

Searching for Dark Matter Using Jets and Jet Substructure at the Large Hadron Collider

by

Siddharth Madhavan Narayanan

Submitted to the Department of Physics
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

February 2019

© Massachusetts Institute of Technology 2019. All rights reserved.

Author
.....

Department of Physics
January 29, 2019

Certified by
.....

Christoph M. E. Paus
Professor of Physics
Thesis Supervisor

Accepted by
.....

Nergis Mavalvala
Associate Department Head of Education, Professor of Physics

**Searching for Dark Matter Using Jets and Jet Substructure
at the Large Hadron Collider**

by

Siddharth Madhavan Narayanan

Submitted to the Department of Physics
on January 29, 2019, in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy

Abstract

Astrophysical observations of gravitational interactions provide strong evidence for the existence of dark matter. Many theories propose and experiments test the hypothesis that dark matter may have a particle physics origin, but this remains unproven. One such experiment is the Compact Muon Solenoid at the Large Hadron Collider. If dark matter couples, at least lightly, to the Standard Model, then it is possible to produce it in collisions at the LHC. Because it would not interact with the detector, we must look for collisions in which dark matter is produced in association with one or more SM particles. This thesis describes two such analyses: dark matter plus one top quark and dark matter plus two light quarks. Both cases result in complicated detector signatures due to the hadronization of final-state quarks. Recently developed jet substructure techniques were applied using novel methods to identify the hadronization products of high-momentum top quarks. In both analyses, the observed data is found to be consistent with SM backgrounds. We translate these results into the most stringent constraints to date on the relevant beyond-SM models.

Thesis Supervisor: Christoph M. E. Paus
Title: Professor of Physics

Contents

1	Introduction	7
1.1	The Standard Model of particle physics	7
1.1.1	Quantum chromodynamics	9
1.1.2	Electroweak interactions	11
1.2	Dark matter	15
1.3	LHC phenomenology	21
1.3.1	Parton distribution functions	23
1.3.2	Hard scattering	23
1.3.3	Parton shower	24
2	The CMS Experiment at the LHC	27
2.1	The Large Hadron Collider	27
2.2	The Compact Muon Solenoid	30
2.2.1	Silicon tracker	31
2.2.2	Electromagnetic calorimeter	37
2.2.3	Hadronic calorimeter	38
2.2.4	Muon chambers	42
2.2.5	Online trigger system	43
2.3	Particle reconstruction and identification	46
2.3.1	Particle flow algorithm	46
2.3.2	Muons	47
2.3.3	Electrons and photons	49
2.3.4	Hadrons and jets	51

2.3.5	Hadronic taus	55
2.3.6	Missing momentum	56
2.3.7	Pileup mitigation	57
2.4	Simulation of CMS	59
3	Hadronic Resonance Identification	61
3.1	Reconstruction	61
3.2	Identification	64
3.2.1	Substructure	66
3.2.2	A combined tagger	71
3.3	Data validation	77
4	The Search for Missing Momentum and a Top Quark	83
4.1	Signal selection	88
4.1.1	Online trigger selection	88
4.1.2	Offline signal selection	90
4.2	Background estimation	93
4.2.1	Visible final states to constrain invisible final states	93
4.2.2	Theoretically-limited extrapolations	105
4.3	Results	113
4.3.1	Constraints on mono-top models	117
5	The Search for Invisible Decays of the Higgs Boson	121
5.1	Signal selection	122
5.1.1	Online trigger selection	122
5.1.2	EW and QCD production of electroweak bosons	130
5.1.3	Sensitivity optimization	133
5.2	Background estimation	134
5.3	Results	136
A	The CL_s Method	141

Chapter 1

Introduction

1.1 The Standard Model of particle physics

The Standard Model (SM) is a quantum field theory governing the kinematics and interactions of a set of fermion fields, respecting Lorentz and local gauge symmetries. The gauge symmetries correspond to the SM gauge group:

$$G = \mathrm{SU}(3) \times \mathrm{SU}(2) \times \mathrm{U}(1) \tag{1.1}$$

This group can be decomposed into two subgroups: the *strong* interaction governed by $\mathrm{SU}(3)$ [1, 2] and the *electroweak* interaction governed by $\mathrm{SU}(2) \times \mathrm{U}(1)$ [3, 4, 5]. The action of G on a fermion depends on the representation of the Lie group in which the fermion resides and the coupling strength. For the special unitary groups, we denote the representation as **N** (fundamental), $\bar{\mathbf{N}}$ (anti-fundamental), or **1** (trivial). The special unitary gauges have the same interaction strength for all fermions in non-trivial representations, known as *universality*. All fermions are in a one-dimensional representation of $\mathrm{U}(1)$, but the weak hypercharge Y distinguishes their transformation under the gauge group $f \mapsto f + iYg'f$, where g' is the coupling strength of $\mathrm{U}(1)$.

Table 1.1 gives a summary of the *first generation* SM fermion fields and the representation of G which acts on them. The SM provides a total of 3 generations of fermions, each of them a copy of the first generation in terms of the field content and

Table 1.1: First generation SM fermions and the action of the SM local gauge symmetry group G . The subscripts L and R refer to left- and right-handed chirality fields. Not shown are the charge conjugated fields $f^C \equiv Cf$, which sit in conjugated representations.

Name	Symbol	Y	SU(2) rep.	SU(3) rep.
Left-handed lepton	ℓ_L	$-1/2$	2	1
Right-handed charged lepton	e_R^-	-1	1	1
Right-handed neutrino	ν_R	0	1	1
Left-handed quark	q_L	$1/6$	2	3
Right-handed up quark	u_R	$2/3$	1	3
Right-handed down quark	d_R	$-1/3$	1	3

gauge group action.

The lepton doublets contain the left-handed charged leptons and neutral neutrinos:

$$\ell_{iL} = \begin{pmatrix} \nu_e \\ e_L^- \end{pmatrix}, \begin{pmatrix} \nu_\mu \\ \mu_L^- \end{pmatrix}, \begin{pmatrix} \nu_\tau \\ \tau_L^- \end{pmatrix} \quad (1.2)$$

where i indexes the generation. The right handed lepton singlets contain the right-handed projections of the same fermions.

The quark (electroweak) doublets contain the left-handed up- and down-type quarks:

$$q_{iL} = \begin{pmatrix} u_L \\ d'_L \end{pmatrix}, \begin{pmatrix} c_L \\ s'_L \end{pmatrix}, \begin{pmatrix} t_L \\ b'_L \end{pmatrix} \quad (1.3)$$

where d' , s' , and b' represent linear combinations of the mass eigenstates d , s , and b (discussed further in Section 1.1.2). All quarks also sit in a strong triplet; we have suppressed its charge above, as the strong representation is orthogonal to the electroweak representation. Where necessary, it will be specified with a superscript, i.e. u^c .

1.1.1 Quantum chromodynamics

The dynamics of quarks under the SU(3) gauge group are commonly referred to as the strong interaction or quantum chromodynamics (QCD). The QCD Lagrangian is:

$$\mathcal{L}_{\text{QCD}} = i\bar{q}_f^a \not{D}^{ab} q_f^b + m_f \bar{q}_f^a q_f^a - \frac{1}{4} G_{\mu\nu}^a G^{a,\mu\nu} \quad (1.4)$$

where repeated indices are contracted; $q_f = u, d, c, s, b, t$ are the spinors for each quark flavor f ; $a, b = r, g, b$ index the basis elements of the triplet representation, called *colors*; m_q is the mass of quark flavor q . D_μ is the QCD covariant derivative:

$$D_\mu^{ab} = \delta^{ab} \partial_\mu - ig_s \sum_c t_c^{ab} G_{c,\mu} \quad (1.5)$$

where g_s is the strong coupling strength; t_c are the 8 generators of the triplet representation of SU(3); G_c are the corresponding 8 gauge boson (gluon) fields. $G_{\mu\nu}^a$ are the gluon field strength tensors:

$$G_{\mu\nu}^a = \partial_\mu G_\nu^a - \partial_\nu G_\mu^a - g_s f^{abc} G_\mu^b G_\nu^c \quad (1.6)$$

where f^{abc} are the structure constants of SU(3), with $a, b, c = r, g, b$.

An additional term can be added to Equation 1.4 without violating any gauge or Lorentz symmetry or renormalizability. This term would violate CP conservation and produce a non-zero electric dipole moment (EDM) for the neutron. No evidence for such a term has yet been found, although it has not been excluded [6] We therefore do not consider it as part of the SM QCD Lagrangian.

Renormalization and running of couplings

Physical quantities in a QFT such as couplings and masses acquire a scale-dependence from higher-order corrections to vertices and propagators. In many cases, the quantum corrections to the bare parameter contain ultraviolet divergent terms. These infinities are absorbed into the Lagrangian by means of adding *counterterms*. The

systematic process of absorbing these infinities and ensuring scale-independence is known as renormalization. A Lagrangian is called *renormalizable* if only a finite number of counterterms are needed to ensure that all observables (i.e. amplitudes) are finite. The SM is a renormalizable theory.

In this discussion, we will focus on the renormalization of the coupling $\alpha_S \equiv g_s^2/4\pi$, but the argument applies broadly to all SM quantities. There are two related consequences of quantum corrections. Firstly, α_S acquires a non-trivial dependence on the probed energy scale $\mu_R^2 = -q^2$, where q^μ is the momentum transferred by the gluon. Secondly, the *bare* α_S as written in the Lagrangian is not the same as the $\alpha_S(\mu_R^2)$ measured in the laboratory. We will refer to the bare coupling as α_{S0} . Note that it does not have a μ_R^2 -dependence. To enforce this, we look for solutions to the differential equation

$$\mu_R^2 \frac{d\alpha_{S0}}{d\mu_R^2} = 0 \quad (1.7)$$

Writing α_{S0} in terms of α_S , which has a μ_R^2 -dependence, and rearranging the terms gives the *β function* for α_S :

$$\beta(\alpha_S) \equiv \mu_R^2 \frac{d\alpha_S}{d\mu_R^2} = -\left(\beta_0 \alpha_S^2 + \beta_1 \alpha_S^3 + \mathcal{O}(\alpha_S^4)\right) \quad (1.8)$$

The β coefficients are:

$$\beta_0 = \frac{33 - 2n_f}{3}, \quad \beta_1 = \frac{153 - 19n_f}{24\pi^2}, \dots \quad (1.9)$$

where n_f is the number of quark flavors with masses below μ_R [7, 1, 2, 8]. To one-loop order, the solution to this differential equation is:

$$\alpha_S(\mu_R) = \frac{2\pi}{\beta_0 \ln \frac{\mu_R}{\Lambda_{\text{QCD}}}} \quad (1.10)$$

where Λ_{QCD} is set by enforcing a measured boundary condition, e.g. $\alpha_S(m_Z^2) = 0.1181 \pm 0.0011$ [7]. Using $n_f = 5$ at $\mu_R = m_Z$ and the $\overline{\text{MS}}$ renormalization scheme [9] gives $\Lambda_{\text{QCD}} = 218$ MeV. $\beta_0 > 0$ implies α_S falls as a function of energy, as illustrated

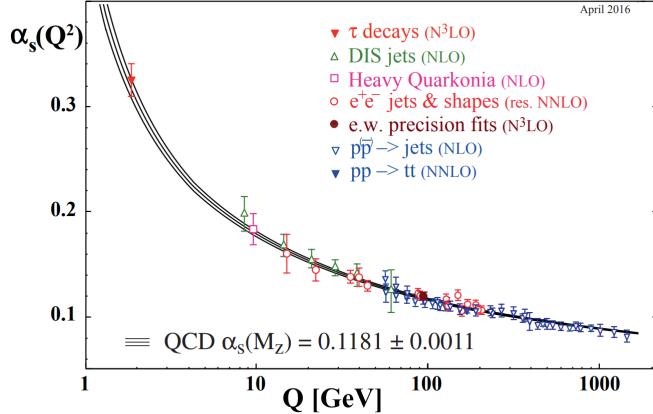


Figure 1.1: Running of the QCD coupling strength α_S as a function of length scale. Reprinted from Reference [7].

in Figure 1.1.

This running is the opposite of theories like quantum electrodynamics, in which $\beta_0 < 0$. It results in an asymptotically free theory as $\mu_R \rightarrow \infty$. Below Λ_{QCD} , the coupling constant is larger than order unity. In the regime where $\alpha_S < 1$, perturbation theory can be applied to calculate observables with α_S as the expansion parameter. This is referred to as perturbative QCD (pQCD). The long-range behavior of QCD also means that quarks and gluons, which have a color charge, cannot be observed below a temperature $\sim 10^{12}$ K [10]. Instead, a spectrum of composite color singlet states, known as *hadrons*, are observed [11].

1.1.2 Electroweak interactions

The electroweak (EW) sector of the SM refers to the $SU(2) \times U(1)$ symmetry group. While all fermion fields in Table 1.1 transform under $U(1)$, only left-handed fermions have non-trivial transformations under $SU(2)$. Ignoring mass terms and the Higgs

sector, the EW Lagrangian is:

$$\begin{aligned}\mathcal{L}_{\text{EW}} = & i\bar{\ell}_{iL} (\not{\partial} - igW^a \tau^a - ig'Y_{\ell_L} \not{B}) \ell_{iL} + i\bar{q}_{iL}^c (\not{\partial} - igW^a \tau^a - ig'Y_{q_L} \not{B}) q_{iL}^c \\ & + i\bar{e}_{iR} (\not{\partial} - ig'Y_{e_R} \not{B}) e_{iR} + i\bar{u}_{iR^c} (\not{\partial} - ig'Y_{u_R} \not{B}) u_{iR}^c + i\bar{d}_{iR}^c (\not{\partial} - ig'Y_{d_R} \not{B}) d_{iR}^c \\ & - \frac{1}{4} W_{\mu\nu}^a W^{a,\mu\nu} - \frac{1}{4} B_{\mu\nu} B^{\mu\nu}\end{aligned}\quad (1.11)$$

where repeated indices are contracted; the subscript i indexes generations; g and g' are respectively the coupling strengths for SU(2) and U(1); Y is the weak hypercharge; W_μ^a are the three gauge fields corresponding to the generators $\tau^a = \sigma^a/2$ of SU(3); B_μ is the gauge field for U(1); and $W_{\mu\nu}^a$ and $B_{\mu\nu}$ are the field strength tensors for the respective gauge fields. The covariant derivative can be written as:

$$D_\mu = \partial_\mu - ig\delta_L W_\mu^a \tau^a - ig'Y B_\mu \quad (1.12)$$

where Y is the particle's hypercharge and δ_L is 1 if the field is in the **2** representation of SU(2) (e.g. left-handed fermions) and 0 otherwise.

EW symmetry breaking

Unlike Equation 1.4, we cannot introduce a quadratic mass term for fermions in Equation 1.11 because $\bar{\psi}\psi = \psi_R^\dagger\psi_L + \psi_L^\dagger\psi_R$ is not invariant under SU(2) rotations. Spontaneous electroweak symmetry breaking remedies this, as well as provides masses for the gauge fields [12, 13, 14, 15, 16, 17]. A complex scalar doublet ϕ (called the complex Higgs field) with $Y_\phi = 1/2$, $\delta_L = 1$ is added to the Lagrangian:

$$\mathcal{L}_{\text{EW}} \mapsto \mathcal{L}_{\text{EW}} + |D_\mu\phi|^2 + \mu^2|\phi|^2 - \lambda|\phi|^4 \quad (1.13)$$

We can write the complex doublet as 4 real fields:

$$\phi = \frac{1}{\sqrt{2}} \begin{pmatrix} \phi_1 + i\phi_2 \\ \phi_3 + i\phi_4 \end{pmatrix} \quad (1.14)$$

The two self-interaction terms create a Higgs potential with a degenerate global minimum at *vacuum expectation value* $v \equiv \langle |\phi| \rangle = \sqrt{\mu^2/\lambda}$. Through gauge rotations, we can fix $\langle \phi_{1,2,4} \rangle = 0$ and, at low energies, expand $\phi_3 = v + H$, where H is the real Higgs field. This is the spontaneous breaking of a symmetry.

By the Nambu-Goldstone theorem [18, 19], these three lost degrees of freedom give rise to three massless bosons. The $|D_\mu \phi|$ term couples the complex Higgs field to the gauge bosons:

$$|D_\mu \phi|_{\phi=\langle \phi \rangle}^2 = \frac{v^2}{8} \left[(gW_\mu^1)^2 + (gW_\mu^2)^2 + (g'B_\mu - gW_\mu^3)^2 \right] \quad (1.15)$$

The diagonalization of this mass term gives 3 massive weak bosons (consuming the 3 massless Nambu-Goldstone bosons) and one massless photon:

$$\begin{aligned} W_\mu^\pm &\equiv \frac{W_\mu^1 \mp iW_\mu^2}{\sqrt{2}} \\ Z_\mu &\equiv \cos \theta_w W_\mu^3 - \sin \theta_w B_\mu \\ A_\mu &\equiv \sin \theta_w W_\mu^3 + \cos \theta_w B_\mu \end{aligned} \quad (1.16)$$

where $\tan \theta_w = g'/g$. The corresponding mass eigenvalues are:

$$m_W = \frac{gv}{2}, \quad m_Z = \frac{v\sqrt{g^2 + g'^2}}{2}, \quad m_A = 0 \quad (1.17)$$

The remaining H field, which has not been consumed, gives rise to the Higgs boson discovered by CMS and ATLAS [20] in 2012. It has mass $m_H = \sqrt{2}\mu$. By expanding ϕ around $\langle \phi \rangle$ in Equation 1.13, we find couplings to the massive gauge bosons:

$$\begin{aligned} \frac{m_Z^2}{v} h Z_\mu Z^\mu, & \quad \frac{2m_W^2}{v} h W^{+\mu} W_\mu^-, \\ \frac{m_Z^2}{2v} h^2 Z_\mu Z^\mu, & \quad \frac{m_W^2}{v} h^2 W^{+\mu} W_\mu^- \end{aligned} \quad (1.18)$$

The breaking of the $SU(2) \times U(1)$ symmetry leaves behind the local $U(1)$ symmetry of electromagnetism (EM), with gauge boson A_μ . Fermions have EM charge $eQ =$

$e(T_3 + Y)$, where $e = g' \cos \theta_w$ and T_3 is the third isospin component. The W^\pm bosons receive charge $\pm e$. After symmetry breaking, the actions of the broken gauge groups on fermions are governed by the following Lagrangian terms:

$$\begin{aligned} \mathcal{L}_{\text{EWSB}} \supset \sum_f & \left[\bar{f} \left(i\partial - eQ_f A \right) f - \frac{g}{2\sqrt{2}} \bar{f}_L \left(T^+ W^+ + T^- W^- \right) f_L \right. \\ & \left. - \frac{g}{2 \cos \theta_w} \bar{f} (g_{Vf} - g_{Af}) Z f \right] \end{aligned} \quad (1.19)$$

where f are all fermion fields; $f_L = \frac{1}{2}(1 - \gamma^5)f$; $g_V = T_3 - 2Q \sin^2 \theta_w$; and $g_A = T_3$.

Fermion masses

The last piece of the EW Lagrangian is the addition of the fermion masses through Yukawa couplings with the Higgs doublet. First, let us add the terms for quark couplings:

$$\mathcal{L}_{\text{EW}} \mapsto \mathcal{L}_{\text{EW}} - y_{ij}^d \bar{q}_i L \phi d_{jR} - y_{ij}^u \bar{q}_i L i \sigma_2 \phi^* u_{uR} i + \text{h.c.} \quad (1.20)$$

where h.c. refers to the Hermitian conjugate of preceding terms; and $y_{ij}^{u,d}$ are the Yukawa matrices for up- and down-type quarks. Breaking the symmetry and collecting terms proportional to v :

$$- \frac{v}{\sqrt{2}} \left(y_{ij}^d \bar{d}'_{iL} d_{jR} + y_{ij}^u \bar{u}'_{iL} u_{jR} \right) \quad (1.21)$$

The mass eigenstates are found by diagonalizing these terms, which are written in terms of the weak eigenstates. Let us denote the unitary transformations from the mass basis to the weak basis as U_u and U_d . If we try to write the rest of $\mathcal{L}_{\text{EWSB}}$ in terms of mass eigenstates, we see that terms of the following form all have trivial transformations:

$$\bar{d}' \gamma^\mu d' \mapsto \bar{d} U_d^\dagger \gamma^\mu U_d d = \bar{d} \gamma^\mu d \quad (1.22)$$

The only non-trivial transformation is in the charged weak interaction:

$$\begin{aligned} \bar{u}_L W^+ d'_L + \bar{d}'_L W^- u_L &\mapsto \bar{u}_L W^+ U_u^\dagger U_d d_L + \bar{d}_L W^- U_d^\dagger U_u u_L \\ &\equiv \bar{u}_L W^+ V_{\text{CKM}} d_L + \bar{d}_L W^- V_{\text{CKM}}^\dagger u_L \end{aligned} \quad (1.23)$$

where V_{CKM} is the Cabibbo-Kobayashi-Maskawa matrix [21, 22]. It is nearly-diagonal, but with non-zero mixing between the generations. The CKM matrix also contains a charge parity (CP) violating phase. When referring to down-type quarks, we typically refer to the mass eigenstate d as opposed to d' .

A similar analysis can be carried out for the lepton sector:

$$\mathcal{L}_{\text{EW}} \mapsto \mathcal{L}_{\text{EW}} - y_{ij}^e \bar{\ell}_{iL} \phi e_{jR} - y_{ij}^\nu \bar{\ell}_{iL} i \sigma_2 \phi^* \nu_{uR} i + \text{h.c.} \quad (1.24)$$

The mixing matrix for leptons is the Pontecorvo-Maki-Nakagawa-Sakata matrix U_{PMNS} [23], which relates the weak eigenstates ν_e, ν_μ, ν_τ with the mass eigenstates ν_1, ν_2, ν_3 . The values of the neutrino masses are known to be non-zero from the observation of neutrino oscillations [24].

After EWSB, each fermion mass eigenstate has a mass term and coupling to the Higgs field:

$$\mathcal{L}_{\text{EWSB}} \supset \sum_f -\frac{y_f v}{\sqrt{2}} (\bar{f} f + \bar{f} H f) \quad (1.25)$$

where we identify the mass as $m_f = y_f v / \sqrt{2}$. Table 1.2 summarizes all SM fermions and some of their properties after EWSB.

1.2 Dark matter

The SM is able to predict the outcomes of many laboratory experiments, ranging from measurements of the electron magnetic moment [25] to the discovery and characterization of the Higgs boson [20, 26]. Nonetheless, there are certain classes of phenomena it cannot explain:

- There is not yet a complete formulation of *quantum gravity* that incorporates

Table 1.2: Summary of the SM fields after electroweak symmetry breaking. All masses are taken from the global fits compiled by the Particle Data Group [7].

Name	Symbol	Spin	Mass	Q_e
gluon	g^{ab}	1	0	0
photon	γ	1	0	0
Z boson	Z	1	91.2 GeV	0
W boson	W^\pm	1	80.4 GeV	± 1
Higgs boson	H	0	125 GeV	0
up quark	u	$1/2$	2.2 MeV	$2/3$
down quark	d	$1/2$	4.7 MeV	$-1/3$
charm quark	c	$1/2$	1.28 GeV	$2/3$
strange quark	s	$1/2$	95 MeV	$-1/3$
top quark	t	$1/2$	173 GeV	$2/3$
bottom quark	b	$1/2$	4.18 GeV	$-1/3$
electron neutrino	ν_e	$1/2$	—	0
electron	e	$1/2$	511 keV	-1
muon neutrino	ν_μ	$1/2$	—	0
muon	μ	$1/2$	105 MeV	-1
tau neutrino	ν_τ	$1/2$	—	0
tau	τ	$1/2$	178 GeV	-1

general relativity [27].

- The observed *matter/antimatter asymmetry* in the universe cannot be explained by the SM’s CP violation and (predicted) baryon number of violation. Additional CP- and B -violating interactions must exist.
- *Neutrino masses* are not completely determined by the SM. While a mass term can be written down as in Section 1.1.2, it does not exclude a Majorana mass term for right-handed neutrinos. Nor does it explain the observed range of masses $m_\nu/m_t \lesssim 10^{-15}$.
- An explanation of *dark energy* accounting for $\sim 68\%$ of the universe’s energy budget, suggested by measurements of the CMB, galaxy clusters, supernovae, and other measurements of the universe expansion rate [28].
- An explanation of *dark matter* accounting for $\sim 27\%$ of the universe’s energy budget, suggested by measurements of galactic rotation curves, the CMB, and

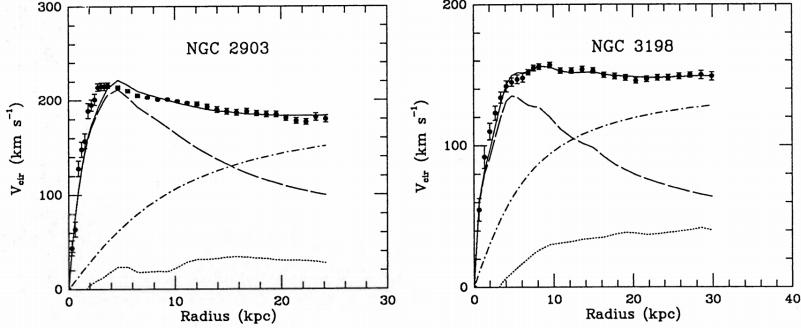


Figure 1.2: Observed and fitted galactic rotational curves for two galaxies. The total fit (solid line) has three components: visible (dashed), gas clouds (dotted), and DM (dash-dotted). Reprinted from Reference [32].

gravitational lensing [7, 29, 30, 31].

In this thesis, we describe tests of certain extensions of the SM which add candidate fields for dark matter (DM).

Astrophysical evidence

All known evidence of DM arises from gravitational measurements. One of the oldest observations is that of galactic rotational curves. The rotational velocities of stars and hydrogen clouds are measured in galaxies, as a function of the distance from the center of the galaxy. It is found that the velocity increases as a function of radius, eventually reaching a plateau that extends well past the bulk of the visible mass of the galaxy. This implies the existence of a massive dark halo that encompasses the galaxy. Indeed, the observed rotational curves $v(r)$ are well-described by a 3 component fit: the visible disk, a gas cloud, and a dark halo. Two galaxies are shown in Figure 1.2, and it is clear that a non-zero DM component is needed to describe the data.

An orthogonal piece of evidence comes from measuring anisotropies in the cosmic microwave background (CMB). The CMB is the remnant of photons after they decoupled from matter in the early universe. The temperature spectrum of the CMB is isotropic to one part in 10^5 , and the anisotropies are driven by matter anisotropies at the time of decoupling. The power spectrum of the temperature anisotropies is modified when there are two matter populations (SM and DM) as opposed to one

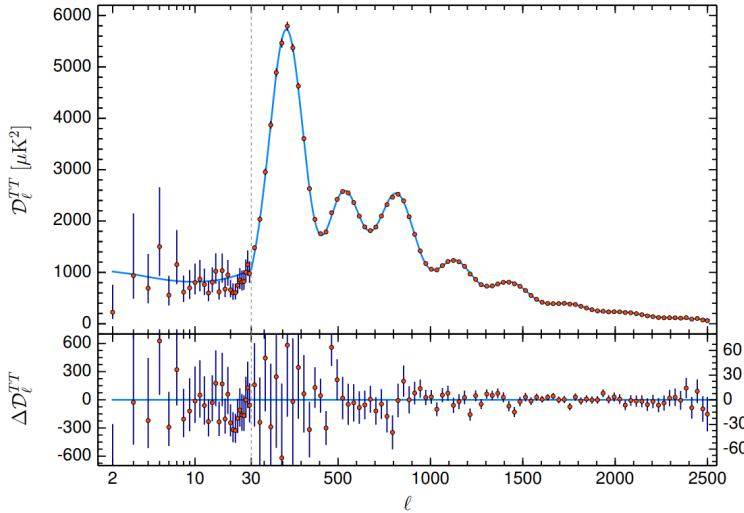


Figure 1.3: Temperature anisotropy power spectrum of the CMB, as measured by Planck. Reprinted from Reference [33].

(just SM), especially when the two populations interact with each other via gravity, but one does not interact with photons (DM). Figure 1.3 shows the power spectrum measured by the Planck experiment [33], compared to the best-fit Λ Cold Dark Matter (Λ CDM) model.

This is hardly an exhaustive list of DM evidence. Other observations include gravitational lensing, cluster collisions, and large-scale structure formation [29, 30].

DM candidates

The *relic density* is the current density of a field that fell out of thermodynamic equilibrium with other fields as the universe expanded. This occurs as the rate of expansion of the universe outpaces the mean particle velocity v (correlated with the temperature T) and the cross sections for annihilation and production of the field. This decoupling of the field is known as *freeze-out*. The relic density is defined:

$$\Omega \cdot h^2 = \frac{\rho}{\rho_c} \cdot \frac{H_0}{100 \text{ km/s/Mpc}} \quad (1.26)$$

where ρ is the energy density of the field, ρ_c is the critical total energy density of flat spacetime, and $H_0 = 73 \text{ km/s/Mpc}$ is the Hubble constant. If σ_A is the annihilation cross section and the field is non-relativistic at freeze-out, then a very simplified approximation [29] is:

$$\Omega h^2 = \frac{10^{-27} \text{ cm}^3 \text{s}^{-1}}{\langle \sigma_A v \rangle} \quad (1.27)$$

Planck [33] measures the relic densities of baryonic and dark matter to be:

$$\Omega_b h^2 = 0.02237 \pm 0.0015, \quad \Omega_{\text{DM}} h^2 = 0.1200 \pm 0.0012 \quad (1.28)$$

As a first candidate, suppose we introduce a fermion χ with mass $m_\chi \sim 100 \text{ GeV}$. Further suppose that it has an interaction with the SM through a new current, and the cross section for the effective four-point interaction $\chi\bar{\chi} \rightarrow f\bar{f}$ (for some SM fermion f) is proportional to g_χ^4 , where $g_\chi \sim g \sim 0.6$. The relic density for such a particle would be $\Omega_\chi h^2 \sim 0.1$, which is quite close to the needed density. A *Weakly-Interacting Massive Particle* (WIMP) such as χ therefore makes a good DM candidate. This coincidence is known as the *WIMP miracle*. It should be noted that the definition of a WIMP is fairly loose. All models considered in these results contain at least one particle falling under the umbrella of a WIMP, but the models themselves are quite diverse.

Many DM models other than WIMPs exist. The obvious SM candidate for DM are neutrinos, as they are known to be non-interacting and massive. However, constraints on the neutrino mass restrict $\Omega_\nu h^2 \leq 0.0062$ [7], which is far too small to entirely account for the observed phenomena. Sterile neutrinos (neutrinos that participate in oscillations but do not couple to the Z boson) remain a viable DM candidate, above a mass of 10 keV. Another DM hypothesis is a scalar *axion* field, which also serves as a solution to the strong CP problem. While axions can have couplings to photons, those interactions would be sufficiently weak for an axion to remain *dark* in the sense of DM. Many other DM candidates have been proposed as well (some of which fall under the generic WIMP umbrella), and this is not an exhaustive list.

Non-collider WIMP searches

Each class of DM models has dedicated experimental signatures, and therefore dedicated experiments to search for them. The philosophy of collider DM searches is to produce DM in the laboratory and directly or indirectly detect it. Here, we focus on searches sensitive to WIMPs, as these are most relevant for the results in this thesis. The mass and coupling ranges of WIMPs allow the possibility for their production at a TeV-scale collider, such as the LHC. In contrast, non-collider searches rely on DM that is present somewhere in the universe and look for its interaction with SM particles.

Indirect detection experiments look for the annihilation process $\chi\bar{\chi} \rightarrow X$, where X is one or more SM particles. The Alpha Magnetic Spectrometer (AMS-02) searches for excess positrons in the cosmic ray flux, arising from $X = e^+e^-$ final states [34]. Similarly, γ ray data from the Fermi Large Area Telescope (Fermi-LAT) is used to look for final states which include high-energy photons [35]. As an example, the 95% CL exclusion from Fermi-LAT is shown in Figure 1.4. The exclusion is drawn as a function of m_χ and $\langle\sigma_A v\rangle$, as the Fermi-LAT analysis is directly dependent on the annihilation rate. Also shown is an exclusion from a search conducted by the CMS experiment at the LHC, looking for the process $pp \rightarrow \chi + \text{jet}$ and assuming a pseudoscalar mediator connecting the DM and SM sectors [36].

Direct detection experiments typically contain a large volume of instrumented material that has a large cross section for interaction with WIMPs. An example is the Large Underground Xenon (LUX) experiment [37], a two phase (liquid and gas) xenon detector. LUX searches for scintillation photons and additional electrons emitted from DM-nucleus interactions. Xenon is chosen because heavy nuclei have larger cross sections for *spin independent* DM interactions. In this context, spin independent refers to any interaction mediated by a scalar or vector current (i.e. there is no coupling to the spin of the DM or nucleon). Figure 1.5 compares the exclusions of LUX to other direct detection experiments. The results are presented as a function of the natural quantities for direct detection: m_χ and $\sigma(\chi N \rightarrow \chi N)$.

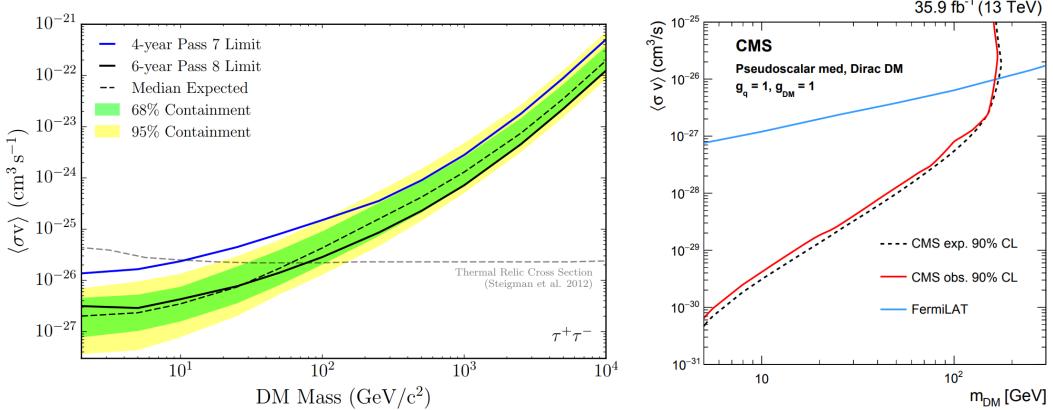


Figure 1.4: Excluded WIMP hypotheses as a function of m_χ , $\langle \sigma v \rangle$ from Fermi-LAT (left) and compared to CMS (right). Reprinted from Reference [35] (left) and Reference [36] (right).

Also shown is the exclusion limit from the CMS invisible Higgs search, which is described in Chapter 5.

The results presented in this section suggest that collider searches have complementary sensitivity for certain DM hypotheses, typically at low masses. The searches described in this thesis target those models, as well as exotic models to which standard direct and indirect detection experiments have very little sensitivity.

1.3 LHC phenomenology

The Large Hadron Collider collides protons at a center of mass energy $\sqrt{s} = 13$ TeV. Section 2.1 provides an overview of the LHC machine. In this section, we describe the methods used to make predictions of observables at the LHC. These observables typically take the form of differential cross sections $d\sigma(pp \rightarrow X)/d\Theta$, where X is some interesting final state with N particles and Θ is a set of interesting kinematics. The differential element of the general cross section for $2 \rightarrow N$ processes is:

$$d\sigma(ab \rightarrow \{c_i\}) = \frac{1}{2s} \left(\prod_i \frac{d^3 p_i}{(2\pi)^3} \frac{1}{2E_i} \right) \cdot (2\pi)^4 \delta^4 \left(k_a + k_b - \sum_i p_i \right) \cdot |\mathcal{M}(ab \rightarrow \{c_i\})|^2 \quad (1.29)$$

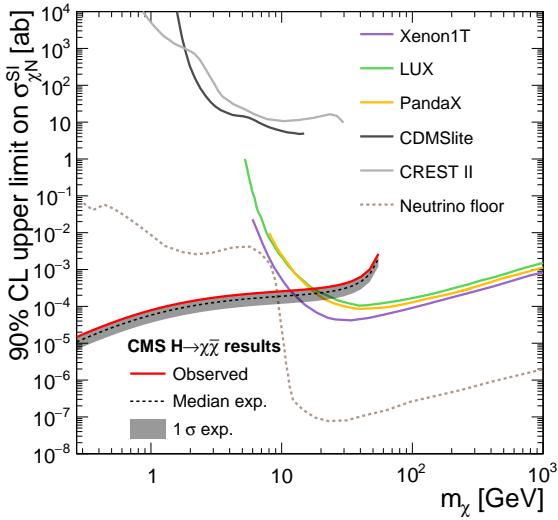


Figure 1.5: Comparison of various direct detection experiments' sensitivity to WIMP models and the sensitivity of the CMS invisible Higgs search. Details of the CMS curve are presented in Chapter 5.

where k_a, k_b are the incoming momenta; $\{p_i\}$ are the outgoing momenta; and \mathcal{M} is the matrix element of this reaction.

