconstructed t' invariant mass), and l represents either an electron or muon.

The constraints are:

- $m(l\nu) = m_W$
- $m(q\bar{q}_2') = m_W$
- $m(l\nu b) = m(q\bar{q}_2 b) \equiv \mathrm{M}_{\text{fit}}$

Here l, ν, b denote either particle or antiparticle.

The final reconstructed objects are the charged lepton, the missing transverse momentum, and the four or more jets. The neutrino's transverse momentum is reconstructed using momentum conservation and its longitudinal momentum can be deduced up to a two-fold ambiguity in the fit.

Taking only the four leading (highest p_T) jets, there are twelve combinations of matching the reconstructed jets to partons. Due to initial and final state radiation with high momenta, considering more jets increases the likelihood that the true combination is included in the set considered. This is at the cost of introducing more combinations from which to pick. A study was done for this analysis, which showed that five leading jets was the best trade off between these two factors [1].

From the set of possible fit combinations of the five leading jets, the one with the smallest χ^2 is chosen, with a χ^2 value for each combination determined by the goodness of fit to the constraints. The mass $M_{\rm fit}$ and H_T for this combination are read off and are shown in Fig. 9.2.

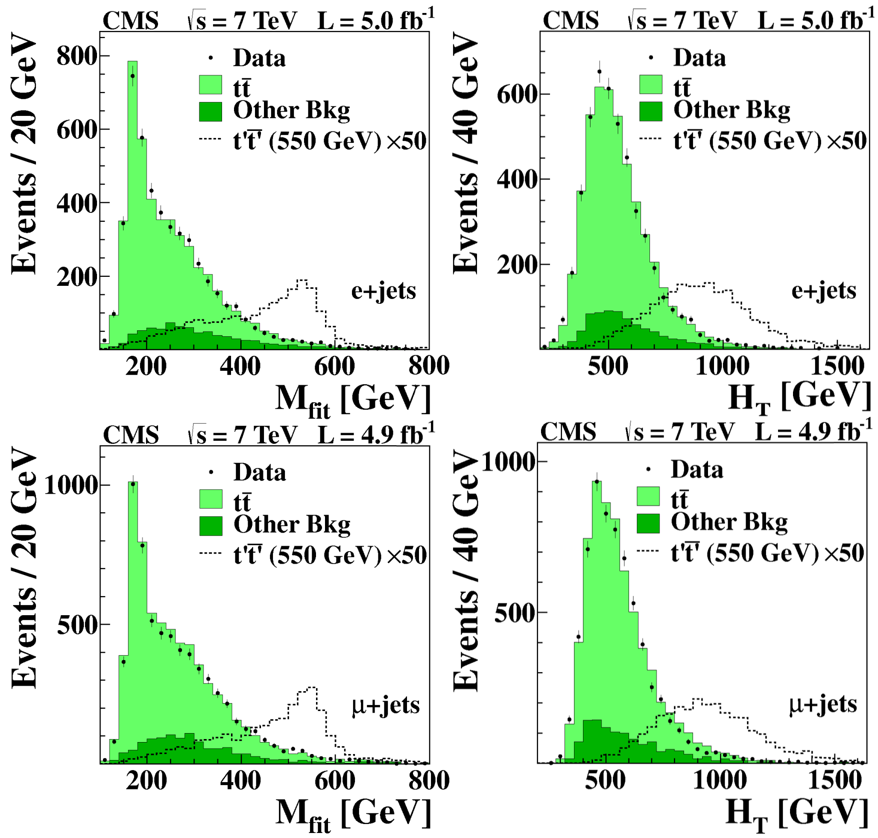


Figure 9.2: Results of kinematics fit: (left) the reconstructed $M_{\rm fit}$ and (right) H_T for the (top) e+jets and (bottom) μ+jets channels. The MC has been normalised to the data and the signal is scaled up by a factor of fifty to its theoretical value, for better visibility.

The fitted mass for the selected signal MC for different generated masses of t' is shown in Fig. 9.3-(a). The bottom panel of Fig. 9.3-(b) shows the relative resolution of the reconstructed t' mass for different generated masses. Finally, Fig. 9.4-(a, b) show the evolution of the mean value and RMS of the relative resolution of the reconstructed mass as a function of the generated mass.

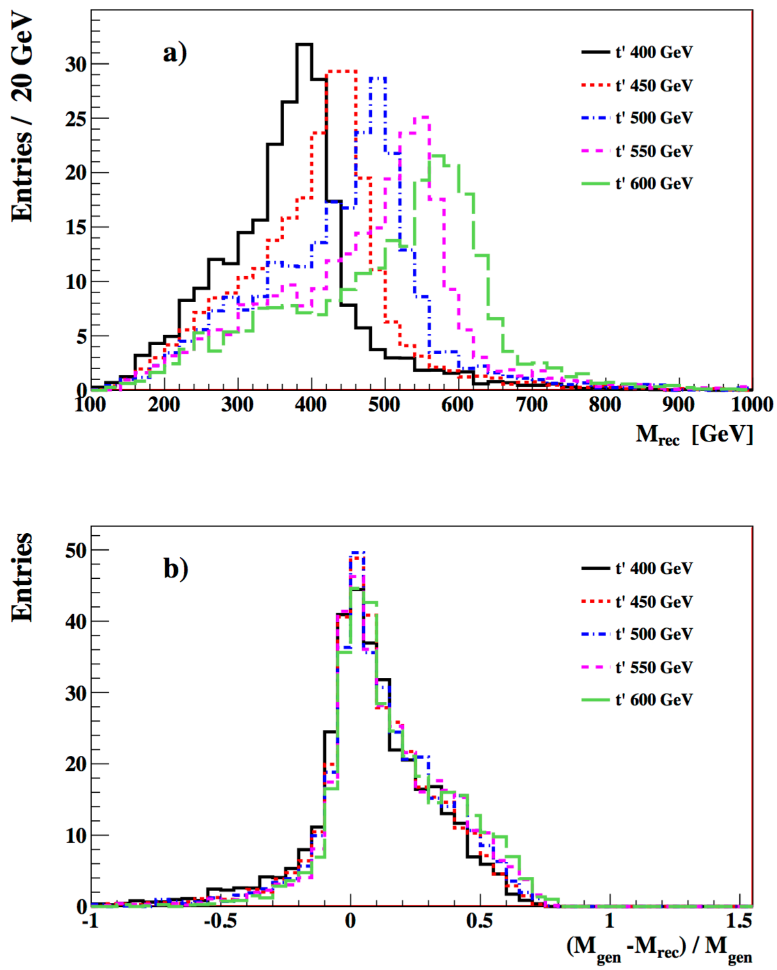


Figure 9.3: (a) Mass distributions for selected t' MC signal for different generated masses of t'. All distributions are normalised to the distribution for mass 400 GeV (for which production cross section 1 pb is assumed). (b) Relative resolution of the reconstructed t' mass for different generated masses.

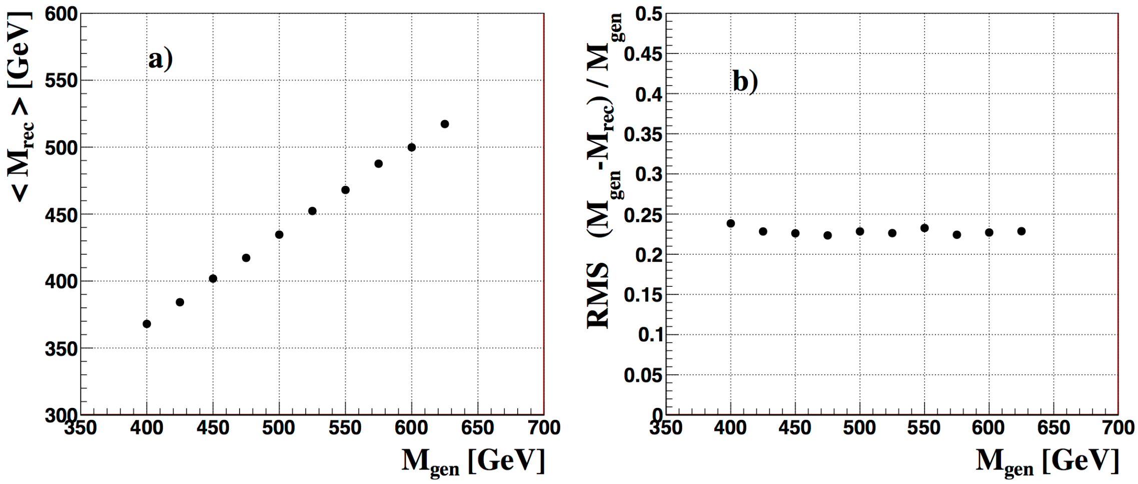


Figure 9.4: (a) Mean value of the reconstructed t' signal mass as a function of the generated mass. (b) RMS of relative resolution of Fig. 9.3-(b).

9.4 Results and Limit Setting

For each sample, a 2-dimensional distribution $M_{\rm fit}$ versus H_T is made to take into account the correlation between the two variables. Due to limited statistics of the simulated samples, the non-$t\bar{t}$ backgrounds are summed into an "other background" category. The input distributions are shown in Fig. 9.5 and Fig. 9.6 for the e+jets and μ+jets channels respectively.

The bins of the two-dimensional histograms of $M_{\rm fit}$ versus H_T are merged by an algorithm which maximally separates signal and background bins, generating final 1-dimensional histograms in which all bins are well populated by simulated background events [1].

Fig. 9.7 shows the 1-dimensional templates after combination of bins for both channels. These are the inputs to the limit setting procedure. The binmaps indicating the origin of each bin are also shown, with each colour corresponding to one particular bins. Since no excess is observed in the data, we compute a lower limit on the cross section for $t'\bar{t}'$ production for each t' mass point.

The main sources of uncertainty in this analysis come from the jet energy scale, selection efficiencies, luminosity, b-tagging, and the number of background events determined by the uncertainty of the cross section. These are accounted for in the limit setting procedure,

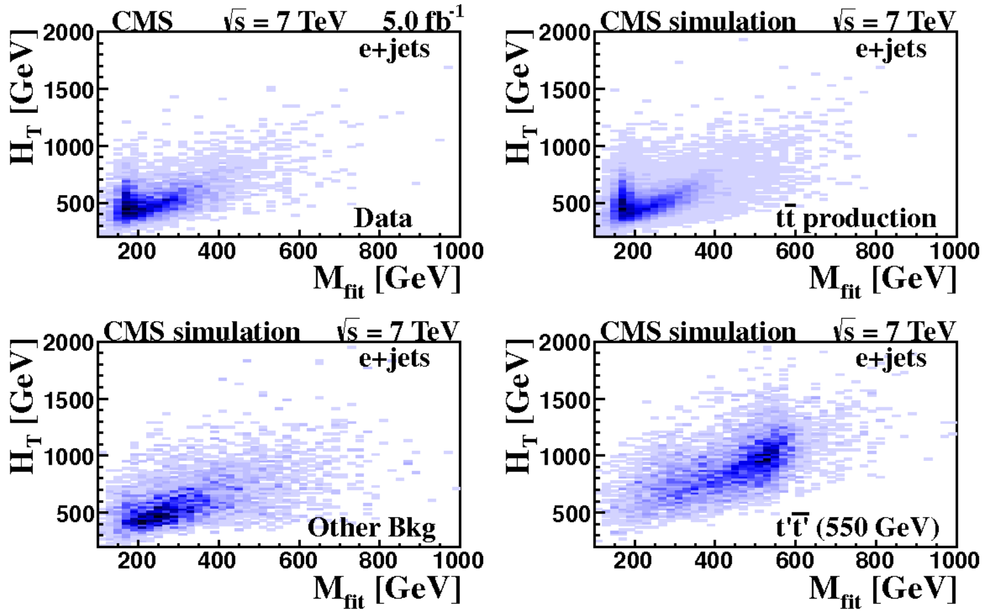


Figure 9.5: Histograms of $\mathrm{M}_{\mathrm{fit}}$ versus H_T for the e+jets channel for data, $\mathrm{t\bar{t}}$ production, other backgrounds and the t' signal.

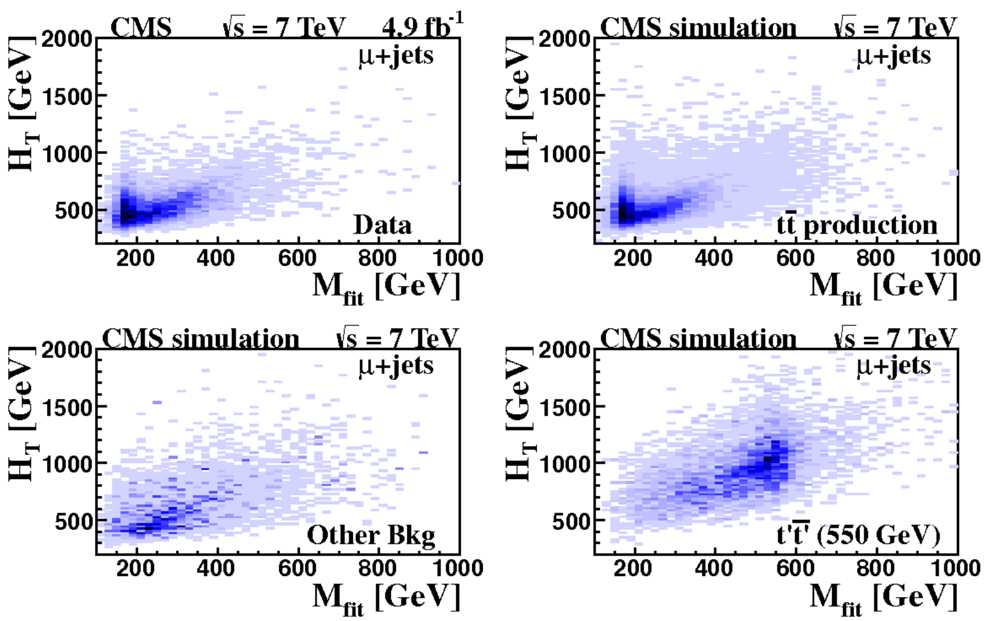


Figure 9.6: Histograms of $\mathrm{M}_{\mathrm{fit}}$ versus H_T for the μ+jetschannel for data, $\mathrm{t\bar{t}}$ production, other backgrounds and the t' signal.

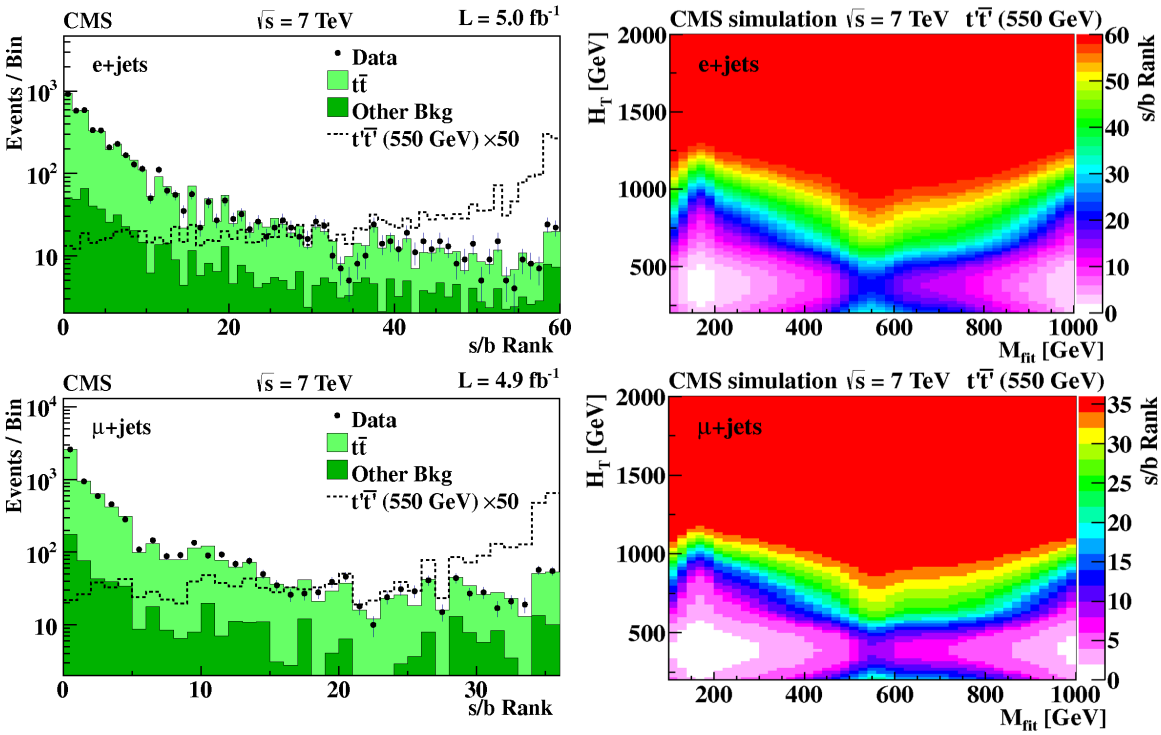


Figure 9.7: (Left) Histogram of 1-dimensional templates after bin-merging for the e+jets and μ+jets channels. (Right) Binmaps for the e+jets and μ+jets channels.

which proceeds by comparing the number of events in the signal, data, and background histograms bin-by-bin by a statistical package [154–156]. The method computes upper limits on the cross section at each t' quark mass point, above which we are sensitive.

Fig. 11.34 shows the observed and expected limits determined in this way. The combination of the two channels is performed by combining the full likelihoods and applying a similar statistical procedure to the resulting likelihood. The full details can be found in [1].

The computed upper limit on the cross section is then compared to the theoretical value. In the e+jets channel, using 4.90 pb^{-1} of data, an observed (expected) lower limit on the t' mass is set at 490 (540) GeV at 95% C.L. For the μ+jets channel, using 4.98 pb^{-1} of data gives an observed (expected) limit of $m_{t'} > 560$ (550) GeV at 95% C.L.

The two lepton channels are then combined and an upper bound cross section limit is again set at each mass point. No excess over the expected SM processes was found and a strongly produced t' quark with 100% branching ratio in the $t' \to bW$ channel is excluded below $m_{t'} < 570(590)$ GeV in the lepton-plus-jets channel at 95% C.L.

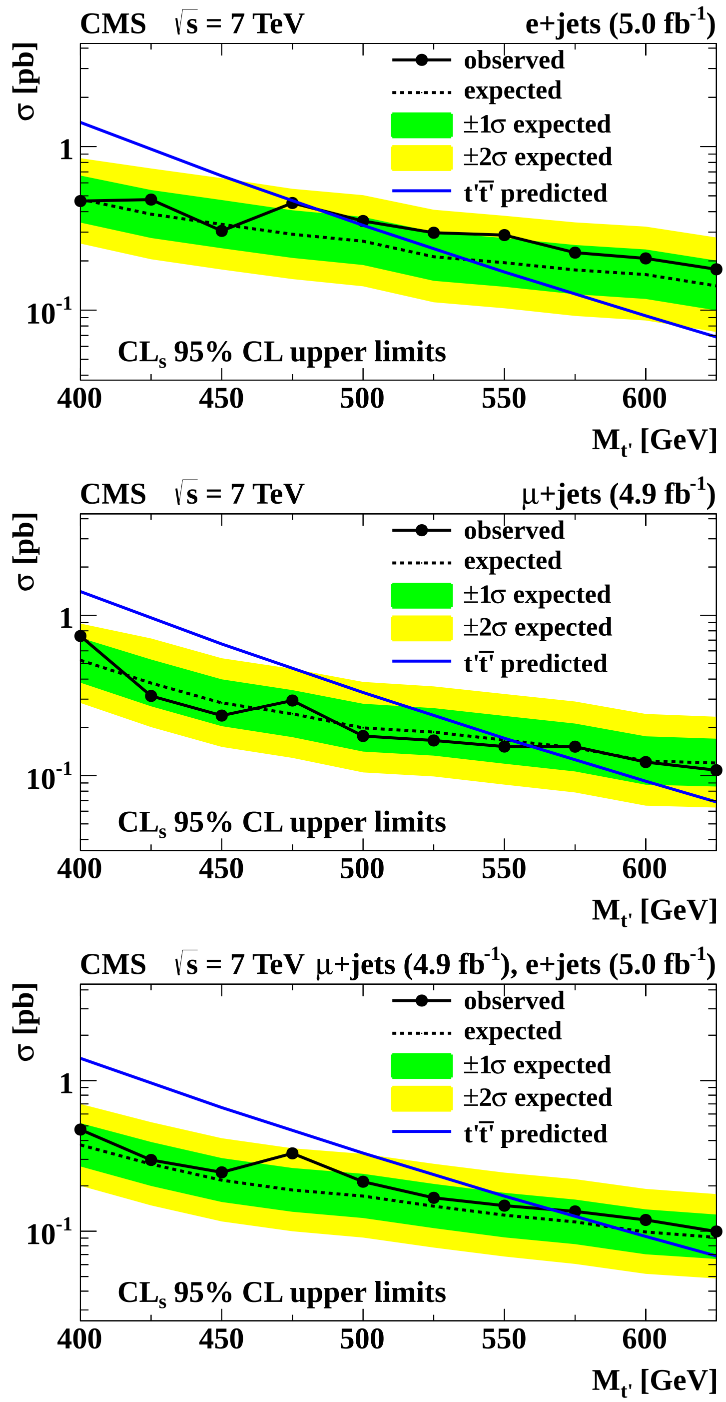


Figure 9.8: The observed (solid line with points) and expected (dotted line) 95% C.L. upper limits on the $t'\bar{t}'$ production cross section as a function of the t' quark mass for $e+$jets (top), $\mu+$ jets (middle), and combined (bottom) channels. The ±1 and ±2 standard deviation ranges for the expected limits are shown by the coloured bands.

Notice that the exclusion limit on the mass of the fourth generation t' quark is entering a region where perturbative calculations for the weak interactions start to fail and non-perturbative effects become significant [157]. This search is equally sensitive to non-chiral heavy quarks decaying to bW. In this case, the results can be interpreted as upper limits on the production cross section times the branching fraction to bW.

We shall extend this search to a general heavy top-like quark T in the next chapter.

Part V

The Little Higgs

Chapter 10

The Little Higgs Models

The Little Higgs models are a class of theories that start from an enlarged gauge group for the electroweak sector. In general, these theories are motivated by resolving the issue of naturalness and the fine-tuning required for the Higgs to be relatively light. An approximate global symmetry is typically introduced that protects the stability of the Higgs mass.

There are, in general, two classifications of Little Higgs models [158–160]: the product group and the simple group. Both introduce a larger gauge group, which breaks down to the SM at some high energy. These two classes have similar phenomenology and they predict additional heavy particles to cancel divergent loop contributions to the Higgs mass from the SM.

The new particles are typically introduced at the TeV scale, whose masses are dependent on the symmetry breaking scale f of the higher gauge group.

We shall discuss the phenomenology of one example known as the Simplest Little Higgs [161] and heuristically describe the experimental predictions made by these theories. We shall closely follow the discussions of Ref. [162].

In Chapter 11, we will move on to the experimental search done at CMS for the heavy top-like quark, which forms a key ingredient in this entire class of theories.

10.1 The Simplest Little Higgs

The Simplest Little Higgs extends the electroweak gauge groups and fits nicely into a regime still permissible from electroweak constraints [163, 164].

The model predicts a whole slew of additional particles (both heavy gauge bosons and heavy fermions) expected to have mass at around the TeV scale. The new particles generate a quartic coupling to the Higgs mass, which provides a radiative correction at low energies to resolve the issue of naturalness. Experimental results indicate that, if these particles exist, they must be relatively heavy and couple very weakly to SM particles.

We proceed by enlarging the $SU(2)_L \times U(1)_Y$ to $SU(3)_W \times U(1)_X$ in the minimal way, where the conserved numbers are given by $X + 2Y = 5(B - L)$ with the canonical weak hypercharge $Y = 2(Q - T_3)$. The fermion $SU(2)_L$ doublets become fermion $SU(3)_W$ triplets, requiring the introduction of new $SU(3)_W$ gauge bosons to make the theory gauge symmetric.

Thus, the full high energy gauge group is

$$SU(3)_C \times SU(3)_W \times U(1)_X \tag{10.1}$$

with fermion fields embedded in the "universal" representation

$$\begin{aligned}
Q_n &= (3,3)_{\frac{1}{3}} \quad , \quad & L_n &= (1,3)_{-\frac{1}{3}} \\
d_n^c &= (\overline{3},1)_{\frac{1}{3}} \quad , \quad & e_n^c &= (1,1)_1 \\
2 \times u_n^c &= (\overline{3},1)_{-\frac{2}{3}} \quad , \quad & N_n^c &= (1,1)_0
\end{aligned}$$

where $n = 1, 2, 3$ labels the generation index. We follow the notation that c denotes the (conjugate) right-handed fields.

Explicitly, the fields can be written out in the **3** and **1** representations of $SU(3)$ for left and right-handed fermions respectively as

$$\begin{aligned}
Q_n^T &= (u, d, iU)_n \quad , \quad & iu_n^c, id_n^c, iU_n^c \\
L_n^T &= (\nu, e, iN)_n \quad , \quad & ie_n^c, iN_n^c
\end{aligned}$$

where we have introduced a factor of i for later convenience.

One of the two additional right-handed u^c fields will become the SM right-handed up-type quarks. The other will couple to the last component of Q_n and will give rise to an electroweak singlet representing a new vector-like particle that acquires a large mass (after breaking the high energy symmetry). Similarly, n^c will couple to the third component of L_n.

The symmetry breaking $SU(3)_W \times U(1)_X \to SU(2)_W \times U(1)_Y$ is achieved with an aligned vacuum expectation value for two complex triplet scalar fields

$$\Phi_1, \Phi_2 = (1,3)_{-\frac{1}{3}} \tag{10.2}$$

which contains the canonical Higgs field.

We parameterise the fields Φ_i with vacuum expectation value f_i as

$$\Phi_1 = e^{i\Theta \frac{f_2}{f_1}} \begin{pmatrix} 0 \\ 0 \\ f_1 \end{pmatrix} \quad , \quad \Phi_2 = e^{-i\Theta \frac{f_1}{f_2}} \begin{pmatrix} 0 \\ 0 \\ f_2 \end{pmatrix} \tag{10.3}$$

where

$$\Theta = \frac{1}{f}\left[\frac{\eta}{\sqrt{2}} + \begin{pmatrix} 0 & 0 & \\ 0 & 0 & H \\ & H^\dagger & 0 \end{pmatrix}\right] \quad \text{and} \quad f^2 = f_1^2 + f_2^2 \tag{10.4}$$

H is the usual Higgs field and η is a diagonal real scalar field that remains unmixed with the unphysical Goldstone bosons that are introduced by the Higgs field after EWSSB.

Note the two scalar fields have aligned vacuum expectation value so as not to break residual $SU(2)_L$ symmetry at this stage.

We also expect that f_1 and f_2 are of the same order as otherwise we would introduce yet another hierarchy problem. Thus, we define the ratio of the vacuum expectation values $f_2/f_1 \equiv \tan\beta \sim \mathcal{O}(1)$, which implies that $\beta \sim \mathcal{O}(1)$.

This higher order symmetry breaking introduces the additional heavy gauge bosons: (Y^0, X^-), which form an $SU(2)_L$ doublet, a Z' gauge boson from the broken linear combination of the T_8 generator of $SU(3)_W$ and the generator of $U(1)_X$, and finally, the model also contains a singlet pseudoscalar η.

10.2 The Gauge and Higgs Sectors

We shall quickly skip over the spectra of new gauge bosons as we are more interested in the new heavy quarks. We shall mainly do so to be complete with our phenomenology but also to show the suppressed couplings between these heavy gauge bosons and the heavy quarks.

The additional gauge bosons from the expanded gauge group (at the TeV scale) typically gain TeV masses.

Recall that $SU(3)$ generators can be written in terms of the 8 Gell-Man matrices, where $T^a = \lambda_a/2$ and λ_a are Gell-mann matrices $a = 1, \cdots 8$.

Then, one can write the $SU(3)_W$ generators as

$$A_a T^a = \frac{A^3}{2}\begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{pmatrix} + \frac{A^8}{2\sqrt{3}} + \frac{1}{\sqrt{2}}\begin{pmatrix} 0 & W^+ & Y^0 \\ W^- & 0 & X^- \\ \overline{Y}^0 & X^+ & 0 \end{pmatrix} \tag{10.5}$$

where we have defined the gauge bosons in the mass basis in a similar manner to the SM ones. The $X^\pm$ and Y^0 are sometimes denoted $W'^\pm$ and W'^0 respectively.

The two remaining gauge fields are formed by linear combinations of the diagonal generators

$$Z'^0 = \frac{\sqrt{3} g A^8 + g_X B^X}{\sqrt{3g^2 + g_X^2}} \tag{10.6a}$$

$$B = \frac{-g_X A^8 + \sqrt{3} g B^X}{\sqrt{3g^2 + g_X^2}} \tag{10.6b}$$

with B the hypercharge gauge boson, with hypercharge $Y = -\frac{1}{\sqrt{3}} T^8 + Q_X$.

Then, the kinetic term for Φ_i is

$$\mathcal{L}_{\Phi_i}^{\text{kin}} = \sum_{i=1,2} \left| \left(\partial_\mu + i g A_{\mu a} T^a - \frac{i}{3} g_X B_\mu^X \right) \Phi_i \right|^2 \tag{10.7}$$

with g the coupling constant of $SU(3)$ that is set equal to that of the canonical SM $SU(2)_L$ coupling, and

$$g_X = \frac{g \tan\theta_W}{\sqrt{1 - \frac{1}{3}\tan^2\theta_W}} \tag{10.8}$$

set equal to the $U(1)_X$ coupling. The generators of the gauge groups $SU(3)_L$ and $U(1)_X$ are T^a and B^X respectively. For convenience, we shall define $\tan\theta_W \equiv t_W$ and $\tan\beta = t_\beta$ (similarly for $\sin \equiv s$ and $\cos \equiv c$).

Explicitly, the scalar fields can be written as

$$\Phi_1 = f c_\beta \left[\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} + \frac{i t_\beta}{f} \begin{pmatrix} H \\ \frac{\eta}{\sqrt{2}} \end{pmatrix} - \frac{t_\beta^2}{2f^2} \begin{pmatrix} \sqrt{2}\eta H \\ H^\dagger H + \frac{\eta^2}{2} \end{pmatrix} + \cdots \right] \tag{10.9a}$$

$$\Phi_2 = f s_\beta \left[\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} - \frac{i}{f t_\beta} \begin{pmatrix} H \\ \frac{\eta}{\sqrt{2}} \end{pmatrix} - \frac{1}{2f^2 t_\beta^2} \begin{pmatrix} \sqrt{2}\eta H \\ H^\dagger H + \frac{\eta^2}{2} \end{pmatrix} + \cdots \right] \tag{10.9b}$$

where η is an additional scalar particle, a real singlet, that remains light. For η not to be mixed with the unphysical (eaten) Goldstone bosons, η is chosen to be diagonal.

One can then find the masses of the gauge bosons by explicitly expanding the scalar field Φ. It is convenient to write

$$\left\langle \Phi_1 \Phi_1^\dagger + \Phi_2 \Phi_2^\dagger \right\rangle = \begin{pmatrix} \langle H H^\dagger \rangle & 0 \\ 0 & f^2 \end{pmatrix} \tag{10.10}$$

taking terms up to H/f where $\langle H^T \rangle = \left(v/\sqrt{2}, 0\right)$. The relevant terms in the covariant derivative is

$$\mathrm{Tr}\left[\left(g A_\mu^a T^a - \frac{1}{3} g_X B_\mu^X \right)^2 \Phi_i \Phi_i^\dagger \right] \tag{10.11}$$

We rotate the gauge fields into the mass eigenstates with the definitions (10.5) and (10.7). After expanding this out around the vacuum expectation value, the masses of the particles can then be read off at leading order:

$$\begin{aligned} m_{W^\pm} &= \frac{g}{2}v \\ m_{Z^0} &= \frac{g}{2c_W}v \\ m_{X^\pm} &= \frac{g}{\sqrt{2}}f \\ m_{Y^0} &= \frac{g}{\sqrt{2}}f \\ m_{Z'} &= \frac{\sqrt{2}g}{\sqrt{3-t_W^2}}f \end{aligned}$$

In summary, the broken gauge generators give rise to the heavy gauge bosons with masses $\mathcal{O}$ TeV. These consist of the Z'_0 boson (linear combination of A^8 and B) and a complex $SU(2)$ doublet (Y^0, X^-).

The Singlet η

The singlet η gains a mass of a few hundred GeV and its nature is very model dependent. In general, its interactions are suppressed by v/f relative to the usual Yukawa couplings and its discussion can be found elsewhere [158–160, 162].

10.3 The Lepton Sector

The scalar interactions of the leptons can be written as

$$\mathcal{L}_Y = i\lambda_{N_m} N_m^c \Phi_2^\dagger L_m + \frac{i\lambda_e^{mn}}{\Lambda} e_m^c \epsilon_{ijk} \Phi_1^i \Phi_2^j L_n^k + \text{h.c.} \tag{10.13}$$

where $m, n = 1, 2, 3$ are the generation indices, $i, j, k = 1, 2, 3$ the $SU(3)_W$ indices, and $h.c.$ denotes the Hermition conjugate terms. $L_m = (\nu_L e_L, iN_L)_m^T$ are the left-handed triplets, with N_m^c the right-handed neutral leptonic fields that will form the heavy leptons with mass $\sim$ TeV.

Expanding the scalar field here yields

$$m_{N_m} = \lambda_{N_m} s_\beta f \tag{10.14}$$

and also gives rise to the terms

$$L_Y \supset -\frac{m_{N_m}}{\sqrt{2} f t_\beta} H N_m^c \nu + \text{h.c.} \tag{10.15}$$

Writing this in the gauge basis leads to couplings between the heavy fermions and all SM leptons [162]. Further, the canonical SM leptons gain their usual masses and couplings after EWSSB.

10.4 The Quark Sector

The quark sector has an additional heavy quark per generation. The masses of these quarks arise from the couplings of these fields to the scalar fields Φ_i. The relevant terms in the Lagrangian are

$$\mathcal{L}_Q = \lambda_1^{un} i u_1^{nc} \Phi_1^\dagger Q_n + \lambda_2^{un} i u_2^{nc} \Phi_2^\dagger Q_n + \frac{\lambda_d^{mn}}{\Lambda} i d_m^c \epsilon_{ijk} \Phi_1^i \Phi_2^j Q_n^k + h.c. \tag{10.16}$$

where $m, n = 1, 2, 3$ are the generations, $i, j, k = 1, 2, 3$ are $SU(3)$ indices. The fields d_m^c run over all down-type right-handed quarks (d^c, s^c, b^c) and the fields $u_{1,2}^{c,n}$ are the conjugate (right-handed) fields that are linear combinations of the conjugate (right-handed) fields of the SM and the heavy up-type quarks.

We can rotate these terms into the mass basis using the following identifications

$$u^c = \frac{-\lambda_2^u s_\beta u_1^c + \lambda_1^u c_\beta u_2^c}{\sqrt{\left(\lambda_1^u\right)^2 c_\beta^2 + \left(\lambda_2^u\right)^2 s_\beta^2}} \tag{10.17a}$$

$$U^c = \frac{\lambda_1^u c_\beta u_1^c + \lambda_2^u s_\beta u_2^c}{\sqrt{\left(\lambda_1^u\right)^2 c_\beta^2 + \left(\lambda_2^u\right)^2 s_\beta^2}} \tag{10.17b}$$

with the inverse transformations are given by

$$u_1^c = \frac{\lambda_1 c_\beta U^c - \lambda_2 s_\beta u^c}{\sqrt{\lambda_1^2 c_\beta^2 + \lambda_2^2 s_\beta^2}} \tag{10.18a}$$

$$u_2^c = \frac{\lambda_2 s_\beta U^c + \lambda_1 c_\beta u^c}{\sqrt{\lambda_1^2 c_\beta^2 + \lambda_2^2 s_\beta^2}} \tag{10.18b}$$

where we have suppressed the generational index n.

Quark Masses

We follow Ref. [161] and proceed by making a few simplifications. To avoid flavour changing neutral currents, we require that one of the coupling matrices λ_i^u that relate the different generations is proportional to the identity (without loss to generality take $\lambda_2 \propto \mathbb{I}$). Now, we still have sufficient gauge freedom to rotate λ_1^u using unitary transformations on the fields Ψ_Q and u_1^c to reproduce the hierarchical SM quark masses.

The 6×6 mass matrix (composed of three families with two types of quarks) can now be split into their separate generations, which gives three sets of 2×2 mass matrices

$$\mathcal{L} = \begin{pmatrix} u_1^c & u_2^c \end{pmatrix} \begin{pmatrix} \lambda_1^u \Phi_1^\dagger \\ \lambda_2^u \Phi_2^\dagger \end{pmatrix} Q \tag{10.19}$$

where $Q^T = (b, t, T)$, one for each generation.

Then, using the definition of Φ_i, we have that the mass matrix is now of the form

$$M \sim \begin{pmatrix} \lambda_1^u \langle H \rangle \frac{f_1}{f} & \lambda_1^u f_1 \\ -\lambda_2^u \langle H \rangle \frac{f_2}{f_1} & \lambda_2^u f_2 \end{pmatrix} \tag{10.20}$$

Writing these terms in the mass eigenbasis, chosen specifically to diagonalise M, and expanding around the vacuum of the Higgs field, where $\langle H \rangle = v/\sqrt{2}$, gives the following

Lagrangian after electroweak spontaneous symmetry breaking

$$\begin{aligned}\mathcal{L}_Q &= -m_{U_n} U_n^c U_n - m_{u_n} u_n^c u_n \\ &\quad + \frac{v}{\sqrt{2}} \frac{s_\beta c_\beta \left[(\lambda_1^{un})^2 - (\lambda_2^{un})^2 \right]}{\sqrt{(\lambda_1^{un})^2 c_\beta^2 + (\lambda_2^{un})^2 s_\beta^2}} U_n^c u_n + \text{h.c.}\end{aligned} \tag{10.21}$$

where $n = 1, 2, 3$ again labels the generations and

$$m_{u_n} = \frac{v}{\sqrt{2}} \frac{\lambda_1^{un} \lambda_2^{un}}{\sqrt{(\lambda_1^{un})^2 c_\beta^2 + (\lambda_2^{un})^2 s_\beta^2}} \tag{10.22a}$$

$$m_{U_n} = \sqrt{\lambda_1^2 c_\beta^2 + \lambda_2^2 s_\beta^2} f \tag{10.22b}$$

To recreate the small mass of the u (c) quark, we require one of λ_1^{u1} (λ_1^{u2}) or λ_1^{u2} (λ_2^{u2}) to be small. Then, taking $\lambda_1^u \ll \lambda_2^u$, the first two generations for the SM up-type quarks are

$$m_u = \lambda_1^{u1} \frac{v}{\sqrt{2}} \quad , \quad m_c = \lambda_1^{u2} \frac{v}{\sqrt{2}} \tag{10.23}$$

and the masses for the first two generations of heavy quarks are

$$m_U = f \lambda_2^{u1} s_\beta \quad , \quad m_C = f \lambda_2^{u2} s_\beta \tag{10.24}$$

For the third generation, the situation is more complicated, we have

$$m_t = \lambda_t v = \frac{\lambda_1 \lambda_2}{\sqrt{2} \sqrt{\lambda_1^2 c_\beta^2 + \lambda_2^2 s_\beta^2}} \tag{10.25a}$$

$$m_T = \sqrt{\lambda_1^2 + \lambda_2^2} f = \sqrt{2} \frac{t_\beta^2 + x_\lambda^2}{(1 + t_\beta^2) x_\lambda} \frac{m_t}{v} f \tag{10.25b}$$

where in the last equality we define $x_\lambda \equiv \lambda_1 / \lambda_2$ to be the ratio of the two couplings and $t_\beta = f_2 / f_1$ where $f^2 = f_1^2 + f_2^2$.

To minimise the fine-tuning that we would otherwise reintroduce, the heavy top-like T-quark must be relatively light and has a lower bound on the mass where $\tan \beta = x_\lambda$, such that

$$m_T \geq 2\sqrt{2}\sin\beta\cos\beta\frac{m_t}{v}f \approxeq f\sin 2\beta \tag{10.26}$$

where $m_t \propto v/\sqrt{2}$. Note that, the parameters f, the symmetry breaking scale, and β (the mixing angle of the two higher vacuum expectation values) governs the mass of the heavy quarks. To avoid fine-tuning, we therefore find $m_T \simeq f\sin 2\beta$ and thus $m_T \leq f$, where f is around the TeV scale. These particles are therefore well within reach at the LHC.

The down type quarks get their masses in the usual SM way from the Yukawa coupling of the scalar fields in the higher energy theory. For simplicity, we will assume that they do not mix significantly with the heavy quarks [162].

Quark Couplings to the Gauge Bosons

One can see the explicit mixing term between the heavy and SM quarks in (10.21). Following Ref. [162], we note that this mixing indications a violation of the $SU(3)$ symmetry. We can rewrite the gauge fields (with subscript 0) in terms of the mass eigenstates U_n and u_n after mixing induced by electroweak spontaneous symmetry breaking as

$$U_{m0} = U_m + \delta_{um}u_m \quad , \quad u_{m0} = \left(1 - \frac{1}{2}\delta_{u_m}^2\right)u_m - \delta_{u_m}U_m \tag{10.27}$$

where

$$\delta_{u_m} = \frac{v}{\sqrt{2}f}\frac{s_\beta c_\beta\left[(\lambda_1^{um})^2 - (\lambda_2^{um})^2\right]}{\left[(\lambda_1^{um})^2 c_\beta^2 + (\lambda_2^{um})^2 s_\beta^2\right]} \tag{10.28}$$

Note that these are just rotations of the mass fields by an $SO(2)$ matrix, with mixing angle δ as defined by the mixing term. One can understand these as small corrections to the canonical SM Lagrangian when rotating the gauge fields into the mass eigenstates.

After both the higher symmetry breaking and the SM EWSSB, the Lagrangian for the heavy quarks coupling to the SM gauge boson reads

$$\mathcal{L}_{U,W} = -\frac{gW_\mu^+}{\sqrt{2}}\left[\left(1-\frac{1}{2}\delta_{ui}^2\right)V_{ij}^{CKM}\overline{u}_i\gamma^\mu d_j - \delta_{u_i}V_{ij}^{CKM}\overline{U}_i\gamma^\mu d_j + h.c.\right] \tag{10.29}$$

$$\begin{aligned}\mathcal{L}_{U,Z} &= -\frac{gZ_\mu}{c_W}\Bigg\{\left(J^\mu - s_W^2 j_{\rm EM}^\mu\right) - \frac{1}{2}\delta_{u_i}^2\overline{u}_i\gamma^\mu u_i - \frac{1}{2}\left[\delta_{u_i}\overline{U}_i\gamma^\mu u_i + \text{h.c.}\right] \\ &+\frac{\delta_Z}{\sqrt{3-4s_W^2}}\Bigg[\left(\frac{1}{2}-\frac{1}{3}s_W^2\right)\left(\overline{u}_i\gamma^\mu u_i + \overline{d}_i\gamma^\mu d_i\right) \\ &-\frac{2}{3}s_W^2\overline{u}_i^c\gamma^\mu u_i^c + \frac{1}{3}s_W^2\overline{d}_i^c\gamma^\mu d_i^c\Bigg]\Bigg\}\end{aligned} \tag{10.30}$$

where the δ_{u_i} are defined as in (10.28), δ_Z is the analogous shift in the gauge field Z^0 defined as

$$\delta_Z = -\frac{(1-t_W^2)}{4c_W\sqrt{3-t_W^2}} \tag{10.31}$$

and the currents are the canonical SM ones from (3.109).

Finally, the coupling of the fermions to the electromagnetic current is the canonical

$$\mathcal{L}_A = -A_\mu e j_{\rm EM}^\mu \tag{10.32}$$

Quark Couplings to X and Y Gauge Bosons

The other decay modes of the heavy quarks are through the heavy gauge bosons (i.e. bX^+,tY^0) and heavy scalar ($t\eta$). These are highly suppressed due to the heavy final state particles. The decay modes to the X, Y gauge bosons is described by the Lagrangian

$$\mathcal{L}_{X,Y} = -\frac{g}{\sqrt{2}}\left[iX_\mu^-\overline{d}_i\gamma^\mu(V_{ji}^{CKM\star}U_j + \delta_{u_j}V_{ji}^{CKM\star}u_j) + iY_\mu^0\overline{u}_i\gamma^\mu(U_i+\delta_{u_i}u_i) + \text{h.c.}\right] \tag{10.33}$$

The couplings give rise to the following possible decay modes

$$U^i \to X^+ d^i \quad \text{and} \quad U^i \to \overline{Y}^0 u^i \tag{10.34}$$

Since the X, Y gauge bosons are heavy in comparison to the SM ones, these modes have a very small branching ratio. In fact, the X and Y bosons may be so heavy that their production from heavy quark decays are completely kinematically forbidden.

Quark Couplings to Scalars

For completeness we shall discuss the couplings of the quarks to the heavy scalars. We shall identify $\lambda_U = \lambda_2^{u1}$,$\lambda_C = \lambda_2^{u2}$, $\lambda_1 = \lambda_1^{u3}$ and $\lambda_2 = \lambda_2^{u3}$. Then the coupling of the heavy quarks to the scalar field η is given by

$$\begin{aligned}\mathcal{L}_{UU,\eta} &= -\left(i\eta U^c U\right)\frac{c_\beta}{\sqrt{2}}\lambda_U - \left(i\eta C^c C\right)\frac{c_\beta}{\sqrt{2}}\lambda_C + \left(i\eta T^c T\right) s_\beta c_\beta \left(\lambda_1^2 - \lambda_2^2\right)\frac{f}{\sqrt{2}m_T} \\ &\quad + \left(hT^c T\right)\frac{v}{f}\left[\left(\lambda_1^2 s_\beta^2 + \lambda_2^2 c_\beta^2\right)\frac{f}{2m_T} - s_\beta^2 c_\beta^2\left(\lambda_1^2 - \lambda_2^2\right)^2 \frac{f^3}{2m_T^3}\right] + \text{h.c.} \qquad (10.35)\end{aligned}$$

The leading order couplings of the scalar to one heavy partner and one up-type SM quark are

$$\begin{aligned}\mathcal{L}_{uU,\eta} &= -\left(hU^c u\right)\frac{c_\beta \lambda_U}{\sqrt{2}} - \left(hC^c c\right)\frac{c_\beta \lambda_C}{\sqrt{2}} \\ &\quad + \left(hT^c t\right)\left(\lambda_1^2 - \lambda_2^2\right)\frac{s_\beta c_\beta f}{\sqrt{2}m_T} - \left(i\eta t^c T\right)\frac{m_t}{v} + \text{h.c.} \qquad (10.36)\end{aligned}$$

where we have neglected terms of order v/f.

Finally, the couplings of the scalars to two SM up-type quarks are

$$\begin{aligned}\mathcal{L}_{uu,\eta} &= \left(hu_i^c u_i\right)\left\{\frac{-m_{u_i}}{v}\left[1 - \frac{v^2}{6f^2}\left(\frac{s_\beta^4}{c_\beta^2} + \frac{c_\beta^4}{s_\beta^2}\right) - \delta_{u_i}\frac{v}{2\sqrt{2}f}\left(\frac{c_\beta^2 - s_\beta^2}{s_\beta c_\beta}\right)\right] + \frac{m_{U_i}}{v}\delta_{u_i}\delta_{u_i^c}\right\} \\ &\quad + \left(i\eta u_i^c u_i\right)\frac{m_{u_i}}{v}\left[\frac{v}{\sqrt{2}f}\left(\frac{s_\beta^2 - c_\beta^2}{s_\beta c_\beta}\right) + \delta_{u_i}\right] + \text{h.c.} \qquad (10.37)\end{aligned}$$

where we have introduced the mixing between u_i^c and U_i^c at order v^2/f^2 given by $U_i^c = U_{i0}^c - \delta_{u_i^c} u_{i0}^c$ with

$$\delta_{u_i^c} = \frac{m_{u_i}}{m_{U_i}}\left[\delta_{u_i} + \frac{v}{2\sqrt{2}f}\left(\frac{c_\beta^2 - s_\beta^2}{s_\beta c_\beta}\right)\right] \qquad (10.38)$$

analogous to (10.28) for the left-handed fields.

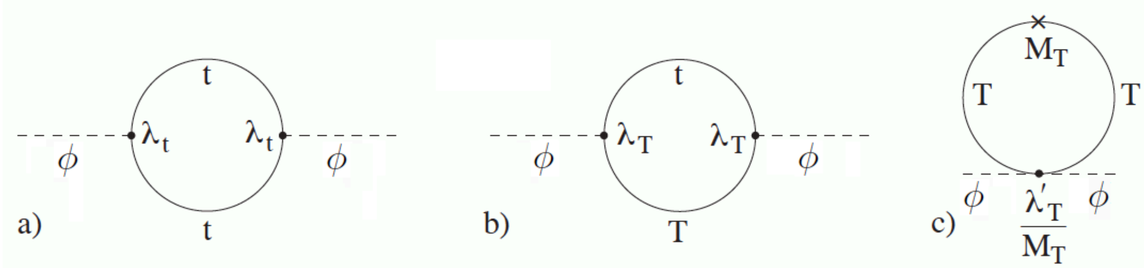


Figure 10.1: Loop diagrams from the top sector that give rise to the quadratic divergences of the Higgs mass.

10.5 The Quadratic Divergence and the Heavy T Couplings to the Higgs

The part of the Lagrangian that describes the coupling of the heavy quark T to the Higgs is

$$\mathcal{L}_{HT} = \lambda_t H t^c t + \lambda_T H T^C t + \frac{\lambda_T'}{2m_T} H H T^C T + \text{h.c.} \tag{10.39}$$

This gives rise to three diagrams that contribute at one loop level to the Higgs mass. One for each of these interaction types. These are shown in Fig. 10.1.

One can immediately observe from the coupling strengths in the Lagrangian that these contributions are of the form

$$(a) \propto \frac{\lambda_t^2}{(4\pi)^2}\Lambda^2 \tag{10.40a}$$

$$(b) \propto \frac{\lambda_T^2}{(4\pi)^2}\Lambda^2 \tag{10.40b}$$

$$(c) \propto -\frac{\lambda_T'}{2m_T\,(4\pi)^2}\Lambda^2 \tag{10.40c}$$

where (c) contains an additional negative sign since it only has one fermion in a closed loop [8]. For these contributions to cancel explicitly the following relation must hold [165],

$$\lambda_T' = \lambda_t^2 + \lambda_T^2 \tag{10.41}$$

In $SU(3)$ simple group models [162],

$$\begin{aligned}
\lambda_t &= m_t/v \\
\lambda_T &= s_\beta c_\beta \left(x_\lambda - x_\lambda^{-1}\right) m_t/v \\
\lambda_T' &= \left[s_\beta^2 c_\beta^2 \left(x_\lambda - x_\lambda^{-1}\right)^2 + 1\right] m_t^2/v^2
\end{aligned} \tag{10.42}$$

with $x_\lambda = \lambda_1/\lambda_2$. Thus, when $x_\lambda = 1$ and λ_T vanishes, the remaining one-loop contributions cancel exactly.

This relationship can be tested by experimental measurements, specifically through independent measurements of λ_T and λ_T' directly. The former governs the cross section of the $T \to bW$ fusion and the latter $T \to tH$ production.

10.6 The Heavy Top-Like Quark T at the LHC

Since the top-partner T is phenomenologically the most interesting we shall concentrate on it here.

The production of the T-quark can occur through QCD interactions and also through single T production via Wb fusion. At high T masses the latter is dominant, as the production of two heavy particles becomes kinematically disfavoured. However, since we are currently probing the regime $m_T \lesssim 1$ TeV, where new physics will likely enter, we shall focus our attention on the pair production mechanism.

The decay widths of the T-quark to the SM products bW, tH, and tZ are the three dominant decay modes in all Little Higgs models since the coupling to the heavier particles is kinematically disfavoured and is suppressed by at least a factor of the mass of the decay product. These couplings to the SM gauge bosons arise from (10.29), (10.30) and (10.39). One can express the field T in terms of the gauge fields T_0 and t_0 before symmetry breaking as

$$T = T_0 - \delta_T t_0 \tag{10.43}$$

where $\delta_T \sim \lambda_T v/m_T$ from (10.28).

Substituting this into (10.29) and (10.30), gives rise to the couplings of T-quark to the SM final states bW and tZ respectively and the coupling to tH comes from (10.39). The decay modes are then of the same form as the corresponding top decays in the SM except that it is suppressed by the factor δ_T. Notice that FCNCs are also introduced into the theory through (10.30).

Then, using our canonical procedure described in Section 2.10.1 and taking into account the additional polarisation factor of $\left(\frac{m_T}{2}\right)^2$ for each boson and a factor 2 for the spin of each fermion, we see that the partial decay widths (at leading order) are

$$\Gamma(T \to tH) = \Gamma(T \to tZ) = \frac{1}{2}\Gamma(T \to bW) = \frac{\lambda_T^2}{32\pi} m_T \tag{10.44}$$

where we have taken the massless limit of the final state SM particles [162].

Assuming that these are the only relevant decay modes, then the total (leading order) decay width is $\frac{5}{64\pi}\lambda_T^2$ with ratios

$$\text{BR}\left(T \to tH\right) = \text{BR}\left(T \to tZ\right) = \frac{1}{4} \quad \text{and} \quad \text{BR}\left(T \to bW\right) = \frac{1}{2} \tag{10.45}$$

Notice that this simple relation, which stems from Goldstone's equivalence principle, is independent of the exact Little Higgs model considered [166]. At high energies and large m_T, the $SU(2)_L$ symmetry holds exactly and the four equivalent degrees of freedom in the Higgs doublet H, which give rise to the longitudinal components of the weak gauge bosons upon SSB, correspond to the three decay modes of the T-quark [162].

Experimental Signatures

The key experimental signatures for the heavy T-quark arise from its dominant decay modes. In contrast to the sequential fourth generation t' (that decays via $t' \to bW$ discussed in Chapter 9 [1]), the Little Higgs T is vector-like and has the additional decay modes through $T \to tH$ and the FCNC $T \to tZ$ usually forbidden by electroweak symmetry.

Further, many other BSM theories predict these heavy sets of quarks [167–180]. In contrast to these models, the exact branching ratios in the class of Little Higgs theories are model independent up to $\mathcal{O}\left(m_t^2/m_T^2\right)$ corrections. Thus, the measurement of these branching ratios (10.45) provide an ideal test for distinguishing the Little Higgs class of theories.

In Chapter 11, we search for this heavy top-like quark that decays into either bW, tH, or tZ final states by scanning across all possible branching ratios. The scan that we preform is thus equally sensitive to all theoretical models that contain such quarks.

Chapter 11

The Search for a Heavy Top-Like Quark

Numerous BSM theories introduce a heavy top-like quark T that typically decays into either a b-quark and a W-boson, a t-quark and a Z-boson, or a t-quark and a Higgs H-boson, with the relative branching ratios set by the specific model.

Along with fourth generation chiral models described in Chapter 8, there are many well-motivated theories that predict vector-like quarks. For example, the Little Higgs Models often include a vector-like $SU(2)_L$ singlet quark to soften the quantum corrections from the SM particles. This class of theories is discussed in Chapter 10 [158–165].

Other models that introduce similar heavy quarks include: composite/holographic Higgs models [167, 168] where the third generation quarks live in an enlarged multiplet of the custodial symmetry of the SM; higher dimensional, Kaluza-Klein [169] and Randall-Sundrum [170] models that generalise the SM, where the SM fields propagate in the additional spatial dimensions; top see-saw models that explain the large mass of the top-quark [171, 172]; top-colour models [173, 174] that provide a mechanism for dynamically breaking the electroweak symmetry; certain supersymmetric models that follow the principle of protecting against baryon violating operators [175]; and even some Grand Unification theories [176]. A general introduction to heavy top-like quark predictions are outlined in Ref. [177–180].

The decay width of the strongly pair-produced T-quark is shared between the six channels shown in Fig. 11.1.

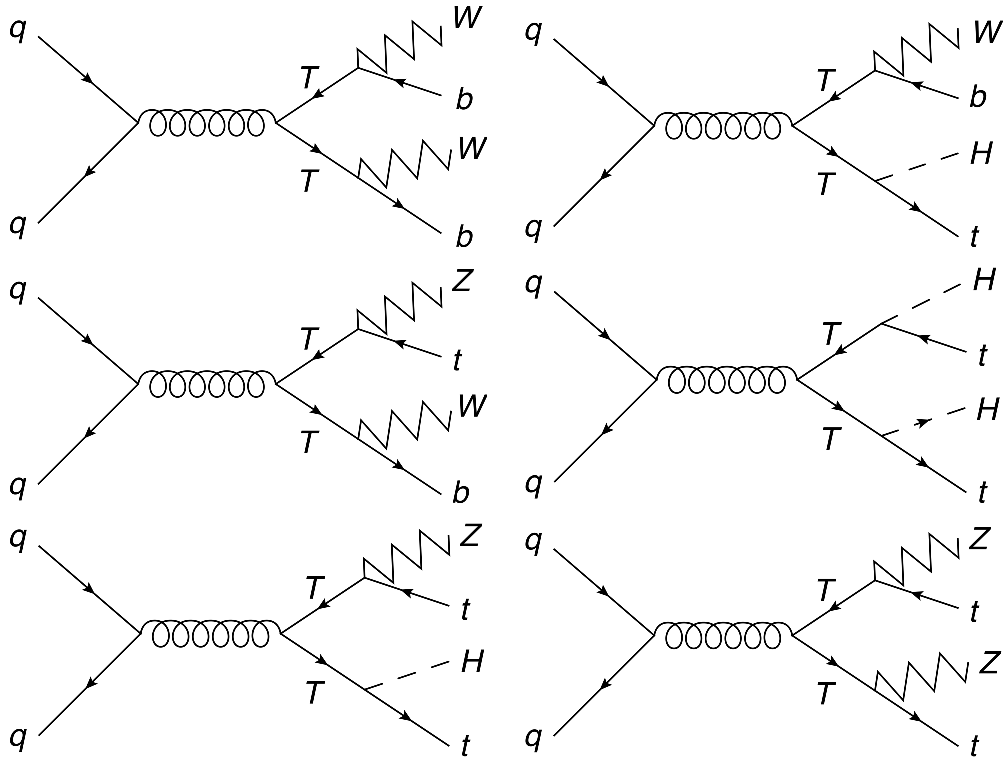


Figure 11.1: The Feynman diagrams for the six leading order decay modes of the heavy vector-like quark T.

In this chapter we present the search for the pair-produced T-quark that decays with a semileptonic topology, consisting of exactly one lepton [2]. To be model independent, we shall scan across the entire phase space of branching ratios by weighting the contributions of the six final states shown in Fig. 11.1.

A partner analysis using multilepton final states, consisting of at least two leptons, can be found in Appendix C.8 [181]. The combination of the semileptonic and multileptonic results is presented in Appendix C.9 [3].

As in the fourth generation t' search (discussed in Chapter 9), the heavy T-quark is assumed to be strongly pair-produced. Thus, the theoretical cross section of the T-quark production is identical to that of the t'; the cross section values at various T-quark masses are tabulated in Table 9.1 and plotted in Fig. 9.1.

11.1 Data

Each recorded event in data is sorted into at least one "dataset" based on the particular trigger path(s) that accepted it. The range of HLT paths per dataset allows us to choose the most relevant trigger and provides the minimum requirements necessary for the reconstruction algorithms to be almost 100% efficient. In this analysis, the single muon and single electron datasets were used, which require events to have at least one isolated muon or one isolated electron. Datasets are further broken down into "runs", corresponding to a continuous period of data taking.

The data used for this analysis were recorded during the whole 2012 running period at a centre of mass energy of $\sqrt{s} = 8\,\text{TeV}$ and amounts to $19.6\,\text{fb}^{-1}$ of integrated luminosity in both the muon and electron channels. The reported luminosity is calculated using only runs that pass a series of quality assurance criteria based on detector performance [140]. The breakdown of the datasets is listed in Table 11.1 and Table 11.2 for the muon and electron channels respectively.

Dataset	Run Range	Lumi [fb^{-1}]
SingleMu/Run2012A-13Jul2012-v1	190456-193621	0.808
SingleMu/Run2012A-recover-06Aug2012-v1	190782-190949	0.082
SingleMu/Run2012B-13Jul2012-v1	193833-196531	4.429
SingleMu/Run2012C-24Aug2012-v1	198022-198913	0.495
SingleMu/Run2012C-PromptReco-v2	198934-203746	6.402
SingleMu/Run2012C-EcalRecover_11Dec2012-v1	201191-201191	0.134
SingleMu/Run2012D-PromptReco-v1	203768-208686	7.274
	Total Luminosity	19.6

Table 11.1: Shows the processed muon datasets recorded during 2012 data taking period at centre of mass energy $\sqrt{s} = 8$ TeV.

The muon (μ+jets) channel uses the single muon triggers named HLT_IsoMu24_eta2p1 over the whole run range 190456-208686. This trigger accepts events with at least one isolated muon with $p_T > 24$ GeV and $|\eta| < 2.1$.

The electron (e+jets) data were collected with the electron trigger HLT_Ele27, which accepts events with at least one isolated electron with $p_T > 27$ GeV.