Hadron collisions do not have two-particle initial states, but rather two composite particles containing partons with varying momenta. The general cross section for $pp \rightarrow X$ is [38]:

$$d\sigma(pp \rightarrow X(\Theta)) = \sum_{a,b} \int dx_a dx_b f_a(x_a, \mu_F) f_b(x_b, \mu_F) \cdot d\sigma(ab \rightarrow \{c_i\}) \cdot D(\{c_i\} \rightarrow X(\Theta)) \quad (1.30)$$

The sum over a, b refers to summing over partons in the initial state. The momentum fractions of the partons, $x_{a,b}$, follow parton distribution functions (PDFs) $f_{a,b}$ that depend on the particle species a and b . We then define an intermediate state $\{c_i\}$ which evolves into the final state $X(\Theta)$. That is, the process $ab \rightarrow X$ is split into $ab \rightarrow \{c_i\} \rightarrow X$. The definition of $\{c_i\}$ is not unique, but is chosen such that the process $ab \rightarrow \{c_i\}$ can be analyzed using perturbation theory. The evolution of $\{c_i\}$ to the final state involves the radiation of strong and collinear radiation, which cannot be described by pQCD. The matrix element for $ab \rightarrow \{c_i\}$ (*hard scattering*) is computed

perturbatively and turned into a cross section by means of Equation 1.30. Other heuristic methods are used to deal with the process $\{c_i\} \rightarrow X$ (*parton shower* (PS)), encoded in the *fragmentation function* D .

The ability to partition the calculation into perturbative (hard scattering) and non-perturbative (PDF and PS) components follows from the collinear factorization theorem [39]. The factorization depends on an arbitrary energy scale μ_F , which defines a lower bound for interactions considered part of the hard scattering. The remainder of this section discusses the use of Monte Carlo (MC) methods to simulate these three factors: the parton distribution functions f_a , the hard scattering matrix element \mathcal{M} , and the fragmentation function D .

1.3.1 Parton distribution functions

The analytic behavior of PDFs as a function of the factorization scale is governed by the DGLAP evolution equation [40, 41, 42]:

$$\begin{aligned} \mu_F \frac{d}{d\mu_F} f_a(x_a, \mu_F) &= \frac{\alpha_S}{\pi} \int_x^1 \frac{dy}{y} f_a(y, \mu_F) P_{qq} \left(\frac{x}{y} \right) \\ \text{where } P_{qq}(z) &= \frac{4}{3} \left[\frac{1+z^2}{1-z} \right]_+ + 2\delta(1-z) \end{aligned} \quad (1.31)$$

The computation of f_a for a fixed scale cannot be done analytically. Instead, data from many experiments are used to fit a parameterization, as is done by the NNPDF collaboration [43]. Results presented in this thesis use the NNPDF3.0 PDF set.

1.3.2 Hard scattering

Monte Carlo generators simulate Equation 1.29 by sampling events with probability proportional to the phase space and matrix element. The dedicated MC generators used in this result are MadGraph5 [44, 45] and Powheg2 [46]. Both generators can simulate to leading order (LO) in EW vertices and up to next-to-leading order (NLO) in QCD vertices. The PS model Pythia8.2 [47] can also generate certain hard scattering processes at LO in QCD.

Table 1.3: Summary of MC generators used for each SM and signal process. Note that the $V + \text{jet}$ processes have two entries each in this table, at different QCD orders. Both will be used to improve the prediction accuracy.

$pp \rightarrow X$	Generator	NLO in QCD?	Notes
$t\bar{t}$	Powheg	yes	
t, tW, tZ	Powheg	yes	
ZZ, WZ, WW	Pythia	no	
$Z (+0,1 \text{ partons})$	MG	yes	
$W (+0,1 \text{ partons})$	MG	yes	
$\gamma (+0,1 \text{ partons})$	MG	yes	
$Z (+0,1,2,3 \text{ partons})$	MG	no	
$W (+0,1,2,3 \text{ partons})$	MG	no	
$\gamma (+0,1,2 \text{ partons})$	MG	no	
Multijet	MG		Jets defined in Section 1.3.3
Resonant DM	MG	no	See Chapter 4
FCNC DM	MG	yes	See Chapter 4
$\text{VBF } H \rightarrow \chi\bar{\chi}$	Powheg	no	See Chapter 5

MC generators also generate additional partons from initial and final state radiation (ISR and FSR) as part of the hard scattering. These quarks, gluons, and photons can be highly energetic in events with large q^2 , and so it is necessary to compute them as part of the hard scattering matrix element. It should be noted that, while NLO event generation is always preferable, it comes with two costs: (a) being able to generate fewer additional partons and (b) additional computational time. Correspondingly, we will use LO simulation in cases where (a) the spectra of ISR and FSR are important and (b) many simulated events are needed for a processes with large cross-sections. Table 1.3 provides a summary of the MC generators and orders used in these results.

1.3.3 Parton shower

Suppose we know $d\sigma(ab \rightarrow \{c_i\})$ and would like to know $d\sigma(ab \rightarrow \{c_i\}j)$, where j is an additional particle radiated from one of the c_i in the soft and/or collinear limit.

Such *splittings* include:

$$\begin{aligned} q \rightarrow qg, \quad g \rightarrow qq, \quad g \rightarrow gg, \\ f \rightarrow f\gamma, \quad \gamma \rightarrow ff \end{aligned} \tag{1.32}$$

Define z to be the fraction of energy kept by the parent parton and θ to be the opening angle between c_i and j . For QCD splittings, the cross section can be written (at LO):

$$d\sigma(ab \rightarrow \{c_i\}j) = P_{c_i \rightarrow c_ij}(z) \cdot \frac{\alpha_S}{2\pi} \cdot \frac{d\theta}{\theta} dz d\sigma(ab \rightarrow \{c_i\}) \tag{1.33}$$

where $P_{c_i \rightarrow c_ij}$ are *splitting functions* analogous to the ones that arise in the DGLAP evolution (Equation 1.3.1). As an example, the splitting function for $q \rightarrow qg$ is [47]:

$$P_{q \rightarrow qg} = \frac{4}{3} \frac{1+z^2}{1-z} \tag{1.34}$$

It should be noted that this cross section grows as $\theta \rightarrow 0$ and $z \rightarrow 1$, i.e. soft and collinear splittings. That is, if a bare quark is produced in an event, it will produce many soft and collinear gluons (which in turn can split to qq and gg) prior to hadronization. This iterative process is known as the *parton shower*. As the width of the parton shower nears $1/\Lambda_{\text{QCD}}$, the partons will hadronize, preventing further splitting. These hadronic endpoints of the shower reach the detector and appear as collimated sprays of hadrons (*jets*). The reconstruction of jets is described in detail in Section 2.3.4 and Chapter 3.

PS models simulate the shower and hadronization processes. The results in this thesis use Pythia8.2 [47] which employs the Lund string model [48] to simulate hadronization.

Chapter 2

The CMS Experiment at the LHC

2.1 The Large Hadron Collider

The Large Hadron Collider (LHC) [49]¹ is a circular particle accelerator, 27 km in circumference and between 40 and 175 m below the surface. The ring cross the French-Swiss border near the city of Geneva. Designed to collide protons at a maximum center-of-mass energy $\sqrt{s} = 14$ TeV, the LHC has delivered collisions at $\sqrt{s} = 7, 8$ TeV (Run 1) and $\sqrt{s} = 13$ TeV (Run 2); the target energy $\sqrt{s} = 14$ TeV will be reached in Run 3. In addition to protons, the LHC accelerates and collides heavy nuclei (Pb and Xe) at lower values of \sqrt{s} . In this thesis, we focus exclusively on data recorded from proton collisions during Run 2.

Protons are brought to the LHC by the multi-stage process [51] depicted in Figure 2.1. Hydrogen atoms are stripped of electrons and accelerated to a kinetic energy of 50 MeV by LINAC2, a linear accelerator. Subsequently, the Booster ring, the Proton Synchrotron (PS) and Super Proton Synchrotron (SPS) successively accelerate the protons to energies of 1.4, 26, and 450 GeV, respectively. Protons exit the SPS and enter the LHC at one of two places, corresponding to two different beams traveling in opposite directions. The two beams intersect in eight places along the LHC, four of which are instrumented by a detector experiment: CMS, ATLAS, LHCb, and

¹Unless otherwise specified, all technical specifications of the LHC are derived from Reference [49]

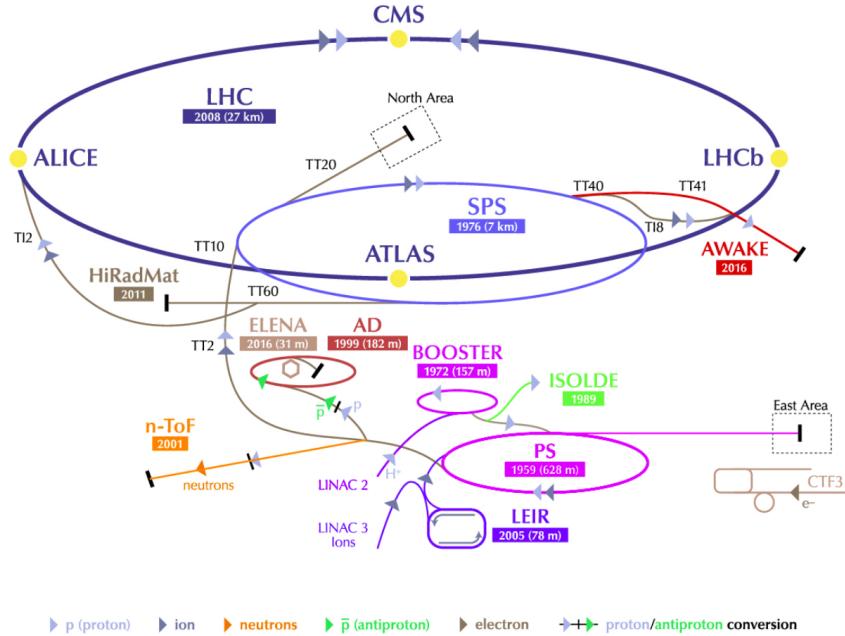


Figure 2.1: Diagram of the CERN accelerator complex. The LHC (dark blue) is fed protons (and heavy ions) by a chain of intermediate accelerators, beginning with LINAC2 (dark pink). Reprinted from Reference [50].

ALICE.

Each proton beam in the LHC is accelerated by eight superconducting cavities exerting 400 MHz oscillating electric fields parallel to the beam line (the *longitudinal* direction). The maximum RF voltage seen by each beam is 16 MV per revolution. The physical and temporal design of the RF system creates bunches of protons, corresponding to nodes of the oscillating field, approximately 7.5 cm in length and separated by 25 ns. Superconducting NbTi dipole magnets bend the two proton beams in opposite directions as they travel around the ring. Each of the 1232 dipoles is 14 m long and exerts a transverse B field between 0.54 and 8.33 T. To achieve such high B fields, the magnets are cooled to 2 K by superfluid helium. In addition, a number of quadrupole magnets are used to focus and match the beams between the dipoles².

The two main figures of merit for a collider the center-of-mass energy \sqrt{s} and the

²Full details on the various quadrupoles can be found in Table 3.7 of Reference [49].

number of events produced per unit time. The latter is defined as:

$$N(pp \rightarrow X) = \int dt L\sigma(pp \rightarrow X) \quad (2.1)$$

where σ is the cross section of the relevant process and L is the instantaneous luminosity of the LHC. The cross section is fixed by nature, and so increasing the luminosity is the only handle to increase N . The instantaneous luminosity of two Gaussian beams is given by [49]:

$$L = \frac{N_b^2 n_b f_{\text{rev}} \gamma F}{4\pi\epsilon\beta^*} \quad (2.2)$$

where:

- N_b = particles per bunch
- n_b = bunches per beam
- f_{rev} = frequency of revolution
- $\gamma = E/m$ of beam
- ϵ = emittance of beam
- β^* = beta function at collision point
- F = factor accounting for beam intersection geometry

The instantaneous luminosity evolves as a function of time, primarily due to n_b and N_b being modified by collisions. The total integrated luminosity after time T is:

$$L_{\text{int}} = \int_0^T dt L(t) = L(0)\tau_L \left(1 - e^{-T/\tau_L}\right) \quad (2.3)$$

where $\tau_L \approx 15$ h is the characteristic beam loss timescale and $L(0)$ is the instantaneous luminosity at $t = 0$. The LHC is designed to deliver $L(0) \sim \mathcal{O}(10^{34}) \text{ cm}^{-2}\text{s}^{-1}$. Figure 2.2 shows the total luminosity delivered by the LHC and recorded by CMS during the 2016 portion of Run 2.

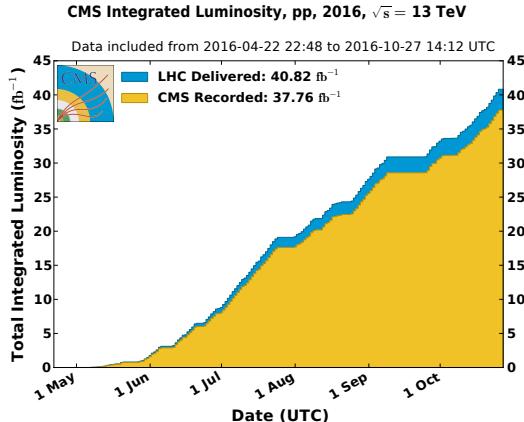


Figure 2.2: Integrated luminosity of the LHC during proton collisions during the 2016 data-taking period [52].

2.2 The Compact Muon Solenoid

The Compact Muon Solenoid (CMS) [53] is one of two general purpose LHC detectors, the other being ATLAS [54]. It is designed to detect and measure stable hadrons, photons, electrons, and muons produced in proton and ion collisions at LHC interaction point 5. From these event descriptions, a number of physics processes can be probed, including SM measurements, BSM searches, and the discovery of the Higgs boson. In what follows, we will use the (r, ϕ, η) coordinate system with respect to the z axis:

- z = distance along beam axis, with $z = 0$ at the center of the detector
- r = distance from the z axis
- ϕ = azimuthal angle in the plane orthogonal to the z axis
- η = pseudorapidity ($-\log \theta/2$), with respect to the polar angle θ

In this coordinate system, we define x and y to lie in the plane perpendicular to z , with x pointing from the center of the detector to the center of the LHC. As with the pseudorapidity, it is convenient to use quantities invariant under z -boosts, and so we

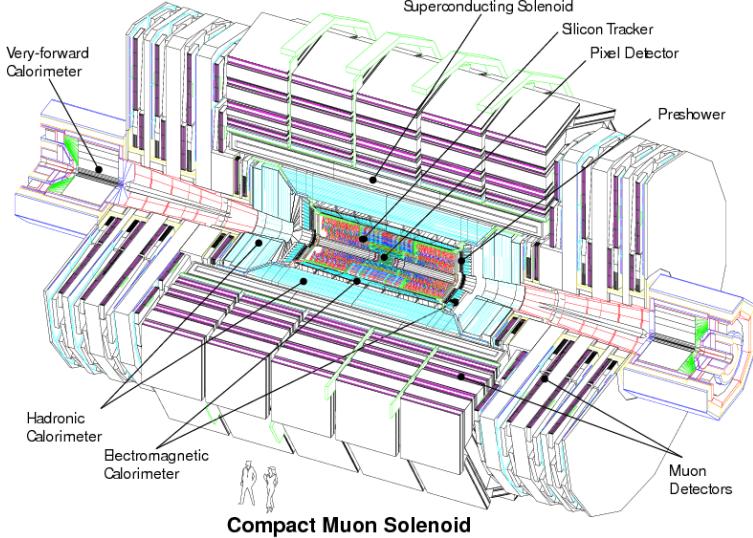


Figure 2.3: Cut-away view of the CMS detector and its subsystems. Reprinted from Reference [53].

define the transverse momentum:

$$\vec{p}_T = \begin{pmatrix} p_x \\ p_y \end{pmatrix} \quad (2.4)$$

We will frequently make use of the magnitude of this vector, p_T . CMS can detect collision products that are within the fiducial volume of $0 \leq \phi < 2\pi$ and $-5 \leq \eta \leq 5$. Several detector subsystems (Figure 2.3) are used to identify and reconstruct muons, electrons, photons, and charged and neutral hadrons.

2.2.1 Silicon tracker

Starting from the beam pipe, the first of these subsystems is the silicon tracker [55], used to identify charged particles and measure their momenta. The tracker consists of silicon detector geometries: pixels (providing 3D position measurement) and strips (2D). The arrangement of the pixel and strip layers are shown in Figure 2.4. A superconducting NbTi solenoid envelopes the tracker, as well as parts of the calorimeters. The magnetic field inside the tracker volume has strength 3.8 T and field lines approximately parallel to the beam direction.

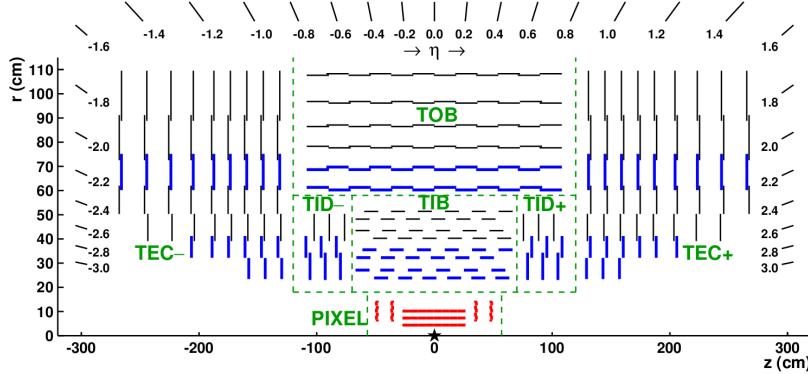


Figure 2.4: Diagram of a slice of the CMS tracking system. The pixel layers are shown in bold red lines. Single-strip (double-strip) layers are indicated by thin black (bold blue) lines. The double-strip modules each consist of two back-to-back strips, rotated with respect to each other, that can provide 3D localization of the hits. Reprinted from Reference [55].

A single silicon pixel has dimensions $285 \times 100 \times 150 \text{ } (\mu\text{m})^3$ (in $r \times r\phi \times z$), leading to a position resolution of $\sim 10 \times 30 \text{ } (\mu\text{m})^2$ (in $r\phi \times z$). The 66 million pixels are arranged into 7 layers: 3 cylindrical *barrels* (at $r = 4.4, 7.3, 10.2 \text{ cm}$) and 2×2 *endcap* annuli (at $z = \pm 34.5, \pm 46.5 \text{ cm}$). Outside the pixel layers are the strip layers, consisting of 9.3 million silicon strips arranged into barrels and endcaps. The resolution in $r\phi$ varies between 10 and 50 μm , depending on the location and pitch of the given strip. Certain strip layers contain two layers of strips, rotated through a *stereo* angle (100 mrad) with respect to each other. By matching adjacent hits, the stereo measurement can add a third dimension (z for barrel, r for endcap) to the strip's 2D measurement, with resolution 100-500 μm . There are a total of 10 barrel layers ($0.2 < r < 1 \text{ m}$) and 24 endcap layers ($0.6 < |z| < 2.8 \text{ m}$). Figure 2.5 shows that the thickness of the tracker is between 0.4 and 2 electron radiation lengths (X_0).

Pixels with a signal greater than a tuneable readout threshold (typically around $3000Q_e$) are read out. These pixels are then aggregated with adjacent signals to form pixel clusters, which are further subjected to readout thresholds ($\sim 4000Q_e$). The exact position of the particle in this layer (known as a *hit*) is inferred by fitting the charge distribution of the pixels in this cluster to pre-determined templates. A similar method is employed to determine the strip hit positions, with some modifications to

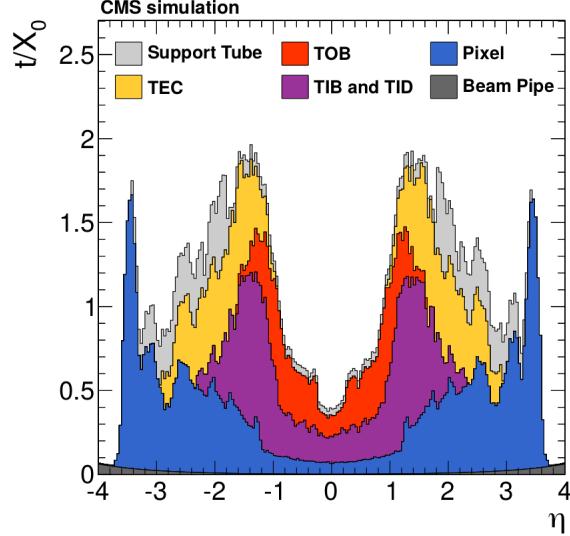


Figure 2.5: The thickness of the various tracker subsystems and other structures as a function of η . The y -axis is in units of electron radiation length X_0 . Reprinted from Reference [55].

account for Lorentz drift of the charges in the silicon detector due to the B -field. The efficiency of reconstructing hits varies with the detector type, location, and particle momentum, but is generally greater than 99% (99.5% if defective modules are not considered).

Tracking

Tracks are found using an iterative *inside-out* process, where each iteration has five steps:

1. Define seeds using pixel hits, double-strip hits (i.e. hits with 3D information), and an estimate of the beam spot (collision point). At least 3 hits are needed for the seed.
2. Use a Kalman filter [56, 57] to evolve track seeds through the rest of the tracker and find hits, accounting for the B -field, energy loss, and multiple scattering.
3. Estimate trajectory parameters after finding all hits.

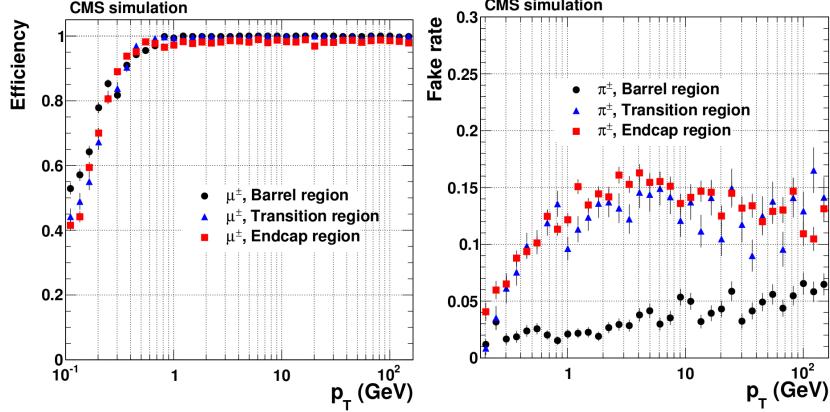


Figure 2.6: Efficiency (fake rate) of the CMS track fit algorithm, evaluated using simulation of muons (charged pions). Reprinted from Reference [55].

4. Decide whether to keep found tracks based on quality requirements (e.g. number of missing hits, track χ^2)
5. Remove hits associated with tracks from hit collection and repeat.

The trajectory parameters referred to in step 3 are the 5 parameters of a helix: ρ (curvature), ϕ_0 (azimuthal angle), λ ($\cot\theta$), d_0 (*impact parameter*, minimum r of track), z_0 (minimum $|z|$ of track). The CMS track fit typically has 5-7 iterations, with each successive iteration loosening the seed and track fit requirements to look for more difficult tracks (e.g. missing hits, large d_0). The efficiency and fake rate of this reconstruction, as a function of track p_T , are shown in Figure 2.6. For muons with $|\eta| < 1.5$ and $p_T > 1$ GeV, the tracking efficiency is over 98%, with a combinatorial fake rate of 2-6%.

Vertexing

The excellent position resolution of the pixel detector is used to accurately measure the position of primary vertices, as well as any secondary vertices from the decays of longer-lived particles with lifetimes $\gtrsim 10^{-13}$ s. In the former case, tracks are first clustered together on the basis of the likelihood that the tracks in a cluster arise from a single primary vertex. This is done using a deterministic annealing algorithm [58], which has as free parameters the number of clusters and the probability of each

track belonging to each cluster. Having determined the clusters, an adaptive fit algorithm [59] is used to determine the vertex for each cluster. The free parameters of this fit are the three spatial coordinates of the vertex. As the LHC collides bunches of $\mathcal{O}(10^{11})$ protons, we expect multiple primary vertices in a single collision, and this is reflected in Figure 2.7. The vertex defined to be the hard scattering interaction (known as *the* primary vertex³) is the vertex which maximizes:

$$\sum_{j \in \text{track jets}} (p_T^j)^2 + (p_T^{\text{miss}})^2 \quad (2.5)$$

where “track jets” refer to jets (Section 2.3.4) clustered from the vertex’s tracks, and p_T^{miss} is defined in Section 2.3.6. If there are more than 2 charged particle tracks in an event, the efficiency of reconstructing and identifying the correct primary vertex (PV) is greater than 0.995. The efficiency does not depend strongly on the total number of reconstructed primary vertices N_{PV} .

Secondary vertexing

The last reconstruction algorithm concerning the tracker alone is the identification of secondary vertices (SVs), which arise from the decay of longer-lived particles. The longer-lived particles in the SM are τ leptons and hadrons containing b and c quarks. The inclusive vertex fitter (IVF) [60] reconstructs such secondary vertices by the following steps:

1. Select a track as a seed if it satisfies $\sqrt{d_0^2 + d_z^2} > 50 \mu\text{m}$ and $d_0 > 1.2\delta d_0$.
2. Choose nearby tracks based on their closest distance to and opening angle with the seed track.
3. Fit the tracks to a displaced vertex using the adaptive fitter [59].
4. Decide which tracks belong to the candidate secondary vertex and which belong to the primary vertex.

³This nomenclature is indeed confusing, defining the singular primary vertex to be one of many primary vertices. However, it is standard terminology in CMS publications, so we will continue to use it. In what follows, the distinction will be clear

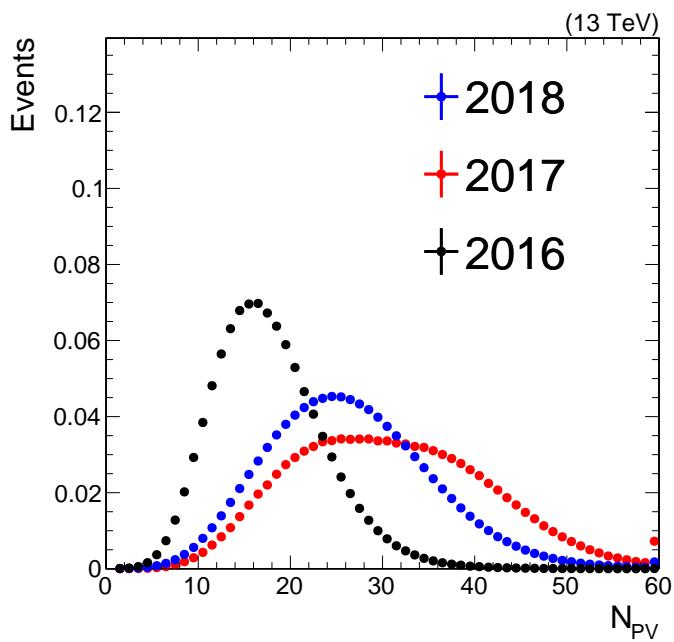


Figure 2.7: The distribution of the number of reconstructed primary vertices in data recorded by CMS during Run 2 of the LHC. While the results in this thesis only concern 2016 data, we show the evolution of N_{PV} as a function of time, as this correlates directly with increased instantaneous luminosity.

5. Re-fit the secondary vertex position only using the former set of tracks from the previous step.

It is important not only to properly determine the location of the secondary vertex, but also to properly assign tracks. Observables that are a function of the *tracks* (e.g. vertex mass) will be critical for b jet tagging.

2.2.2 Electromagnetic calorimeter

The CMS electromagnetic calorimeter [61] (ECAL) is a homogenous detector with good energy and angular resolution, composed of 76,000 PbWO₄ crystals. The crystals are arranged in two sections: a cylindrical barrel (EB) covering $|\eta| < 1.44$ and two endcap annuli (EE) extending to $|\eta| < 3$. This provides slightly more coverage than the tracking volume. Each crystal in the EB (EE) has dimensions $2.2 \times 2.2 \times 23$ ($2.68 \times 2.68 \times 22$) (cm³), with the long dimension pointing towards the beam. This can be compared to a Moli  re radius $r_M = 2.19$ cm and a radiation length of $X_0 = 0.89$ cm. A cross-sectional area comparable to $r_M \times r_M$ facilitates the differentiation of different electromagnetic (EM) showers arising from electrons and photons. The depth of the crystal (in units of X_0) drives the excellent energy resolution, which is determined using an electron beam:

$$\frac{\sigma_E}{E} = \frac{2.8\%}{\sqrt{E/\text{GeV}}} \oplus \frac{12\%}{E/\text{GeV}} \oplus 0.3\% \quad (2.6)$$

Scintillation photons from the PbWO₄ crystals are collected by avalanche photodiodes (APDs) in the EB and vacuum phototriodes (VPTs) in the EE, which provide amplification factors of 50 and 10, respectively.

At high momenta, the two photons from a π^0 decay may merge into a single ECAL crystal. This primarily occurs at high $|\eta|$ due to the z -boost of the initial state. To differentiate one- and two-photon deposits, a *preshower* detector sits in front of the EE ($1.6 < |\eta| < 2.5$). The preshower detector consists of a lead absorber and silicon strips. A photon (or photon pair) initiates a shower in the lead. The shower can be

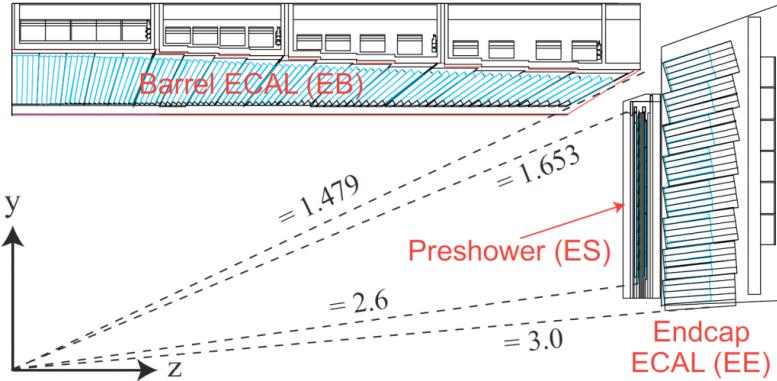


Figure 2.8: One quadrant of the CMS ECAL (symmetric with rotation around z and reflection across $z = 0$). The dashed lines indicate values of η . Reprinted from Reference [61].

resolved in the silicon strips, which have resolution $\mathcal{O}(1\text{--}10)$ mm.

The physical placement of all three ECAL components is shown in Figure 2.8.

Due to the bending of a charged particle's trajectory in the solenoidal B -field and interactions with the tracker material, bremsstrahlung photons will be emitted at similar values of η , but spread along ϕ . A *supercluster* (SC) is defined by clustering nearby ECAL energy depositions, allowing for a wider spread in ϕ than in η (Figure 2.9). The particle's EM energy is defined to be the weighted sum of the energies of all crystals in the SC, where the coefficients account for crystal-specific calibration effects [62]. For an electron or photon, the EM energy is typically the energy of the particle, whereas for other particles (charged hadrons and muons), it is only a fraction of the total energy.

2.2.3 Hadronic calorimeter

The hadronic calorimeter (HCAL) [64, 65, 66] is used to identify and measure the energy of hadrons. It consists of 4 calorimeters: barrel (abbreviated HB, covering $|\eta| < 1.4$), endcap (HE, $1.3 < |\eta| < 3$), forward (HF, $3 \lesssim |\eta| < 5$), and outer (HO, $|\eta| < 1.3$). Their arrangement is shown in Figure 2.10.

The HB and HE are both composed of alternating absorber and plastic scintillator layers. The absorber is a non-magnetic brass alloy with an interaction length $\lambda_I = 1.5$

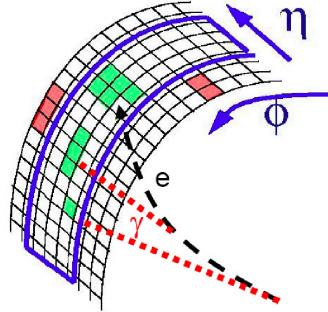


Figure 2.9: The combination of multiple ECAL crystals into a single supercluster, intended to capture energy depositions from bremsstrahlung photons. Reprinted from Reference [63].

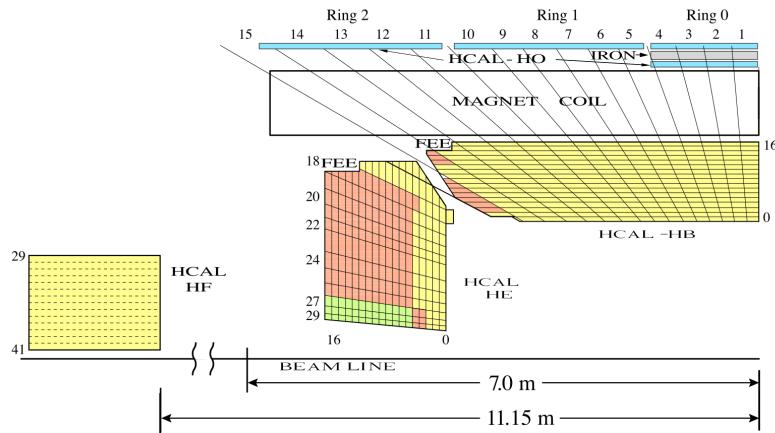


Figure 2.10: One quadrant of the CMS HCAL (symmetric with rotation around z and reflection across $z = 0$). Note that the slight overlap of the detectors in η ensures the hermeticity of the detector. Reprinted from Reference [66].

cm. Absorber layers range in thickness from 40 to 75 mm in the HB and HE, providing a total of 5.8-10.6 λ_I of material, depending on the η of the particle. These dimensions are limited by the constraint that the HB and HE be inside the solenoid. To augment the number of interaction lengths, additional layers of plastic scintillator are located outside of the solenoid. These comprise the HO, which use the magnet as an absorber, providing an additional $\sim 1.1 \lambda_I$. The light from the scintillator tiles is read out by hybrid photodiodes (HPDs) in the HB and HE and by silicon photomultipliers (SiPMs) in the HO.

Beyond the HE, at 11 m from the interaction point, is the HF. The HF is also a sampling calorimeter, made of steel absorbers instrumented with quartz fibers. Charged particles from the nuclear shower in the steel emit Cerenkov radiation as they traverse the quartz fibers, which transports the light to photomultiplier tubes (PMTs). The HF uses more radiation-hard materials than the rest of the HCAL because this η range is subject to significantly more radiation from collisions than the central part of the detector (Figure 2.11).

Each HCAL subsystem is read out in *towers*. A tower corresponds to a small segment in η and ϕ and traverses the perpendicular dimension. The direction of a particle is determined from the tower location. For $|\eta| < 1.74$, the segmentation is $\eta \times \phi = 0.087 \times 0.087$. Beyond this, the towers vary in size from 0.09×0.175 to 0.35×0.175 .

The energy resolution of the HCAL must be considered in conjunction with the ECAL, as energy is deposited in both detectors. This is calibrated using test beams of various charged particles. For the HB and HE, the combined ECAL+HCAL resolution is:

$$\frac{\sigma_E}{E} = \frac{0.847}{\sqrt{E/\text{GeV}}} \oplus 0.074 \quad (2.7)$$

The region instrumented by the HF does not have ECAL coverage, and so the resolution for the HF alone is:

$$\frac{\sigma_E}{E} = \frac{1.98}{\sqrt{E/\text{GeV}}} \oplus 0.09 \quad (2.8)$$

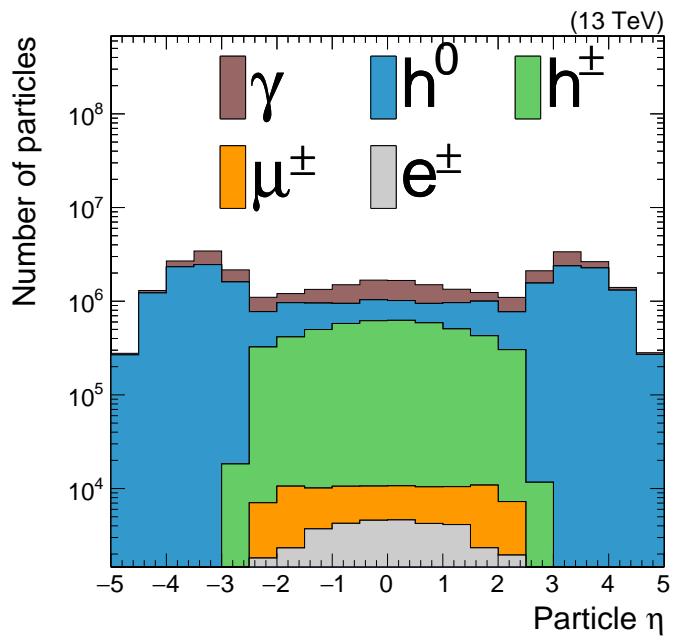


Figure 2.11: Flux of various particle types as a function of η in the detector. The symbol h refers to any type of hadron. Note that h^0 (γ) is used to refer to any neutral or charged hadron (photon or electron) in the forward region, where the tracker and ECAL do not have coverage.

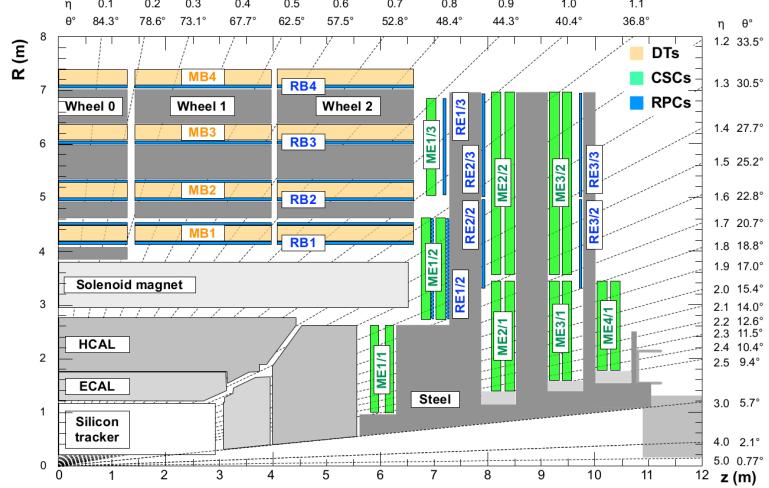


Figure 2.12: One quadrant of the CMS muon detection system, with the DTs (MBs), CSCs (MEs), and RPCs (RBs/REs) labeled. Reprinted from Reference [67].

2.2.4 Muon chambers

The muon detectors are gas ionization chambers and are the outermost component of CMS [67]. The chambers are interleaved with the steel return yoke of the CMS magnet, ensuring a B -field which runs anti-parallel to the field inside the solenoid. This results in the characteristic S -shape of muon trajectories, as the bending changes direction across the solenoid. 3 types of ionization detectors are used: drift tubes (DTs; barrel), cathode strip chambers (CSCs; endcaps), and resistive plate chambers (RPCs; barrel and endcaps). The physical placement of the detectors is shown in Figure 2.12.

The DT cells are filled with a 85:15 mix of argon and carbon dioxide, with a gold/steel anode wire held at a voltage of 3600 V. The cell is a rectangular prism, with transverse dimensions of $42 \times 13 \text{ mm}^2$ and a longitudinal dimension ranging from 1.9 to 4.1 m. The dimensions and the drift speed of $55 \mu\text{m}/\text{s}$ result in a maximum response (drift) time of 400 ns. The DTs are organized into cylindrical *stations* (MB). With the exception of the outermost barrel station (MB4), each MB consists of 3 *superlayers* (SLs). Each SL has 4 parallel drift cells, and so measure position in a particular plane. Of the 3 SLs, two measure r - ϕ position and one measures r - z

position. MB4 does not have an r - z SL.

The endcap is instrumented with CSCs as these have a faster response time and better spatial resolution than DTs. This is needed in the forward region, where both the muon and background fluxes are higher. While the CSC dimensions vary depending on the position of the chamber, each is instrumented with 80 cathode strips, held at voltages (relative to the anode) of 2.9-3.6 kV. The CSC wires have a separation of 2.5-3.16 mm, which governs the position resolution. A 50:40:/10 mix of CO₂/Ar/CF₄ is used to fill the chambers.

RPCs are interspersed among the DTs and CSCs in both the barrel and endcap. These serve as a very fast muon detector (~ 1 ns) for the online trigger system. The spatial resolution of the RPC hits is worse than the DTs and CSCs.

The hit position resolution of the DTs is 78-120 μm (140-390 μm) in r - ϕ (r - z). For CSCs, the resolution varies from 40 to 152 μm . The efficiency of individual reconstructed hits (*rechits*) is over 95%.

2.2.5 Online trigger system

Figure 2.13 shows various cross sections of pp collisions as a function of \sqrt{s} . While it is clear that interesting SM processes increase as a function of \sqrt{s} , their cross sections are still several orders of magnitude below the inclusive cross section $\sigma_{\text{tot}} \sim 10^8$ nb. To produce an appreciable number of rare events, the LHC has bunch crossings every 25 ns (40 MHz). However, it is not possible to read out the entire detector at this rate, much less reconstruct the data and store it. CMS uses a two-stage trigger system [68] to refine the events to keep for permanent storage and analysis. First, a Level 1 (L1) hardware-based trigger selects interesting events at a rate of 100 kHz. The L1 makes these decisions using incomplete detector information, in order to reduce the read-out and computation time. Selected events are then fed to the high level trigger (HLT) to partially reconstruct the events on a CPU farm. The HLT uses the full granularity of the detector. The final selected data rate from the HLT is 400 Hz.

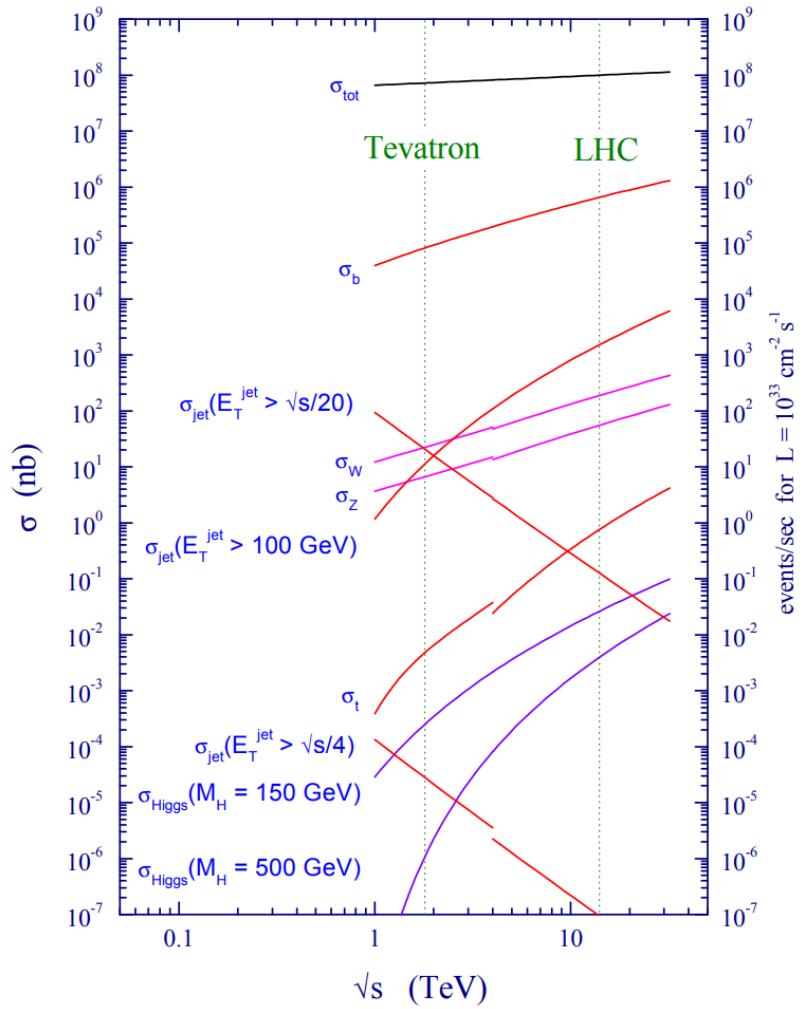


Figure 2.13: Various cross sections of interesting SM processes compared to the inclusive cross section of pp collisions, at various values of \sqrt{s} . Reprinted from Reference [69].

Level 1

The L1 trigger makes decisions within $4 \mu\text{s}$ of collisions using field programmable gate arrays (FPGAs) and application specific integrated circuits (ASICs), which offer significant speed advantages as compared to CPUs for certain tasks. Individual detector systems (ECAL, HCAL, muon detectors) feed simple reconstructed objects (*trigger primitives* (TPs)) to a series of regional trigger decisions. Quality selections are placed on calorimeter towers, and they are aggregated into clusters of energy deposits. A simple segment-finding and tracking algorithm is run on hits in the muon chambers to produce muon tracks. Note that the inner tracker is not included in the L1 decision: at the time of the construction of the CMS detector, the detector readout and reconstruction algorithms were not fast enough for the L1's requirements.⁴ Regional trigger decisions are sent to the global trigger (GT), which correlates the TPs it receives. Some GT trigger decisions are localized, such as requiring a high-energy ECAL deposit (e.g. e, γ). Others require computing event-wide observables, such as a large momentum imbalance (e.g. neutrinos, DM). The GT also computes simple jets. If the GT decides to select the event, the detector is read out and forwarded to the data acquisition system (DAQ). Figure 2.14 shows a schematic description of this process.

High level trigger

The full detector readout is picked up by the HLT. The HLT computing farm consists of 20k CPU cores (this number has grown from 13k in 2012 and continues to grow). Unlike the L1, the HLT has access to the entire event information. Therefore, a full reconstruction of the event is performed, using algorithms similar to those used in offline reconstruction. To be selected by the HLT, an event must pass an HLT *path*. A path consists of a series of filters, each making simple decisions that can be chained into complex decisions. For example, three filters may be: (1) two forward jets, (2) large momentum imbalance, and (3) an electron and a muon. These filters could

⁴It should be noted that implementing tracking for hardware triggers is a significant goal for the community in the next few years.

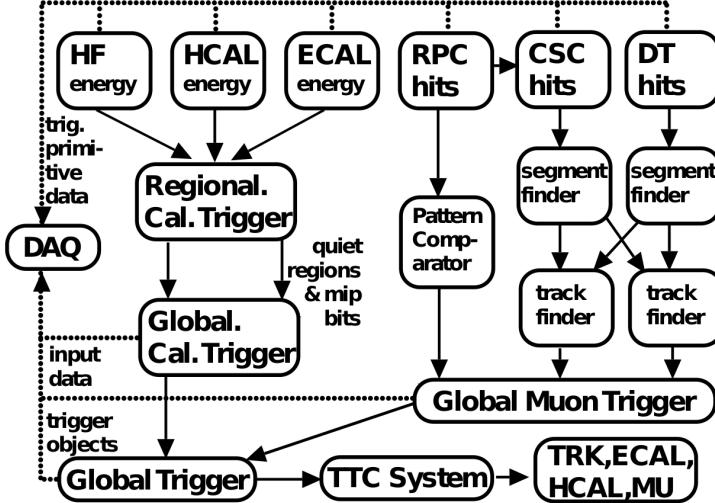


Figure 2.14: Schematic diagram of the CMS L1 trigger system. In addition to the region and global trigger decisions, the flow of data also includes the trigger, timing, and control system (TTC). Reprinted from Reference [68].

be chained into two separate paths, targeting vector boson fusion production of the Higgs, where the Higgs decays to DM (1&2) or two taus (1&3). Simple decisions (calorimeter and muons) are computed before complex decisions (tracking). Events that are selected by the HLT are passed on to the Tier 0 computing farm for full offline reconstruction, and then onwards to disk and tape resources for analysis and storage, respectively.

2.3 Particle reconstruction and identification

2.3.1 Particle flow algorithm

The excellent angular granularity of the CMS detector and the momentum resolution of the tracker are leveraged by the *particle flow* [65] algorithm, which correlates information from all detector subsystems to build a global description of each event. Particle flow (PF) algorithms date back to ALEPH [?].

The key feature of the PF algorithm is to *link* multiple detector signals together into a single PF candidate. This linkage combines inner tracks, ECAL clusters, HCAL

clusters, and muon tracks based on their proximity in the (η, ϕ) plane. Inner track helices are extended into the calorimeters, searching for clusters compatible with the trajectory. Similarly, clusters from the ECAL, the ECAL preshower, and the HCAL can be linked without a track present. The remainder of this section is organized according to *blocks* in the PF algorithm. At the end of each block, any detector signature (track, calorimeter cell) which has been assigned to a PF candidate in that block is removed from the set of objects passed to the next block. For example, tracks associated with muons will not be considered when reconstructing charged hadrons.

2.3.2 Muons

The first block of the PF algorithm links inner tracks and muon chamber hits to identify muons. While *standalone* muons can rely solely on muon chamber hits, we also consider muons that use the inner tracker. To construct outside-in *global muons*, standalone muon tracks are first reconstructed using a Kalman filter fit using muon chamber rechits [70, 67]. These standalone tracks are extrapolated inwards to the inner tracker, accounting for effects from the magnetic field and the material between the tracker and the muon chambers. A second Kalman filter fit is run to combine the inner track with the standalone track to form a global muon. To reject backgrounds, requirements are placed on the χ^2/N_{dof} (poorly fit tracks, charged hadrons) and d_0, d_z (muons from hadron decays and cosmic rays). The total efficiency of global muon reconstruction is 99%. *Tracker* muons are reconstructed from the inside-out, extrapolating inner tracks to the muon chambers and re-fitting with muon rechits. Figure 2.15 compares the momentum resolution of muons reconstructed using the different algorithms. The PF selection consumes all three types of reconstructed muons: global, tracker, and standalone. In this thesis, we will only use global muons.

The PF selection is intentionally non-restrictive, so as to minimize the rate of not identifying a real muon. To reduce the rate of mis-identified charged hadrons, we define two muon identification (ID) criteria that are slightly stronger than the PF selection. The first is a loose ID, which is used to veto muons when defining muon-free data samples. A loose muon is required to be either a global or tracker muon and

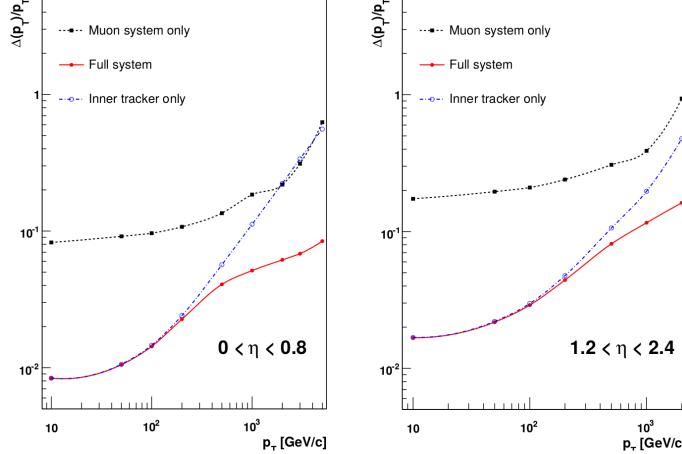


Figure 2.15: Resolution of muon p_T , compared between the different reconstruction algorithms. The salient feature is that the muon systems become important to the momentum measurement at approximately 200 GeV. Reprinted from Reference [67].

have PF isolation less than 0.25. PF isolation is defined as:

$$\left(\sum_{i \in \text{PV charged had.}} p_T^{(i)} + \max \left\{ 0, \sum_{i \in \text{neut. had.}} p_T^{(i)} + \sum_{j \in \gamma} p_T^{(j)} - \frac{1}{2} \sum_{k \in \text{PU charged had.}} p_T^{(k)} \right\} \right) / p_T^{(\mu)} \quad (2.9)$$

where the sums are over PF candidates within $\Delta R < 0.4$ of the muon. PV and PU refer to the primary vertex and pileup vertices, respectively. The tight muon ID requires a global muon, PF isolation less than 0.15, and has selections on the criteria in Table 2.1.

To account for differences between data and simulation in the performance of the muon IDs, corrections (known as scale factors) are derived:

$$\text{SF} = \frac{\epsilon_{\text{Data}}}{\epsilon_{\text{MC}}} \quad (2.10)$$

where ϵ refers to the fraction of muons that are correctly identified by the ID criteria. While ϵ_{MC} is computed directly from MC truth information, ϵ_{Data} is extracted from data using the *tag-and-probe* method [71]. The standard approach starts with $Z \rightarrow \mu\mu$ candidate events with exactly two PF muons. One muon (the *tag*) is selected using some stringent ID criteria. Then, the events are categorized by whether the second PF

Table 2.1: Observables used in identifying muons and rejecting backgrounds

Observable	Notes
p_T	Backgrounds grow at low p_T .
χ^2/N_{dof}	Ensure a good track fit.
$N_{\text{hit}}^{\mu\mu}$	At least one hit in the muon chamber in the global fit.
N_{stations}	At least two muon chamber stations contain segments of the track.
d_0	Track impact parameter, remove cosmic rays and hadron decays.
d_z	Track impact parameter, as above but also for pileup.
$N_{\text{hit}}^{\text{pixel}}$	At least one pixel hit.
$N_{\text{hit}}^{\text{tracker}}$	At least 5 hits in the tracker.

muon candidate (the *probe*) in the event also passes the ID criteria with efficiency. The number of $Z \rightarrow \mu\mu$ events in each category is measured by fitting the $m_{\mu\mu}$ distribution, which strongly distinguishes between the falling charged hadron background and the Z peak. The desired ID efficiency is then:

$$\epsilon_{\text{Data}} = \frac{N_{\text{pass}}^Z}{N_{\text{pass}}^Z + N_{\text{fail}}^Z} \quad (2.11)$$

Note that the efficiency of the tag criteria does not affect ϵ_{Data} as the two muons are statistically independent. The tag definition is chosen to balance the statistical power and the purity of the sample.

2.3.3 Electrons and photons

Both electrons and photons are seeded using ECAL SCs, with electrons also being associated with an isolated track. Due to the significant material budget of the silicon tracker, electrons will lose a significant amount of energy through bremsstrahlung. Although this energy is partially recovered through the supercluster algorithm, the standard Kalman filter tracking fit does not properly account for the non-Gaussian hit uncertainties induced by bremsstrahlung. Therefore, a modified algorithm based on a Gaussian Sum Filter (GSF) [55, 72] is used to fit the track. Hits in the pixel and strip endcap layers consistent with the ECAL SC are used to seed the GSF fit. GSF differs

Table 2.2: Observables used in identifying electrons and rejecting hadron and photon backgrounds

Observable	Notes
p_T	Backgrounds grow at low p_T and brem. photons can make low- p_T tracking difficult.
$\sigma_{i\eta i\eta}$	Energy-weighted width of cell η in SC. Small for electrons.
$ \Delta\eta(\text{track, SC}) $	$\Delta\eta$ between SC seed crystal and GSF track at PV. Small for electrons.
$ \Delta\phi(\text{track, SC}) $	$\Delta\phi$, as above.
E_H/E_{EM}	Ratio of HCAL and ECAL energies. Large for hadrons.
PF isolation	Sum of energies of other PF candidates near electron. Large for particles in jets.
$ 1/E - 1/p $	Checks if ECAL and tracker agree ($m_e \sim 0$)
$N_{\text{hit}}^{\text{miss}}$	Conversions or bad tracks will have multiple missing hits in the inner tracker.
Conversion veto	Check for a pair of tracks originating at a displaced vertex.

from a Kalman filter by modeling the uncertainties as a Gaussian mixture model instead of a single Gaussian. The primary backgrounds for electron identification are (1) the overlap of a charged hadron with a neutral hadron or photon and (2) a photon which converts into a e^-/e^+ pair in the tracker. Table 2.2 lists the observables used to reject these backgrounds; two cut-based IDs are defined. The *loose* ID selects electrons with $\sim 90\%$ signal efficiency and $\sim 0.5\%$ background acceptance (strongly dependent on the electron phase space); it is used to *veto* electrons. The *tight* ID selects electrons with $\sim 70\%$ signal efficiency and $\sim 0.1\%$ background acceptance; it is used to *select* electron-pure samples.

Photons are defined as ECAL SCs without matched tracks. Table 2.3 describes the selection variables used to define the loose veto ID ($\epsilon_{\text{sig}} \approx 90\%$, $\epsilon_{\text{bkg}} \approx 17\%$) and the tight selection ID ($\epsilon_{\text{sig}} \approx 82\%$, $\epsilon_{\text{bkg}} \approx 12\%$). The direction and energy of a photon are defined by the ECAL SC position and energy, respectively.

To calibrate ECAL SCs, we define the $R_9 = E_{3\times 3}/E_{\text{SC}}$. $E_{3\times 3}$ is the sum of the energies of the crystals in a 3×3 square centered on the most energetic crystal in the SC. Much like $\sigma_{i\eta i\eta}$, it is sensitive to the width of the shower shape. A regression

Table 2.3: Observables used in identifying photons and rejecting hadron backgrounds

Observable	Notes
p_T	Backgrounds grow at low p_T .
η	EB resolution is better than EE; less hadron background.
$\sigma_{\eta\eta}$	Defined in Table 2.2.
E_H/E_{EM}	Defined in Table 2.2.
PF isolations	Defined in Table 2.2. For photons, separate isolation criteria are placed on each PF type (photon, charged hadron, neutral hadron).

to correct the energy scale [62] is trained as a function of SC energy, η , R_9 , and the width of the SC in ϕ . Differences between data and MC in the efficiencies of the electron and photon IDs are corrected using scale factors using $Z \rightarrow ee$ events, as was described for muon IDs. In the case of the photon ID, the electron SC is used as a proxy for the photon.

2.3.4 Hadrons and jets

The remaining final states considered by the PF algorithm are hadrons. These are primarily found in jets, along with nonisolated photons from π^0 decays and muons from heavy hadron decays. If a calorimeter cluster is within the tracker acceptance and not already consumed by an earlier PF block, and it is not linked to a track, then it is assumed to have arisen from a photon or neutral hadron. If the cluster is in the ECAL, then the constructed PF candidate is a photon; if in the HCAL, it is a neutral hadron. Proximal ECAL and HCAL clusters are not linked together when constructing neutral hadrons. This is because neutral hadrons are expected to leave very little energy in the ECAL, and the tracking efficiency for charged hadrons is over 90%. To find charged hadrons, HCAL clusters, ECAL clusters, and tracks are linked together. The calorimeter energy of the clusters is estimated using a calibrated sum of the HCAL and ECAL deposits. If this energy is significantly (500 MeV) above the momentum estimate from the tracks, then it is assumed there are additional neutral hadrons or photons in the cluster. The number of charged hadrons is the number of

reconstructed tracks. Outside of the tracker acceptance, ECAL clusters not linked to HCAL clusters are assigned to PF photons. Linked HCAL and ECAL clusters are assumed to arise from a shower containing neutral and charged hadrons.

Jet clustering

Jets arise from the hadronization and fragmentation of colored particles, as described in Section 1. The precise definition of a jet is dependent on the algorithm used to cluster particles, and so we will use *jet* to refer to the final states of hadronization/fragmentation and to the output of a jet finding algorithm. LHC experiments typically use *sequential recombination* algorithms [73, 74, 75]. Given a set of PF candidates in the event E , we compute two metrics:

$$d_{iB} = p_{T,i}^2$$

$$d_{ij} = \min\{p_{T,i}^{2q}, p_{T,j}^{2q}\} \frac{\Delta R(p_i^\mu, p_j^\mu)^2}{R} \quad (2.12)$$

where $i, j \in E$; B represents the beam; and p and R are tunable parameters. Let $\tilde{E} = E \cup \{B\}$ and find $i, j \in \tilde{E}$ that minimize d_{ij} . If $j = B$, then we remove particle i from \tilde{E} and add it to the set of candidate jets. Otherwise, we combine i and j into a new particle k by defining $p_k^\mu = p_i^\mu + p_j^\mu$. This *pseudojet* k is added to \tilde{E} , while i and j are removed. This process is repeated until \tilde{E} is exhausted. The exact value of R is an approximate measure of the ΔR radius of the jet. The value of q defines the relationship between the momentum and angular factors; $q < 0$ enforces clustering circular jets around hard seeds. For standard single-parton jets, CMS uses $q = -1$ (referred to as anti- k_T) and $R = 0.4$ (AK4). In Section 3.1, we will discuss the use of $q = 0$ and $R \gg 0.4$ for multi-parton jets. CMS software uses the FastJet library [76] for efficient implementations of sequential jet clustering. In particular, FastJet reduces the $\mathcal{O}(N^2)$ computation of Equation 2.12 to $\mathcal{O}(N \log N)$, where $N \sim \mathcal{O}(10^2 - 10^3)$.

Jet calibration

While individual PF candidates are individually calibrated using detector subsystem-specific corrections, it is still necessary to calibrate each jet as a whole [77]. There are 3 steps in the additional calibration:

1. A pileup correction (L1) is derived from MC truth information and randomly-triggered data events. It subtracts energy from jets, as a function of the jet p_T , η , A (area), and ρ . A is computed by adding a uniform distribution of infinitesimally soft particles (*ghosts*) to the event prior to the jet clustering, and counting how many are clustered with each jet. A wide jet will have many such particles and will be more susceptible to pileup contamination. The event-wide quantity ρ is a measure of the median p_T per unit area in the event, also computed using ghosts.
2. The L2L3⁵ correction accounts for detector response biases by using the GEANT4 simulation of the detector (see Section 2.4). Two jet collections are computed in the simulation: PF jets (using PF candidates) and truth jets (using the particle collection from the hard scattering and showering simulation). The energy scale of PF jets is corrected (as a function of p_T, η) to match that of truth jets.
3. The L2L3 residual correction is only applied to data and corrects for small differences between the real detector and the simulation used in MC. It is derived using $Z(\rightarrow \ell\ell) + \text{jet}$, $\gamma + \text{jet}$, and dijet events. In each case, one well-measured object (Z , γ , jet) is used to calibrate the recoiling jet. While the Z events provide the best reference object (μ energy uncertainties are very small), the small cross section introduces large statistical uncertainties at high p_T and η .

The p_T dependence of the corrections and the associated uncertainties are shown in Figure 2.16.

⁵The terminology is historical, referring to previously-factorized corrections L2 and L3

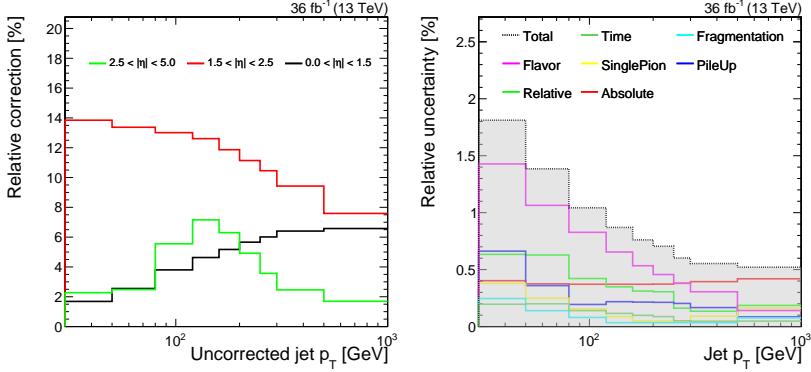


Figure 2.16: Left: the total effect of the L1 and L2L3 jet energy scale corrections in MC, as a function of p_T, η . Right: the corresponding uncertainties on the JES corrections, in the same p_T binning and averaged over η .

Heavy flavor jet ID

It is frequently useful to identify the type of parton that induces a jet. In the context of the results presented in this thesis, identifying b quarks (b -tagging) is important, but identifying other initial states is also possible (c quarks, gluons vs quarks). The hadronization of a b quark involves the production of a b hadron, which have lifetime of $\tau \sim 1.5$ ps. The lab frame displacement is $\gamma\beta c\tau = p\tau/m$. A B meson with $p_T = 50$ GeV will have a mean transverse displacement of 4.3 mm, which is well within the vertexing resolution of the pixel detector. Therefore, the signature of a b jet is the identification of an SV displaced from the PV, with properties consistent with a heavy hadron (Figure 2.17).

The reconstruction of the SV is described in Section 2.2.1. The Combined Secondary Vertex (CSVv2) tagger [60] is an artificial neural network trained to distinguish b jets from $u/d/c/s/g$ jets. Nineteen characteristics of the jet (e.g. mass of tracks associated with the SV, presence of soft leptons from the SV, d_0 of the SV) are used to train the NN⁶. For the analyses described in this thesis, a jet is considered b -tagged if it satisfies $\text{CSVv2} > 0.54$. Discrepancies between data and MC in the efficiency of the b jet ID are corrected for using SFs, derived separately for b and non- b jets. Two b -enriched samples are used for the former task. The first is a relatively pure,

⁶See Section 4.1.2.1 of Reference [60] for a full list.

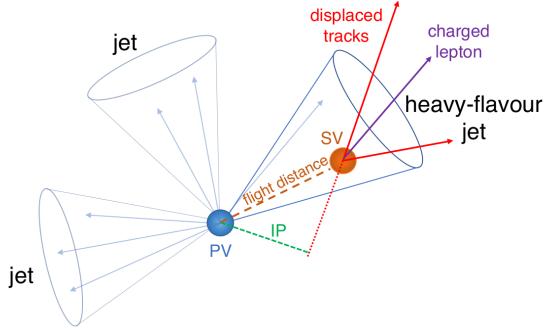


Figure 2.17: Location of a secondary vertex inside a b jet. Reprinted from Reference [60].

but statistically limited, sample of $t\bar{t}$ events, in which the b quarks from the t quark decays are used. The second is a much more statistically powerful sample of $g \rightarrow b\bar{b}$ events, which is selected by triggering on and identifying a non-isolated muon from the decay of one of the b hadrons. Figure 2.18 compares the CSVv2 response in data and simulation, as well as the SFs used to correct this distribution.

2.3.5 Hadronic taus

A hadronic tau (τ_h) is defined as a tau lepton which decays to one ν_τ and one or more hadrons. A dedicated algorithm [65, 78] is used to extract τ_h candidates from the AK4 jet collection. Only jets within the tracker acceptance and with $p_T > 13$ GeV are used. Combinations of particles within the jet are considered for compatibility with decay modes consisting of a tau neutrino, 0-2 neutral pions, and 1 or 3 charged hadrons. While other hadronic final states exist, they have a fairly small branching ratio and are not considered. The π^0 will present in the jet as two photons; in some cases, a photon may convert into an e^-e^+ pair. In the case of multi-hadron final states, compatibility with intermediate resonance masses is checked. The highest $p_T \tau_h$ candidate in the jet is selected as a PF τ_h . To reduce the large combinatorial background from quark and gluon jets, stringent requirements are made on the PF isolation. A true τ_h should be fairly well-isolated, whereas a combinatorial fake will be surrounded by a parton shower.

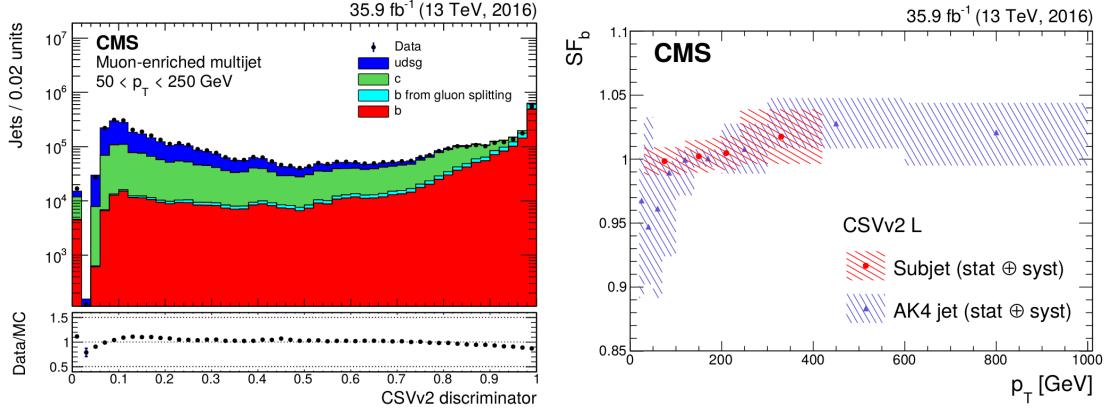


Figure 2.18: Left: The distribution of the CSVv2 response of AK4 jets in a dijet sample enriched with $g \rightarrow b\bar{b}$ events. Right: data/MC scale factors to correct for discrepancies in the simulation of the b jet ID. AK4 jet SFs and uncertainties are shown in blue; overlaid in red are the SFs for subjets (see Chapter 3). Reprinted from Reference [60].

2.3.6 Missing momentum

The PF missing momentum (p_T^{miss}) is defined as the magnitude of the missing momentum vector:

$$\vec{p}_T^{\text{miss}} = - \sum_{i \in \text{PF cands.}} \begin{pmatrix} p_x^{(i)} \\ p_y^{(i)} \end{pmatrix} \quad (2.13)$$

As the initial state has no net transverse momentum, conservation of momentum requires the same to be true of the final state. In a perfectly reconstructed event, non-zero p_T^{miss} implies the presence of non-interacting particles such as neutrinos or DM candidates.

Fake energy reconstruction or failure to properly reconstruct energy deposits can lead to large, fake p_T^{miss} . Such events are removed through a series of filters:

- HCAL and ECAL filters, identifying events with calorimeter clusters caused by noise. The shape and timing of the energy distribution can be used to identify noise, as well as detector-specific information (such as known problematic ECAL crystals).
- Beam halo filter, identifying energy deposits from muons traveling parallel to

the beam. These muons are produced from pre-collision interactions between the beam and the machine. They are identified by their localization in ϕ and longitudinal signature in the ECAL/CSCs.

- Reconstruction filters, identifying failures of the PF algorithm to properly reconstruct particles. In some cases, muons can be double-counted or mis-identified as charged hadrons and muons.

These filters remove effectively all anomalous p_T^{miss} events while rejecting less than 1% of events with real p_T^{miss} .

As the p_T^{miss} indirectly depends on the momentum of each PF candidate that enters the sum, any *ad hoc* calibrations must be propagated. Most energy calibrations are intrinsic to the PF algorithm, but the jet energy scale corrections are not. Therefore, the missing momentum is accordingly updated:

$$\vec{p}_T^{\text{miss}} \mapsto \vec{p}_T^{\text{miss}} + \sum_{j \in \text{jets}} \left(\vec{p}_T^{j,\text{corr.}} - \vec{p}_T^{j,\text{raw}} \right) \quad (2.14)$$

2.3.7 Pileup mitigation

Two algorithms are used to mitigate the effects of pileup in jets: charged hadron subtraction (CHS) and pileup per-particle identification (PUPPI) [79]. CHS simply removes charged particles not from the primary vertex from the set of PF candidates fed into the jet clustering. This is only able to remove charged pileup contamination for jets within the tracker acceptance.

The PUPPI algorithm defines a local shape α_i for every particle i in the event that is independent of particle location or charge. The distribution of α_i is determined for central, charged particles from the PV and from PU. This is then extrapolated to neutral particles and forward particles, to assign a probability $P(\text{PU}|\alpha_i)$. More concretely, the local shape is defined:

$$\alpha_i = \log \sum_{j \neq i} \frac{p_{T,j}}{\Delta R_{ij}} H(\Delta R_{ij} - R_{\min}) H(R_0 - \Delta R_{ij}) \quad (2.15)$$

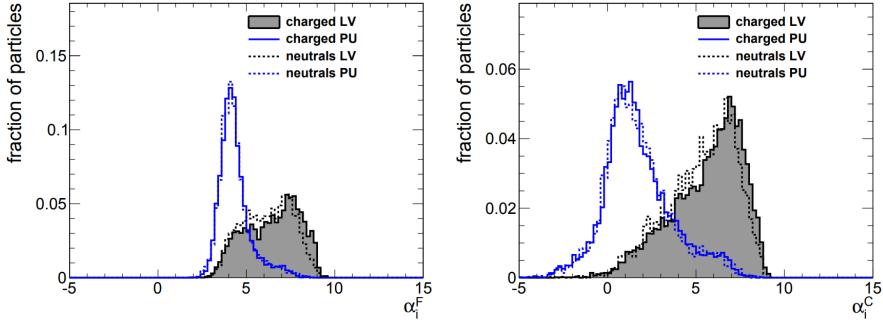


Figure 2.19: Distribution of PUPPI local shapes for charged and neutral particles from the PV and PU vertices. Reprinted from Reference [79].

where R_{\min}, R_0 are tunable parameters and H is the Heaviside step function. Up to rescaling, α_i is the sum of p_{TS} of particles in an annulus around particle i . It is expected to be larger for PV particles than PU particles, as PU radiation is uniformly distributed, whereas PV radiation is centered around hard partons. This is illustrated in Figure 2.19, which shows α_i^F (same as α_i) and α_i^C (same as α_i , except sum only charged particles).