Dataset	Run Range	Lumi [fb^{-1}]
SingleElectron/Run2012A-13Jul2012-v1	190456-193621	0.808
SingleElectron/Run2012A-recover-06Aug2012-v1	190782-190949	0.082
SingleElectron/Run2012B-13Jul2012-v1	193833-19653	4.429
SingleElectron/Run2012C-24Aug2012-v1	198022-198913	0.495
SingleElectron/Run2012C-PromptReco-v2	198934-203746	6.402
SingleElectron/Run2012C-EcalRecover_11Dec2012-v1	201191-201191	0.134
SingleElectron/Run2012D-PromptReco-v1	203768-208686	7.274
	Total Luminosity	19.6

Table 11.2: Shows the processed electron datasets recorded during 2012 data taking period at centre of mass energy $\sqrt{s} = 8$ TeV.

11.2 Simulated Samples

Simulated Monte Carlo (MC) samples are used to estimate signal efficiencies and SM background contributions.

11.2.1 Signal Samples

In this analysis, the $pp \to T\overline{T}$ parton process is simulated, with up to two additional hard partons, using the MadGraph event generator and the CTEQ6l1 parton distribution functions (PDFs), with each T decaying to either a b-quark and a W-boson, a t-quark and a Higgs H boson or a t-quark and a Z-boson. The six possible combinations of $T\overline{T}$ decay modes ($bWbW$,$bWtH$, $bWtZ$, $tHtH$, $tHtZ$, and $tZtZ$) are generated separately.

The parton level events are then passed to pythia for hadronisation and showering of final state partons [182]. Simulation of particles in the CMS detector is done with geant4 [183] and CMS specific software known as CMSSW_5_3_3.

T benchmark mass points between 500 and 1500 GeV in 100 GeV intervals, are used in this analysis. The signal samples are listed in Table 11.3. The various branching ratios are initially set to the “nominal” branching ratios of BR(bW)=$\frac{1}{2}$, BR(tH)=$\frac{1}{4}$ and BR(tZ)=$\frac{1}{4}$. Unless otherwise stated, all plots will contain a stacked signal template at 800 GeV using these nominal branching ratios and normalised to the integrated luminosity in data using the NNLO $T\overline{T}$ production cross section shown in Table 9.1.

11.2.2 Standard Model Background Samples

Samples of the SM background processes: $t\bar{t}$+jets, W+2jets, W+3jets, W+4jets, Z+jets, $t\bar{t}$W+jets, and $t\bar{t}$Z+jets are generated using the combination of MadGraph interfaced to the pythia parton shower (PS) simulation using the MLM ME-PS matching scheme [182].

We use the MCFM theoretical cross sections of 225.2 pb [184, 185] for $t\bar{t}$+jets and 3532.8 pb for Z+jets. The W+nJets samples (where $n = 2, 3, 4$) are initially normalised to the theoretical leading order value with a correction factor k corresponding to the ratio between the measured W+Jets cross-section 35640 pb [152] and the leading order W+Jets theoretical cross-section 30400 pb [186].

Diboson samples (WW, WZ, and ZZ hereafter collectively labelled as VV) and the $t\bar{t}$H+jets are produced use stand-alone pythia [182]. The diboson cross sections are measured from CMS Data and Monte Carlo samples are 69.9 pb for W^+W^- and 8.4 pb for ZZ [187]. The cross section calculated for $W^{\pm}Z$ at next-to leading-order (NLO) using MCFM is 33 pb [184, 185].

QCD processes are generated by standalone pythia and have LO pythia cross sections.

The remaining single top samples are generated using the event generator Powheg, which generates the hardest (largest momenta) radiation first using exact NLO matrix elements [188]. These are normalised to an approximate NNLO, calculated in [189].

Finally, the hadronic decays of tau leptons that form jets are handled separately by tauola [190].

The cross sections for all SM background processes and the samples used are summarised in Table 11.4. The listed cross-sections are used to normalise the contributions from the different background processes to the luminosity in data. In all figures, unless otherwise stated, the W+nJets, Z+Jets, and all $t\bar{t}$X+jets (where X = H,W,Z) are combined into one histogram, denoted as V+jets. Similarly, the QCD, dibosons, and single-top backgrounds are each summed into their own histograms, denoted QCD, VV, and single-top respectively.

The uncertainty bands on the MC in all plots include the luminosity error of 4.4%, the $t\bar{t}$ normalisation 10% uncertainty and the 30% normalisation uncertainty on all other backgrounds (discussed in Section 11.6). The "pull" on each of these plots takes into account the statistical uncertainty and is defined as

$$\sigma(\text{Data} - \text{MC}) = \frac{\text{Data} - \text{MC}}{\sqrt{\text{MC} + \text{Syst(MC)}}} \tag{11.1}$$

where Syst(MC) accounts for the same systematic uncertainties as the bands on the MC.

mode	Summer 2012 MC Samples
bWbW	TprimeTprimeToBWBW_M-*_TuneZ2star_8TeV-madgraph
bWtH	TprimeTprimeToTHBW_M-*_TuneZ2star_8TeV-madgraph
bWtZ	TprimeTprimeToBWTZ_M-*_TuneZ2star_8TeV-madgraph
tHtH	TprimeTprimeToTHTH_M-*_TuneZ2star_8TeV-madgraph
tHtZ	TprimeTprimeToTHTZ_M-*_TuneZ2star_8TeV-madgraph
tZtZ	TprimeTprimeToTZTZ_M-*_TuneZ2star_8TeV-madgraph

Table 11.3: $T\overline{T}$ simulated signal samples used in this analysis, where * denotes the mass range $500 - 1500$ in 100 (GeV) intervals. All samples are produced using version Summer12_DR53X-PU_S10_START53_V7A-v1.

process	σ (pb)	Summer 2012 MC Samples
$t\bar{t}$ + jets	225.2	TTJets_MassiveBinDECAY_TuneZ2star_8TeV-madgraph-tauola
Single top (t t-channel)	56.4	T_t-channel-DR_TuneZ2star_8TeV-powheg-tauola
Single top (tbar t-channel)	30.7	Tbar_t-channel_TuneZ2star_8TeV-powheg-tauola
Single top (t s-channel)	3.79	T_t-channel_TuneZ2star_8TeV-powheg-tauola
Single top (tbar s-channel)	1.76	Tbar_t-channel_TuneZ2star_8TeV-powheg-tauola
Single top (tW)	11.1	T_tW-channel-DR_TuneZ2star_8TeV-powheg-tauola
Single top (tbarW)	11.1	Tbar_tW-channel-DR_TuneZ2star_8TeV-powheg-tauola
W+2Jets	2051.65	W2JetsToLNu_TuneZ2Star_8TeV-madgraph
W+3Jets	608.46	W3JetsToLNu_TuneZ2Star_8TeV-madgraph
W+4Jets	250.89	W4JetsToLNu_TuneZ2Star_8TeV-madgraph
Z+Jets	3532.8	DYJetsToLL_M-50_TuneZ2Star_8TeV-madgraph-tarball
TTH+jets	1.0	TTH_Inclusive_M-125_8TeV_pythia6
TTW+jets	0.232	TTWJets_8TeV-madgraph
TTZ+jets	0.208	TTZJets_8TeV-madgraph_v2
WW	69.9	WW_TuneZ2star_8TeV_pythia6_tauola
WZ	33.2	WZ_TuneZ2star_8TeV_pythia6_tauola
ZZ	8.4	ZZ_TuneZ2star_8TeV_pythia6_tauola
QCD μ enriched	84679	QCD_Pt-20_MuEnrichedPt-15_TuneZ2_8TeV-pythia6
QCD EM enriched	2920632	QCD_Pt_20_30_EMEnriched_TuneZ2star_8TeV_pythia6
QCD EM enriched	4615893	QCD_Pt_30_80_EMEnriched_TuneZ2star_8TeV_pythia6
QCD EM enriched	183722	QCD_Pt_80_170_EMEnriched_TuneZ2star_8TeV_pythia6
QCD EM enriched	4588	QCD_Pt_170_250_EMEnriched_TuneZ2star_8TeV_pythia6
QCD EM enriched	556.75	QCD_Pt_250_350_EMEnriched_TuneZ2star_8TeV_pythia6
QCD EM enriched	89.1	QCD_Pt_350_EMEnriched_TuneZ2star_8TeV_pythia6

Table 11.4: SM background processes, cross sections, and names of the simulated samples used for background predictions in muon and electron channels. All samples are produced using version Summer12_DR53X-PU_S10_START53_V7A-v1.

11.3 Event Reconstruction

Events are reconstructed using the PF algorithm [191] and all recommended corrections are applied. In this procedure, PF objects are used. All charged PF candidates which are not associated with the main primary vertex are assumed to come from pileup events and are ignored.

The standard jets used in this analysis are found by the anti-k_T jet clustering algorithm with the distance parameter $\Delta R = 0.5$ cone as implented in FASTJET [133]. Jets reconstructed by this algorithm shall be labelled ak5-jets.

To select isolated leptons, the lepton relative isolation I_{rel} is defined as the scalar sum of the p_T of tracks with $p_T > 1\,\mathrm{GeV}$ (excluding the lepton) and all transverse energy in the calorimeter (where transverse energy $E_T = E\sin(\theta)$) within a cone of radius $\Delta R = \sqrt{(\Delta\phi)^2 + (\Delta\eta)^2} < 0.4$ (0.3) for muons (electrons) around the lepton, divided by the lepton p_T -

$$I_{\mathrm{rel}} = \left(I_{\mathrm{track}} + I_{\mathrm{ECAL}} + I_{\mathrm{HCAL}}\right)/p_T \tag{11.2}$$

where $I_{\mathrm{track}} = \sum_{\mathrm{track}} p_T$, $I_{\mathrm{ECAL}} = \sum_{\mathrm{ECAL}} E_T$, and $I_{\mathrm{HCAL}} = \sum_{\mathrm{HCAL}} E_T$.

A muon candidate is considered to be isolated if $I_{\mathrm{rel}} < 0.12$, and an electron candidate if $I_{\mathrm{rel}} < 0.10$. We use the *tight* identification setting for both channels as defined by [192,193].

The probability of b-tagging a jet from a hadronising b-quark is $66.1\pm0.3\%$ [194] and the probability for mistagging jets from light quarks and gluons is $1.3\pm0.2\%$ calculated directly from the entire $t\bar{t}$ MC sample with a scale factor of 0.96 to adjust for the disagreement between data and MC [195].

The p_T and η dependent data-to-MC b-tagging efficiency scale factors and mistag rate scale factors for light jets were applied on a jet-by-jet basis to all b-jets, c-jets and light jets in the MC samples. In order to implement these corrections, the b-tag status was randomly downgraded for heavy flavour jets and upgraded for light flavour jets as recommended by the b-tagging Physics Operating Group (POG) [194].

11.3.1 Jet Corrections

In this analysis, jet energy corrections (JEC) are applied throughout before any other processing occurs. All jets in this analysis are corrected for nonlinear responses in η and p_T, as well as for a residual data-to-simulation discrepancy, the so-called "$L1FastJet$", "$L2Relative$", "$L3Absolute$" and "$L2L3Residual$" corrections labeled for MC and data respectively [134], using version `jec12_V6` and `CMSSW_5_3_3`.

The relative correction takes a jet in a central control region ($|\eta| < 1.3$, the uniformity of the detector) as a reference to remove variations in jet response. The absolute corrections remove variations in jet response as a function of jet p_T. The "$L2L3Residual$" are applied to data to take into account residual differences to simulated jets [134, 196]. Furthermore, we use Type-I corrections for the `PFMET`, which take into account the "$L2L3$" JEC for the jets as described in Ref. [197].

11.3.2 Preselection

Before discussing lepton specific effects, we shall first motivate the set of preselection cuts that are equivalent in both channels; these are kinematic cuts for the jets, lepton and missing energy $\not{E}_T$ that will maximise our sensitivity to the signal. In this section we have relaxed all selection to the minimum required cuts to account for detector effects. The 800 GeV signal templates are stacked using the nominal branching ratios and normalised to the NNLO value with a factor of 5000 for visibility.

Jets coming from heavy T-quarks are expected to have significantly higher transverse momenta than jets typically produced in the background SM processes. The jet momenta of the four leading jets before pre-selection are shown in Fig. 11.2 for the muon channel and Fig. 11.3 for the electron channel.

From these, we determine the following cuts on the four leading (highest p_T) ak5-jets:

- Leading jet $p_T > 120\,\mathrm{GeV}$
- 2nd leading jet $p_T > 90\,\mathrm{GeV}$
- 3rd leading jet $p_T > 50\,\mathrm{GeV}$

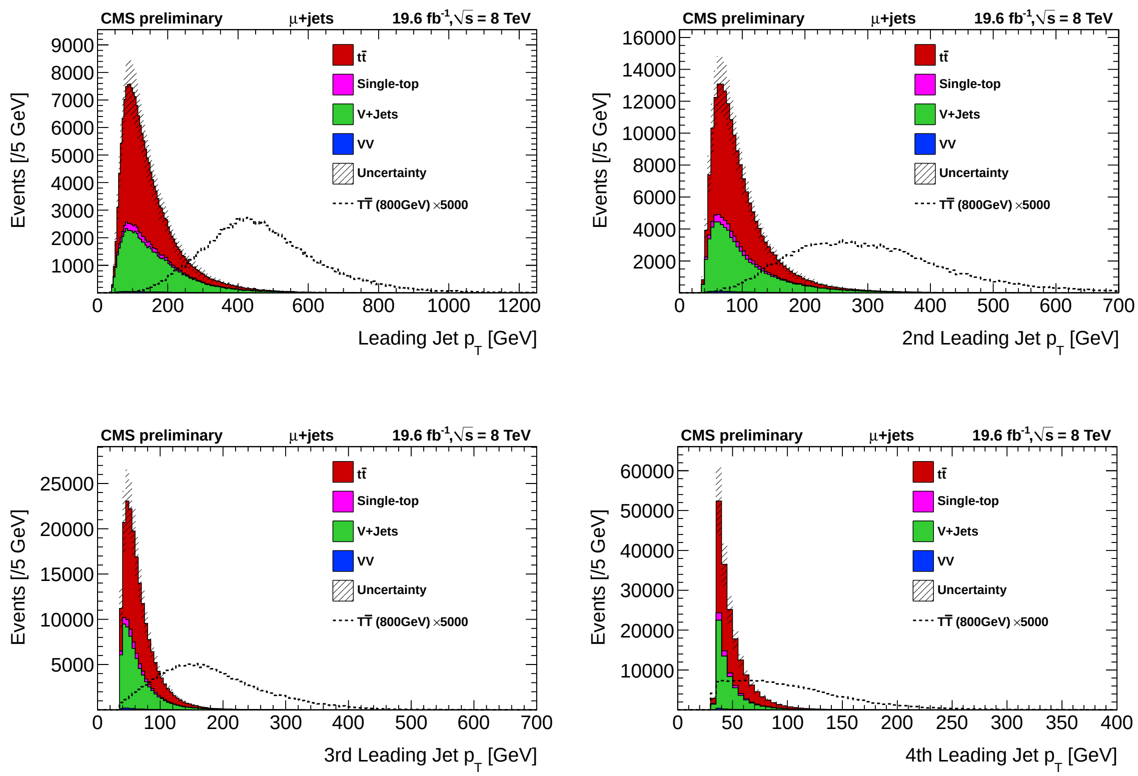


Figure 11.2: Transverse momenta of the four leading jets in the μ+jets channel before preselection. The $T\overline{T}$ signal templates have been stacked using the nominal branching ratios and are scaled to a factor of 5000 of the predicted cross section for visibility.

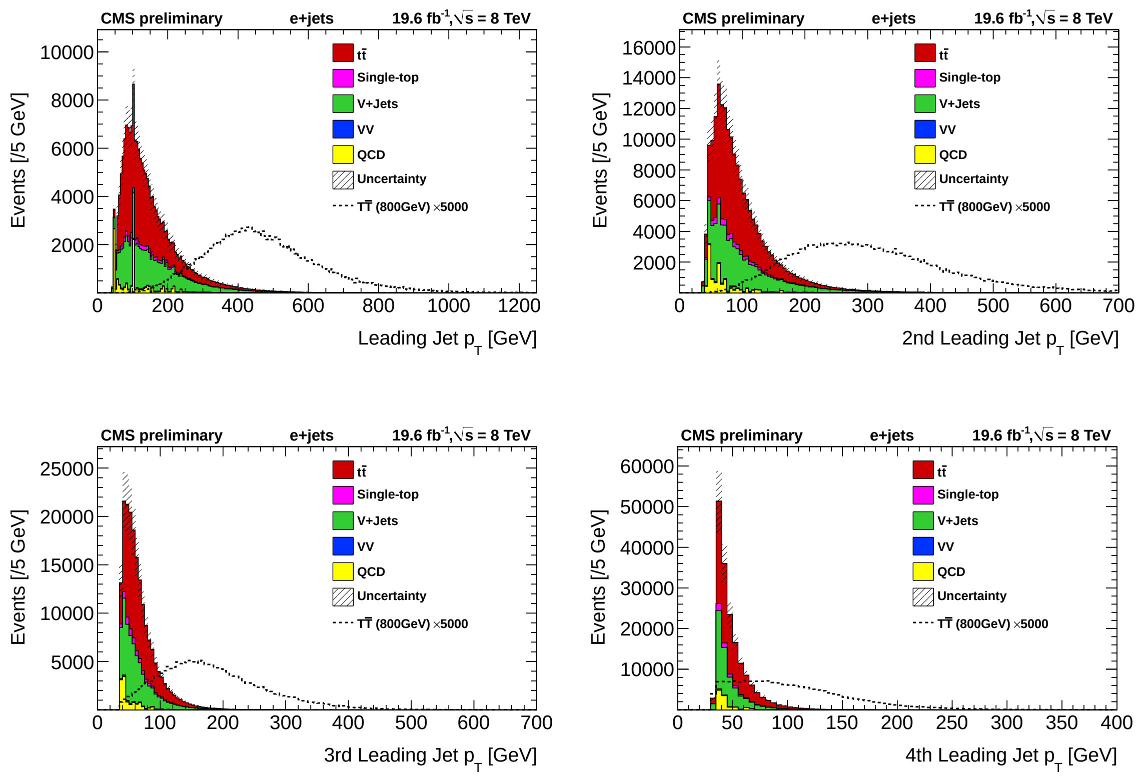


Figure 11.3: Transverse momenta of the four leading jets in the e+jets channel before preselection. The $T\overline{T}$ signal templates have been stacked using the nominal branching ratios and are scaled to a factor of 5000 of the predicted cross section for visibility.

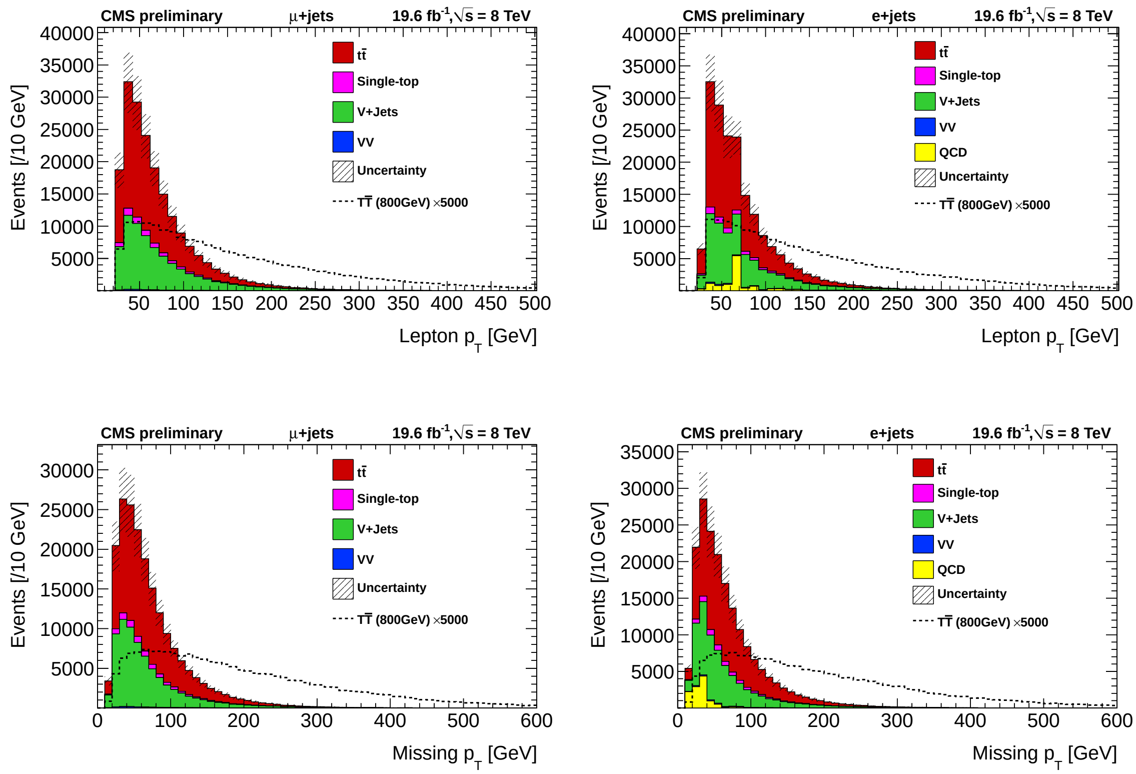


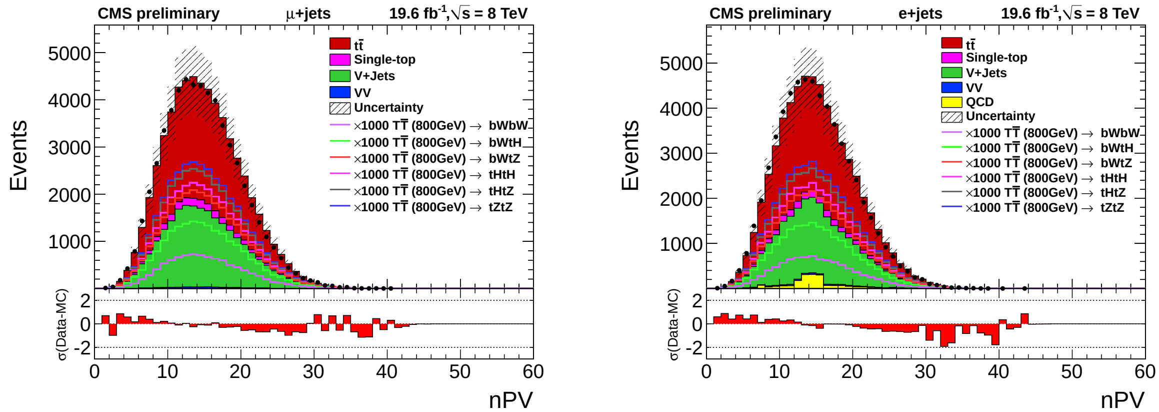
Figure 11.4: The lepton p_T and $\not\!\!E_T$ in the (left) μ+jets and (right) e+jets channels before preselection. The $T\overline{T}$ signal templates have been stacked using the nominal branching ratios and are scaled to a factor of 5000 of the predicted cross section for visibility.

- 4th leading jet $p_T > 35\,\mathrm{GeV}$, when a fourth ak5-jet requirement is used.

Fig. 11.4 shows the distributions of $\not\!\!E_T$ and the lepton p_T before preselection in both channels.

The decays of a pair-produced T-quark via the lepton+jets channel are expected to exhibit significant missing transverse energy $\not\!\!E_T$ due to the undetected neutrino from the leptonically decaying W. A missing transverse energy cut $\not\!\!E_T > 20\,\mathrm{GeV}$ is thus applied to the selected events. In addition, a cut on the lepton transverse momenta is imposed with $p_T > 32\,\mathrm{GeV}$ to account for the HLT trigger turn on curve. In both channels, this cut puts us in the region where the identification of leptons by the trigger is well understood in both the data and MC simulations. We shall discuss the trigger in each channel separately.

Finally, detector constraints require that all jets have $|\eta| < 2.5$, and the muon (electron) $|\eta| < 2.1$ (2.4). These selection cuts, along with the momenta of the jets, lepton and $\not\!\!E_T$



(a) (b)

Figure 11.5: Number of primary vertices after pileup reweighting for (a) μ+jets and (b) e+jets channels.

criteria are imposed throughout the rest of this analysis. The full selection is described in Section 11.4 for the muon channel and Section 11.5 for the electron channel.

11.3.3 Pileup

A pileup correction is applied to the MC events to reproduce the distribution of the number of pileup interactions as measured in data.

The number of added pileup interactions in the MC samples is generated using a flat distribution of pileup interactions plus a Poisson tail for higher number of interactions. This distribution does not correspond to the one observed in data.

Thus, a "pile-up reweighting" procedure is implemented for the MC samples according to the distribution of the number of pile up interactions from the simulation in order to match the pile-up distribution in data [198]. This is estimated by using the per-bunch-crossing instantaneous luminosity. After the reweighting, the agreement between data and MC in the distribution of the number of reconstructed primary vertices is good. The resulting distribution of the number of vertices is shown in Fig. 11.5.

Throughout this analysis, the pileup reweighting is applied to all MC events on an event by event basis.

11.3.4 Substructure

It is expected that at higher T masses the number of reconstructed jets will decrease. This is because the mass of the decaying quark is converted to energy and if the decay products are light in relation to the parent particle, they will have very large momenta. This in turn causes a high degree of collimation of their decay products making it difficult to reconstruct them separately using the standard anti-k_T algorithm. High momenta *boosted* particles coming from the same parent will thus be reconstructed as a single jet.

Therefore, we remove the requirement of a fourth jet and investigate the effect of jet merging as a function of mass. Fig. 11.6 and Fig. 11.7 show the number of jets in muon and electron channels respectively for all six channels after event selection. These figures are area normalised and indicate that a significant proportion of signal events would be lost from requiring a fourth jet in our analysis.

Notice that since the $T \to bW$ has the lightest final state particles, the bW particles are likely to be the most boosted objects and suffer the greatest from the merging of jets.

To take this into account, we shall consider two situations: boosted "W-jets" and boosted "top-jets", collectively known as CA8-jets.

W-jets

The W-boson may arise directly from the T-quark or it may be from a moderately boosted top-quark. We consider the case where the b-quark is reconstructed independently of the W-boson. In the scenario of a boosted "W-jet", the decay products of the W-boson are highly collimated and reconstructed as a single jet. The jet-clustering algorithm that we use is the Cambridge-Aachen (CA) with a cone size $\Delta R = 0.8$. This algorithm implements the procedure from the original theoretical paper Ref. [199]. From these jets, we then use a jet pruning algorithm [137], [138] to identify the hadronic decay of the W-boson.

Jets are required to have loose `PF`-jet identification [196] and the following additional requirements:

- $\mathrm{N}_{\mathrm{subjets}} = 2$, where $\mathrm{N}_{\mathrm{subjets}}$ is the number of subjets that make up the W-jet.
- $p_T > 200\,\mathrm{GeV}$, and $|\eta| < 2.4$ of the W-jet.

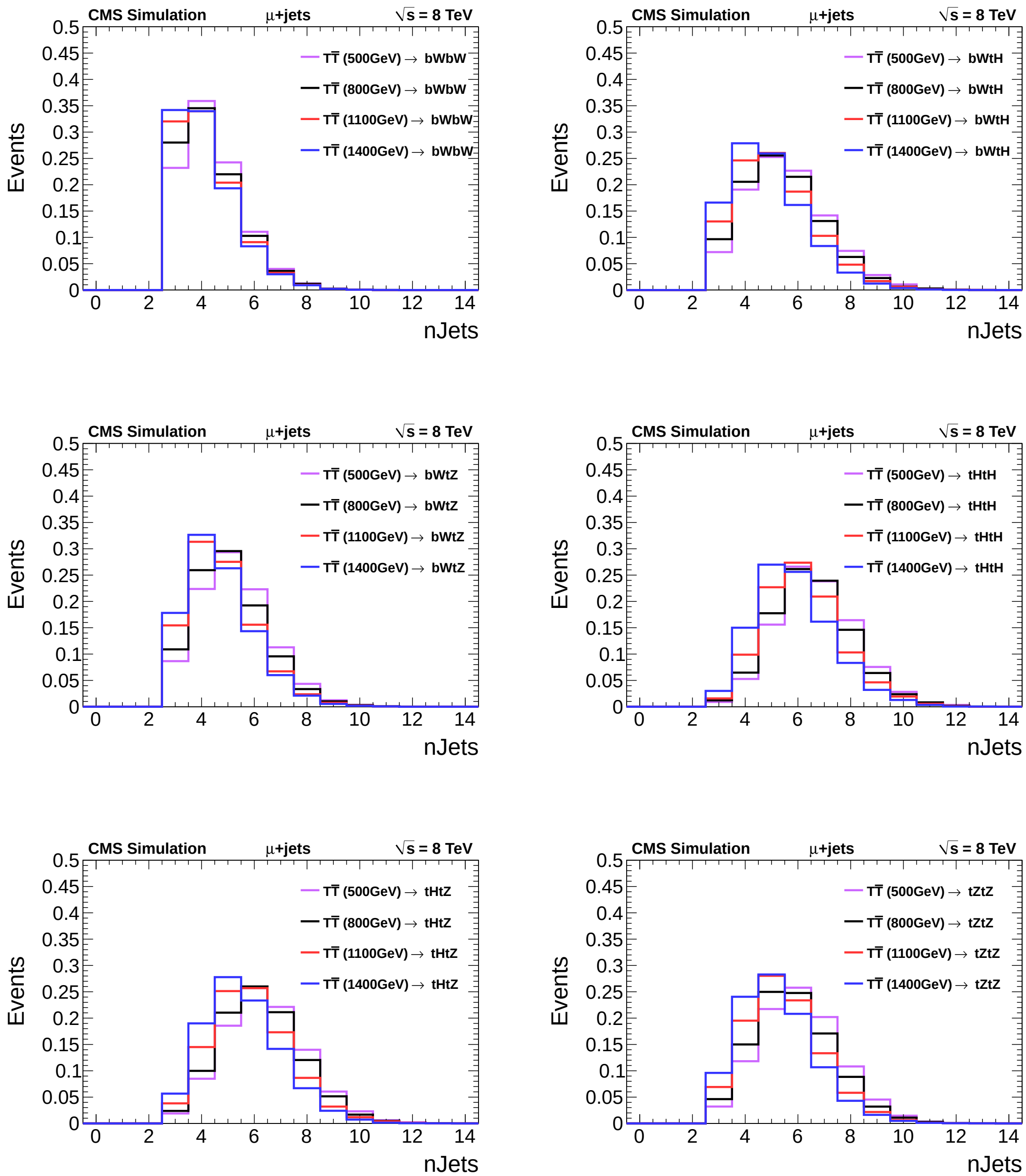


Figure 11.6: Area normalised distribution of the number of ak5-jets in μ+jets events for all decay modes separately.

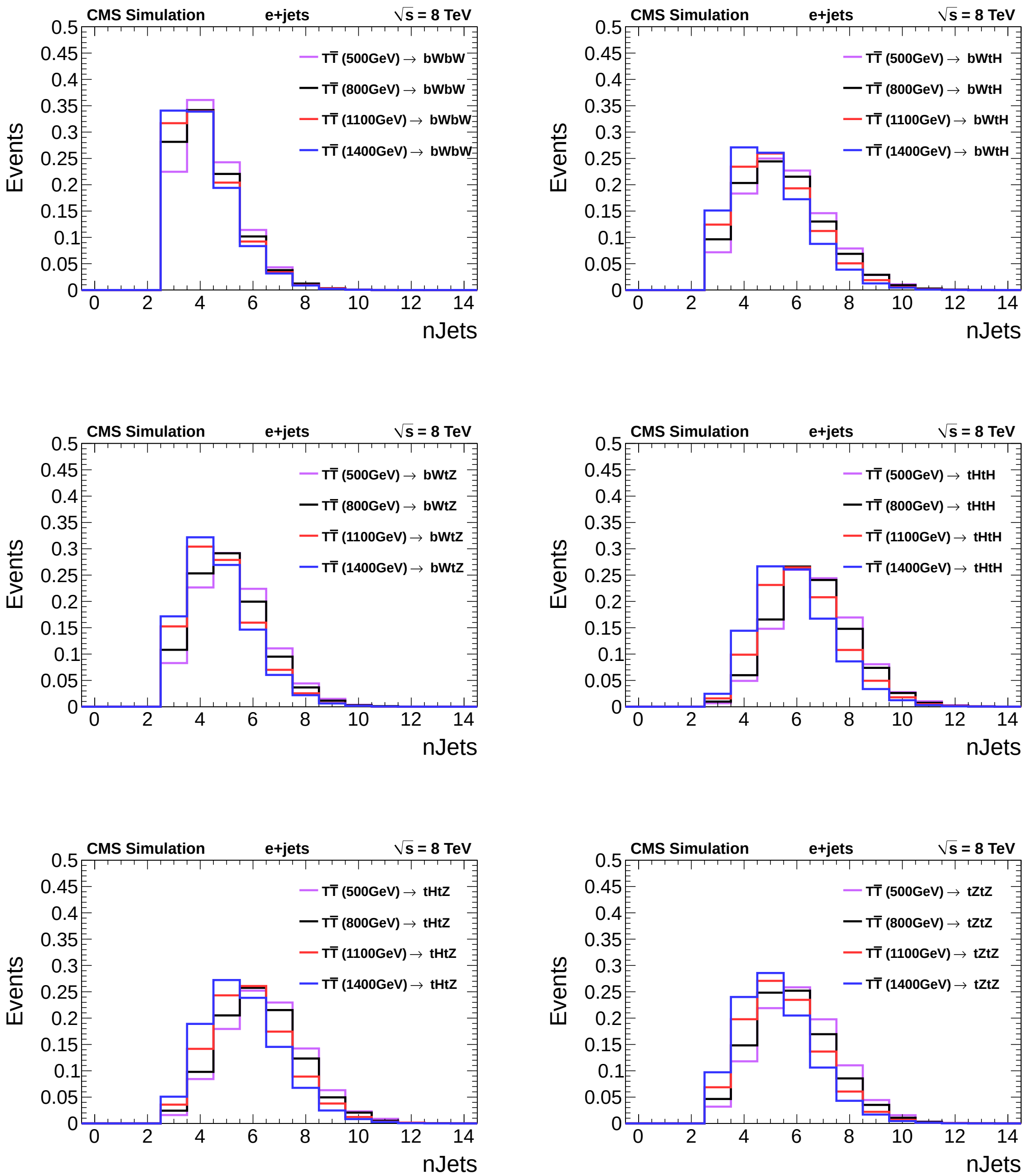
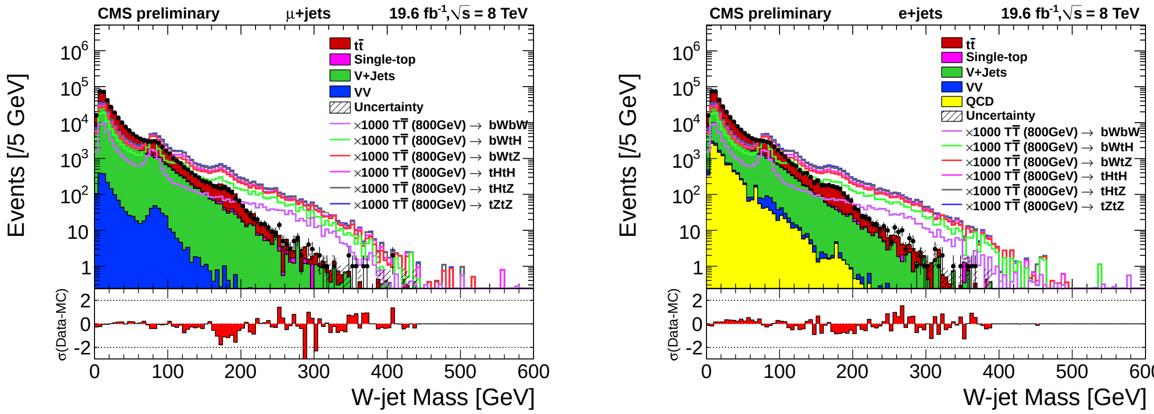


Figure 11.7: Area normalised distribution of the number of ak5-jets in e+jets events for all decay modes separately.



(a) (b)

Figure 11.8: Invariant mass of all W-tagged jets in the (a) μ+jets and (b) e+jets channels.

- $60 < m_{\text{jet}} < 130$, where m_{jet} is the invariant mass of the four-vector sum of the constituents of the hard CA W-jet.

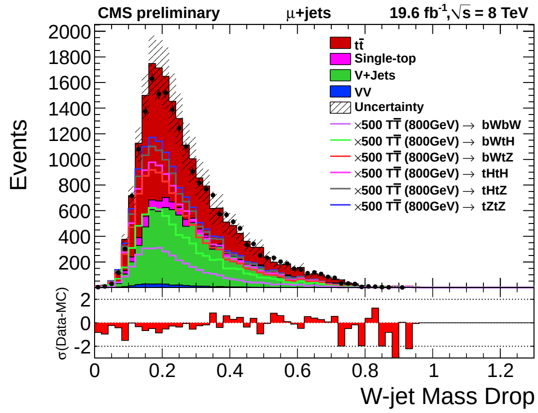
The invariant mass of all W-jets is shown in Fig. 11.8 after relaxing the mass requirement. There is a clear peak at the W-boson mass 80.4 GeV, and also at the t-quark mass at 172.9 GeV. A small excess in the signal distribution is also visible at around the Higgs mass (125 GeV).

The mass-drop defined as the ratio of the mass of the leading (highest p_T) subjet to the mass of the W-jet is shown in Fig. 11.9. An additional cut on the mass-drop in the W-tagged jets was considered but does not provide good discrimination between background and signal and therefore was not included in our selection criteria.

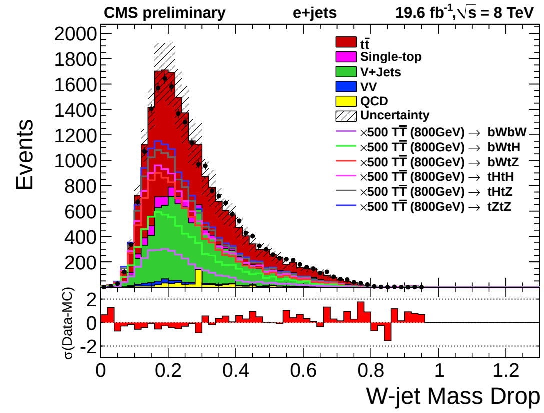
With the definition described above, the number of W-tagged jets and the leading W-jet p_T is shown in Fig. 11.10 for both channels. These figures include all events with and without substructure.

As the collimation increases with higher T-quark masses, the number of W-tagged jets also increases. This is shown for the different signal decay channels in Fig. 11.11 for the muon channel and Fig. 11.12 for the electron channel.

Using the definition of W-tagging, we split all selected events in each channel into two mutually exclusive categories. The first set, which we shall call "substructure events", contain ≥ 3 standard ak5-jets and ≥ 1 W-tagged jets. The second, "without substructure",

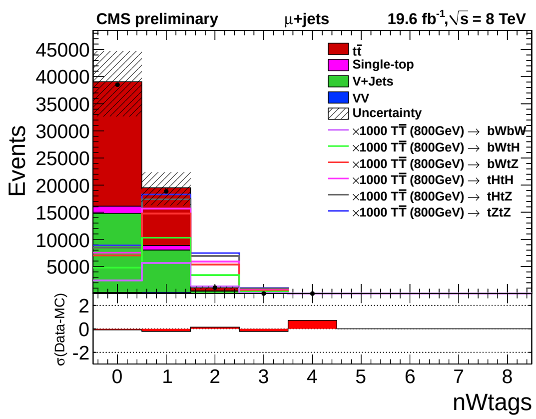


(a)

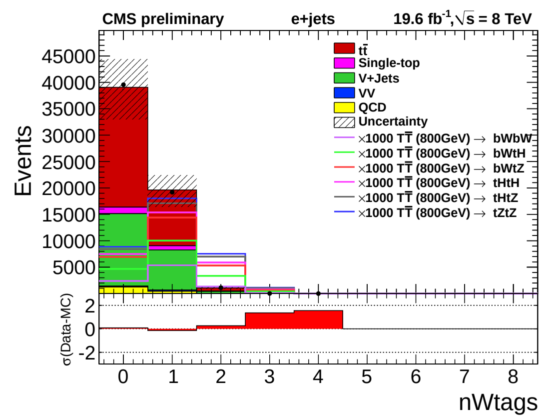


(b)

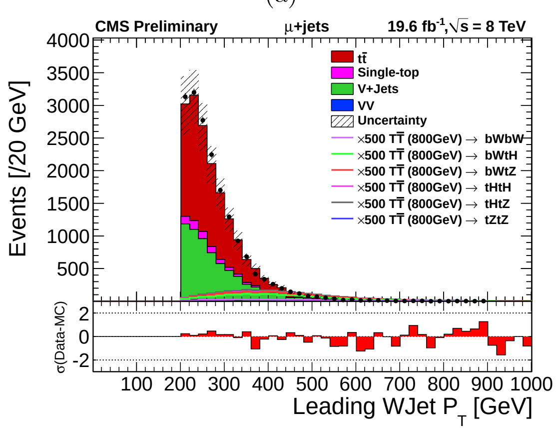
Figure 11.9: Mass-drop, defined as ratio of the mass of the leading subjet of the leading W-jet, in the (a) μ+jets and (b) e+jets channels.



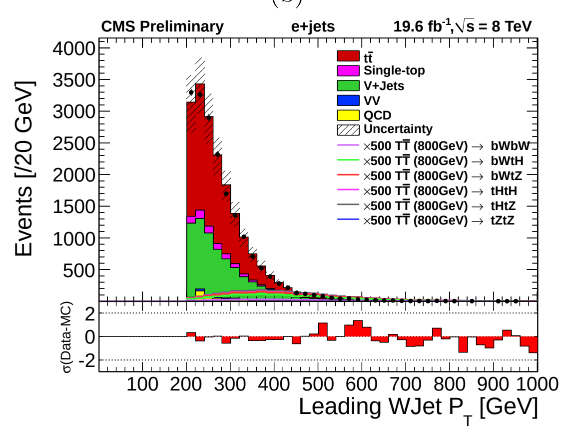
(a)



(b)



(c)



(d)

Figure 11.10: Number of W-tagged jets in the (a) μ+jets and (b) e+jets channels and the leading W-jet transverse momentum p_T in the (c) μ+jets and (d) e+jets channels.

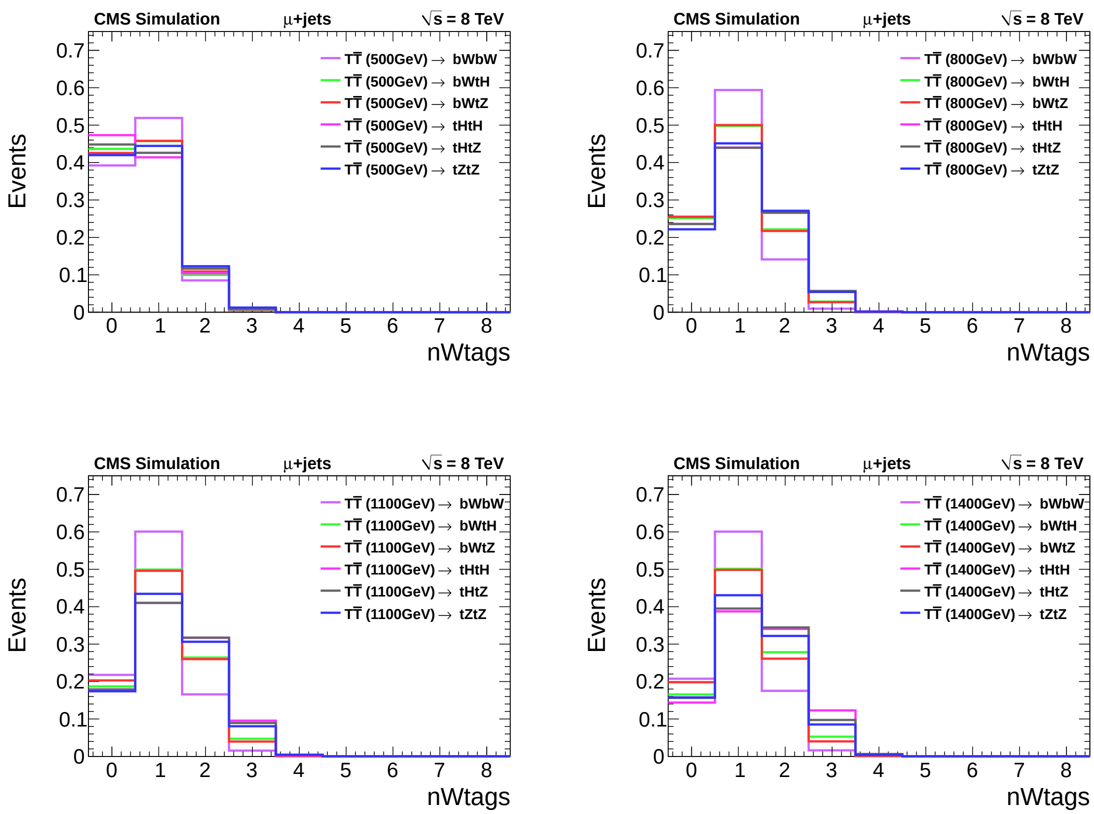


Figure 11.11: Area normalised number of W-tagged jets for all signal decay modes at 500, 800, 1100, 1400 GeV in the μ+jets channel.

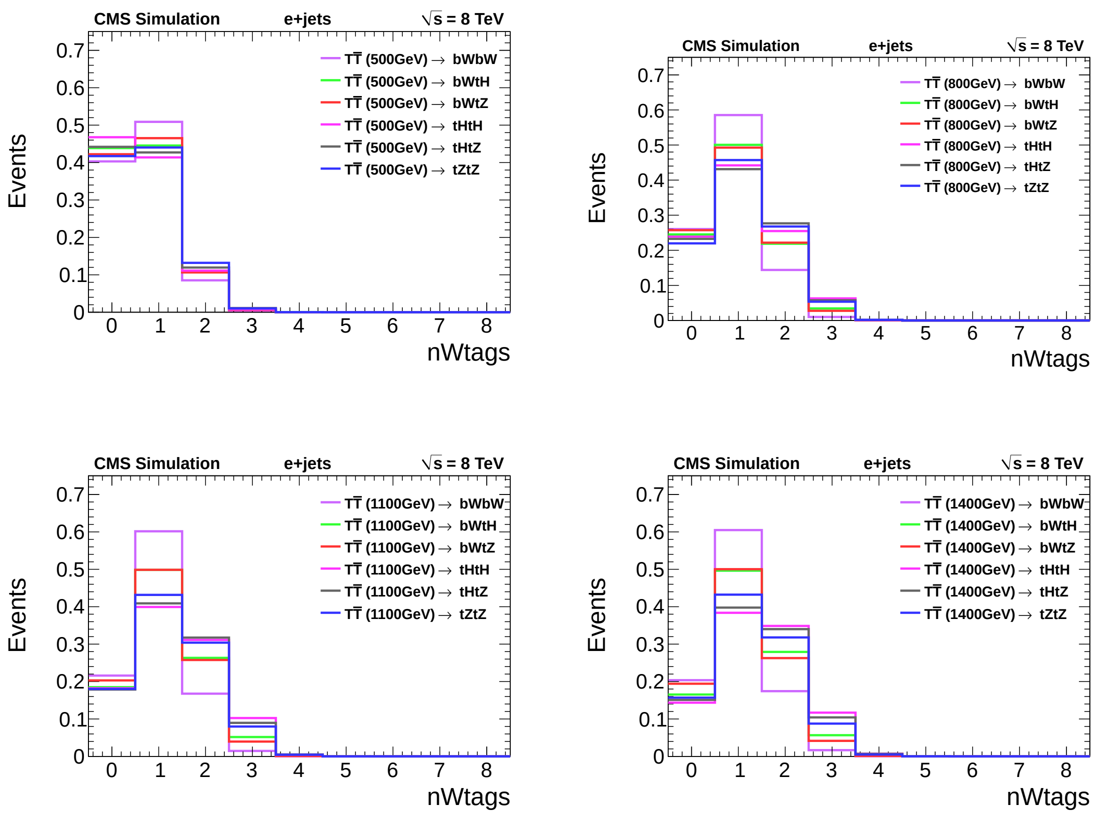


Figure 11.12: Area normalised, number of W-tagged jets for all signal decay modes at 500, 800, 1100, 1400 GeV in the e+jets channel.

corresponds to events with $\geq$ 4 ak5-jets and no W-tagged jets. These distinct categories of events shall be statistically combined at the very end in the limit setting procedure. However, for conciseness, all plots unless otherwise stated will contain all events with and without substructure per lepton channel.

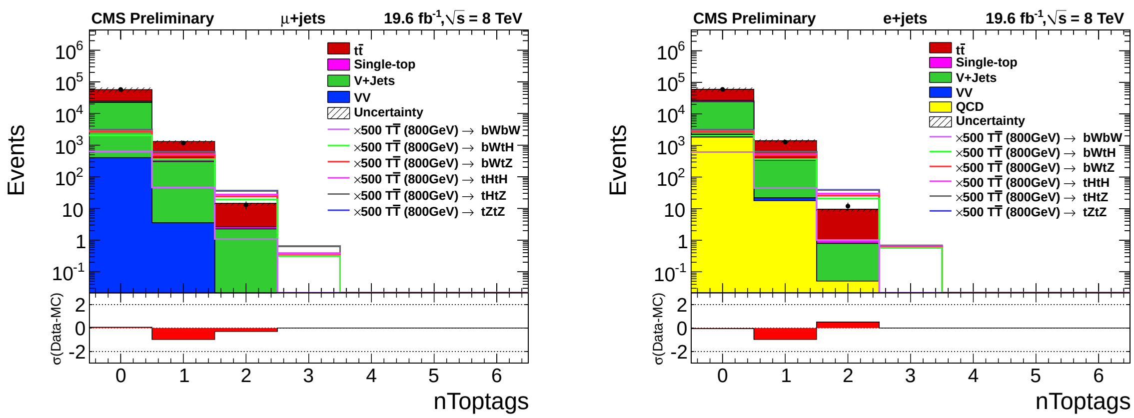
Top-jets

The number of "top-jets" in the event will be used to discriminate signal from background in the statistical procedure. In this case, all decay products of a top-like quark are merged into one jet. We use a top-tagging algorithm based on identifying jet substructure [137] and follow the implementation in [138].

Top-jets are clustered using the Cambridge-Aachen clustering algorithm with a cone of size $\Delta R = 0.8$ [199] as implemented in `FASTJET` version 3 [133]. As with the W-jets, top-jet candidates are required to have loose `PF`-jet identification [196] and the following additional requirements:

- $\mathrm{N_{subjets}} \geq 3$, where $\mathrm{N_{subjets}}$ is the number of subjets of the top-jet.
- $p_T > 200$ GeV, $|\eta| < 2.4$ of the top-jet.
- $140 < m_{\mathrm{jet}} < 250$ GeV where m_{jet} is the invariant mass of the four-vector sum of the constituents of the hard CA top-jet.
- $m_{\mathrm{min}} > 50$ GeV where m_{min} is the minimum pair-wise mass. For defining m_{min} the three highest p_T subjets are taken pairwise. For each pair the invariant mass is calculated as $m_{ij} = \sqrt{(E_i + E_j)^2 - (\bar{p}_i + \bar{p}_j)^2}$ and the minimum $m_{\mathrm{min}} = \mathrm{min}[m_{12}, m_{13}, m_{23}]$ is taken.

With the above definition, the number of top-tagged jets is shown in Fig. 11.13 for both channels. The expected number of background events with large top-jet multiplicity is small corresponding to the small number of decay process that contain a top-quark. These figures include all events with and without substructure.



(a) (b)

Figure 11.13: The number of top-tagged jets in the (a) μ+jets and (b) e+jets channels.

The number of top-tagged jets in the different signal decay channels are shown in Fig. 11.14 and Fig. 11.15 for the muon and the electron channels respectively. As in the case for the W-jets, there are more merged top-jets scenarios with higher T-quark masses.

Finally, since no CA8 jet corrections are currently available, the W-jets and top-jets used in this analysis use the (additional) `ak7-CHS` jet correction [200] and a flat 3% uncertainty on the CA8-jet tagging efficiency [201].

11.3.5 W+Jets Scale Factors

The branching ratios of the W+nJets samples into its decay products is poorly understood in the corner of phase space of our analysis. W+Jets scale factors are therefore applied to all W+nJets MC events to match the events observed in data.

We start with the SHyFT scale factors for light, b, and c-jet events calculated at $\sqrt{s} = 7$ TeV in 2011 [202]. We then apply a correction to account for the increase of centre of mass energy to $\sqrt{s} = 8$ TeV in 2012. This additional correction consists of two separate scale factors applied using the MC truth information: a heavy scale factor is applied to W+nJets MC events with at least one b-quark and a light scale factor is applied to all other W+nJets events.

To derive these corrections, we define a (mutually exclusive) control region with exactly 3 ak5-jets and exactly 0 W-tagged jets, where all other selection requirements are the same

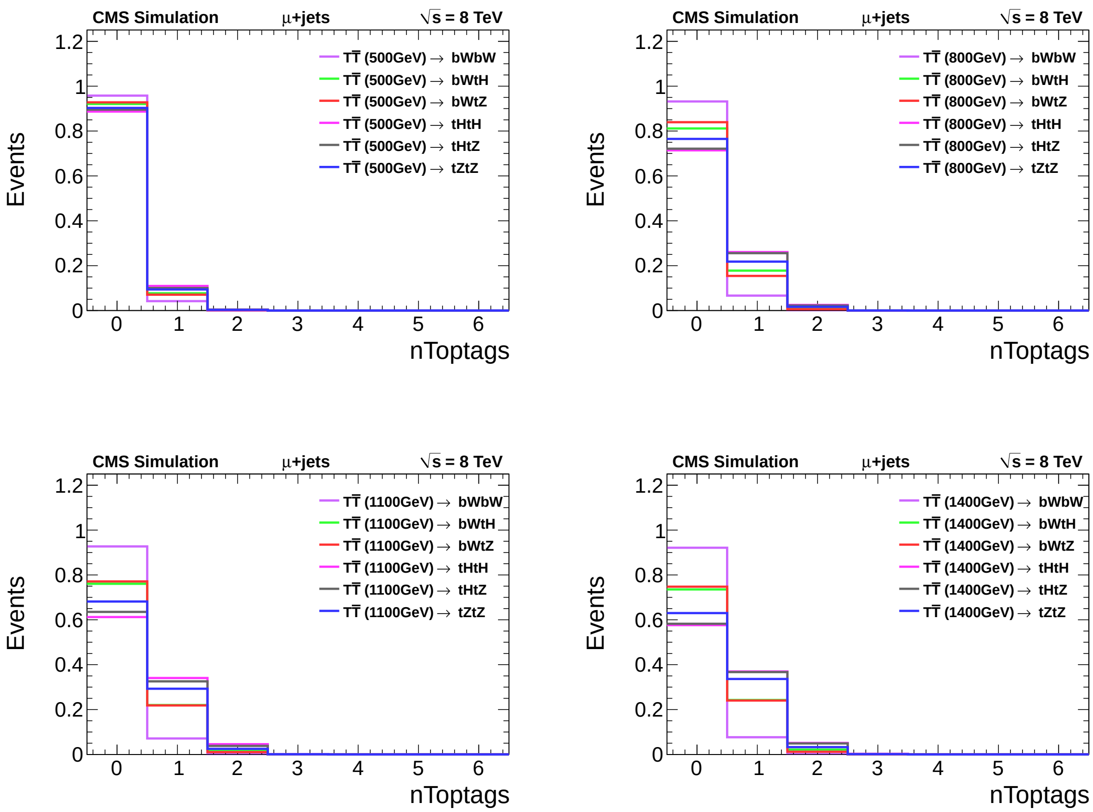


Figure 11.14: Area normalised number of top-tagged jets for all signal decay modes at $500, 800, 1100, 1400$ GeV in the muon channel.

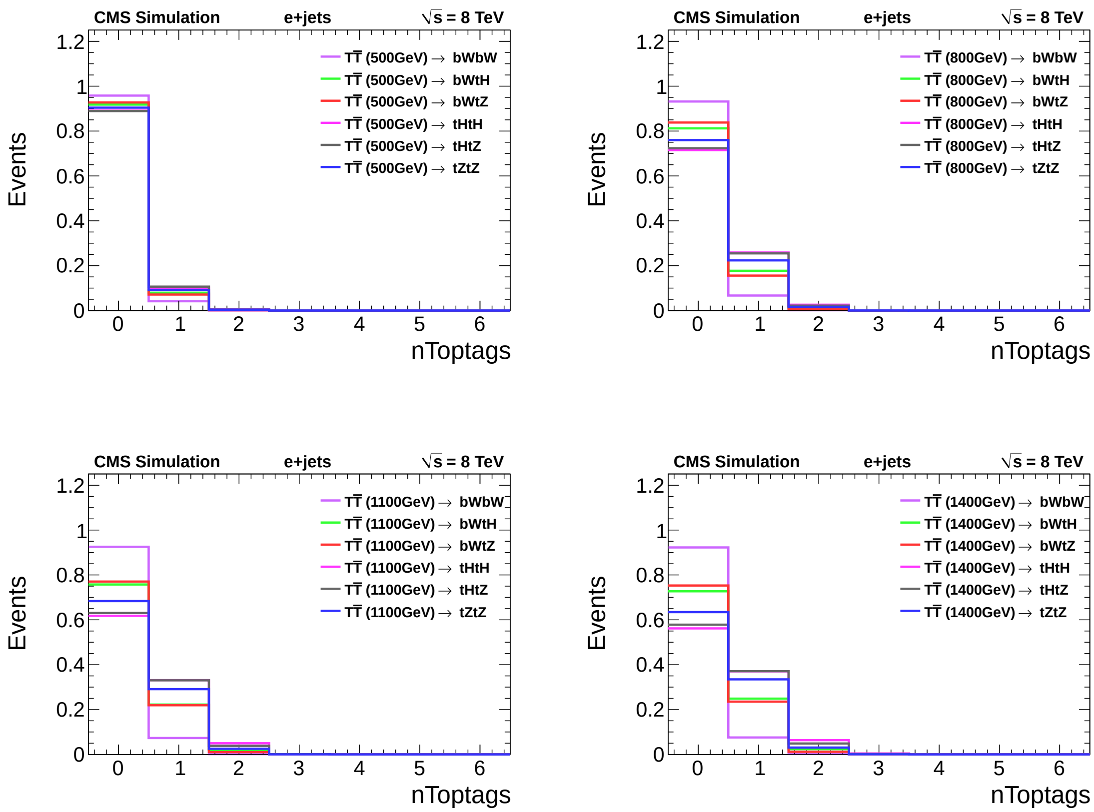


Figure 11.15: Area normalised number of top-tagged jets for all signal decay modes at 500, 800, 1100, 1400 GeV in the electron channel.

as the signal region. We shall derive the corrections using the number of b-tagged jets in the event, using the CSV algorithm *medium* working point. We analytically solve for the two scale factors using the two constraints that the number of events with 0 b-tagged jets and the total number of events in the simulated MC background matches the numbers of events in the data.

This is calculated in both channels independently to account for any differences in the lepton trigger efficiencies and ID efficiencies. There is roughly a 10% uncertainty in these scale-factors, derived by computing the values for the b-tag "up" and "down" template systematics (discussed later in Section 11.6) in a similar manner.

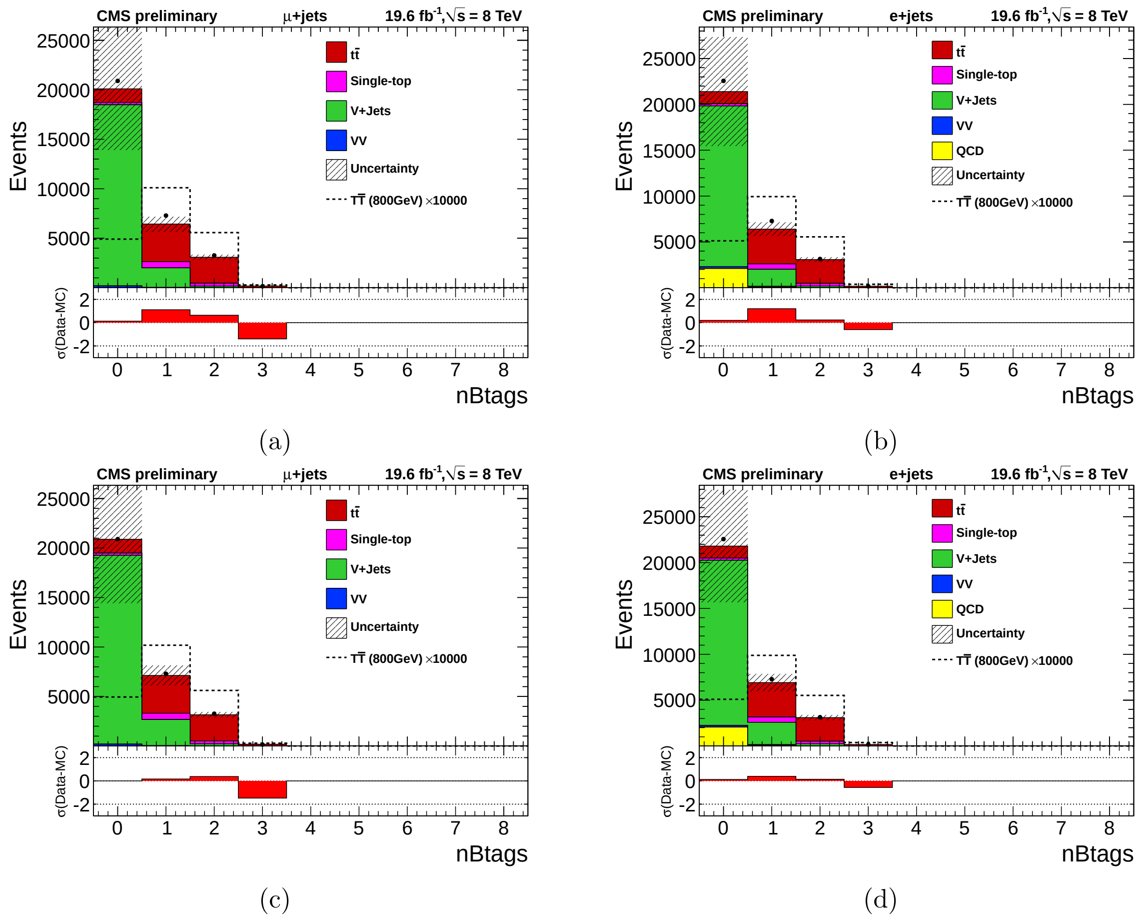
The exact values of the final scale factors applied to the W+nJets samples are listed here for completeness:

- W+light jets : 0.93
- W+c jets : 1.57
- W+b jets : 1.14

The resulting b-tag multiplicities for both channels before and after applying the complete set of scale factors are shown for the control region in Fig. 11.16.

We then test these scale-factors on mutually exclusive events in a second control region. This region has the same selection requirement as the first, except that it has either two or three jets (if the third jet has *less* than 35 GeV). The b-tag multiplicity is shown before and after the application of the W-jet scale-factors in Fig. 11.17. As one can see, it significantly improves the agreement between data and MC simulation.

These scale factors are applied throughout the analysis.



(a) (b) (c) (d)

Figure 11.16: Number of jets with CSVM b-tag in the control region before applying the scale factors in the (a) μ+jets, (b) e+jets channel and after applying the scale factors in the (c) μ+jets, (d) e+jets channel.

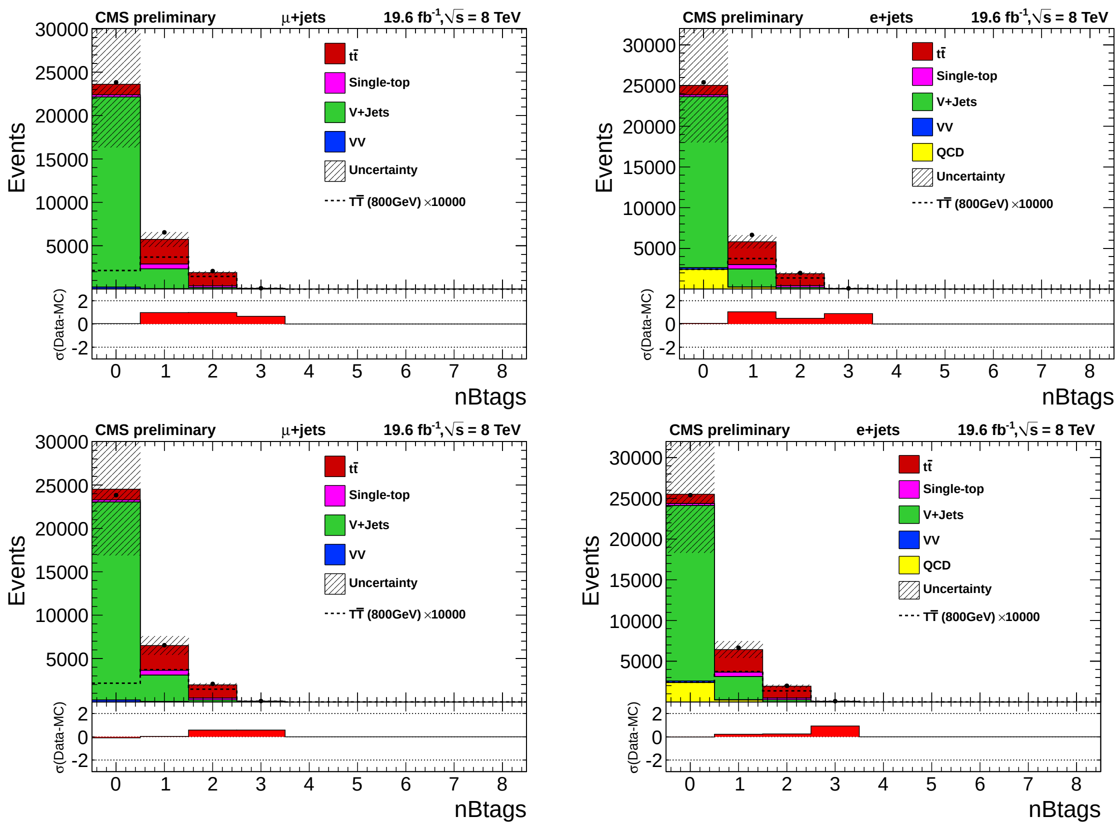


Figure 11.17: Number of jets with CSVM b-tag in the second control region before applying the scale factors in the (a) μ+jets, (b) e+jets channel and after applying the scale factors in the (c) μ+jets, (d) e+jets channel.

11.4 Muon Channel

11.4.1 Trigger and Muon ID

The trigger HLT_IsoMu24_eta2p1 used in the muon (μ+jets) channel is based on the presence of at least one isolated muon with minimum transverse momentum p_T of 24 GeV. This trigger was unprescaled throughout all data-taking in 2012.

The trigger, isolation, and identification efficiencies for simulated MC samples are scaled to the scale factors listed in Table 11.5. These are derived from the full 2012 $ABCD$ datasets and have a flat 3% uncertainty associated with them [192]. This accounts for the reconstruction differences between data and MC of the muon.

11.4.2 Muon Selection

Events in the muon channel are required to satisfy the selection criteria described in detail in the following paragraphs:

- The use of particle-flow (`PF`) muons obtained from the `PF2PAT` algorithm [191]
- Events are required to have at least one primary pp interaction vertex where vertices are identified by applying an adaptive fit to clusters of reconstructed tracks. The primary vertex must be within ± 24 cm of the nominal interaction point in the direction along the proton beams and less than 2 cm away in the plane perpendicular to the beam.
- Require exactly one tight muon in the event [203]
- Muon $p_T > 32\,\text{GeV}$.

$p_T > 20$ GeV	$0 < \lvert\eta\rvert < 0.9$	$0.9 < \lvert\eta\rvert < 1.2$	$1.2 < \lvert\eta\rvert < 2.1$
Trigger SF	0.981	0.961	0.983
Isolation SF	1.000	1.003	1.005
Identification	0.994	0.990	0.997

Table 11.5: The scale factors applied throughout the muon channel provided by the muon POG [192]. A flat 3% uncertainty is applied for these numbers in the statistical procedure.

- Muon $|\eta| < 2.1$, motivated by detector coverage.

- Muons must extrapolate to within $|\mathrm{d}_0| < 0.02\,\mathrm{cm}$ of the interaction vertex in the plane perpendicularlar to the beams and to within $|\mathrm{d}_z| < 1$ cm in the direction along the beam-axis

- Require `PF` based isolation $I_{\mathrm{rel}} < 0.12$ for the muon with 'DeltaBeta' corrections (11.2) [192]. This ensures that the muons from W bosons in top-like decays are isolated from other high-p_T particles in the event.

- Dilepton Veto:

  We veto events with an identified second loose muon. The loose muon selection criteria is defined as Global muons with $p_T > 10.0\,\mathrm{GeV}$ and $|\eta| < 2.5$. They are required to have a relative combined `PF` isolation $I_{\mathrm{rel}} < 0.2$ within a cone of $\Delta R < 0.4$ around the Global muon candidate momentum direction.

  We veto events with an identified loose electron candidate. The loose electron selection criteria is defined using the Cut Based electron identification 'Loose' working point [193]. There is a relative isolation requirement of $I_{\mathrm{rel}} < 0.15$. No conversion rejection cuts are applied. These loose electron candidates are required to have a $E_T > 20.0$ GeV and $|\eta| < 2.5$ with the exclusion of the transition region of the calorimeter.

- We require events that contain at least three ak5-jets with leading jet $p_T > 120\,\mathrm{GeV}$, second leading jet $p_T > 90\,\mathrm{GeV}$, third leading jet $p_T > 50\,\mathrm{GeV}$. In events without W-jets a fourth ak5-jet with $p_T > 35\,\mathrm{GeV}$ is required.

- We use the anti-k_T clustering algorithm for the standard `PFJets`, which must be within the η-range of $|\eta| < 2.4$. The η-range is driven by the acceptance of the silicon tracker to allow the use of b-tagging. Additionally, jets that are close to the muon, within $\Delta R < 0.4$, are discarded.

- The selection criteria for W-jets and top-jets are described in Section 11.3.4 and Section 11.3.4 respectively.

- b-tagging: We use the Combined Secondary Vertex (CSV) b-jet tagging algorithm [136]. The use of b-tagging information can further discriminate signal and background events due to the high expected b jet multiplicity in signal. The CSV b-tagging algorithm with the medium working point was used.

- To suppress QCD backgrounds we apply a $\not\!\!E_T > 20\,\text{GeV}$ cut. No QCD events pass all selection criteria in the muon channel.

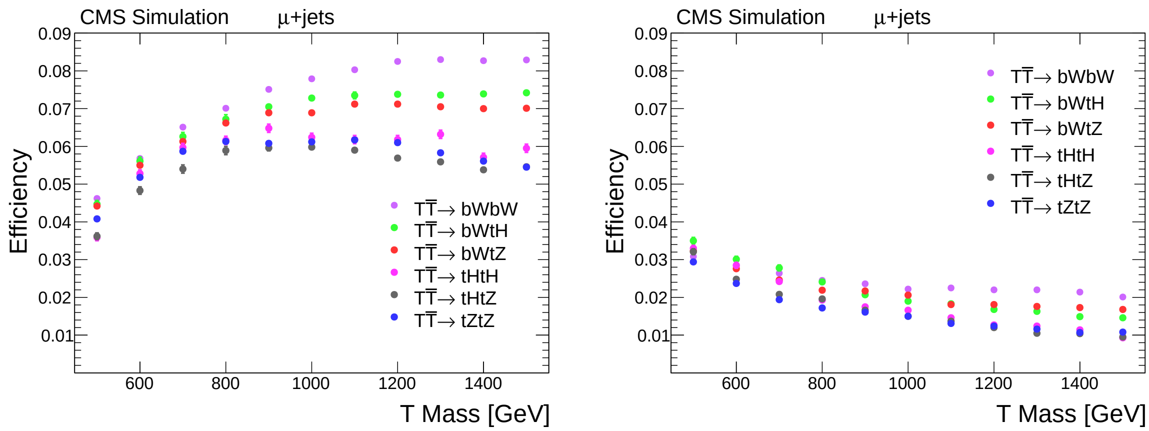
The event yields for events surviving the total selection in data and background MC predictions are shown on Table 11.6 for events with and without substructure. Data and expected SM backgrounds, based on the MC samples listed in Table 11.4, are normalised to an integrated luminosity of 19.6 fb^{-1}.

process	events with substructure	events without substructure
t t-channel	74.37 ± 4.59	232.97 ± 8.18
$\bar{t}$ t-channel	35.45 ± 3.26	113.92 ± 5.99
tW	372.18 ± 12.51	451.38 ± 13.96
$\bar{t}W$	378.34 ± 12.00	474.59 ± 13.39
t s-channel	15.12 ± 2.12	32.15 ± 2.95
$\bar{t}$ s channel	4.28 ± 1.01	12.36 ± 1.73
t$\bar{\text{t}}$+jets	11022.09 ± 66.37	22477.67 ± 94.95
W+2Jets	2.68 ± 2.68	27.04 ± 8.15
W+3Jets	939.16 ± 38.25	106.50 ± 13.42
W+4Jets	6567.94 ± 73.22	13361.33 ± 106.47
Z+Jets	832.80 ± 59.95	1434.08 ± 78.12
ttH+Jets	197.27 ± 1.96	368.26 ± 2.68
ttW+Jets	61.63 ± 0.72	82.60 ± 0.84
ttZ+Jets	43.74 ± 0.91	64.65 ± 1.10
WW	206.47 ± 5.26	145.92 ± 4.44
WZ	71.22 ± 1.99	55.73 ± 1.75
ZZ	5.60 ± 0.30	7.03 ± 0.33
QCD μ enriched	0.00 ± 0.00	0.00 ± 0.00
total MC	20830.35 ± 123.31	39448.19 ± 164.99
total data	20017	38664

Table 11.6: Number of events for background MC and data after selection in μ+jets events with and without substructure. The SM backgrounds are normalised to predicted cross sections for 19.6 fb^{-1}. Errors are purely statistical.

The selection efficiency for the $T\bar{T}$ signal is shown in Fig. 11.18 and tabulated in Table 11.7 for events with substructure and Table 11.8 for events without substructure.

The b-jets in signal events originate from the T-quark decays. Additionally, signal events often have gluon jets that arise from initial or final state radiation. Distributions for the number of jets and the number of b-tagged jets in the μ+jets sample are shown in Fig. 11.19.



(a) (b)

Figure 11.18: Signal selection efficiency in the μ+jets channel in events (a) with and (b) without substructure.

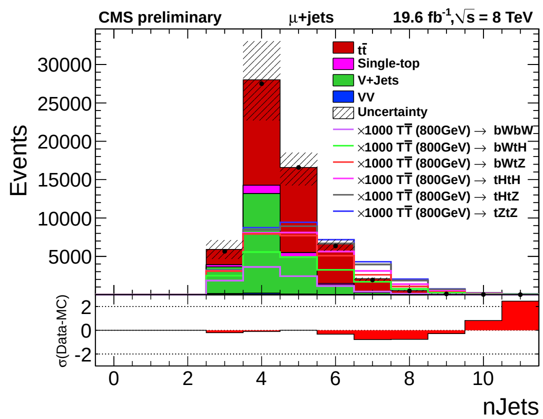
Mass [GeV]	bWbW	bWtH	bWtZ	tHtH	tHtZ	tZtZ
500	$4.62 \pm 0.06\%$	$4.47 \pm 0.10\%$	$4.42 \pm 0.04\%$	$3.59 \pm 0.09\%$	$3.62 \pm 0.09\%$	$4.08 \pm 0.05\%$
600	$5.68 \pm 0.06\%$	$5.61 \pm 0.11\%$	$5.50 \pm 0.06\%$	$5.28 \pm 0.10\%$	$4.83 \pm 0.10\%$	$5.18 \pm 0.06\%$
700	$6.51 \pm 0.07\%$	$6.26 \pm 0.11\%$	$6.13 \pm 0.06\%$	$5.97 \pm 0.11\%$	$5.40 \pm 0.11\%$	$5.87 \pm 0.06\%$
800	$7.01 \pm 0.06\%$	$6.72 \pm 0.12\%$	$6.62 \pm 0.07\%$	$6.16 \pm 0.11\%$	$5.89 \pm 0.11\%$	$6.13 \pm 0.06\%$
900	$7.51 \pm 0.07\%$	$7.05 \pm 0.07\%$	$6.89 \pm 0.05\%$	$6.48 \pm 0.11\%$	$5.96 \pm 0.08\%$	$6.08 \pm 0.06\%$
1000	$7.79 \pm 0.07\%$	$7.28 \pm 0.07\%$	$6.89 \pm 0.07\%$	$6.24 \pm 0.11\%$	$5.98 \pm 0.06\%$	$6.12 \pm 0.06\%$
1100	$8.03 \pm 0.07\%$	$7.35 \pm 0.09\%$	$7.12 \pm 0.07\%$	$6.18 \pm 0.11\%$	$5.90 \pm 0.06\%$	$6.17 \pm 0.06\%$
1200	$8.25 \pm 0.07\%$	$7.38 \pm 0.07\%$	$7.12 \pm 0.07\%$	$6.18 \pm 0.11\%$	$5.69 \pm 0.06\%$	$6.10 \pm 0.06\%$
1300	$8.30 \pm 0.07\%$	$7.36 \pm 0.07\%$	$7.05 \pm 0.07\%$	$6.32 \pm 0.11\%$	$5.59 \pm 0.06\%$	$5.83 \pm 0.06\%$
1400	$8.27 \pm 0.05\%$	$7.39 \pm 0.07\%$	$7.00 \pm 0.05\%$	$5.71 \pm 0.11\%$	$5.38 \pm 0.06\%$	$5.61 \pm 0.06\%$
1500	$8.29 \pm 0.07\%$	$7.42 \pm 0.07\%$	$7.01 \pm 0.05\%$	$5.95 \pm 0.11\%$	$5.46 \pm 0.04\%$	$5.45 \pm 0.06\%$

Table 11.7: Signal selection efficiency with statistical uncertainties in the μ+jets channel for different T masses in events with substructure (≥ 3 ak5-jets and ≥ 1 W-jets).

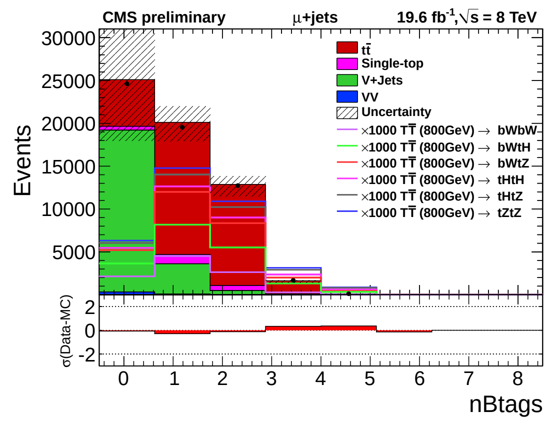
Mass [GeV]	bWbW	bWtH	bWtZ	tHtH	tHtZ	tZtZ
500	$3.08 \pm 0.04\%$	$3.5 \pm 0.09\%$	$3.23 \pm 0.03\%$	$3.3 \pm 0.08\%$	$3.21 \pm 0.08\%$	$2.94 \pm 0.04\%$
600	$2.83 \pm 0.04\%$	$3.01 \pm 0.08\%$	$2.76 \pm 0.04\%$	$2.85 \pm 0.08\%$	$2.48 \pm 0.07\%$	$2.37 \pm 0.04\%$
700	$2.64 \pm 0.04\%$	$2.78 \pm 0.08\%$	$2.46 \pm 0.04\%$	$2.42 \pm 0.07\%$	$2.08 \pm 0.07\%$	$1.94 \pm 0.03\%$
800	$2.45 \pm 0.04\%$	$2.41 \pm 0.07\%$	$2.19 \pm 0.04\%$	$1.93 \pm 0.06\%$	$1.96 \pm 0.06\%$	$1.72 \pm 0.03\%$
900	$2.36 \pm 0.04\%$	$2.07 \pm 0.04\%$	$2.17 \pm 0.03\%$	$1.75 \pm 0.06\%$	$1.66 \pm 0.04\%$	$1.61 \pm 0.03\%$
1000	$2.22 \pm 0.04\%$	$1.90 \pm 0.04\%$	$2.06 \pm 0.04\%$	$1.66 \pm 0.06\%$	$1.50 \pm 0.03\%$	$1.50 \pm 0.03\%$
1100	$2.25 \pm 0.04\%$	$1.83 \pm 0.04\%$	$1.81 \pm 0.03\%$	$1.46 \pm 0.06\%$	$1.37 \pm 0.03\%$	$1.31 \pm 0.03\%$
1200	$2.20 \pm 0.04\%$	$1.68 \pm 0.03\%$	$1.81 \pm 0.03\%$	$1.27 \pm 0.05\%$	$1.20 \pm 0.03\%$	$1.24 \pm 0.03\%$
1300	$2.20 \pm 0.04\%$	$1.63 \pm 0.03\%$	$1.76 \pm 0.03\%$	$1.24 \pm 0.05\%$	$1.05 \pm 0.03\%$	$1.16 \pm 0.03\%$
1400	$2.14 \pm 0.03\%$	$1.49 \pm 0.03\%$	$1.73 \pm 0.02\%$	$1.14 \pm 0.05\%$	$1.04 \pm 0.03\%$	$1.07 \pm 0.03\%$
1500	$2.01 \pm 0.03\%$	$1.46 \pm 0.03\%$	$1.68 \pm 0.02\%$	$0.93 \pm 0.04\%$	$0.95 \pm 0.02\%$	$1.08 \pm 0.02\%$

Table 11.8: Signal selection efficiency with statistical uncertainties in the μ+jets channel for different T masses in events without substructure (≥ 4 ak5-jets and $= 0$ W-jets).

The most important kinematic distributions for this analysis are the jet p_T of the four leading ak5-jets, shown in Fig. 11.20, the lepton p_T and the neutrino momenta in the form of the missing transverse energy $\not{E}_T$, both shown in Fig. 11.21. Finally, the H_T the scalar sum of the transverse momentum of all jets, $\not{E}_T$ and lepton p_T, is shown in Fig. 11.22. These variables are used in the statistical procedure described in Section 11.7.



(a)



(b)

Figure 11.19: (a) Number of jets and (b) number of CSVM b-tagged jets in the μ+jets channel.

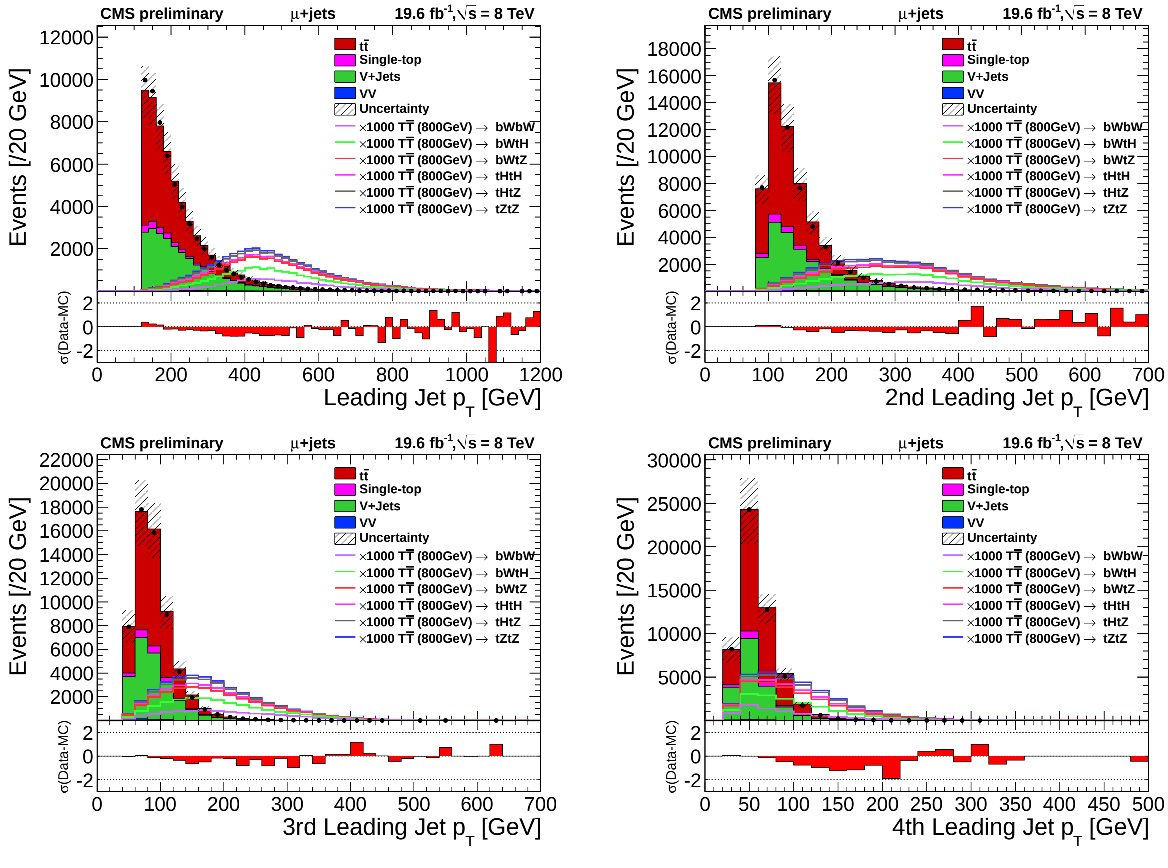
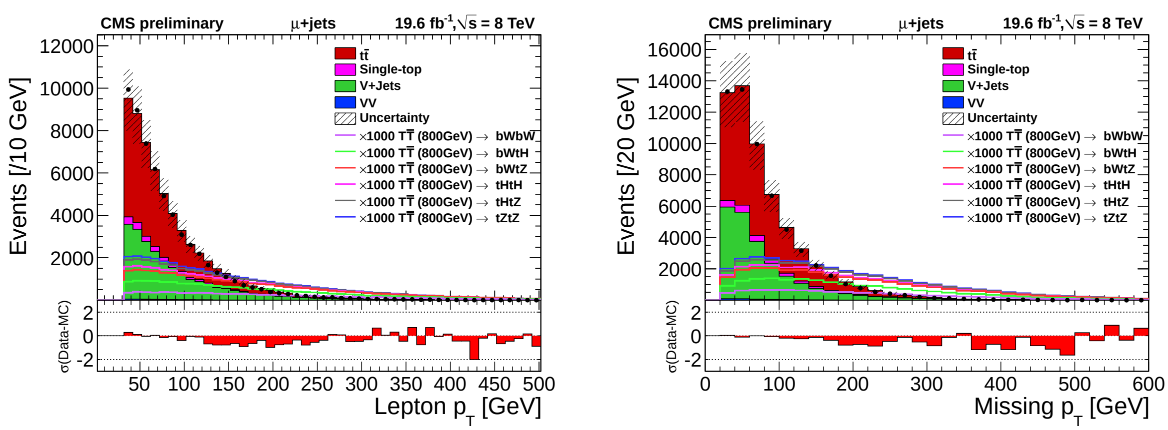


Figure 11.20: Transverse momenta for the four jets with highest p_T in the μ+jets channel.



(a) (b)

Figure 11.21: (a) Muon p_T and (b) the missing transverse energy $\not{E}_T$ in the μ+jets channel.

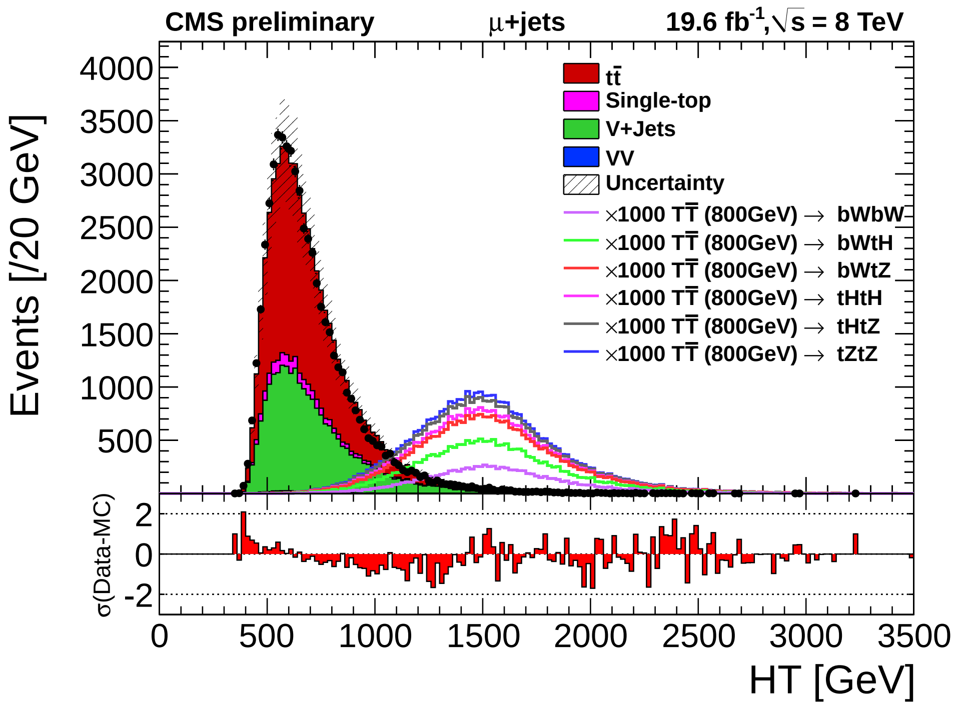


Figure 11.22: H_T scalar sum of p_T of all jets, the lepton and the missing transverse energy $\not{E}_T$ in the μ+jets channel.

11.5 Electron Channel

11.5.1 Trigger and Electron ID

The trigger for the electron channel (e+jets) in this analysis, is the single electron trigger HLT_Ele27, which requires at least one isolated electron with $p_T > 27$ GeV. This trigger was unprescaled for the whole 2012 data taking period.

We derive a flat scale factor to account for the difference of efficiencies of triggering on electrons between MC and data. This is derived by using a tag-and-probe method which proceeds as follows. In the single electron data and the Drell-Yann (Z+jets) MC sample, we look for events with exactly two *tight* electrons. We match one of them, the "tag", to a HLT_Ele27 trigger object. The second electron is used as the "probe". We check if the probe is matched to a second HLT_Ele27 trigger object. The invariant mass of the electron pairs for all events, and also for the subsets of all events in which the probe is matched to a second trigger object ("passing") and in which it is not ("failing"), are shown

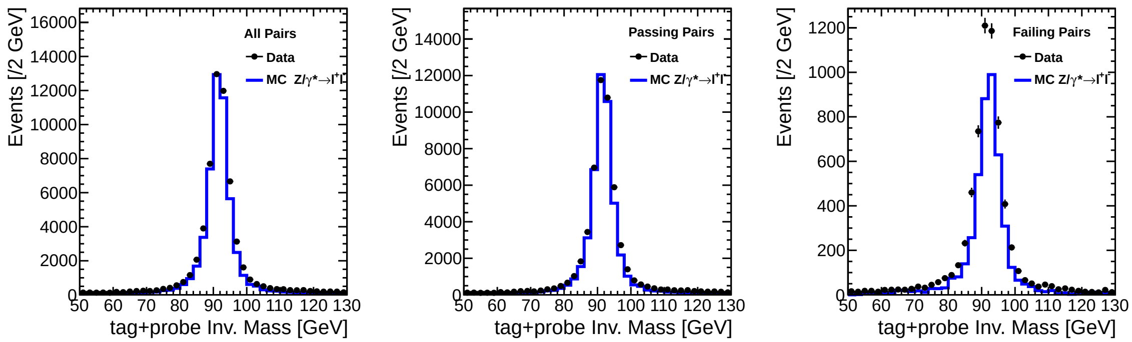
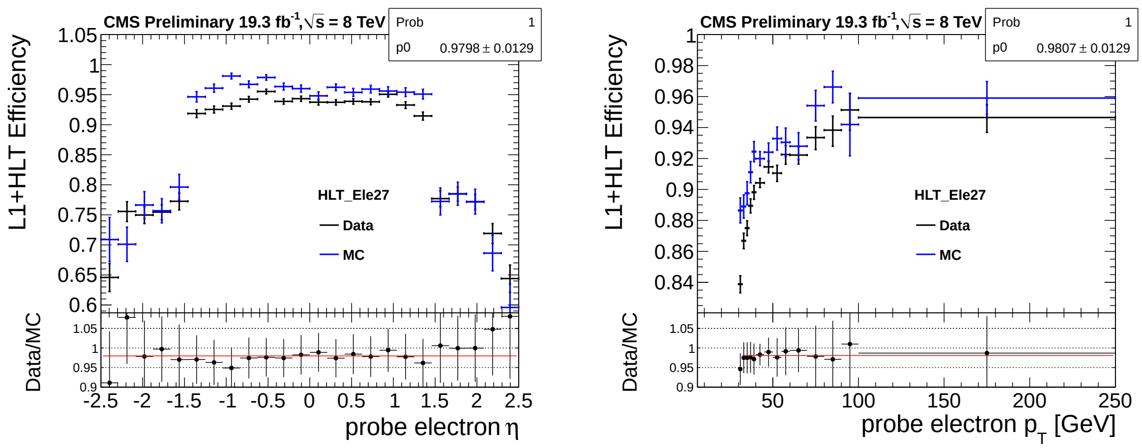


Figure 11.23: The invariant mass of the two electrons in the tag-and-probe for (left) all lepton pairs, (centre) passing lepton pairs and (right) failing lepton pairs.



(a) (b)

Figure 11.24: Trigger efficiency for the single electron trigger Ele_27 as a function of (a) η and (b) p_T of the electron probe.

in Fig. 11.23. We then count the number of events in the invariant mass window of the dileptons $80 < M_{ll} < 100$ for all probes and all passing probes as a function of p_T and η. The ratio of the passing probes to all events with two electrons is shown for both data and Z+jets MC simulation in Fig. 11.24. Finally, we take the ratio of data and Monte Carlo simulation efficiencies. For simplicity, we select events with one electron with $p_T > 32$ GeV and use a flat scale factor of 0.980 ± 0.013 across all η and p_T ranges.

The electron ID for this analysis uses the Cut Based Electron ID working point tight [193]. For the ID and isolation efficiency we use the centrally produced scale factors, listed in Table 11.9, to account for differences between data and MC [126].

p_T	$0.0 < \|\eta\| < 0.8$	$0.8 < \|\eta\| < 1.442$	$1.442 < \|\eta\| < 1.556$	$1.556 < \|\eta\| < 2.0$	$2.0 < \|\eta\| < 2.5$
$10 - 15$	0.818 ± 0.042	0.840 ± 0.024	1.008 ± 0.203	0.906 ± 0.031	0.991 ± 0.084
$15 - 20$	0.928 ± 0.044	0.914 ± 0.008	0.877 ± 0.049	0.907 ± 0.019	0.939 ± 0.046
$20 - 30$	0.973 ± 0.005	0.948 ± 0.005	0.983 ± 0.003	0.957 ± 0.008	1.001 ± 0.013
$30 - 40$	0.979 ± 0.002	0.961 ± 0.005	0.983 ± 0.002	0.962 ± 0.002	1.002 ± 0.003
$40 - 50$	0.984 ± 0.000	0.972 ± 0.001	0.957 ± 0.000	0.985 ± 0.001	0.999 ± 0.006
$50 - 200$	0.983 ± 0.000	0.977 ± 0.004	0.978 ± 0.004	0.986 ± 0.001	0.995 ± 0.009

Table 11.9: Electron SF for electron identification and isolation for the tight working point used to correct the differences between data and MC.

In total, we assign a conservative 3% uncertainty to account for triggering, identification and isolation efficiencies. This identification uncertainty is used in the limit setting tool, `theta`, described in Section 11.8.1 [204].

11.5.2 Electron Selection

Events in the electron channel are required to satisfy the same requirements as the muon selection except for the differences described in the following paragraphs:

- Instead of a tight muon, we require exactly one tight electron in the event using the Cut Based ID tight working point from EGamma recommendations [193].
- Electron $p_T > 32$ GeV is required.
- Electron $|\eta| < 2.5$. We exclude the barrel-endcap transition region in the ECAL, $1.44 < |\eta_{SC}| < 1.566$, since it is not instrumented and thus the quality of the reconstructed superclusters (SC) used for identifying electrons is degraded.
- Require `PF` ρ-corrected isolation for the electron with a cone size of $\Delta R = 0.3$ (11.2) [193].
- Conversion rejection: electrons with no hits in the innermost pixel layer are considered to arise from photon conversions and are rejected. Furthermore, electrons are rejected when a partner track is found that is consistent with a photon conversion, based on the vertex fit probability being less that 10^{-6}.

- In addition, specific to the e+jets channel, jets have to pass `jetId` cuts to reject most fake jets from calorimeter and/or readout electronics noise, provided by the JetMET group [205]:

  `CHF` > 0.00,

  `NHF` < 0.99,

  `CEF` < 0.99,

  `NEF` < 0.99,

  `NCH` > 0,

  `Nconstituents` > 1

  where `CHF, NHF, CEF, NEF` are the fractions of the uncorrected `PFJet` energy carried by charged hadronic energy, neutral hadronic energy, charged electromagnetic energy, and neutral electromagnetic energy, respectively. `NCH` and `Nconstituents` is the number of charged hadrons in the `PFJet` and the number of `PF` constituents in the jet, respectively.

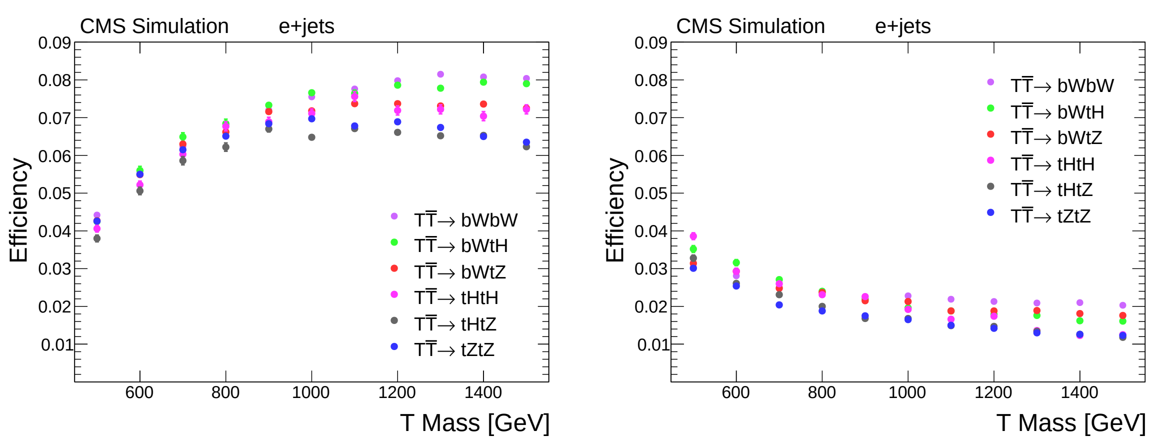
The event yields for the total selection in data and background MC predictions are shown in Table 11.10 for events with and without substructure. Data and expected SM backgrounds, based on the MC samples listed in Table 11.4, are normalised to an integrated luminosity of 19.6 fb^{-1}.

The selection efficiency for the $T\bar{T}$ signal is shown in Fig. 11.25 and tabulated in Table 11.11 and Table 11.12 for events with and without substructure respectively.

Distributions for the number of jets and the number b-tagged jets in e+jets events are shown in Fig. 11.26. The most important kinematic distributions for this analysis are: the jet p_T of the four leading jets, shown in Fig. 11.27; the lepton p_T and the missing transverse energy, both shown in Fig. 11.28; and H$_T$, the scalar sum of the transverse momentum of all jets, $\not{E}_T$ and lepton p_T shown in Fig. 11.29. These variables are used in the statistical procedure described in Section 11.7.

process	events with substructure			events without substructure		
t t-channel	69.18	$\pm$	4.38	234.17	$\pm$	8.17
$\bar{t}$ t-channel	28.85	$\pm$	2.86	119.20	$\pm$	5.96
tW	382.20	$\pm$	12.69	423.70	$\pm$	13.35
$\bar{t}W$	350.19	$\pm$	11.25	436.81	$\pm$	12.81
t s-channel	11.98	$\pm$	1.79	27.21	$\pm$	2.73
$\bar{t}$ s channel	6.16	$\pm$	1.21	8.72	$\pm$	1.40
$t\bar{t}$+jets	11058.51	$\pm$	65.83	22484.31	$\pm$	94.16
W+2Jets	3.27	$\pm$	3.27	28.74	$\pm$	8.30
W+3Jets	1052.11	$\pm$	40.83	64.27	$\pm$	9.37
W+4Jets	6400.03	$\pm$	71.64	12466.56	$\pm$	101.49
Z+Jets	667.74	$\pm$	53.12	1388.80	$\pm$	76.68
ttH+Jets	207.34	$\pm$	1.99	379.72	$\pm$	2.70
ttW+Jets	59.20	$\pm$	0.70	80.80	$\pm$	0.82
ttZ+Jets	44.82	$\pm$	0.92	65.33	$\pm$	1.10
WW	208.18	$\pm$	5.24	137.44	$\pm$	4.25
WZ	67.78	$\pm$	1.90	58.76	$\pm$	1.80
ZZ	6.72	$\pm$	0.33	7.02	$\pm$	0.34
QCD EM 20 $< p_T <$ 80 GeV	0.00	$\pm$	0.00	0.00	$\pm$	0.00
QCD EM 80 $< p_T <$ 170 GeV	103.61	$\pm$	103.61	727.64	$\pm$	275.02
QCD EM 170 $< p_T <$ 250 GeV	229.39	$\pm$	24.59	493.14	$\pm$	37.07
QCD EM 250 $< p_T <$ 350 GeV	161.52	$\pm$	7.02	168.77	$\pm$	7.28
QCD EM 350 GeV $< p_T$	57.97	$\pm$	1.70	59.97	$\pm$	1.73
total MC	21176.74	$\pm$	160.35	39861.06	$\pm$	320.55
total Data	20836			40364		

Table 11.10: Number of events for background MC and data after selection in e+jets events with and without substructure. The SM backgrounds are normalised to predicted cross sections for 19.6 fb^{-1}. Errors are purely statistical.



(a)

(b)

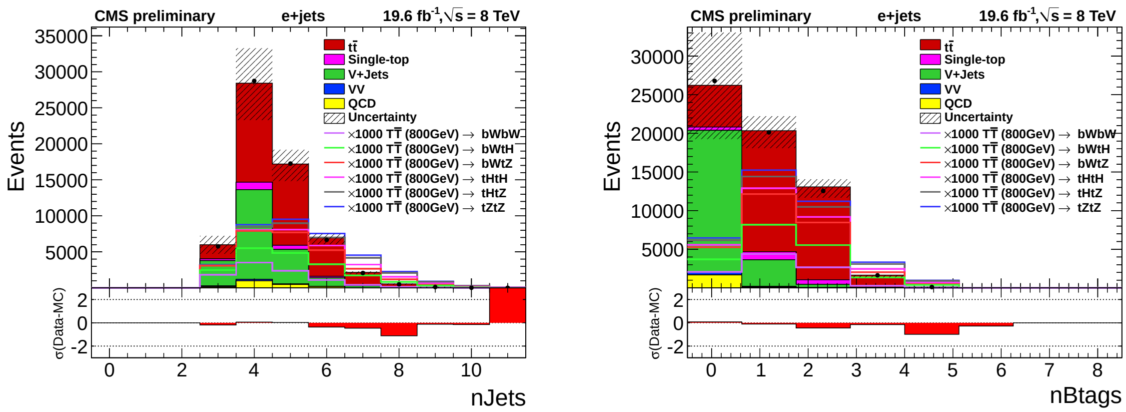
Figure 11.25: Signal efficiency of in the e+jets channel in events (a) with and (b) without substructure.

Mass [GeV]	bWbW	bWtH	bWtZ	tHtH	tHtZ	tZtZ
500	$4.42 \pm 0.05\%$	$4.25 \pm 0.09\%$	$4.27 \pm 0.04\%$	$4.06 \pm 0.09\%$	$3.8 \pm 0.09\%$	$4.26 \pm 0.05\%$
600	$5.51 \pm 0.06\%$	$5.60 \pm 0.11\%$	$5.49 \pm 0.06\%$	$5.22 \pm 0.1\%$	$5.06 \pm 0.1\%$	$5.50 \pm 0.06\%$
700	$6.24 \pm 0.06\%$	$6.49 \pm 0.11\%$	$6.30 \pm 0.06\%$	$6.03 \pm 0.11\%$	$5.86 \pm 0.11\%$	$6.15 \pm 0.06\%$
800	$6.82 \pm 0.06\%$	$6.84 \pm 0.12\%$	$6.62 \pm 0.07\%$	$6.77 \pm 0.12\%$	$6.22 \pm 0.11\%$	$6.51 \pm 0.07\%$
900	$7.20 \pm 0.07\%$	$7.33 \pm 0.07\%$	$7.16 \pm 0.05\%$	$6.89 \pm 0.12\%$	$6.70 \pm 0.08\%$	$6.84 \pm 0.07\%$
1000	$7.55 \pm 0.07\%$	$7.66 \pm 0.07\%$	$7.18 \pm 0.07\%$	$7.11 \pm 0.12\%$	$6.48 \pm 0.07\%$	$6.97 \pm 0.07\%$
1100	$7.76 \pm 0.07\%$	$7.64 \pm 0.09\%$	$7.37 \pm 0.07\%$	$7.56 \pm 0.13\%$	$6.71 \pm 0.07\%$	$6.78 \pm 0.07\%$
1200	$7.98 \pm 0.07\%$	$7.86 \pm 0.07\%$	$7.37 \pm 0.07\%$	$7.19 \pm 0.12\%$	$6.61 \pm 0.07\%$	$6.89 \pm 0.06\%$
1300	$8.15 \pm 0.07\%$	$7.78 \pm 0.07\%$	$7.31 \pm 0.07\%$	$7.22 \pm 0.12\%$	$6.52 \pm 0.07\%$	$6.74 \pm 0.07\%$
1400	$8.08 \pm 0.05\%$	$7.94 \pm 0.07\%$	$7.36 \pm 0.05\%$	$7.04 \pm 0.12\%$	$6.53 \pm 0.07\%$	$6.50 \pm 0.06\%$
1500	$8.04 \pm 0.07\%$	$7.90 \pm 0.07\%$	$7.25 \pm 0.05\%$	$7.22 \pm 0.12\%$	$6.23 \pm 0.04\%$	$6.35 \pm 0.06\%$

Table 11.11: Signal selection efficiency with statistical uncertainties in the e+jets channel for different T masses in events with substructure (≥ 3 ak5-jets and ≥ 1 W-jets).

Mass [GeV]	bWbW	bWtH	bWtZ	tHtH	tHtZ	tZtZ
500	$3.04 \pm 0.04\%$	$3.52 \pm 0.09\%$	$3.14 \pm 0.03\%$	$3.86 \pm 0.09\%$	$3.28 \pm 0.08\%$	$3.01 \pm 0.04\%$
600	$2.81 \pm 0.04\%$	$3.16 \pm 0.08\%$	$2.93 \pm 0.04\%$	$2.93 \pm 0.08\%$	$2.61 \pm 0.07\%$	$2.54 \pm 0.04\%$
700	$2.50 \pm 0.04\%$	$2.71 \pm 0.07\%$	$2.48 \pm 0.04\%$	$2.60 \pm 0.07\%$	$2.31 \pm 0.07\%$	$2.04 \pm 0.03\%$
800	$2.38 \pm 0.04\%$	$2.40 \pm 0.07\%$	$2.37 \pm 0.04\%$	$2.31 \pm 0.07\%$	$2.00 \pm 0.06\%$	$1.88 \pm 0.04\%$
900	$2.24 \pm 0.04\%$	$2.19 \pm 0.04\%$	$2.15 \pm 0.03\%$	$2.26 \pm 0.07\%$	$1.68 \pm 0.04\%$	$1.75 \pm 0.03\%$
1000	$2.28 \pm 0.04\%$	$1.96 \pm 0.04\%$	$2.13 \pm 0.04\%$	$1.92 \pm 0.06\%$	$1.68 \pm 0.03\%$	$1.65 \pm 0.03\%$
1100	$2.19 \pm 0.04\%$	$1.88 \pm 0.04\%$	$1.88 \pm 0.04\%$	$1.66 \pm 0.06\%$	$1.49 \pm 0.03\%$	$1.50 \pm 0.03\%$
1200	$2.13 \pm 0.04\%$	$1.79 \pm 0.04\%$	$1.88 \pm 0.04\%$	$1.74 \pm 0.06\%$	$1.47 \pm 0.03\%$	$1.42 \pm 0.03\%$
1300	$2.09 \pm 0.04\%$	$1.76 \pm 0.03\%$	$1.89 \pm 0.03\%$	$1.36 \pm 0.05\%$	$1.34 \pm 0.03\%$	$1.30 \pm 0.03\%$
1400	$2.10 \pm 0.03\%$	$1.62 \pm 0.03\%$	$1.81 \pm 0.02\%$	$1.23 \pm 0.05\%$	$1.26 \pm 0.03\%$	$1.26 \pm 0.03\%$
1500	$2.03 \pm 0.03\%$	$1.61 \pm 0.03\%$	$1.76 \pm 0.02\%$	$1.25 \pm 0.05\%$	$1.18 \pm 0.02\%$	$1.23 \pm 0.03\%$

Table 11.12: Signal selection efficiency with statistical uncertainties in the e+jets channel for different T masses in events without substructure (≥ 4 ak5-jets and $= 0$ W-jets).



(a) (b)

Figure 11.26: (a) Number of jets and (b) Number of CSVM b-tagged jets in the e+jets channel.

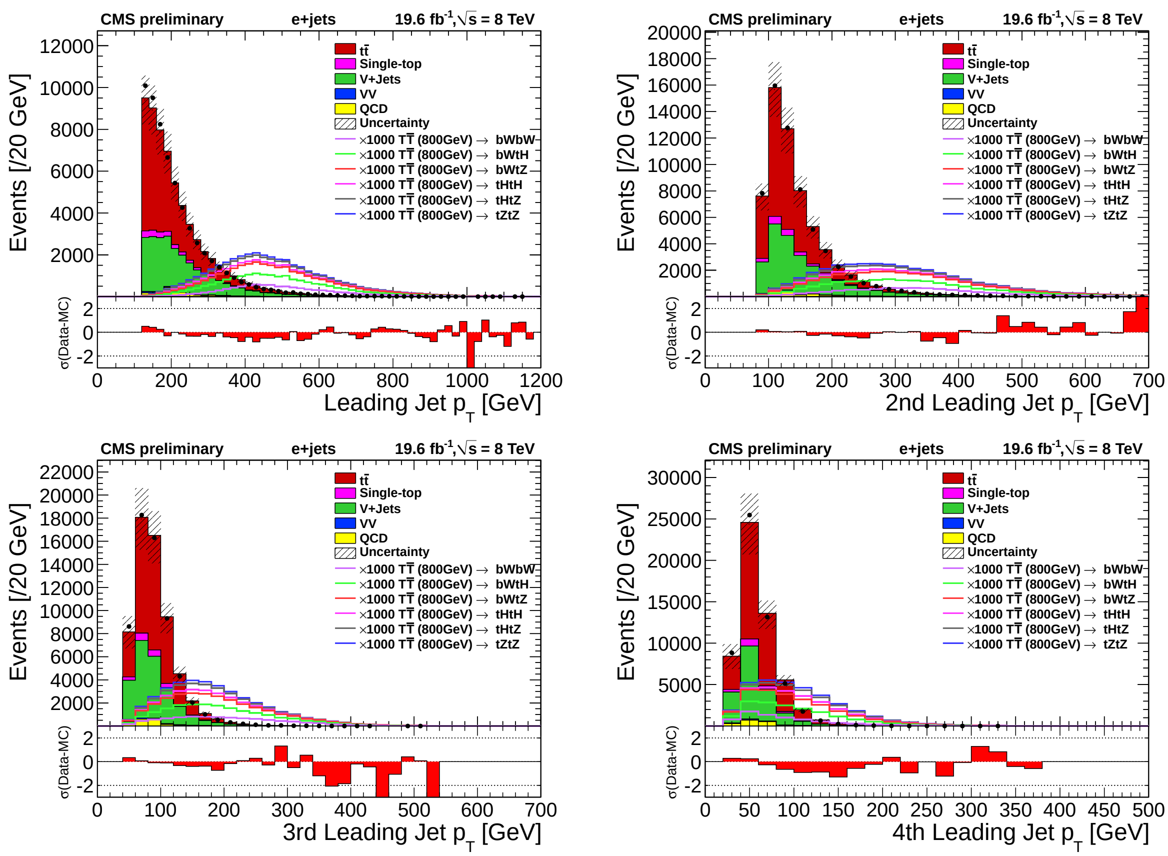
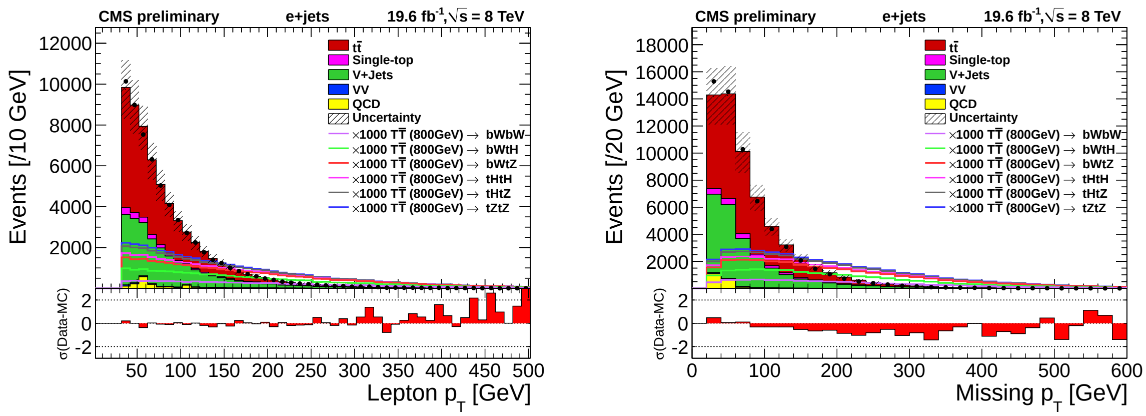


Figure 11.27: Transverse momentum for the four jets with highest p_T in the e+jets channel.



(a) (b)

Figure 11.28: (a) Electron p_T and (b) missing transverse energy $\not{E}_T$ in the e+jets channel.

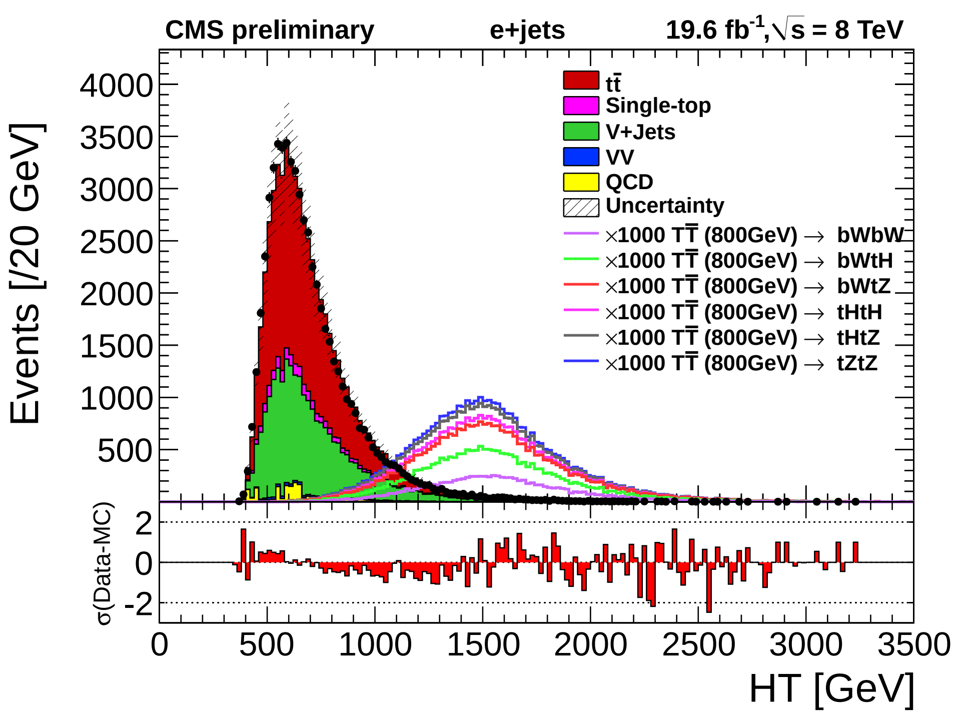


Figure 11.29: H$_T$ scalar sum of p_T of all jets, the lepton and the missing transverse energy $\not{E}_T$ in the e+jets channel.

11.6 Systematic Uncertainties

In general, the presence of a systematic uncertainty affects both the number of selected events and the distribution of the investigated discriminating observables, resulting in modified distributions for the different processes.

Uncertainties in the lepton scale factors, background yields, and luminosity are applied as Gaussian constraints in the statistical procedure. To incorporate jet energy correction, resolution, b-tagging, scale and match uncertainties in the computation of the T cross section limits, we create systematic variations in the templates corresponding to the up and down variations described in this section.

The resulting templates are integrated into the final limit calculation by applying a Gaussian constraint within $\pm 1\sigma$ using vertical morphing techniques.

11.6.1 Jet Energy Scale

The jet energy systematic (JES) uncertainty is the main systematic uncertainty. The signal, $T\overline{T}$, and the largest background, $\mathrm{t\bar{t}}$, have very steeply falling spectra for the p_T of the third and fourth jets in our events. Any uncertainty in the energy scale would have a considerable impact on the behaviour of these steeply falling distributions and hence influence our measurements significantly.

We take the jet energy correction uncertainty directly from the Global Tag provided by the JetMET group [200]. The missing transverse energy is modified simultaneously since the effect of changing the four momentum of our jets is 100% correlated with the measured $\not\!\!E_T$. We extract the uncertainty in the nominal scale creating new up and down uncertainty templates.

11.6.2 Jet Energy Resolution

As studies done by the JetMET group have shown, the jet energy resolution (JER) in the simulation is different than in data. The procedure developed to correct for these differences carries an uncertainty. The jets in this analysis are corrected in the simulation to match the resolution in data.

To account for JER correction systematic, we follow prescriptions on systematic uncertainties from [206]. We prepare -1σ, nominal and $+1\sigma$ templates by scaling the difference of the reconstructed jet p_T and the matched generator-level p_T by an η dependent factor. The $\not{E}_T$ value is simultaneously modified according the changes of the jets' momenta.

11.6.3 b-tagging

The uncertainty in the b-tagging scale factors, which account for the difference between data and simulated MC, has a smaller impact on the shape of the kinematic distributions. However, it is important for the b-tag multiplicity in all events and thus will influence the final discrimination of signal against background. This systematic uncertainty is taken into account by varying the b-tagging scale factor by one sigma up and down. The b-tagging working group provides these uncertainties that are parameterised for jet p_T and η [195].

11.6.4 Scale (Q^2)

The uncertainty in the factorisation/normalisation "scale" Q, used for the strong coupling constant $\alpha_s(Q^2)$, is estimated by using two sets of $t\bar{t}$ MC samples in which the scale Q is increased and decreased by factors of two relative to its nominal value [152].

11.6.5 Matching thresholds

PYTHIA is used to hadronise all partons in the $t\bar{t}$ MC sample; it then uses a "matching" criteria to filter out events where any jet is generated too far away from their original partons.

Two $t\bar{t}$ MC samples, generated with varied matching up and down thresholds (corresponding to the situations where the default criteria is adjusted by a factor of 0.5 and 2 respectively), are used to estimate the uncertainties introduced by the matching criteria [207].

11.6.6 Lepton Scale Factors

Differences in the lepton identification efficiencies between data and MC are corrected by scale factors. These correction factors have been applied throughout this analysis to match the trigger and lepton selection efficiencies in simulated samples to those in data; these are obtained from data-driven techniques using decays of Z bosons to dileptons.

In the μ+jets channel, the trigger efficiency scale factor is measured as an average over the 2012 A, B, C and D runs for the single muon HLT_IsoMu24 trigger. The trigger, muon ID and muon `PF` isolation data-to-MC scale factors are listed in Table 11.5 as provided by the muon POG [192].

In the e+jets channel, the trigger scale factor applied to the MC is 0.980 ± 0.01. The combined electron ID and isolation scale factors are shown in Table 11.9 as provided by the EGamma POG [126].

We assign a 3% flat uncertainty in both channels to account for the three scale factors that account for ID, isolation and trigger efficiencies [2].

11.6.7 W-jet Uncertainties

A flat 3% uncertainty is applied to account for the uncertainty in tagging W-jets [201].

11.6.8 Background Yields

The uncertainty in the background yields is accounted for by flat normalisation uncertainties. Due to the limited statistics in the simulated samples, all non-$\mathrm{t\bar{t}}$ backgrounds are added together. A conservative 30% uncertainty is applied to account for the cross section uncertainty from the W+Jets scale factors. This is calculated by combining (in quadrature) the 20% variation in the scale factors between channels and an additional 20% efficiency difference observed when using the Q^2 scale variations of the W+Jets MC samples. Further, the uncertainty in cross sections of all other non-$\mathrm{t\bar{t}}$ SM backgrounds is also roughly 30% [2].

The systematic uncertainty for the $\mathrm{t\bar{t}}$ cross section is taken to be 10%, following the prescription in Ref. [208]. This is applied as a flat systematic in the statistical procedure and covers the difference between theoretical calculations in Refs [209] and [210].

11.6.9 Integrated Luminosity

The uncertainty of the integrated luminosity is 4.4% for the 2012 data [211]. We therefore apply a 4.4% Gaussian constraint to all backgrounds, and $T\overline{T}$ normalisations.

11.7 Statistics

All events from data and from the MC simulated samples are required to pass the same trigger and event reconstruction algorithms. Scale factors that account for the differences in the performance of these algorithms between data and simulations are computed and applied in this analysis.

A multivariate (TMVA) classification algorithm [212] is trained on the remaining signal and background events that pass selection (described in Sections 11.4 and 11.5). This technique uses a Boosted Decision Tree (`BDT`) [213, 214] and is implemented in the TMVA framework [212]. TMVA enhances the statistical power of the analysis by taking into account the correlations between discriminating variables in signal and background events.

We "train" our multivariate algorithm on (approximately) half of all signal and background events that pass selection. This constructs a binary tree, known as a decision tree with each node containing information about a single input variable. The remaining signal, background, and all data events are now "tested". The algorithm traverses the tree for each event by making a (left / right) decision at each node until a stop condition is met. Once reached, the event is classified as either signal-like or background-like based on where the majority of training events ended up. In this way, each tested event is assigned a value of the (continuous) parameter known as the `BDT` discriminant (or simply `BDT`). By convention, this value is $+1$ for extremely signal-like events and -1 for extremely background-like events.