Then, we define:

$$x_i = H(\alpha_i - \bar{\alpha}) \frac{(\alpha_i - \bar{\alpha})^2}{\sigma^2} \quad (2.16)$$

where $\bar{\alpha}$ and σ are the median and RMS, respectively, of the charged PU α_i distribution, with some extrapolation for η_i and q_i . It is found that α_i has a Gaussian distribution for pileup, and so x_i should have a χ^2 distribution with 1 degree of freedom. Therefore, we define the probability of pileup as:

$$w_i = 1 - P(\text{PU}|\alpha_i) = \begin{cases} 1 & \text{if } |\eta| < 2.5, q_i \neq 0, \text{ from PV} \\ 0 & \text{if } |\eta| < 2.5, q_i \neq 0, \text{ not from PV} \\ P(\chi^2 < x_i | N_{\text{dof}} = 1) & \text{otherwise} \end{cases} \quad (2.17)$$

The four momenta of particles are scaled by the PUPPI probability, i.e. $p_i^\mu \mapsto w_i p_i^\mu$.

In the context of the results described in this analysis, it is found that CHS provides sufficient pileup mitigation for AK4 jets, but as described in Chapter 3,

PUPPI will be necessary for CA15 jets.

2.4 Simulation of CMS

The CMS offline software suite (CMSSW [80]) uses Geant4 [81, 82] to simulate the detector response to particles produced in collisions. As described in Section 1, particle-level MC simulates the hard scattering and parton shower. Multiple simulated proton-proton collisions are overlaid into a single event to mimic the effect of pileup. The generated particles are then interfaced to Geant4, which simulates the passage of a particle through the magnetic field, the energy deposited as the particle interacts with the detector material, and the evolution of any additional particles produced (e.g. EM showers). The readout electronics' response to the detector signature is then simulated. The software used to reconstruct the simulated electronic response is identical to that used for real data. This minimizes any reconstruction differences between MC and data. Additional truth information from the generators is retained for use in data analysis tasks.

Chapter 3

Hadronic Resonance Identification

In this chapter, we describe the reconstruction and identification of heavy ($\gtrsim 100$ GeV) resonances that decay to two or more quarks. Within the Standard Model, the only such resonances are the massive vector bosons ($W, Z \rightarrow q\bar{q}'$), the Higgs boson (typically $H \rightarrow b\bar{b}$), and the top quark ($t \rightarrow bW(\rightarrow q\bar{q}')$). The focus of this chapter is on the cases in which the resonance is boosted and the decay products merge, such that they cannot be identified as 2 or 3 distinct AK4 jets. In preparation for Chapter 4, we will take the top quark as a concrete example. The studies presented here can, and in some cases have been, applied to other heavy resonances, both within and beyond the Standard Model.

3.1 Reconstruction

The need for a dedicated boosted top jet reconstruction arises from the momentum-dependence of the angular separation between the quarks from a heavy resonance decay [83]:

$$\Delta R \sim \frac{2M}{p_T} \tag{3.1}$$

where M is the resonance mass and p_T is the resonance transverse momentum. If $p_T \gtrsim 300$ GeV and $M = m_t$, then $\Delta R = 1.2$. At this separation, three AK4 jets will begin to overlap, reducing the efficiency of reconstructing each of the daughter

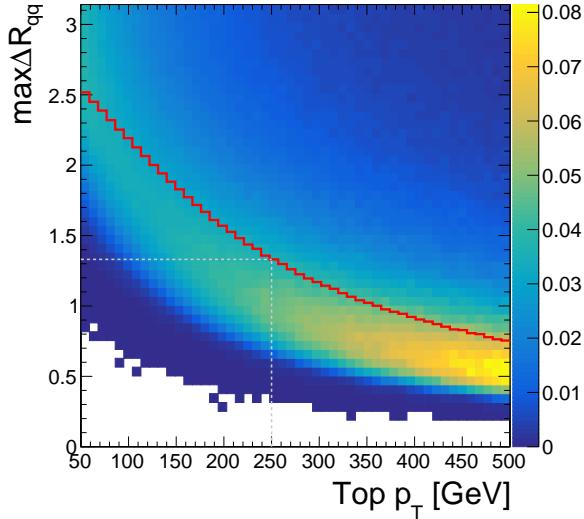


Figure 3.1: Distribution of top quark momenta versus decay radii in a simulated top quark pair sample. The events are weighted such that the inclusive momentum distribution is uniform. The z -axis units are arbitrary, but proportional to the distribution of jets. The solid red line marks the 50% quantile of jets at each value of p_T .

quarks separately. This is verified by checking the distribution of the *decay radius* in top quark simulation. Here, we define decay radius as:

$$\max \Delta R_{qq} \equiv \max_{0 \leq i < j \leq 2} \{\Delta R(q_i, q_j)\}, \text{ where } t \rightarrow q_0 q_1 q_2 \quad (3.2)$$

Using a broad spectrum of generated top quark p_T , Figure 3.1 shows the dependence of the decay radius on the top quark p_T .

The analysis described in Chapter 4 targets top jets with $p_T > 250$ GeV. Figure 3.1 and Equation 3.1 both indicate that at this momentum, a three-jet reconstruction is not optimal. In particular, Figure 3.1 shows that for $p_T \approx 250$ GeV, it is equally likely that a top quark's decay products will fall within a single $R = 1.5$ jet or that they will be resolvable as three separate jets. However, past this threshold momentum, the large-radius jet becomes the preferred reconstruction option. This motivates the use of $R = 1.5$ jets to reconstruct hadronic top quarks with $p_T > 250$ GeV.

There are two tunable parameters in jet reconstruction. We have specified the jet radius, but we must also choose the jet algorithm. The anti- k_T algorithm tends

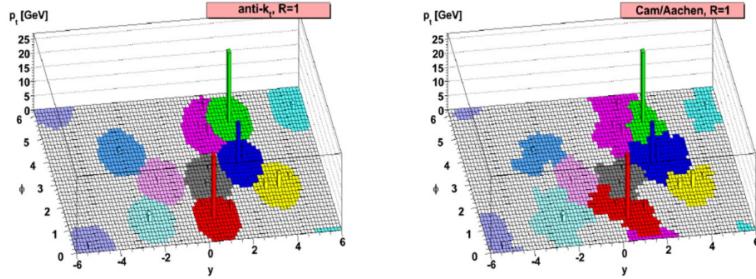


Figure 3.2: Jets clustered using the anti- k_{T} (left) and CA (right) algorithms. Shown is the y - ϕ plane of a hypothetical calorimeter, unrolled onto a flat surface. The height of each cell represents the p_{T} of the particle. The anti- k_{T} jets tend to be more circular when compared to the CA jets. Figures are adapted from [73].

to pick circular jets, whereas the Cambridge-Aachen (CA) algorithm results in more geometric shapes (Figure 3.2). As the top jets we seek to reconstruct are the sum of three light quark jets, we do not necessarily expect the $R = 1.5$ jet to be circular. Figure 3.3 compares the jet mass distribution for top and light quark/gluon (LQG) jets, where the jets are clustered using both algorithms. CA produces a top jet mass distribution with a narrower peak closer to m_t than anti- k_{T} , which is why we choose the CA algorithm. Hereafter, we will refer to Cambridge-Aachen $R = 1.5$ jets as CA15 jets.

The distance parameter of $R = 1.5$ corresponds approximately to a maximal azimuthal angle separation of $\pi/2$, which can cover half of the detector’s fiducial volume. As the jet is so large, particles from pile-up interactions may be clustered into a jet from the primary vertex. Fundamental quantities (like top quark momentum) are uncorrelated with N_{PV} , but reconstructed quantities acquire such a dependence due to the extra radiation. These additional particles bias the energy scale of the jet, as well as other observables. To mitigate these effects, we scale the particles’ 4-momenta by their corresponding PUPPI scores (described in Chapter 2) prior to clustering the jet. Jets clustered without PUPPI weighting have a jet mass and τ_{32}^{SD} distributions (Figure 3.4) in which both the mean and variance are dependent on N_{PV} . Subjecting particles to PUPPI prior to clustering strongly reduces this dependence.

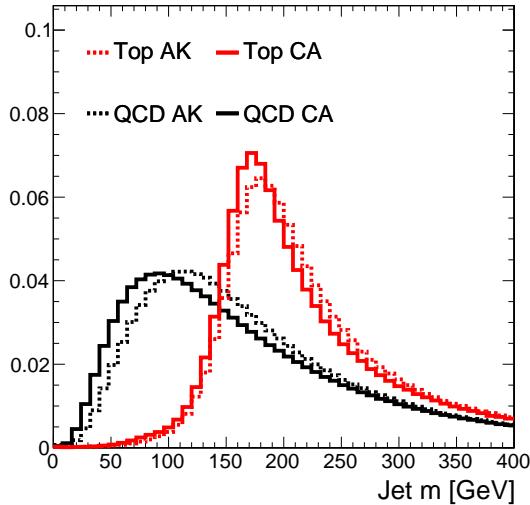


Figure 3.3: Mass distribution for jets clustered using the anti- k_T (dashed) and CA (solid) algorithms. QCD refers to jets originating in QCD multijet events, i.e. from the hadronization of light quarks or gluons. The CA top peak is slightly sharper than the anti- k_T peak, is closer to m_t , and the QCD background is slightly lower near the peak.

3.2 Identification

Having reconstructed the candidate top quark jets, we turn to the problem of *identifying* which CA15 jets originate from top quarks as opposed to light q/g hadronization. As indicated in Figure 3.3, the jet mass is a powerful observable, but top (LQG) jets do not necessarily have a mass of m_t ($m_q, m_g \sim 0$). While some of this discrepancy is caused by mismeasurement of the jet energy scale, a substantial fraction originates from extra radiation being absorbed into the jet. These extra particles arise from pile-up, ISR, and underlying event (UE). Pile-up particles are already accounted for by PUPPI. Many algorithms exist to *groom* ISR and UE particles from a jet after it has been clustered; here, we will discuss and use the soft drop (SD) method [84]. SD removes parts of the CA clustering tree which represent very wide-angle or asymmetric splittings, which are atypical for a parton shower. More formally, at each node in the clustering tree, the softer subjet of the node will be removed if it satisfies the

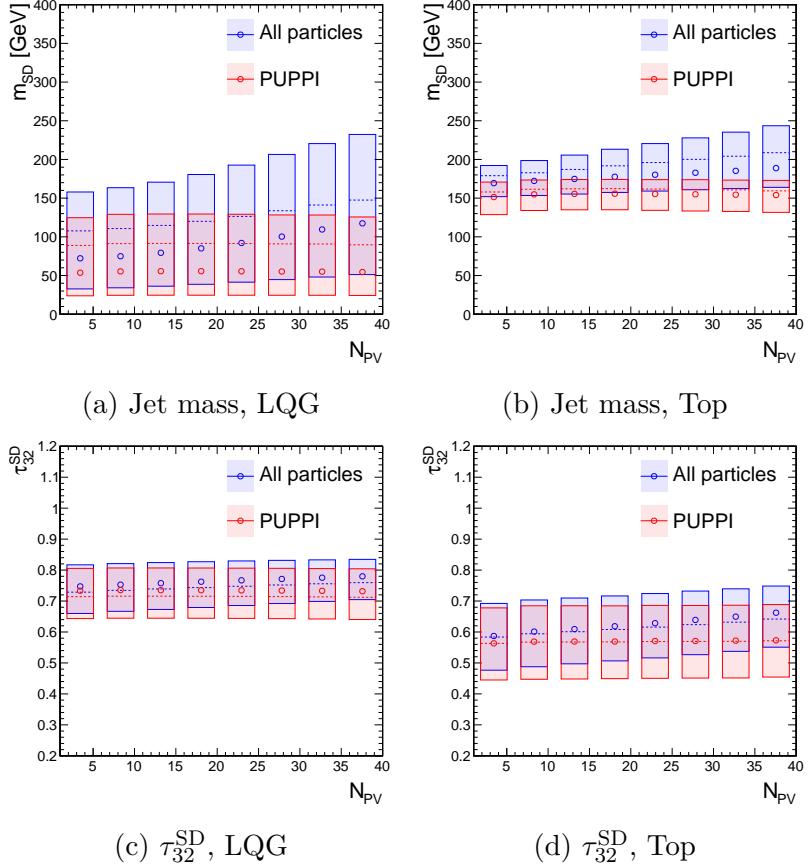


Figure 3.4: Stability of two CA15 jet observables (described in Section 3.2) as a function of N_{PV} . The median (mean) of each N_{PV} bin is represented by an open circle (dashed line), while the [25%, 75%] percentile range is shown with a box.

condition:

$$\frac{\min(p_{T,1}, p_{T,2})}{p_{T,1} + p_{T,2}} < \left(\frac{\Delta R_{12}}{R}\right)^\beta \quad (3.3)$$

where $p_{T,i}$ refers to the p_T of the i -th subjet of the node; ΔR_{12} is the R -distance between the two subjets; and R and β are tunable parameters. This process starts at the root node of the clustering tree (i.e. the whole jet) and proceeds iteratively to the leaves (i.e. individual particles). This condition is satisfied if the two subjets are very far apart (assuming $\beta \geq 0$) or if the splitting is very asymmetric in momentum. We define the *SD subjets* (or where clear, simply *subjets*) of a jet to be the two branches of the root node, after branches failing the SD condition have been removed. The particles remaining after this grooming procedure are combined to make the *groomed* or SD jet.

We then define m_{SD} as the mass of the SD jet. Observables may also be defined in terms of the groomed or ungroomed jet. Figure 3.5 compares the ungroomed and groomed mass distributions in top and LQG jets, as a function of jet momentum. It is immediately clear that grooming provides (a) a sharper mass peak in top jets at m_t and (b) a smoothly falling mass distribution in LQG jets that goes to 0. Furthermore, SD ensures the stability of the mass distribution as a function of jet p_T , especially in LQG jets. For these reasons, m_{SD} will be our standard definition of jet mass.

3.2.1 Substructure

A substructure observable is any function of a jet's constituents that is sensitive to the multi-pronged structure of a heavy resonance decay. In addition to jet mass and b -tagging, substructure is used to reject LQG jets as top decay candidates.

N -subjettiness

The N -subjettiness (τ_N) is a measure of the compatibility of a jet with an N -axis hypothesis [85]. It is defined as:

$$\tau_N = \frac{\sum_{i \in \text{jet}} p_{T,i} \min\{\Delta R_{ia} | a \in A\}}{\sum_{i \in \text{jet}} p_{T,i} R} \quad (3.4)$$

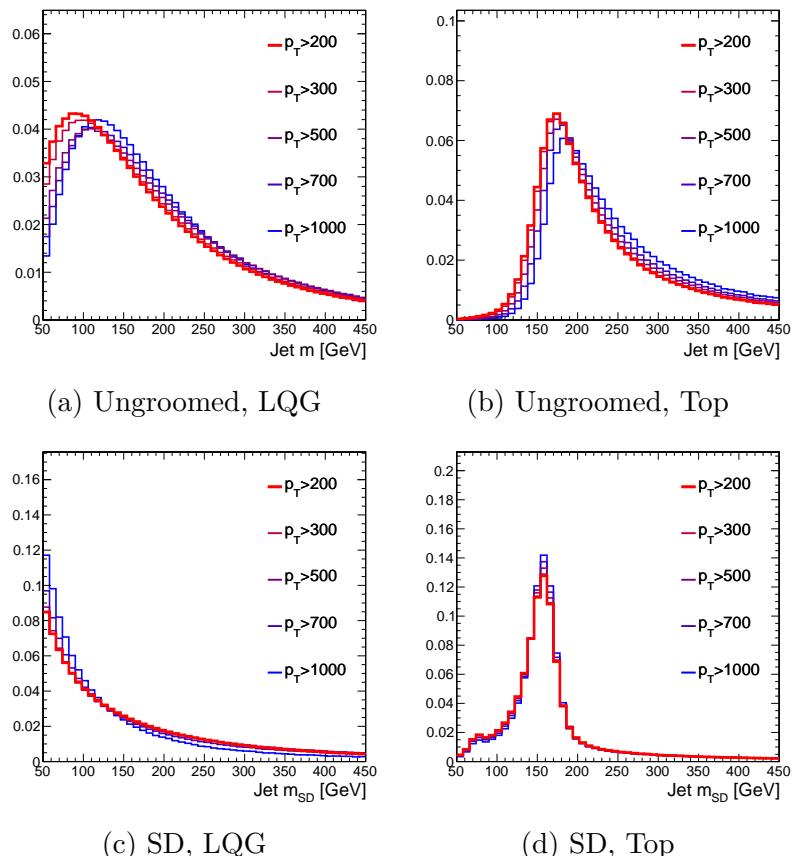


Figure 3.5: Distribution in simulation of ungroomed and groomed jet mass in CA15 jets originating from LQG or hadronic top decays. The multiple histograms represent increasingly stringent p_T requirements on the truth parton that initiates the jet.

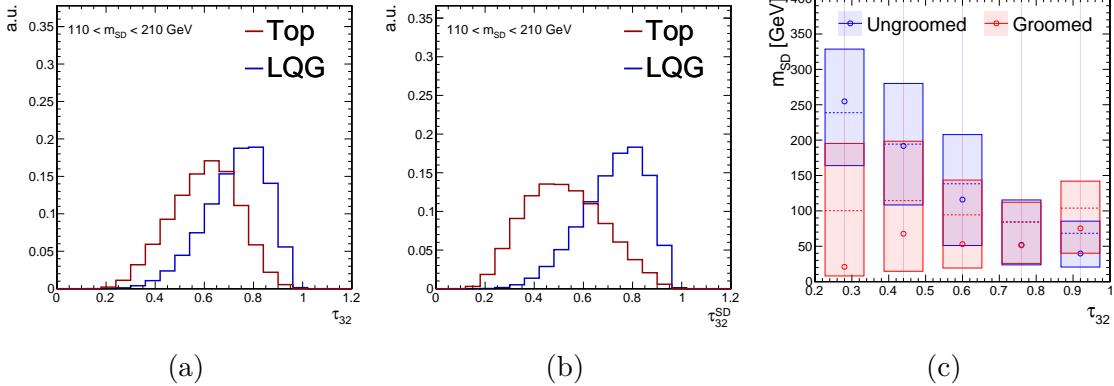


Figure 3.6: Shape of ungroomed (left) and groomed (center) τ_{32} distributions in top and LQG jets, with a mass selection consistent with m_t . Right: the correlation between τ_{32} and m_{SD} in LQG jets, comparing groomed and ungroomed jets.

where $R = 1.5$ (the jet radius); ΔR_{ia} is the ΔR distance between the particle i and the axis a ; and A is a set of N axes. Ideally, A would be defined to be the set of axes that minimize τ_N for each jet, but this minimization problem is computationally difficult. Instead, the exclusive k_T algorithm is used to partition the jet's constituents into N subjets¹. The set of axes A is taken to be the directions of the N k_T subjets. Since the k_T distance metric is proportional to $\Delta R^2/R^2$, this approximates the ideal minimization. Assuming the axes are chosen to minimize Equation 3.4, $\tau_N \leq \tau_M$ for $N > M$. A small τ_N indicates a high degree of compatibility with the N -axis hypothesis. Therefore, we expect a top quark jet with three prongs to satisfy $\tau_3 \ll \tau_2$, whereas a 1-pronged LQG jet should satisfy $\tau_3 \lesssim \tau_2$. Correspondingly, we take $\tau_{32} \equiv \tau_3/\tau_2$ to be the tagging observable.

Figure 3.6 shows the distribution of τ_{32} . As with jet mass, we may calculate τ_{32} either on the whole jet or on the SD jet. The discrimination between top and LQG jets is similar in both cases, but as Figure 3.6c demonstrates, τ_{32}^{SD} has the weaker correlation with m_{SD} in LQG jets. This feature will be critical to validate any tagger in data, as described in Section 3.3.

¹These are distinct from the SD subjets discussed above.

HEPTopTagger

The HEPTopTagger algorithm de-clusters the jet into many subjets and attempts to reconstruct the W and t decay products out of these subjets [86]. The computation of the tagging variable f_{rec} can be simplified into three steps (a more detailed description is found in the appendix of Reference [86]):

1. Compute subjets of the CA15 jet. This is done in a fashion similar to the SD subjets discussed above, but instead of taking the two subjets of the root node, the tree is traversed until some lower p_T bound is crossed.
2. Test all triplet combinations of the found subjets and define the m_{123} as the groomed mass of the trijet system.
3. Choosing the triplet most consistent with a 3-body top decay (see Equation 12 in Reference [86]), define:

$$f_{\text{rec}} = \min_{0 \leq i < j \leq 2} \left| \frac{m_{ij}/m_{123}}{m_W/m_t} - 1 \right| \quad (3.5)$$

where the indices i, j index elements of the selected triplet.

Figure 3.7 shows the distribution of the selected m_{123} and f_{rec} , although we will only use the latter as a tagging observable. Note that we do not define distinct groomed and ungroomed versions of these observables, as grooming is already applied when defining subjets.

Energy Correlation Functions

Energy correlation functions measure the correlation of the positions of hard particles in a jet [87]. Heuristically, an N -point ECF is small if the hard particles can be grouped into fewer than N prongs and large if they arise from N or more prongs. An N -point ECF, with angular parameters α and β , is defined as:

$$e(\alpha, N, \beta) \equiv {}_o e_N^\beta = \sum_{K \subset J, |K|=N} \left[\prod_{i \in K} \frac{p_T^{(i)}}{p_T^{(j)}} \right] \times \min \left\{ \prod_{i,j \in P} \Delta R_{ij}^\beta \mid P \subset \tilde{K}^2, |P| = o \right\} \quad (3.6)$$

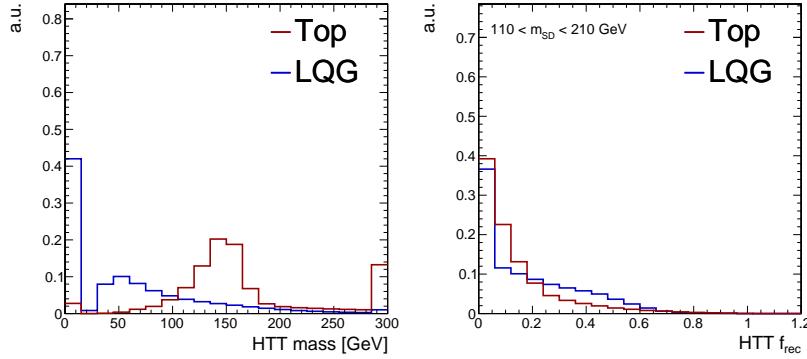


Figure 3.7: Shape of the m_{123} and f_{rec} observables computed by the HEPTopTagger algorithm. The value $m_{123} = 0$ indicates the cases in which a subjet triplet consistent with a top decay could not be found.

where $0 < o \leq N^2 - N$ and \tilde{K}^2 indicates all pairs of distinct particles in K . The proposed tagger in Reference [87] is:

$$N_3^{(\beta)} = \frac{e(2, 4, \beta)}{(e(1, 3, \beta))^2} \quad (3.7)$$

Figure 3.8 shows N_3 for various values of β ; given our desire for stability as a function of jet p_T and mass, we only consider ECFs computed on the SD jet. The discrimination power of this ECF ratio is roughly comparable to that of τ_{32}^{SD} . N_3 is motivated by the behavior of 3- and 4-point ECFs in top and LQG jets:

- In top jets, $e(N = 4) \ll e(N = 3)$, since 3-point correlation functions are large in a 3-pronged jet
- In QCD jets, $e(N = 3) \sim e(N = 4)$, since both 3- and 4-point ECFs are weak in a 1-pronged jet

Therefore, taking the ratio $e(N = 4)/e(N = 3)$ constructs a useful observable.

While N_3 has a strong theoretical motivation, it is possible that other functions of ECFs distinguish between top and LQG jets. In order to construct observables that do not have a strong dependence on the jet p_T , we restrict ourselves to ratios of the form:

$$\psi(a, N, \alpha, b, M, \beta) = \frac{e(a, N, \alpha)}{(e(b, M, \beta))^x}, \text{ where } M \leq N \text{ and } x = \frac{a\alpha}{b\beta} \quad (3.8)$$

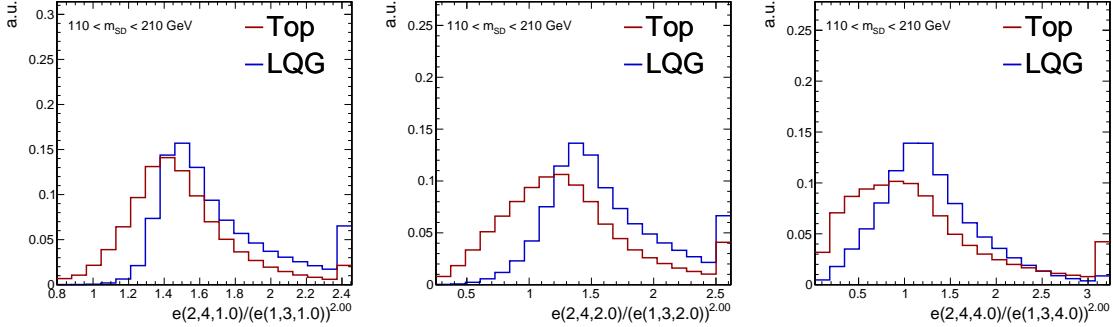


Figure 3.8: Shape of the N_3 observables in top and LQG jets, for various values of β .

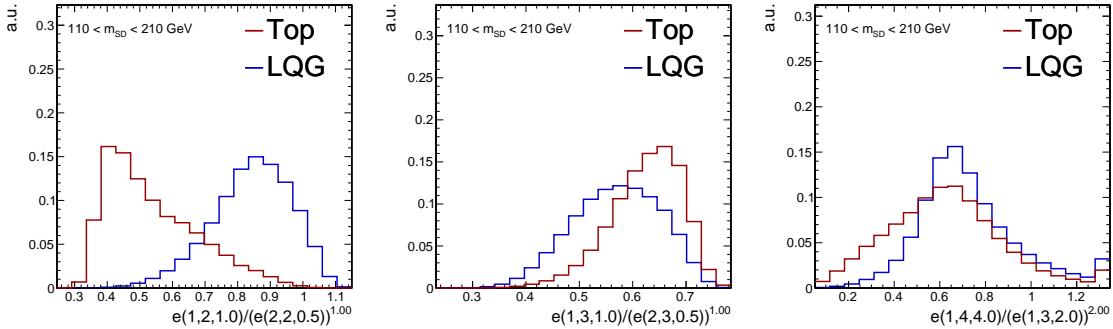


Figure 3.9: Examples of non-trivial ECF ratios other than N_3 that separate top and LQG jet distributions.

A large subset of this broader class of ECF observables are found to be useful (Figure 3.9), including ratios not of the form $e(N = 4)/e(N = 3)$.

3.2.2 A combined tagger

In principle, we have constructed an infinitely large space of jet substructure observables. In practice, we only consider a finite sampling of ECF parameters:

$$N \in \{1, 2, 3, 4\}$$

$$o \in \{1, 2, 3\}$$

$$\beta \in \{0.5, 1, 2, 4\} \quad (3.9)$$

This grid results in ~ 900 ψ observables.

Boosted decision trees

To build a single optimal observable out of all the $\{\psi_i\}$ s, we will use a boosted decision tree (BDT). A simplified algorithm to train a single decision tree node n for classification is as follows:

1. Choose a ψ_j , either by sampling randomly or selecting the one most optimal for the next step.
2. Based on the training data fed to the node, select a decision boundary d_n to optimize a loss function. For classification, we use cross-entropy:

$$\ell(X, y; j, d_n) = -\hat{\pi}_B \ln \hat{\pi}_B - \hat{\pi}_S \ln \hat{\pi}_S \quad (3.10)$$

$$\hat{\pi}_c = P(y = c | \psi_j < d_n) \quad (3.11)$$

A tree is built iteratively:

1. Train a node n using the above criteria.
2. If a stopping condition is not met, train one node on the samples that pass n and another on the samples that fail. Stopping conditions can take into account features including:
 - Number of nodes
 - Depth of tree
 - Relative change in $\ell(n)$

Figure 3.10 provides a pictorial example of how a decision tree can be built.

Decision trees can very accurately describe training data, but they also pathologically overfit the data. *Boosting* many shallow trees is a standard method to mitigate this and retain descriptive power. The simplicity of individual trees prevents overfitting, while boosting many trees allows for a complex model. The result of a BDT is a classifier $f_n(x) = \sum_{i=0}^n \nu^i T_i(x)$, where $\nu \leq 1$ is tunable and each T_i is a decision tree.

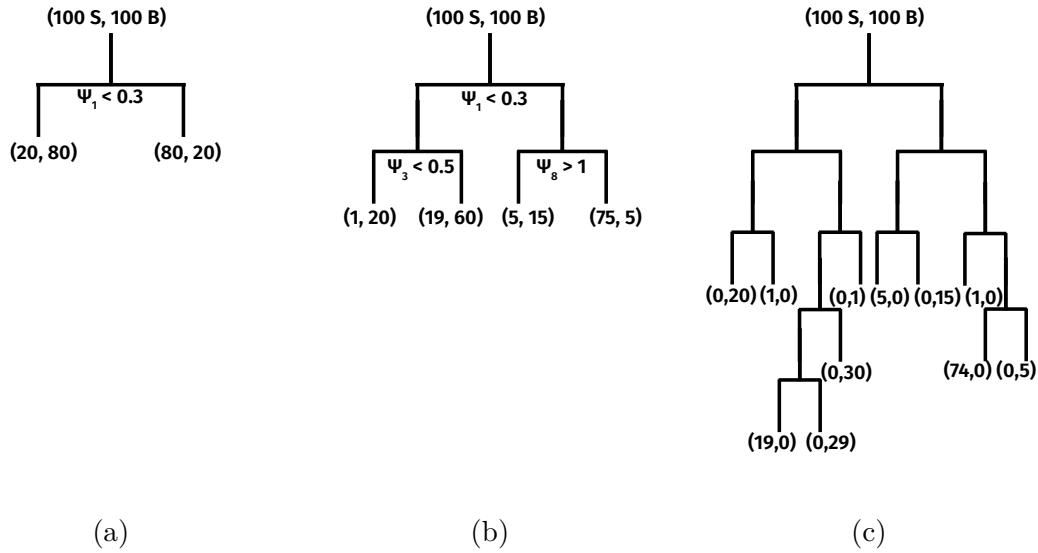


Figure 3.10: Steps in greedily training a simple decision tree. At each node, the number of signal and background samples is indicated by the tuple (S, B) . For example, the root node in step (a) is fed 100 signal and 100 background samples. The learning rule decides the optimal decision at this node is $\psi_i < 0.3$, which subdivides the 200 samples into two subsamples with composition (20, 80) and (80, 20). Step (b) shows a further partitioning of the training samples using additional features. Step (c) shows a fully trained tree, in which each leaf node only contains signal or background samples. Such a tree cannot be further refined, but earlier stopping conditions may also be used.

Each T_j is trained to correct for the residual error from the previously trained trees $f_j(x) = \sum_{i=0}^j \nu^i T_i(x)$. A simplified algorithm to train a BDT is as follows:

1. Define a global loss function, e.g.:

$$L(y_i; f_i) = \ln(1 + \exp(-y_i f_i)) \quad (3.12)$$

2. Train a single tree T_0 and initialize classifier $f_0 = T_0$

3. Until some stopping condition (index $m = 1, \dots, n$):

- 3.1. Compute the *residual* as the gradient of the global loss with respect to the classifier f_{m-1} :

$$r_{mi} = -\nabla_f L(y_i; f)|_{f=f_{m-1}(\psi_i)} \quad (3.13)$$

- 3.2. Fit a regression tree T_m to predict r_{mi} as a function of x_i :

$$\ell(X, r_m; j, d, \hat{r}) = \sum_{i|\psi_{ji} < d} (r_{mi} - \hat{r})^2 \quad (3.14)$$

- 3.3. Update $f_m = f_{m-1} + \nu T_m$

The decision tree training and boosting algorithms we use are provided by the Toolkit for Multivariate Analysis [88].

Training the BDT

While we would like to train a BDT on the entire space of $\{\psi\}$, there are two issues to be solved: poorly modeled ratios and a large feature space. Firstly, the descriptions of many ECF ratios in LQG jets are poorly-simulated by MC (Figure 3.11). More systematically, we can compute the CDF of each ψ and define a score:

$$-\log_{10} \text{KS}(F_{\text{data}}(\psi_i), F_{\text{MC}}(\psi_i)) = -\log_{10} \max |F_{\text{data}}(\psi_i) - F_{\text{MC}}(\psi_i)| \quad (3.15)$$

where F represents the CDF and KS denotes the Kolmogorov-Smirnov metric on probability distributions. The score is close to 0 for poorly-simulated distributions

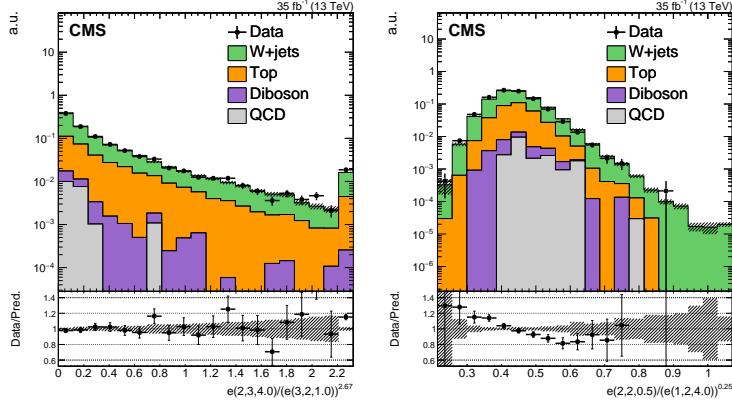


Figure 3.11: Two different ECF ratios in a W +jets selection, heavily enriched in LQG jets. One is fairly well-modeled, while the other is not.

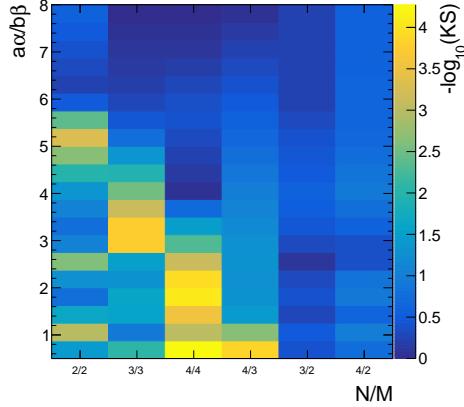


Figure 3.12: The $-\log_{10}$ KS metric as a function of N/M and $a\alpha/b\beta$, computed using events enriched in LQG jets.

and approaches ∞ as $F_{\text{MC}} \rightarrow F_{\text{data..}}$. Figure 3.12 parameterizes this as a function of N/M (ratio of the number of particles) and $a\alpha/b\beta$ (ratio of the angular powers) and shows an interesting structure. It is found that $3/2$ and $4/2$ ratios are uniformly poorly modeled, as are ratios with large $a\alpha/b\beta$. In what follows, we only consider ECF ratios with $-\log_{10} \text{KS} > 1$, which corresponds to the bottom-left of Figure 3.12.

Even after filtering poorly-modeled ratios, we are left with a sampled ψ grid of ~ 400 points. It is desirable to reduce the size of the feature space, as computing each ECF is somewhat computationally-intensive: an N -point ECF on a p -particle jet has $\binom{N}{p}$ terms. Note that standard pre-processing techniques, like principal component

analysis, do not reduce the number of features to be computed. Methods like L1 regularization [89] do force sparsity in the feature space, but cannot be trivially applied to BDTs. Therefore, we introduce a targeted iterative training method to solve this problem:

1. Train a BDT with trees T_1, \dots, T_n
2. For each ψ_i , define a score:

$$s_i = \sum_{m=1}^n \nu^{m-1} \sum_{\text{nodes using } \psi_i \text{ in } T_m} N_{\text{samples}}(\text{node}) \times (\ell(\text{node}) - \ell(\text{parent}))^2 \quad (3.16)$$

3. Remove one or more ψ_i with smallest s_i and repeat.

Iterative training is expensive and can require the training of $\mathcal{O}(50)$ BDTs. It is semi-parallelizable, and the entire process typically takes a few hours. However, as the inference samples are 1-2 orders of magnitude larger than the training samples, this method reduces the total CPU time needed to run an analysis. Figure 3.13 shows background acceptance rate at $\epsilon_{\text{sig}} = 0.5$ (a proxy for the global loss) as a function of feature space size. For illustrative purposes, we only show the range [1, 50]. The inputs for this training are the ECF ratios, as well as τ_{32}^{SD} and f_{rec} , which provide additional information.