"Boosting" extends the concept to several trees, where each tree is derived from the same training sample by reweighting events. The results of testings from each tree are then combined into a single classifier by taking an average of the individual decision trees. This stabilises the responses of the decision tree with respect to statistical fluctuations in the training sample.

Due to the limited amount of statistics that pass selection, only the $t\bar{t}$, W+nJets, Z+Jets backgrounds are used in training; these are the three dominant backgrounds contributing more than 90% of the total SM prediction. Further, only half of the statistics for the background and signal samples are used in the training, ensuring that we can use the

Variable		Description
2nd Jet p_T	:	transverse momentum of second leading jet
3rd Jet p_T	:	transverse momentum of the third leading jet
Lepton p_T	:	transverse momentum of the lepton
$\not{E}_T$	:	missing transverse energy
H_T	:	scalar sum of the transverse momenta of all jets, the lepton, and $\not{E}_T$
nBtags	:	multiplicity of b-tagged jets
nJets	:	multiplicity of jets in the event
nTopJets*	:	multiplicity of top-jets
nWJets*	:	multiplicity of W-jets
WJet p_T*	:	leading W-jet p_T

Table 11.13: Shows the variables used for training of the BDT. * indicates that the variable is present only in events with substructure.

remaining events to construct a boosted decision tree discriminant without the risk of overtraining and biasing our statistical procedure.

We use the BDT input variables listed in Table 11.13, which include the multiplicity of jets and b-tagged jets, and various kinematic properties of the event. The training is repeated at each mass point, for each lepton channel, and for each of the two substructure categories. Events with substructure that contain at least one W-tagged jet have three additional variables added to the training. To avoid any excess computational complexity, we train only once assuming the nominal relative branching ratios of the signal samples.

Variables are chosen based on their importance calculated by the BDT algorithm and to avoid of avoiding over-correlation ($> 75\%$) between inputs so as to limit statistical noise in the training. The correlation matrices may be found Fig. 11.30 for the μ+jets channel and Fig. 11.31 for the e+jets channel.

The BDT discriminant is then constructed for all data and simulated MC samples. The nominal templates are shown in Fig. 11.32 for both channels for events with and without substructure. These four mutually exclusive sets of events, along with all systematics discussed in Section 11.6, are passed to the statistical limit setting tool discussed in Section 11.9. BDT distributions for the signal at multiple mass points are shown in Appendix C.3.

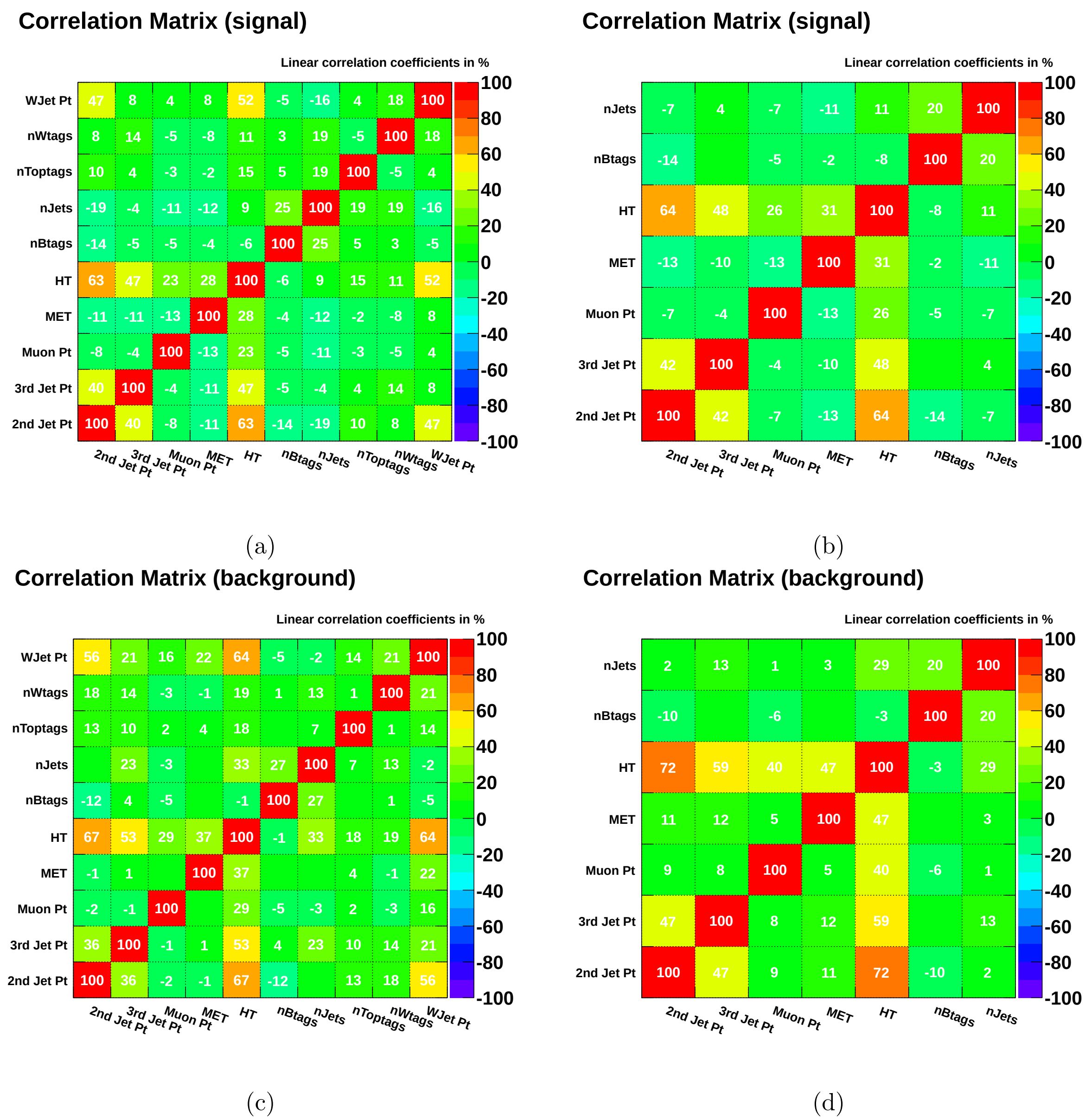


Figure 11.30: The correlation matrix for the BDT input variables in the μ+jets channel for the $m_T = 800$ GeV mass point. Correlation matrices are shown for signal events (a) with substructure and (b) without structure and for the background events (c) with substructure and (d) without substructure.

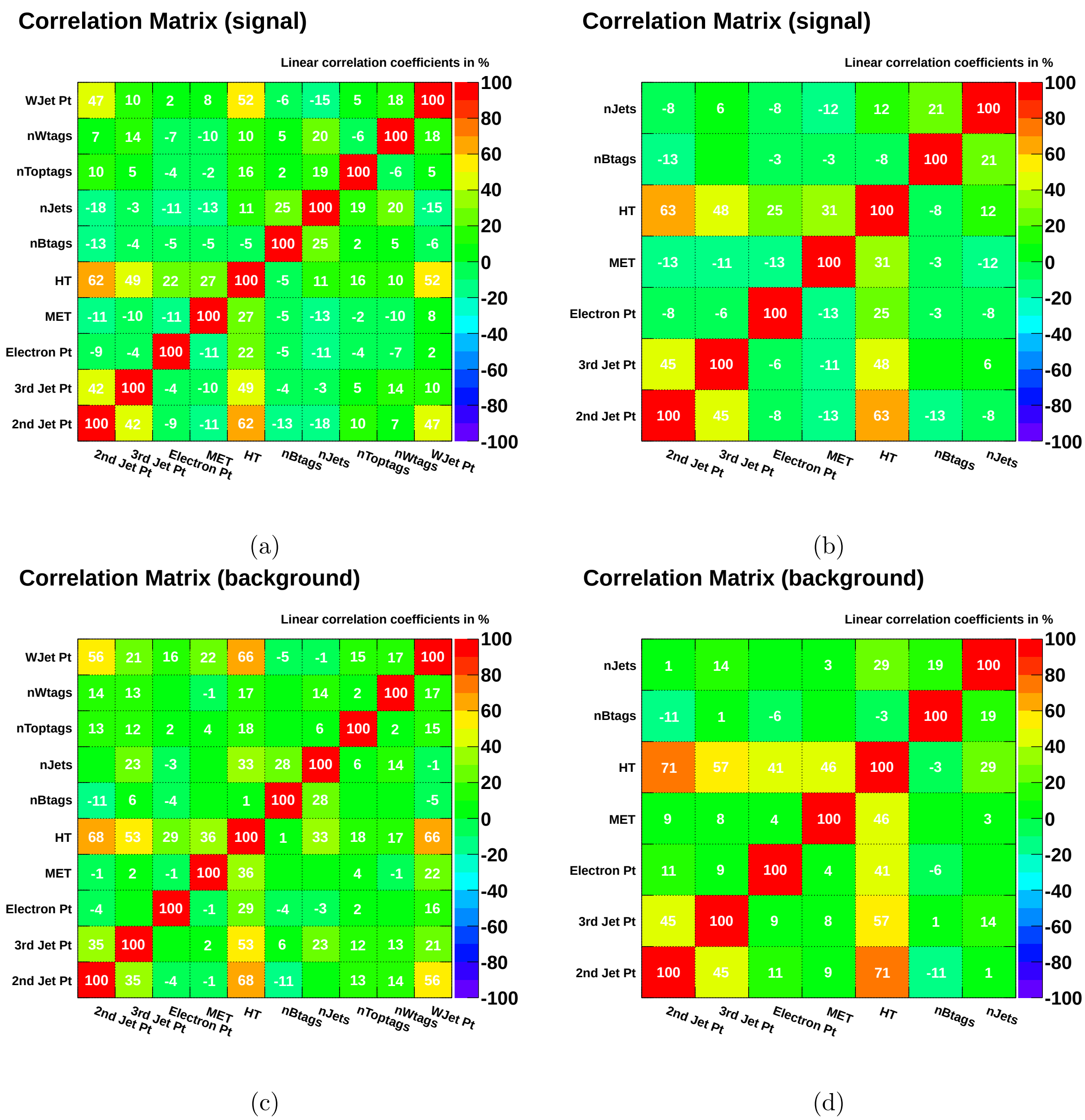


Figure 11.31: The correlation matrix for the BDT input variables in the e+jets channel for the $m_T = 800$ GeV mass point. Correlation matrices are shown for signal events (a) with substructure and (b) without structure and for the background events (c) with substructure and (d) without substructure.

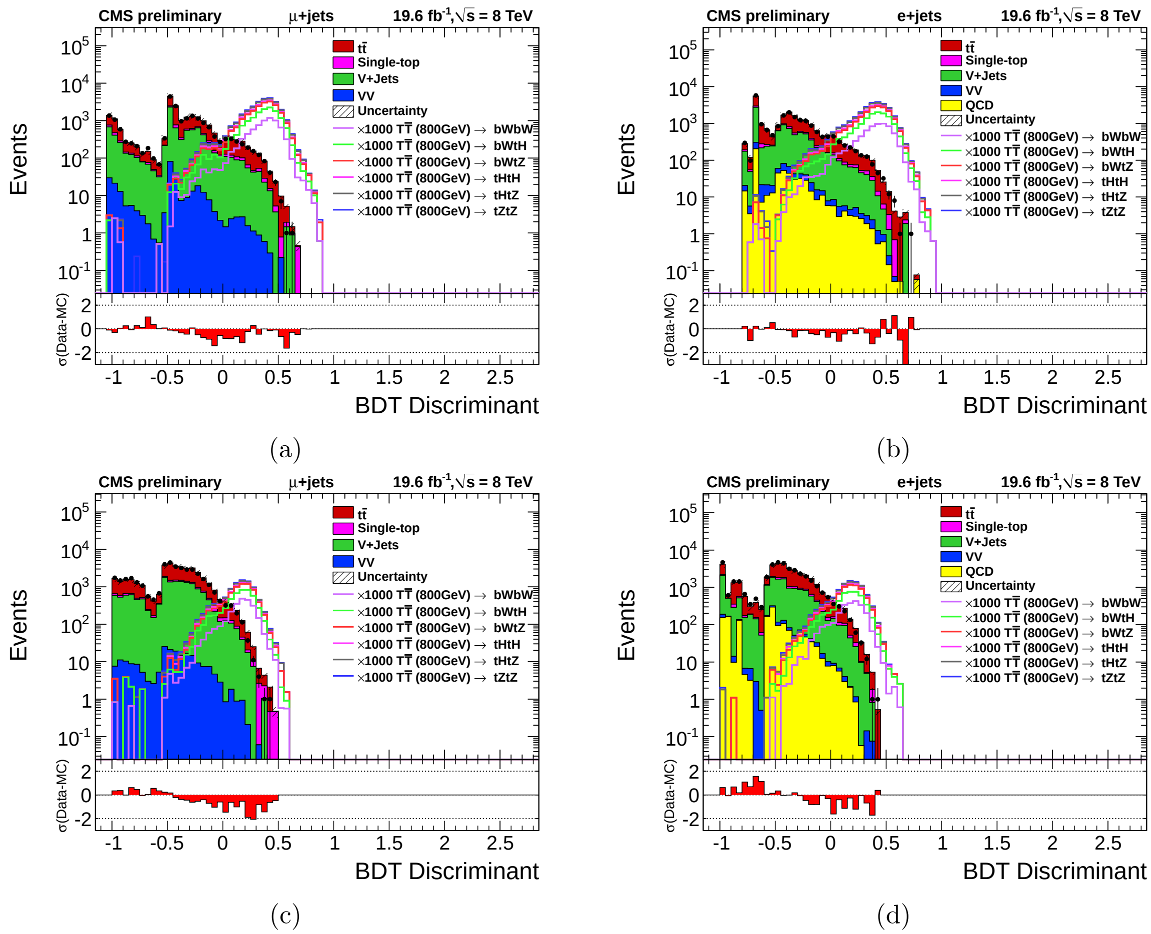


Figure 11.32: The BDT discriminant for events with substructure in the (a) μ+jets, (b) e+jets channels and for events without substructure (c) μ+jets, (d) e+jets channels.

Note that the double hump structure in the BDT distribution is due to the training on multiple signal samples simultaneously. This is evident when training solely on one decay mode as shown in Appendix C.4.

Furthermore, we investigated the apparent deficit of data in the tails of the BDT distribution (most noticeable in the μ+jets channel without substructure). This appears to arise from an inherent mismodelling of the MC samples and is accounted for by the Q^2 scale systematics of the $t\bar{t}$ sample. This study is shown in Appendix C.6.

Since the observed data are consistent with the background predictions within uncertainties, we set exclusion limits on the T quark pair production cross section at each T mass value.

11.8 Limit Calculation Procedure

The expected sensitivity is computed at each mass point using the separation of background and signal. This procedure computes an upper limit on the cross section, known as the "expected limit", using pseudo datasets based on a background only hypothesis and charachterises the expected sensitivity if there is no signal present.

An "observed limit" is computed at each mass point by comparing the data to the background+signal hypothesis. We then calculate an upper limit on the cross section above which we can rule out at 95% confidence level. Thus, a deviation up of the observed limit from the expected limit could potentially indicate a T-quark signal.

For both expected and observed limits, we use the statistical package `theta` [204], which computes the posterior likelihood to calculate upper limits for the signal production cross section. The `BDT` distributions for T-quark masses between 500–1500 GeV are used as the input templates to the limit calculations.

11.8.1 Template Systematics

We investigate the impact on the expected sensitivity of the template systematics used in this analysis. These include the b-tag, JES, JER, matching, and scale up/down templates described in detail in Section 11.6. Using the nominal branching ratios scenario, we determine the effect of each systematic by the following prescription. We calculate the expected limit (for the combined lepton-plus-jets channels) using the flat systematics and only the JES template uncertainties. We then find the difference in the expected limit between this JES only case and JES + X, where X represents one of the other template systematics. This difference is normalised to the JES only case and is shown in Fig. 11.33. An upwards (positive) difference thus indicates that the expected limits are weaker when including the additional template uncertainties.

Since there is more freedom in the fit, it is possible for a stronger expected limit to occur with the addition of more systematics. This results in a downward difference in the limit, as seen in the inclusion of the JER systematics.

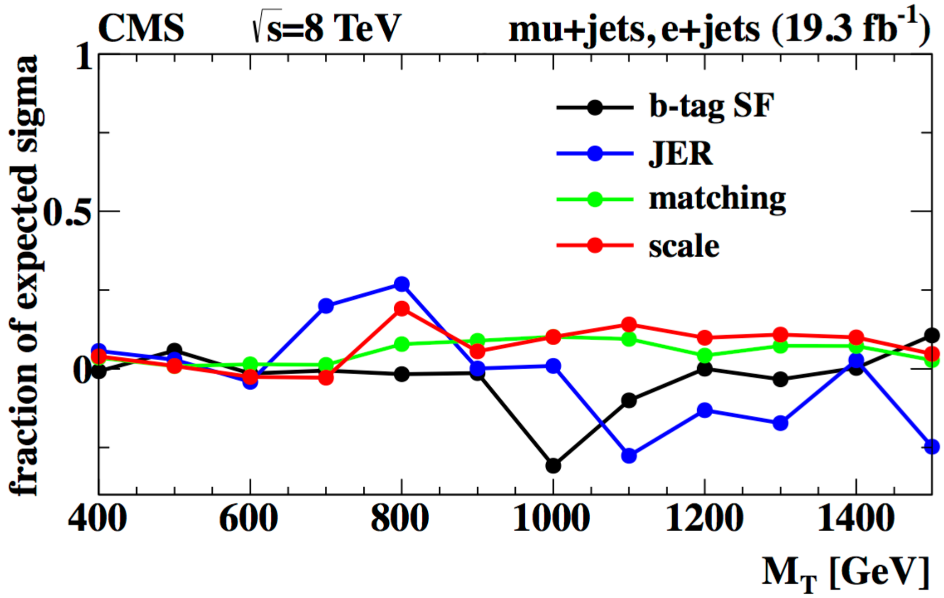


Figure 11.33: Shows the fractional difference between the expected limits using only the JES systematic templates and using the JES+X systematic templates, where X represents any of the other template uncertainties considered (in the combined lepton-plus-jets channel). The difference is normalised to the expected limits of the JES only case at each T-quark mass point.

In the limit setting procedure all possible template uncertainties are taken into account by vertical morphing.

11.9 Results

The largest contribution of the SM background is from $\mathrm{t\bar{t}}$ production (with a 10% uncertainty). Other electroweak processes (W+jets, Z+jets, single-top, diboson production, $\mathrm{t\bar{t}}$V+jets, and QCD) also contribute events to this final state but, due to the high jet p_T cuts, are highly suppressed. The number of events from these electroweak backgrounds that pass our selection cannot be reliably predicted from theory and simulation. The normalisation of the non-$\mathrm{t\bar{t}}$ background is therefore only loosely constrained to their predicted cross section (with a 30% uncertainty). Due to the limited statistics, all prompt non-$\mathrm{t\bar{t}}$ backgrounds contributions to the `BDT` discriminant are combined before being passed to `theta` [204].

Table 11.14 summarises the sources of systematic uncertainty that are accounted for and Table 11.15 shows how the nuisance parameters affect the simulated samples.

source of systematics	label	uncertainty
electron efficiency	ϵ_e	3%
muon efficiency	ϵ_μ	3%
CA8 jet efficiency	ϵ_{CA8}	3%
jet energy scale	jes	$\pm1\sigma$ (shape)
jet energy resolution	jer	$\pm1\sigma$ (shape)
btag	btag	$\pm1\sigma$ (shape)
scale	scale	$\pm1\sigma$ (shape)
match	match	$\pm1\sigma$ (shape)
integrated luminosity	$\int \mathcal{L}\, dt$	4.4%
k factor for number of non-$\mathrm{t\bar{t}}$ background events	k_{ewk}	30%
$t\bar{t}$ cross section	$\sigma(t\bar{t})$	10%

Table 11.14: Summary of systematic uncertainties used for the limit calculation.

The $\mathrm{t\bar{t}}$ and electroweak contributions are modelled by simulated events. In principle, a third component arises from non-prompt lepton backgrounds. In the e+jets channel these come from events in which the fragmentation of a jet fluctuates such that its signature in the calorimeter resembles that of an electron. In the μ+jets channel the primary source of non-prompt lepton background is from events with hadron decays to muons that happen to

source	$\sigma_{t\bar{t}}$	k_{ewk}	α_{jes}	α_{jer}	α_{btag}	α_{scale}	α_{match}	$\epsilon_{e/\mu}$	ϵ_{CA8}	$\int \mathcal{L}dt$
$t\bar{t}$	x		x	x	x	x	x	*	x	x
electroweak		x	x	x	x			*	x	x
$T\overline{T}$			x	x	x			*	x	x

Table 11.15: Effect of nuisance parameters on different signal and background contributions. An "x" means the parameter affects the e+jets and μ+jets channel in a correlated fashion. A "*" means the parameters for the two channels are uncorrelated. No entry means the parameter does not affect the source.

appear isolated from jets. The number of events from non-prompt backgrounds is so small that we can safely neglect them [2].

The $T\bar{T}$ signal is modelled by simulated events and its cross section is the single parameter of interest in the limit calculation. The signal templates for the different decay modes are merged into one histogram with particular branching ratios listed in Table 11.16. The limits are computed for each set of branching ratios.

For the limit calculation, we merge adjacent bins in the `BDT` distribution of all background, signal, and data samples, to generate a final set of 1-dimensional histograms in which all bins are populated by simulated events, each with a maximum of 40% statistical uncertainty in any bin. These 1-dimensional histograms are used to test for the presence of a $T\bar{T}$ signal in the data and upper limits on the $T\overline{T}$ production cross section are obtained from fits to the shape of the `BDT` distributions using `theta`.

11.9.1 Limits

The combination of the four sets of templates, corresponding to the two lepton channels, with and without substructure, is performed by combining the full likelihoods and applying the statistical procedure to the resulting likelihood. We apply all uncertainties and template systematics described in Section 11.6 and compute the limit using `theta`.

The final combined limit for nominal signal branching ratios is shown in Fig. 11.34. By comparing the expected and observed sensitivities to the theoretically predicted cross section at each mass point, we can rule out a mass range. For all T-quark mass values below the intersection of the cross section limit curve with the theoretical line, our sensitivity is greater than the predicted production rate of the T-quark. In this way, the intersection point of the expected and observed lines with the theoretical values corresponds to a lower bound on the mass, below which we can rule out a T-quark with a 95% confidence level. For the nominal branching ratios the pair produced T-quark a lower bound on the observed (expected) limit is set at 698 (770) GeV.

The breakdown of the μ+jets and e+jets channels and the improvements due to the inclusion of jet substructure information can be seen in Fig. 11.35 and Fig. 11.36 respectively.

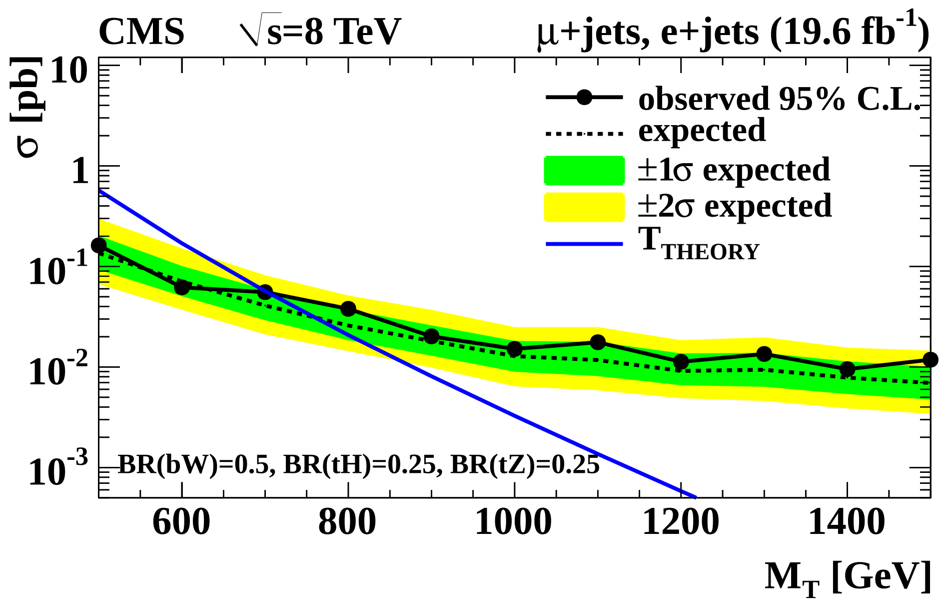
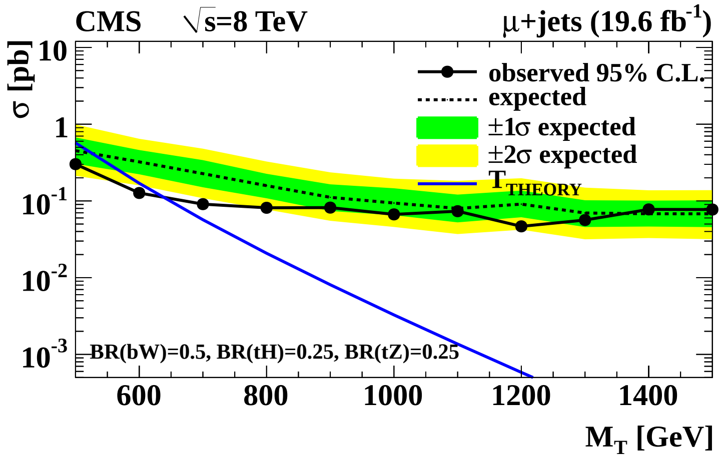
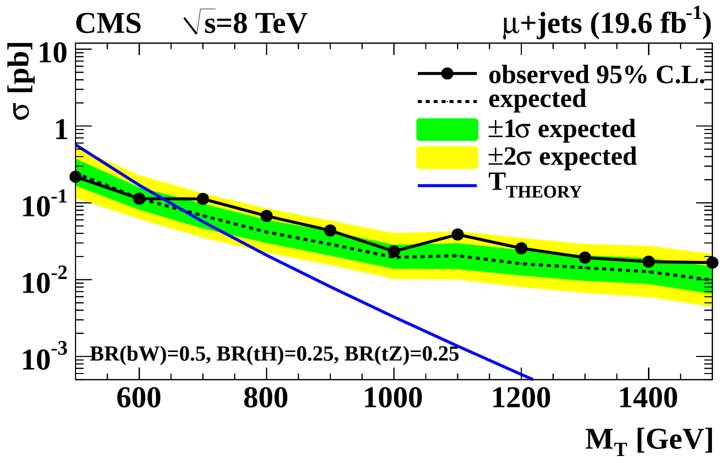


Figure 11.34: Shows the total lepton-plus-jets observed and expected limits on the $T\bar{T}$ production cross section as a function of T-quark mass. The data correspond to 19.6 fb^{-1} for both the e+jets channel and μ+jets channels. The branching ratios correspond to version 0 nominal listed in Table 11.16.

We perform a few cross-checks of our sensitivity and the robustness of our procedure. An additional b-tag selection requirement of ≥ 1 b-tags is considered in Appendix C.1 as we expect a large b-jet multiplicity in signal events. In Appendix C.2, we compare the sensitivity of our `BDT` procedure with a naïve H_T vs number of b-tag techniques. Finally, the change of sensitivity by training on the different signal samples individually is shown in Appendix C.5. No significant improvement over the standard method is observed in any of these situations.





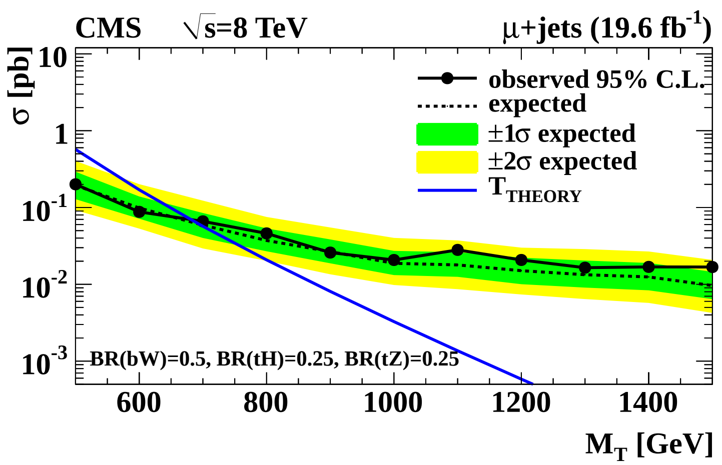


Figure 11.35: The μ+jets observed and expected limits on the $T\bar{T}$ production cross section as a function of T-quark mass: (top) shows the limits of the μ+jets events without substructure, (middle) μ+jets with substructure and (bottom) the combined μ+jets channel. The data correspond to 19.6 fb^{-1} for both the e+jets channel. The branching ratios correspond to version 0 nominal listed in Table 11.16

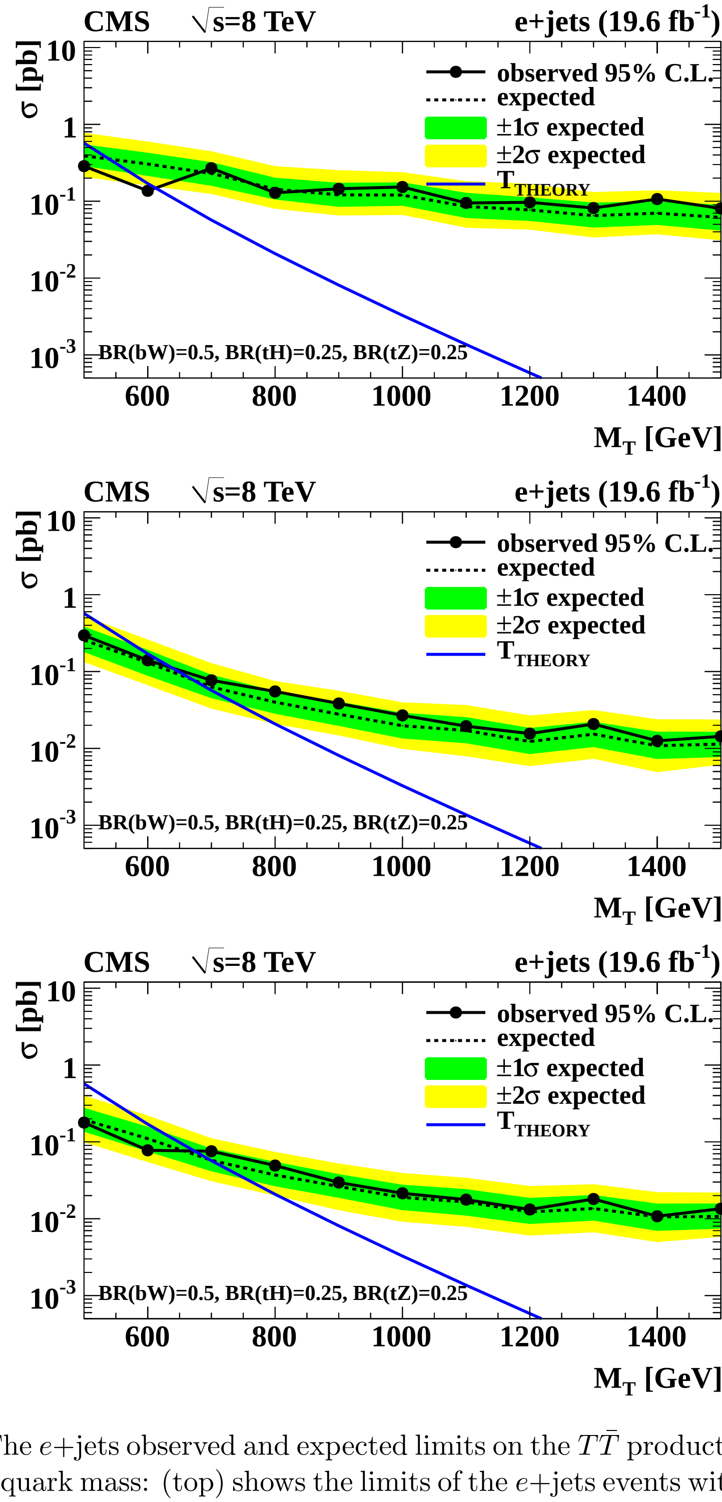


Figure 11.36: The e+jets observed and expected limits on the $T\bar{T}$ production cross section as a function of T-quark mass: (top) shows the limits of the e+jets events without substructure, (middle) e+jets with substructure and (bottom) the combined e+jets channel. The data correspond to 19.6 fb^{-1} for both the e+jets channel. The branching ratios correspond to version 0 nominal listed in Table 11.16

11.10 Limits for Different Branching Ratios

The relative branching ratios of the different T-quark decay modes are now varied. As mentioned, the `BDT` training from the nominal branching ratio is used to create the `BDT` discriminants for all branching ratio scenarios.

We set limits on a total of 66 sets of branching ratio values by scanning across the phase space in 0.1 BR steps. The expected and observed limits for the combined lepton-plus-jets channel for the 0.2 step intervals, as well as the nominal scenario, are listed in Table 11.16. The 1D limits for the non-nominal branching ratios, including separate electron and muon channel limits, can be found in Appendix C.7.

The expected limits of the combined lepton-plus-jets channel for the scenarios with branching ratios $(0), (1), (6), (21)$ corresponding to the nominal branching fractions, 100% branching fraction to bW, 100% branching fraction to tH, and 100% branching fraction to tZ respectively are compared in Fig. 11.37.

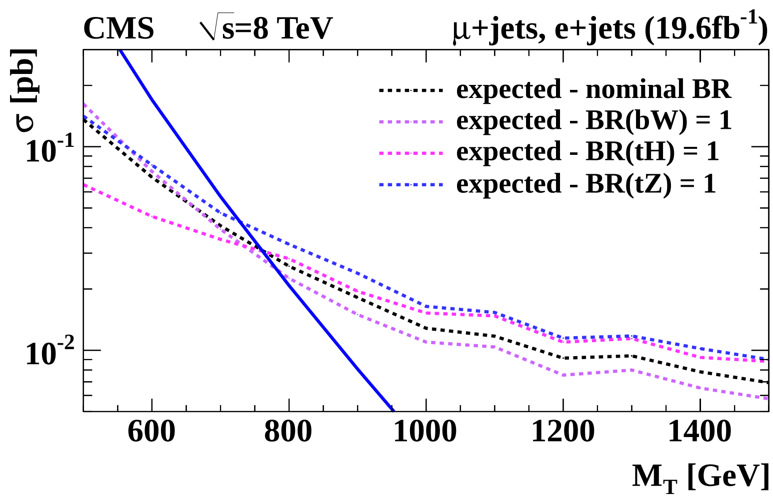


Figure 11.37: Comparison of the expected limits for the total lepton-plus-jets channel for the scenarios with branching ratios $(0), (1), (6), (21)$ corresponding to the nominal, 100% bW, 100% tH, and 100% tZ.

We notice that at low masses the expected limit for tH is strongest, likely due to the large number of b-jets in the signal. At higher masses, the kinematics dominate the importance in the `BDT` with $T \to bW$ giving the most sensitive expected limit. Events with a $T \to tZ$ decay have the lowest sensitivity since this final state has only one b-jet (in comparison

	Branching Fractions			expected	observed
Scenario	T→bW	T→tH	T→tZ	limit	limit
(0) Nominal	0.5	0.25	0.25	775 GeV	707 GeV
(1) Full tZ	0.0	0.0	1.0	743 GeV	691 GeV
(2)	0.0	0.2	0.8	745 GeV	695 GeV
(3)	0.0	0.4	0.6	754 GeV	697 GeV
(4)	0.0	0.6	0.4	761 GeV	704 GeV
(5)	0.0	0.8	0.2	770 GeV	715 GeV
(6) Full tH	0.0	1.0	0.0	774 GeV	730 GeV
(7)	0.2	0.0	0.8	751 GeV	693 GeV
(8)	0.2	0.2	0.6	758 GeV	694 GeV
(9)	0.2	0.4	0.4	765 GeV	697 GeV
(10)	0.2	0.6	0.2	772 GeV	708 GeV
(11)	0.2	0.8	0.0	780 GeV	718 GeV
(12)	0.4	0.0	0.6	763 GeV	696 GeV
(13)	0.4	0.2	0.4	770 GeV	698 GeV
(14)	0.4	0.4	0.2	778 GeV	704 GeV
(15)	0.4	0.6	0.0	781 GeV	718 GeV
(16)	0.6	0.0	0.4	770 GeV	699 GeV
(17)	0.6	0.2	0.2	778 GeV	713 GeV
(18)	0.6	0.4	0.0	786 GeV	724 GeV
(19)	0.8	0.0	0.2	782 GeV	721 GeV
(20)	0.8	0.2	0.0	788 GeV	732 GeV
(21) Full bW	1.0	0.0	0.0	790 GeV	728 GeV

Table 11.16: Shows the 22 sets of branching ratio values and the corresponding limits for the combined lepton-plus-jets channel in 0.2 interval steps.

to tH decays that have two) and also softer kinematics (in comparison to the $T \to bW$ channel, due to the higher masses final state particles).

The information from Table 11.16 is combined with similar limits set at all other 0.1 BR steps. We thus present the lower bounds on the T-quark mass across the whole of (branching fraction) phase space in Fig. 11.38, below which we can rule out the T-quark at 95% C.L. These triangle plots represent the expected and observed mass limits for all 66 BR scenarios in the μ+jets, e+jets, and combined lepton+jets channels separately.

The muon and electron channels exhibit similar behaviour, with the strongest exclusion limit set in the high BR(bW) regime and the weakest limits being set in the high BR(tZ) regime. This is reflective of the `BDT` signal distributions shown in Appendix C.3 and is due to the hard kinematics of the $T \to bW$ mode and the high b-tag multiplicity of the $T \to tH$ decay, which give strong discrimination power in their respective decay channels.

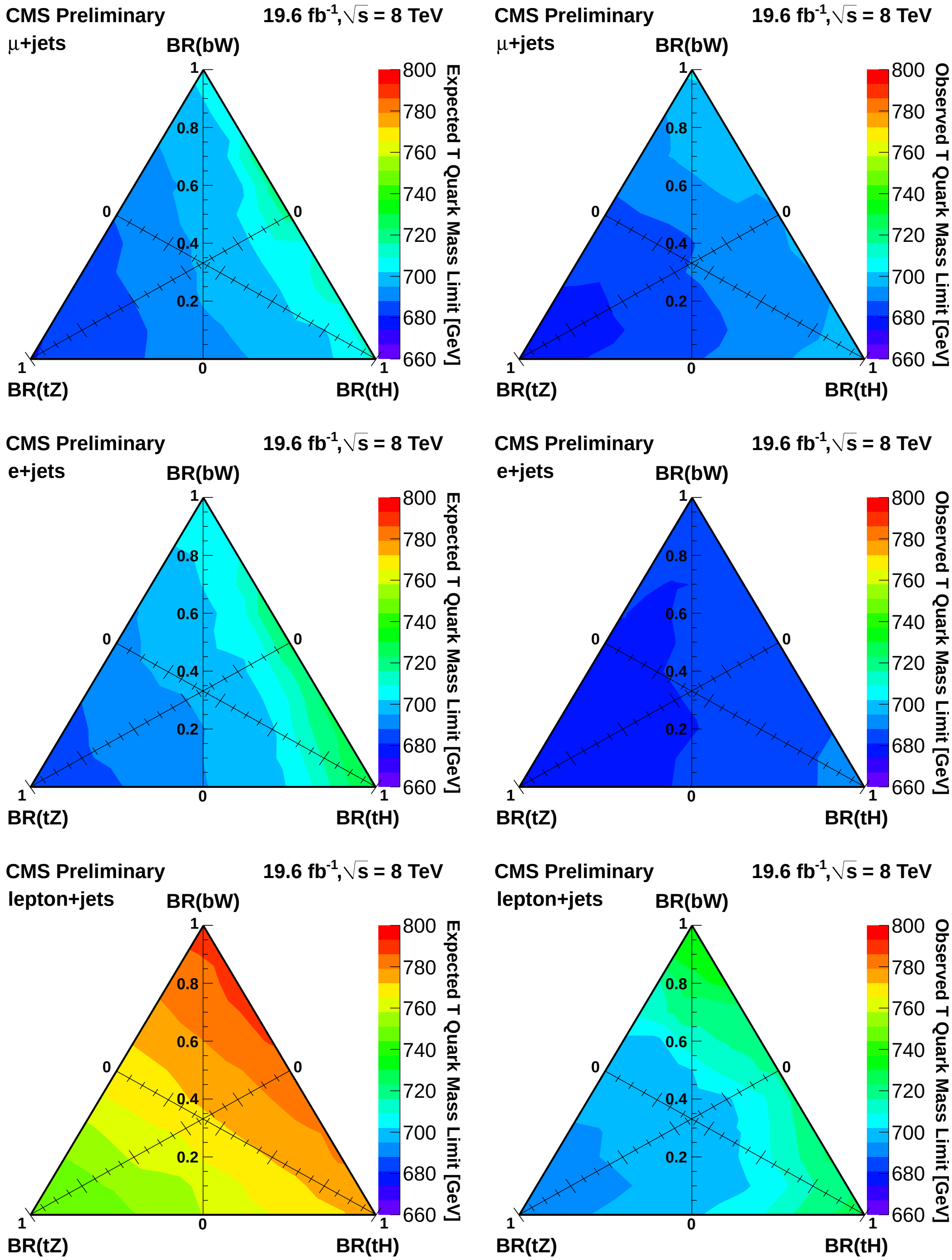


Figure 11.38: Expected and observed mass limits in the (top) μ+jets, (middle) e+jets, and (bottom) combined lepton+jets channels for the $T\bar{T}$ production. The data correspond to 19.6 fb^{-1} for both the μ+jets and e+jets channels.

11.11 Conclusions

We have searched for the pair-production of a heavy vector-like T-quark, with charge $\frac{2}{3}$, in events with exactly one isolated lepton using $19.6\,\mathrm{fb}^{-1}$ of data collected at CMS at $\sqrt{s} = 8\,\mathrm{TeV}$. Each T-quark is assumed to decay promptly to one of the bW, tH, or tZ final states. Our data exhibit no excess over SM expectations and upper limits on the production cross section have been set across the entire phase space of branching ratios, shown in Fig. 11.38. The existence of the T-quark has been ruled out in the combined lepton-plus-jets channel below a mass of approximately 700 GeV dependent on the exact BR at 95% C.L.

A brief summary of the partner multilepton analysis is discussed in Appendix C.8 and a combined result for the inclusive search for a vector-like T-quark, decaying into at least one lepton, is shown in Appendix C.9 [3]. This enhances our sensitivity by approximately 30 GeV across all of phase space, ruling out masses below approximately 730 GeV, again dependent on branching ratio. The limits that we have set place the most stringent constraints on the existence of such a quark to date.

In general, the Little Higgs Models that predict these vector-like quarks fall into two classes of theories [162]: Littlest Higgs Models that are built on product group extensions to the SM, and simple $SU(3)$ group models, such as the Simplest Little Higgs (described in Chapter 10), that minimally extend the electroweak gauge group. In both cases, the predicted branching ratios of the T-quark is model independent, with values BR(bW)=$\frac{1}{2}$, BR(tH)=$\frac{1}{4}$ and BR(tZ)=$\frac{1}{4}$. In this nominal scenario, the T-quark is ruled out below 707 GeV in the lepton-plus-jets analysis and 702 GeV in the combined lepton-plus-jets and multilepton analysis at 95% C.L.

Furthermore, the mass of the T-quark is predicted to be:

$$m_T = \begin{cases} \sqrt{\lambda_1^2 + \lambda_2^2} f & \text{in the Littlest Higgs Models} \\ \sqrt{\lambda_1^2 c_\beta^2 + \lambda_2^2 s_\beta^2} f & \text{in the SU(3) simple group models} \end{cases} \tag{11.3}$$

For a perturbative theory, we require $\lambda \lesssim 1$. This implies that the higher order vacuum, where new physics arises, must have an expectation value above $f \gtrsim 700\,\mathrm{GeV}$.

This is the first analysis of its kind and rules out new physics in this entire class of models below 700 GeV at 95% C.L.

Appendices

Appendix A

Classification of Particles

In this appendix, we shall give a brief introduction to the Lie groups that define the Standard Model. We shall describe the manner in which symmetries can be used to define physically equivalent states and notably, we shall identify the light quarks as living in a particular representation of an $SU(3)$ group.

Finally, we shall define a procedure to classify all known composite particles known as hadrons. These particles are made up of quarks and will be further subdivided into the set baryons, made of linear combinations of three quarks (or three antiquarks), and the mesons, made of linear combinations of a quark-antiquark pair. We shall in general follow Ref [5], where the measured values of the hadron masses may be found in Ref [4].

A.1 $U(1)$

The group $G \equiv U(1)$ corresponds to the one-dimensional unitary transformation. In the matrix representation D, the group elements can be written as $D(g) = e^{i\alpha T}$ where $g \in G$, $\alpha \in \mathbb{R}$ is some arbitrary constant and T is a pure complex unit vector known as the generator of the symmetry.

In physical terms, the single group element can be viewed as a complex phase. Irreducible representations of the group contain just one basis state, known as the "singlet" of $U(1)$, which is invariant under $D(g)$ by construction.

In general, the set of generators span a vector space V and together with the commutator

$$[\cdot,\cdot] \quad : \quad V \times V \to V \tag{A.1}$$

$$: \quad x \times y \ \to [x,y] \equiv x \cdot y - y \cdot x \tag{A.2}$$

define the Lie algebra.

By Lie's theorems [215,216] there is a one-to-one correspondence between the group and its algebra. It is often simpler to consider the Lie algebra of a Lie group, as it provides an obvious linear vector space on which to do calculations.

A.2 $SU(2)$

The $SU(2)$ group and corresponding Lie algebra represent the simplest non-trivial case. The group elements are unitary $\exp(i\alpha \cdot J) \in G$ where the generators J are now 2×2 matrices with determinant $+1$. For notational convenience, we shall absorb the complex i factor into the generator J and define the vector space as pure real.

The commutation relations that govern the Lie algebra is

$$[J_i, J_j] = i\epsilon_{ijk} J_k \tag{A.3}$$

with completely antisymmetric tensor ϵ_{ijk}, for $i,j,k = 1,2,3$. The constant factors on the right of the commutator are known as the structure constants and completely define the Lie algebra and consequently the Lie group as well.

A.2.1 J_3 eigenstates

We cannot, in general, find a common set of eigenstates for all generators in the algebra. In the matrix representation, this corresponds to not being able to find simultaneously diagonalisable matrices that represent the generators. The best we can do is diagonalise as many as we can that share a common set of eigenstates defined by their eigenvalues. These eigenvalues of the diagonal operators are known as weights.

One then constructs "ladder" operators by linear combinations of the remaining non-diagonal generators. These operators transform the state into an equivalent one. One then constructs all physically equivalent states in an irreducible representation by repeated application of these generators.

In general, there can be more than one set of ladder operators and one needs to take all linear combinations of the lowering operators to construct all possible states that furnish an irreducible representation.

In physical cases, the representations are assumed finite and happen to fall into relatively simple irreducible representations of the theory.

Since matrices commute if and only if they are simultaneously diagonalisable, we see that we can only diagonalise one of the generators of $SU(2)$. Canonically, this is the J_3 generator. We can now proceed to construct that ladder operators that will define the states in the irreducible representations of $SU(2)$.

Since we assume that the representation is finite dimensional, there must be some unique *highest weight* state(s) [5], which has the largest eigenvalue of J_3 with eigenvalue labelled j such that

$$J_3 \, |j, \alpha\rangle = j \, |j, \alpha\rangle \tag{A.4}$$

where α labels the degeneracy of the state that is not affected by group theory but are measurable physical properties. We shall often suppress α as they are irrelevant in regards to the group theory of the system.

As we are attempting to construct unique irreducible representations, we require all states $|j, \alpha\rangle$ to form an orthonormal basis where $\langle j, \alpha | j, \beta\rangle = \delta_{\alpha\beta}$.

The $SU(2)$ raising and lowering ladder operators are defined as

$$J^{\pm} \equiv \frac{1}{\sqrt{2}}(J_1 \pm iJ_2) \tag{A.5}$$

It immediately follows from (A.3) that

$$\begin{aligned} \left[J_3 \, , J^{\pm}\right] &= \pm J^{\pm} \\ \left[J^{+}, J^{-}\right] &= J_3 \end{aligned} \tag{A.6}$$

On a general state $|m,\alpha\rangle$ with $J_3\,|m,\alpha\rangle = m\,|m,\alpha\rangle$ we find

$$
\begin{aligned}
J_3 J^{\pm}\,|m,\alpha\rangle &= ([J^3, J^{\pm}] + J^{\pm}J_3)\,|m,\alpha\rangle \\
&= (m \pm 1) J^{\pm}\,|m,\alpha\rangle
\end{aligned}
\tag{A.7}
$$

We can therefore define $J^{\pm}\,|m\rangle \equiv N\,|m+1\rangle$ up to some normalisation N. The normalisation is straight forward to compute directly from the orthonormality condition. Thus,

$$
J^{\pm}\,|m\rangle = \frac{1}{\sqrt{2}}\sqrt{(j \mp m)(j \pm m + 1)}\,|m \pm 1\rangle \tag{A.8}
$$

Continuing iteratively from the highest weight state, there are $2j+1$ states with $m = -j, -j+1, \cdots, j$ in integer steps defining the states in the irreducible representation.

The largest eigenvalue j, corresponding to the highest weight state is necessarily half-odd integer and uniquely defines the representation. As such, it is often convenient to denote a general state in the representation j as $|j,m\rangle$.

The matrix elements of the generators are

$$
\langle jm'|\,J_i\,|jm\rangle = [J_i^j]_{mm'} \tag{A.9}
$$

where a labels the generator, j the weight of the highest weight state, and $m, m' \in \{-j, \cdots, +j\}$.

In particular,

$$
\langle jm'|\,J_3\,|j,m\rangle = m\delta_{mm'} \tag{A.10}
$$

by construction and using the definitions (A.5), the $J^{\pm}$ matrix elements are

$$
\begin{aligned}
\langle jm'|\,J^{+}\,|jm\rangle &= \sqrt{(j+m+1)(j-m)/2}\,\delta_{m',m+1} \\
\langle jm'|\,J^{-}\,|jm\rangle &= \sqrt{(j+m)(j-m+1)/2}\,\delta_{m',m-1}
\end{aligned}
\tag{A.11}
$$

The $j = \frac{1}{2}$ irreducible representation thus has the following generators,

$$J_i^{\frac{1}{2}} = \begin{pmatrix} \langle j, -\frac{1}{2} | \, J_i \, | j, -\frac{1}{2} \rangle & \langle j, -\frac{1}{2} | \, J_i \, | j, +\frac{1}{2} \rangle \\ \langle j, +\frac{1}{2} | \, J_i \, | j, -\frac{1}{2} \rangle & \langle j, +\frac{1}{2} | \, J_i \, | j, +\frac{1}{2} \rangle \end{pmatrix}$$

Explicitly, using (A.5) and (A.11), we find

$$\begin{aligned} J_1^{\frac{1}{2}} &= \frac{1}{2} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \\ J_2^{\frac{1}{2}} &= \frac{1}{2} \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \\ J_3^{\frac{1}{2}} &= \frac{1}{2} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \end{aligned} \tag{A.12}$$

Expressing this in terms of the Pauli spin matrices σ_i, we have

$$J_i^{\frac{1}{2}} = \frac{i}{2} \sigma_i \tag{A.13}$$

for $i = 1, 2, 3$.

From the Lie algebra, we have now explicitly constructed the generators of the group in matrix form as well as ennummerated the states in the irreducible representations. The irreducible representations with highest weight j, correspond to set with $m = \{-j, \ldots, j\}$ equivalent states, which transform amongst themselves under the ladder operators.

In the simplest case $SU(2)$ representation $j = \frac{1}{2}$ there are two states. This irreducible representation is known as the fundamental representation of $SU(2)$, often referred to as the doublet, and is denoted **2**.

Physically, The degeneracy m can be lifted by a perturbation. Indeed, intrinsic spin was historically discovered by the Stern and Gerlach in 1922 when electrons travelled through a magnetic field along two different paths, indicative of an physically unobservable degree of freedom that lived in a doublet of $SU(2)$.

Reducibility

A representation is not necessarily irreducible, but one can always decompose the space into irreducible components. The prescription is to start with the highest weight state and construct an irreducible representation using the ladder operators. One then removes these states, making use of the fact that they are orthogonal to the remaining states, and iterates on the remaining space until all states fall into exactly one irreducible representation.

The tensor product space

We define a tensor product space, whose states are formed from the product of states $|m\rangle \equiv |m_1\rangle \otimes |m_2\rangle$. We then define, $D(g) \equiv D_{1\otimes 2}(g)$ as follows,

$$\begin{aligned}
D(g)\,|m_1, m_2\rangle &= \left|m_1' m_2'\right\rangle [D_{1\otimes 2}(g)]_{m_1' m_2' m_1 m_2} \\
&= \left|m_1', m_2'\right\rangle [D_1(g)]_{m_1' m_1} [D_2(g)]_{m_2' m_2} \\
&= \left(\left|m_1'\right\rangle [D_1(g)]_{m_1' m_1}\right)\left(\left|m_2'\right\rangle [D_2(g)]_{m_2' m_2}\right) \qquad \text{(A.14)}
\end{aligned}$$

with both state $|m_1\rangle$ and $|m_2\rangle$ transforming independently.

Expanding the group element around the identity, we have the following

$$\begin{aligned}
D(g)\,|m_1 m_2\rangle &= (1 + i\alpha^a J_a)\,|m_1 m_2\rangle \\
&= \sum_{m_1', m_2'} \underbrace{\left|m_1' m_2'\right\rangle \left\langle m_1' m_2'\right|}_{=\mathbb{I}} (1 + i\alpha_a J_a\,|m_1 m_2\rangle) \\
&= \left|m_1' m_2'\right\rangle \delta_{m_1' m_1} \delta_{m_2' m_2} + i\alpha^a \left|m_1' m_2'\right\rangle \underbrace{\left\langle m_1' m_2'\right| J_a \,|m_1 m_2\rangle}_{=[J_a^{1\otimes 2}(g)]_{jyix}} \\
&= \left|m_1' m_2'\right\rangle \left(\delta_{m_1' m_1} + i\alpha^a [J_a^1]_{m_1' m_1}\right)\left(\delta_{m_2' m_2} + i\alpha^a [J_a^2]_{m_2' m_2}\right) \\
&\equiv \left(\left|m_1'\right\rangle [D_1(g)]_{m_1' m_1}\right)\left(\left|m_2'\right\rangle [D_2(g)]_{m_2' m_2}\right) \qquad \text{(A.15)}
\end{aligned}$$

the third equality follows if and only if we take the infinitesimal transformation with $\alpha_a \to 0$ with

$$\left[J_a^{1\otimes 2}(g)\right]_{m_1' m_1 m_2' m_2} \equiv [J_a^1]_{m_1' m_1} \delta_{m_2' m_2} + \delta_{m_2' m_2} [J_a^2]_{m_1' m_1} \qquad \text{(A.16)}$$

The operation on a product space is equivalent to the addition of the two transformations separately.

For $SU(2)$,

$$
\begin{aligned}
J_3(|j_1, m_1\rangle\, |j_2, m_2\rangle) &= (J_3\, |j_1, m_1\rangle)\, |j_2, m_2\rangle + |j_1, m_1\rangle\, (J_3\, |j_2, m_2\rangle) \\
&= (m_1 + m_2)\, |j_1, m_1\rangle\, |j_2, m_2\rangle \qquad \text{(A.17)}
\end{aligned}
$$

In this way, the J_3 values add in a tensor product space.

Note that there is still a (unique) highest weight state denoted

$$
|j_1 + j_2, m_1 + m_2\rangle = |j_1, m_1\rangle\, |j_2.m_2\rangle \qquad \text{(A.18)}
$$

which has $J_3^{1\otimes 2}$ eigenvalue $j_1 + j_2$. Acting repeatedly with $J^-_{1\otimes 2}$ yields the states with different $J_3^{1\otimes 2}$ values and furnishes the irreducible representation. One continues as before, removing these states and iterating.

A.2.2 The $SU(2)$ weight diagram

The weights of the group correspond to the eigenvalues of the eigenstates. Fig. A.1 shows the weight diagrams for the $j = \frac{1}{2}$ and $j = 1$ representations of $SU(2)$. In both cases there are $2j + 1$ states as required. Note that $j = 1$, is not irreducible and in fact can be formed by the tensor product of two $j = \frac{1}{2}$ irreducible representations.

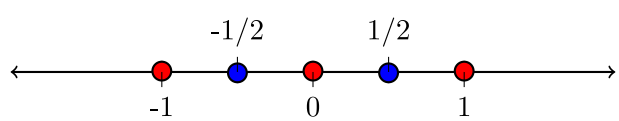


Figure A.1: Weight diagram for (blue) $j = \frac{1}{2}$ and (red) $j = 1$ representations of the $SU(2)$ Lie algebra.

The weight diagrams of the more complicated product representation $(j_1 = \frac{1}{2}) \otimes (j_2 = \frac{1}{2}) \otimes (j_3 = \frac{1}{2})$ is given in Fig. A.2.

These types of weight diagrams can be enumerated from the number of combinations of lowering operators that a certain J_3 eigenvalue can appear. For example, the state of weight $+\frac{1}{2}$ can be constructed in three ways by three combinations of lowering operators,

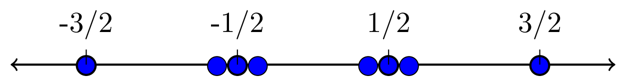


Figure A.2: Weight diagram of the $\frac{1}{2} \otimes \frac{1}{2} \otimes \frac{1}{2}$ product representation.

- $\left|j = \frac{1}{2}, m = -\frac{1}{2}\right\rangle \otimes \left|\frac{1}{2}, \frac{1}{2}\right\rangle \otimes \left|\frac{1}{2}, \frac{1}{2}\right\rangle$
- $\left|j = \frac{1}{2}, m = \frac{1}{2}\right\rangle \otimes \left|\frac{1}{2}, \frac{1}{2}\right\rangle \otimes \left|\frac{1}{2}, -\frac{1}{2}\right\rangle$
- $\left|j = \frac{1}{2}, m = \frac{1}{2}\right\rangle \otimes \left|\frac{1}{2}, \frac{1}{2}\right\rangle \otimes \left|-\frac{1}{2}, \frac{1}{2}\right\rangle$

More formally, one takes the highest weight state $\left|\frac{3}{2}, \frac{3}{2}\right\rangle = |1, 1\rangle \otimes \left|\frac{1}{2}, \frac{1}{2}\right\rangle$, and acts with J^- until we get to a lowest weight state that is annihilated by the lowering operator. Then, one uses $|1, 1\rangle = \left|\frac{1}{2}, \frac{1}{2}\right\rangle \otimes \left|\frac{1}{2}, \frac{1}{2}\right\rangle$ to find the complete weight diagram. There is thus three ways of lowering the state $\left|\frac{3}{2}, \frac{3}{2}\right\rangle$ to get to $\left|\frac{3}{2}, \frac{1}{2}\right\rangle$.

The $SU(2)$ group will play several crucial roles in particle physics, most notably of the description of spin and as part of the Electroweak force in the Standard Model.

A.3 $SU(3)$

The $SU(3)$ Lie algebra is defined by the commutation relation

$$[T_i, T_j] = f^{ijk} T_k \tag{A.19}$$

with structure constants

- $f^{123} = 1$
- $f^{147} = f^{165} = f^{246} = f^{257} = f^{345} = f^{376} = \frac{1}{2}$
- $f^{458} = f^{678} = \frac{\sqrt{3}}{3}$

and all else zero.

The matrix generators that satisfy these relations can be constructed from the Gell-Mann matrices, 3×3 analogues of the Pauli matrices.

We define $T_a = -\frac{i}{2}\lambda_a$ where λ_a are the Gell-Mann matrices

$$\lambda_1 = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \; \lambda_2 = \begin{pmatrix} 0 & -i & 0 \\ i & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \; \lambda_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

$$\lambda_4 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}, \; \lambda_5 = \begin{pmatrix} 0 & 0 & -i \\ 0 & 0 & 0 \\ i & 0 & 0 \end{pmatrix}, \; \lambda_6 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}$$

$$\lambda_7 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -i \\ 0 & i & 0 \end{pmatrix}, \; \lambda_8 = \frac{1}{\sqrt{3}} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -2 \end{pmatrix} \tag{A.20}$$

For $SU(3)$ two matrices can be simultaneously diagonalised; canonically, $h_1 \equiv \lambda_3$ and $h_2 \equiv \lambda_8$. Therefore, the states in each representation are defined by two weights, which are eigenvalues of (h_1, h_2).

From the remaining non-diagonal generators, we can construct three sets of raising and lowering ladder operators

$$\begin{aligned} e_1^{\pm} &= \frac{1}{2\sqrt{2}}(\lambda_1 \pm i\lambda_2) \\ e_2^{\pm} &= \frac{1}{2\sqrt{2}}(\lambda_6 \pm i\lambda_7) \\ e_3^{\pm} &= \frac{1}{2\sqrt{2}}(\lambda_4 \pm i\lambda_5) \end{aligned} \tag{A.21}$$

A.3.1 The Fundamental Representation 3 and The Light Quarks

The lowest dimensional irreducible representation of $SU(3)$ are $1,\ 3,\ 6,\ 10, \ldots$ dimensional. The trivial singlet representation **1** consists of only one state of weight $(0, 0)$.

The fundamental representation, defined by $D(g) = g \forall g \in G$, of the $L(SU(3))$ algebra is denoted **3** and has eigenstates

$$\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \text{ and } \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \tag{A.22}$$

which form a 3 dimensional set of states that span the representation's vector space.

The weights of these states are: where we have identified these states as the light

state	weight
$\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \equiv u$	$\left(\frac{1}{2}, \frac{1}{2\sqrt{3}}\right)$
$\begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \equiv d$	$\left(-\frac{1}{2}, \frac{1}{2\sqrt{3}}\right)$
$\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \equiv s$	$\left(0, -\frac{1}{2\sqrt{3}}\right)$

quark states u,d, and s. In this way, we have related the light quarks to the states in the fundamental representation of $SU(3)$. Note that in reality, there is only a non-exact symmetry that relates the quarks, as their masses are not identical. The breaking of this approximate global symmetry that governs these three light flavours of quarks gives rise to the massive Pions as pseudo-Goldstone bosons.

We can relate the states to each other using the lowering operators from the highest weight state $u \equiv \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$. The weight diagram can easily be constructed and is shown in Fig. A.3.

The complex conjugate representation $\overline{D}(g) \equiv (D(g))^{\star}$ is known as the anti-fundamental representation denoted $\overline{\mathbf{3}}$. From this definition, the weights can be obtained by a transformation of those in the fundamental representation by the identification:

$$(p, q) \to -(p, q) \tag{A.23}$$

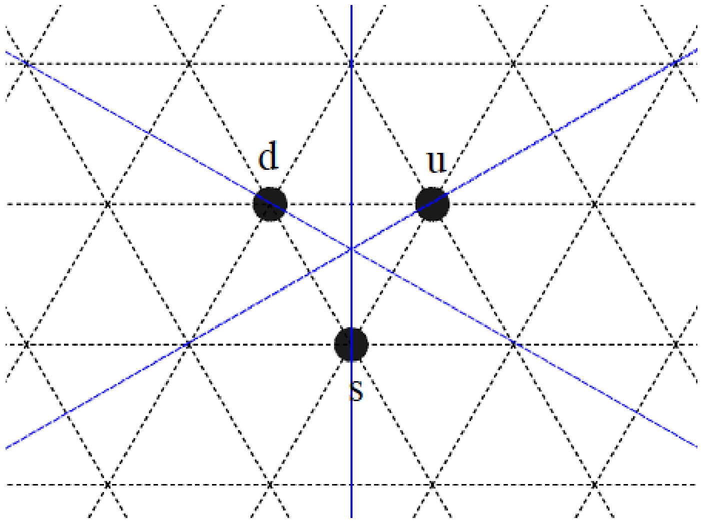


Figure A.3: The light quarks in the fundamental representation of $SU(3)$. [5]

The basis states of the anti-fundamental representation are:

state	weight
$\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \equiv \overline{s}$	$\left(0, \frac{1}{2\sqrt{3}}\right)$
$\begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \equiv \overline{d}$	$\left(\frac{1}{2}, -\frac{1}{2}s\right)$
$\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \equiv \overline{u}$	-$\left(\frac{1}{2}, \frac{1}{2\sqrt{3}}\right)$

Again, all weights are related through the complex conjugate of the lowering and raising operators defined in (A.21) and give the weight diagram Fig. A.4.

A.3.2 The Adjoint Representation

There is an eight dimensional representation of $SU(3)$ known as the adjoint representation, defined by

$$\mathrm{ad}(v)w = [v, w] \quad \forall v, w \in L(G) \tag{A.24}$$

Note that this is merely a restating of the commutation relation and is a very natural representation of the Lie algebra with the structure constants completely defining this rep-

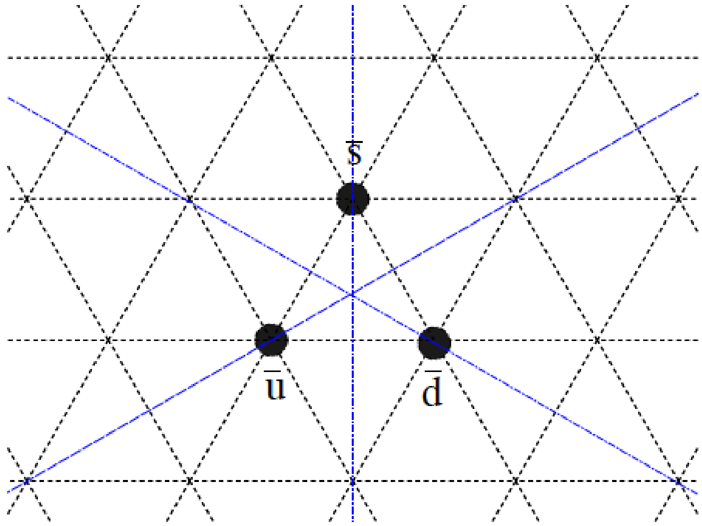


Figure A.4: Weight diagram for the anti-fundamental representation of $SU(3)$. [5]

resentation. Note that we are just taking the generators of the group as acting on the vector space spanned by the Lie algebra basis elements w using the commutator.

To determine the weight (p, q) of any state w, one just computes the commutator with the diagonalised generators h_1, h_2. The states and corresponding weights are:

State w	$[h_1, w]$	$[h_2, w]$	weights
h_1	0	0	(0,0)
h_2	0	0	(0,0)
e^1_+	e^1_+	0	(1,0)
e^1_-	$-e^1_-$	0	(-1,0)
e^2_+	$-\frac{1}{2}e^2_+$	$\frac{\sqrt{3}}{2}e^2_+$	$(-\frac{1}{2}, \frac{\sqrt{3}}{2})$
e^2_-	$\frac{1}{2}e^2_-$	$-\frac{\sqrt{3}}{2}e^2_-$	$(\frac{1}{2}, -\frac{\sqrt{3}}{2})$
e^3_+	$\frac{1}{2}e^3_+$	$\frac{\sqrt{3}}{2}e^3_+$	$(\frac{1}{2}, \frac{\sqrt{3}}{2})$
e^3_-	$-\frac{1}{2}e^3_-$	$-\frac{\sqrt{3}}{2}e^3_-$	$(-\frac{1}{2}, -\frac{\sqrt{3}}{2})$

The highest weight state is e^3_+ the highest weight state, of weight $(\frac{1}{2}, \frac{\sqrt{3}}{2})$ and the weight diagram forms a regular hexagon, shown in Fig. A.5.

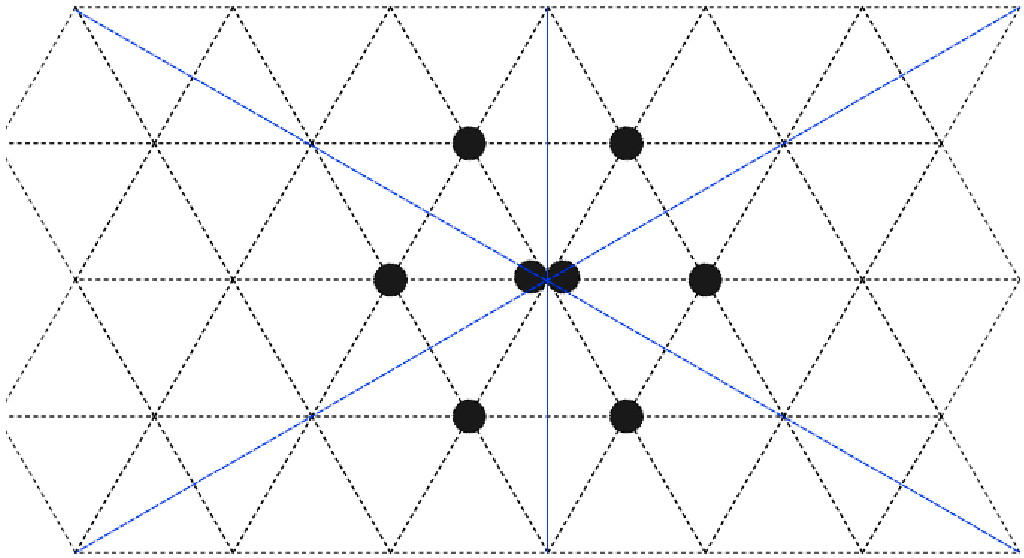


Figure A.5: Hexagonal weight diagram of the adjoint representations hexagonal. [5]

As there are 8 generators in $SU(3)$, we see that the adjoint representation is 8 dimensional and is often denoted **8**.

A.3.3 The Eightfold Way

Murray Gell-Man, first noted that composite particles called mesons and baryons, composed of a quark-antiquark and three quarks (or three antiquarks) respectively, can be grouped into different representations of $SU(3)$.

This follows as a direct consequence from the strong nuclear force being invariant under the flavour symmetry between the three kinds of light quark. Due to their different masses, the symmetry is not exact. However, we are still able to classify the collective set of baryons and mesons using product represents of this approximate symmetry group.

A.3.4 $\mathbf{3} \otimes \overline{\mathbf{3}}$

By decomposing the tensor product representations $\mathbf{3} \otimes \overline{\mathbf{3}}$, we can find the physically relevant irreducible representations.

The states and corresponding weights of this product representation are:

state	weight
$u \otimes \overline{s}$	$(\frac{1}{2}, \frac{\sqrt{3}}{2})$
$u \otimes \overline{d}$	$(1, 0)$
$d \otimes \overline{s}$	$(-\frac{1}{2}, -\frac{\sqrt{3}}{2})$
$u \otimes \overline{u}, d \otimes \overline{d}, s \otimes \overline{s}$	$(0, 0)$
$d \otimes \overline{u}$	$(-\frac{1}{2}, -\frac{\sqrt{3}}{2})$
$s \otimes \overline{d}$	$(\frac{1}{2}, -\frac{\sqrt{3}}{2})$

The weight diagram is shown in Fig. A.6.

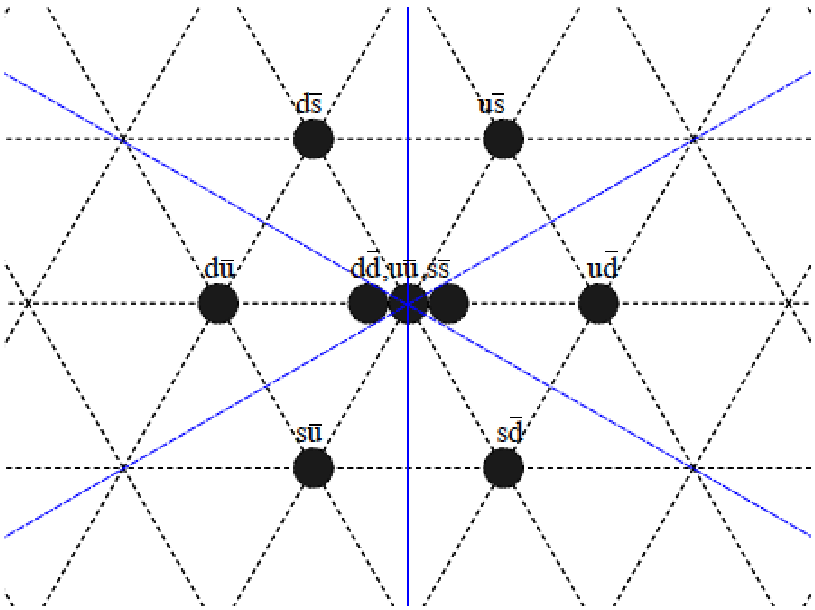


Figure A.6: $\mathbf{3} \times \overline{\mathbf{3}}$ weight diagram. [5]

Following the prescription of using the lowering operators on the highest weight states, we find that the $\mathbf{3} \otimes \overline{\mathbf{3}}$ representation can be reduced to the $\mathbf{8}$ adjoint and $\mathbf{1}$ singlet irreducible representations using the ladder operators

$$E_{\pm}^{m} = e_{\pm}^{m} \otimes \mathbb{I} + \mathbb{I} \otimes \overline{e}_{\pm}^{m} \tag{A.25}$$

where $m = 1, 2, 3$ and $\overline{e}_{\pm}^{m} = -\left(e_{\mp}^{m}\right)^{\star}$. [5]

The $\mathbf{8}$ in $\mathbf{3} \otimes \overline{\mathbf{3}}$ has unit normalised states and weights:

state	weight
$u \otimes \overline{s}$	$(\frac{1}{2}, \frac{\sqrt{3}}{2})$
$u \otimes \overline{d}$	$(1, 0)$
$d \otimes \overline{s}$	$(-\frac{1}{2}, -\frac{\sqrt{3}}{2})$
$\frac{1}{\sqrt{2}}\left(d \otimes \overline{d} - u \otimes \overline{u}\right), \frac{1}{\sqrt{6}}\left(d \otimes \overline{d} + u \otimes \overline{u} - 2s \otimes \overline{s}\right)$	$(0, 0)$
$d \otimes \overline{u}$	$(-\frac{1}{2}, -\frac{\sqrt{3}}{2})$
$s \otimes \overline{d}$	$(\frac{1}{2}, -\frac{\sqrt{3}}{2})$

Removing these states from the original tensor product representation we are left with the **1** state

$$\frac{1}{\sqrt{3}}\left(u \otimes \overline{u} + s \otimes \overline{s} + d \otimes \overline{d}\right) \tag{A.26}$$

We have therefore decomposed the product representation as $\mathbf{3} \otimes \overline{\mathbf{3}} = \mathbf{8} \oplus \mathbf{1}$, where the states in **8** are antisymmetric and the singlet is symmetric.

A.3.5 $\mathbf{3} \otimes \mathbf{3} \otimes \mathbf{3}$

The weight content of the $\mathbf{3} \otimes \mathbf{3} \otimes \mathbf{3}$ is:

state	weight
$u \otimes u \otimes u$	$(\frac{1}{2}, \frac{\sqrt{3}}{2})$
$s \otimes s \otimes s$	$(0, -\sqrt{3})$
$d \otimes d \otimes d$	$(-\frac{3}{2}, -\frac{\sqrt{3}}{2})$
$u \otimes u \otimes s, u \otimes s \otimes u, s \otimes u \otimes u$	$(1, 0)$
$u \otimes u \otimes d, u \otimes d \otimes u, d \otimes u \otimes u$	$(\frac{1}{2}, \frac{\sqrt{3}}{2})$
$u \otimes s \otimes s, s \otimes u \otimes s, s \otimes s \otimes u$	$(\frac{1}{2}, -\frac{\sqrt{3}}{2})$
$d \otimes s \otimes s, s \otimes d \otimes s, s \otimes s \otimes d$	$(-\frac{1}{2}, -\frac{\sqrt{3}}{2})$
$s \otimes d \otimes d, d \otimes s \otimes d, d \otimes d \otimes s$	$(-1, 0)$
$u \otimes d \otimes d, d \otimes u \otimes d, d \otimes d \otimes u$	$(-\frac{1}{2}, \frac{\sqrt{3}}{2})$
$u \otimes d \otimes s, u \otimes s \otimes d, d \otimes u \otimes s, d \otimes s \otimes u, s \otimes u \otimes d, s \otimes d \otimes u$	$(0, 0)$

The weights can again be organised into the weight diagram is shown in Fig. A.7.

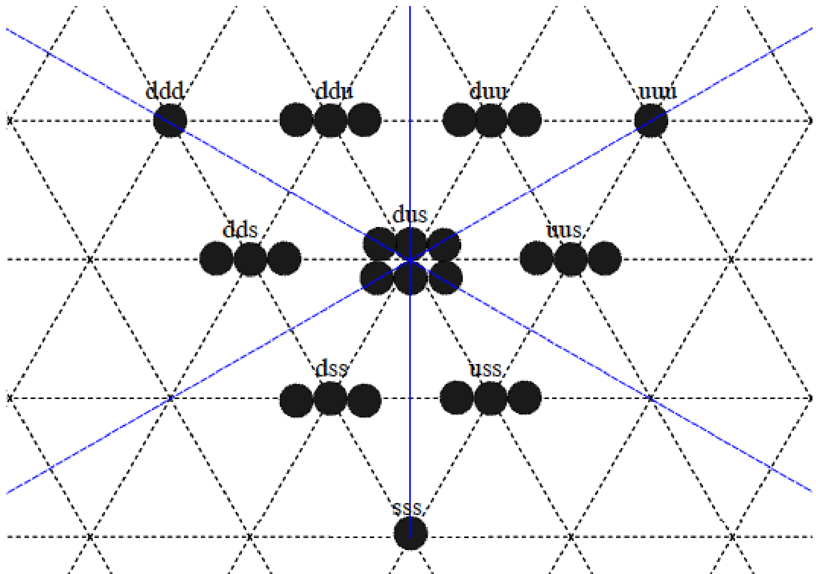


Figure A.7: Weights of the $\mathbf{3} \times \mathbf{3} \times \mathbf{3} = \mathbf{10} + \mathbf{8} + \mathbf{8} + \mathbf{1}$ product representation. [5]

Acting on the highest weight state $u \otimes u \otimes u$, which has weight $(\frac{1}{2}, \frac{\sqrt{3}}{2})$, with all combinations of ladder operators

$$E_{\pm}^{m} = e_{\pm}^{m} \otimes \mathbb{I} \otimes \mathbb{I} + \mathbb{I} \otimes e_{\pm}^{m} \otimes \mathbb{I} + \mathbb{I} \otimes \mathbb{I} \otimes e_{\pm}^{m} \tag{A.27}$$

constructs an irreducible representation. Removing these states and iterating, we can decompose the $\mathbf{3} \otimes \mathbf{3} \otimes \mathbf{3}$ into

$$\mathbf{3} \otimes \mathbf{3} \otimes \mathbf{3} = \mathbf{10} \oplus \mathbf{8} \oplus \mathbf{8} \oplus \mathbf{1} \tag{A.28}$$

where $\mathbf{10}$ is symmetric, $\mathbf{1}$ antisymmetric, and both $\mathbf{8}$'s have mixed symmetry.

A.4 The Quark Model

We are now ready to identify the linear combinations of light quarks with the observed particles in nature. We arrange the mesons and baryons into *multiplets* of $SU(3)$ that correspond to the states that span various tensor product spaces on which the irreducible representations of the symmetry group act.

All observed mesons and baryons can be naturally grouped into multiplets with the same baryon number B and total angular momentum $J = I + L$ where I the spin and L orbital angular momentum both live in representations of $SU(2)$.

The weights of the states are identified as

$$(p, q) \equiv (I_3, \frac{\sqrt{3}}{2} Y) \tag{A.29}$$

where

$$I_3 = \frac{1}{2}\left[(n_u - n_{\overline{u}}) - \left(n_d - n_{\overline{d}}\right)\right] \tag{A.30}$$

with n_u ($n_{\overline{u}}$) the number of up-type (anti) quarks and n_u ($n_{\overline{u}}$) the number of down-type (anti) quarks

Further, $Y = S + B$ is known as the (light) hypercharge; $S = n_s - n_{\overline{s}}$ is known as the *strangeness* where n_s ($n_{\overline{s}}$) is the number of strange (anti) quarks and $B \equiv \frac{1}{3}\left(n_q - n_{\overline{q}}\right)$ is the baryon number for n_q ($n_{\overline{q}}$) is the number of (anti) quarks.

A.4.1 Meson Multiplets

The meson multiplets are composed of a light quark-antiquark pair and can thus be placed in irreducible representations in the $\mathbf{3} \otimes \overline{\mathbf{3}}$ of $SU(3)$. The (light) meson states come from a $\mathbf{3} \otimes \overline{\mathbf{3}} = \mathbf{8} \oplus \mathbf{1}$ and make up sets of octets and singlets of mesons.

The pseudoscalar mesons $B = 0$, $J = 0$

The octet states in the pseudoscalar mesons with baryon number $B = 0$ and total angular momentum $J = 0$, correspond to the following meson octet:

meson	quark content	rest mass [MeV/c^2]
k^0	$d\bar{s}$	497.614 ± 0.024
k^+	$u\bar{s}$	493.677 ± 0.016
k^-	$\bar{u}s$	493.677 ± 0.016
$\bar{k}^0$	$\bar{d}s$	497.614 ± 0.024
π^+	$u\bar{d}$	139.57018 ± 0.00035
π^-	$\bar{u}d$	139.57018 ± 0.00035
π^0	$\frac{1}{\sqrt{2}}\left(u\bar{u} - d\bar{d}\right)$	134.9766 ± 0.0006
η	$\frac{1}{\sqrt{6}}\left(u\bar{u} + d\bar{d} - 2s\bar{s}\right)$	547.843 ± 0.024

The orthogonal singlet is:

$$\eta' = \frac{1}{3}\left(u\bar{u} + d\bar{d} + s\bar{s}\right) \tag{A.31}$$

with rest mass 957.78 ± 0.06 MeV/c^2.

These states can be organised in the (I_3, Y) weight diagram Fig. A.8.

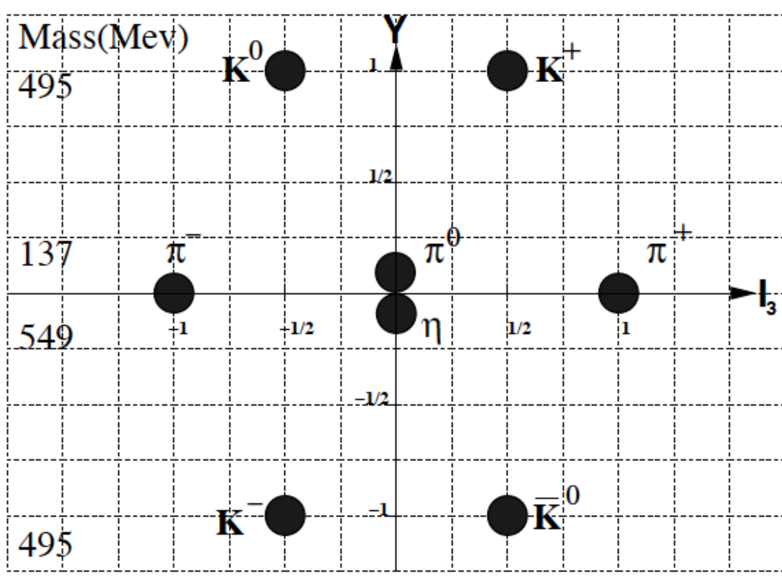


Figure A.8: Meson octet with $B = 0$, $J = 0$. There is also a meson singlet η'. [5]

The vector mesons $B = 0$, $J = 1$

The vector mesons with $B = 0$, $J = 1$ are also composed of an octet and a singlet.

The vector meson octet correspond to the following observed particles:

meson	quark content	rest mass [MeV/c^2]
$k^{0\star}$	$d\overline{s}$	895.94 ± 0.22
$k^{+\star}$	$u\overline{s}$	891.66 ± 0.026
$k^{-\star}$	$\overline{u}s$	891.66 ± 0.026
$\overline{k}^{0\star}$	$\overline{d}s$	895.94 ± 0.22
ρ^+	$u\overline{d}$	775.11 ± 0.34
ρ^-	$\overline{u}d$	775.11 ± 0.34
ρ^0	$\frac{1}{\sqrt{2}}\left(u\overline{u} - d\overline{d}\right)$	775.49 ± 0.34
ω	$\frac{1}{\sqrt{6}}\left(u\overline{u} + d\overline{d} - 2s\overline{s}\right)$	775.49 ± 0.34

and the vector meson singlet:

$$\phi = s\overline{s} \tag{A.32}$$

with mass $1019.445 \pm 0.020\,\mathrm{MeV}/c^2$.

The weight diagram for the vector meson octet is thus shown in Fig. A.9.

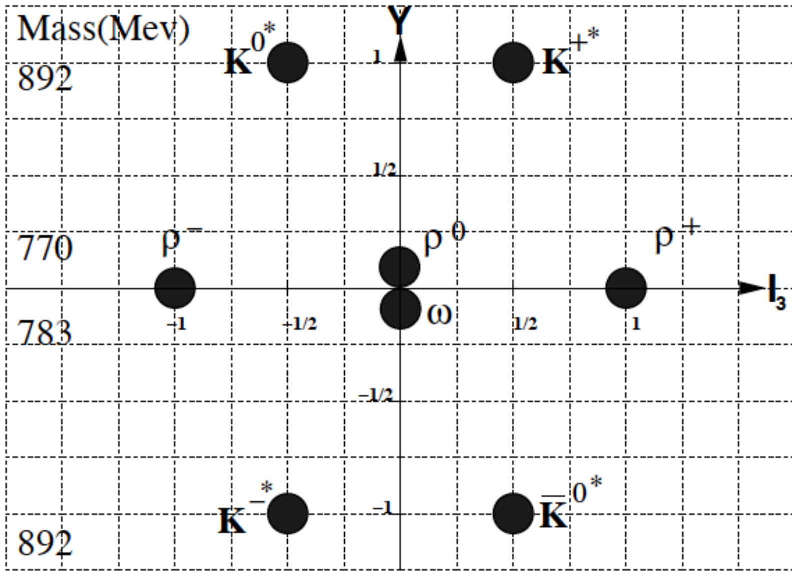


Figure A.9: Vector meson octet with $B = 0$, $J = 1$. [5] There is also a singlet meson ϕ.

A.4.2 Baryon Multiplets

The baryon multiplets are composed of particles made of three light quarks. Baryons can thus be placed in irreducible representations of $\mathbf{3} \otimes \mathbf{3} \otimes \mathbf{3}$ of $SU(3)$.

As these particles are composed of three quarks, they can either have a total angular momentum $J = \frac{1}{2}$ or $J = \frac{3}{2}$. However there is an additional constraint that baryons composed of one flavour of quark can only exist in $J = \frac{3}{2}$ configuration. The $J = \frac{1}{2}$ configuration is forbidden by the Pauli exclusion principle as otherwise these two quarks, which are fermions, would necessarily have the same spin and therefore the exact same state. Baryons of two or three types of quarks can exist in either $J = \frac{1}{2}$ or $J = \frac{3}{2}$ configurations.

The light baryon states live in the $\mathbf{3} \otimes \mathbf{3} \otimes \mathbf{3} = \mathbf{10} \oplus \mathbf{8} \oplus \mathbf{8} \oplus \mathbf{1}$. The baryons are grouped in the decouplet $\mathbf{10}$, the octet, which is a linear combination of elements of both $\mathbf{8}$ and the singlet.

The baryon octet $B = 1$, $J = \frac{1}{2}$

There is a baryon octet with $B = 1, J = \frac{1}{2}$. The states correspond to the following particles: and comes with the singlet state:

$$\Sigma^{0\star} = uds \tag{A.33}$$

baryons	quark content	rest mass [MeV/c^2]
n	uud	939.565346 ± 0.000023
p	udd	938.272013 ± 0.000023
Σ^+	uus	1189.37 ± 0.07
Σ^0	uds	1192.642 ± 0.024
Σ^-	dds	1197.449 ± 0.030
Ξ^0	uss	1314.86 ± 0.20
Ξ^-	dss	1321.71 ± 0.07
Λ^0	uds	1115.683 ± 0.006

with mass $1387.7 \pm 1.0\,\mathrm{MeV}/c^2$.

The weight diagram for the octet is shown Fig. A.10.

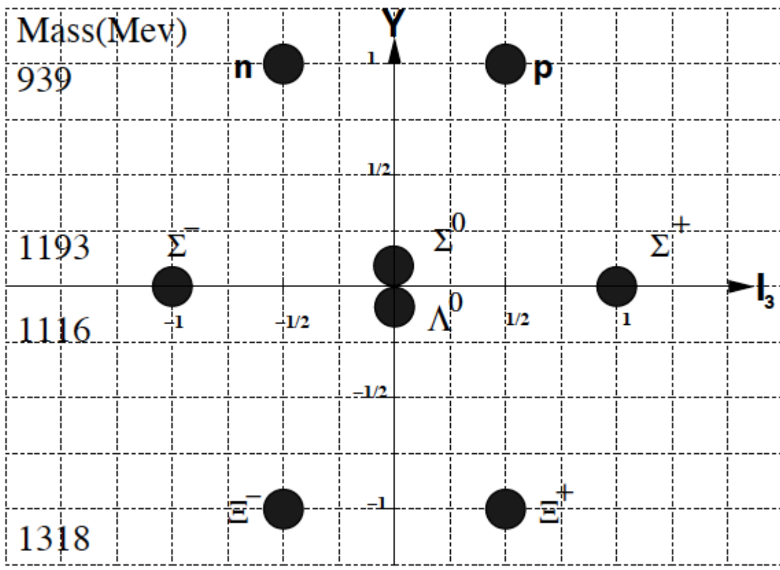


Figure A.10: Baryon octet with $B = 1$, $J = \frac{1}{2}$. [5]

The baryon decuplet $B = 1$, $J = \frac{3}{2}$

The states in the baryon decuplet **10** has $B = 1$,$J = \frac{3}{2}$ with states:

baryons	quark content	rest mass [MeV/c^2]
Δ^-	ddd	1232 ± 1
Δ^0	udd	1232 ± 1
Δ^+	uud	1232 ± 1
Δ^{++}	uuu	1232 ± 1
$\Sigma^{-\star}$	dds	1387.2 ± 0.5
$\Sigma^{0\star}$	uds	1383.7 ± 1.0
$\Sigma^{+\star}$	uus	1382.8 ± 0.4
$\Xi^{-\star}$	uss	1535.0 ± 0.6
$\Xi^{0\star}$	dss	1531.80 ± 0.32
Ω^-	sss	1672.45 ± 0.29

The baryon decuplet weight diagram is shown in Fig. A.11.

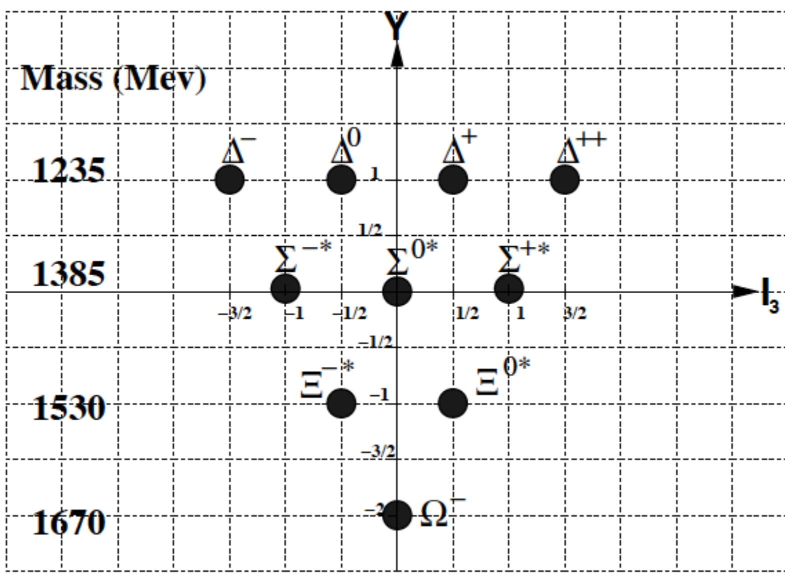


Figure A.11: Baryon decuplet with $B = 1$, $J = \frac{3}{2}$. [5]

There is also an antibaryon decouplet $(I_3, Y) \to -(I_3, Y)$ with $B = 1, J = \frac{1}{2}$, in the anti-fundamental representation.

Note that the $\Delta^{++} = u \otimes u \otimes u$ particle is the lowest mass state in the baryon decouplet. We assume therefore it has vanishing angular momentum $L = 0$ and thus spin $I = \frac{3}{2}$. This requires all the up type quarks to have up-type spin, which would violate the Pauli exclusion principle. The existence of this particle therefore implies the additional degree of freedom, colour.

A.4.3 Colour Symmetry

Colour is an additional internal degree of freedom that forms an exact $SU(3)$ symmetry (analogous to spin in $SU(2)$) responsible for strong nuclear interactions. Each quark has an associated colour charge, red, green or blue, with all observed baryons composed of colourless linear combinations. The wave functions that describe quark states contain this additional colour factor, which lies in the $\mathbf{3}$ of the $SU(3)_C$. Similarly, the antiquark states have anticolour that live in $\overline{\mathbf{3}}$.

Furthermore, it was noticed that the hadrons (any composite particle composed of quarks) are limited to the observed mesons and baryons. An ad hoc requirement called *confinement* is imposed to explain this observation. Confinement states that only particles that are colourless singlets are seen in nature. Other bound states, such as $qqqq$, $qq\bar{q}$ etc. are ruled out as they do not contain a singlet.

This also explains some observed discrepancies of factors of three between what is experimentally measured and what would otherwise have been predicted. For example, the ratio between $e^+e^- \to$ had and $e^+e^- \to \mu^+\mu^-$ would be a factor of three off in the theory.

A.4.4 Heavy Flavours

One would naïvely expect that this procedure could be extended to encompass all 6 observed quarks in the Standard Model, simply by extending the group from $SU(3)$ to $SU(6)$. Unfortunately, the heavy quarks c, t, and b are all significantly heavier than the others. This mass difference breaks the approximate symmetry that would otherwise relate these quarks. Further, the heavier quarks are inherently unstable, decaying promptly into the lighter making measurement of stable bound states difficult.

Whilst there is no symmetry governing the three heavy quarks, one can still construct mesons from one heavy and one light quark by viewing this meson as living in the $\mathbf{3}$ of $SU(3)$.

Similarly, baryons composed of two light quarks and one heavy quark can be thought of as living in a $\mathbf{3} \otimes \mathbf{3} = \mathbf{6} \oplus \mathbf{3}$ multiplets.

Other, heavier hadrons are theoretically predicted but are very difficult to detect due to their extremely short lifetime.

Appendix B

Lorentz Transformations

To define the Lorentz transformation ${\Lambda^{\nu}}_{\mu}$ on the space-time coordinate x^{μ}, we consider two inertial frames S and S', the latter moving at a velocity v away from S. This set of transformations govern the theory of Special Relativity and form the proper orthochronous Lorentz group $SO^{+}(1,3)$ as we have shown in Section 1.1.1. We shall construct the equations of Special Relativity in this section.

Without loss of generality, we can define (using the rotations) v to be in the positive z direction, with the origin of S ($t=0, z=0$) coinciding with the the origin of S' ($t'=0$,$z'=0$). We shall omit x, y indices in our discussion for the sake of brevity. We assume that the motion is linear, such that the transformations are also linear. We can then write $x'^{\mu} = {\Lambda^{\mu}}_{\nu} x^{\nu}$; explicitly,

$$\begin{pmatrix} t' \\ z' \end{pmatrix} = \begin{pmatrix} \lambda & \delta \\ \alpha & \gamma \end{pmatrix} \begin{pmatrix} t \\ z \end{pmatrix} \tag{B.1}$$

where $\lambda, \delta, \alpha, \gamma \in \mathbb{R}$ are currently arbitrary constants, such that we can choose $\lambda \geq 1$.

Consider the motion of the origin of frame S'. In S', we will be at the coordinates $(t', z'=0)$ for a general t'. However, in S we shall be at $(t, z=vt)$. These points are related by

$$\begin{pmatrix} t' \\ 0 \end{pmatrix} = \begin{pmatrix} \lambda & \delta \\ \alpha & \gamma \end{pmatrix} \begin{pmatrix} t \\ vt \end{pmatrix} \tag{B.2}$$

and thus $\alpha = -\gamma v$. Analogously, if we consider the position of the origin of S in the frame S', we have

$$\begin{pmatrix} t' \\ -vt' \end{pmatrix} = \begin{pmatrix} \lambda & \delta \\ \alpha & \gamma \end{pmatrix} \begin{pmatrix} t \\ 0 \end{pmatrix} \tag{B.3}$$

which in turn implies $-vt' = \alpha t$ and $t' = \lambda t$. Together with the previous result, we find $\alpha = -v\lambda$ and thus $\lambda = \gamma$.

Putting this together, we have

$$\begin{pmatrix} t' \\ z' \end{pmatrix} = \begin{pmatrix} \gamma & \delta \\ -v\gamma & \gamma \end{pmatrix} \begin{pmatrix} t \\ z \end{pmatrix} \tag{B.4}$$

The inverse transformation of the Lorentz group can be constructed as

$$\begin{pmatrix} t \\ z \end{pmatrix} = \frac{1}{\gamma^2 + v\delta\gamma} \begin{pmatrix} \gamma & -\delta \\ v\gamma & \gamma \end{pmatrix} \begin{pmatrix} t' \\ z' \end{pmatrix} \tag{B.5}$$

We shall now make use of the equivalence principle, which states that there are no preferred reference frames. Consider S as moving at $-v$ away from S', we then have

$$\begin{pmatrix} t \\ z \end{pmatrix} = \begin{pmatrix} \gamma(-v) & \delta(-v) \\ v\gamma(-v) & \gamma(-v) \end{pmatrix} \begin{pmatrix} t' \\ z' \end{pmatrix} \tag{B.6}$$

If we assume that we are living in an isotropic world, then it is clear that γ cannot be dependent on the direction of v and $\gamma(-v) = \gamma(v)$. Now, we can make the main diagonals of (B.5) and (B.6) consistent by requiring $\gamma^2 + v\delta\gamma = 1$. In this case, γ defines both the relativistic length contraction and time dilation (if $\gamma < 1$) of the moving frame when viewed from the stationary one. Furthermore, notice that this condition implies that $\det \Lambda = +1$.

Performing two consecutive transformations gives

$$\begin{aligned}\begin{pmatrix} t'' \\ z'' \end{pmatrix} &= \begin{pmatrix} \gamma(v') & \delta(v') \\ -v'\gamma(v') & \gamma(v') \end{pmatrix} \begin{pmatrix} \gamma(v) & \delta(v) \\ -v\gamma(v) & \gamma(v) \end{pmatrix} \begin{pmatrix} t \\ z \end{pmatrix} \\ &= \begin{pmatrix} \gamma'\gamma - \delta' v\gamma & \gamma'\delta + \delta'\gamma \\ -v'\gamma'\gamma - \gamma' v\gamma & -v'\gamma'\delta + \gamma'\gamma \end{pmatrix} \begin{pmatrix} t \\ z \end{pmatrix} \end{aligned} \tag{B.7}$$

Notice that this is another (Lorentz) transformation and therefore satisfies the group axiom of closure.

In the original transformation matrix, the main diagonal elements are both equal to γ and therefore

$$\gamma'\gamma - v\delta'\gamma = -v'\gamma'\delta + \gamma'\gamma \tag{B.8}$$

which requires $\frac{\delta'}{v\gamma'} = \frac{\delta}{v\gamma}$. For a general Lorentz transformation with $v \neq 0 \Rightarrow \gamma \neq 1$, this condition is consistent with $\gamma^2 + v\delta\gamma = 1$. The case $v = 0$ is the identity transformation of the Lorentz group and can be recovered from the general case at the end of the derivation.

For $v \neq 0$, we thus have a frame-invariant quantity

$$\kappa \equiv \frac{\delta'}{v'\gamma'} = \frac{\delta}{v\gamma} \tag{B.9}$$

which has dimensions $[\kappa] = [v]^{-2}$ (since γ is dimensionless $[t] = [\gamma t]$, $[\delta] = [t]/[z] = [v]^{-1}$).

Then $1 = \gamma^2 + v\delta\gamma = \gamma^2(1 + \kappa v^2)$ and thus

$$\gamma = +\frac{1}{\sqrt{1 + \kappa v^2}} \tag{B.10}$$

where we have taken the positive root since we wish to considering the proper orthochronous Lorentz subgroup.

Finally, the transformation can be written

$$\Lambda = \frac{1}{\sqrt{1 + \kappa v^2}} \begin{pmatrix} 1 & \kappa v \\ -v & 1 \end{pmatrix} \tag{B.11}$$

and

$$\begin{pmatrix} t' \\ z' \end{pmatrix} = \frac{1}{\sqrt{1+\kappa v^2}} \begin{pmatrix} 1 & \kappa v \\ -v & 1 \end{pmatrix} \begin{pmatrix} t \\ z \end{pmatrix} \tag{B.12}$$

If $\kappa > 0$, in the limit $\kappa v^2 \gg 1$ the spatial coordinates transform with a time coordinates and vice-versa since $t' \to \sqrt{\kappa} z$ and $z \to (\sqrt{\kappa})^{-1} t$. We exclude this possibility as unphysical, since we are taking the proper orthochronous subgroup where time is taken to run only in the positive direction. Therefore, either $\kappa = 0$ or $\kappa < 0$.

Taking $\kappa = 0$, we recover Galilean/Newtonian kinematics and transformation laws:

$$\begin{pmatrix} t' \\ z' \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ -v & 1 \end{pmatrix} \begin{pmatrix} t \\ z \end{pmatrix} \tag{B.13}$$

In this case, the relative velocities of inertial frames are not limited and since $t' = t$, time is absolute.

The final case $\kappa < 0$ corresponds to a general Lorentz transformation. We set the (finite) frame invariant speed of light in a vacuum to be

$$c = \frac{1}{\sqrt{-\kappa}} \tag{B.14}$$

or equivalently $\kappa = -\frac{1}{c^2} = -\frac{1}{\mu_0 \epsilon_0}$, where $c = 299,792,458\ \mathrm{ms}^{-1}$ is the speed of light, which sets an upper limit on physical speeds [217]. c is thus treated as a universal constant and defines the highest possible relative velocity between inertial frames. Further, Michelson-Morley showed that the speed of light is an absolute and finite speed [218] such that κ is constant. Historically, Einstein used a slightly stronger assumption that the speed of light was constant to derive Special Relativity [219], whilst we have just used equivalence of inertial frames in our derivation.

If $v \ll c$, we recover the set of Galilean transformation as required. The case $v = 0$ occurs when S' and S are identical and the transformation is merely the trivial identity transformation.

We have finally arrived at the Lorentz transformations that govern Einstein's theory of Special Relativity

$$\begin{pmatrix} t' \\ z' \end{pmatrix} = \frac{1}{\sqrt{1-\frac{v^2}{c^2}}} \begin{pmatrix} 1 & -\frac{v^2}{c^2} \\ -v & 1 \end{pmatrix} \begin{pmatrix} t \\ z \end{pmatrix} \tag{B.15}$$

Notice that the Lorentz transformations do indeed form a group as they satisfy the group axioms of closure (B.7), the existence of the identity (the $v = 0$ case) and inverse (B.5), and finally associativity follows directly from matrix multiplication being associative.

The matrix representation of this group governs one time and three spatial dimensions, where each group element $\Lambda \in SO^+(1,3)$ has $\det \Lambda = +1$ [54], where the $+$ indicates that the transformation is the (proper orthochronous) subgroup that preserves the direction of time. In particular, the equivalence of reference frames in flat Minkowski space implies that its metric $\eta^{\mu\nu}$ must be invariant under the Lorentz transformations, such that ${\Lambda^\mu}_\sigma \eta_{\mu\nu} {\Lambda^\nu}_\rho = \eta_{\sigma\rho}$.

B.1 Lorentz Transformations on Scalar Fields

We would like the Schrödinger equation of Quantum Mechanics (QM) and thus the Klein-Gordon Lagrangian to be relativistic. Indeed, the unification of Special Relativity (described by the Lorentz transformations) and QM was one of the main motivations for developing a Quantum Field Theory (QFT) [6].

The Lorentz transformation of the space-time coordinates is given by:

$$x^\mu \to \left(x'\right)^\mu = {\Lambda^\mu}_\nu x^\nu \tag{B.16}$$

To be consistent with the transformation leaving the path length invariant, we require the metric $\eta_{\mu\nu}$ to transform as

$${\Lambda^\mu}_\sigma \eta_{\mu\nu} {\Lambda^\nu}_\tau = \eta_{\sigma\tau} \tag{B.17}$$

For the scalar field $\phi(x)$, the Lorentz transformation is thus

$$\phi(x) \longrightarrow \phi'(x) = \phi(\Lambda^{-1}x) \tag{B.18}$$

where we have used the inverse Λ^{-1} to indicate an "active" transformation, which shifts the field by expressing the old field configuration in terms of the transformed frame of reference. In contrast, the "passive" transformation is defined with a Λ and represents only a change of coordinates [6].

For Lorentz invariance to hold, we require that if ϕ satisfies the equations of motion then $\phi(\Lambda^{-1}x)$ must also. Further, the derivative transforms as a vector

$$\partial_\mu \longrightarrow \frac{\partial}{\partial x'^\mu} = \frac{\partial x^\nu}{\partial x'^\mu}\frac{\partial}{\partial x^\nu} = (\Lambda^{-1})^\nu{}_\mu \partial_\nu \tag{B.19}$$

In the particular case of the Klein-Gordon Lagrangian, the derivative term becomes

$$\begin{aligned}\mathcal{L}_{\mathrm{KG,kin}}(x) = \partial_\mu\phi(x)\partial_\nu\phi(x)\eta^{\mu\nu} \quad &\longrightarrow \quad \left(\Lambda^{-1}\right)^\rho{}_\mu \left(\partial_\rho\phi\right)(y) \left(\Lambda^{-1}\right)^\sigma{}_\nu \left(\partial_\sigma\phi\right)(y)\eta^{\mu\nu} \\ &= \quad \left(\partial_\rho\phi\right)(y)\left(\partial_\sigma\phi\right)(y)\eta^{\rho\sigma} \\ &= \quad \mathcal{L}_{\mathrm{KG,kin}}(y)\end{aligned} \tag{B.20}$$

where we have defined $y \equiv \Lambda^{-1}x$.

Together with the potential term, which transforms as $\phi^2(x) \to \phi^2(y)$, the Klein-Gordon action

$$S_{\mathrm{KG}} = \int d^4x\, \mathcal{L}(x) \longrightarrow \int d^4x\, \mathcal{L}(y) = \int d^4y\, \mathcal{L}(y) \tag{B.21}$$

is thus Lorentz invariant, where the additional Jacobian in the last equality is assumed to be unity ($\det\Lambda = 1$).

B.2 Lorentz Transformations on Spinor Fields

The free space Dirac equation on spinors

$$\left(i\gamma^\mu\partial_\mu - m\right)\psi = 0 \tag{B.22}$$

in the Weyl representation of γ^μ matrices, with $\psi = \begin{pmatrix}\psi_L \\ \psi_R\end{pmatrix}$.

This is the equation of motion of spinor fields, which we would like to use to describe the spin-$\frac{1}{2}$ fermions. For this to be a physical theory, we require Lorentz invariance, which can be imposed in two ways:

- view $\gamma^\mu \partial_\mu$ as Lorentz scalar. Since ∂_μ is a covariant vector, this is equivalent to requiring γ^μ to transform contravariantly under the Lorentz transformation. In this case ψ is considered a scalar.

- assume γ^μ to be fixed, and have ψ transform in some non-trivial way under the Lorentz group, to counteract the transformation of ∂_μ.

In 1955 R. H. Good showed that any choice of γ^μ are equivalent up to some unitary transformation [220] and thus the γ matrices can be fixed for any frames of reference.

We therefore proceed with the latter case where $\psi^\alpha(x) \to S[\Lambda]^\alpha{}_\beta \psi(\Lambda^{-1}x)$, with $S[\Lambda]$ some 4×4 matrices to be determined, where $\alpha, \beta = 1, 2, 3, 4$ run over the indices of the matrix.

Since we wish $S[\Lambda]$ to form a representation of the Lorentz group, the inverse transformation must exist $S[\Lambda^{-1}] = \psi'(\Lambda^{-1}x)S[\Lambda]^{-1}\psi'(x) = \psi(\Lambda^{-1}x)$ such that $S[\Lambda]S[\Lambda]^{-1} = \mathbb{I}$. As in the scalar field case Λ^{-1} is required as we are considering active transformations of the field. For convenience, we shall define $y \equiv \Lambda^{-1}x$.

In particular, we would like the Dirac equation to be invariant under this representation of the Lorentz group, such that if $\psi(x)$ satisfies the equation of motion, then so will $S[\Lambda]\psi(y)$.

We start from the Dirac equation in the transformed reference frame:

$$\left(i\gamma^\mu \partial'_\mu - m\right)\psi'(x) = 0 \tag{B.23}$$

The partial derivative transforms as a vector with $\left(\partial_\mu \phi\right)(x) \to \left(\Lambda^{-1}\right)^\nu{}_\mu \left(\partial_\nu \phi\right)(y)$.

From (B.23), we proceed as follows

$$\begin{aligned} i\gamma^\mu \Lambda_\mu{}^\nu \partial_\nu S[\Lambda]\psi(y) &= mS[\Lambda]\psi(y) \\ &= iS\gamma^\nu \partial_\nu \psi \end{aligned} \tag{B.24}$$

where the last line follows from the Dirac equation (B.22).

We thus have the condition $S[\Lambda]^{-1}\gamma^{\mu}\Lambda_{\mu}^{\ \nu}S[\Lambda] = \gamma^{\nu}$, for the spinor field to have to correct transformation property. Note that since $[S[\Lambda], \partial_{\nu}] = 0 = [\Lambda, S[\Lambda]] = 0$, this condition becomes

$$S[\Lambda]^{-1}\gamma^{\mu}S[\Lambda] = \Lambda^{\mu}_{\ \nu}\gamma^{\nu} \tag{B.25}$$

We make the ansatz that $S[\Lambda] = \exp\left(k\Omega_{\mu\nu}\mathcal{S}^{\mu\nu}\right)$ where

$$\mathcal{S}^{\mu\nu} \equiv \frac{1}{4}\left[\gamma^{\mu}, \gamma^{\nu}\right] = \frac{1}{2}\gamma^{\mu}\gamma^{\nu} - \frac{1}{2}\eta^{\mu\nu} \tag{B.26}$$

The matrices $\mathcal{S}^{\mu\nu}$ are themselves generators and form a representation of the Lorentz algebra in (1.13) as they obey the same commutation relations:

$$[\mathcal{S}^{\rho\sigma}, \mathcal{S}^{\tau\nu}] = \eta^{\sigma\tau}\mathcal{S}^{\rho\nu} - \eta^{\rho\tau}\mathcal{S}^{\sigma\nu} + \eta^{\rho\nu}\mathcal{S}^{\sigma\tau} - \eta^{\sigma\nu}\mathcal{S}^{\rho\tau} \tag{B.27}$$

In the infinitesimal limit,

$$S[\Lambda]^{\alpha}_{\ \beta} \approxeq \left(1 + k\Omega_{\mu\nu}\mathcal{S}^{\mu\nu}\right)^{\alpha}_{\ \beta} \tag{B.28a}$$

$$\left(S[\Lambda]^{-1}\right)^{\alpha}_{\ \beta} \approxeq \left(1 - k\Omega_{\mu\nu}\mathcal{S}^{\mu\nu}\right)^{\alpha}_{\ \beta} \tag{B.28b}$$

$$\Lambda^{\alpha}_{\ \beta} \approxeq \left(1 + \frac{1}{2}\Omega_{\mu\nu}\mathcal{M}^{\mu\nu}\right)^{\alpha}_{\ \beta} \tag{B.28c}$$

where the last line uses the generators of the Lorentz group defined in (1.14). Although the basis generators $\mathcal{S}^{\mu\nu}$ and $\mathcal{M}^{\mu\nu}$ are different, we shall use the same six numbers $\Omega_{\mu\nu}$ to ensure the same transformation is done on both the coordinate x and the field ψ. We require

$$\left(1 - k\Omega_{\alpha\beta}\mathcal{S}^{\alpha\beta}\right)\gamma^{\mu}\left(1 + k\Omega_{\alpha\beta}\mathcal{S}^{\alpha\beta}\right) = \frac{1}{2}\left(1 + \frac{1}{2}\Omega_{\alpha\beta}\mathcal{M}^{\alpha\beta}\right)\gamma^{\mu} \tag{B.29}$$

where we have suppressed the matrix indices. Rearranging this equation, we find

$$-k\Omega_{\alpha\beta}\left[S^{\alpha\beta}, \gamma^{\mu}\right] = \frac{1}{2}\left(\mathcal{M}^{\alpha\beta}\right)\gamma^{\mu} \tag{B.30}$$

where $[\mathcal{S}^{\mu\nu}, \gamma^\rho] = \gamma^\mu \eta^{\nu\rho} - \gamma^\nu \eta^{\rho\mu}$; this result follows directly from the definition of $S^{\mu\nu}$ and the anticommutation rules for $\mu \neq \nu$ in (2.30).

Finally, since

$$\begin{aligned} (\mathcal{M}^{\rho\sigma})^\mu{}_\nu \, \gamma^\nu &= (\eta^{\rho\mu}\delta^\sigma{}_\nu - \eta^{\sigma\mu}\delta^\rho{}_\nu) \, \gamma^\nu \\ &= \eta^{\rho\mu}\gamma^\sigma - \eta^{\sigma\mu}\gamma^\rho \end{aligned} \tag{B.31}$$

where the first line is just the definition of $\mathcal{M}$ in (1.11), we see that $k = \frac{1}{2}$ and

$$S[\Lambda] = 1 + \frac{1}{2}\Omega_{\mu\nu}\mathcal{S}^{\mu\nu} \tag{B.32}$$

Thus, the total transformation on the field ψ, found by repeated application of the infinitesimal one, is given by

$$S[\Lambda] = \lim_{n\to\infty}\left(1 + \frac{1}{2}\frac{\Omega_{\mu\nu}}{n}\mathcal{S}^{\mu\nu}\right)^n = \exp\left(\frac{1}{2}\Omega_{\mu\nu}\mathcal{S}^{\mu\nu}\right) \tag{B.33}$$

This is indeed a unique representation and the spinor fields ψ are defined by this (spinor) representation's transformation rule. In particular, we require a 4π rotation of Ω for the spinor field to return to the original state, rather than the usual 2π in the case of the scalar field.

Additionally, it is possible to put $S[\Lambda]$ back into block-diagonal form by:

$$\begin{aligned} \frac{1}{2}\Omega_{\mu\nu}\mathcal{S}^{\mu\nu} &= \frac{1}{4}\Omega_{\mu\nu}\gamma^\mu\gamma^\nu \\ &= \frac{1}{2}\left(\Omega_{0k}\gamma^0\gamma^k + \gamma_{12}\gamma^1\gamma^2 + \Omega_{13}\gamma^1\gamma^3 + \Omega_{23}\gamma^2\gamma^3\right) \\ &= \frac{1}{2}\phi_k \begin{pmatrix} \sigma^k & 0 \\ 0 & -\sigma^k \end{pmatrix} + \frac{1}{2}i\theta_k \begin{pmatrix} \sigma^k & 0 \\ 0 & \sigma^k \end{pmatrix} \end{aligned} \tag{B.34}$$

where we have used the definitions of γ^μ from (2.33) and ϕ_k, θ_k correspond to the Lorentz boosts and rotations respectively.

This gives the block diagonal form of S as

$$S[\Lambda] = \exp\begin{pmatrix} \frac{1}{2}\phi_k\sigma^k + \frac{1}{2}i\theta_k\sigma^k & 0 \\ 0 & -\frac{1}{2}\phi_k\sigma^k + \frac{i}{2}\theta_k\sigma^k \end{pmatrix} \tag{B.35}$$

Notice that this separates out two inequivalent (due to the additional negative sign in the second block) irreducible representations.

Explicitly, taking $S[\Lambda]$ to act on $\psi = \begin{pmatrix} \psi_L \\ \psi_R \end{pmatrix}$, we find

$$\begin{aligned} \psi'_L &= \exp\left(\frac{1}{2}i\sigma\cdot\theta + \frac{1}{2}\sigma\cdot\phi\right)\psi_L \\ \psi'_R &= \exp\left(\frac{1}{2}i\sigma\cdot\theta - \frac{1}{2}\sigma\cdot\phi\right)\psi_R \end{aligned} \tag{B.36}$$

with ψ_L, ψ_R irreducible "Weyl" spinors. These are related by the Hermition conjugate, such that $\psi_L^\dagger$ transforms as a right-handed spinor and $\psi_R^\dagger$ as a left-handed one.

Notice that the Weyl spinors do not preserve parity $x \to -x$, since the left-handed spinor transforms into the right-handed one. This is in direct contrast to the Dirac spinors, which contain both left and right handed components and are full parity conserving. Weyl spinors are therefore used in the theories such as the electroweak theory where parity symmetry has been observed to be violated [221].

By construction, the Dirac equation (B.22) is now Lorentz invariant for the field ψ, which transforms under $\psi \to S[\Lambda]\psi$. However, we cannot simply write down the term $\psi^\dagger\psi$ in the Lagrangian since the matrices $S[\Lambda]$ are in general not unitary (in particular the boosts of (B.34)) [6].

This means that the term

$$\psi^\dagger\psi \longrightarrow \psi^\dagger S[\Lambda]^\dagger S[\Lambda]\psi \neq \psi^\dagger\psi \tag{B.37}$$

is not Lorentz invariant. Instead, we construct the term

$$\overline{\psi}\psi \tag{B.38}$$

where we define $\overline{\psi} = \psi^\dagger \gamma^0$ as the "Dirac adjoint".

It is simple to check that such a term is a Lorentz scalar, and therefore permissible in equations that govern our physical theories. Following [6], we start by noting that $(\gamma^0)^\dagger = \gamma^0$ and $(\gamma^i)^\dagger = -\gamma^i$ for $i = 1, 2, 3$. Then,

$$\gamma^0 \gamma^\mu \gamma^0 = (\gamma^\mu)^\dagger \tag{B.39}$$

for $\mu = 0, 1, 2, 3$. Thus,

$$(S^{\mu\nu})^\dagger = \frac{1}{4}\left[(\gamma^\nu)^\dagger, (\gamma^\mu)^\dagger\right] = -\gamma^0 S^{\mu\nu} \gamma^0 \tag{B.40}$$

We see that

$$S[\Lambda]^\dagger = \exp\left(\frac{1}{2}\Omega_{\rho\sigma}\left(\mathcal{S}^{\rho\sigma}\right)^\dagger\right) = \gamma^0 S[\Lambda]^{-1} \gamma^0 \tag{B.41}$$

Then,

$$\begin{aligned}
\overline{\psi}(x)\psi(x) &= \psi^\dagger(x)\gamma^0\psi(x) \\
&\rightarrow \psi^\dagger\left(\Lambda^{-1}x\right) S[\Lambda]^\dagger \gamma^0 S[\Lambda] \psi\left(\Lambda^{-1}x\right) \\
&= \psi^\dagger\left(\Lambda^{-1}x\right)\gamma^0\psi\left(\Lambda^{-1}x\right) \\
&= \overline{\psi}(\Lambda^{-1}x)\psi(\Lambda^{-1}x)
\end{aligned} \tag{B.42}$$

as required for a Lorentz scalar.

In a similar manner, more complicated terms involving multiple spinor fields and γ matrices can also be constructed. It is then a simple exercise to use the infinitesimal Lorentz transformation rules to determine how these new quantities behave [8]. We reiterate that only combinations of fields that are collectively Lorentz scalars (invariant under Lorentz transformations) are permissible as terms in the physical Lagrangians that describe nature.

We shall check the transformation rules of the term $\bar{\psi}\gamma^\mu\psi$. This term transforms as

$$\overline{\psi}(x)\gamma^\mu\psi(x) \longrightarrow \Lambda^\mu{}_\nu \overline{\psi}\left(\Lambda^{-1}x\right)\gamma^\nu\psi\left(\Lambda^{-1}x\right) \tag{B.43}$$

Notice that we have already shown that $S[\Lambda]^{-1}\gamma^\mu S[\Lambda] = \Lambda^\mu{}_\nu\gamma^\nu$ in (B.25), and so $\bar{\psi}\gamma^\mu\psi$ transforms as a vector

$$\bar{\psi}\gamma^\mu\psi \to \Lambda^\mu{}_\nu\bar{\psi}\left(\Lambda^{-1}x\right)\gamma^\nu\psi\left(\Lambda^{-1}x\right) \tag{B.44}$$

Now, since $\bar{\psi}\partial_\mu\gamma^\mu\psi$ is a scalar (since $\partial_\mu \to (\Lambda^{-1})^\nu{}_\mu\partial_\nu$ from (B.19)), we can construct the full Lorentz invariant Lagrangian, which gives rise to the Dirac equation:

$$\mathcal{L}_{\text{Dirac}} = \overline{\psi}\left(i\gamma^\mu\partial_\mu - m\right)\psi \tag{B.45}$$

The factor i is included to make the action real; upon complex conjugation, it cancels the additional minus sign that comes from integration by parts. We have now constructed the Lorentz invariant Dirac Lagrangian, which describes spin-$\frac{1}{2}$ particles (and their antiparticles) of mass m.

B.3 Lorentz Transformations on Vector Fields

The Lorentz transformation of vector fields is straight forward since it is similar to the transformation of the space-time coordinates. In particular, the active transformation of the vector field A^μ is

$$A^\mu(x) \longrightarrow \frac{\partial x'^\mu}{\partial x^\nu}A^\nu(\Lambda^{-1}x) = \Lambda^\mu{}_\nu A^\nu(\Lambda^{-1}x) \tag{B.46}$$

These types of fields will typically represent the bosons of the Standard Model and arise in equations of the form

$$\mathcal{L}_{\text{Dirac+vector}} = \overline{\psi}\left(\gamma^\mu\left(i\partial_\mu - eA_\mu\right) - m\right)\psi \tag{B.47}$$

with e a constant. Since the A_μ transforms in a similar manner to the derivation ∂_μ, it trivially follows from the Lorentz invariance of (B.45) that this Lagrangian is also Lorentz invariant.

Similarly, one can have arbitrary tensor product fields, where each index transforms as a vector-like one. For example,

$$B^{\mu\nu} \to \Lambda^{\mu}{}_{\rho}\Lambda^{\nu}{}_{\sigma}B^{\rho\sigma}(\Lambda) \tag{B.48}$$

The stress-energy tensor and the metric tensor are two examples of tensor product fields. These are required to have this well-defined transformation law to allow them to couple to all other fields in a Lorentz invariant manner. These coupling terms can then be included into the physical equations of motion.