The background acceptance reaches a minimum with 13 features, and so we choose this configuration as the final BDT. The features are:

$$\begin{aligned} & \frac{e(1, 4, 20)}{e(1, 3, 10)^2}, \frac{e(1, 4, 40)}{e(1, 3, 20)^2}, \frac{e(2, 4, 05)}{e(1, 3, 05)^2}, \frac{e(2, 4, 10)}{e(1, 3, 10)^2}, \frac{e(2, 4, 10)}{e(2, 3, 05)^2}, \frac{e(2, 4, 20)}{e(1, 3, 20)^2} \\ & \frac{e(1, 2, 20)}{e(1, 2, 10)^2}, \frac{e(1, 3, 40)}{e(2, 3, 20)}, \frac{e(3, 3, 10)}{e(1, 3, 40)^{3/4}}, \frac{e(3, 3, 10)}{e(2, 3, 20)^{3/4}}, \frac{e(3, 3, 20)}{e(3, 3, 40)^{1/2}} \end{aligned} \quad (3.17)$$

$$\tau_{32}^{\text{SD}}, f_{\text{rec}}$$

While a number of N_3 or other 4/3 ratios appear in this list, we find a number of 2/2 and 3/3 ratios to contribute meaningfully to the classification task as well. Figure 3.14 shows the distributions of all selected features. Figure 3.15 shows the

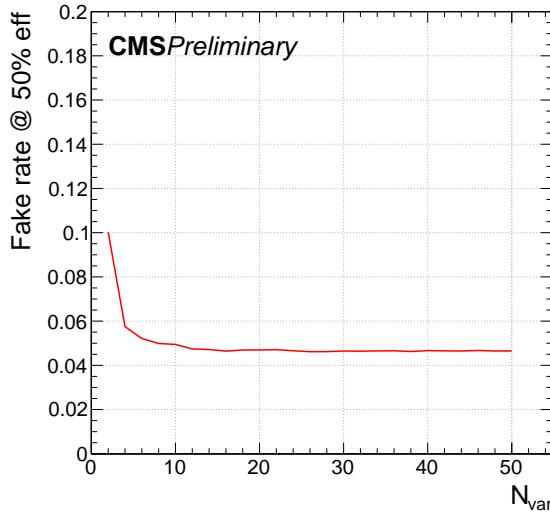


Figure 3.13: Performance of BDTs as a function of the number features used in the training. The performance is defined by the fraction of LQG jets accepted by the BDT when the signal efficiency is fixed at 50%.

background acceptance as a function of signal efficiency, comparing the final BDT (“Combined BDT”) to several other taggers. The standard LHC tagger prior to this work is τ_{32}^{SD} , which, by itself, has a worse performance than the BDTs. The Combined BDT’s performance is near that of a BDT trained with many ECFs, including poorly modeled ones (“50 ECF”). Removing τ_{32}^{SD} and f_{rec} from Equation 3.17 results in the “11 ECF” BDT. Therefore, the addition of these two features is critical to reaching the maximal performance achievable with the ECF ratio set, if description in simulation were not a concern. At fixed signal efficiency $\epsilon_{\text{sig}} = 0.5$, the combined BDT reduces the background acceptance by 30% relative to τ_{32}^{SD} .

3.3 Data validation

Prior to using the top BDT to identify top jets and reject LQG jets, we must verify that the simulation describes the BDT distribution properly as compared to data, and correct for any residual discrepancies. Figure 3.16 shows the BDT response and m_{SD} in top- and LQG-enriched selections. Top quarks are isolated by selecting events

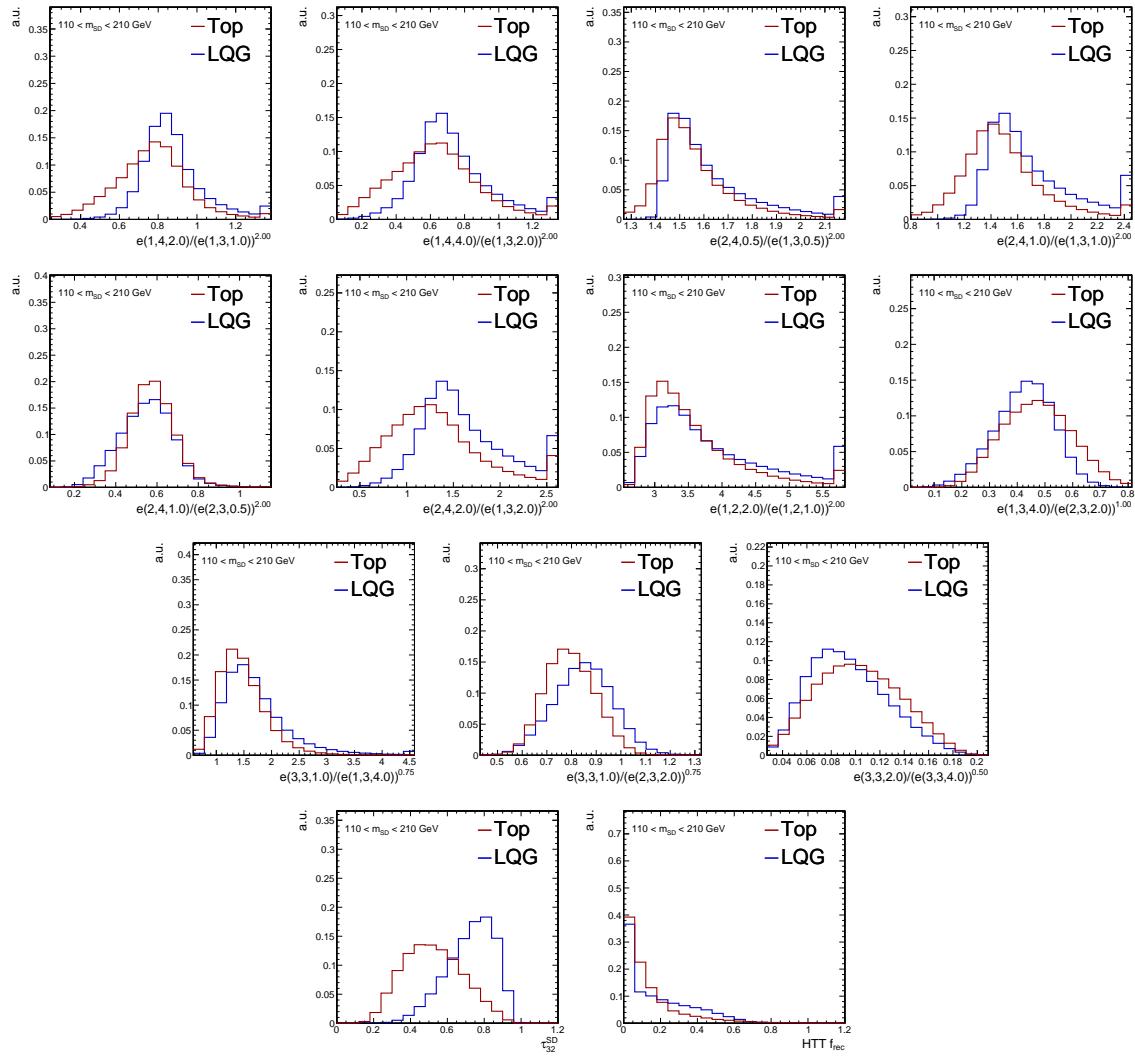


Figure 3.14: Distributions of the 13 features selected by the iterative BDT training

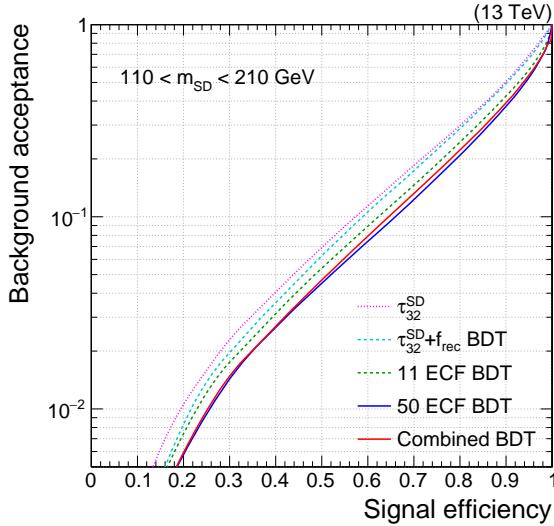


Figure 3.15: Receiver Operating Characteristic (ROC) curve comparing various top identification methods. The Combined BDT is the ID method chosen as the final tagger.

that produce $t\bar{t}$ pairs, in which one top quark decays hadronically (the top jet) and the other decays muonically ($t \rightarrow b\mu^+\nu_m u$). The leptonic t is selected by identifying the muon and b jet. We further require that the CA15 jet have $110 < m_{\text{SD}} < 210$ GeV and at least one SD subjet to be b -tagged. LQG jets are selected by using $Z(\rightarrow \mu\mu) + \text{jet}$ events. We require two opposite sign muons, with $|m_{\mu\mu} - m_Z| < 30$ GeV; this selection selects a $\gtrsim 95\%$ pure $Z + \text{jet}$ sample. In both samples, we observe reasonably good agreement between data and simulation.

To account for any remaining differences, we define a scale factor:

$$\text{SF}(x) = \frac{\epsilon_{\text{Data}}(\text{BDT} > x \text{ and } 110 < m_{\text{SD}} < 210)}{\epsilon_{\text{MC}}(\text{BDT} > x \text{ and } 110 < m_{\text{SD}} < 210)} \quad (3.18)$$

where x is a particular decision boundary and ϵ is the fraction of data or MC events passing this BDT and mass selection. These are chosen to optimize sensitivity to the mono-top analysis, as described in Chapter 4. The SF is strongly dependent on the type of jet; in particular, we expect different SFs for top and LQG jets. In what follows, we will define two decision boundaries: loose ($x > 0.1$) and tight ($x > 0.45$) categories.

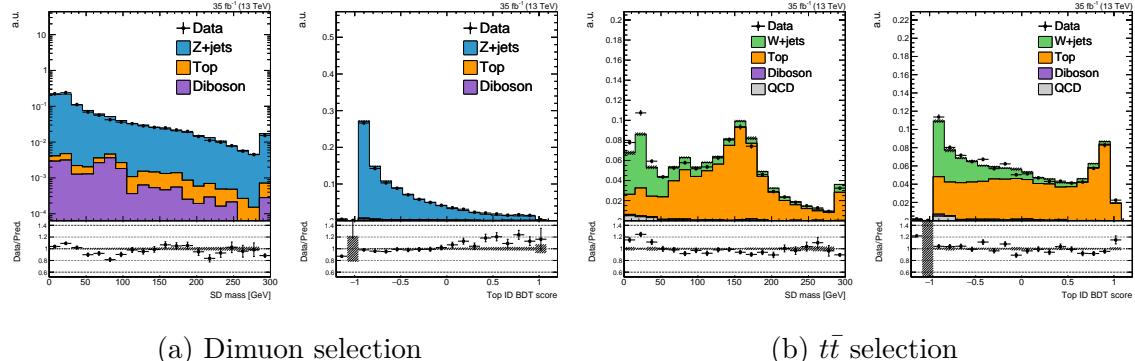


Figure 3.16: Comparison of the BDT response and jet mass in data and simulation, in top and LQG jets.

To compute SF_{LQG} , we use the dimuon selection in Figure ??, as this contains an essentially pure selection of LQG jets. Two sources of uncertainty are considered: the statistical uncertainties present in the data and MC, and the uncertainties on the theoretical prediction of the cross section of the small non-LQG backgrounds ($t\bar{t}$ and diboson events). The measured SFs are:

$$\begin{aligned}
 SF_{LQG}(0.1) &= 1.02 \pm 0.05(\text{total}) \\
 &\quad \pm 0.04(\text{statistical}) \pm 0.03(t\bar{t} + \text{diboson}) \\
 SF_{LQG}(0.45) &= 0.97 \pm 0.07(\text{total}) \\
 &\quad \pm 0.06(\text{statistical}) \pm 0.03(t\bar{t} + \text{diboson})
 \end{aligned} \tag{3.19}$$

The process for top jets is complicated by the fact that the top pair selection in Figure 3.16 is not sufficiently pure in merged top jets. There is significant contamination from $W+jets$ events. Furthermore, we cannot ensure that every $t\bar{t}$ event selected produces a *merged* top jet - some events may contain jets in which only part of the top's decay products are clustered into the CA15 jet. Therefore, we extract the efficiency by means of a template fit to the mass distribution of passing and failing events, which can separate the top and LQG components in the selection. It is for this reason that we only use groomed observables in the BDT: grooming prevents a strong correlation between the observables and m_{SD} . Such a correlation would cause

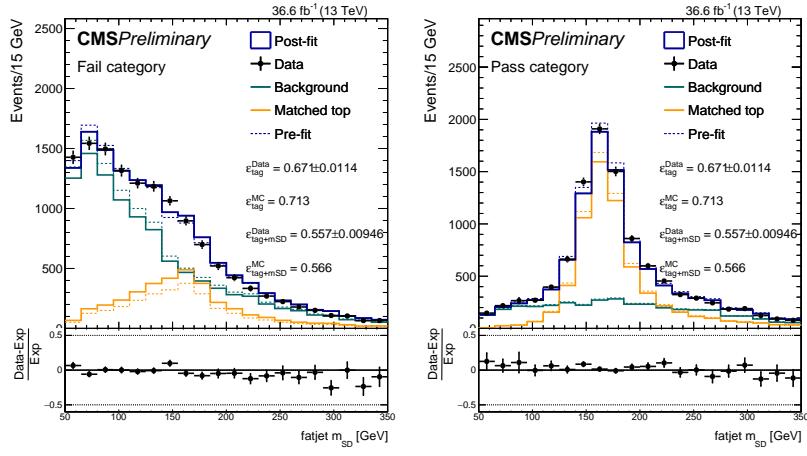
the mass distribution of passing LQG jets to be indistinguishable from that of passing top jets. Figure 3.17 show the fits in the passing and failing regions for both decision boundaries.

Several sources of uncertainty are considered for this measurement:

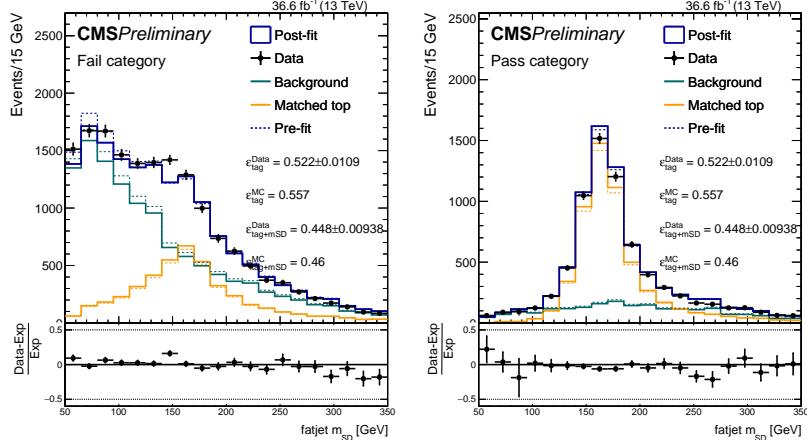
- Poisson uncertainties in the data and simulation
- CA15 jet energy scale
- Definition used to select *merged top* jets, allowing $\max \Delta R_{qq'}$ to vary between 1 and 1.5 (nominal value is 1.2)
- Efficiency of selecting b jets

The resultant SFs and associated uncertainties are:

$$\begin{aligned} \text{SF}_{\text{top}}(0.1) &= 1.08 \pm 0.04(\text{total}) \\ &\quad \pm 0.03(\text{statistical}) \pm 0.02(\text{JES + JER}) \pm 0.02(\text{merging}) \pm 0.002(b) \\ \text{SF}_{\text{top}}(0.45) &= 1.07 \pm 0.06(\text{total}) \\ &\quad \pm 0.03(\text{statistical}) \pm 0.02(\text{JES + JER}) \pm 0.014(\text{merging}) \pm 0.000(b) \end{aligned} \tag{3.20}$$



(a) Loose BDT-tagged



(b) Tight BDT-tagged

Figure 3.17: Fits to the m_{SD} distribution in a $t\bar{t}$ sample to extract the efficiency in data of the BDT and mass selections. All uncertainties plotted and quoted are statistical in nature.

Chapter 4

The Search for Missing Momentum and a Top Quark

In this chapter, we discuss the search for dark matter produced in association with a single top quark (*mono-top*). Since the initial state of pp collisions do not contain any appreciable contribution from top quarks, any process that produces a single top quark must involve some flavor violation. In the Standard Model, any such process is heavily suppressed by off-diagonal elements of the CKM matrix. The SM production mechanism for the mono-top signature (Figure 4.1) involves a b quark in the final state, and thus does not couple the third generation with the first or second. True production of mono-top must introduce some such coupling as an extension to the SM, in addition to one (or more) invisible particle to serve as a DM candidate.

We introduce two DM models that can produce this final state: a flavor-changing neutral current V and a charged, colored scalar ϕ . These models will also be used to benchmark the sensitivity of the analysis. However, it should be emphasized that the search is motivated and designed without reliance on any specific model; the assumption is that the mono-top final state alone is indicative of new physics, regardless of the specific production mechanism.

The flavor-changing neutral current (FCNC) V is assumed to couple to a fermionic

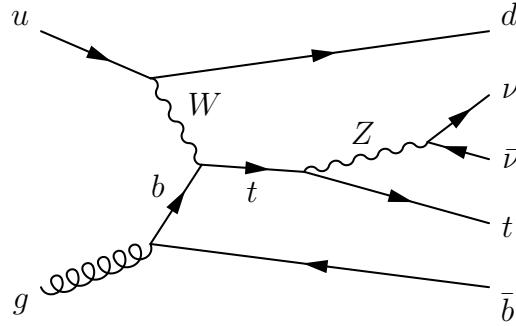


Figure 4.1: Production of mono-top in the SM, in which a top quark is produced in addition to a Z boson and bottom quark. The Z decays to neutrinos, providing large p_T^{miss} .

DM candidate χ . A partial Lagrangian of the interaction terms is given by:

$$\mathcal{L}_{\text{int}} = V_\mu \bar{\chi} \gamma^\mu (g_\chi^V + g_\chi^A \gamma_5) \chi + \bar{q}_u \gamma^\mu (g_u^V + g_u^A \gamma_5) q_u V_\mu + \bar{q}_d \gamma^\mu (g_d^V + g_d^A \gamma_5) q_d V_\mu + \text{h.c.}, \quad (4.1)$$

The model comes with 22 free parameters, broadly organized in three sets:

- The masses m_V and m_χ . (2)
- The couplings g_χ^V and g_χ^A . These, respectively, control the strength of the vector and axial interactions between V and χ . (2)
- The four coupling matrices g_q^X , where $q = u, d$ and $X = V, A$. As before, X determines the type of spin-1 interaction. In principle, different coupling strengths can be permitted for up- and down-type quarks, so this indexed by q . Each g_q^X is a 3×3 matrix, cross-coupling the three quark generations. To preserve $SU(2)_L$ symmetry, we require $g_u^V - g_u^A = g_d^V - g_d^A$. ($3 \times 6 = 18$)

It is the $g_{u,d}^{V,A}$ matrices that determine whether the model can produce mono-top, or mono-bottom, or mono-up, etc. If $g_{u,d}$ is strongly diagonal, in the sense that the cross-generation couplings are weak, then mono-light quark production will dominate, resulting in the mono-jet final state (Figure 4.2a). On the other hand, if we assume the only non-zero elements are those that couple the first and third generations, then mono-top production at the LHC is the best way to probe this model (Figure 4.2b). It

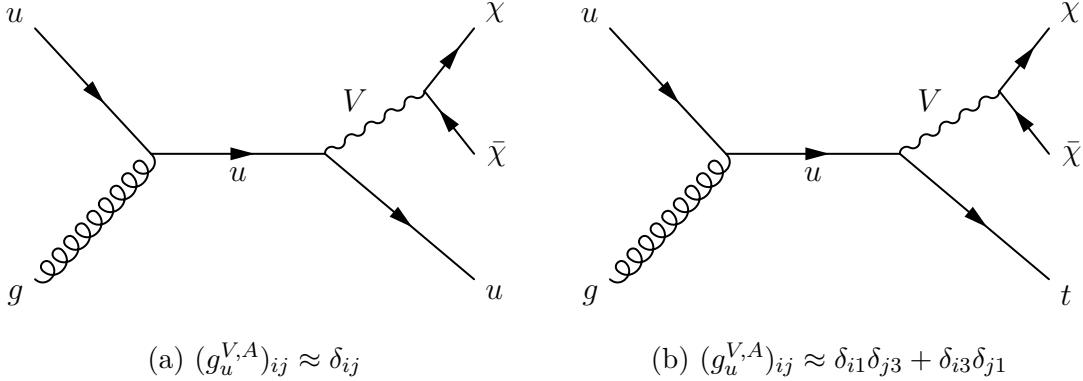


Figure 4.2: Possible DM production at the LHC, assuming a simplified spin-1 extension to the SM. The shown diagrams are at LO, although the MC simulation of these processes is at NLO in QCD vertices.

is this latter choice that will be made in the rest of this chapter; other choices are best probed using a combination of multiple DM channels, which is left as future work. Furthermore, to respect $SU(2)_L$ symmetry, we make the assumption that $g_u^V = g_d^V$ and $g_u^A = g_d^A$.

In the second benchmark model, the charged, colored scalar ϕ couples to down-type quarks, or to a fermionic DM candidate ψ and a top quark. The interaction terms of the Lagrangian is given by:

$$\mathcal{L}_{\text{int}} = \phi \bar{d}_i^C [(a_q)^{ij} + (b_q)^{ij} \gamma^5] d_j + \phi \bar{t} [a_\psi + b_\psi \gamma^5] \psi + \text{h.c.} \quad (4.2)$$

There are 16 free parameters in this model, broadly organized in three categories:

- The masses m_ϕ and m_ψ . (2)
- The couplings at the $\phi \bar{t} \psi$ vertex a_ψ and b_ψ , which respectively control the strength of the scalar and pseudoscalar interactions. (2)
- The couplings at the $\phi \bar{d}_i d_j$ vertex a_q^{ij} and b_q^{ij} where $i, j = 1, 2, 3$. Again, a and b refer the scalar and pseudoscalar couplings, respectively. (12)

In this model, mono-top production primarily occurs through the resonant decay of ϕ to ψ and t , as shown in Figure 4.3.

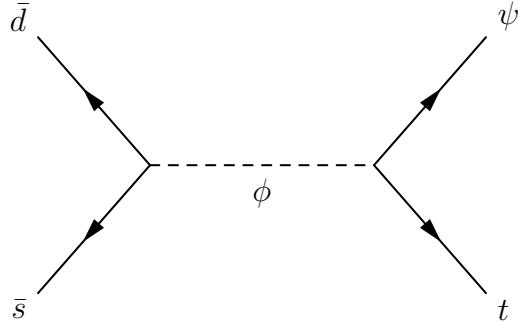


Figure 4.3: Possible DM production at the LHC, assuming the existence of a charged, color scalar that couples to DM and the top quark. The diagrams and the MC simulation of this process are at LO.

The two benchmark models show markedly different spectra in Figure 4.4, motivating their use to test different modes of mono-top production. The FCNC produces a falling p_T^{miss} distribution. The decay of a scalar resonance produces a distribution peaking at approximately $p_T^{\text{miss}} \approx m_\phi/2$. This is because p_T^{miss} is the transverse component of the momentum of one of the decay products of the resonance. In the region of interest, m_ϕ is quite large and the scalar is produced near rest, and so $|\vec{p}_\psi| \sim m_\phi/2$.

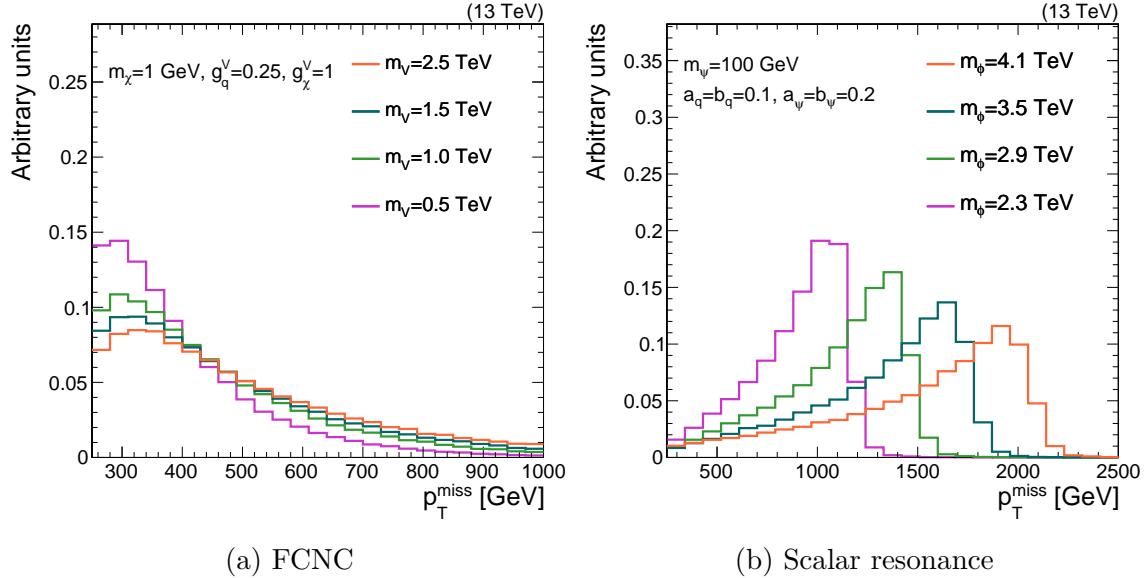


Figure 4.4: Spectra of DM (missing) momentum under various signal hypothesis. In the FCNC case, the spectra become harder as m_V increases, as the momentum transfer needed to produce a particle of mass m_V increases. In the scalar resonance case, the p_T^{miss} distribution is a Jacobian shape, with a peak near half the resonance mass.

4.1 Signal selection

In this section, we will describe how events with large missing momentum and a top jet candidate are selected, and events with other signatures, such as leptons and photons, are rejected. The selection begins with a relatively loose trigger definition, to record collision events, followed by a more stringent set of *offline* criteria and categorization.

When looking at events that pass a simple set of criteria (moderate p_T^{miss} and one CA15 jet), it is clear (Figure 4.5) that the highest signal sensitivity is found in regions of high p_T^{miss} and jet p_T . After this loose selection, the three primary backgrounds, from most to least important, are:

- $Z \rightarrow \nu\nu$. When the Z is produced in association with one or more jets, the jet system will, in rare cases, pass the criteria used to select a top jet. The neutrinos manifest as p_T^{miss} .
- $W \rightarrow \ell\nu$. As in the case of the Z , additional jets mimic the signature of a top jet. Typically, the charged lepton in the final state is vetoed, but if it is out of acceptance (e, μ) or fails ID criteria (τ_h), then it is not identified.
- $t\bar{t} \rightarrow bqq' + \bar{b}\ell\nu$. As in the case of the W , a charged lepton in the final state may not be properly identified. Unlike the previous two processes, a semi-leptonic $t\bar{t}$ event contains a real hadronic top quark decay.

4.1.1 Online trigger selection

Data events are first selected by the L1 trigger system by requiring $p_{T,\text{L1}}^{\text{miss}} > 70 \text{ GeV}$, where:

$$p_{T,\text{L1}}^{\text{miss}} = - \left(\sum_{i \in C} \vec{p}_i \right)_T, \quad C = \{\text{calorimeter deposits with } |\eta| < 3.0\} \quad (4.3)$$

Events that pass this selection are sent to the HLT system, where we place requirements on the both the missing momentum ($p_{T,\text{HLT}}^{\text{miss}}$) and the missing hadronic momen-

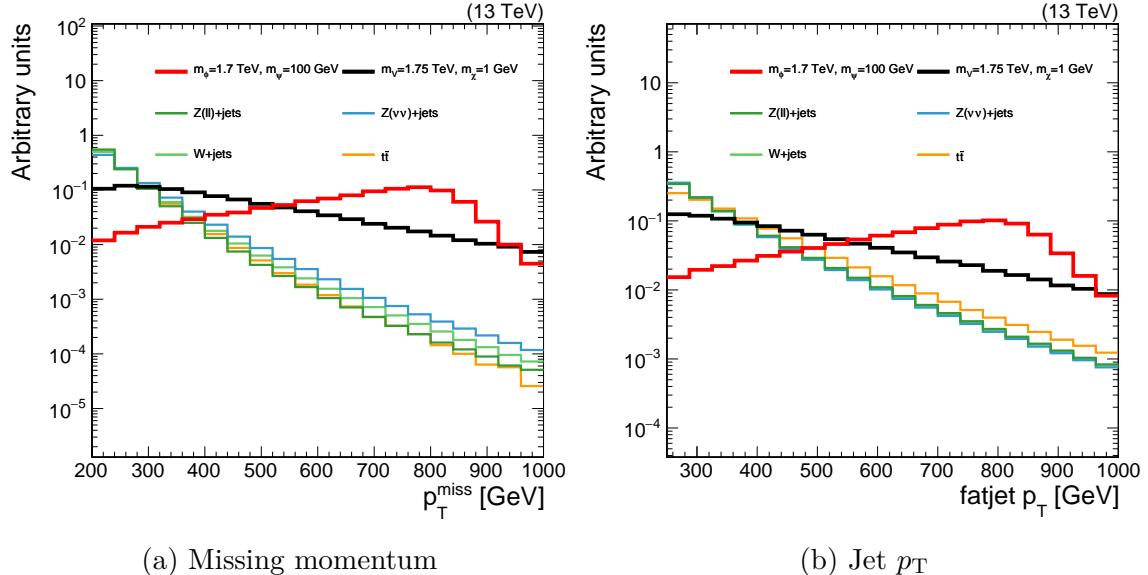


Figure 4.5: Comparison of missing and jet momenta in various backgrounds and signal models.

tum ($H_{T,\text{HLT}}^{\text{miss}}$). These are defined as:

$$p_{T,\text{HLT}}^{\text{miss}} = - \left(\sum_{i \in \text{particles}} \vec{p}_i \right)_T, \text{ all particles except muons} \quad (4.4)$$

$$H_{T,\text{HLT}}^{\text{miss}} = - \left(\sum_{i \in \text{jets}} \vec{p}_i \right)_T, \text{ jets passing noise-rejection ID} \quad (4.5)$$

The HLT decides to keep an event if $\min(p_{T,\text{HLT}}^{\text{miss}}, H_{T,\text{HLT}}^{\text{miss}})$ is higher than a specified threshold. Over the course of the data-taking period considered in this chapter, this threshold varied from 90 to 120 GeV. The use of noise-rejection in the definition of $H_{T,\text{HLT}}^{\text{miss}}$ makes it more resilient to detector noise than $p_{T,\text{HLT}}^{\text{miss}}$. Therefore, we place symmetric selections on both quantities.

Note that in all trigger decisions, muons are excluded from the missing momentum calculations. This means that an event which produces high-momentum muons can be selected using these triggers. This flexibility will be exploited in Section 4.2. Missing momentum with and without muons are treated on equal footing by defining

the *hadronic recoil* U :

$$\vec{U} = \vec{p}_T^{\text{miss}} + \sum_{\ell \in \text{lep.}} \vec{p}_T^\ell + \sum_{\gamma \in \text{photons}} \vec{p}_T^\gamma \quad (4.6)$$

where *lep* refers to all identified electrons and muons. In an event without electrons, muons and photons, $U = p_T^{\text{miss}}$.

Since the online environment and reconstruction are significantly limited as compared to the offline reconstruction of U , we do not expect the trigger decision to be a step function at $U = 120$ GeV. Therefore, we define and measure a trigger efficiency:

$$\epsilon_{\text{trig}}(U) = \frac{N_{\text{pass trig}}(U)}{N(U)} \quad (4.7)$$

This is measured using $W \rightarrow \mu\nu$ events containing one or more high- p_T jets. The events are triggered using single- μ triggers, which have lower thresholds and efficiencies close to 1 in this phase space. We then require events have exactly one well-identified muon and at least one jet with $p_T > 100$ GeV. Figure 4.6 shows the efficiency as a function of U . A threshold of $U > 250$ GeV is chosen to optimize the tradeoff between the number of selected signal events and the increased systematic uncertainties incurred from the steeply-rising part of the efficiency curve.

4.1.2 Offline signal selection

The signal regions (SRs) are defined by a further set of offline selection criteria, which are listed and motivated in Table 4.1. The selections are optimized on the basis of sensitivity to a benchmark signal hypothesis, an FCNC with $m_V = 1.7$ TeV. This optimization is approximately ideal for all relevant signal models, as the main differences are in the tail of the p_T^{miss} distribution, which is not affected by the criteria in Table 4.1. As described in Section 3.1, two working points (WPs) are defined for the top ID BDT. The signal events passing all other selection criteria are partitioned into a *loose* SR and a *tight* SR on the basis of which WP the top candidate jet satisfies.

Figure 4.7 shows the p_T^{miss} distributions, as predicted by MC and as observed in

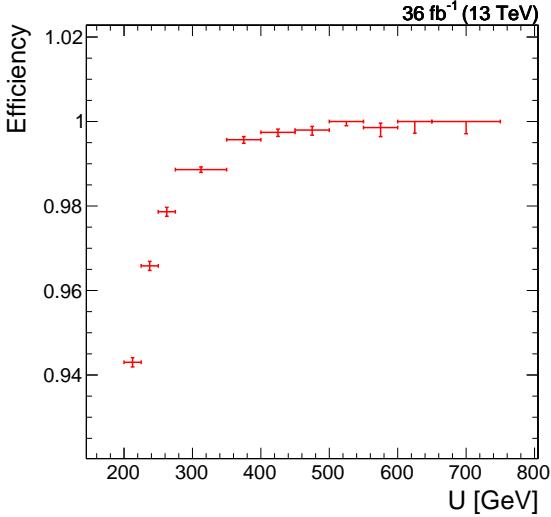


Figure 4.6: Efficiency of the p_T^{miss} trigger measured in single-muon events.

Table 4.1: Criteria used to select events for the mono-top search signal regions. Note that two SRs are defined, based on the BDT score.

Criterion	Notes
$p_T^{\text{miss}} > 250 \text{ GeV}$	Signal events should have large missing momentum. Exact threshold is chosen to maximize online trigger efficiency.
1 CA15 jet with $p_T > 250 \text{ GeV}$	Top quark candidate. Recoils against p_T^{miss} , so threshold is set at 250 GeV.
CA15 jet $110 < m_{\text{SD}} < 210 \text{ GeV}$	Consistency with top quark mass.
At least one b -tagged sub-jet	Identifying b hadron produced from top decay/hadronization.
No b -tagged narrow jets	Rejecting semi-leptonic $t\bar{t}$ decays.
No identified e, μ, τ_h	Suppress $W+\text{jet}$ and $t\bar{t}$ processes.
No identified γ	Suppress $\gamma+\text{jet}$ processes.
$\min_{\text{jets}} \Delta\phi(\text{jet}, p_T^{\text{miss}}) > 0.5$	Remove events with large p_T^{miss} caused by mismeasured jets.
CA15 jet BDT	Identifying top decay structure. If the jet passes the tight WP, it is placed in the <i>tight</i> SR. Otherwise, if it only passes the loose WP, it is placed in the <i>loose</i> SR.

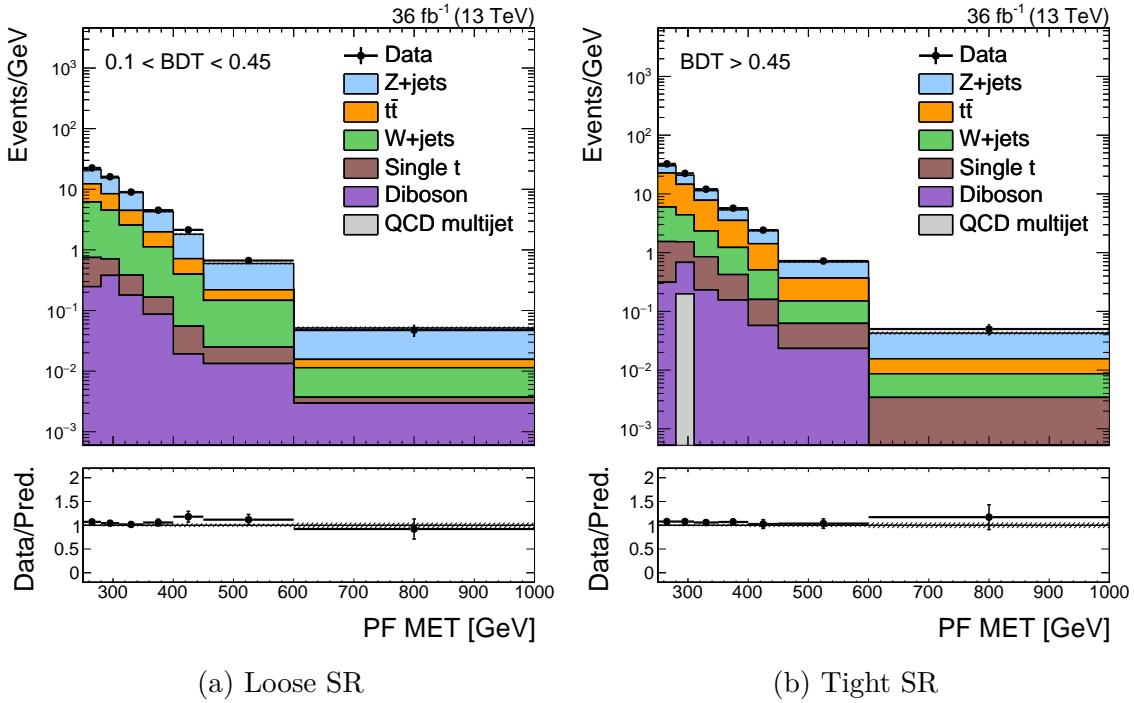


Figure 4.7: p_T^{miss} distributions in the two mono-top signal regions. The bottom section of each figure shows the ratio of the data and the prediction. The only uncertainties plotted in these figures are those arising from Poisson fluctuations in data (black bars) and MC (grey band). Enough MC events are produced such that the MC statistical uncertainties are smaller than those in data.

collected data, in the two signal regions.