Appendix C

The Search for a Heavy Top-Like Quark : Additional Material

C.1 Sensitivity Cross Check : Additional ≥ 1 b-tag Requirement

We separate out the events with 0 b-tags, dominated by background, and events with ≥ 1 b-tag, dominated by signal. Using the nominal `BDT` training, we find the `BDT` distributions shown in Fig. C.1 for the two categories of events with $= 0$ b-tags and with ≥ 1 b-tags.

We compare the standard procedure to one that includes an additional cut on the b-tag multiplicity in the event (using a minimal set of systematic uncertainties). Using only events with ≥ 1 b-tags, we find a comparable expected limit to the situation using standard selection in the combined lepton-plus-jets channel. This is shown in Fig. C.2. The `BDT` therefore accounts for background dominated 0 b-tag events and gives no significant impact on sensitivity.

This test is done using a simplified `BDT` training with all non-substructure variables. An asymptotic $\mathrm{CL_S}$ calculator described in [156, 222] was used and the sources of systematic uncertainties included in this test were all non-template systematics and the JES shape uncertainty.

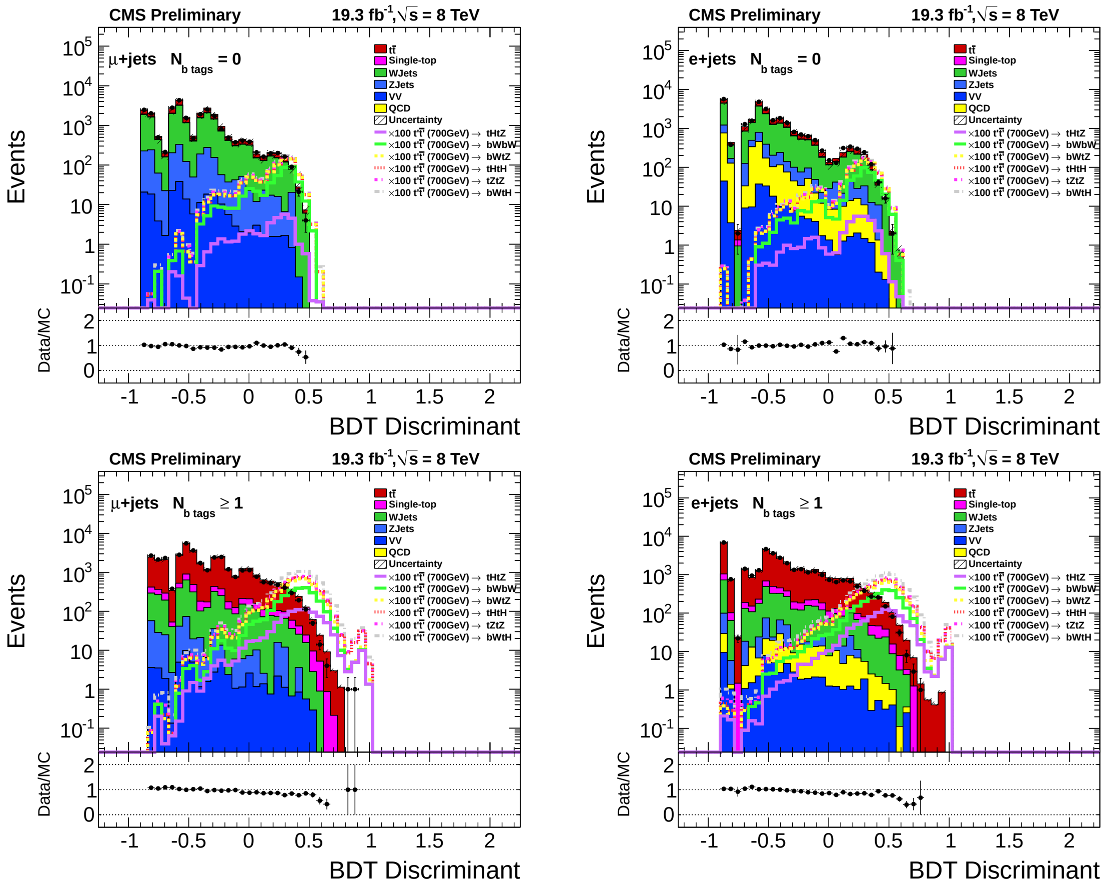


Figure C.1: Top (Bottom) shows the 700 GeV BDT distribution in the $= 0$ (≥ 1) b-tag regime for both channels.

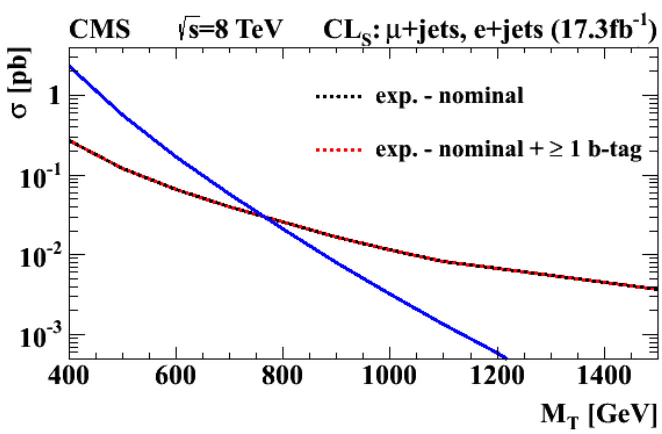


Figure C.2: Expected limit of (black) standard selection and (red) standard + ≥ 1 b-tag selection.

C.2 Sensitivity Cross Check : H_T vs number of b-tagged jets

We check the limits in this nominal BR case against a "linearised" H_T vs number of b-tagged jets distribution. We split the H_T distribution into four categories of events with different b-tag multiplicities $(0, 1, 2, \geq 3)$ and then linearise them. The distributions that are after bin merging (requiring all bins to be well populated, with a maximum of 40% statistical uncertainty by the MC) are shown in Fig. C.3.

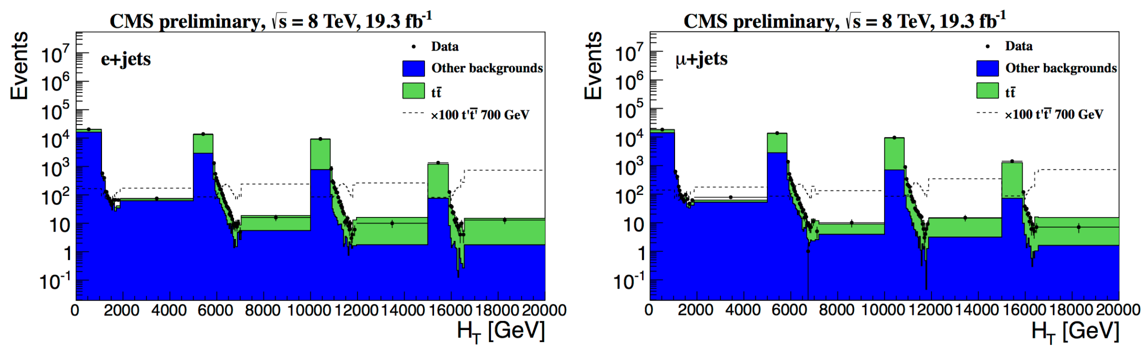


Figure C.3: Linearised H_T vs b-tag multiplicity after bin merging (to ensure statistics in all MC templates). The bins are merged right to left until there is less than a 40% statistical uncertainty in the bin. Note that the H_T range is 5 TeV per b-tag category such that $0 - 5$ TeV corresponds to 0 b-tags, $5 - 10$TeV is 1 b-tag, $10 - 15$ TeV is 2 b-tags and $15 - 20$ TeV is ≥ 3 b-tags, where each category spans $0 - 5$ TeV in H_T .

These templates are fed into the statistical procedure along with a minimal set of (JES and all non-template) uncertainties and the expected limit is compared to the nominal `BDT` procedure in Fig. C.4. We see good improvement over the entire mass range with the `BDT` prescription.

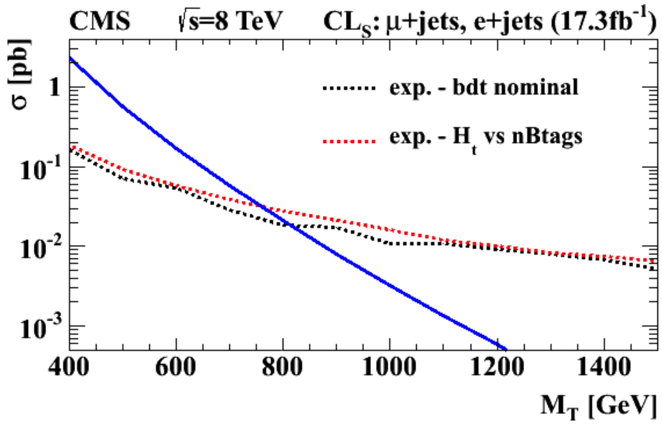


Figure C.4: The expected limits of the nominal `BDT` procedure is compared to the linearised H_T vs b-tag multiplicity after bin merging.

C.3 Signal `BDT` Distributions

`BDT` distributions, using nominal BR, for the signal samples with 500, 800, 1100, and 1400 GeV mass hypotheses. This shows the expected behaviour, with heavier masses appearing more signal like, where a `BDT` discriminant of +1 indicates a pure signal-like event and −1 a pure background-like event. The signal `BDT` distributions are shown in Fig. C.5 and Fig. C.6 for μ+jets events with and without substructure and Fig. C.7 and Fig. C.8 for e+jets events with and without substructure.

All decay modes at each mass point overlap, indicating that the training on different branching ratios separately will be unlikely to yield significant improvement over the nominal case we have considered. Therefore, we train only on the nominal branching ratios to avoid excessive computation time.

As a cross-check a comparison is made between training on the individual signal channels and the nominal branching ratios in Appendix C.5, and indeed the expected sensitivity is very comparable to the nominal training. Furthermore, the double-hump nature can be seen to be due to the simultaneous training on multiple signal samples Appendix C.4.

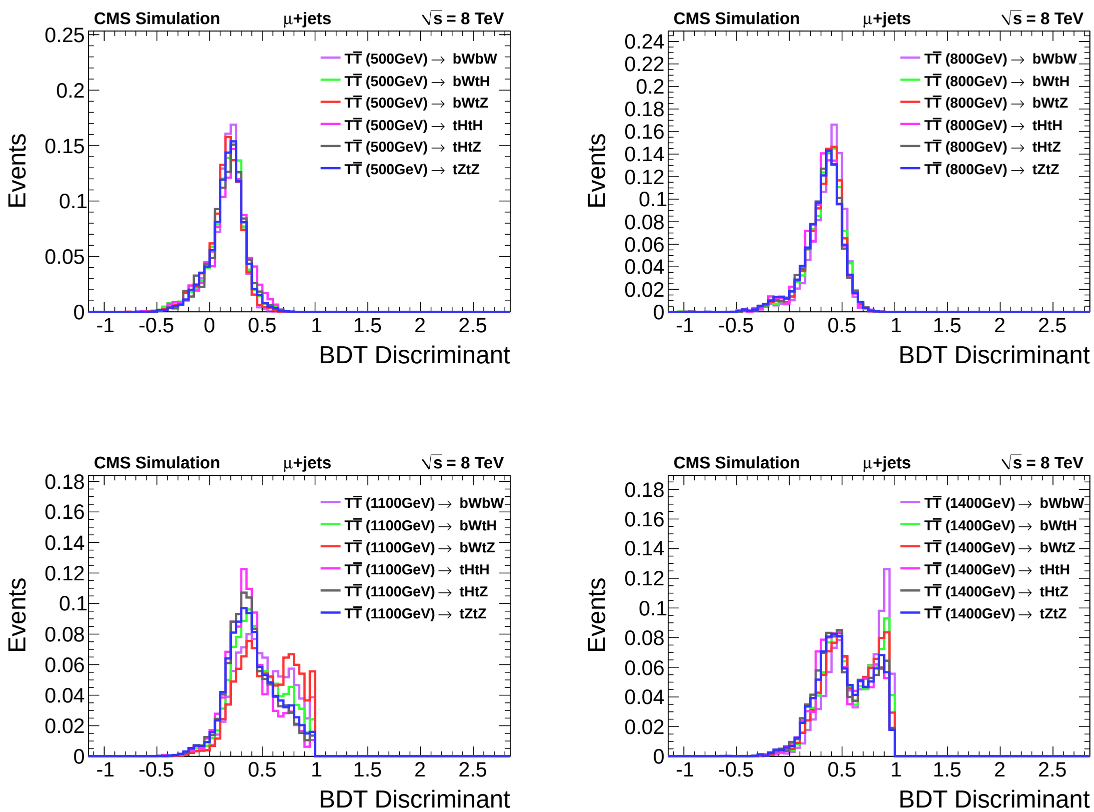


Figure C.5: Area normalised signal BDT for the 500, 800, 1100, 1400 GeV mass points for the μ+jets channel events with substructure corresponding to events with ≥ 3 jets and at least one W-tagged jet.

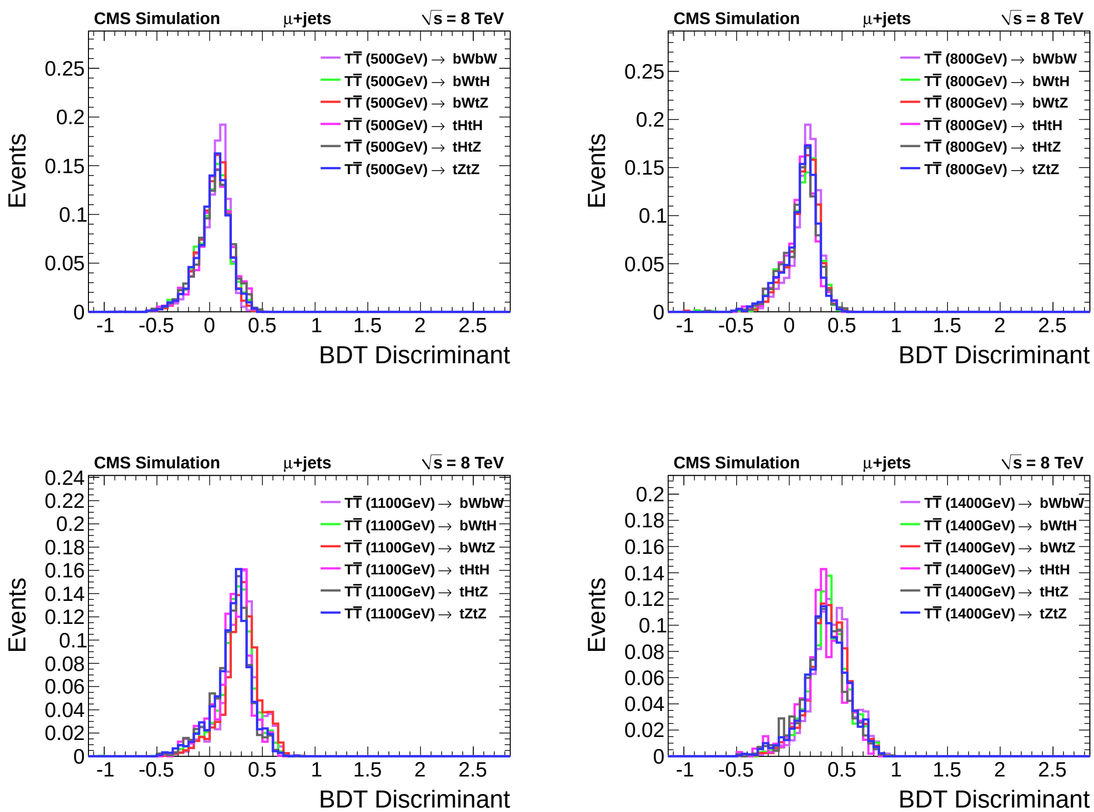


Figure C.6: Area normalised signal BDT for the 500, 800, 1100, 1400 GeV mass points in the μ+jets channel events without substructure corresponding to events with ≥ 4 jets and no W-tagged jet.

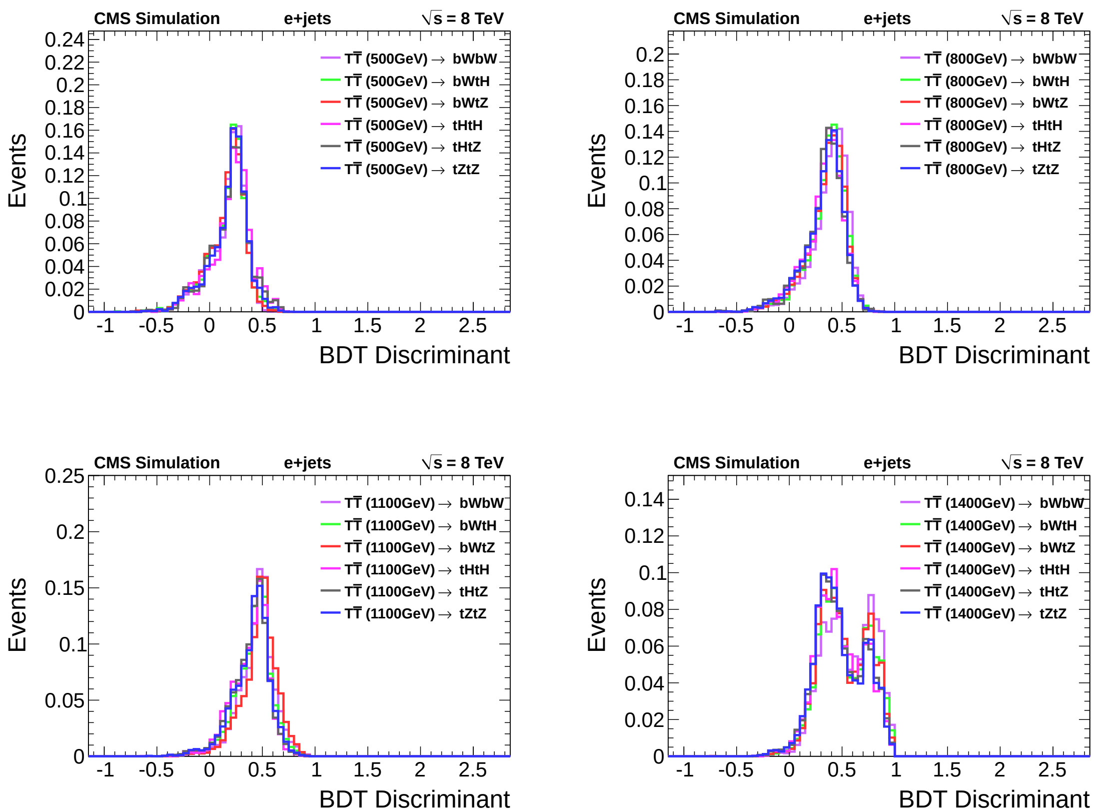


Figure C.7: Area normalised signal BDT for the 500, 800, 1100, 1400 GeV mass points in the e+jets channel events with substructure corresponding to events with ≥ 3 jets and at least one W-tagged jet.

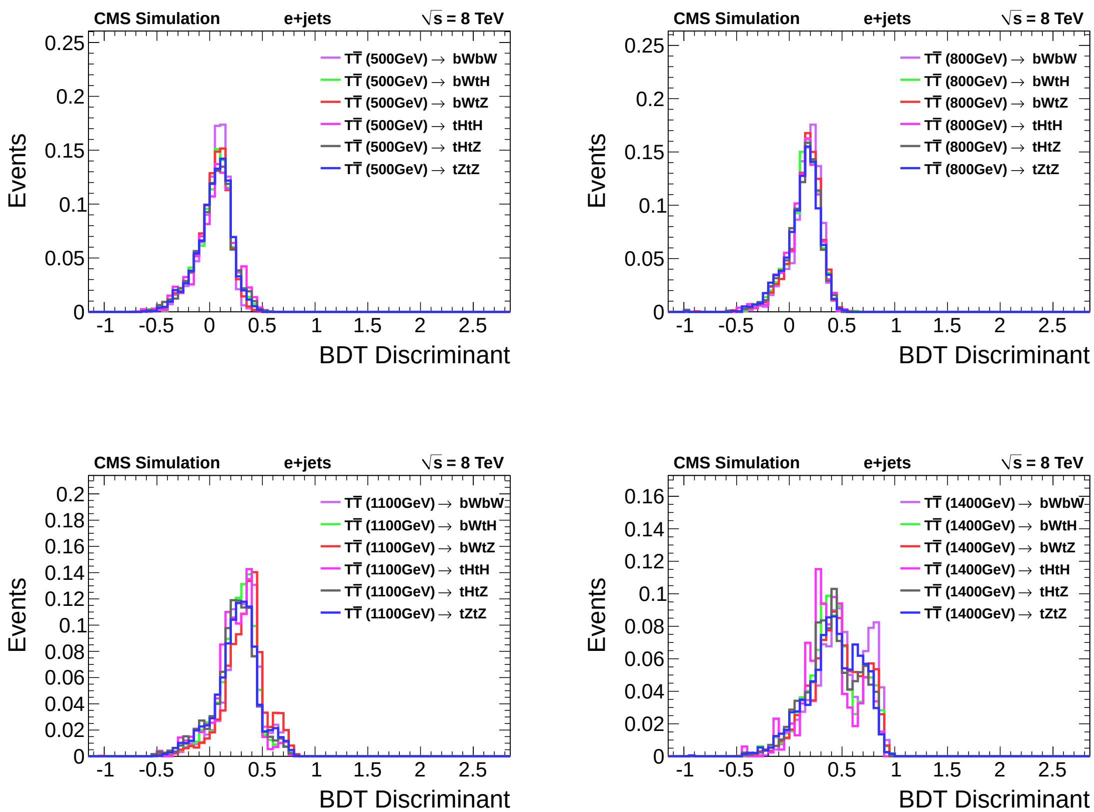


Figure C.8: Area normalised signal BDT for the 500, 800, 1100, 1400 GeV mass points in the e+jets channel events without substructure corresponding to events with ≥ 4 jets and no W-tagged jet.

C.4 BDT Distributions Training on Individual Signal Samples

We train on the 100% branching ratio modes separately, where both T-quarks decay via $T \to bW$, $T \to tH$, or $T \to tZ$ exclusively. The resulting BDT distributions are shown in Fig. C.9, Fig. C.10, and Fig. C.11 for $T \to bW$, $T \to tH$ and $T \to tZ$ respectively. As one can see, there is no longer a distinct second hump as exhibited when training using the nominal BRs. We conclude that the two bump structure seen in Fig. 11.32 are therefore due to the simultaneous training on the different signal samples.

The limit computation using these templates is shown in Appendix C.5. No significant improvement is seen in the limits.

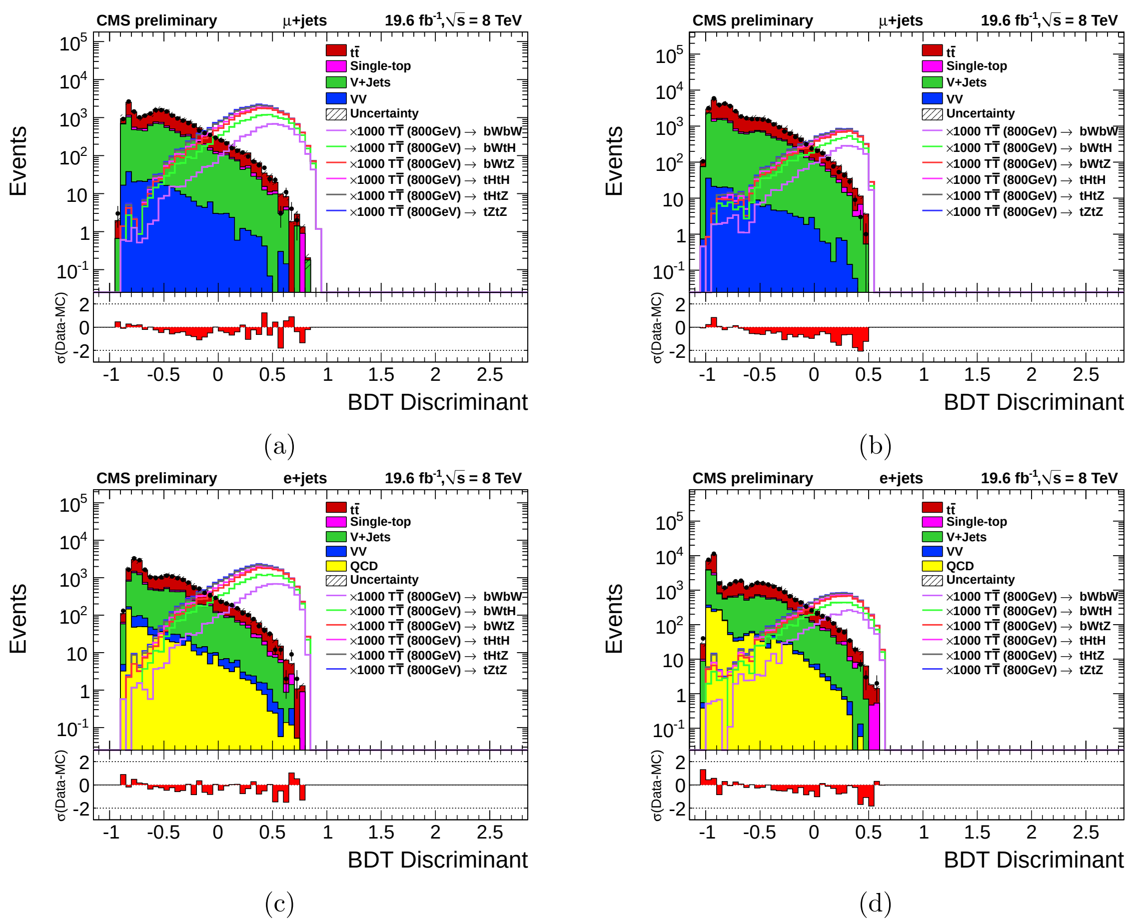
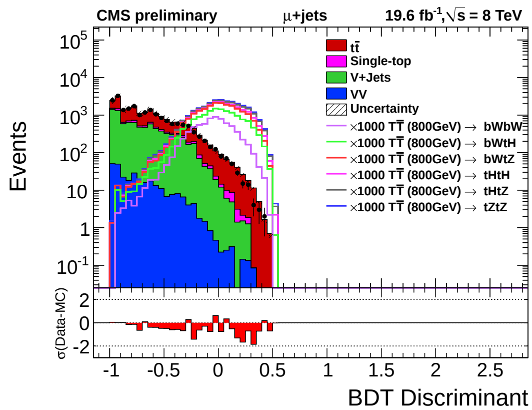
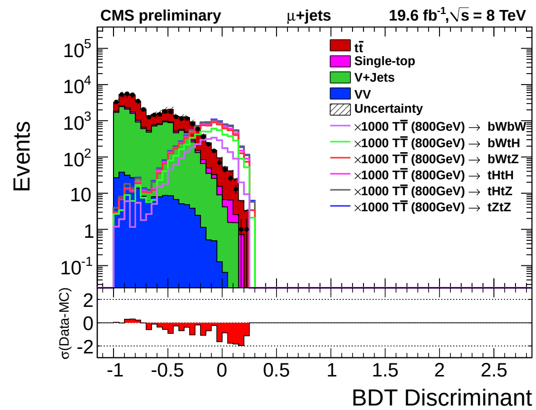


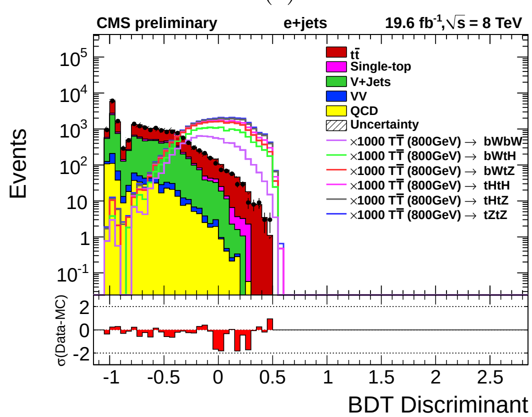
Figure C.9: The BDT distribution for training on the 100% $T \to bW$ scenario. The μ+jets channel for events (a) with and (b) without substructure and in the e+jets channel for events (c) with and (d) without substructure.



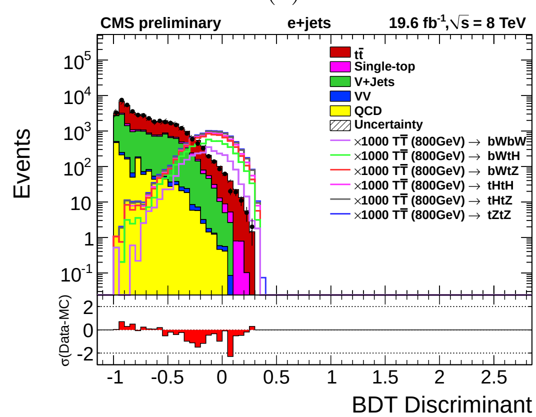
(a)



(b)

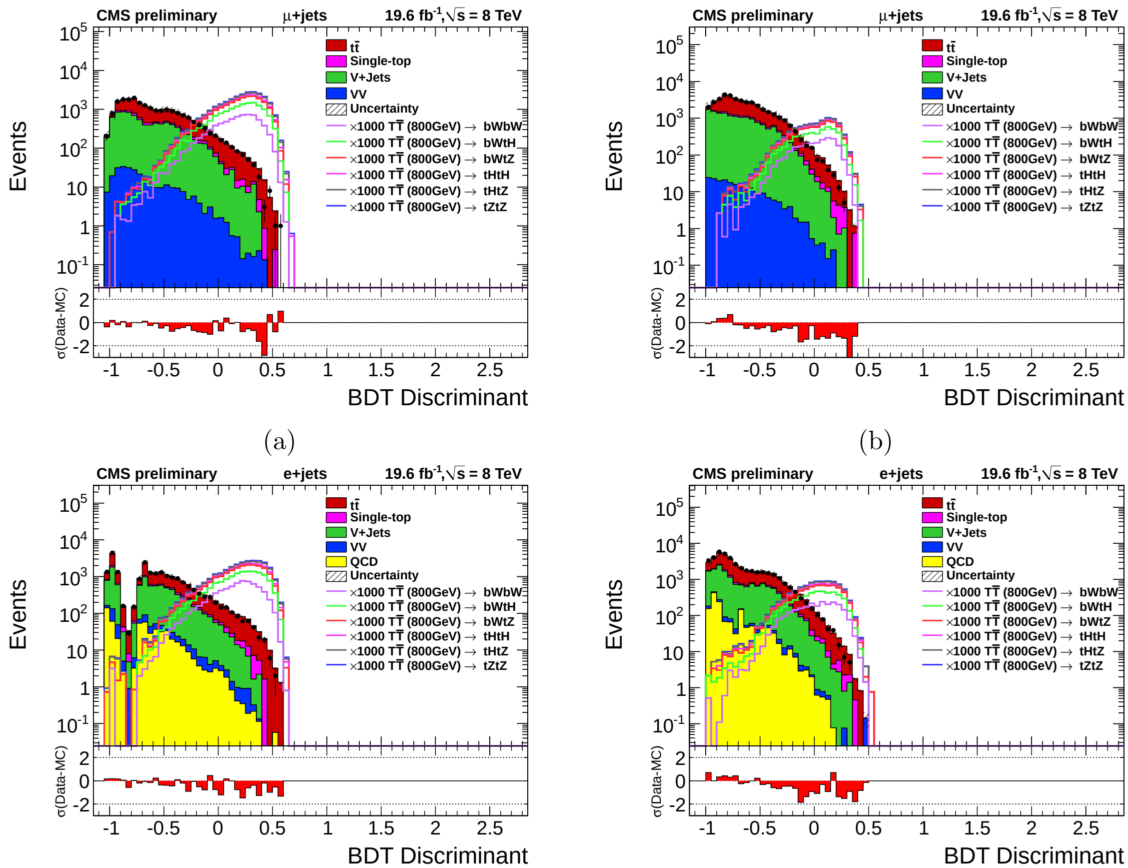


(c)



(d)

Figure C.10: The BDT distribution for training on the 100% $T \to tH$ scenario. The μ+jets channel for events (a) with and (b) without substructure and in the e+jets channel for events (c) with and (d) without substructure.



(a) (b) (c) (d)

Figure C.11: The BDT distribution for training on the 100% $T \to tZ$ scenario. The μ+jets channel for events (a) with and (b) without substructure and in the e+jets channel for events (c) with and (d) without substructure.

C.5 Sensitivity Cross Check : Training on Individual Signal Samples

We here check the affect of sensitivity on training on each of the full branching ratio channels independently. We compare the `BDT` distribution using training with the nominal branching ratio to the distribution created by training against the 100% decay modes in bW, tH, or tZ exclusively. The comparisons are shown in Fig. C.12.

For simplicity, this cross check was done with training using only the one mass 800 GeV and without separation of events into substructure categories. This training was then applied to all the different mass points. Further, this test uses an asymptotic $\mathrm{CL_S}$ technique described in [156, 222] with all non-template systematics but only the JES shape systematic for computational simplicity.

No significant improvement is observed in any of these channels at the 800 GeV mass point where our limits fall. For simplicity, we shall only train on the nominal branching ratios for each mass point.

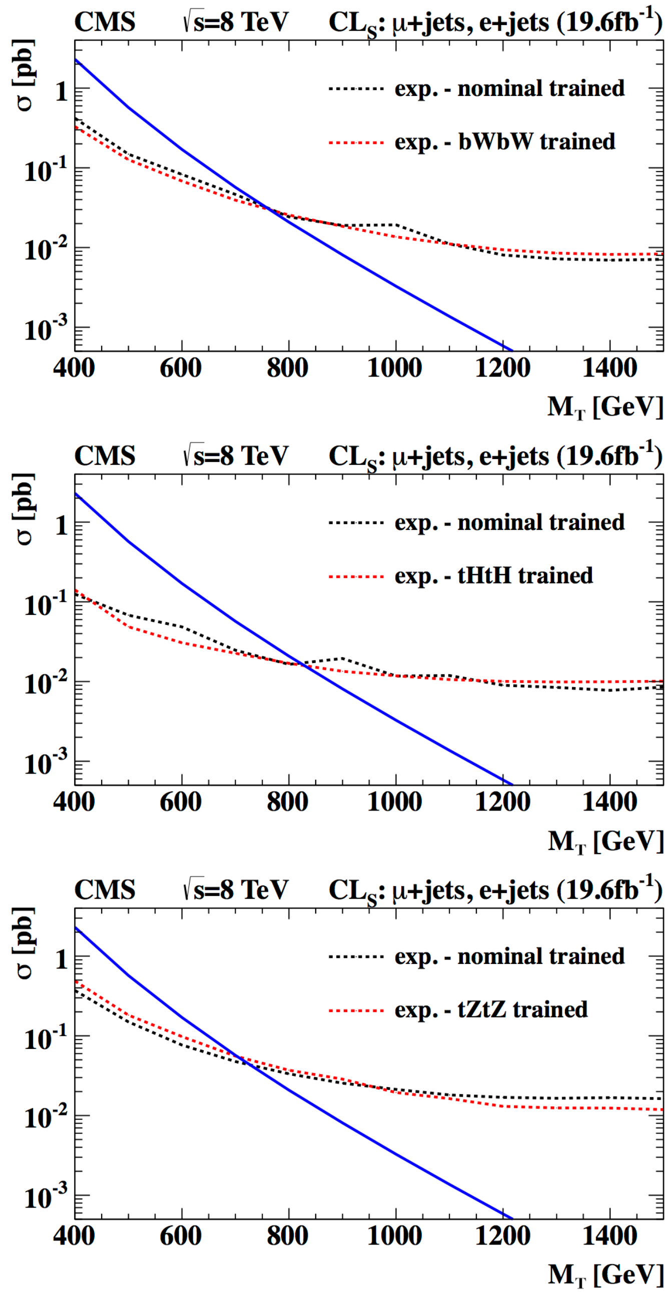
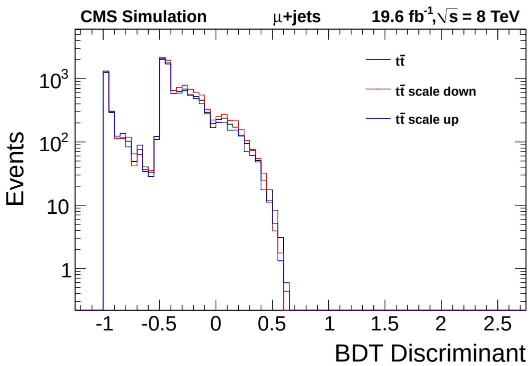


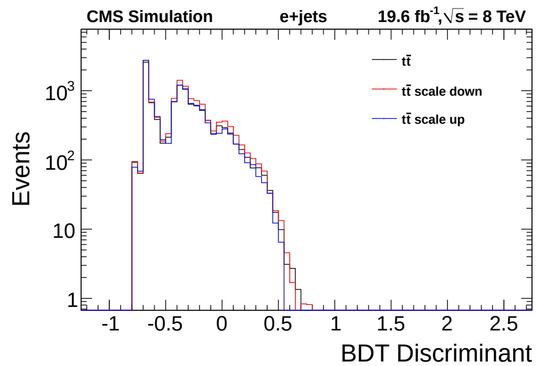
Figure C.12: Expected limits for the of the nominal training compared to training on the 100% BR in (top) $T \to bW$, (middle) $T \to tH$ and (bottom) $T \to tZ$ decay modes.

C.6 Cross Check : The Data Deficit in the BDT Distribution

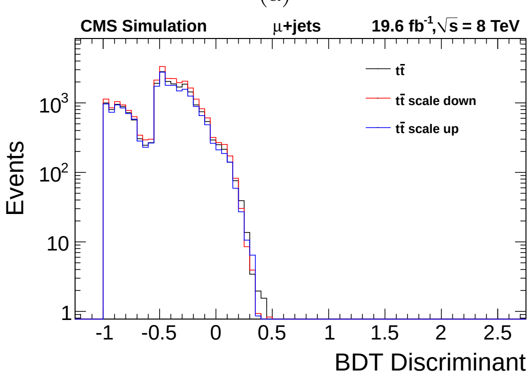
We account for the mismodelling of the data in the high BDT range by Q^2 (scale) systematics of the $t\bar{t}$ background described in Section 11.6. In particular, we see in Fig. C.13 the comparison between the nominal $t\bar{t}$ background and the scale templates in the BDT distribution.



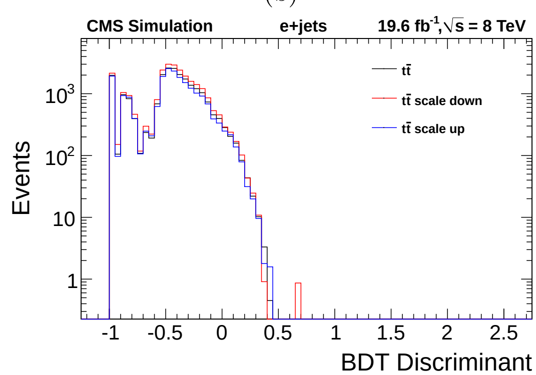
(a)



(b)



(c)

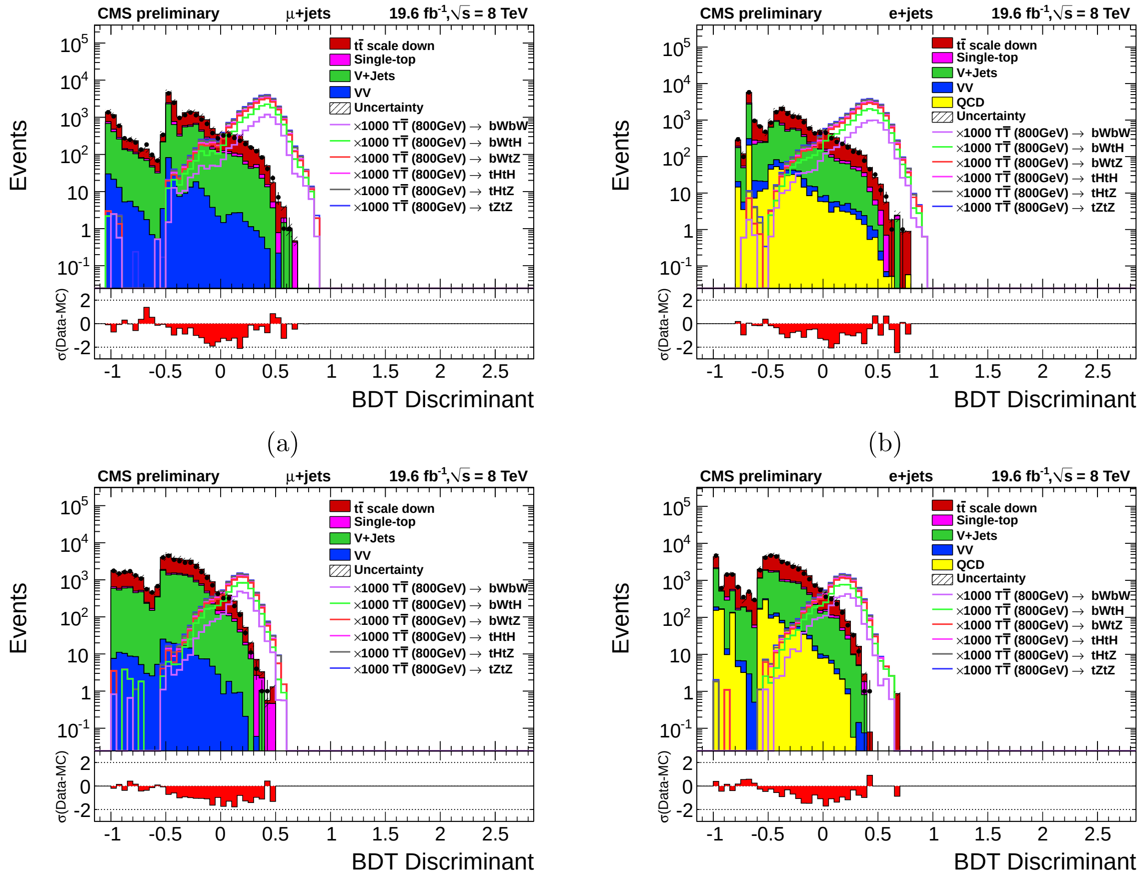


(d)

Figure C.13: BDT distributions for the $t\bar{t}$ background and its Q^2 scale template systematics in the signal region: for events with substructure in the (a) μ+jets and (b) e+jets channels and similarly for events without substructure in the (c) μ+jets and (d) e+jets channels.

Fig. C.14 shows the one-sigma down variation replacing the nominal $t\bar{t}$ contribution with the scale down template. As one can see, there is no longer a deficit in the data at the high BDT end. For completeness Fig. C.15 shows the BDT distribution using the scale up $t\bar{t}$

distribution. Since we use these scale template systematics in the limit setting procedure, we have thus accounted for the observed deficit in the data.

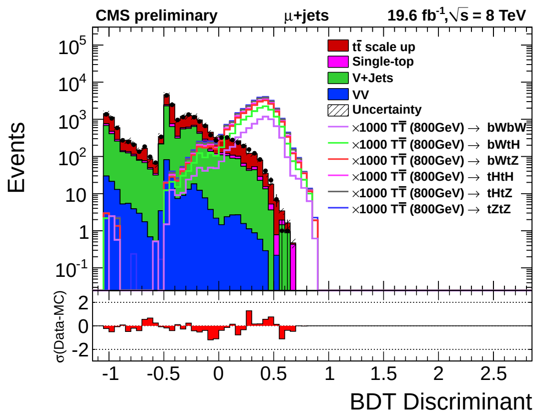


(a) (b) (c) (d)

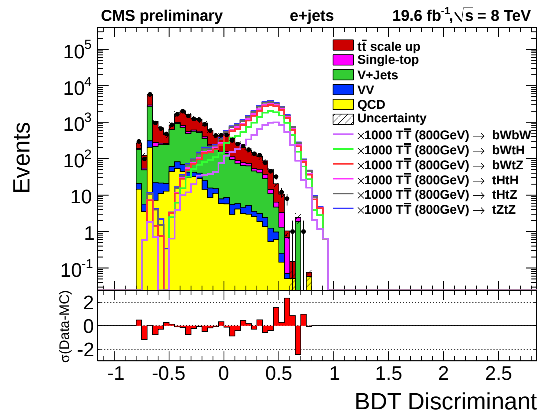
Figure C.14: BDT distributions with the $t\bar{t}$ background replaced with its Q^2 scale down systematic template in the signal region: for events with substructure in the (a) μ+jets and (b) e+jets channels and similarly for events without substructure in the (c) μ+jets and (d) e+jets channels.

The scale systematic also accounts for the data-MC disagreement in the H_T distributions (Fig. 11.22 and Fig. 11.29 for the μ+jets and e+jets channels respectively). Fig. C.16 shows the comparison between the nominal $t\bar{t}$ background and the $\pm 1\sigma$ scale templates in the H_T distribution.

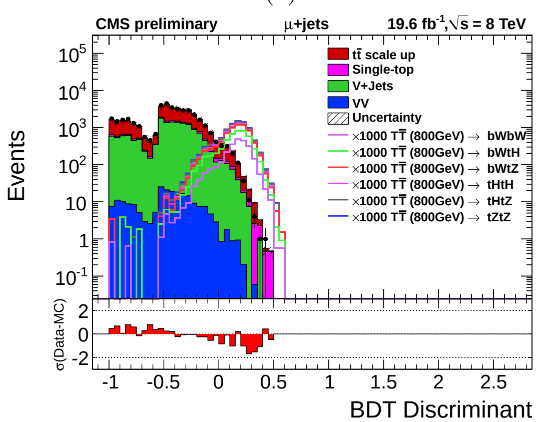
Finally, we substitute the nominal $t\bar{t}$ background H_T histogram with its scale systemic templates. This is shown in Fig. C.17. As one can see this systematic uncertainty by itself accounts for the majority of the disagreement between data and MC.



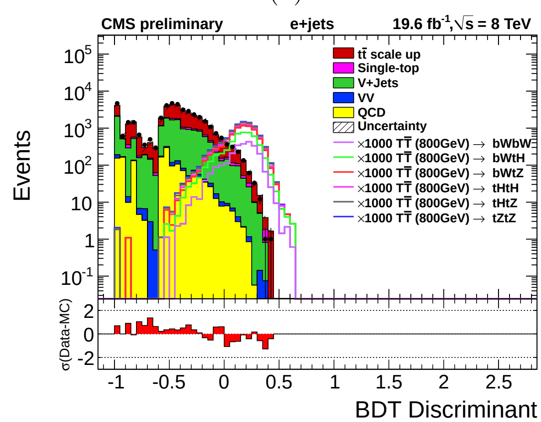
(a)



(b)

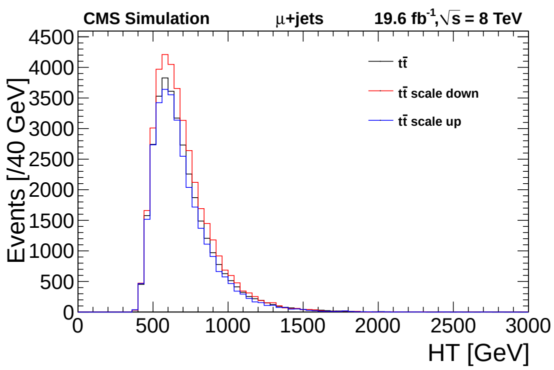


(c)

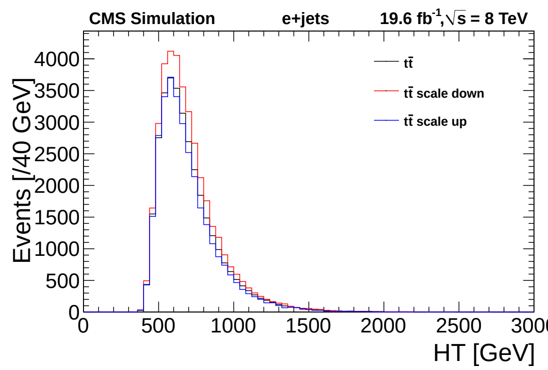


(d)

Figure C.15: BDT distributions with the $t\bar{t}$ background replaced with its Q^2 scale up systematic template in the signal region: for events with substructure in the (a) μ+jets and (b) e+jets channels and similarly for events without substructure in the (c) μ+jets and (d) e+jets channels.

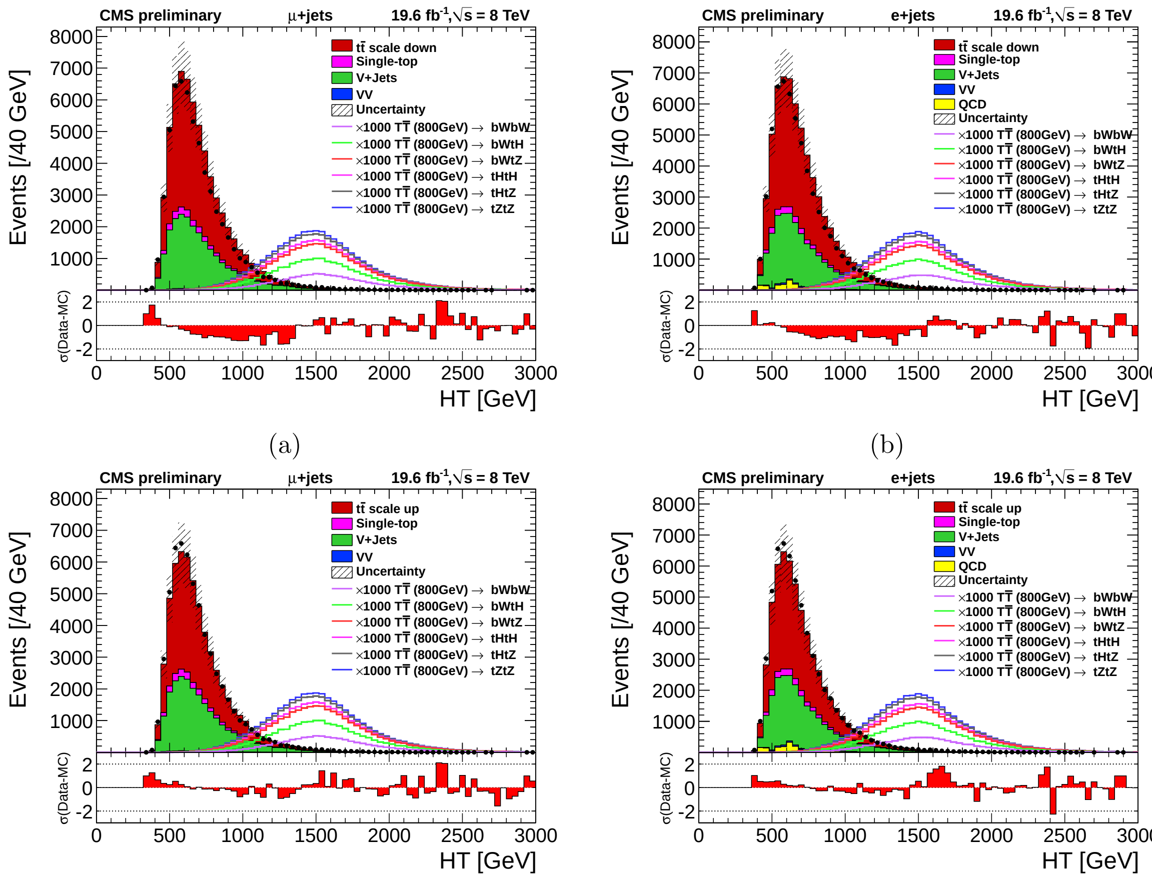


(a)



(b)

Figure C.16: H_T distributions for the $\mathrm{t\bar{t}}$ background and its Q^2 scale template systematics in the signal region: for events with substructure in the (a) μ+jets and (b) e+jets channels and similarly for events without substructure in the (c) μ+jets and (d) e+jets channels.



(a) (b) (c) (d)

Figure C.17: H_T distribution with the $t\bar{t}$ background replaced with its Q^2 scale systematic template in the signal region: using the $t\bar{t}$ scale down templates for the (a) μ+jets and (b) e+jets channels and similarly, using the $t\bar{t}$ scale up templates for the (c) μ+jets and (d) e+jets channels.

C.7 1D Limits for All Branching Ratios

We set limits in the muon and electron channels separately for the 66 branching ratio points, 22 of which (corresponding to the 0.2 interval steps in BR) are listed in Table C.1. For conciseness, only the combined electron, combined muon and total combined lepton+jets channels are shown.

Scenario	$T \to bW$	$T \to tH$	$T \to tZ$	expected μ	observed μ	expected e	observed e
(0)	0.5	0.25	0.25	697 GeV	689 GeV	698 GeV	683 GeV
(1)	0.0	0.0	1.0	680 GeV	675 GeV	681 GeV	675 GeV
(2)	0.0	0.2	0.8	685 GeV	679 GeV	685 GeV	677 GeV
(3)	0.0	0.4	0.6	690 GeV	684 GeV	692 GeV	678 GeV
(4)	0.0	0.6	0.4	694 GeV	690 GeV	696 GeV	682 GeV
(5)	0.0	0.8	0.2	698 GeV	695 GeV	706 GeV	685 GeV
(6)	0.0	1.0	0.0	712 GeV	699 GeV	731 GeV	692 GeV
(7)	0.2	0.0	0.8	682 GeV	680 GeV	685 GeV	676 GeV
(8)	0.2	0.2	0.6	688 GeV	682 GeV	692 GeV	678 GeV
(9)	0.2	0.4	0.4	694 GeV	686 GeV	695 GeV	680 GeV
(10)	0.2	0.6	0.2	697 GeV	691 GeV	702 GeV	683 GeV
(11)	0.2	0.8	0.0	708 GeV	694 GeV	726 GeV	688 GeV
(12)	0.4	0.0	0.6	686 GeV	685 GeV	690 GeV	679 GeV
(13)	0.4	0.2	0.4	691 GeV	687 GeV	696 GeV	681 GeV
(14)	0.4	0.4	0.2	698 GeV	691 GeV	699 GeV	682 GeV
(15)	0.4	0.6	0.0	712 GeV	695 GeV	721 GeV	685 GeV
(16)	0.6	0.0	0.4	688 GeV	689 GeV	696 GeV	680 GeV
(17)	0.6	0.2	0.2	696 GeV	694 GeV	701 GeV	683 GeV
(18)	0.6	0.4	0.0	720 GeV	695 GeV	725 GeV	682 GeV
(19)	0.8	0.0	0.2	696 GeV	696 GeV	699 GeV	682 GeV
(20)	0.8	0.2	0.0	711 GeV	699 GeV	709 GeV	684 GeV
(21)	1.0	0.0	0.0	706 GeV	702 GeV	705 GeV	680 GeV

Table C.1: Shows the branching ratios and the limits for the μ+jets and e+jets channels for the 0.2 interval points in phase space where the limit computation is done.

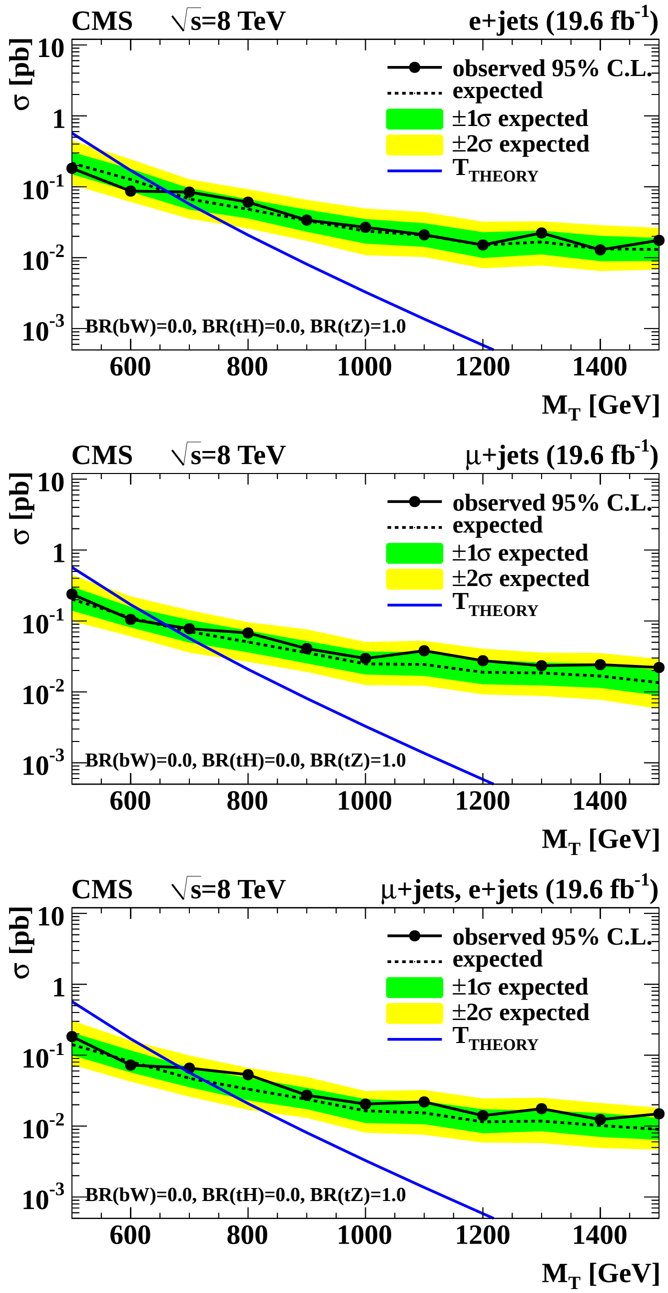


Figure C.18: Observed and expected limits in the (top) e+jets, (middle) μ+jets and combined channel on the $T\overline{T}$ production cross section as a function of T quark mass. The data correspond to $19.6\,\mathrm{fb}^{-1}$ for both the e+jets channel and μ+jets channels. The branching ratios BR(bW:tH:tZ) = 0.0 : 0.0 : 1.0 correspond to version 1 listed in Table 11.16

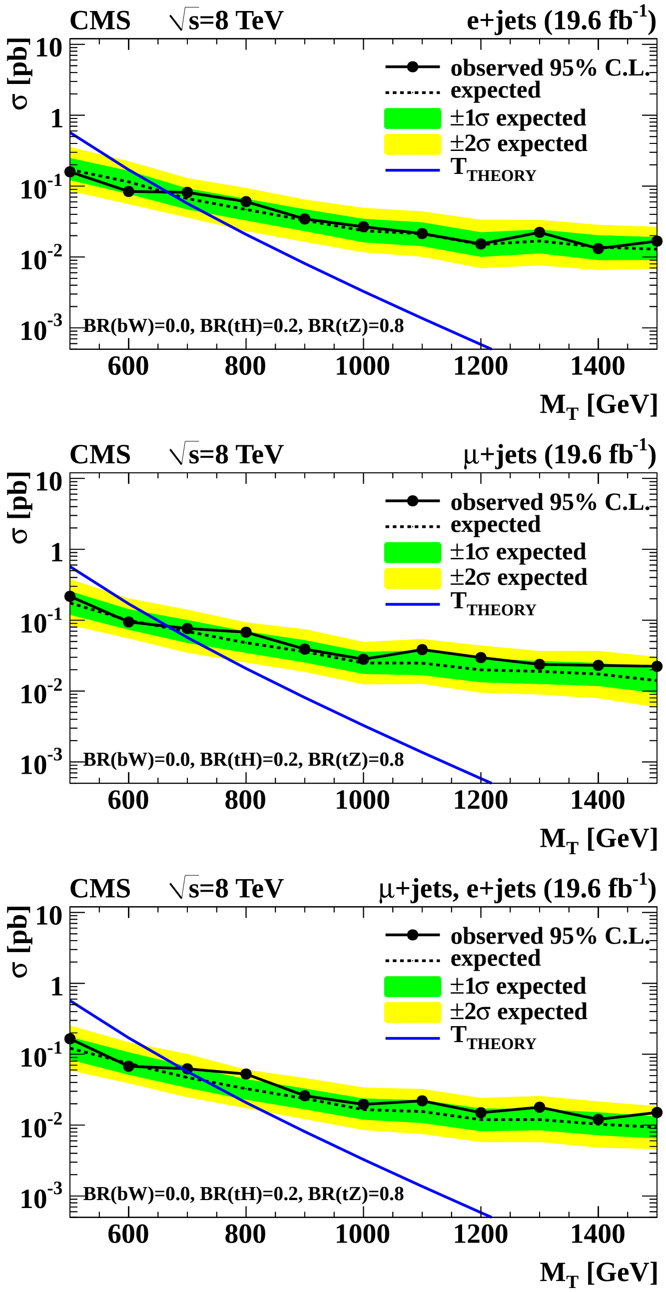


Figure C.19: Observed and expected limits in the (top) e+jets, (middle) μ+jets and combined channel on the $T\overline{T}$ production cross section as a function of T quark mass. The data correspond to $19.6\,\mathrm{fb}^{-1}$ for both the e+jets channel and μ+jets channels. The branching ratios correspond to version BR(bW:tH:tZ) $= 0.0 : 0.2 : 0.8$ correspond to version 2 listed in Table 11.16

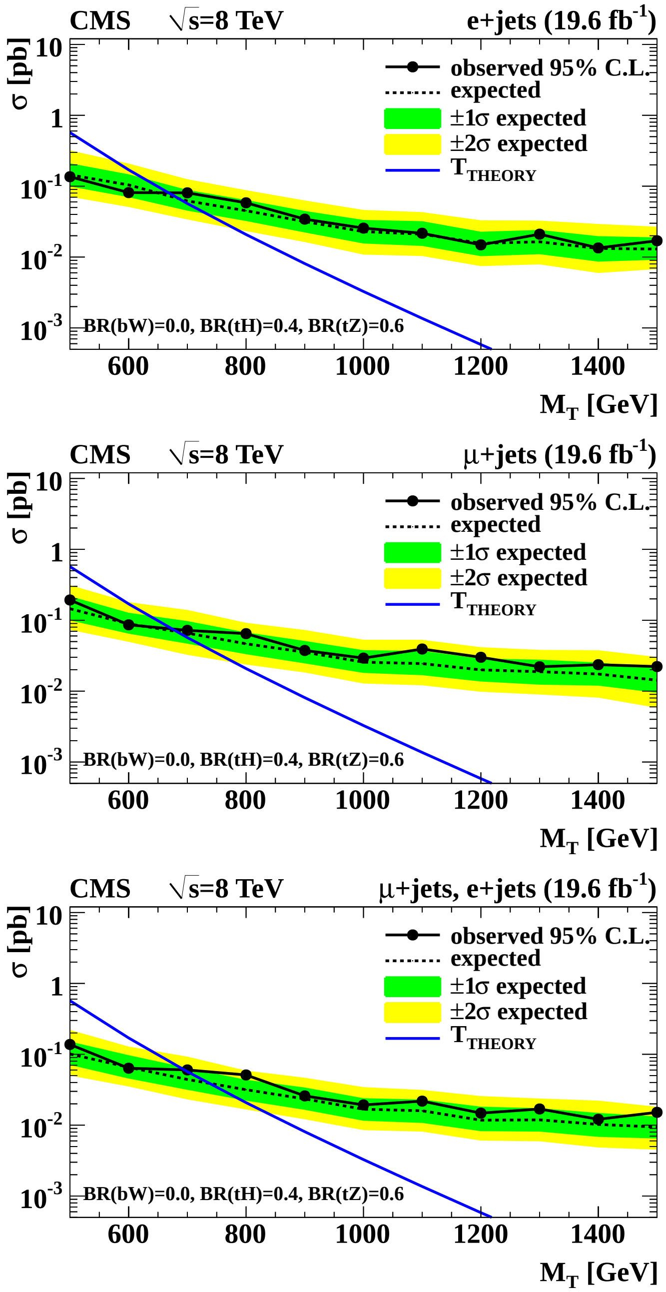


Figure C.20: Observed and expected limits in the (top) e+jets, (middle) μ+jets and combined channel on the $T\overline{T}$ production cross section as a function of T quark mass. The data correspond to $19.6\,\mathrm{fb}^{-1}$ for both the e+jets channel and μ+jets channels. The branching ratios BR(bW:tH:tZ) = 0.0 : 0.4 : 0.6 correspond to version 3 listed in Table 11.16

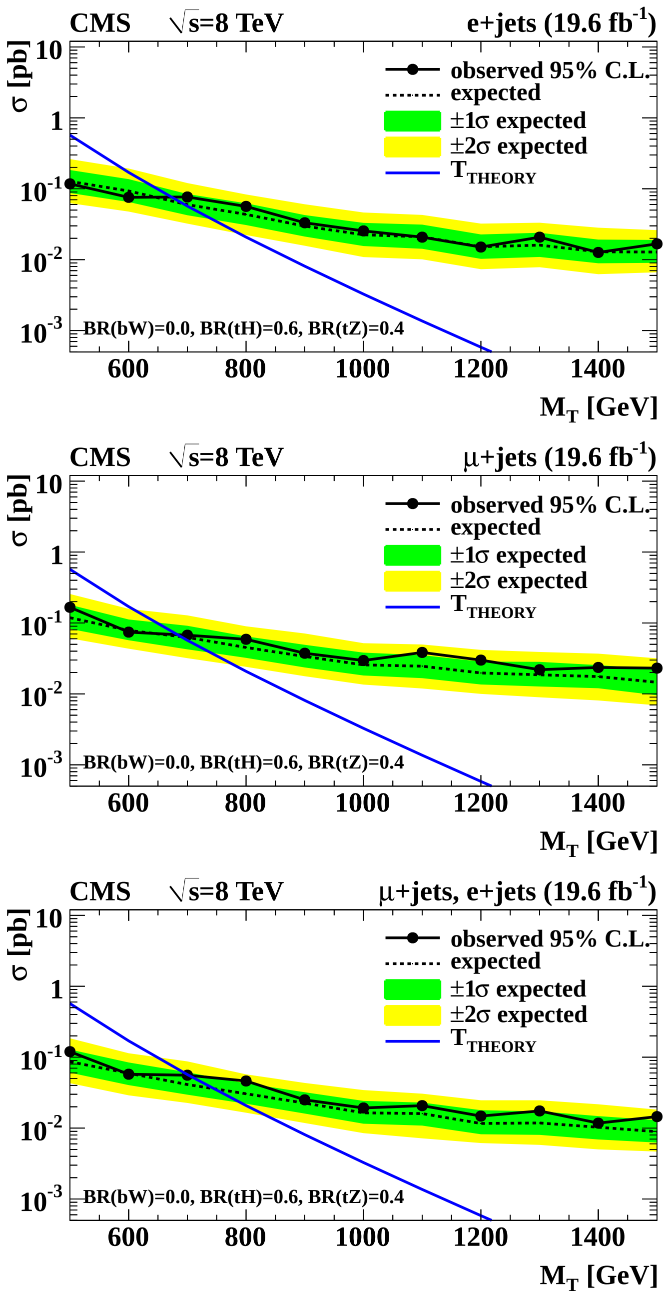


Figure C.21: Observed and expected limits in the (top) e+jets, (middle) μ+jets and combined channel on the $T\overline{T}$ production cross section as a function of T quark mass. The data correspond to $19.6\,\mathrm{fb}^{-1}$ for both the e+jets channel and μ+jets channels. The branching ratios correspond to version 4 listed in Table 11.16

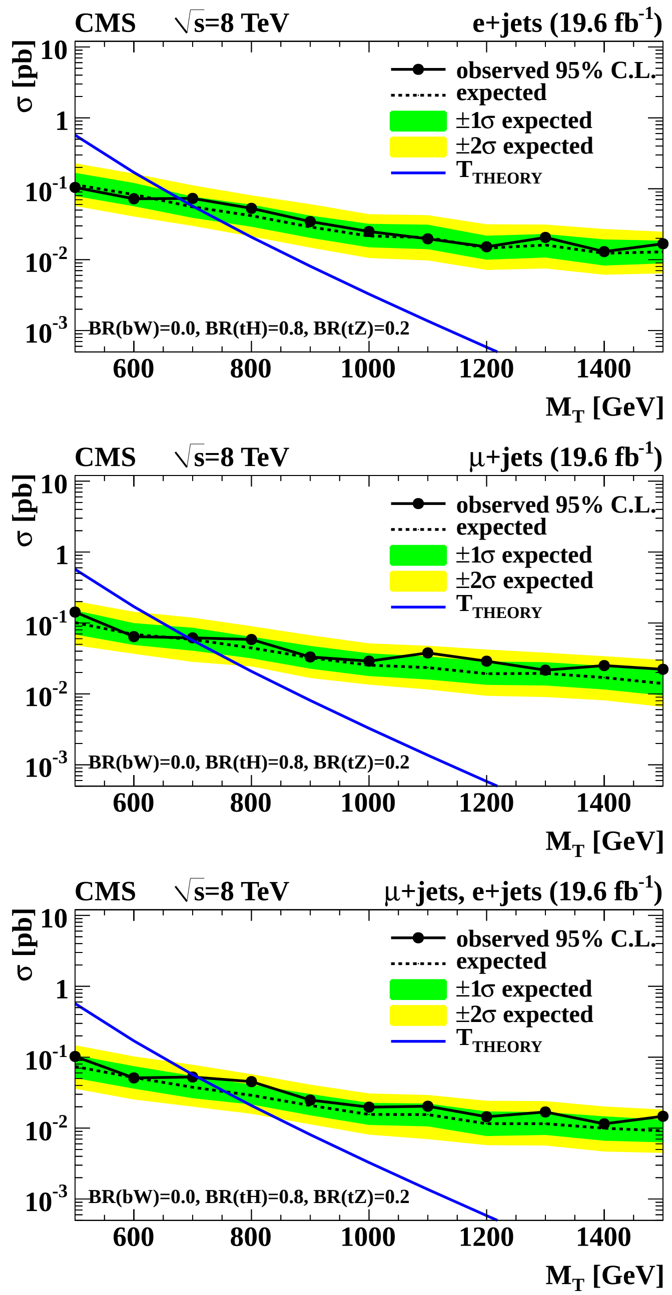


Figure C.22: Observed and expected limits in the (top) e+jets, (middle) μ+jets and combined channel on the $T\overline{T}$ production cross section as a function of T quark mass. The data correspond to $19.6\,\mathrm{fb}^{-1}$ for both the e+jets channel and μ+jets channels. The branching ratios BR(bW:tH:tZ) = 0.0 : 0.8 : 0.2 correspond to version 5 listed in Table 11.16

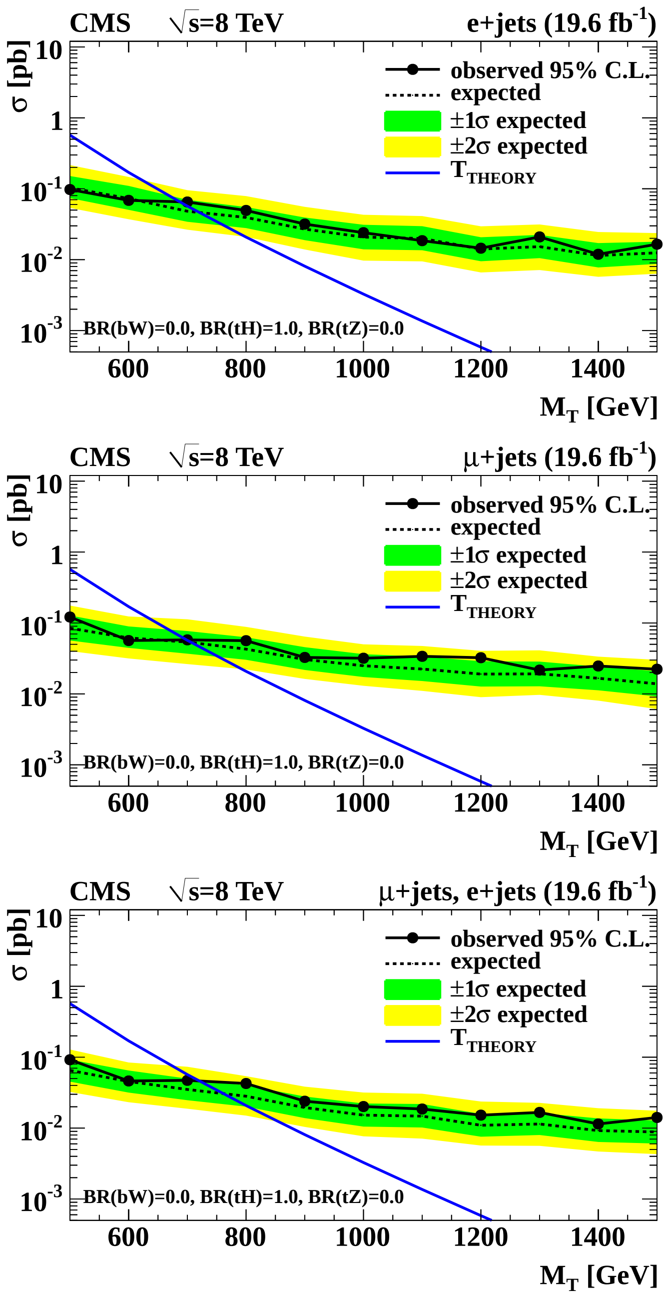


Figure C.23: Observed and expected limits in the (top) e+jets, (middle) μ+jets and combined channel on the $T\overline{T}$ production cross section as a function of T quark mass. The data correspond to 19.6 fb^{-1} for both the e+jets channel and μ+jets channels. The branching ratios BR(bW:tH:tZ) = 0.0 : 1.0 : 0.0 correspond to version 6 listed in Table 11.16

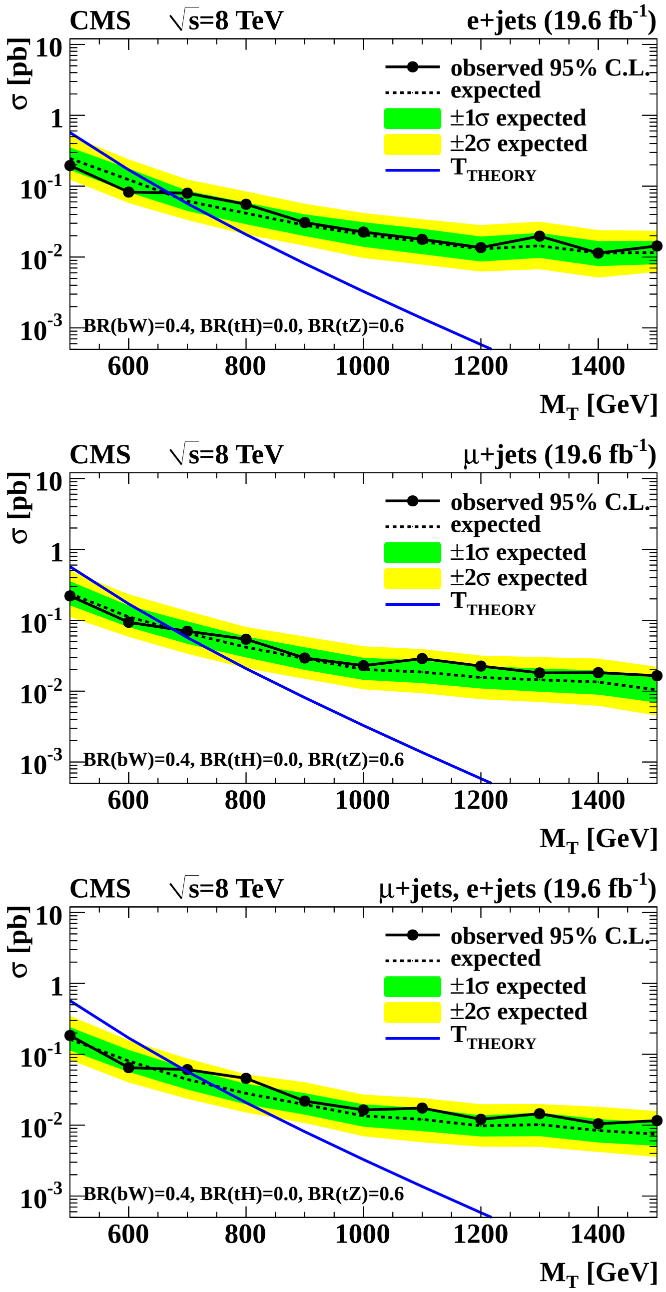


Figure C.24: Observed and expected limits in the (top) e+jets, (middle) μ+jets and combined channel on the $T\overline{T}$ production cross section as a function of T quark mass. The data correspond to 19.6 fb^{-1} for both the e+jets channel and μ+jets channels. The branching ratios BR(bW:tH:tZ) = 0.2 : 0.0 : 0.8 correspond to version 12 listed in Table 11.16

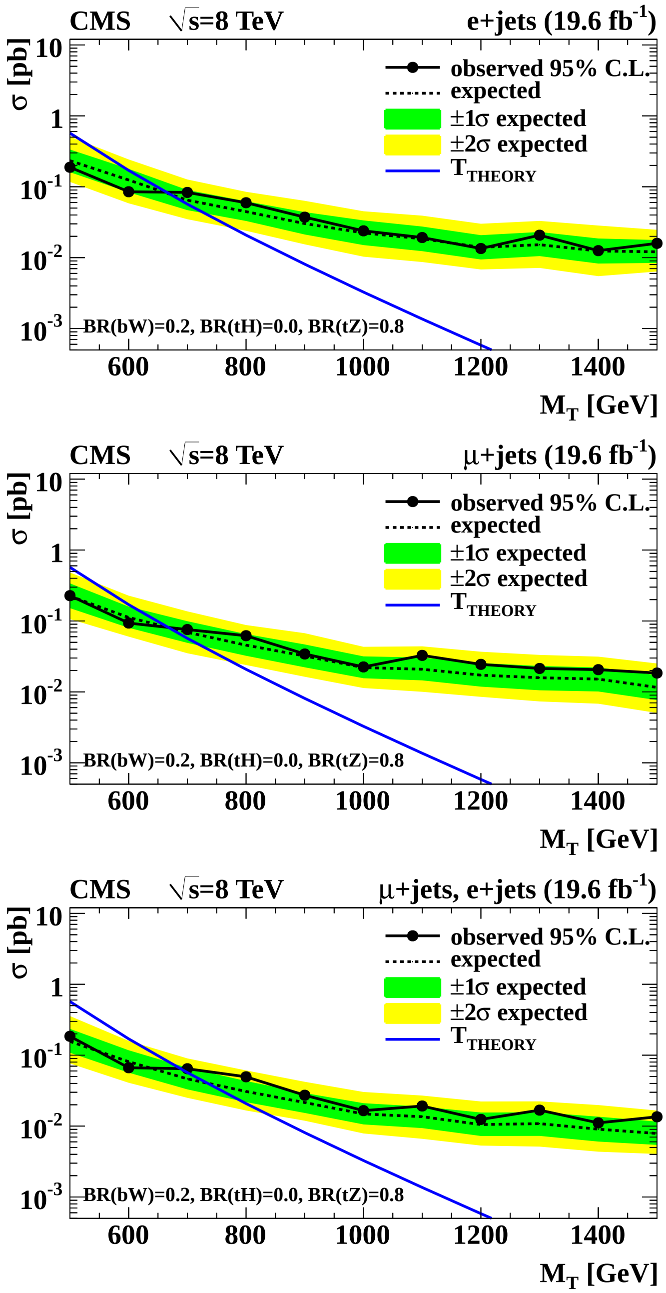


Figure C.25: Observed and expected limits in the (top) e+jets, (middle) μ+jets and combined channel on the $T\overline{T}$ production cross section as a function of T quark mass. The data correspond to $19.6\,\mathrm{fb}^{-1}$ for both the e+jets channel and μ+jets channels. The branching ratios BR(bW:tH:tZ) $= 0.4 : 0.0 : 0.6$ correspond to version 7 listed in Table 11.16

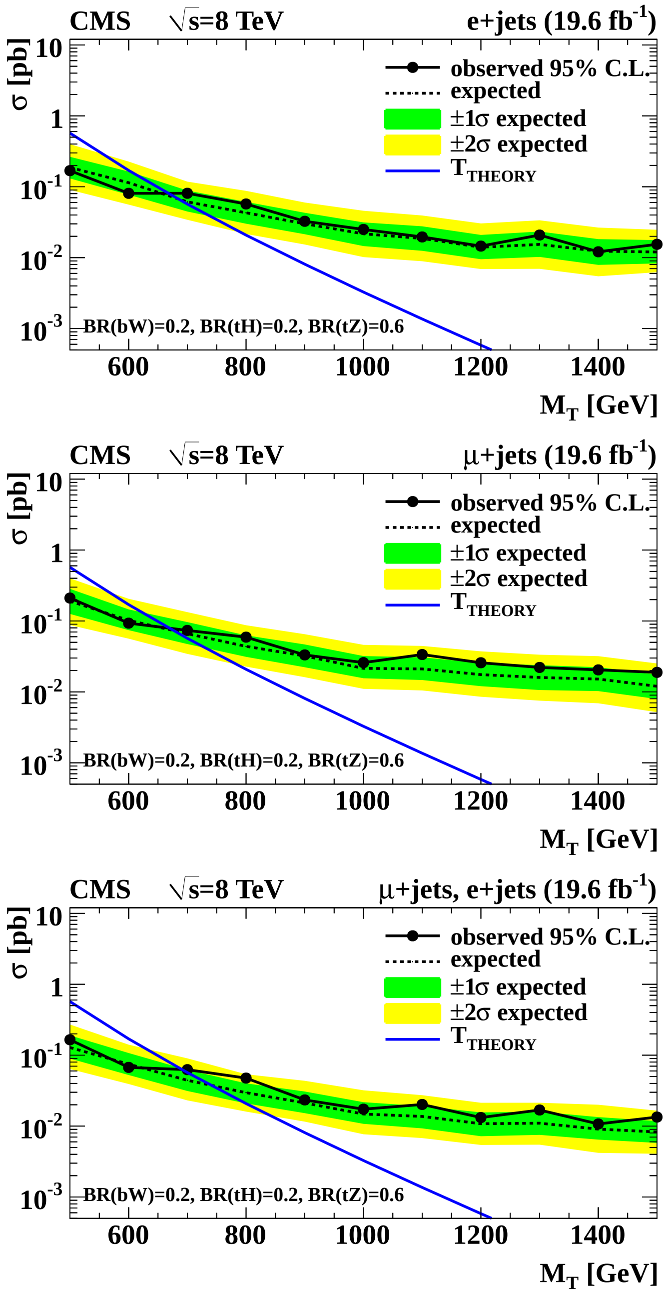


Figure C.26: Observed and expected limits in the (top) e+jets, (middle) μ+jets and combined channel on the $T\overline{T}$ production cross section as a function of T quark mass. The data correspond to $19.6\,\mathrm{fb}^{-1}$ for both the e+jets channel and μ+jets channels. The branching ratios BR(bW:tH:tZ) = 0.4 : 0.2 : 0.4 correspond to version 8 listed in Table 11.16

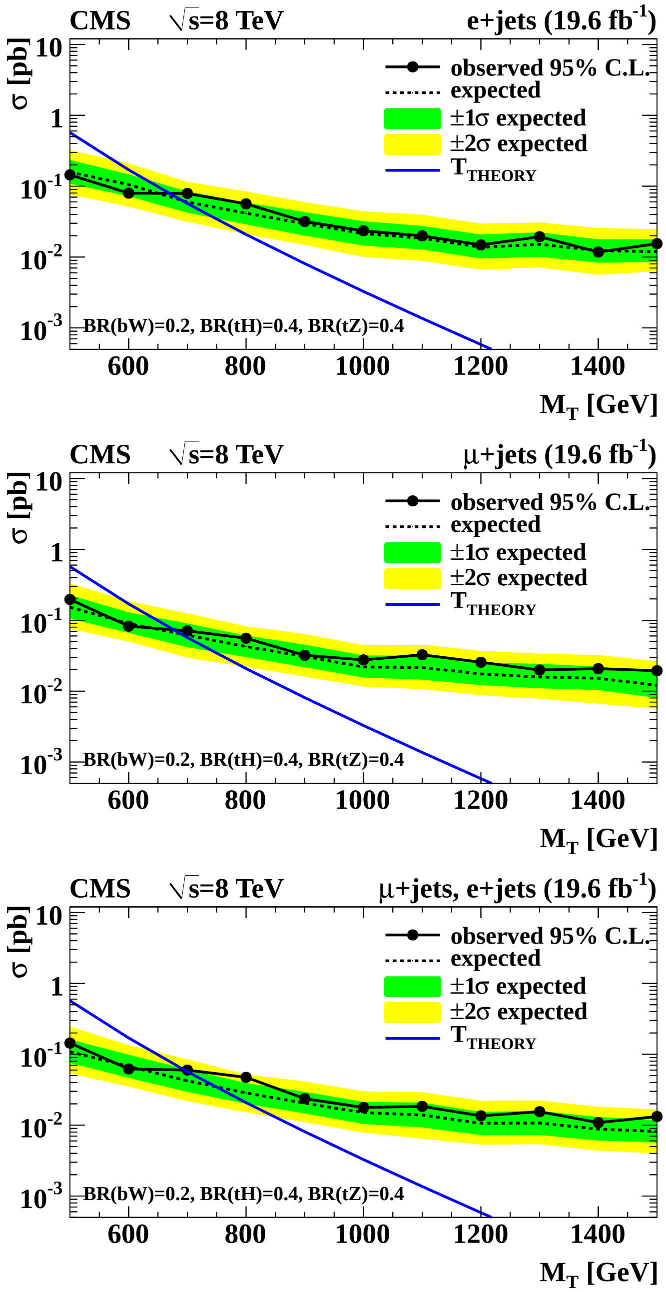


Figure C.27: Observed and expected limits in the (top) e+jets, (middle) μ+jets and combined channel on the $T\overline{T}$ production cross section as a function of T quark mass. The data correspond to 19.6 fb^{-1} for both the e+jets channel and μ+jets channels. The branching ratios BR(bW:tH:tZ) = 0.4 : 0.4 : 0.2 correspond to version 9 listed in Table 11.16

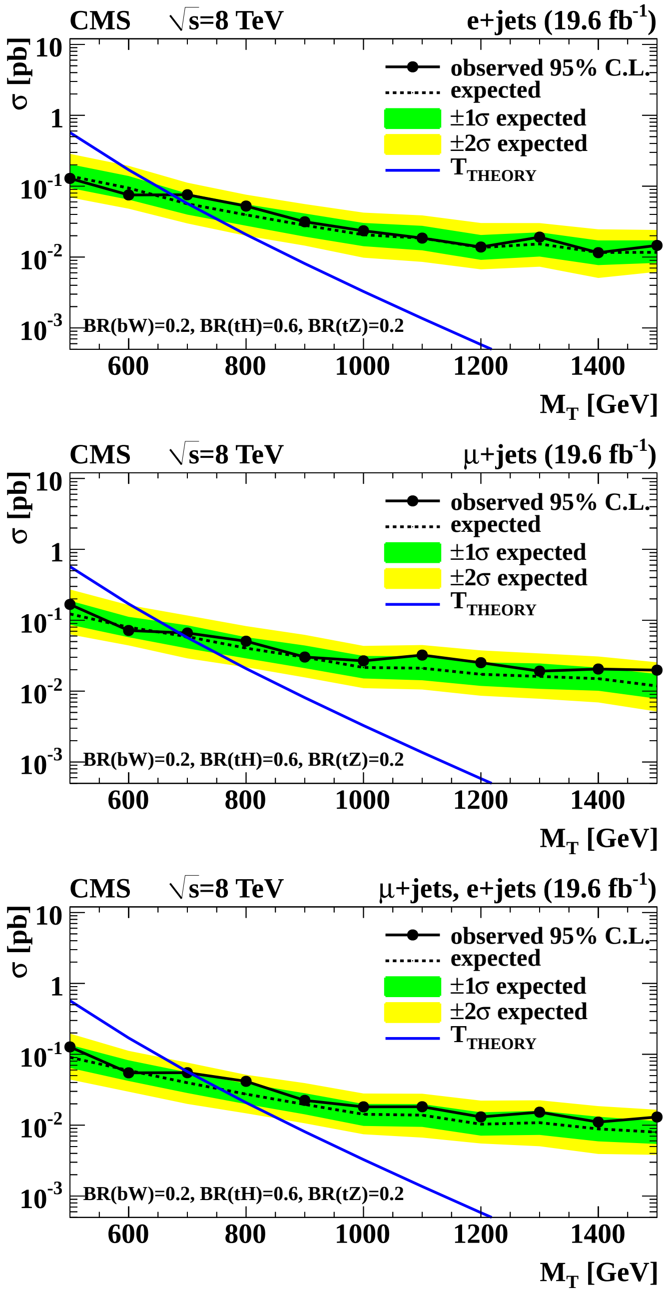


Figure C.28: Observed and expected limits in the (top) e+jets, (middle) μ+jets and combined channel on the $T\overline{T}$ production cross section as a function of T quark mass. The data correspond to 19.6 fb^{-1} for both the e+jets channel and μ+jets channels. The branching ratios BR(bW:tH:tZ) = 0.4 : 0.6 : 0.0 correspond to version 10 listed in Table 11.16

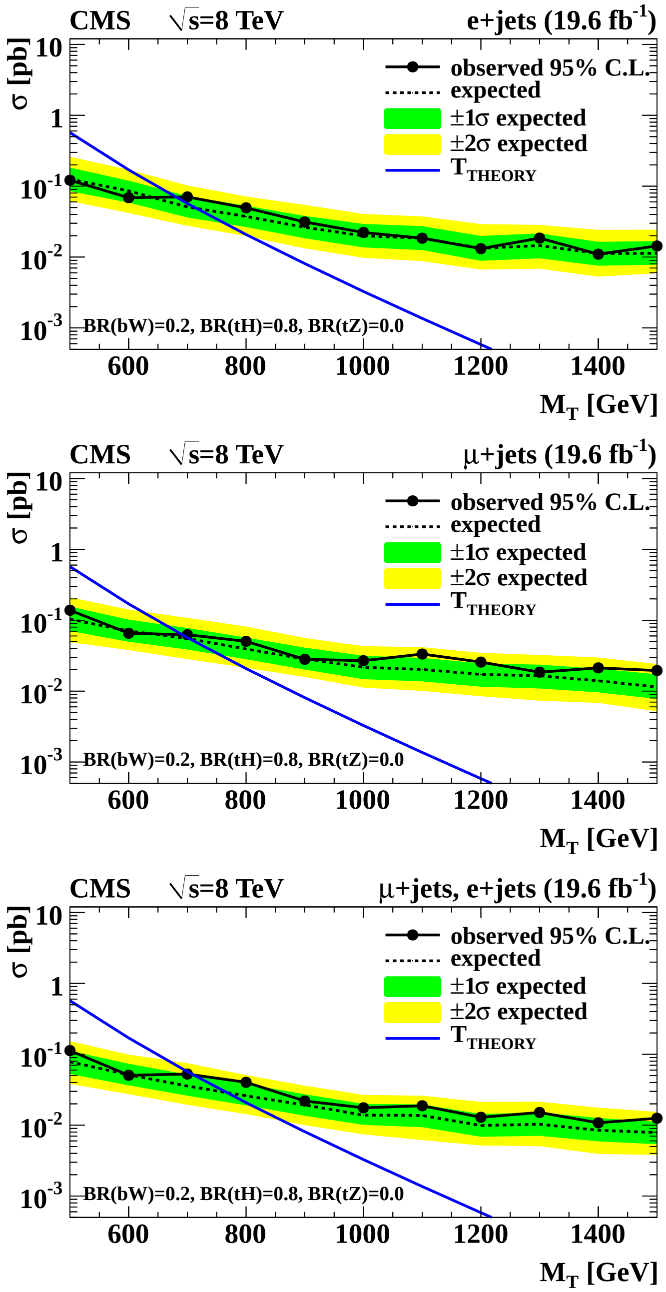
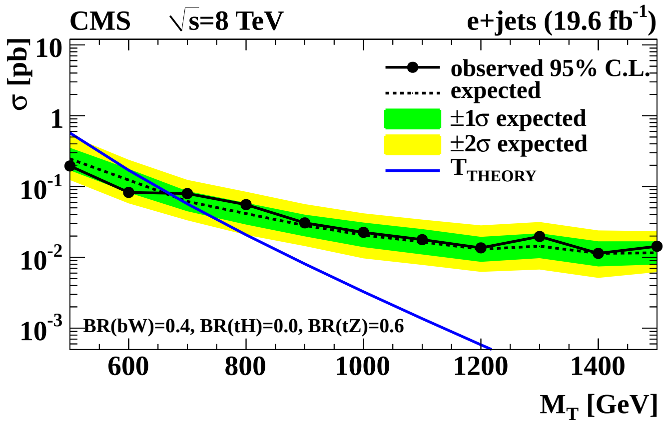
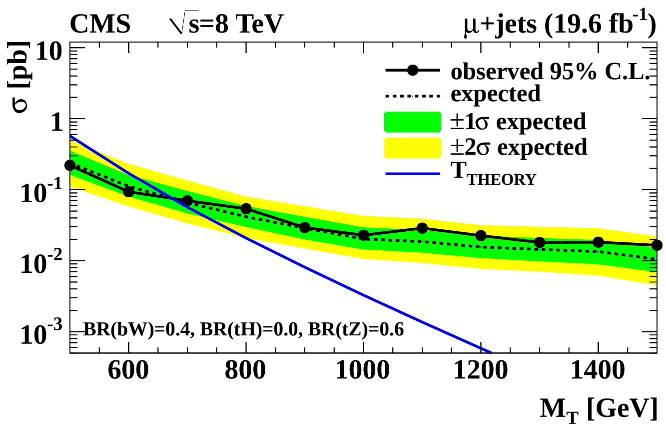


Figure C.29: Observed and expected limits in the (top) e+jets, (middle) μ+jets and combined channel on the $T\overline{T}$ production cross section as a function of T quark mass. The data correspond to 19.6 fb^{-1} for both the e+jets channel and μ+jets channels. The branching ratios BR(bW:tH:tZ) = 0.2 : 0.8 : 0.0 correspond to version 11 listed in Table 11.16





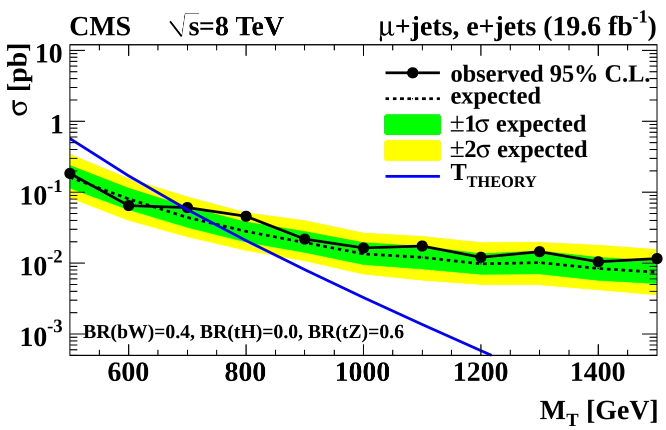


Figure C.30: Observed and expected limits in the (top) e+jets, (middle) μ+jets and combined channel on the $T\overline{T}$ production cross section as a function of T quark mass. The data correspond to $19.6\,\text{fb}^{-1}$ for both the e+jets channel and μ+jets channels. The branching ratios BR(bW:tH:tZ) $= 0.4 : 0.0 : 0.6$ correspond to version 12 listed in Table 11.16.

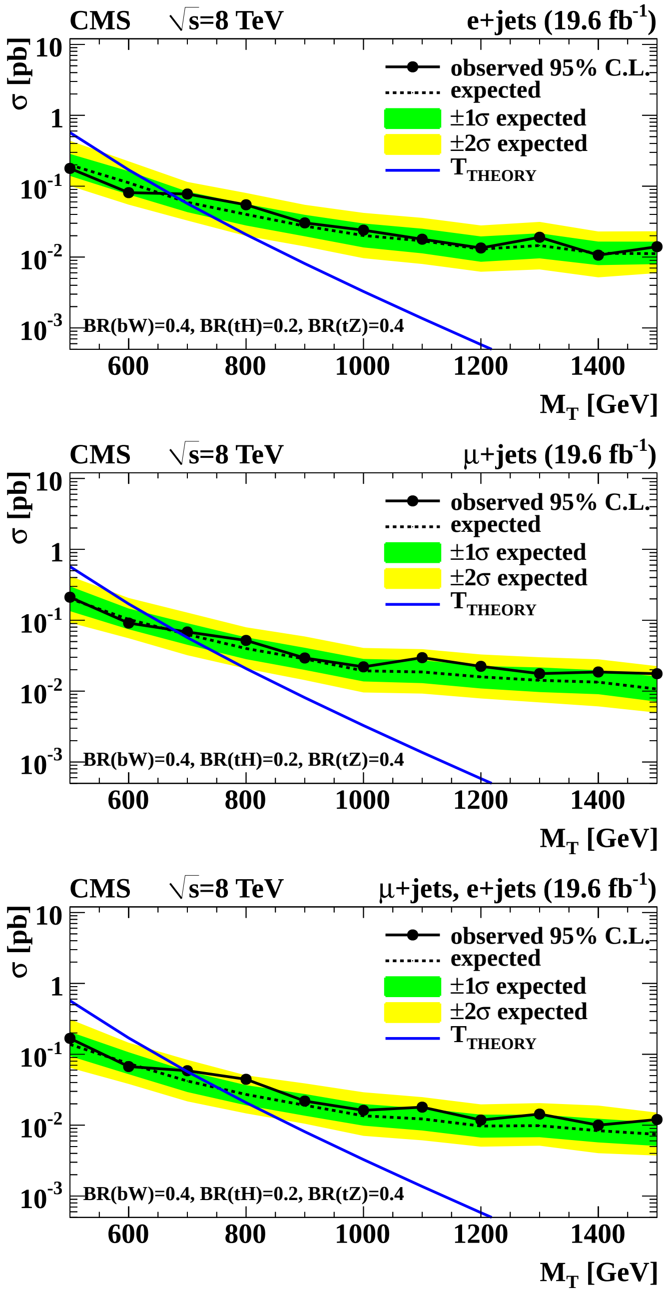


Figure C.31: Observed and expected limits in the (top) e+jets, (middle) μ+jets and combined channel on the $T\overline{T}$ production cross section as a function of T quark mass. The data correspond to $19.6\,\mathrm{fb}^{-1}$ for both the e+jets channel and μ+jets channels. The branching ratios BR(bW:tH:tZ) = 0.4 : 0.2 : 0.4 correspond to version 13 listed in Table 11.16.

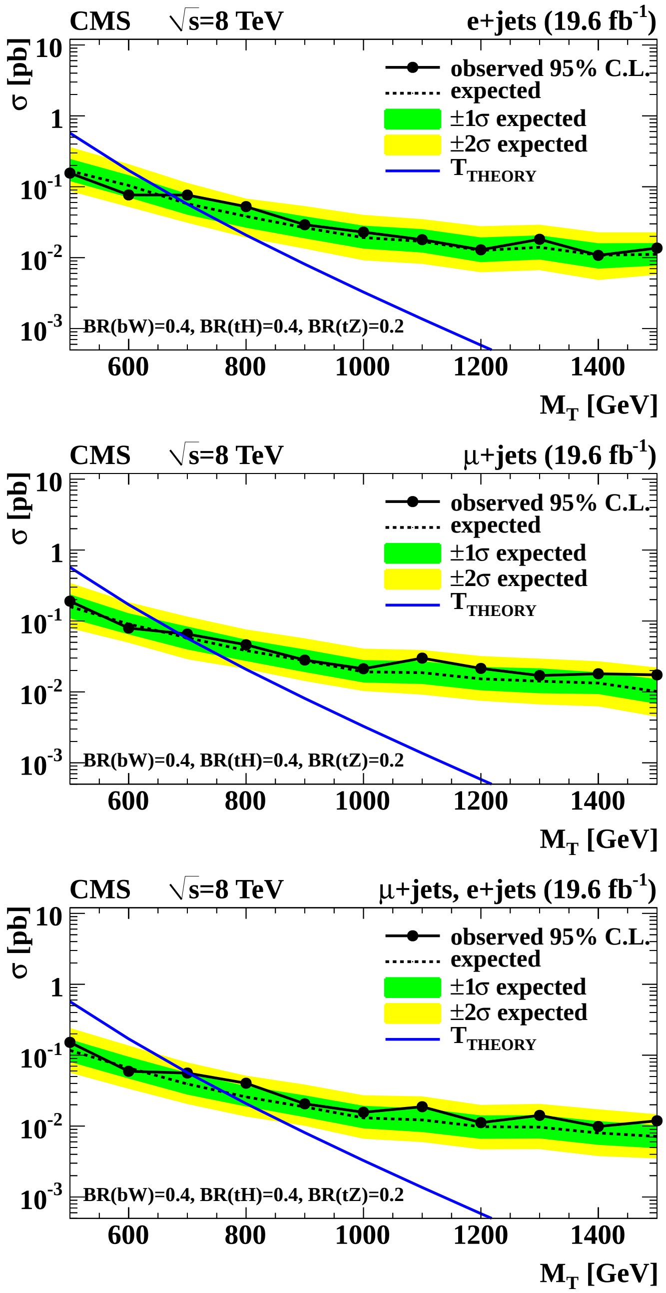


Figure C.32: Observed and expected limits in the (top) e+jets, (middle) μ+jets and combined channel on the $T\overline{T}$ production cross section as a function of T quark mass. The data correspond to 19.6 fb^{-1} for both the e+jets channel and μ+jets channels. The branching ratios BR(bW:tH:tZ) = 0.4 : 0.4 : 0.2 correspond to version 14 listed in Table 11.16

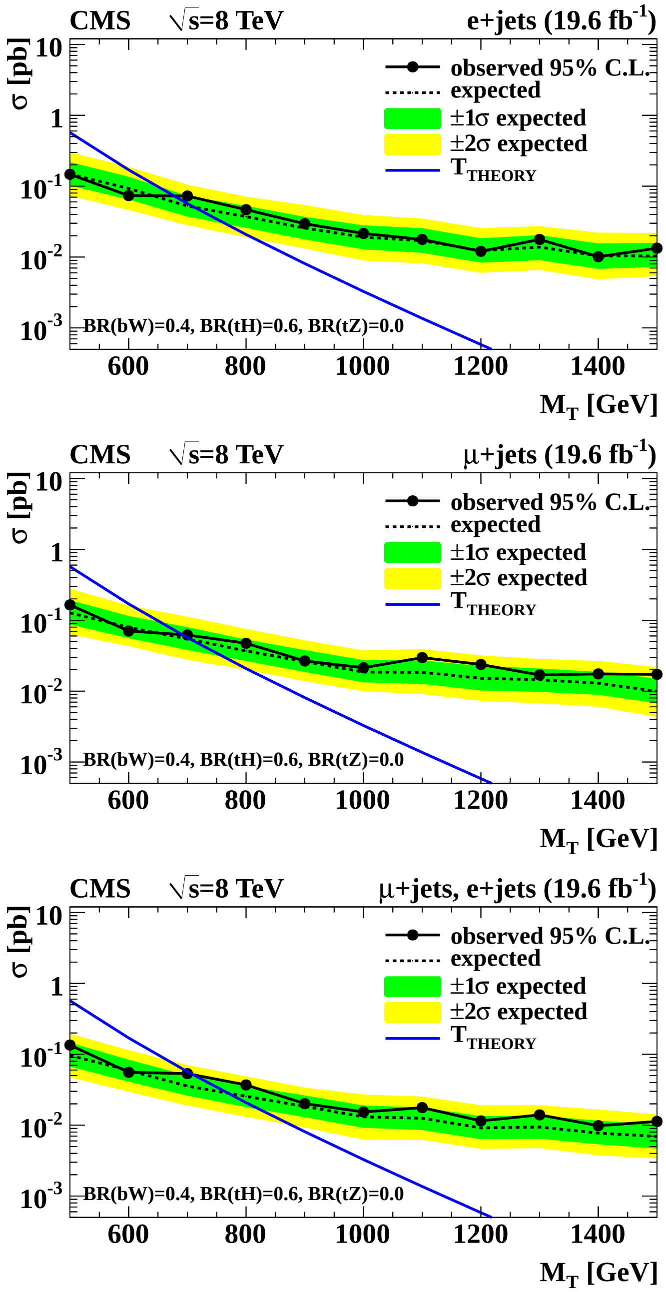


Figure C.33: Observed and expected limits in the (top) e+jets, (middle) μ+jets and combined channel on the $T\overline{T}$ production cross section as a function of T quark mass. The data correspond to $19.6\,\mathrm{fb}^{-1}$ for both the e+jets channel and μ+jets channels. The branching ratios BR(bW:tH:tZ) = 0.4 : 0.6 : 0.0 correspond to version 15 listed in Table 11.16

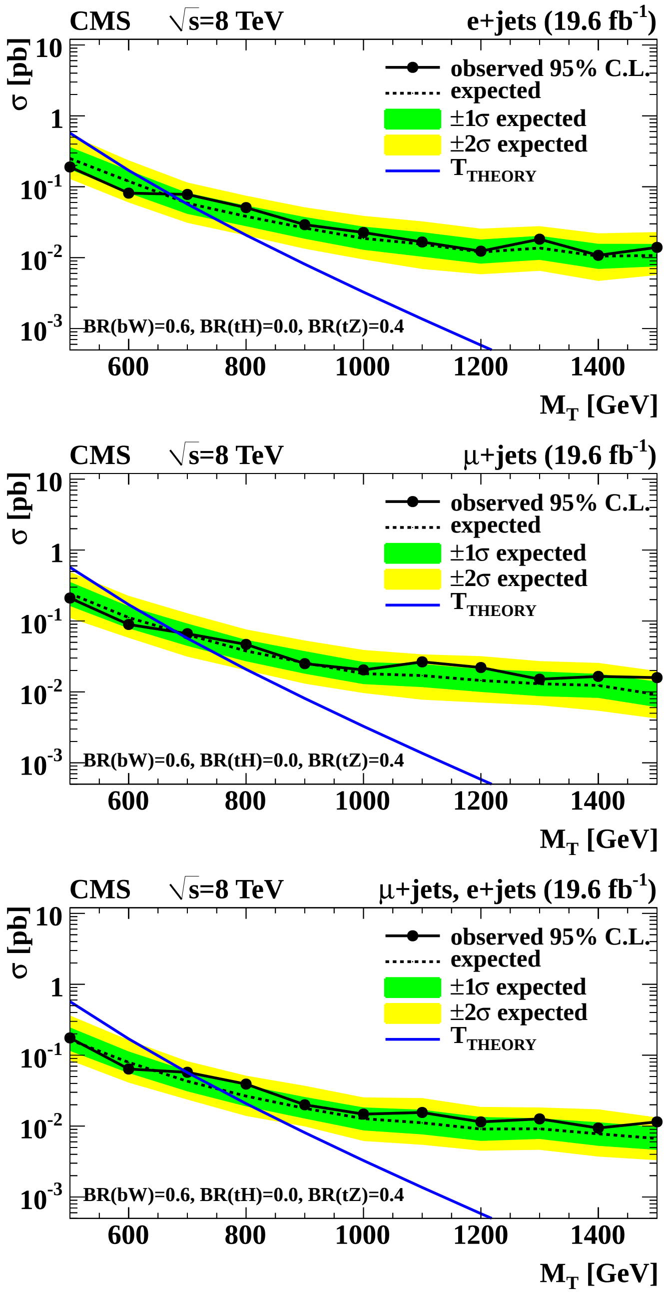


Figure C.34: Observed and expected limits in the (top) e+jets, (middle) μ+jets and combined channel on the $T\overline{T}$ production cross section as a function of T quark mass. The data correspond to $19.6\,\mathrm{fb}^{-1}$ for both the e+jets channel and μ+jets channels. The branching ratios BR(bW:tH:tZ) = 0.6 : 0.0 : 0.4 correspond to version 16 listed in Table 11.16

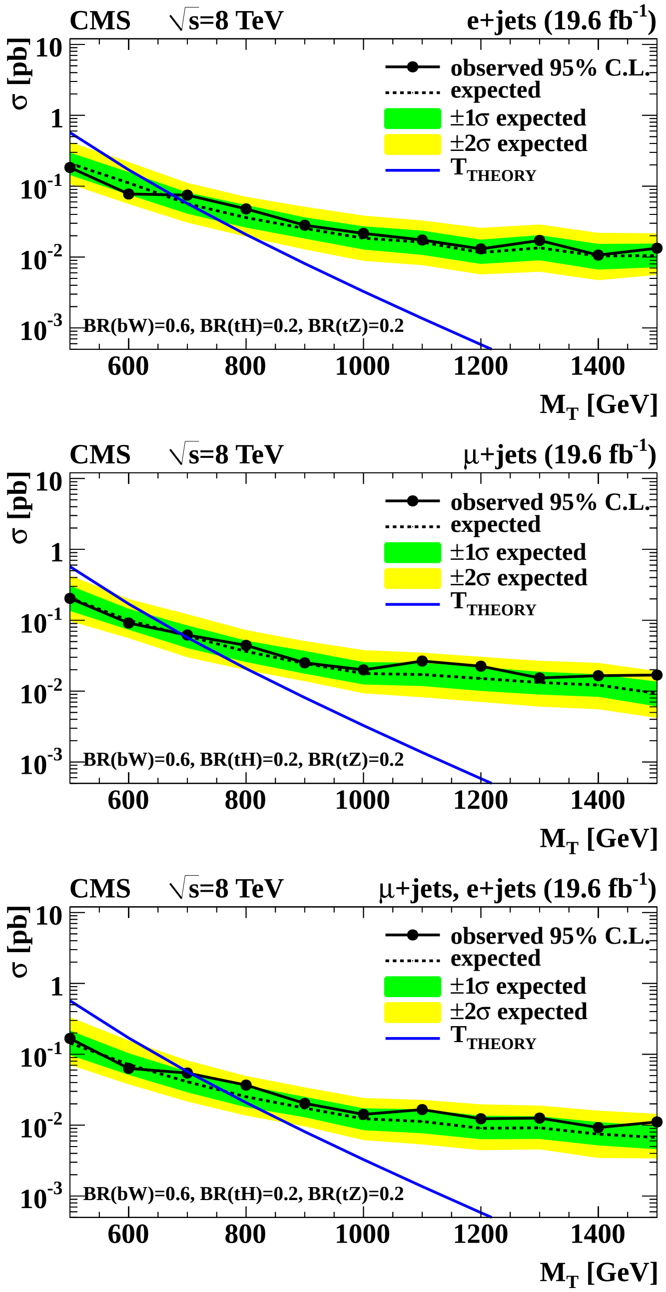


Figure C.35: Observed and expected limits in the (top) e+jets, (middle) μ+jets and combined channel on the $T\overline{T}$ production cross section as a function of T quark mass. The data correspond to $19.6\,\mathrm{fb}^{-1}$ for both the e+jets channel and μ+jets channels. The branching ratios BR(bW:tH:tZ) = 0.6 : 0.2 : 0.2 correspond to version 17 listed in Table 11.16

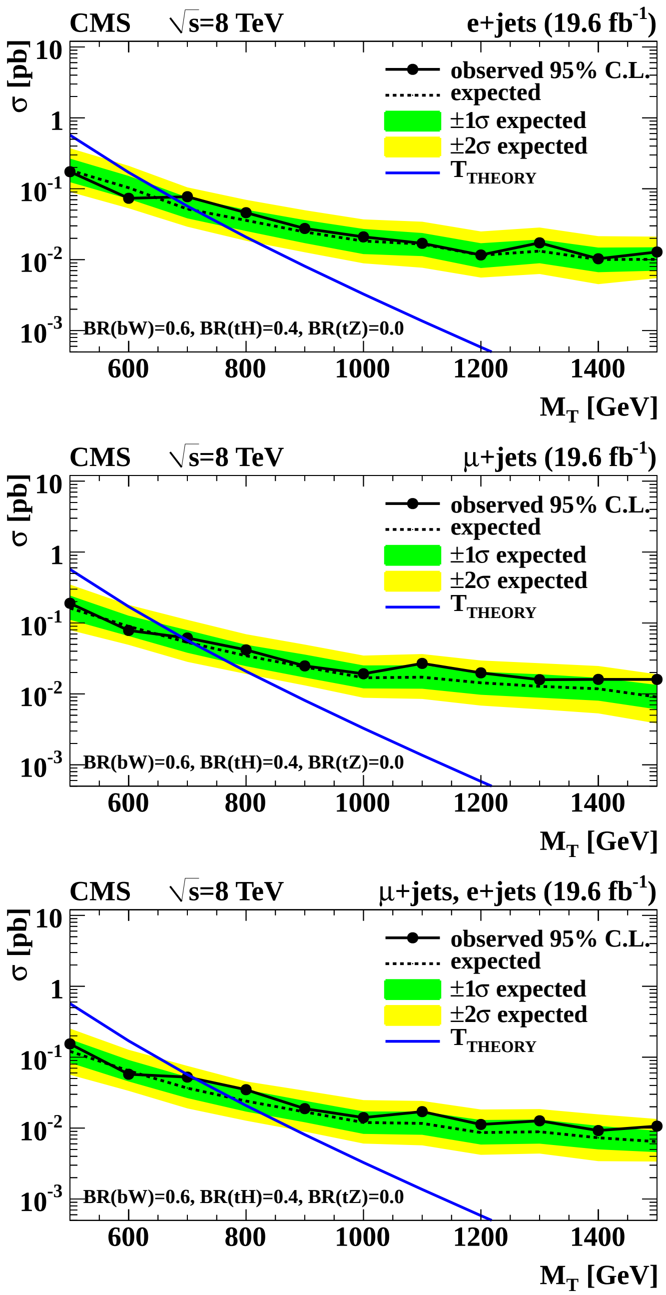


Figure C.36: Observed and expected limits in the (top) e+jets, (middle) μ+jets and combined channel on the $T\overline{T}$ production cross section as a function of T quark mass. The data correspond to $19.6\,\mathrm{fb}^{-1}$ for both the e+jets channel and μ+jets channels. The branching ratios $\mathrm{BR}(bW{:}tH{:}tZ) = 0.6 : 0.4 : 0.0$ correspond to version 18 nominal listed in Table 11.16

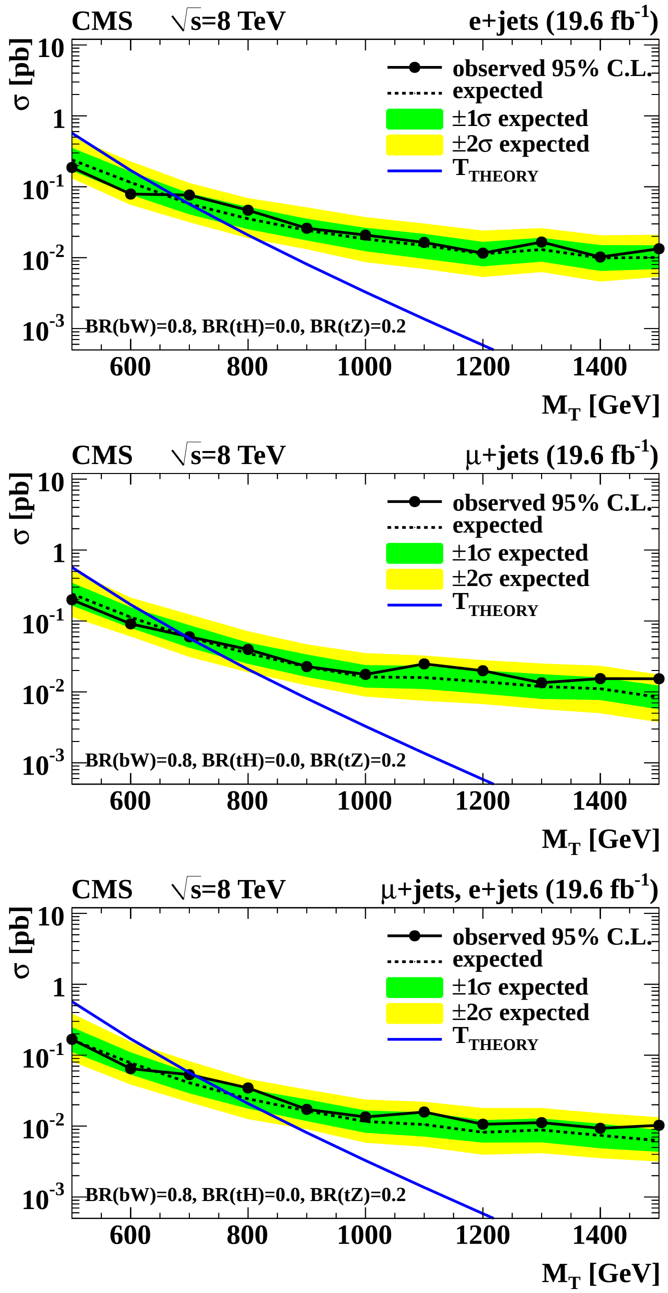


Figure C.37: Observed and expected limits in the (top) e+jets, (middle) μ+jets and combined channel on the $T\overline{T}$ production cross section as a function of T quark mass. The data correspond to $19.6\,\mathrm{fb}^{-1}$ for both the e+jets channel and μ+jets channels. The branching ratios $\mathrm{BR}(bW{:}tH{:}tZ) = 0.8 : 0.0 : 0.2$ correspond to version 19 nominal listed in Table 11.16

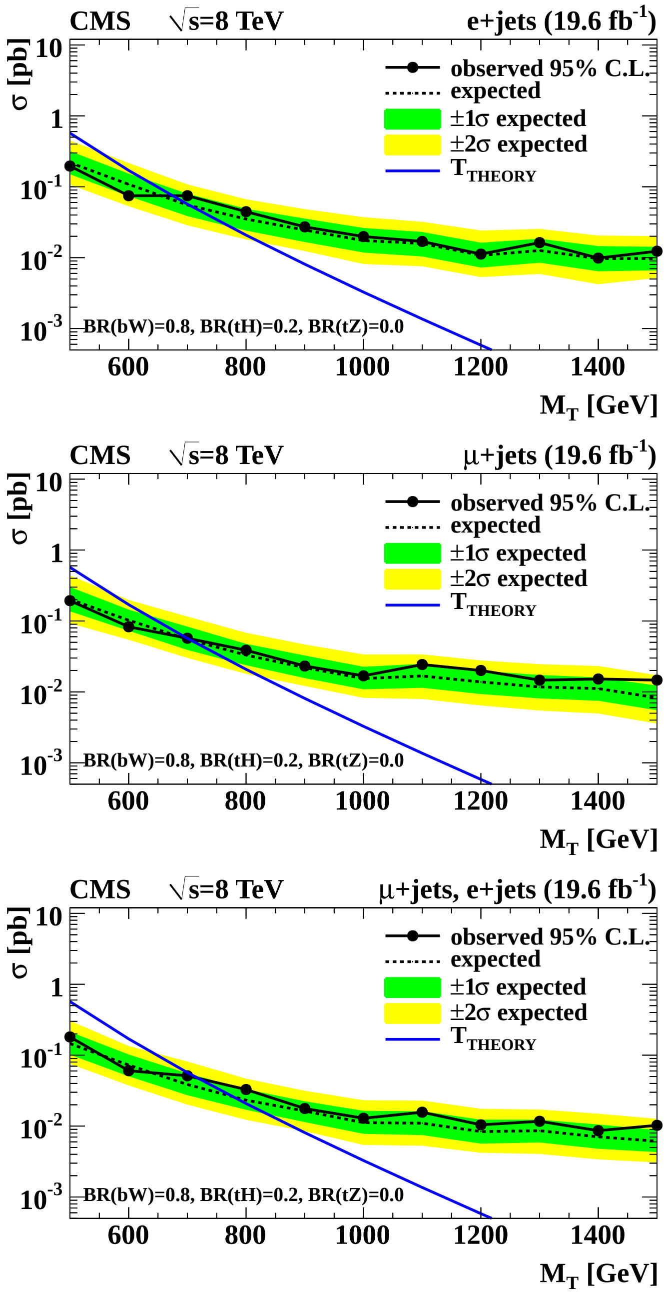


Figure C.38: Observed and expected limits in the (top) e+jets, (middle) μ+jets and combined channel on the $T\overline{T}$ production cross section as a function of T quark mass. The data correspond to 19.6 fb^{-1} for both the e+jets channel and μ+jets channels. The branching ratios BR(bW:tH:tZ) = 0.8 : 0.2 : 0.0 correspond to version 20 nominal listed in Table 11.16

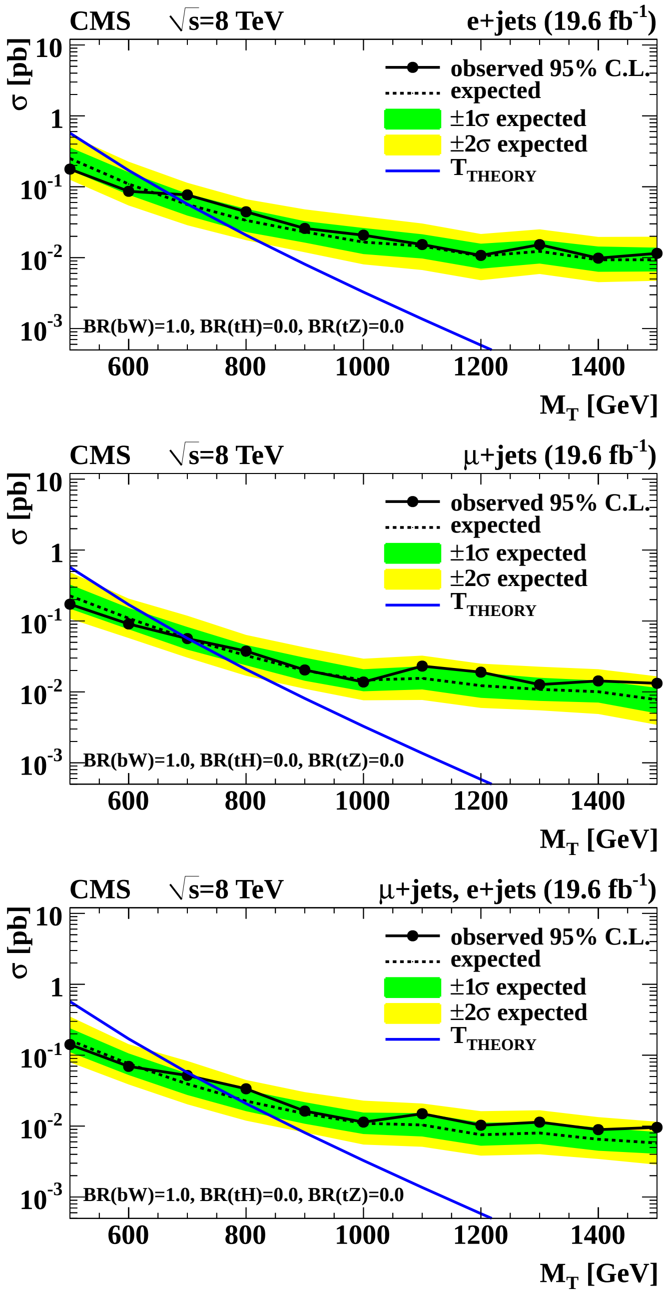


Figure C.39: Observed and expected limits in the (top) e+jets, (middle) μ+jets and combined channel on the $T\overline{T}$ production cross section as a function of T quark mass. The data correspond to $19.6\,\text{fb}^{-1}$ for both the e+jets channel and μ+jets channels. The branching ratios BR$(bW{:}tH{:}tZ) = 1.0 : 0.0 : 0.0$ correspond to version 21 listed in Table 11.16

C.8 Multilepton Channel

The multilepton final state is sorted into four mutually exclusive categories described below. Dilepton events have exactly two leptons with $p_T > 20$ GeV. The dilepton events are further subdivided into opposite and same sign dilepton categories, with the former consisting of two types of samples. Finally, trilepton events have at least three leptons with $p_T > 20$ GeV. Full details of this analysis may be found in [181].