4.2 Background estimation

Searching for DM amounts to looking for an excess of data events over the SM prediction at large values of p_T^{miss} . Therefore, the p_T^{miss} distribution of the three primary SM backgrounds described in Section 4.1 must be predicted with small uncertainty. The MC simulation provides a reasonable description of the data, but the theoretical uncertainties inherent in the MC, primarily due to higher-order QCD effects, can range up to 20%. To reduce the prediction uncertainty, data samples that cannot contain a DM signal (*control* data) is used to directly estimate or supplement the estimation of SM processes in the SR. This approach is known as *data-driven* estimation.

The basic strategy is as follows: each invisible background will be constrained by one or more visible processes for which we can measure the U spectrum. The visible processes are chosen such that they are sufficiently similar to the targeted background, minimizing the uncertainties in the extrapolation. The $t\bar{t}$ background is estimated using $t\bar{t}$ events in which the charged lepton is properly reconstructed. The Z and W backgrounds are simultaneously constrained using a combination of $Z \rightarrow \ell\ell$, $W \rightarrow \ell\nu$, and γ events. Extrapolations which link very similar processes, such as $Z \rightarrow \mu\mu$ and $Z \rightarrow \nu\nu$, can be made with relatively small experimental uncertainties. The linkage of dissimilar processes, such as γ and $Z \rightarrow \nu\nu$, additionally brings relatively large theoretical uncertainties, which must be carefully estimated.

4.2.1 Visible final states to constrain invisible final states

As a starting point, let us tackle the estimation of $Z \rightarrow \nu\nu$ in the SR. Since the momentum imbalance (up to experimental effects) in a $Z \rightarrow \nu\nu$ event is just the transverse momentum of the Z boson (p_T^Z), we must estimate p_T^Z . To good approximation, the p_T^Z distribution is independent of the decay mode of the Z boson. Therefore, it is natural to estimate $p_T^{\text{miss}}(Z \rightarrow \nu\nu)$ by measuring $p_T^Z(Z \rightarrow \mu\mu)$, as muons are easily identifiable and reconstructible.

However, there is one important distinction between $\nu\nu$ and $\mu\mu$ events. In the latter, p_T^Z can be directly measured, whereas in the former it must be inferred through a

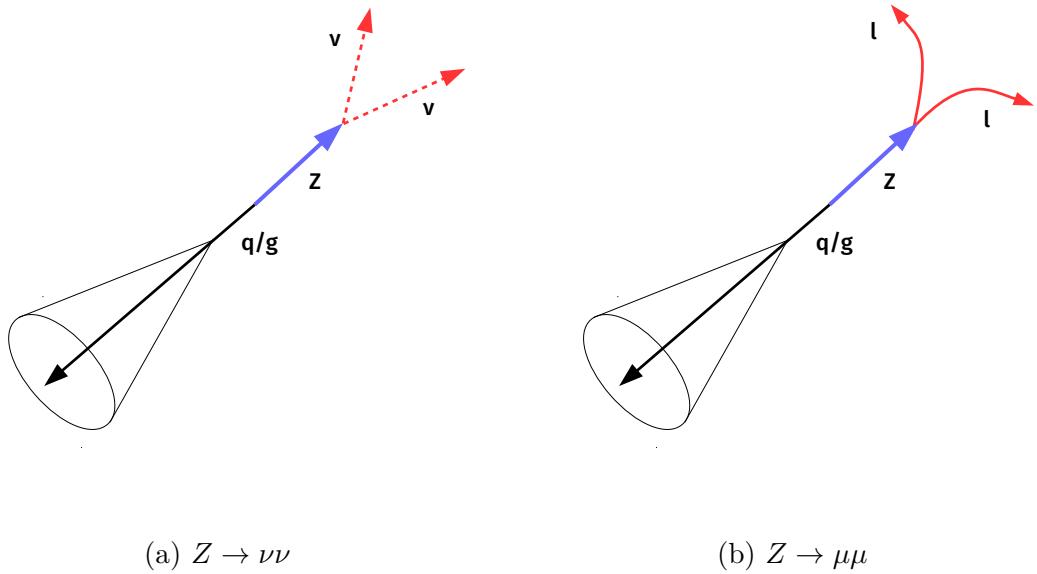


Figure 4.8: Schematic representation of two Z decay modes: to neutrinos (as in the SR) and to muons (as in the CRs). Note that in both cases, U is sensitive to the same effects arising from the measurement of the jet recoiling against the Z boson, whereas $p_T^{\mu\mu}$ is largely independent of the jet.

momentum imbalance. Effects like jet energy scale impact p_T^{miss} , but not $p_T^{\mu\mu}$. Therefore, instead of directly measuring $p_T^{\mu\mu}$ in $\mu\mu$ events, we use the hadronic recoil U . In SR events, in which there are no e, μ, γ , $U = p_T^{\text{miss}}$. In $Z \rightarrow \mu\mu$ events, U mimics the momentum imbalance, if we had pretended the identified muons did not exist when computing p_T^{miss} . Therefore, U is an exact analogy for p_T^{miss} in the SR. Figure 4.8 makes the same argument in a schematic fashion.

Table 4.2 describes the criteria used to define events in the $\mu\mu$ control regions (CRs). Figure 4.9 shows the distribution of U in these CRs, as well as the $m_{\mu\mu}$ and p_T^μ distributions.

The control data is used to constrain the SR prediction by means of *transfer factors* $T_{Y,i}^X$, where X refers to a particular CR (e.g. $\mu\mu$), Y refers to a particular process (e.g. Z), and i refers to a particular bin in the CR (e.g. $200 < U < 250$ GeV in the tight category). Formally:

$$T_{Z,i}^{\mu\mu} = \frac{N_i^{\text{SR}}(Z \rightarrow \nu\nu)}{N_i^{\mu\mu}(Z \rightarrow \mu\mu)} \quad (4.8)$$

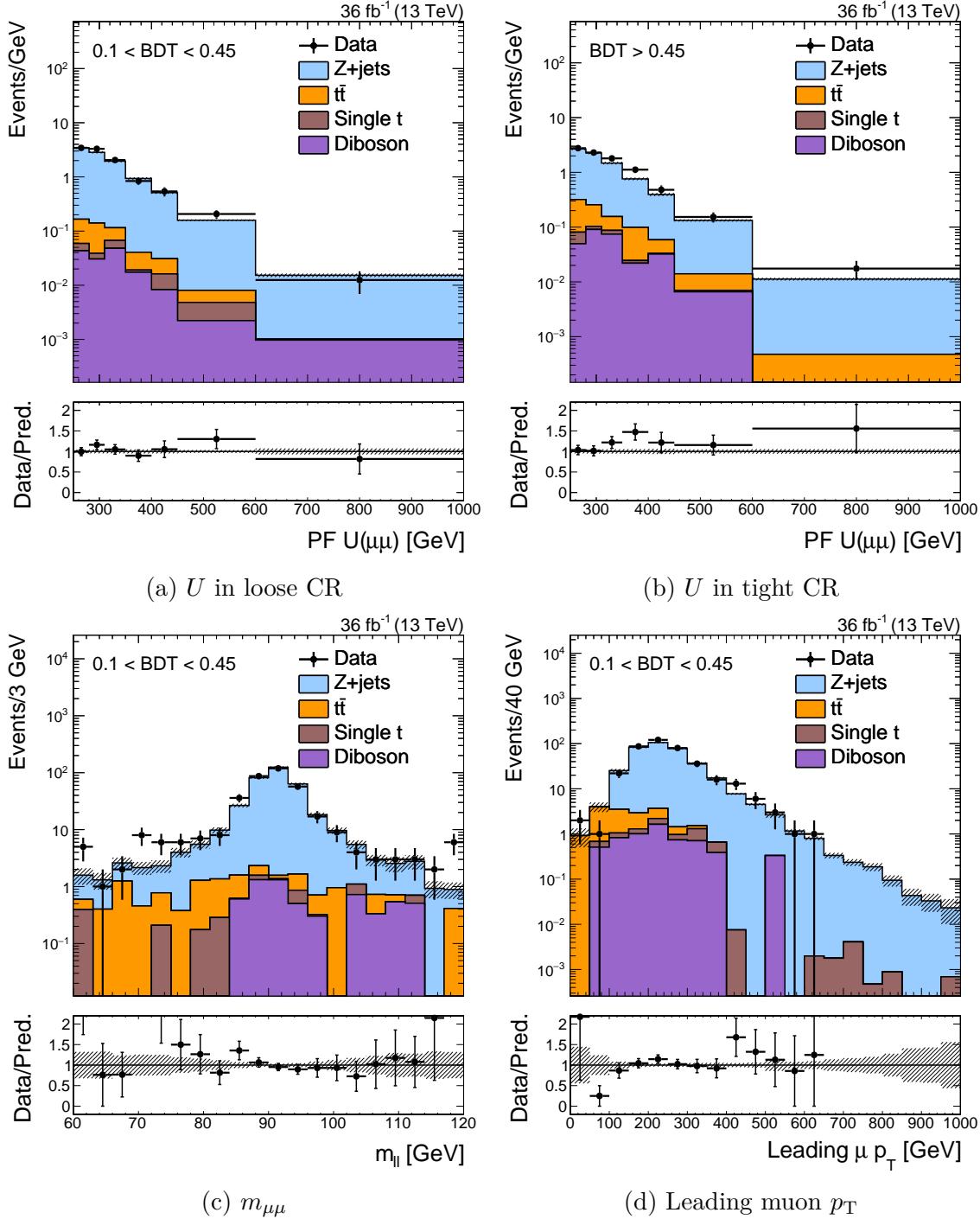


Figure 4.9: Various kinematic distributions in the two mono-top $\mu\mu$ CRs. All predicted distributions are prior to the maximization of the likelihood, and the grey band refers only to the statistical uncertainty of the MC. The pre-fit MC describes the observed data reasonably well.

Table 4.2: Criteria used to select events for the mono-top $Z \rightarrow \mu\mu$ CR. As in the SR, the region is further subdivided based on the jet BDT score.

Criterion	Notes
$U > 250$ GeV	Same as the SR selection, so as to select Z s in the same phase space.
1 CA15 jet with $p_T > 250$ GeV	Same as SR
CA15 jet $110 < m_{SD} < 210$ GeV	Same as SR
Well-identified μ^-, μ^+ pair, with $ m_{\mu\mu} - m_Z < 30$ GeV	Identifying the $Z \rightarrow \mu\mu$ resonance.
No identified e, τ_h	Same as SR.
No identified γ	Same as SR
$\min_{\text{jets}} \Delta\phi(\text{jet}, U) > 0.5$	Same as SR
CA15 jet BDT	Same as SR

The transfer factors are estimated using MC simulation. To encode the effects of various uncertainties, we introduce nuisance parameters $\boldsymbol{\theta}$. That is:

$$T_{Y,i}^X \rightarrow T_{Y,i}^X(\boldsymbol{\theta}) \equiv T_{Y,i}^X \times \prod_{j=0}^{n_\theta} (1 + \theta_j) \quad (4.9)$$

$$\theta_j \sim p_j(\theta_j) \quad (4.10)$$

where n_θ is the number of nuisance parameters and $p_j(\theta_j)$ is some prior distribution for each nuisance. The priors are typically chosen to have a central value at 0, with a finite variance that encodes the uncertainty. In this chapter, we assume p_j is either a normal distribution centered at 0 or a log-normal distribution, in cases where negative values are undesirable. We will use the terms *uncertainty* and *nuisance parameter* interchangeably.

Let $\text{Pois}(d|\lambda)$ refer to the Poisson probability of observing d with an expected mean of λ . In terms of these transfer factors, the likelihood for the data observed in the signal and $\mu\mu$ control regions is:

$$\begin{aligned} \mathcal{L}(\mathbf{d} \mid \mu, \boldsymbol{\mu}_{\text{SR}}^{Z \rightarrow \nu\nu}, \boldsymbol{\theta}) = \prod_{i \in \text{bins}} & \left[\text{Pois}\left(d_i^{\text{SR}} \mid \mu S_i^{\text{SR}}(\boldsymbol{\theta}) + \mu_{\text{SR},i}^{Z \rightarrow \nu\nu} + B_i^{\text{SR}}(\boldsymbol{\theta})\right) \right. \\ & \times \text{Pois}\left(d_i^{\mu\mu} \mid \frac{\mu_{\text{SR},i}^{Z \rightarrow \nu\nu}}{T_{Z,i}^{\mu\mu}(\boldsymbol{\theta})} + B_i^{\mu\mu}(\boldsymbol{\theta})\right) \left. \right] \times \prod_{j=0}^{n_\theta} p_j(\theta_j) \end{aligned} \quad (4.11)$$

where the following notation is used:

- d_i^X : The data observed in bin i of region X . For now, $X = \text{SR}, \mu\mu$.
- S_i^{SR} : The predicted number of signal events in bin i of the SR, under some fixed signal hypothesis.
- μ : The *signal strength*. Essentially an unconstrained nuisance parameter that scales up or down the total signal yield.
- $\mu_{\text{SR},i}^P$: The expected number of events from process P in bin i of the SR. This is also an unconstrained nuisance parameter.
- B_i^X : The predicted number of *minor* background events in bin i of region X . For now, minor refers to all SM processes that are not estimated using a data-driven method. Although it has not yet been described, all major SM backgrounds have data-driven estimations. The remainder are sufficiently small that a precise prediction is not critical.

The signal and background yields \mathbf{S} and \mathbf{B} are estimated using MC. The purpose of the priors is to encode our hypothesis for the distribution of the nuisance parameters. A maximum likelihood estimate will fix $\boldsymbol{\theta}$ to balance between maximizing the Poisson probabilities and maximizing $\prod_j p_j(\theta_j)$.

Ignoring small minor backgrounds, the null hypothesis sets $B_i = \mu = 0$. In this scenario, there is an intuitive picture of the likelihood maximization. The parameters $\mu_{\text{SR}}^{Z \rightarrow \nu\nu}$ float freely to satisfy $d_{\text{SR},i} \sim \mu_{\text{SR},i}^{Z \rightarrow \nu\nu}$ and $d_{\mu\mu,i} \sim \mu_{\text{SR},i}^{Z \rightarrow \nu\nu}/T_{Z,i}^{\mu\mu}(\boldsymbol{\theta})$. If both constraints cannot be satisfied simultaneously by scaling $\mu_{\text{SR}}^{Z \rightarrow \nu\nu}$, the (constrained) nuisance parameters $\boldsymbol{\theta}$ modify the transfer factor $T_Z^{\mu\mu}$. Table 4.6 shows the relevant uncertainties for $T_Z^{\mu\mu}$, and Figure 4.10 shows the shape of uncertainties that evolve as a function of U .

The transfer factors are shown in Figure 4.11. The exact values of $T_Z^{\mu\mu}(\boldsymbol{\theta})$ have two salient features:

1. The values are strictly greater than one. This is due to (a) $\mathcal{B}(Z \rightarrow \nu\nu) > \mathcal{B}(Z \rightarrow \mu\mu)$ and (b) a non-100% efficiency in reconstructing and identifying

Table 4.3: Uncertainties affecting the $\mu\mu \leftrightarrow \nu\nu$ extrapolation. *Shape* uncertainties have different priors for each bin, but are assumed to be correlated across bins.

Uncertainty	1 s.d.	Notes
μ ID	2%	
μ track	1%	
τ_h veto	3%	
$Z +$ heavy flavor	3%	
Trigger	0-2%	Shape
b -tag	$\sim 0.5\%$	Shape
$udcsg$ -mistag	3-4%	Shape

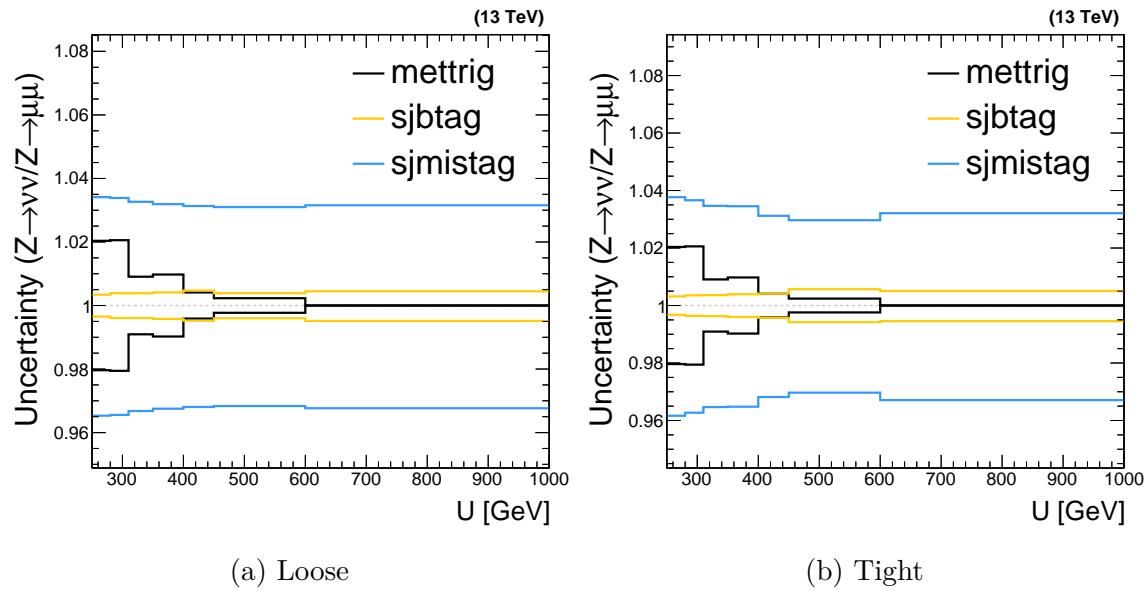


Figure 4.10: Shape uncertainties affecting $T_i^{\mu\mu}$ in both categories, as a function of U .

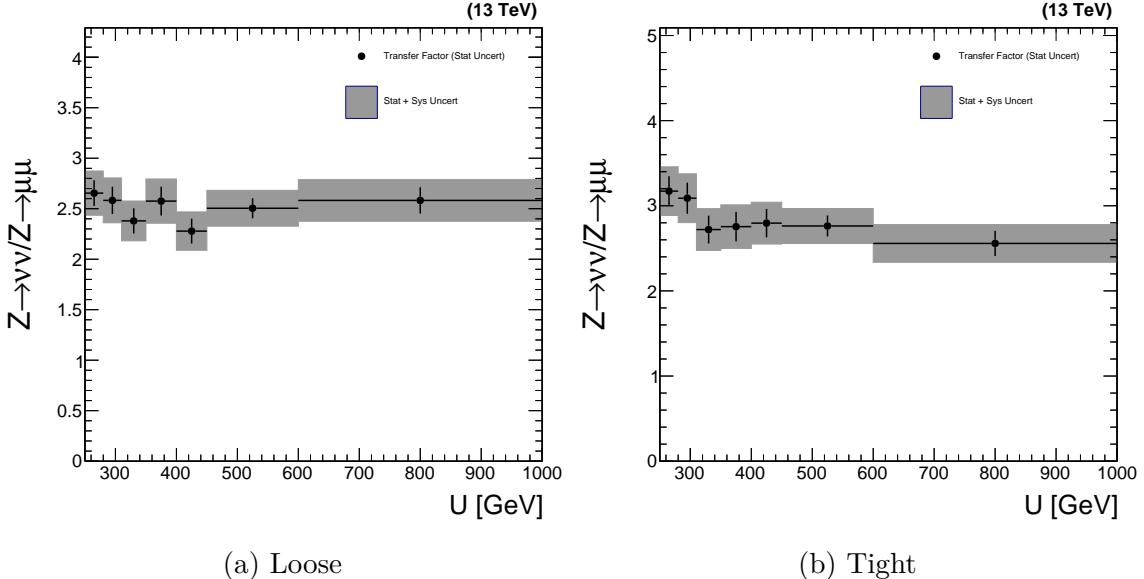


Figure 4.11: The transfer factors $T_{Z,i}^{\mu\mu}$ as a function of recoil and BDT score. The vertical black bars represent the Poisson uncertainties in the MC simulation, while the grey bands represent the sum of Poisson uncertainties and other, systematic, uncertainties. All uncertainties are represented as one standard deviation.

muons. This implies that the statistical uncertainties of the $\mu\mu$ CR are greater than of the process being estimated, especially at high U .

2. The one standard deviation variation of all uncertainties that impact $T_{Z,i}^{\mu\mu}$ are contained within a 10% envelope. This is already a factor of two smaller than the $\sim 20\%$ theoretical uncertainties in the MC simulation.

To account for point (1), we can simply add more control data by also looking at $Z \rightarrow ee$ decays. Essentially all of the arguments used for the $\mu\mu$ CRs applies to the ee CRs. Figures 4.12-4.13 show the data/simulation agreement and the transfer factors for the new dielectron regions. A further set of statistical constraints to improve the estimate at high U (which is where the signals are most enhanced) is described in Section 4.2.2.

Similar methods are used to predict the $W+jets$ and $t\bar{t}$ contributions in the SRs; these three backgrounds comprise at least 95% of the SM processes. In both cases, the momentum imbalance in the SR is a proxy for the momentum of the W boson, since

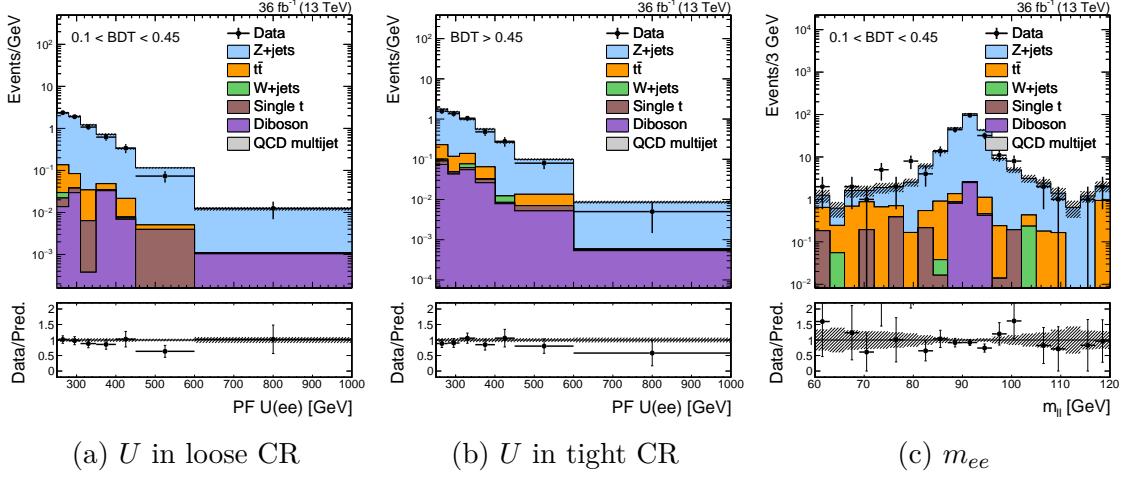


Figure 4.12: Various kinematic distributions in the two mono-top ee CRs. All predicted distributions are prior to the maximization of the likelihood, and the grey band refers only to the statistical uncertainty of the MC. The pre-fit MC describes the observed data reasonably well.

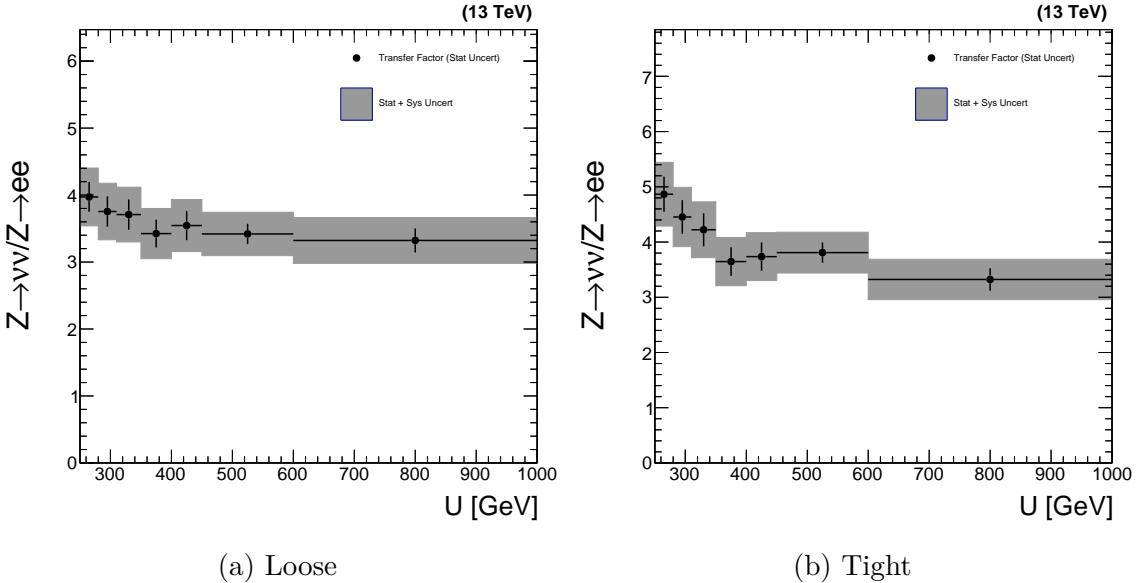


Figure 4.13: The transfer factors T_i^{ee} as a function of recoil and BDT score. The vertical black bars represent the Poisson uncertainties in the MC simulation, while the grey bands represent the sum of Poisson uncertainties and other, systematic, uncertainties. All uncertainties are represented as one standard deviation.

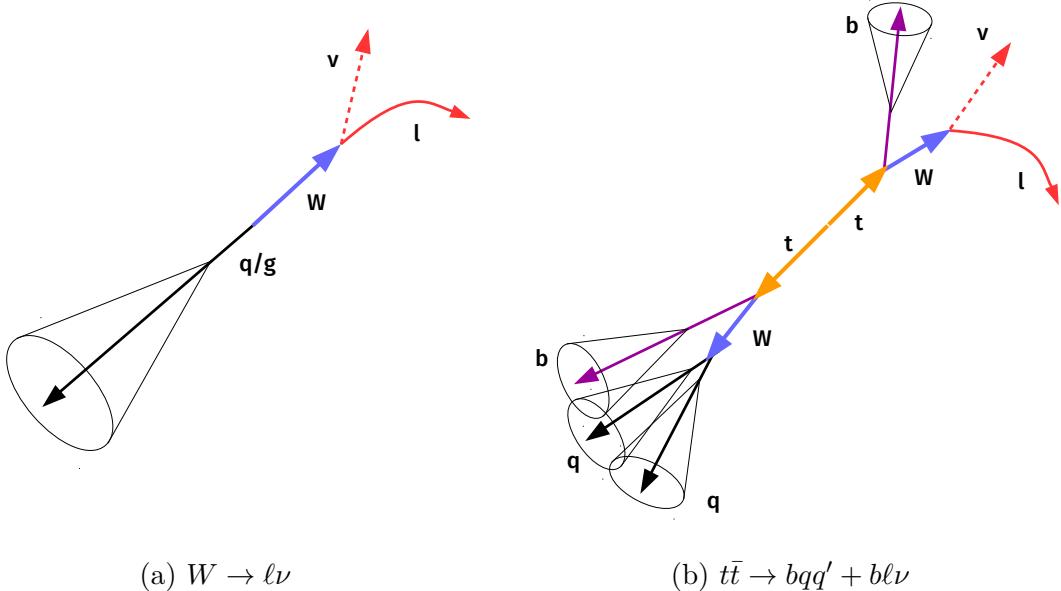


Figure 4.14: Schematic representation of the W and $t\bar{t}$ SM processes. In both cases, the measured U is an approximation of p_T^W . Furthermore, if the charged lepton is lost, $U = p_T^{\text{miss}} \approx p_T^W$.

the charged lepton is lost. A sketch of the event topologies is shown in Figure 4.14. Following the same arguments as used for the $Z \rightarrow ll$ CRs, we can use the hadronic recoil U in CRs that measure visible final states of W and $t\bar{t}$.

Starting with muon final states (electrons will follow naturally), we define two sets of CRs based on the number of identified B hadrons. The selection for the $b\mu$ CRs (to measure $t\bar{t}$) is shown in Table 4.4. The selection for the μ CRs (to measure W) is shown in Table 4.5. Figures 4.15-4.16 show various kinematic distributions in these regions.

Each CR gets a set of transfer factors to constrain the targeted process in the SR: $\mathbf{T}_{t\bar{t}}^{b\mu}$ and \mathbf{T}_W^μ . In the tight μ CR (Figures 4.16b-4.16c), the stringent top ID requirement enhances the $t\bar{t}$ and suppresses the W contribution. Since we cannot create a pure W in the tight category, we introduce an additional set of transfer factors $\mathbf{T}_{t\bar{t}}^\mu$. This extra constraint uses the $b\mu$ CRs to estimate the $t\bar{t}$ component in the μ CRs, thereby leaving only one large degree of freedom in the μ CRs. These three sets of transfer factors, and the corresponding uncertainties, are shown in Figure 4.17.

Table 4.4: Criteria used to select events for the mono-top $b\mu$ CR. As with the SR, the region is further subdivided based on the jet BDT score.

Criterion	Notes
$U > 250 \text{ GeV}$	Same as the SR selection, so as to select t quarks in the same phase space.
1 CA15 jet with $p_T > 250 \text{ GeV}$	Same as SR
CA15 jet $110 < m_{SD} < 210 \text{ GeV}$	Same as SR
At least one b -tagged sub-jet	Identifying B hadron produced from hadronic top decay.
Exactly one b -tagged narrow jet	Identifying B hadron produced from leptonic top decay.
Exactly one well-identified μ	Produced from $W \rightarrow \mu\nu$
No identified e, τ_h	Same as SR.
No identified γ	Same as SR
$\min_{\text{jets}} \Delta\phi(\text{jet}, U) > 0.5$	Same as SR
CA15 jet BDT	Same as SR

Table 4.5: Criteria used to select events for the mono-top μ CR. As with the SR, the region is further subdivided based on the jet BDT score.

Criterion	Notes
$U > 250 \text{ GeV}$	Same as the SR selection, so as to select Ws in the same phase space.
1 CA15 jet with $p_T > 250 \text{ GeV}$	Same as SR
CA15 jet $110 < m_{SD} < 210 \text{ GeV}$	Same as SR
No b -tagged sub-jets	Suppressing semi-leptonic $t\bar{t}$ decays
No b -tagged narrow jets	Suppressing semi-leptonic $t\bar{t}$ decays
Exactly one well-identified μ	Produced from $W \rightarrow \mu\nu$
No identified e, τ_h	Same as SR.
No identified γ	Same as SR
$\min_{\text{jets}} \Delta\phi(\text{jet}, U) > 0.5$	Same as SR
CA15 jet BDT	Same as SR

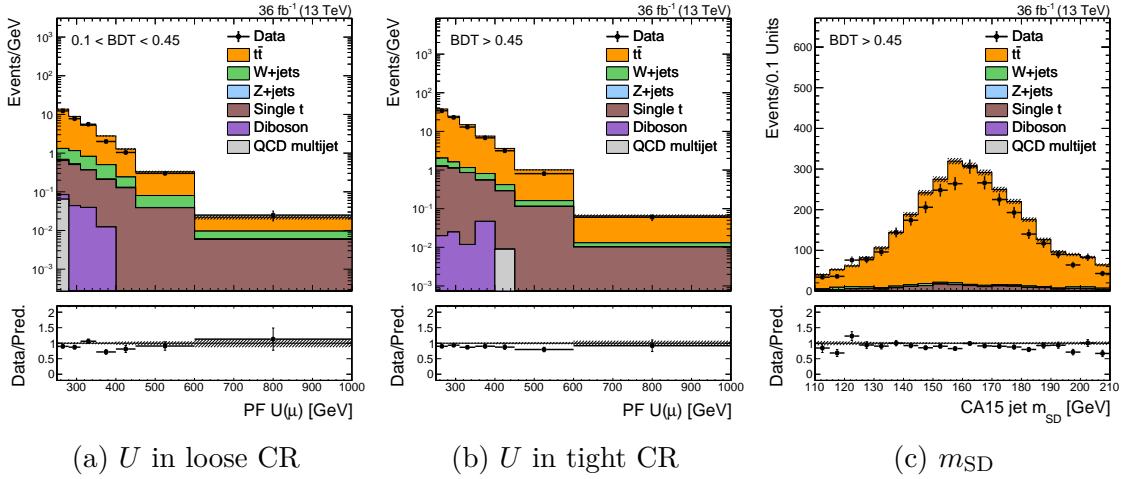


Figure 4.15: Various kinematic distributions in the two mono-top $b\mu$ CRs. All predicted distributions are prior to the maximization of the likelihood, and the grey band refers only to the statistical uncertainty of the MC. Although a flat offset is observed in all spectra, it is also observed in the corresponding electron final state (Figure 4.18). Such an effect cancels in the transfer factor and is corrected for by the fit (Figure 4.26).

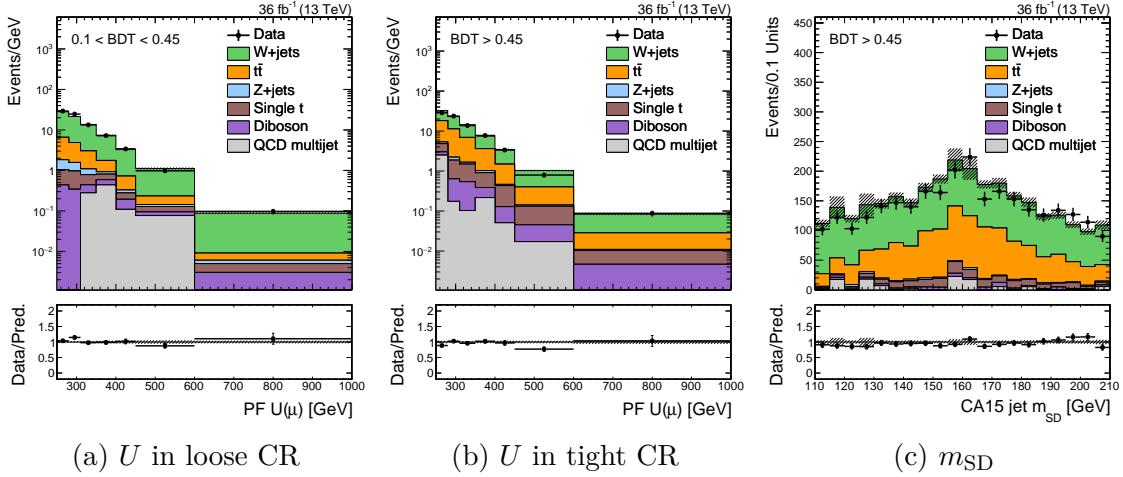


Figure 4.16: Various kinematic distributions in the two mono-top μ CRs. All predicted distributions are prior to the maximization of the likelihood, and the grey band refers only to the statistical uncertainty of the MC. The pre-fit MC describes the observed data reasonably well.