For all categories of multilepton events, we require that there is a dilepton pair with a mass above 20 GeV, $p_T > 30$ GeV, and that at least one jet is identified as a b jet.

In the multilepton case, we redefine H_T and introduce S_T as:

$$
\begin{aligned}
\mathrm{H}_T &\equiv \sum_i p_T \qquad \mathrm{i} = \text{jets} \\
\mathrm{S}_T &\equiv \sum_j p_T \qquad \mathrm{j} = \text{jets, leptons, } \not{E}_T
\end{aligned}
$$

The first of the two opposite sign dilepton samples (denoted OS1) mostly accepts events in which both the T and the $\overline{T}$ quarks decay to bW, resulting in a $bWbW$ final state. The main irreducible background are the $\mathrm{t\bar{t}}$ and Drell-Yan production. To improve sensitivity, we impose the following requirements. We suppress the Drell-Yan background by excluding the dilepton pair mass window around the Z-boson mass, $76 < \mathrm{m}_{ll} < 106$ GeV. Further, the invariant mass of the lepton and b-jet system, m_{lb}, from top quark decays is necessarily smaller than the top quark mass. We suppress the $\mathrm{t\bar{t}}$ background by requiring $\mathrm{m}_{lb} >$ 170 GeV, shown in Fig. C.40. Finally, events are required to have either two or three jets with $\mathrm{H}_T > 300$ GeV, and $\mathrm{S}_T > 900$ GeV. The residual Drell-Yan background is scaled to the observed event count in the Z mass peak.

Events in the second opposite sign dilepton sample are required to have at least five jets (denoted OS2), of which two must be b-tagged. This sample accepts final states in which both leptons come from the decay of a Z boson but is insensitive to the $bWbW$ final state. The selection $\mathrm{H}_T > 500$ GeV, and $\mathrm{S}_T > 1000$ GeV is imposed. In this channel, the dominant background in this channel is $\mathrm{t\bar{t}}$ production.

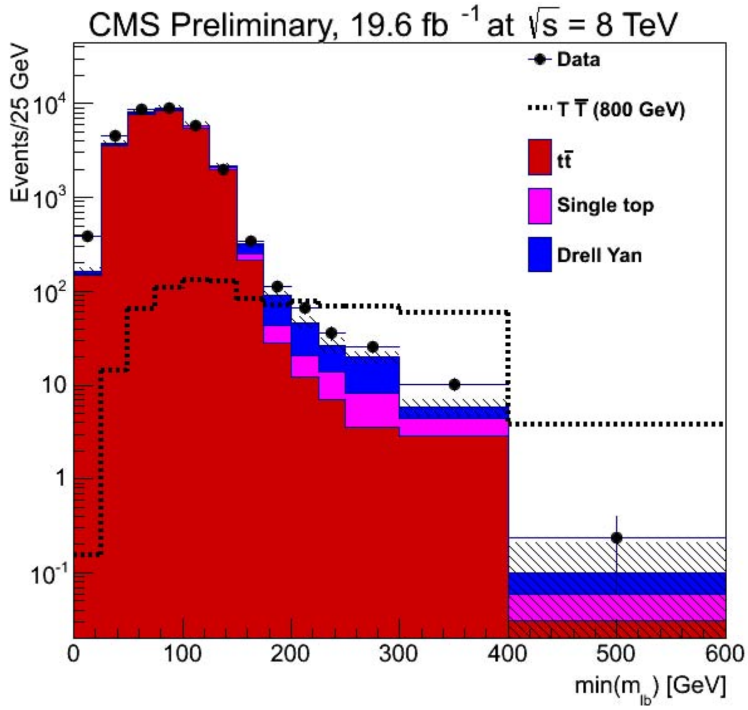


Figure C.40: Observed and expected distributions of the smallest m_{lb} from the opposite sign dilepton sample. The signal distribution is dominated by the $bWbW$ final state for a T quark with a mass of 800 GeV.

The same sign dilepton category is motivated to accept events in which at least one T-quark decaying to either a tZ or tH final state. These events are required to have at least three jets, $\mathrm{H}_T > 500$ GeV, and $\mathrm{S}_T > 700$ GeV, shown in Fig. C.41. In this channel, the main backgrounds fall into three categories. SM prompt backgrounds with same-sign dilepton signatures have very small cross section and are determined from simulation. Events with two prompt opposite-sign leptons can contribute if one lepton has its charge mis-reconstructed. The charge sign mis-reconstruction of muons in the p_T range considered is negligible. The probability to mis-reconstruct the charge of an electron is determined from a simulated sample of Z decays. Finally, we also determine instrumental backgrounds arising from jets faking one or both leptons from control data samples.

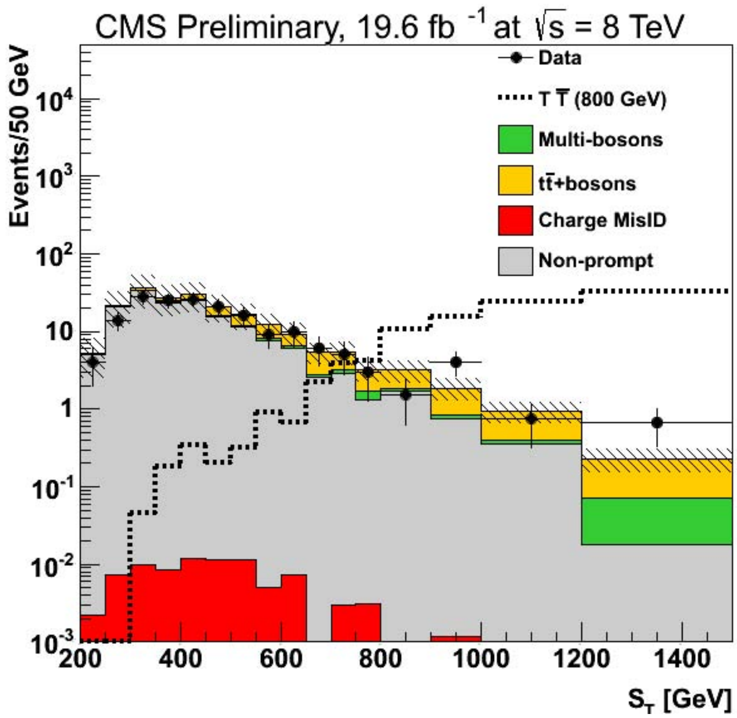


Figure C.41: Observed and expected distributions of S_T from the same sign dilepton sample.

The trilepton event category is also motivated to accept events where at least one T-quark decays to either a tH or tZ final state. These events are required to have at least three leptons, and at least three jets with $\mathrm{H}_T > 500$ GeV and $\mathrm{S}_T > 700$ GeV. The prompt sources of SM background originate from processes with three or more final state leptons; these include diboson and triboson productions, which are modelled by simulation. Non-prompt backgrounds are characterised by one or more misidentified leptons. These are determined in a similar manner to the dilepton event category.

The numbers of events in the simulated background samples and data for all multilepton event categories samples are tabulated in Table C.2. The selection efficiencies and expected numbers of events for the T-quark signal assuming nominal branching fractions are tabulated in Table C.3.

sample	OS1	OS2	SS	Trileptons
$t\bar{t}$	4.75 ± 0.25	75.00 ± 3.70	-	-
single top	2.54 ± 0.31	2.00 ± 0.28	-	-
Z	7.50 ± 1.70	2.47 ± 0.74	-	-
$t\bar{t}$W	-	-	5.80 ± 1.20	0.24 ± 0.05
$t\bar{t}$Z	-	-	1.85 ± 0.58	1.80 ± 0.53
WW	-	-	0.53 ± 0.20	-
WZ	-	-	0.35 ± 0.04	0.40 ± 0.13
ZZ	-	-	0.03 ± 0.00	0.07 ± 0.00
WWW/WWZ/ZZZ/WZZ	-	-	1.85 ± 0.58	0.08 ± 0.03
$t\bar{t}$WW	-	-	-	0.05 ± 0.02
charge mis-ID	-	-	0.01 ± 0.00	-
non-prompt	-	-	8.50 ± 3.20	3.00 ± 1.20
total background	14.8 ± 1.8	79.4 ± 3.8	17.2 ± 3.4	5.7 ± 3.0
data	16	88	18	1

Table C.2: Number events predicted for background processes and observed in collision data in the opposite sign dilepton samples with 2 or 3 jets (OS1) and with at least 5 jets (OS2), the same sign dilepton sample (SS), and the trilepton sample.

channel	OS1		OS2		SS		trileptons	
T quark mass	efficiency	events	efficiency	events	efficiency	events	efficiency	events
500 GeV	0.15%	17.0	0.30%	33.5	0.19%	21.2	0.17%	18.9
600 GeV	0.27%	9.0	0.47%	15.7	0.22%	7.5	0.25%	8.3
700 GeV	0.36%	4.0	0.57%	6.3	0.25%	2.8	0.28%	3.1
800 GeV	0.40%	1.6	0.58%	2.4	0.25%	1.0	0.31%	1.3
900 GeV	0.44%	0.69	0.58%	0.92	0.25%	0.39	0.32%	0.51
1000 GeV	0.45%	0.29	0.54%	0.35	0.23%	0.14	0.33%	0.21
1100 GeV	0.45%	0.12	0.50%	0.13	0.22%	0.06	0.32%	0.09
1200 GeV	0.46%	0.05	0.45%	0.05	0.20%	0.02	0.31 %	0.04

Table C.3: Efficiencies and number of events predicted for the T-quark signal assuming nominal branching fractions in the opposite sign dilepton samples with 2 or 3 jets (OS1) and with at least 5 jets (OS2), the same sign dilepton sample (SS), and the trilepton sample.

In all multilepton event categories, the number of events are of similar magnitudes. Therefore, a simple event count is used in the limit computation. The event count is separated into four separate categories: OS1, OS2, SS, and trileptons. These categories are linearised into a single set of histograms, with each histogram containing all multilepton events for the data, backgrounds or signal samples. These histograms are fed into `theta` and an upper bound on the cross section using all events with multilepton final states is computed at each mass point for 66 different branching ratio scenarios.

Fig. C.42 shows the observed and expected limits determined in this way for a T-quark that decays with the nominal branching ratios.

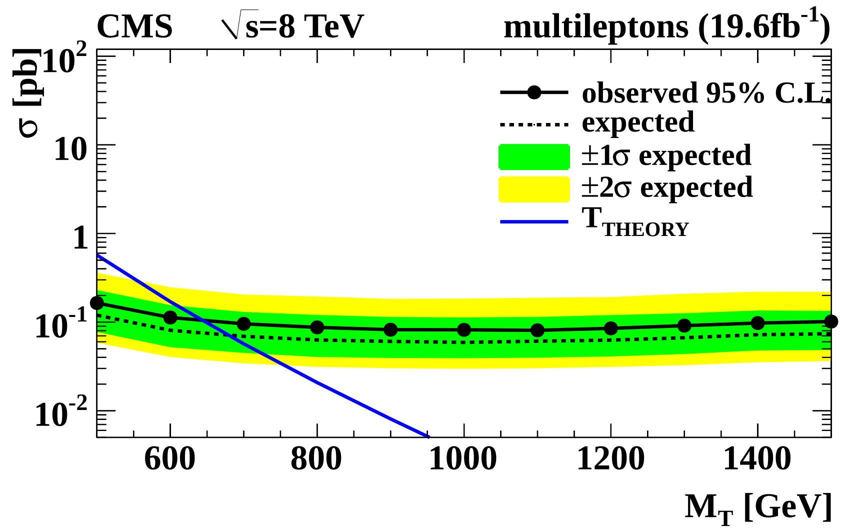


Figure C.42: Observed and expected upper limits for the T-quark production cross section using nominal branching fractions in the multilepton channel at 95% C.L.

The remaining 65 points correspond to a scan across the entire BR phase space in intervals of 0.1. Comparing this to the theoretical cross section of $T\bar{T}$ production, a lower bound on the mass of the T-quark is set at 95% C.L. This information is summarised in Table C.4 for the 22 scenarios that are separated by 0.2 BR intervals. In Section C.9, we will show the exclusion limits for the multilepton analysis using the full 66 BR scenarios.

Out of the multilepton final states, the strongest limit is set in the high tZ BR scenarios as the events in this decay mode have a higher selection efficiency (from the leptonically decaying Z-bosons).

	Branching Fractions			expected	observed
Scenario	T→bW	T→tH	T→tZ	limit	limit
(0)	0.5	0.25	0.25	688 GeV	658 GeV
(1)	0.0	0.0	1.0	802 GeV	781 GeV
(2)	0.0	0.2	0.8	787 GeV	763 GeV
(3)	0.0	0.4	0.6	769 GeV	736 GeV
(4)	0.0	0.6	0.4	739 GeV	697 GeV
(5)	0.0	0.8	0.2	698 GeV	678 GeV
(6)	0.0	1.0	0.0	676 GeV	649 GeV
(7)	0.2	0.0	0.8	783 GeV	757 GeV
(8)	0.2	0.2	0.6	761 GeV	726 GeV
(9)	0.2	0.4	0.4	726 GeV	691 GeV
(10)	0.2	0.6	0.2	689 GeV	663 GeV
(11)	0.2	0.8	0.0	656 GeV	613 GeV
(12)	0.4	0.0	0.6	757 GeV	715 GeV
(13)	0.4	0.2	0.4	715 GeV	685 GeV
(14)	0.4	0.4	0.2	682 GeV	649 GeV
(15)	0.4	0.6	0.0	643 GeV	591 GeV
(16)	0.6	0.0	0.4	715 GeV	684 GeV
(17)	0.6	0.2	0.2	684 GeV	646 GeV
(18)	0.6	0.4	0.0	652 GeV	590 GeV
(19)	0.8	0.0	0.2	692 GeV	660 GeV
(20)	0.8	0.2	0.0	675 GeV	622 GeV
(21)	1.0	0.0	0.0	697 GeV	685 GeV

Table C.4: Shows the expected and observed limits in the multilepton channel for the 22 scenarios separated by 0.2 BR intervals.

C.9 The Search for the Heavy Top-Like Quark in Events with at Least One Lepton

The full likelihoods of the lepton-plus-jets and multilepton channels are statistically combined using `theta` [204]. Upper limits on the cross section for the pair-produced T-quark is computed for each mass point between 500 and 1500 GeV in 100 GeV steps. The full details of this combination can be found in [3].

In the lepton-plus-jets channel a fit to the `BDT` discriminant distributions for signal and background is made to the data at each mass point. For the multilepton channels we fit the predicted numbers of events in four categories to the observed counts. The parameters of the fit are the nuisance parameters that characterise the systematic uncertainties and the $T\overline{T}$ production cross section, which is the sole parameter of interest.

Systematic uncertainties affect both the normalisation and the distribution of the discriminating observable. There is a 4.4% uncertainty in the integrated luminosity for 8 TeV data taken by CMS in 2012. We apply a 10% uncertainty on the $\mathrm{t\bar{t}}$ cross section normalisation and assign a systematic uncertainty of 30% to the diboson backgrounds, single top quark production, and the W and Z-boson backgrounds. An uncertainty of 3% in the normalisation of the signal and background samples are assigned to the e+jets, μ+jets and multilepton channels separately, to account for errors in the lepton trigger and identification scale factors that correct the simulation to agree with the performance observed in the data. An additional 3% uncertainty is applied to account for the uncertainty of finding W-jets in the lepton-plus-jets channel.

Jet energy corrections, jet energy resolution, b-tagging, scale and matching template systematics are common in both the lepton-plus-jets and multilepton channels; these correspond to one standard deviation up and down within the uncertainty margin. These are incorporated into the limit procedure using vertical morphing.

We compute 95% C.L. upper limits using Bayesian statistics with `theta` [204]. We fit all samples simultaneously and then multiply the likelihoods to derive the limits. Fig. C.43 shows the observed and expected limits determined in this way for the nominal signal branching fractions.

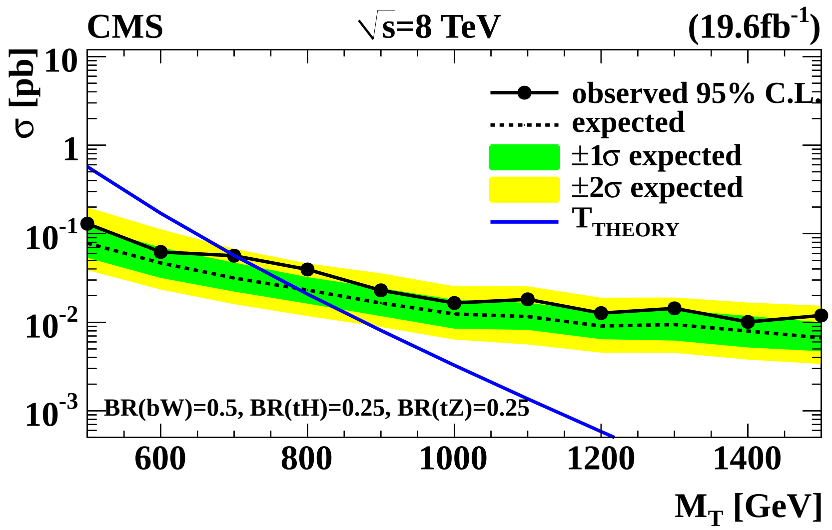


Figure C.43: Observed and expected upper limits for the T-quark production cross section using nominal branching fractions in the combined lepton-plus-jets and multilepton channel at 95% C.L.

We scan across the entire phase space of branching ratios in 0.1 intervals by weighting the contributions of the six final states for the signal samples. At each of these 66 branching ratios (including the nominal scenario) we compute the expected and observed limits; these are tabulated in Table C.5.

Finally, the expected and observed lower limits on the mass of the T-quark are represented graphically for the lepton+jets, multilepton, and combined analyses in Fig. C.44.

We rule out a heavy vector-like T-quark with a mass below 700 GeV across the whole of phase space.

Scenario	Branching Fractions T→bW	T→tH	T→tZ	expected limit	observed limit	Scenario	Branching Fractions T→bW	T→tH	T→tZ	expected limit	observed limit
(0)	0.5	0.25	0.25	791 GeV	702 GeV	(34)	0.3	0.3	0.4	797 GeV	720 GeV
(1)	0.0	0.0	1.0	843 GeV	779 GeV	(35)	0.3	0.4	0.3	795 GeV	713 GeV
(2)	0.0	0.1	0.9	841 GeV	773 GeV	(36)	0.3	0.5	0.2	790 GeV	700 GeV
(3)	0.0	0.2	0.8	829 GeV	768 GeV	(37)	0.3	0.6	0.1	789 GeV	702 GeV
(4)	0.0	0.3	0.7	820 GeV	758 GeV	(38)	0.3	0.7	0.0	790 GeV	712 GeV
(5)	0.0	0.4	0.6	804 GeV	750 GeV	(39)	0.4	0.0	0.6	811 GeV	740 GeV
(6)	0.0	0.5	0.5	798 GeV	746 GeV	(40)	0.4	0.1	0.5	804 GeV	728 GeV
(7)	0.0	0.6	0.4	796 GeV	737 GeV	(41)	0.4	0.2	0.4	798 GeV	721 GeV
(8)	0.0	0.7	0.3	793 GeV	734 GeV	(42)	0.4	0.3	0.3	793 GeV	706 GeV
(9)	0.0	0.8	0.2	788 GeV	719 GeV	(43)	0.4	0.4	0.2	792 GeV	700 GeV
(10)	0.0	0.9	0.1	787 GeV	723 GeV	(44)	0.4	0.5	0.1	788 GeV	699 GeV
(11)	0.0	1.0	0.0	787 GeV	729 GeV	(45)	0.4	0.6	0.0	790 GeV	706 GeV
(12)	0.1	0.0	0.9	840 GeV	773 GeV	(46)	0.5	0.0	0.5	799 GeV	737 GeV
(13)	0.1	0.1	0.8	833 GeV	766 GeV	(47)	0.5	0.1	0.4	796 GeV	720 GeV
(14)	0.1	0.2	0.7	822 GeV	754 GeV	(48)	0.5	0.2	0.3	793 GeV	703 GeV
(15)	0.1	0.3	0.6	810 GeV	747 GeV	(49)	0.5	0.3	0.2	791 GeV	698 GeV
(16)	0.1	0.4	0.5	801 GeV	736 GeV	(50)	0.5	0.4	0.1	791 GeV	698 GeV
(17)	0.1	0.5	0.4	795 GeV	729 GeV	(51)	0.5	0.5	0.0	791 GeV	705 GeV
(18)	0.1	0.6	0.3	793 GeV	723 GeV	(52)	0.6	0.0	0.4	796 GeV	719 GeV
(19)	0.1	0.7	0.2	788 GeV	716 GeV	(53)	0.6	0.1	0.3	793 GeV	706 GeV
(20)	0.1	0.8	0.1	789 GeV	711 GeV	(54)	0.6	0.2	0.2	791 GeV	705 GeV
(21)	0.1	0.9	0.0	787 GeV	720 GeV	(55)	0.6	0.3	0.1	792 GeV	699 GeV
(22)	0.2	0.0	0.8	831 GeV	763 GeV	(56)	0.6	0.4	0.0	792 GeV	707 GeV
(23)	0.2	0.1	0.7	821 GeV	754 GeV	(57)	0.7	0.0	0.3	795 GeV	716 GeV
(24)	0.2	0.2	0.6	811 GeV	743 GeV	(58)	0.7	0.1	0.2	796 GeV	706 GeV
(25)	0.2	0.3	0.5	801 GeV	733 GeV	(59)	0.7	0.2	0.1	795 GeV	703 GeV
(26)	0.2	0.4	0.4	797 GeV	722 GeV	(60)	0.7	0.3	0.0	795 GeV	720 GeV
(27)	0.2	0.5	0.3	792 GeV	718 GeV	(61)	0.8	0.0	0.2	797 GeV	709 GeV
(28)	0.2	0.6	0.2	791 GeV	709 GeV	(62)	0.8	0.1	0.1	797 GeV	716 GeV
(29)	0.2	0.7	0.1	788 GeV	711 GeV	(63)	0.8	0.2	0.0	798 GeV	734 GeV
(30)	0.2	0.8	0.0	787 GeV	715 GeV	(64)	0.9	0.0	0.1	798 GeV	723 GeV
(31)	0.3	0.0	0.7	824 GeV	753 GeV	(65)	0.9	0.1	0.0	799 GeV	738 GeV
(32)	0.3	0.1	0.6	810 GeV	744 GeV	(66)	1.0	0.0	0.0	803 GeV	749 GeV
(33)	0.3	0.2	0.5	799 GeV	732 GeV						

Table C.5: Expected and observed mass limits on the mass of the T-quark in the combined lepton-plus-jets and multilepton channel for all points in phase space where the limit computation is performed.

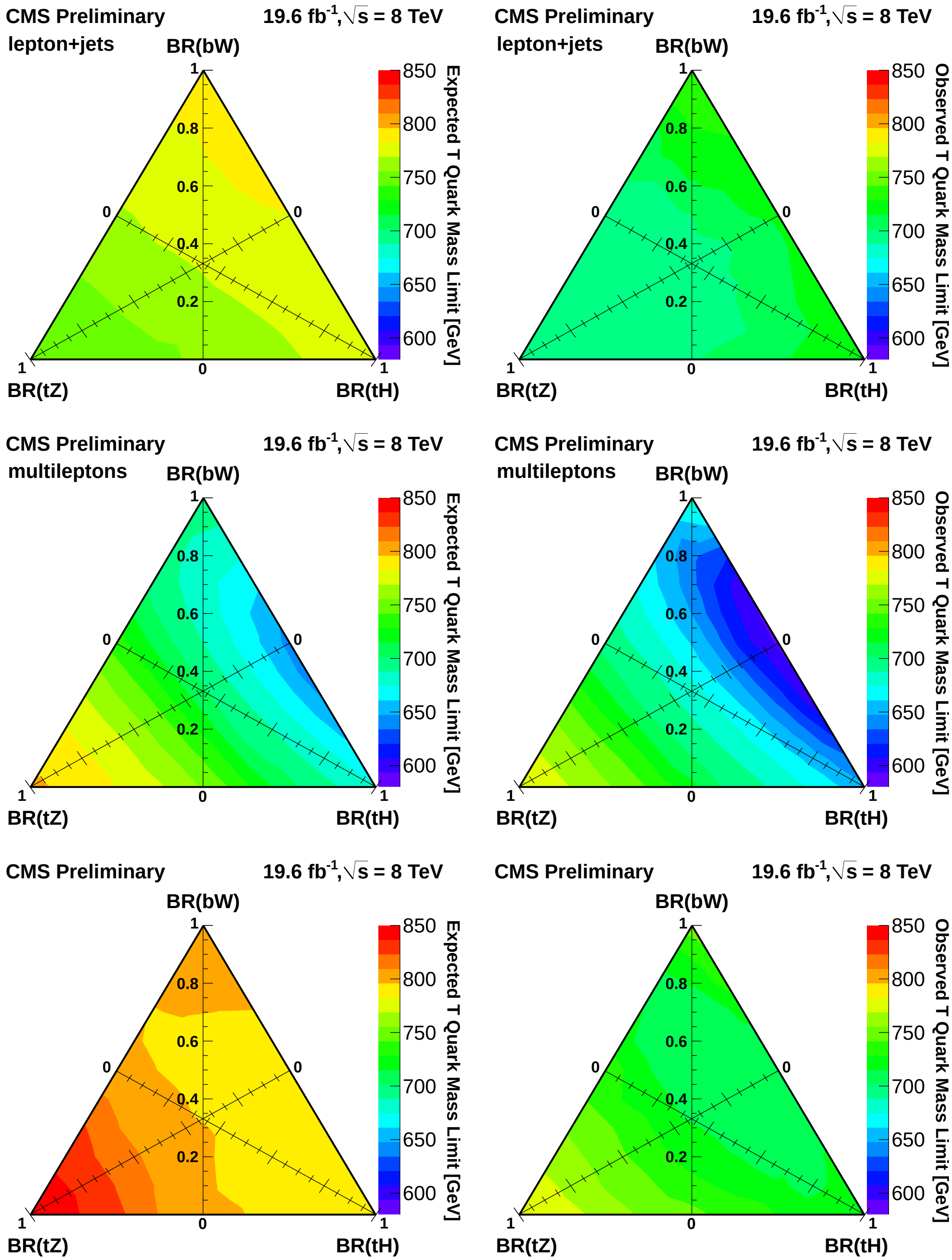


Figure C.44: Expected and observed mass limits in the (top) lepton+jets, (middle) multilepton, and (bottom) the combined channels on the $T\bar{T}$ production. These plots are populated by the mass limits set in each of the 66 (including the nominal) BR scenarios separated by 0.1 BR interval, with information summarised in Table 11.16, Table C.4, and Table C.5.

Bibliography

[1] CMS Collaboration. Search for pair produced fourth-generation up-type quarks in pp collisions at with a lepton in the final state. *Physics Letters B*, 718(2):307 – 328, 2012.

[2] Michael Luk et al. The inclusive Search for a Heavy Top-Like Quark T in the Lepton-plus-Jets Channel at CMS. AN-13-028, March 2013.

[3] CMS Collaboration. Inclusive search for a vector-like T quark in events with at least one lepton at $\sqrt{s} = 8$ TeV. B2G-12-015, March 2013.

[4] J. Beringer et al. The Review of Particle Physics. *Physical Review D*, 86, 2012.

[5] Jan Gutowski. Symmetry and Particle Physics 2006-2007 (Lecture Notes). Department of Mathematics and Theoretical Physics, Cambridge University.

[6] David Tong. Quantum Field Theory (Lecture Notes). Department of Mathematics and Theoretical Physics, Cambridge University, 2006-2007.

[7] Xuan-Yem Pham Quang Ho-Kim. *Elementary Particles and Their Interactions.* Springer, 1998.

[8] Dan V. Shroeder Michael E. Peskin. *An Introduction To Quantum Field Theory.* Westview Press, 1995.

[9] B.C. Allanach. The Standard Model (Lecture Notes). Department of Mathematics and Theoretical Physics, Cambridge University, 2006-2007.

[10] CMS Collaboration. *CMS Physics Technical Design Report Volume I: Detector Performance and Software*. Technical Design Report CMS. CERN, Geneva, 2006.

[11] CMS Collaboration. *CMS Physics Technical Design Report Volume II: Physics Performance*. Technical Design Report CMS. CERN, Geneva, 2006.

[12] CMS Collaboration. The CMS experiment at the CERN LHC. *JINST*, 3, 2008.

[13] Enrico Fermi. Tentativo di una Teoria Dei Raggi β. *Il Nuovo Cimento*, 11(1):1–19, 1934.

[14] Abdus Salam and J.C. Ward. Weak and electromagnetic interactions. *Il Nuovo Cimento Series 10*, 11(4):568–577, 1959.

[15] John Archibald Wheeler Charles W. Misner, Kip S. Thorne. *Gravitation*. W. H. Freeman, 1973.

[16] David Griffits. *Introduction to Elementary Particles*. Wiley & Sons, USA, 1987.

[17] Mary Boas. *Mathematical Methods in the Physical Sciences*. Wiley & Sons, 3 edition, 2005.

[18] J. Clerk Maxwell. A Dynamical Theory of the Electromagnetic Field. *Phil. Trans. R. Soc. Lond.*, 155:459–512, January 1865.

[19] Erwin Schrödinger. Quantisierung als Eigenwertproblem. *Annalen der Physik*, 386(18):109–139, 1926.

[20] Alastair Rae. *Quanutm Mechanics*. Taylor & Francis, 2007.

[21] Lewis H. Ryder. *Quantum Field Theory*. Cambridge University Press, 2 edition, June 1996.

[22] Walther Gerlach and Otto Stern. Das magnetische Moment des Silberatoms. *Zeitschrift für Physik A Hadrons and Nuclei*, 9(1):353–355, December 1922.

[23] M. Goldhaber, L. Grodzins, and A. W. Sunyar. Helicity of Neutrinos. *Phys. Rev.*, 109:1015–1017, Feb 1958.

[24] Thomas Young. *Experimental Demonstration of the General Law of the Interference of Light*, volume 94. Philosophical Transactions of the Royal Society of London, 1804.

[25] Niels Bohr. The Quantum Postulate and the Recent Development of Atomic Theory. *Nature*, 121:580–590, 1928.

[26] R. P. Feynman. Space-Time Approach to Non-Relativistic Quantum Mechanics. *Rev. Mod. Phys.*, 20:367–387, Apr 1948.

[27] Hugh Osborn. Quantum Field Theory II (Lecture Notes). Department of Mathematics and Theoretical Physics, Cambridge University, 2006-2007.

[28] Werner Heisenberg. Ueber den anschaulichen Inhalt der quantentheoretischen Kinematik und Mechanik. *European Physical Journal A*, 43(3):172–198, 1927.

[29] E. H. Kennard. Zur Quantenmechanik einfacher Bewegungstypen. *Zeitschrift für Physik*, 44(4-5):326, 1927.

[30] Zenios Avanis and Michael Luk. Black Holes in General Relativity. March 2007.

[31] Tanya Kuzmenko Moshe Carmel. Value of the Cosmological Constant in the Cosmological Relativity Theory. *International Journal of Theoretical Physics*, 41(1):131–135, 2002.

[32] S. Perlmutter et al. Measurements of Ω and Λ from 42 High-Redshift Supernovae. *Astrophysics Journal*, 517(565), 1999.

[33] Murray Gell-Mann. The Eightfold Way: A Theory of Strong Interaction Symmetry. California Instituyde of Technology Laboratory, March 1961.

[34] G Zweig. An SU_3 model for strong interaction symmetry and its breaking; Part I. (CERN-TH-401):24 p, Jan 1964.

[35] Wikipedia: The free encyclopedia, April 2013.

[36] Albert Einstein. Lichtgeschwindigkeit und Statik des Gravitationsfeldes - The Speed of Light and the Statics of the Gravitational Field. *Annalen der Physik*, 38, 1912.

[37] Albert Einstein. Theorie des statischen Gravitationsfeldes - On the Theory of the Static Gravitational Field. *Annalen der Physik*, 38:443–458, 1912.

[38] S.L. Glashow. Partial Symmetries of Weak Interactions. *Nucl.Phys.*, 22:579–588, 1961.

[39] Steven Weinberg. A Model of Leptons. *Phys. Rev. Lett.*, 19:1264–1266, Nov 1967.

[40] F. Englert and R. Brout. Broken Symmetry and the Mass of Gauge Vector Mesons. *Phys. Rev. Lett.*, 13:321–323, Aug 1964.

[41] Guralnik, G. S. and Hagen, C. R. and Kibble, T. W. B. Global conservation laws and massless particles. *Phys. Rev. Lett.*, 13:585–587, Nov 1964.

[42] Peter W. Higgs. Broken Symmetries and the Masses of Gauge Bosons. *Phys. Rev. Lett.*, 13:508–509, Oct 1964.

[43] CMS Collaboration. Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC. *Physics Letters B*, 716(1):30–61, September 2012.

[44] Albert Einstein. Über einen die Erzeugung und Verwandlung des Lichtes betreffenden heuristischen Gesichtspunkt. *Annalen der Physik*, 17(6):132–148, 1905.

[45] The UA1 Collaboration. Experimental observation of lepton pairs of invariant mass around 95 GeV/c^2 at the CERN SPS collider. *Phys. Lett. B*, 126(CERN-EP-83-073):398–410. 17 p, Jun 1985.

[46] UA2 Collaboration. Evidence for Z $\rightarrow$ ee at the CERN pp Collider. *Physics Letters. Section B: Nuclear, Elementary Particle and High-Energy Physics*, 129(1-2):130–140, 1983.

[47] The UA1 Collaboration. Experimental observation of isolated large transverse energy electrons with associated missing energy at $\sqrt{s} = 540$ GeV. *Physics Letters B*, 122(1):103 – 116, 1983.

[48] The UA2 Collaboration. Observation of single isolated electrons of high transverse momentum in events with missing transverse energy at the CERN pp collider. *Physics Letters B*, 122:476 – 485, 1983.

[49] PLUTO Collaboration. Evidence for gluon bremsstrahlung in *ee* annihilations at high energies. *Physics Letters B*, 86:418 – 425, 1979.

[50] Steven Weinberg. *The Quantum Theory of Fields : Foundations*, volume 1. Cambridge University Press, 2005.

[51] Chen-Ning Yang and Robert L. Mills. Conservation of Isotopic Spin and Isotopic Gauge Invariance. *Phys.Rev.*, 96:191–195, 1954.

[52] T. Muta. *Foundations of Quantum Chromodynamics*, volume 57. World Scientific, 2000.

[53] L.D. Faddeev and V.N. Popov. Feynman diagrams for the Yang-Mills field. *Physics Letters B*, 25(1):29 – 30, 1967.

[54] Morton Hamermesh. *Group Theory and Its Application to Physical Problems*. Dover Publications Inc., 1989.

[55] F.J. Hasert et al. Observation of Neutrino Like Interactions Without Muon Or Electron in the Gargamelle Neutrino Experiment. *Phys.Lett.*, B46:138–140, 1973.

[56] F.J. Hasert et al. Observation of Neutrino Like Interactions without Muon or Electron in the Gargamelle Neutrino Experiment. *Nucl.Phys.*, B73:1–22, 1974.

[57] T. D. Lee and C. N. Yang. Question of Parity Conservation in Weak Interactions. *Phys. Rev.*, 104:254–258, Oct 1956.

[58] C. S. Wu, E. Ambler, R. W. Hayward, D. D. Hoppes, and R. P. Hudson. Experimental Test of Parity Conservation in Beta Decay. *Phys. Rev.*, 105:1413–1415, Feb 1957.

[59] G.'t Hooft. Renormalizable Lagrangians for massive Yang-Mills fields. *Nuclear Physics B*, 35(1):167 – 188, 1971.

[60] G. 't Hooft and M. Veltman. Regularization and renormalization of gauge fields. *Nuclear Physics B*, 44(1):189 – 213, 1972.

[61] Yoichiro Nambu. Quasi-Particles and Gauge Invariance in the Theory of Superconductivity. *Phys. Rev.*, 117:648–663, Feb 1960.

[62] Jeffrey Goldstone, Abdus Salam, and Steven Weinberg. Broken Symmetries. *Phys.Rev.*, 127:965–970, 1962.

[63] Sidney R. Coleman and Erick J. Weinberg. Radiative Corrections as the Origin of Spontaneous Symmetry Breaking. *Phys.Rev.*, D7:1888–1910, 1973.

[64] Francis Halzen and Alan Martin. *Quarks and Leptons: An Introductory Course in Modern Particle Physics.* Wiley & Sons, USA, 1984.

[65] C et al. Amsler. Review of Particle Physics*. *Physics Letters B*, 667:1–1340, 2008.

[66] Paul Villard. Sur la Réflexion et la Réfraction des Rayons Cathodiques et des Rayons Déviables du Radium. *Comptes Redus de l'Académie des Science*, 130(1010), 1900.

[67] The ATLAS Collaboration. Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC. *Physics Letters B*, 716(1):1 – 29, 2012.

[68] C.S. Wu et al. Experimental Test of Parity Conservation in Beta Decay. *Physical Review*, 105(4), 1957.

[69] J.J. Thomson. Cathode Rays. *Philosophical Magazine*, 44:293, 1897.

[70] Jr. F. Reines, C.L. Cowan. The Neutrino. *Nature*, 4531:178, 1956.

[71] S. H. Neddermeyer and C. D. Anderson. Note on the Nature of Cosmic-Ray Particles. *Physical Review*, 51:884–886, May 1937.

[72] G. Danby et al. Observation of high-energy neutrino reactions and the existence of two kinds of neutrinos. *Physical Review Letters*, 9:36, 1962.

[73] M. L. Perl et al. Evidence for Anomalous Lepton Production in *ee* Annihilation. *Physical Review Letters*, 35:22, 1975.

[74] DONUT Collaboration. Observation of tau neutrino interactions. *Physics Letters B*, 504:3, 2001.

[75] SLAC Collaboration. High-Energy Inelastic e p Scattering at 6-Degrees and 10-Degrees. *Phys.Rev.Lett.*, 23:930–934, 1969.

[76] SLAC Collaboration. Observed Behavior of Highly Inelastic electron-Proton Scattering. *Phys.Rev.Lett.*, 23:935–939, 1969.

[77] SLAC Collaboration. Discovery of a Narrow Resonance in e+ e- Annihilation. *Phys.Rev.Lett.*, 33:1406–1408, 1974.

[78] BNL Collaboration. Experimental Observation of a Heavy Particle J. *Phys.Rev.Lett.*, 33:1404–1406, 1974.

[79] Fermilab E288. Observation of a Dimuon Resonance at 9.5-GeV in 400-GeV Proton-Nucleus Collisions. *Phys.Rev.Lett.*, 39:252–255, 1977.

[80] D0 Collaboration. Observation of top quark production in $\bar{p}p$ collisions. *Phys.Rev.Lett.*, 74:2626–2631, 1995.

[81] Gary Steigman. When clusters collide: constraints on antimatter on the largest scales. *Journal of Cosmology and Astroparticle Physics*, 2008(10):001, 2008.

[82] A.D. Sakharov. Violation of CP Invariance, C Asymmetry, and Baryon Asymmetry of the Universe. *Pisma Zh.Eksp.Teor.Fiz.*, 5:32–35, 1967.

[83] M. Fukugita and T. Yanagida. Baryogenesis Without Grand Unification. *Phys. Lett.*, B174:45, 1986.

[84] E.C.G. Stueckelberg. Remarks on the creation of pairs of particles in the theory of relativity. *Helv.Phys.Acta*, 14:588–594, 1941.

[85] Nicola Cabibbo. Unitary symmetry and leptonic decays. *Physical Review Letters*, 10:531–533, 1963.

[86] J.H. Christenson et al. Evidence for the 2pi Decay of the K0 Meson. *Physical Review Letters*, 13(138), 1964.

[87] T. Maskawa M. Kobayashi. CP Violation in the Renormalizable Theory of Weak Interaction. *Progress of Theoretical Physics*, 49(2):652–657, 1973.

[88] Gregory Gabadadze and M. Shifman. QCD vacuum and axions: What's happening? *Int.J.Mod.Phys.*, A17:3689–3728, 2002.

[89] T. Das, G. S. Guralnik, V. S. Mathur, F. E. Low, and J. E. Young. Electromagnetic mass difference of pions. *Phys. Rev. Lett.*, 18:759–761, May 1967.

[90] F. Capra. Current algebra and k+-k0 mass difference. *Il Nuovo Cimento A*, 54(2):479–486, 1968.

[91] Roger F. Dashen. Chiral SU(3) x SU(3) as a symmetry of the strong interactions. *Phys.Rev.*, 183:1245–1260, 1969.

[92] John F. Donoghue, Barry R. Holstein, and Daniel Wyler. Electromagnetic selfenergies of pseudoscalar mesons and Dashen's theorem. *Phys.Rev.*, D47:2089–2097, 1993.

[93] Johan Bijnens. Violations of Dashen's theorem. *Phys.Lett.*, B306:343–349, 1993.

[94] Kim Maltman and David Kotchan. Chiral Log Corrections to Pseudoscalar Electromagnetic Self-Energies and the Reliability of Dashen's Theorem. *Mod.Phys.Lett.*, A5:2457–2464, 1990.

[95] Robert Baur and Res Urech. On the corrections to Dashen's theorem. *Phys.Rev.*, D53:6552–6557, 1996.

[96] Christine Davies. Lattice QCD. pages 105–146, 2002.

[97] F. Wilczek D.J. Gross. Ultraviolet behavior of non-abelian gauge theories. *Physical Review Letters*, 30(26):1343–1346, 1973.

[98] H.D. Politzer. Reliable perturbative results for strong interactions. *Physical Review Letters*, 30(26):1346–1349, 1973.

[99] Daniel Treille. Searches at LEP. *Physics Reports*, (0):221 – 239, 2004.

[100] Fernando Quevedo. Supersymmetry and Extra Dimension (Lecture Notes). Department of Mathematics and Theoretical Physics, Cambridge University, 2006-2007.

[101] J. H. Schwarz M. B. Green and E. Witten. *Superstring theory*, volume I & II. Cambridge University Press, 1988.

[102] M. B. Green. String Theory (Lecture Notes). Department of Mathematics and Theoretical Physics, Cambridge University, 2006-2007.

[103] Carlo Rovelli. *Quantum Gravity.* Cambridge University Press, 2007.

[104] Jr Davis, Raymon. Solar neutrinos ii, experimental. *Physical Review Letters*, 12, 1964.

[105] P.A.R. Ade et al. Planck 2013 results. XVI. Cosmological parameters. 2013.

[106] Joshua Frieman, Michael Turner, and Dragan Huterer. Dark Energy and the Accelerating Universe. *Ann.Rev.Astron.Astrophys.*, 46:385–432, 2008.

[107] Adam G. Riess et al. Observational evidence from supernovae for an accelerating universe and a cosmological constant. *Astron.J.*, 116:1009–1038, 1998.

[108] Ivaylo Zlatev, Li-Min Wang, and Paul J. Steinhardt. Quintessence, cosmic coincidence, and the cosmological constant. *Phys.Rev.Lett.*, 82:896–899, 1999.

[109] W. Ford Vera C. Rubin. Rotation of the Andromeda Nebula from a Spectroscopic Survey of Emission Regions. *Astrophysics Journal*, 159(02):379, 1970.

[110] H. Babcock. *The rotation of the Andromeda Nebula.* PhD thesis, University of California, Berkeley, 1940.

[111] Lyndon Evans and Philip Bryant. LHC Machine. *Journal of Instrumentation*, 3(08):S08001, 2008.

[112] *LHC Design Report.* CERN, Geneva, 2004.

[113] CERN Communications Group. CERN FAQ: LHC the Guide. *CERN-Brochure-2008-001-Eng*, January 2008.

[114] Claudia-Elisabeth Wulz. The CMS experiment at CERN. *SPIE*, 2005.

[115] CMS Collaboration. The CMS experiment at the CERN LHC. *JINST*, 2008.

[116] CMS Collaboration. Precise mapping of the magnetic field in the CMS barrel yoke using cosmic rays. *Journal of Instrumentation*, 5:3021, March 2010.

[117] The CMS Collaboration. CMS Technical Design Report for the Phase 1 Upgrade of the Hadron Calorimeter. Technical report, CMS, 2012.

[118] CMS Collaboration. Design, Performance, and Calibration of CMS Hadron Endcap Calorimeters. *CMS NOTE*, 2008.

[119] S Chatrchyan et al. Calibration of the CMS Drift Tube Chambers and Measurement of the Drift Velocity with Cosmic Rays. *JINST*, 5:T03016, 2010.

[120] CMS Collaboration. Vertex fitting with the Kalman filter formalism in the ORCA reconstruction program. CMS IN 2003/008, February 2003.

[121] Aruna Kumar Nayak. Reconstruction of physics objects in the CMS detector. Technical Report CMS-CR-2013-019. CERN-CMS-CR-2013-019, CERN, Geneva, Jan 2013.

[122] W Erdmann. Vertex reconstruction at the CMS experiment. *J. Phys, IOPscience*, 2008.

[123] CMS Collaboration. Particle-Flow Event Reconstruction in CMS and Performance for Jets, Taus, and MET. Apr 2009.

[124] CMS Collaboration. Electron reconstruction and identification at sqrt(s) = 7 TeV. *CMS-PAS*, (CMS-DP-2010-032), 2010.

[125] CMS Collaboration. Reconstruction of Electrons with Gaussian-Sum Filter in the CMS Tracker at the LHC. *CMS-Note*, CMS-NOTE-2005-001, 2005.

[126] CMS Collaboration. https://twiki.cern.ch/twiki/bin/view/Main/EGammaScaleFactors2012.

[127] A. Everett CMS Collaboration. Muon reconstruction and identification in cms. *AIP Conference Proceedings*, 1200(1):701–704, 2010.

[128] CMS Collaboration. Performance of τ-lepton reconstruction and identification in CMS. *JINST*, P01001, 2012.

[129] CMS Collaboration. https://twiki.cern.ch/twiki/bin/view/CMS/B2GPublicPlots.

[130] Matteo Cacciari and Gavin P. Salam. The anti-k_t jet clustering algorithm. *JHEP*, 063, 2008.

[131] Yu. L. Dokshitzer et al. Better Jet Clustering Algorithms. *JHEP*, 001, 1998.

[132] The CMS Collaboration. A Cambridge-Aachen (CA) based Jet Algorithm for boosted top-jet tagging. *CMS Physics Analysis Summary*, JME-09-001, 2009.

[133] G.P. Salam M. Cacciari and G. Soyez. FastJet User Manual. *Eur.Phys.J.*, C72, 2012.

[134] CMS Collaboration. Jet Energy Corrections determination at 7 TeV. *CMS PAS*, JME-10-010, 2010.

[135] S. Banerjee V. Lefbure and I. Gonzlez. CMS Simulation Software Using Geant4. *Technical Report*, CMS-NOTE-1999-072, 1999.

[136] CMS Collaboration. Algorithms for b Jet Identification in CMS. *CMS Physics Analysis Summary*, BTV-09-001, 2009.

[137] CMS Collaboration. Study of Jet Substructure in pp Collisions at 7 TeV in CMS. *Physics Analysis Summary CMS-PAS-JME-10-013*, 2010.

[138] CMS Collaboration. Search for BSM $t\bar{t}$ Production in the Boosted All-Hadronic Final State using pp Collisions at $\sqrt{s} = 8$TeV. *Physics Analysis Summary CMS-PAS-B2G-12-005*, 2012.

[139] Measurement of CMS Luminosity. Technical Report CMS-PAS-EWK-10-004, CERN, 1900. Geneva, 1900.

[140] CMS Collaboration. https://twiki.cern.ch/twiki/bin/view/CMSPublic/LumiPublicResults.

[141] Taipei, Second Workshop on Beyond 3 Generation Standard Model. *Constraints on CKM matrix elements within a fourth generation*, January 2010.

[142] H. Lacker and A. Menzel. PSimultaneous Extraction of the Fermi constant and PMNS matrix elements in the presence of a fourth generation. *HU-EP*, 10(http://arxiv.org/abs/1003.4532), 2010.

[143] A. Menzel. Constraints on CKM matrix elements of the Quark and Neutrino Mixing Matrix with a Four Generation Standard Model with CKMfitter. Master's thesis, Humboldt University, 2010.

[144] J. Charles et al. CP violation and the CKM matrix; assessing the impact of the asymmetric B factories. *The European Physical Journal C - Particles and Fields*, 41(1-131), 2005.

[145] Dennis Wendland. Search for a signal of quarks of a fourth generation in same-sign final states with ATLAS. Master's thesis, Humboldt University, 2010.

[146] Otto Eberhardt, Alexander Lenz, and Jurgen Rohrwild. Less space for a new family of fermions. *Phys. Rev.*, D82:095006, 2010.

[147] Amarjit Soni, Ashutosh Kumar Alok, Anjan Giri, Rukmani Mohanta, and Soumitra Nandi. SM with four generations: Selected implications for rare B and K decays. *Phys. Rev.*, D82:033009, 2010.

[148] Graham D. Kribs, Tilman Plehn, Michael Spannowsky, and Timothy M.P. Tait. Four generations and Higgs physics. *Phys. Rev.*, D76:075016, 2007.

[149] M. Aliev et al. HATHOR - HAdronic Top and Heavy quarks crOss section calculatoR. *Comput. Phys. Commun.*, 2011.

[150] T T D Choudhury and C. Wagner. Beautiful Mirrors and Precision Electroweak Data. *Physical Review D*, 2002.

[151] M Schmaltz. Physics beyond the standard model (Theory): Introducing the Little Higgs. *Nuclear Physics B - Proceedings Supplements 117 (2003) 40 – 49. 31st International Conferences on High Energy Physics.*, 2003.

[152] Fabio Maltoni and Tim Stelzer. MadEvent: Automatic event generation with MadGraph. *JHEP*, 02:027, 2003.

[153] D0 Collaboration. Measurement of the top quark mass in the lepton+jets channel using the ideogram method. *Phys. Rev. D*, 75:092001, May 2007.

[154] M. N. Rosenbluth et al. N. Metropolis, A. W. Rosenbluth. Equation of state calculations by fast computing machines. *J. Chem. Phys.*, 1953.

[155] W K Hastings. Monte Carlo Sampling Methods Using Markov Chains and Their Applications. *Biometrika*, 1970.

[156] L Moneta et al. K Kranmer, G Schott. Roostats user's guide v0. 2010.

[157] I. Hinchliffe M. Chanowitz, M. Furman. Weak interactions of ultraheavy fermions. *Phys. Lett. B*, 78(285), 1978.

[158] Martin Schmaltz and David Tucker-Smith. Little Higgs review. *Ann.Rev.Nucl.Part.Sci.*, 55:229–270, 2005.

[159] Maxim Perelstein. Little Higgs models and their phenomenology. *Prog.Part.Nucl.Phys.*, 58:247–291, 2007.

[160] Tao Han, Heather E. Logan, Bob McElrath, and Lian-Tao Wang. Phenomenology of the little Higgs model. *Phys.Rev.*, D67:095004, 2003.

[161] Martin Schmaltz. The Simplest little Higgs. *JHEP*, 0408:056, 2004.

[162] Tao Han, Heather E. Logan, and Lian-Tao Wang. Smoking-gun signatures of little Higgs models. *JHEP*, 0601:099, 2006.

[163] Can Kilic and Rakhi Mahbubani. Precision electroweak observables in the minimal moose little Higgs model. *JHEP*, 0407:013, 2004.

[164] Mu-Chun Chen. Models of Little Higgs and Electroweak Precision Tests. *Modern Physics Letters A*, 21(8):621–637, 2006.

[165] Maxim Perelstein, Michael E. Peskin, and Aaron Pierce. Top quarks and electroweak symmetry breaking in little Higgs models. *Phys.Rev.*, D69:075002, 2004.

[166] H J He, Kuang Yu Ping, and C P Yuan. Longitudinal/Goldstone boson equivalence and phenomenology of probing the electroweak symmetry breaking sector. (hep-ph/9507454. MSU-HEP-50712. TUIMP-TH-95-67. VPI-IHEP-95-12):6 p, Jul 1995.

[167] Kaustubh Agashe, Roberto Contino, Leandro Da Rold, and Alex Pomarol. A Custodial symmetry for Zb anti-b. *Phys.Lett.*, B641:62–66, 2006.

[168] Roberto Contino, Yasunori Nomura, and Alex Pomarol. Higgs as a holographic pseudoGoldstone boson. *Nucl.Phys.*, B671:148–174, 2003.

[169] Andrzej J. Buras, Bjoern Duling, and Stefania Gori. The Impact of Kaluza-Klein Fermions on Standard Model Fermion Couplings in a RS Model with Custodial Protection. *JHEP*, 0909:076, 2009.

[170] Roberto Contino, Thomas Kramer, Minho Son, and Raman Sundrum. Warped/composite phenomenology simplified. *JHEP*, 0705:074, 2007.

[171] Hael Collins, Aaron K. Grant, and Howard Georgi. The phenomenology of a top quark seesaw model. *Phys.Rev.*, D61:055002, 2000.

[172] R. Sekhar Chivukula, Bogdan A. Dobrescu, Howard Georgi, and Christopher T. Hill. Top quark seesaw theory of electroweak symmetry breaking. *Phys.Rev.*, D59:075003, 1999.

[173] Bogdan A. Dobrescu and Christopher T. Hill. Electroweak symmetry breaking via top condensation seesaw. *Phys.Rev.Lett.*, 81:2634–2637, 1998.

[174] Kenneth D. Lane. An Introduction to technicolor. 1993.

[175] Mfv susy: A natural theory for r-parity violation. 2011.

[176] E. Masso J. Bagger, S. Dimopoulos. Heavy families: Masses and mixings. *Nucl.Phys.B.*, 253(397), 1985.

[177] Giacomo Cacciapaglia, Aldo Deandrea, Daisuke Harada, and Yasuhiro Okada. Bounds and Decays of New Heavy Vector-like Top Partners. *JHEP*, 1011:159, 2010.

[178] J.A. Aguilar-Saavedra. Identifying top partners at LHC. *JHEP*, 0911:030, 2009.

[179] Roberto Franceschini Robert Contino and Michael Peskin. A module for vectorlike heavy top quarks. Nov 2010.

[180] Kanishka Rao and Daniel Whiteson. Triangulating an exotic T quark. *Phys.Rev.*, D86:015008, 2012.

[181] Aram Avetisyan et al. Inclusive Search for Heavy Vector-Like Top Quarks in Multiple-Lepton Final States at $\sqrt{s}$=8 TeV. AN-13-006, March 2013.

[182] S. Mrenna S. Sjostrand and P. Z. Skands. A Brief Introduction to PYTHIA 8.1. *Comput.Phys.Commun.*, 178:852–867, 2008.

[183] John Allison et al. Geant4 developments and applications. *IEEE Trans. Nucl. Sci.*, 53:270, 2006.

[184] J.M. Campbell and R.K. Ellis. MCFM for the TeVatron and the LHC. *Nucl. Phys. Proc. Suppl.*, 10(205-206), 2010.

[185] J.M. Campbell and R.K. Ellis. An update on vector boson pair production at hadron colliders. *Phys. Rev. D60*, 113006, 1999.

[186] CMS Collaboration. Measurements of inclusive W and Z cross sections in pp collisions at $\sqrt{s} = 8$ TeV. *CMS-PAS*, SMP-12-011, 2012.

[187] CMS Collaboration. Measurement of the WW and ZZ production cross sections in pp collisions at $\sqrt{s} = 8$ TeV. *CMS-PAS*, SMP-12-024, 2012.

[188] Stefano Frixione, Paolo Nason, and Carlo Oleari. Matching NLO QCD computations with parton shower simulations: the POWHEG method. *JHEP*, 11:070, 2007.

[189] N. Kidonakis and R. Vogt. The theoretical top quark cross section at the TeVatron and the LHC. *arXiv:0805.3844*, 2005.

[190] N. Davidson, G. Nanava, T. Przedzinski, E. Richter-Was, and Z. Was. Universal interface of TAUOLA technical and physics documentation. 2010.

[191] CMS Collaboration. https://twiki.cern.ch/twiki/bin/view/CMS/SWGuidePF2PAT.

[192] CMS Collaboration. https://twiki.cern.ch/twiki/bin/view/CMS/MuonPOG.

[193] CMS Collaboration. https://twiki.cern.ch/twiki/bin/view/CMS/EgammaCutBasedIdentification.

[194] CMS Collaboration. Identification of b-quark jets with the CMS experiment. *PAS*, 001(BTV12-001), 2012.

[195] CMS Collaboration. Measurement of the b-tagging efficiency in semi-leptonic decays of tt events using Flavour-tag Consistency Method in $\sqrt{s} = 8$ TeV data. *AN-12-187*, 2012.

[196] CMS Collaboration. Jet Performance in pp Collisions at $\sqrt{s} = 7$ TeV. *Physics Analysis Summary CMS-PAS-JME-10-013*, 2010.

[197] CMS Collaboration. https://twiki.cern.ch/twiki/bin/view/CMS/METType1Type2Formulae, 2013.

[198] CMS Collaboration. https://twiki.cern.ch/twiki/bin/view/CMS/PileupMCReweightingUtilities, 2013.

[199] CMS Collaboration. Better jet clustering algorithms. *JHEP*, 9708(001), 1997.

[200] CMS Collaboration. https://twiki.cern.ch/twiki/bin/view/CMSPublic/WorkBookJetEnergyCorrections.

[201] CMS Collaboration. Search for BSM $t\bar{t}$ Production in the Boosted All-Hadronic Final State using pp Collisions at $\sqrt{s} = 8$ TeV. *PAS*, (B2G-12-005), 2012. B2G-12-005.

[202] L. Baeni et al. Simultaneous Heavy Flavor and Top (SHyFT) Cross Section Measurement. *CMS Note 2011/210, (2011).*, 2011.

[203] CMS Collaboration. https://twiki.cern.ch/twiki/bin/view/CMSPublic/SWGuideMuonId.

[204] J. Ott T. Müller and J. Wagner-Kuhr. theta : a framework for template-based modelling and inference.

[205] CMS Collaboration. https://twiki.cern.ch/twiki/bin/viewauth/CMS/JetID.

[206] CMS Collaboration. https://twiki.cern.ch/twiki/bin/viewauth/CMS/JECUncertaintySources.

[207] Torbjorn Sjöstrand, Stephen Mrenna, and Peter Skands. PYTHIA 6.4 physics and manual. *JHEP*, 05:026, 2006.

[208] CMS Collaboration. Measurement of the top/antitop production ratio in single top t-channel production. *TOP-12-038*, 2012.

[209] Nikolaos Kidonakis. Differential and total cross sections for top pair and single top production. 2012.

[210] Matteo Cacciari, Michal Czakon, Michelangelo Mangano, Alexander Mitov, and Paolo Nason. Top-pair production at hadron colliders with next-to-next-to-leading logarithmic soft-gluon resummation. *Phys.Lett.*, B710:612–622, 2012.

[211] CMS Collaboration. Absolute Calibration of the Luminosity Measurement at CMS: Winter 2012 Update. 2012.

[212] A. Hoecker et al. TMVA - toolkit for multivariate data analysis. *Arxiv*, (0703039), 2007.

[213] B. P. Roe, H.-J. Yang, J. Zhu, Y. Liu, I. Stancu, and G. McGregor. Boosted decision trees as an alternative to artificial neural networks for particle identification. *Nuclear Instruments and Methods in Physics Research A*, 543:577–584, May 2005.

[214] H.-J. Yang, B. P. Roe, and J. Zhu. Studies of stability and robustness for artificial neural networks and boosted decision trees. *Nuclear Instruments and Methods in Physics Research A*, 574:342–349, May 2007.

[215] Sophus Lie. *Theorie der transformationsgruppen.* Leipzig, B.G. Teubner, 1888.

[216] Robert Gilmore. *Lie Groups, Lie Algebras, and Some of Their Applications.* Dover Publications Inc., 2001.

[217] Peter J. Mohr et al. CODATA recommended values of the fundamental physical constants 2010. *Reviews of Modern Physics*, 84, 2012.

[218] Albert A. Michelson and Edward W. Morley. Influence of Motion of the Medium on the Velocity of Light. *American Journal of Science*, 31(185):377–386, 1886.

[219] A. Einstein. Zur Elektrodynamik bewegter Körper. *Annalen der Physik*, 17(10):891–921, 1905.

[220] R. H. Good. Properties of the Dirac Matrices. *Reviews of Modern Physics*, 27:187–211, April 1955.

[221] C. S. Wu, E. Ambler, R. W. Hayward, D. D. Hoppes, and R. P. Hudson. Experimental Test of Parity Conservation in Beta Decay. *Phys. Rev.*, 105:1413–1415, Feb 1957.

[222] Lorenzo Moneta Akira Shibata Wouter Verkerke Kyle Cranmer, George Lewis. HistFactory: A tool for creating statistical models for use with RooFit and RooStats. *CERN preprint*, CERN-OPEN-2012-016, Jan 2012.