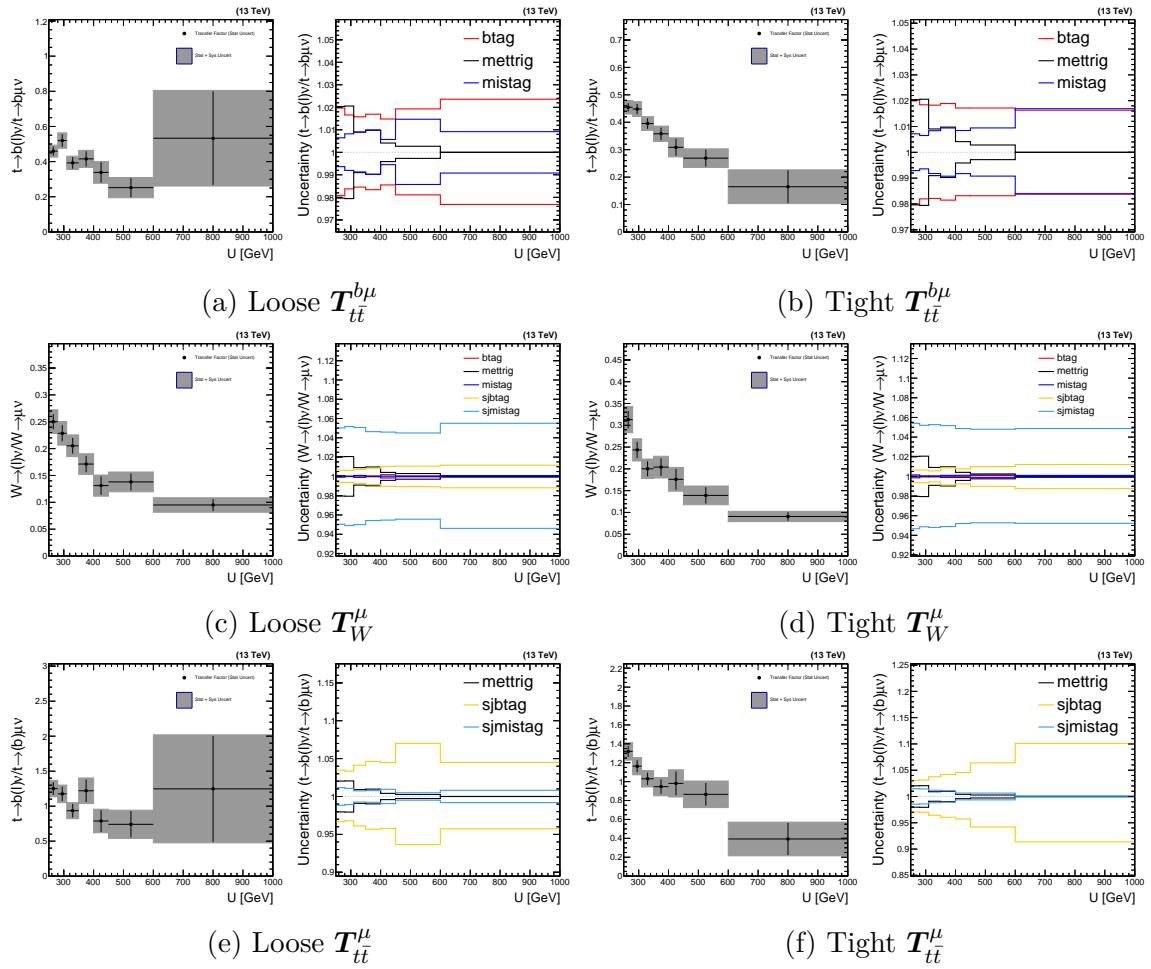


Figure 4.17: The transfer factors $\mathbf{T}_{t\bar{t}}^{b\mu}$, \mathbf{T}_W^μ , and $\mathbf{T}_{t\bar{t}}^\mu$; and corresponding shape uncertainties.

Table 4.6: Uncertainties affecting the various single-muon extrapolations. *Shape* uncertainties have different priors for each bin, but are assumed to be correlated across bins.

Uncertainty	1 s.d. ($\mathbf{T}_{t\bar{t}}^{b\mu}$)	1 s.d. (\mathbf{T}_W^μ)	1 s.d. ($\mathbf{T}_{t\bar{t}}^\mu$)	Notes
μ ID	1%	1%	1%	
μ track	0.5%	0.5%	0.5%	
τ_h veto	3%	3%	3%	
W +heavy flavor		3%		
Trigger	0-2%	0-2%	0-2%	Shape
b -tag	2%	$\sim 0.5\%$	3-6%	Shape
$udcsg$ -mistag	1%	5%	1%	Shape

As we added the ee CRs to complement the $\mu\mu$ CRs, we also add be (e) CRs to augment the statistical power of the $b\mu$ (μ) CRs, especially at high recoil. Figures 4.18 and 4.19 respectively show some kinematic distributions and the transfer factors corresponding to these electron constraints.

Having defined (almost all of) the CRs and transfer factors, we can write down a complete likelihood for the mono-top search:

$$\begin{aligned} \mathcal{L}(\mathbf{d} \mid \mu, \boldsymbol{\mu}_{\text{SR}}^{Z \rightarrow \nu\nu}, \boldsymbol{\mu}_{\text{SR}}^{W \rightarrow \ell\nu}, \boldsymbol{\mu}_{\text{SR}}^{t\bar{t}}, \boldsymbol{\theta}) = \\ \prod_{i \in \text{bins}} \left[\text{Pois}\left(d_i^{\text{SR}} \mid \mu S_i^{\text{SR}}(\boldsymbol{\theta}) + \mu_{\text{SR},i}^{Z \rightarrow \nu\nu} + \mu_{\text{SR},i}^{W \rightarrow \ell\nu} + \mu_{\text{SR},i}^{t\bar{t}} + B_i^{\text{SR}}(\boldsymbol{\theta})\right) \right. \\ \times \prod_{X=\mu\mu, ee} \text{Pois}\left(d_i^X \mid \frac{\mu_{\text{SR},i}^{Z \rightarrow \nu\nu}}{T_{Z,i}^X(\boldsymbol{\theta})} + B_i^X(\boldsymbol{\theta})\right) \\ \times \prod_{X=b\mu, be} \text{Pois}\left(d_i^X \mid \frac{\mu_{\text{SR},i}^{t\bar{t}}}{T_{t\bar{t},i}^X(\boldsymbol{\theta})} + B_i^X(\boldsymbol{\theta})\right) \\ \left. \times \prod_{X=\mu, e} \text{Pois}\left(d_i^X \mid \frac{\mu_{\text{SR},i}^{W \rightarrow \ell\nu}}{T_{W,i}^X(\boldsymbol{\theta})} + \frac{\mu_{\text{SR},i}^{t\bar{t}}}{T_{t\bar{t},i}^X(\boldsymbol{\theta})} + B_i^X(\boldsymbol{\theta})\right)\right] \times \prod_{j=0}^{n_\theta} p_j(\theta_j) \quad (4.12) \end{aligned}$$

4.2.2 Theoretically-limited extrapolations

Despite the combination of the $\mu\mu$ and ee regions, there are still large statistical uncertainties in the estimate of $Z \rightarrow \nu\nu$ at high U . This is apparent in Figure 4.12, in which exactly one event is observed in the last bin of the tight CR. The dilepton CRs are limited by $\sigma(pp \rightarrow Z \rightarrow \nu\nu) > \sigma(pp \rightarrow Z \rightarrow \ell^+ \ell^-)$; accordingly, to alleviate this limitation, we look to a process with a much bigger cross-section. In similar regions of final-state phase space and detector acceptance, $\sigma(pp \rightarrow \gamma) \sim 30 \times \sigma(pp \rightarrow Z \rightarrow \nu\nu)$ and $\sigma(pp \rightarrow W \rightarrow \ell\nu) \sim 2 \times \sigma(pp \rightarrow Z \rightarrow \nu\nu)$. Therefore, it is natural to use the W +jet and γ +jet production spectra to further estimate the Z +jet spectrum. For concreteness, we will begin with the latter, more powerful, case. First, we select γ +jet events using the selection in Table 4.7.

As before, let us define another transfer factor:

$$T_{\gamma,i}^\gamma = \frac{N_i^{\text{SR}}(Z \rightarrow \nu\nu)}{N_i^\gamma(\gamma)} \quad (4.13)$$

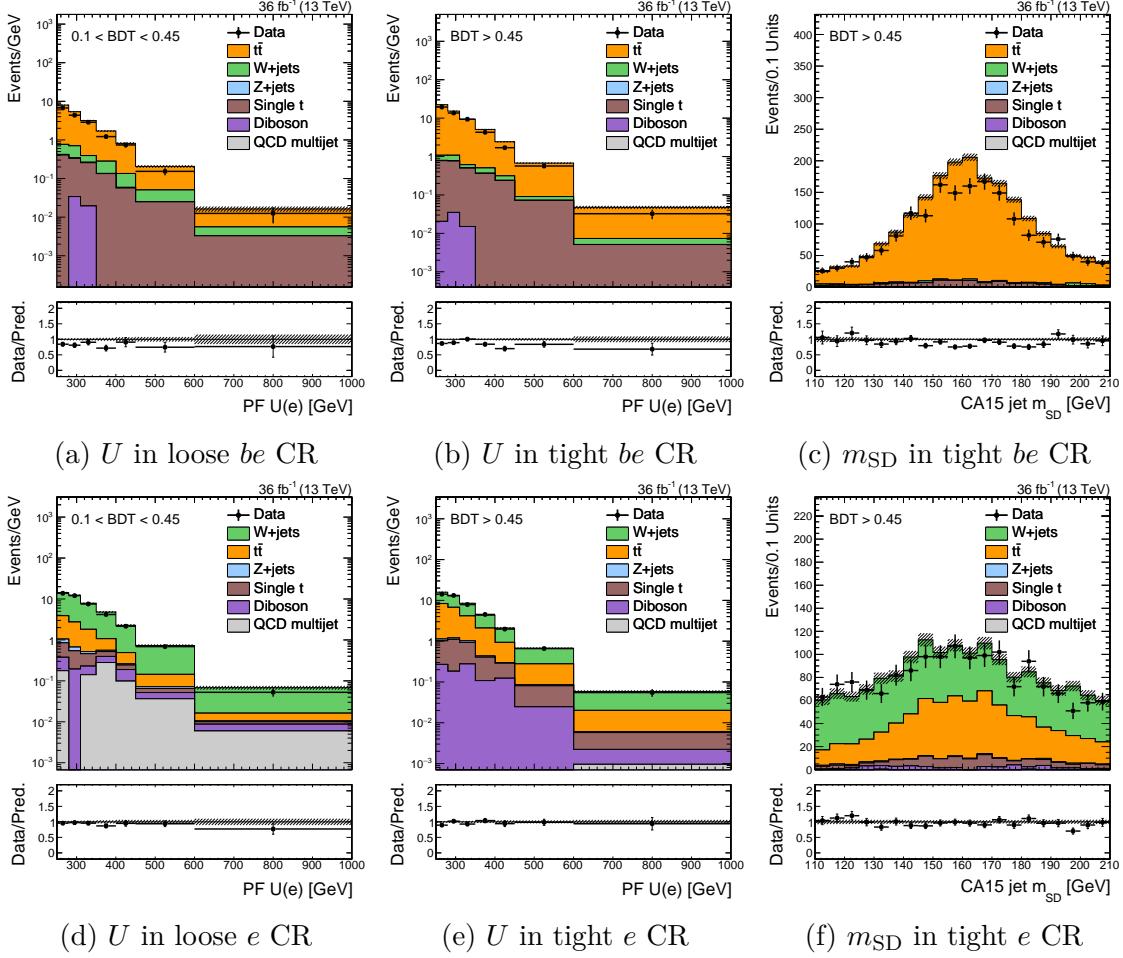


Figure 4.18: Various kinematic distributions in the mono-top be CRs (top) and e CRs (bottom). All predicted distributions are prior to the maximization of the likelihood, and the grey band refers only to the statistical uncertainty of the MC. Although a flat offset is observed in all $t\bar{t}$ spectra, it is also observed in the corresponding muon final state (Figure 4.15). Such an effect cancels in the transfer factor and is corrected for by the fit (Figure 4.26).

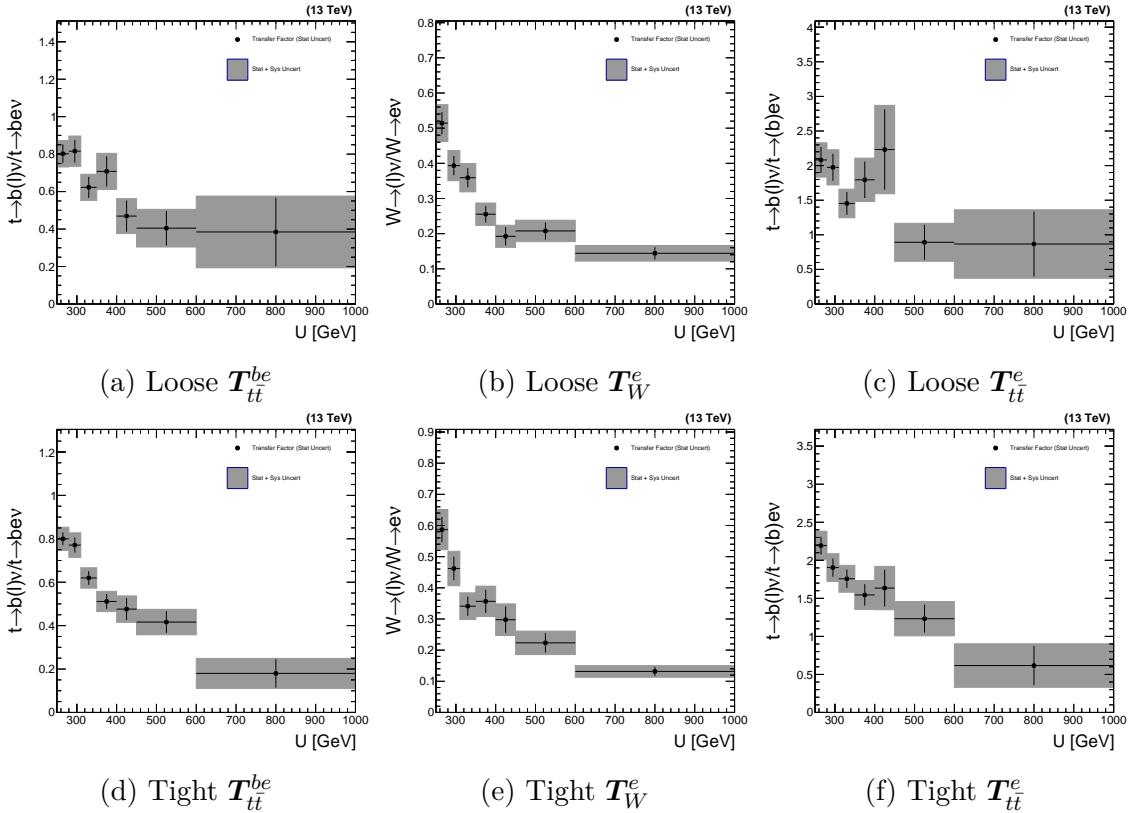


Figure 4.19: The transfer factors \mathbf{T}_{tt}^{be} , \mathbf{T}_W^e , and \mathbf{T}_{tt}^e

Table 4.7: Criteria used to select events for the mono-top γ CR. As in the SR, the region is further subdivided based on the jet BDT score.

Criterion	Notes
$U > 250$ GeV	Same as the SR selection, so as to select γ s in the same phase space as the targeted Z s.
1 CA15 jet with $p_T > 250$ GeV	Same as SR
$CA15 \text{ jet } 110 < m_{SD} < 210 \text{ GeV}$	Same as SR
No identified e, μ, τ_h	Same as SR.
Well-identified γ with $p_T > 175$ GeV	Same as SR, minimum p_T set by trigger threshold.
$\min_{\text{jets}} \Delta\phi(\text{jet}, U) > 0.5$	Same as SR
CA15 jet BDT	Same as SR

where $N_i^\gamma(\gamma)$ is the number of γ +jet events in bin i of the photon CR. The other \mathbf{T}_Y^X we have discussed so far correlate similar processes, such as $Z \rightarrow \nu\nu$ and $Z \rightarrow \mu\mu$, and theoretical uncertainties in the predictions of the numerator and denominator cancel in the transfer factor.. In contrast, \mathbf{T}_γ^γ is sensitive to the theoretical predictions of the hadronic recoil spectra in Z and γ events.

LO-to-NLO corrections

To reduce the impact of higher-order effects on \mathbf{T}_γ^γ , we ensure that the numerator and denominator are predicted to as high an order as possible. Leading-order MC is used to simulate V +jet processes, because higher-order simulations are much more computationally intensive. We therefore choose to produce a less accurate LO simulation, as opposed to a more accurate, but statistically-limited, NLO QCD simulation. However, certain inclusive distributions can be computed at NLO. Since $U \approx p_T^V$ is the quantity of interest in this analysis, we want to ensure this distribution is accurately predicted. It is clear from Figure 4.20 that adding an additional QCD order induces corrections up to 50% across the p_T^V spectrum. In the LO simulation, we can obtain an estimate of the uncertainty due to NLO effects by varying the renormalization and factorization scales (μ_R, μ_F) by factors of two. This is represented by the red envelope and grey band in Figure 4.20 and is insufficient to cover all NLO effects. Therefore, we compute a simple correction for the NLO QCD effects, known as a k -factor:

$$k_{Z,\text{QCD}}(p_T^Z) = \frac{d\sigma_{\text{NLO QCD}}(Z)/dp_T^Z}{d\sigma_{\text{LO}}(Z)/dp_T^Z} \quad (4.14)$$

We include another k -factor, k_{EW} , to correct for higher-order EW effects. Unlike k_{QCD} , k_{EW} is derived using a theoretical calculation [90, 91, 92] instead of NLO simulation. The factor k_{EW} covers NLO EW terms, as well as large Sudakov logarithms that appear at high p_T^V in the NNLO expansion (NLL).

Figure 4.21 compares the k -factors for all three V +jet processes. While there are similar trends as a function of p_T^V , it is clear that the corrections are quite different

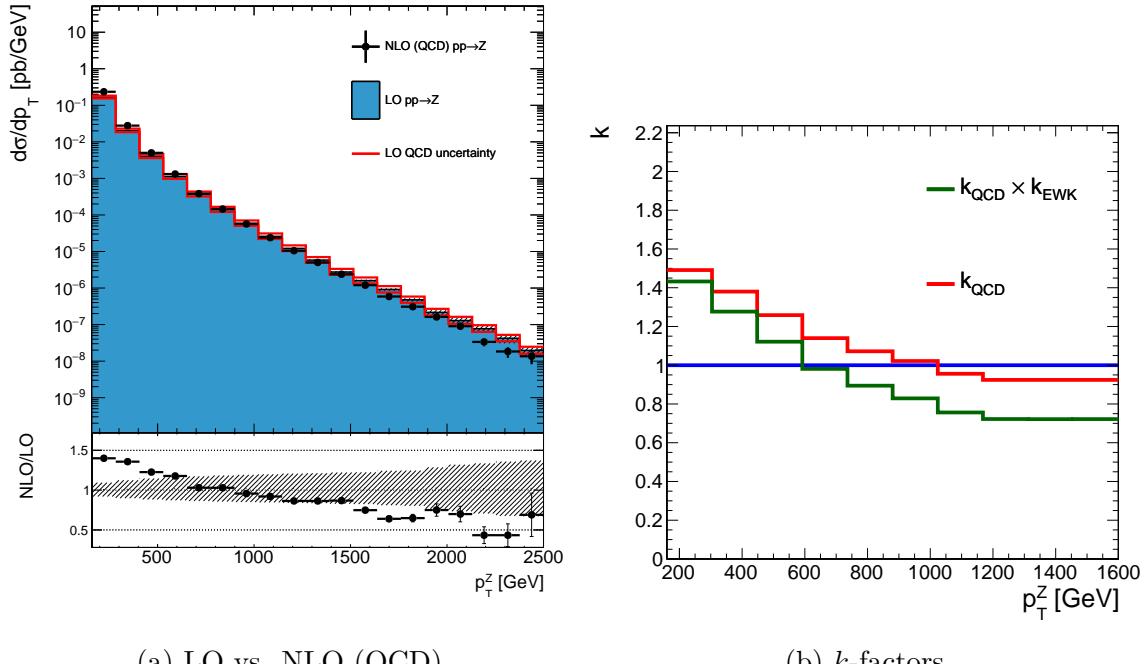


Figure 4.20: Theoretical predictions for p_T^Z in $Z + \text{jet}$ events and the corresponding k -factors. No detector simulation is applied in these figures; all quantities are directly from MC simulation of the physics process. “LO QCD uncertainty” refers to an estimate of the effect of the QCD renormalization and factorization scales on the LO simulation. The grey band in the ratio is the quadrature sum of the QCD and statistical uncertainties.

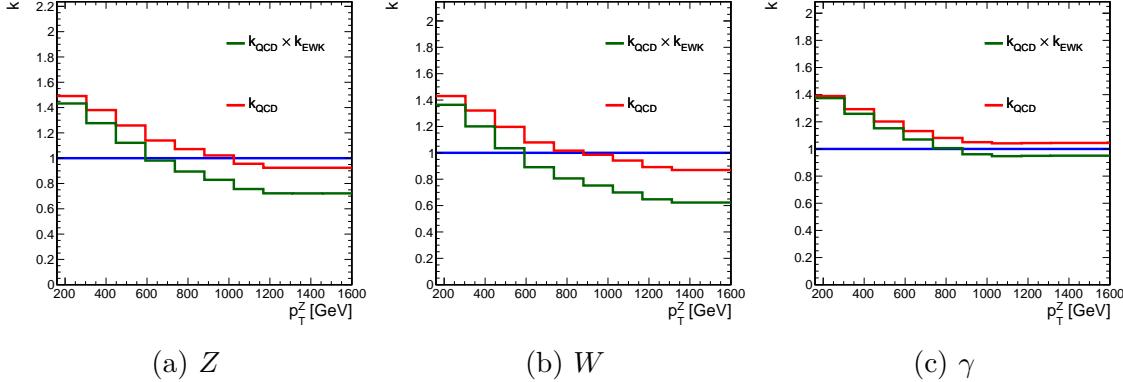


Figure 4.21: Values of k -factors as a function of p_T^V for each of the $V + \text{jet}$ processes. While the trend of the k -factors is similar between the three processes, the values are not identical. This leads to a non-trivial modification of the transfer factors.

for each process. Therefore, transfer factors like T_γ^γ are sensitive to NLO effects, i.e.:

$$T_{\gamma,i}^\gamma = \frac{N_i^{\text{SR}}(Z \rightarrow \nu\nu)}{N_i^\gamma(\gamma)} \neq \frac{N_i^{\text{SR,LO}}(Z \rightarrow \nu\nu)}{N_i^{\gamma,\text{LO}}(\gamma)} \quad (4.15)$$

The distributions in Sections 4.1-4.2 are all corrected using these k -factors. Figure 4.22 shows the equivalent distributions for CRs that target $\gamma + \text{jet}$ events.

Higher-order uncertainties

Now that we can describe \mathbf{T}_γ^γ at NLO, we must assess the impact of unknown higher-order terms on the transfer factors. We account for variations caused by uncertainties in the PDF model by taking the RMS of the 100 parameter variations prescribed for the NNPDF3.0 set [43]. By varying μ_F and μ_R by factors of 0.5 and 2, we assess the effect of the integration scale choices on \mathbf{T} . These scale and PDF uncertainties cover all unknown QCD effects on the production of electroweak bosons. To be conservative, they are assumed to be uncorrelated between processes. However, the uncertainties are correlated between all bins (i.e. as a function of p_T^V). A second set of uncertainties is included for higher-order EW effects, following what is suggested in References [93, 94, 95, 90, 91, 96, 97, 92, 98] and agreed upon in the LHC Dark Matter Working Group. These EW uncertainties break down into three categories:

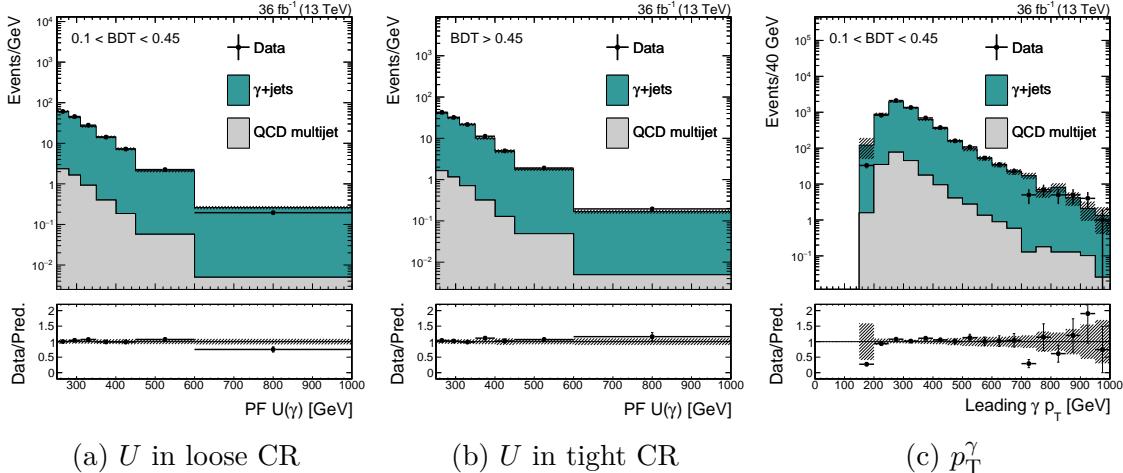


Figure 4.22: Various kinematic distributions in the two mono-top γ CRs. All predicted distributions are prior to the maximization of the likelihood, and the grey band refers only to the statistical uncertainty of the MC. The pre-fit MC describes the observed data reasonably well.

- Unknown Sudakov logarithms in the NLL correction. These uncertainties are correlated across processes (Z, W, γ).
- Missing NNLO EW effects not covered by the NLL correction. These are not correlated across processes.
- The full difference between the NLL correction and an exponentiation of the NLO correction; also not correlated across processes.

It should be stressed that while these uncertainties apply to the prediction of each V +jet process, they do not affect transfer factors that correlate processes differing only in decay mode or acceptance [91, 36]. This is simply because these uncertainties primarily deal with the initial state or the production of an electroweak boson, which are not related to the description of the decay to leptons or the experimental identification of leptons and b -jets. That is:

$$T_{Z,i}^{\mu\mu} = \frac{N_i^{\text{SR}}(Z \rightarrow \nu\nu)}{N_i^{\mu\mu}(Z \rightarrow \mu\mu)} \approx \frac{N_i^{\text{SR,LO}}(Z \rightarrow \nu\nu)}{N_i^{\mu\mu,\text{LO}}(Z \rightarrow \mu\mu)} \quad (4.16)$$

Now that we have tools to construct transfer factors of the form $N(V)/N(V')$ with

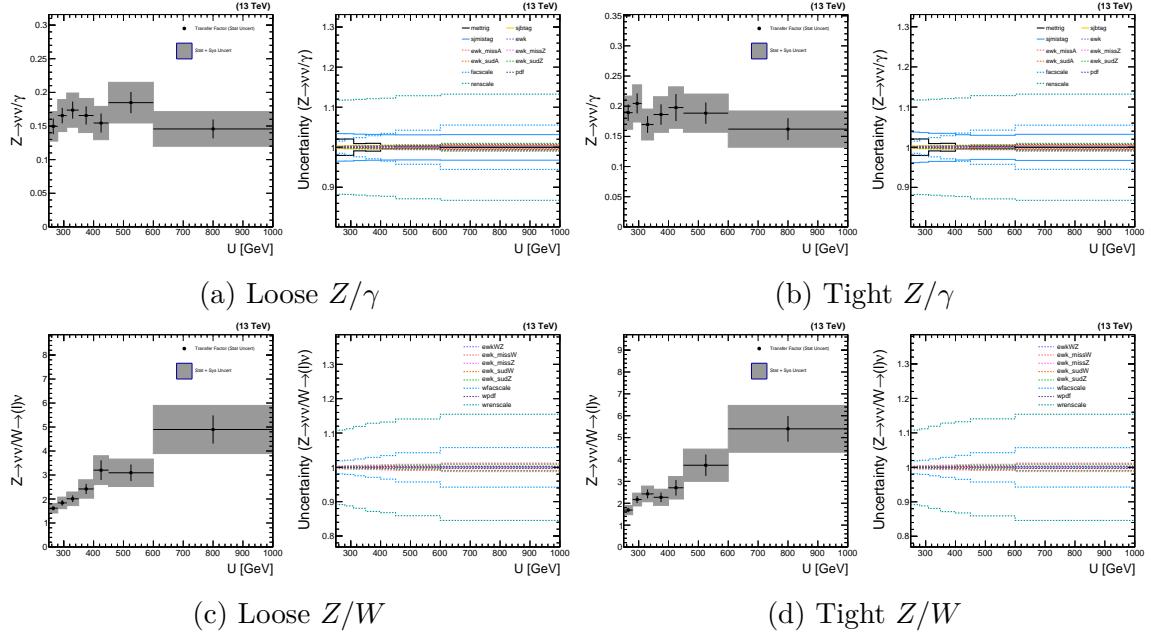


Figure 4.23: The transfer factors T_γ^γ and $T_{Z/W}^{\text{SR}}$; and corresponding shape uncertainties. The theoretical uncertainties dominate these extrapolations and are bigger than the MC Poisson and experimental uncertainties.

uncertainties smaller than the statistical uncertainty of the data, we can also add the Z/W constraint to the likelihood:

$$T_{Z/W,i}^{\text{SR}} = \frac{N_i^{\text{SR}}(Z \rightarrow \nu\nu)}{N_i^{\text{SR}}(W \rightarrow \ell\nu)} \quad (4.17)$$

This allows us to use the e, μ CRs (which target $W+\text{jet}$ production) to further reduce the uncertainty in the estimation of $Z \rightarrow \nu\nu$ in the SR. For technical reasons, the transfer factor is defined as the ratio Z/W in the SR. However, the SR and the μ CR are connected through a product of transfer factors:

$$N_i^{\text{SR}}(Z \rightarrow \nu\nu) = T_{Z/W,i}^{\text{SR}}(\hat{\boldsymbol{\theta}}) \times T_{W,i}^\mu(\hat{\boldsymbol{\theta}}) \times N_i^\mu(W \rightarrow \ell\nu) \quad (4.18)$$

where $\hat{\boldsymbol{\theta}}$ is the maximum-likelihood estimate of $\boldsymbol{\theta}$.

Figure 4.23 shows these additional transfer factors and their shape uncertainties. It is clear from inspection that $T_\gamma^\gamma \ll 1$, and the same holds for the effective transfer factor $T_{Z/W}^{\text{SR}} \times T_W^\mu$. This indicates that the CR data to which the transfer factor is

linked has greater statistical power than the SR data.

Having included all of these components, the likelihood can be written as:

$$\begin{aligned}
\mathcal{L}(\mathbf{d} \mid \mu, \boldsymbol{\mu}_{\text{SR}}^{Z \rightarrow \nu\nu}, \boldsymbol{\mu}_{\text{SR}}^{t\bar{t}}, \boldsymbol{\theta}) = & \\
\prod_{i \in \text{bins}} \left[\text{Pois}\left(d_i^{\text{SR}} \mid \mu S_i^{\text{SR}}(\boldsymbol{\theta}) + \mu_{\text{SR},i}^{Z \rightarrow \nu\nu} + \frac{\mu_{\text{SR},i}^{Z \rightarrow \nu\nu}}{T_{Z/W,i}^{\text{SR}}(\boldsymbol{\theta})} + \mu_{\text{SR},i}^{t\bar{t}} + B_i^{\text{SR}}(\boldsymbol{\theta})\right) \right. & \\
\times \prod_{X=\mu\mu,ee} \text{Pois}\left(d_i^X \mid \frac{\mu_{\text{SR},i}^{Z \rightarrow \nu\nu}}{T_{Z,i}^X(\boldsymbol{\theta})} + B_i^X(\boldsymbol{\theta})\right) & \\
\times \prod_{X=b\mu,be} \text{Pois}\left(d_i^X \mid \frac{\mu_{\text{SR},i}^{t\bar{t}}}{T_{tt,i}^X(\boldsymbol{\theta})} + B_i^X(\boldsymbol{\theta})\right) & \\
\times \prod_{X=\mu,e} \text{Pois}\left(d_i^X \mid \frac{\mu_{\text{SR},i}^{Z \rightarrow \nu\nu}}{T_{W,i}^X(\boldsymbol{\theta}) T_{Z/W,i}^{\text{SR}}(\boldsymbol{\theta})} + \frac{\mu_{\text{SR},i}^{t\bar{t}}}{T_{t\bar{t},i}^X(\boldsymbol{\theta})} + B_i^X(\boldsymbol{\theta})\right) & \\
\times \text{Pois}\left(d_i^\gamma \mid \frac{\mu_{\text{SR},i}^{Z \rightarrow \nu\nu}}{T_{\gamma,i}^\gamma(\boldsymbol{\theta})}\right] \times \prod_{j=0}^{n_\theta} p_j(\theta_j) & \quad (4.19)
\end{aligned}$$

The discussion in this section has largely relied on arguments from simulation and calculation. We can, however, validate that our estimates of \mathbf{T} and the corresponding uncertainties are reasonable by using CR data and appropriate proxies. For example, to test $\mathbf{T}_\gamma^\gamma \sim N(Z \rightarrow \nu\nu)/N(\gamma)$, we can look at $N(Z \rightarrow \mu\mu)/N(\gamma)$. Up to differences in branching ratio and muon identification, these ratios are identical. Figure 4.24 shows some examples of these proxy ratios. In particular, we see that the Z/γ , Z/W and W/γ ratios are well-described the MC, especially as compared to the systematic uncertainties that are assigned.

4.3 Results

Having built this likelihood, we perform a maximum likelihood fit to the data in all regions simultaneously. The minimization of the negative log likelihood is done by the Minuit2 algorithm, through an interface provided by the RooStats package [99]. The results of a ML fit under a background-only hypothesis, with $\mu = 0$, is shown in Figures ??-??.

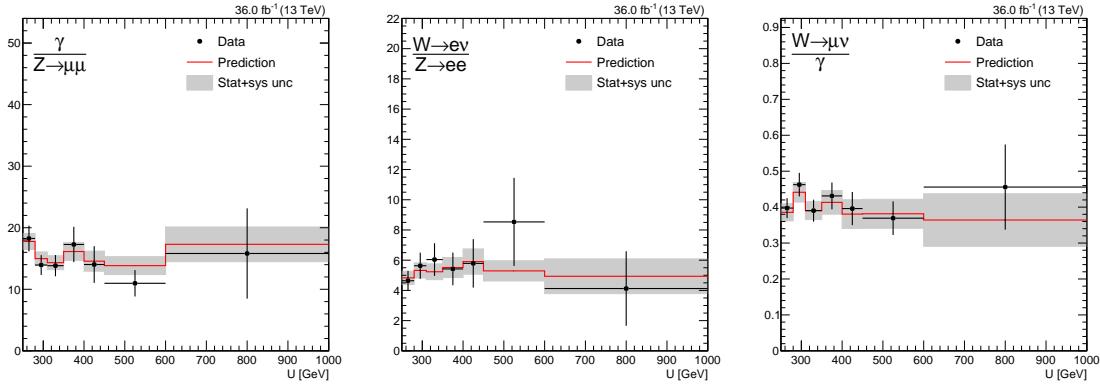


Figure 4.24: Data validation of CR-to-CR transfer factors in the loose category. Only ratios with theoretically-limited systematic uncertainties are shown.

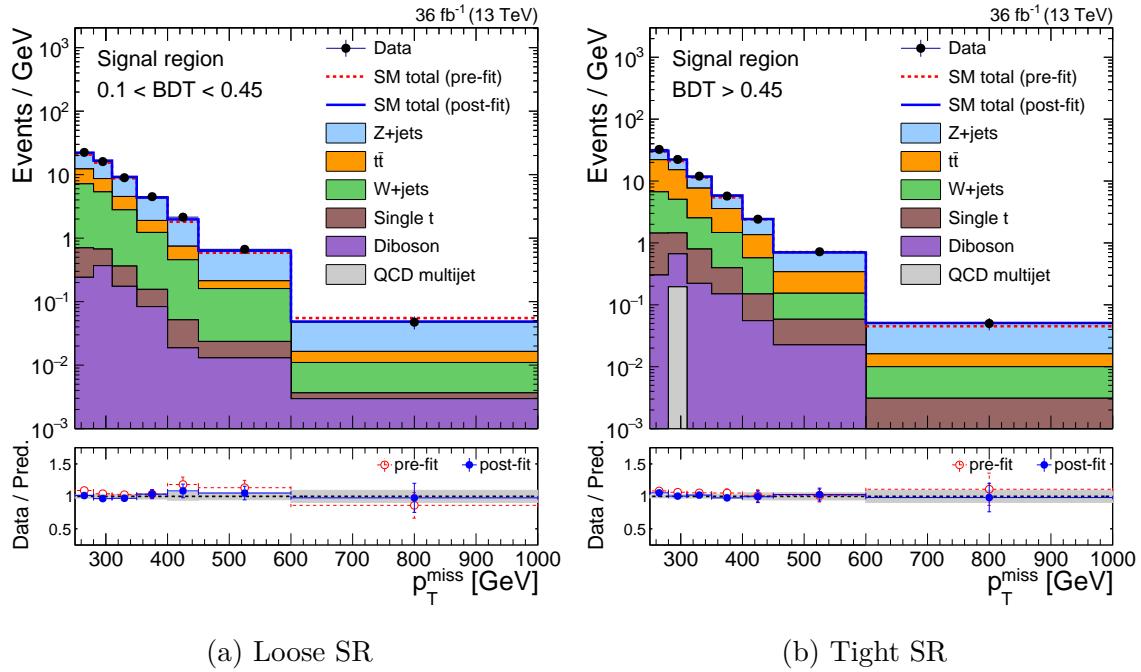


Figure 4.25: Comparison of pre- and post-fit results in the SRs, after simultaneously fitting all channels. The fit is performed having fixed $\mu = 0$.

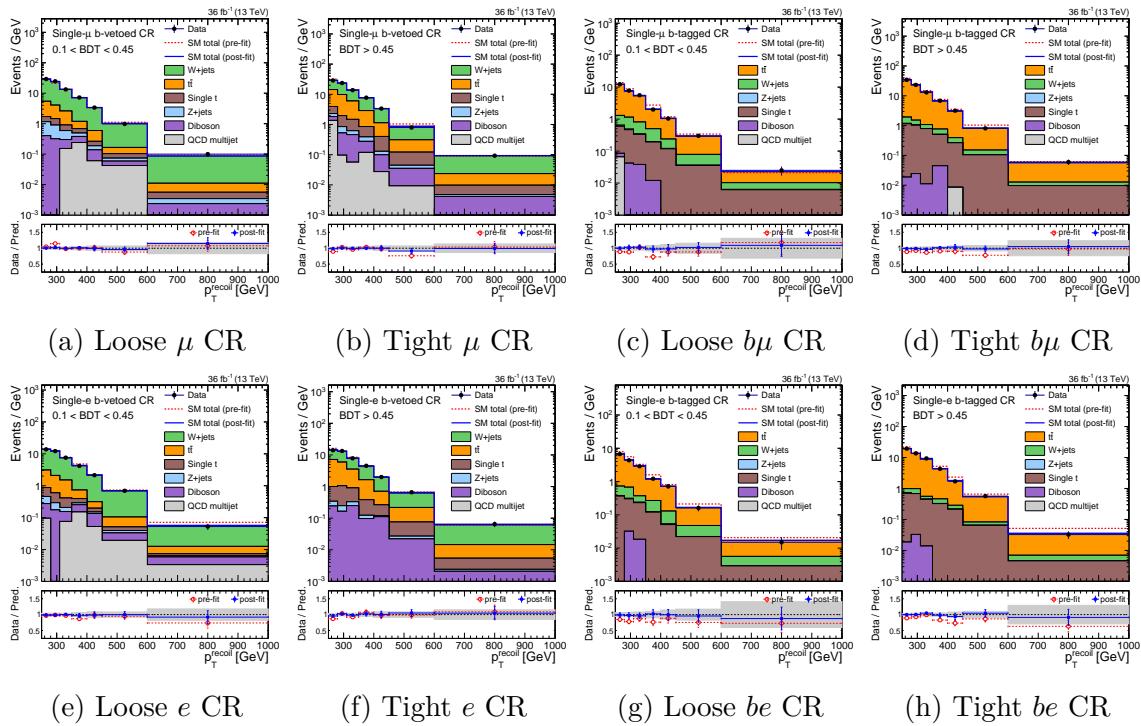


Figure 4.26: Comparison of pre- and post-fit results in the single-lepton CRs, after simultaneously fitting all channels. The fit is performed having fixed $\mu = 0$.

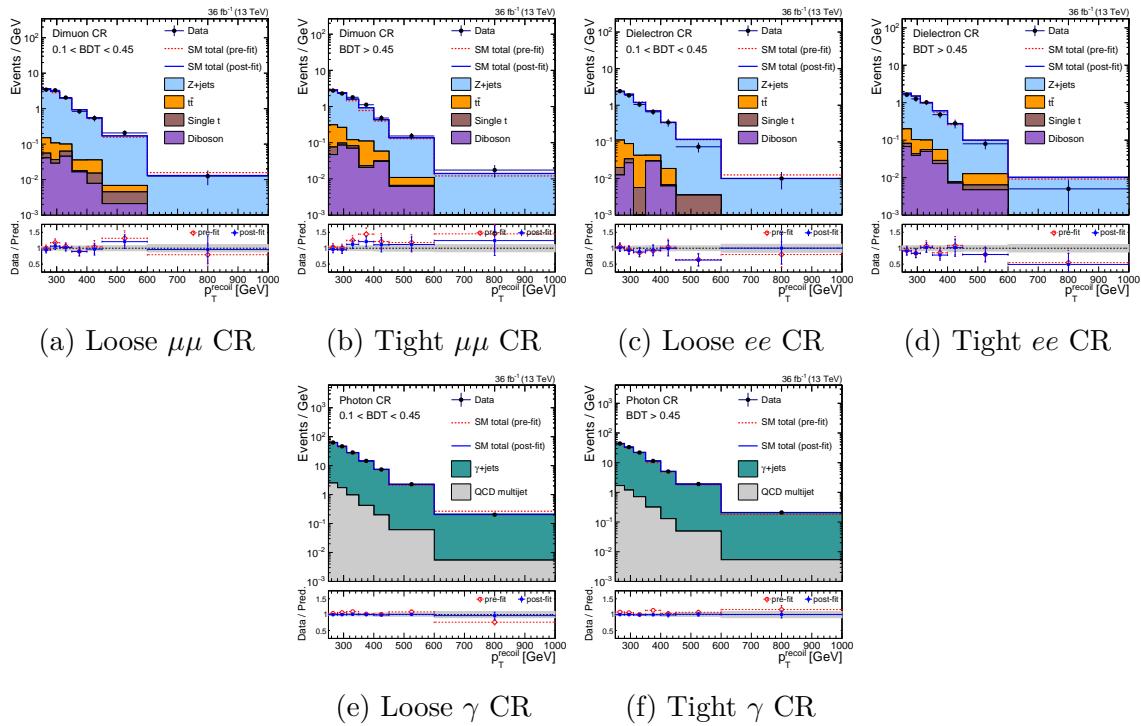


Figure 4.27: Comparison of pre- and post-fit results in the di-lepton and photon CRs, after simultaneously fitting all channels. The fit is performed having fixed $\mu = 0$.

4.3.1 Constraints on mono-top models

The results of the previous section show that the background-only (B-only) model is able to describe the data. To quantify the preference of the B-only model to a signal+background model (S+B), we compute 95% CLs upper limits on the signal strength μ for each signal hypothesis. A description of CLs upper limits and the asymptotic profile likelihood method is given in Appendix A.

We begin with the simpler of the two models: the resonant scalar. Figure 4.28 shows the upper limit on the cross section of ϕ production as a function of m_ϕ . A fixed value of m_ψ is chosen such that $m_\psi \ll m_\phi$. In this regime, the exact value of m_ψ does not strongly affect the kinematics or cross section; in the opposite regime, the decay $\phi \rightarrow t\psi$ is strongly suppressed. The values of the couplings a_q, b_q and a_ψ, b_ψ are similarly fixed to reasonable values, as agreed upon by the LHC Dark Matter Working Group (DMWG) [100]. The cross section, and therefore μ , roughly scales as $(a_q + b_q)^2$ and $(a_\psi + b_\psi)^2$. Given these assumptions, the observed data excludes scalars with mass $m_\phi < 3.4$ TeV. This can be compared to the expected exclusion, which is also 3.4 TeV.

The FCNC model contains many more non-trivial parameters, and so we do not simply fix all but one. First, fixing the couplings at DMWG benchmarks [100], we show the upper limits as a function of (m_V, m_χ) in Figure 4.29. Both vector-like and axial vector-like couplings are probed independently. In either scenario, assuming $m_\chi < 50$ GeV, we are able to exclude $0.2 < m_V < 1.75$ TeV. The lower bound arises from measurements of m_t and Γ_t ; allowing $m_V \lesssim 200$ GeV modifies Γ_t beyond measured bounds as the $t^{(*)} \rightarrow Vu$ channel opens. As m_χ crosses the $m_V/2$ boundary, σ_{theory} drops off rapidly, reducing the strength of the exclusion. In the vector case, this defines a clear exclusion triangle bounded by $m_V < 1.75$ TeV and $2m_\chi < m_V$. In the axial-vector case, Γ_V is much narrower, modifying the transition to the off-shell region. On the entire plane, the observed exclusion is consistent with the median expected exclusion within one standard deviation.

Sticking to two-dimensional projections of the parameter space, Figure 4.30 shows

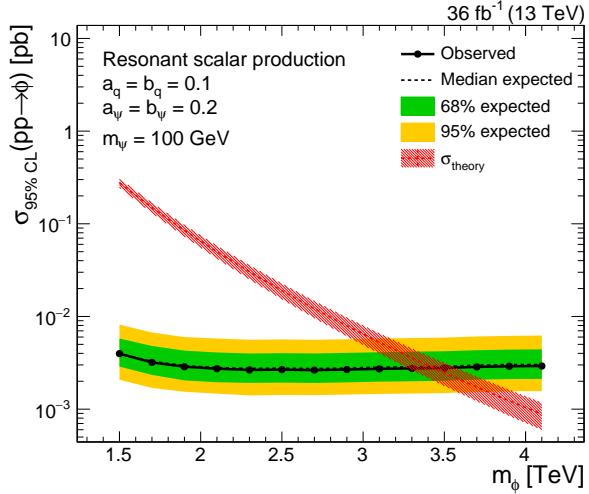


Figure 4.28: 95% CLs upper limits on the cross section of the production of ϕ , where the resonance mass m_ϕ is scanned. Values of m_ϕ for which $\sigma_{95\% \text{ CLs}} < \sigma_{\text{theory}}$ are excluded at 95% confidence level. The dashed black line represents the median expected exclusion, with the green and yellow bands representing the 1 and 2 standard deviation envelopes on the expected exclusion. The red band represents the 1 standard deviation of the uncertainty on the theory cross section (driven by higher-order QCD terms).

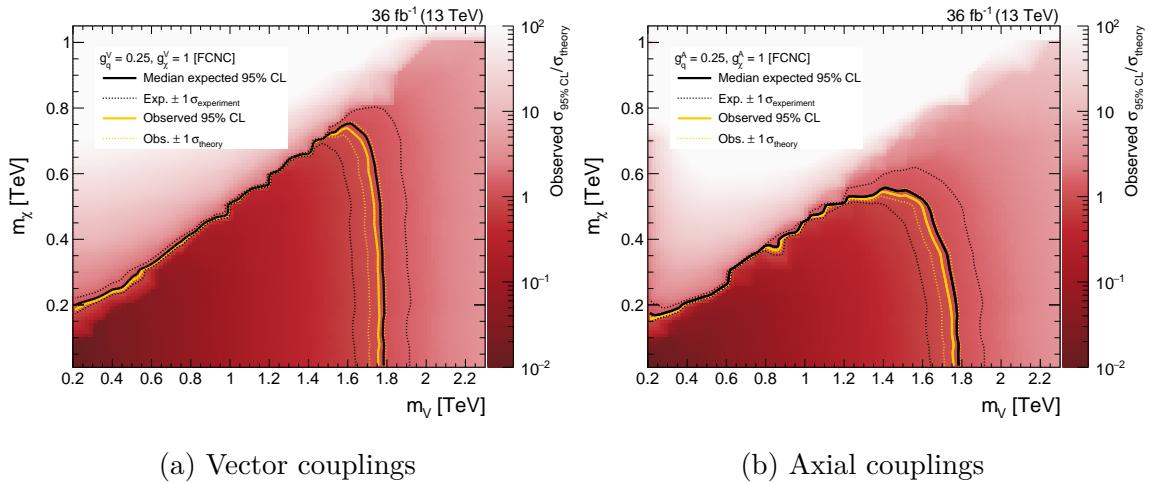


Figure 4.29: 95% CLs upper limits on the signal strength ($\mu = \sigma/\sigma_{\text{theory}}$) of the flavor-violating process $pp \rightarrow t\chi\bar{\chi}$ as a function of m_V and m_χ . Two hypotheses are tested: assuming $g_\chi^V, g_q^V \neq 0$ (vector-like) and assuming $g_\chi^A, g_q^A \neq 0$ (axial vector-like).

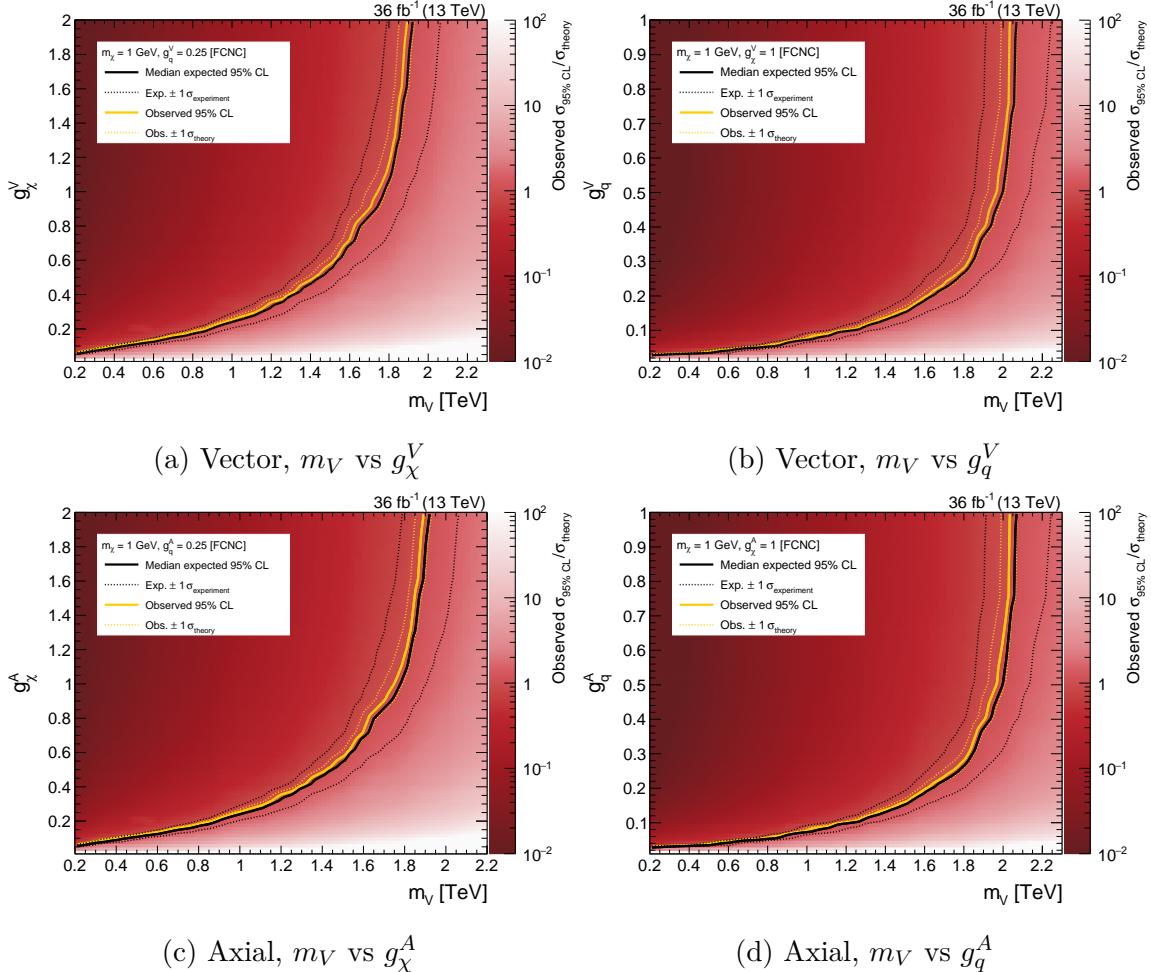


Figure 4.30: 95% CLs upper limits on the signal strength ($\mu = \sigma/\sigma_{\text{theory}}$) of the flavor-violating process $pp \rightarrow t\chi\bar{\chi}$ as a function of m_V , g_q , and g_χ . Two hypotheses are tested: assuming $g_\chi^V, g_q^V \neq 0$ (vector-like) and assuming $g_\chi^A, g_q^A \neq 0$ (axial vector-like). The DM mass m_χ is fixed at 1 GeV.

the excluded regions as a function of mediator mass m_V and all four free couplings $g_q^V, g_q^A, g_\chi^V, g_\chi^A$. As the DM mass is fixed to be strictly on-shell regardless of m_V , there are no visible differences between the vector and axial-vector scenarios. Assuming sufficiently low masses, we exclude couplings as weak as $g_q \cdot g_\chi \sim 0.01$.

It is difficult to fully visualize more than two dimensions of the parameter space at a time. Figure 4.31 shows the largest mediator mass that is excluded as a function of g_q^V, g_χ^V (assuming $g_q^A = g_\chi^A = 0$). Given sufficiently large couplings, we exclude FCNCs as massive as 2.5 TeV, while only assuming that $2m_\chi < m_V$.

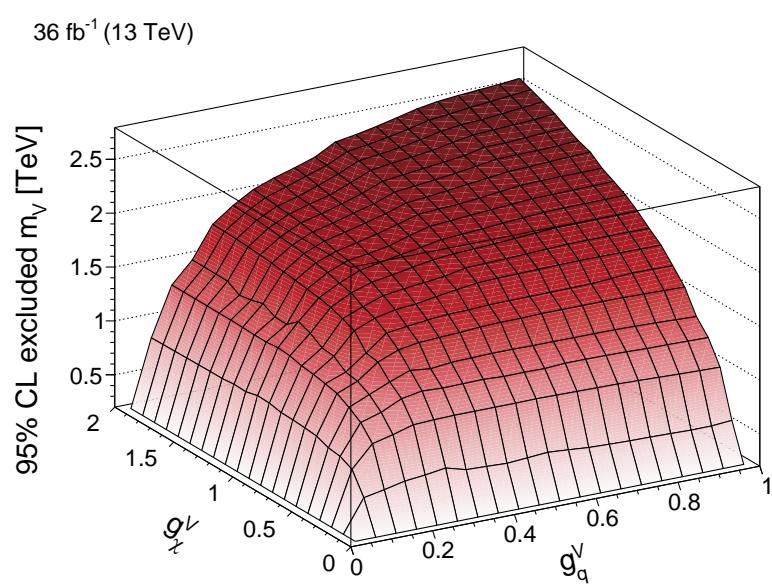


Figure 4.31: Maximum excluded value of m_V for each set of couplings, assuming vector-only couplings.

Chapter 5

The Search for Invisible Decays of the Higgs Boson

In this chapter, we use events with two forward jets and large missing momentum to probe the coupling of the Higgs boson to invisible particles with sufficiently low masses. The discovery of the SM Higgs boson [20] involved multiple production modes. Gluon fusion has the largest cross section (49 pb) at the LHC because of the large gluon PDF, followed by vector boson fusion (VBF) (3.8 pb), WH (1.4 pb), and ZH (0.89 pb) [101]. While gluon fusion is the most frequent mode, the unique detector signatures of the other production modes can be combined with various Higgs decay signatures to define signal topologies with smaller backgrounds.

Many DM models [?] allow for DM fermions or scalars to acquire mass through the Higgs mechanism, coupling to the SM Higgs boson. If the DM candidate χ satisfies $2m_\chi < m_H$, then we expect to observe $H \rightarrow \chi\bar{\chi}$. From measurements of the visible branching fractions, we can indirectly place an upper bound of $\mathcal{B}(H \rightarrow \chi\bar{\chi}) < 0.2$ [?]. In this chapter, we describe a direct search for $H \rightarrow \chi\bar{\chi}$ decays.

As with the case of the mono-top search, the $H \rightarrow \chi\bar{\chi}$ process manifests as p_T^{miss} . Each of the aforementioned Higgs production modes translates into a $p_T^{\text{miss}} + X$ signature, where X refers to one or more SM particles. Figure 5.1 shows each of the signatures; in this chapter, we will focus on the VBF production mode. This unique final state topology provides the best sensitivity to $H \rightarrow \chi\bar{\chi}$.

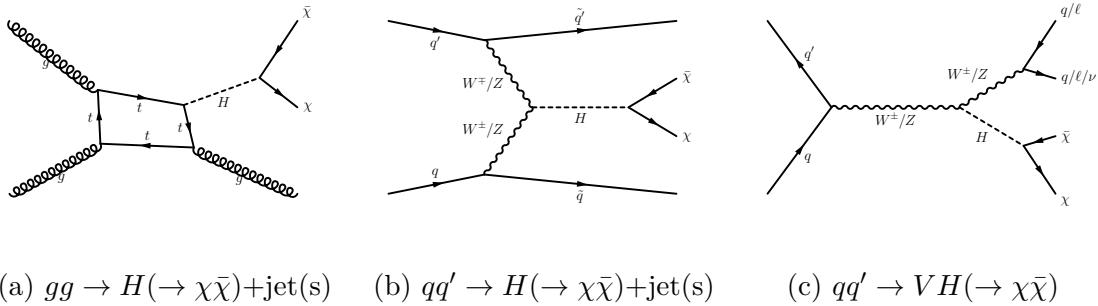


Figure 5.1: Diagrams that contribute to the production of the SM Higgs boson at the LHC, with the subsequent decay to DM candidates. The shown diagrams are all chosen to generate large p_T^{miss} through the presence of one or more SM particles in the final state.

5.1 Signal selection

VBF $H \rightarrow \chi\bar{\chi}$ events are characterized by large p_T^{miss} and two jets. These jets are typically:

- Fairly forward in the detector
- Far apart from each other in η
- Highly energetic (large E , moderate p_T)
- Close together in ϕ

A candidate VBF $H \rightarrow \chi\bar{\chi}$ event displaying these properties is shown in a CMS event display in Figure 5.2.

5.1.1 Online trigger selection

The same trigger decisions (L1 and HLT) as described in Section 4.1.1 are used to select events in this analysis. However, the L1 seeds for the 2016 data run were designed with mono-top-like analyses in mind; i.e., searches where the momentum imbalance is created by central objects. To avoid noise and resolution issues in the forward calorimeters, the L1 seed only considers energy deposits in the region $|\eta| < 3$. Therefore, VBF events in which both jets are in the forward region are not selected.

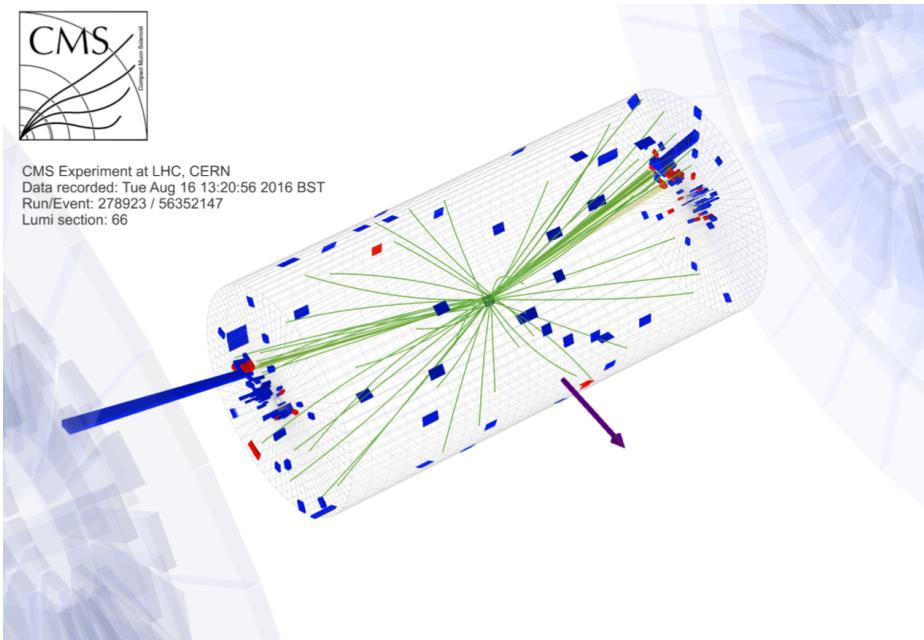


Figure 5.2: Candidate VBF $H \rightarrow \chi\bar{\chi}$ event with two energetic forward jets ($p_T = 180, 107$ GeV) and large p_T^{miss} (360 GeV). Red (blue) towers represent deposits in the hadronic (electromagnetic) calorimeter. Green lines are tracks reconstructed from hits of charged particles in the tracker. The blue arrow represents the direction and magnitude of the p_T^{miss} .

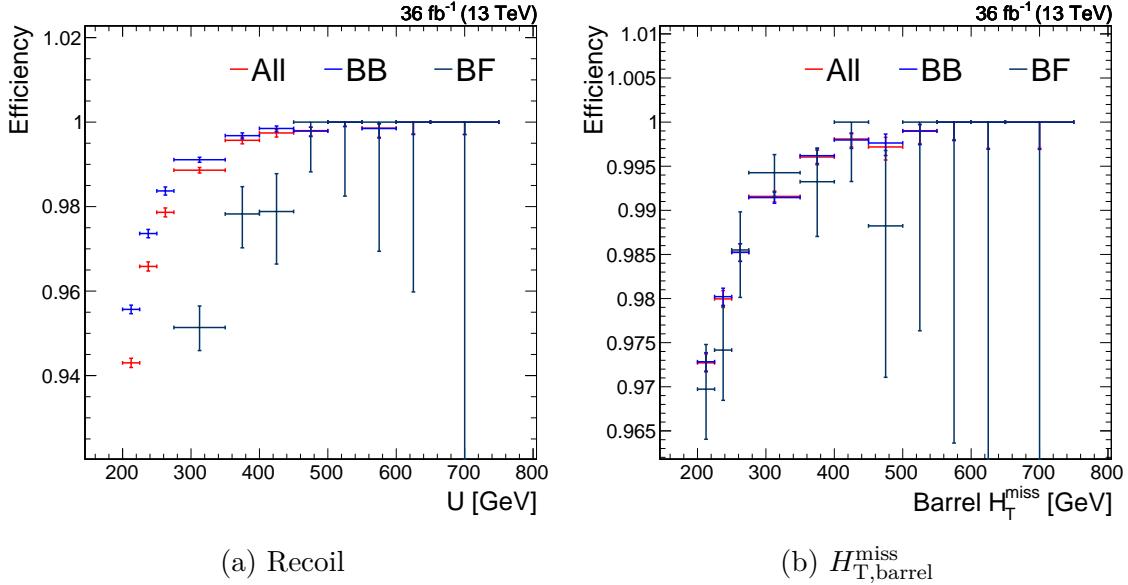


Figure 5.3: Trigger efficiency of events with a VBF-like topology (two jets with $p_T > 80, 40$ GeV) as a function of two different observables. Events are split into two categories: those where both jets have $|\eta| < 3$ (BB) and those where exactly one jet has $|\eta| > 3$ (FF). "All" refers to the sum of these categories.

This is visible in Figure 5.3a, where events are classified based on the location of the two highest- p_T jets. Events with both jets in the barrel (BB) have a higher efficiency than events with one jet in the forward detector (BF). Note that events with two forward jets (FF) are not considered at all, as the efficiency for such events is essentially zero.

The trigger efficiency is truly characterized by the energy deposited in the $|\eta| < 3$ region of the detector, and will be dominated in VBF events by the energy of jets. Accordingly, we define the *missing barrel hadronic transverse momentum*:

$$H_{T,\text{barrel}}^{\text{miss}} = \left| \left(\sum_{j \in \text{barrel}} \vec{p}_j \right)_T \right|, \text{ where barrel refers to jets with } |\eta| < 3 \quad (5.1)$$

As shown in Figure 5.3b, the three categories (BB, BF, All) have similar behavior as a function of $H_{T,\text{barrel}}^{\text{miss}}$. Therefore, we use this parameterization of the efficiency to correct MC simulation to match data.

A second L1-related issue that plagues the 2016 dataset is caused by a *pre-firing*

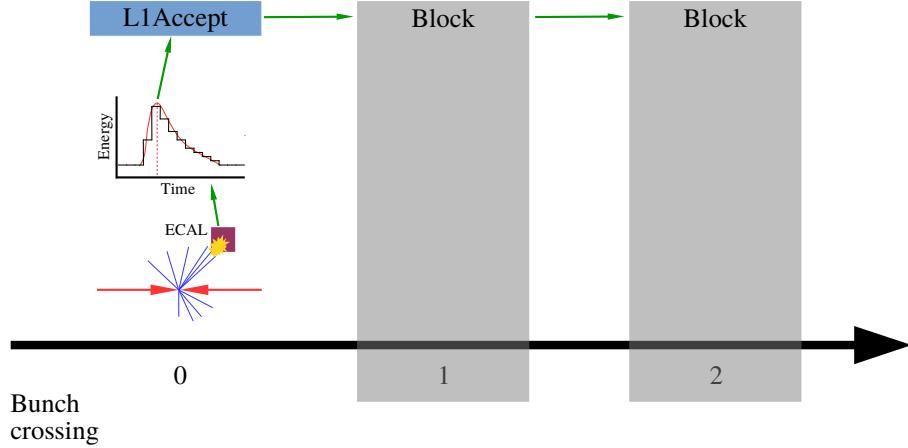


Figure 5.4: A normal event in which an ECAL seed triggers the L1A signal. The subsequent two bunch crossings are blocked. BX₀ refers to the event containing the physics object of interest. Green arrows indicate causality.

effect. When an L1 seed is triggered to accept an event (*L1Accept* or *L1A*), the following two bunch crossings (not necessarily corresponding to collisions) are blocked from firing L1As. At most, two in four consecutive events can fire L1A (i.e. the sequence L1A, blocked, blocked, L1A). Figure 5.4 is an example of a normal ECAL L1 seed accepting an event and blocking the subsequent bunch crossings. In what follows, we will refer to the bunch crossing with an interesting collision (i.e. the one we would like the trigger to select) as BX₀.

A pre-fire refers to the case in which a malformed detector signal is mis-reconstructed, so that the peak of the pulse appears to have occurred in the previous bunch crossing (BX₋₁). In this particular case, a region of the ECAL ($2.5 < |\eta| < 3$) suffered from a loss in transparency due to radiation damage and would produce pulse shapes that are poorly described by the model used to extract the pulse energy and time. When this happens, the L1 seeds for ECAL-based signatures (e.g. electron triggers) can fire an L1A for BX₋₁. This ECAL L1 seed in BX₋₁ will zero out the corresponding ECAL clusters in BX₀ (known as zero suppression), further biasing the event description. So, we would have an L1A for an arbitrary event (BX₋₁), and the interesting event (BX₀) would be blocked from passing the L1 altogether. This is depicted in Figure 5.5. Typically, BX₋₁ contains uninteresting physics signatures, and so is not

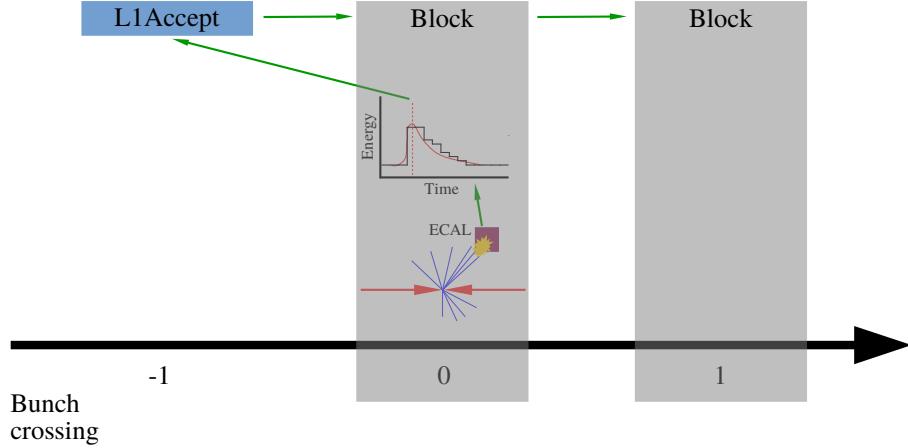


Figure 5.5: A pre-fired event in which an ECAL seed triggers the L1A signal for BX_{-1} . The subsequent two bunch crossings (including the one of interest) are blocked. BX_0 refers to the event containing the physics object of interest. Green arrows indicate causality.

accepted by the HLT.

To measure how often an ECAL energy deposit (typically left by a jet) causes an event to be lost by pre-firing, we need to compute the following efficiency:

$$\epsilon_{\text{pre-fire}}(p_T, \eta, \phi) = \frac{N_{\text{pre-fire}}(p_T, \eta, \phi)}{N_{\text{all events}}(p_T, \eta, \phi)} \quad (5.2)$$

However, by definition, pre-fired events cannot be recorded, and therefore $N_{\text{pre-fire}}(p_T, \eta, \phi)$ is difficult to measure. A very small subset of the recorded dataset (0.2%) consists of *un-pre-fireable* events. These are recorded events (BX_0) in which an L1A fired 3 bunch crossings prior (BX_{-3}). Due to the blocking rules, L1A cannot fire in BX_{-2} and BX_{-1} . Even if there is an ECAL seed in BX_0 that pre-fires, it will be blocked from firing an L1A, and therefore BX_0 is protected. If some other object in BX_0 manages to pass L1 and HLT decisions, then BX_0 will be recorded and can be studied. A schematic of such events is shown in Figure 5.6.

The L1 trigger system records trigger primitive (TP) information (4-vectors of physics objects considered in an L1 selection) for BX_{-1} if BX_0 is triggered. This means we can identify the cases in which a physics object in BX_0 coincides with a TP in BX_{-1} , indicating a pre-fire. Therefore (using the bunch crossing numbering in

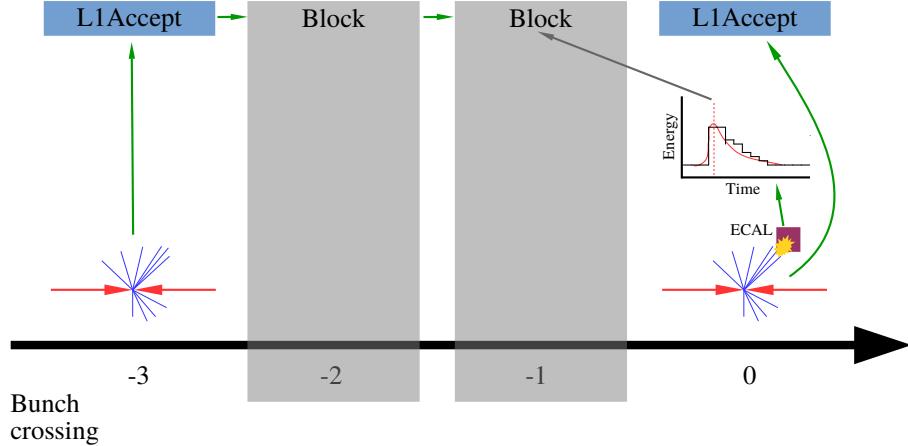


Figure 5.6: An un-pre-fireable event in which BX_{-3} protects BX_0 from being pre-fired. Green arrows indicate causality.

Figure 5.6), we re-define the efficiency:

$$\epsilon_{\text{pre-fire}}(p_T, \eta, \phi) = \frac{N_{\text{pre-fire } \text{BX}_0 | \text{BX}_{-3}}(p_T, \eta, \phi)}{N_{\text{BX}_0 | \text{BX}_{-3}}(p_T, \eta, \phi)} \quad (5.3)$$

By definition, all events in this ratio will be recorded. Figure 5.7 shows this efficiency as a function of jet location. We observe there is a *hot* ECAL tower near the location $\eta = -2.8$ and $\phi = 2$. Not only does this tower fire very frequently (leading to many particles, leading to many jets), but it almost always pre-fires. To first order, events with a jet in this crystal should be rejected. Beyond this, there is very little localization in the pre-fire probability (besides restriction to the ECAL endcap).

In Figure 5.8 we see $\epsilon_{\text{pre-fire}}$ as a function of p_T in a restricted η range. Firstly, we observe that $\epsilon_{\text{pre-fire}}$ increases as a function of p_T , and the turn-on is sharper as a function of EM p_T . This is explained by the mechanism of the pre-fire: the individual ECAL trigger seeds have a threshold of 30 GeV. The higher the jet p_T , the higher the probability of the jet depositing 30 GeV of EM energy in a localized area, setting off an L1 seed. Secondly, we observe a strong dependence on the reference triggers used to select BX_0 . For example, jet-based triggers (JetHT) lead to a much higher efficiency than p_T^{miss} -based triggers (MET). This is a consequence of zero suppression biasing the BX_0 triggers, as shown diagrammatically in Figure 5.9. Muon-based

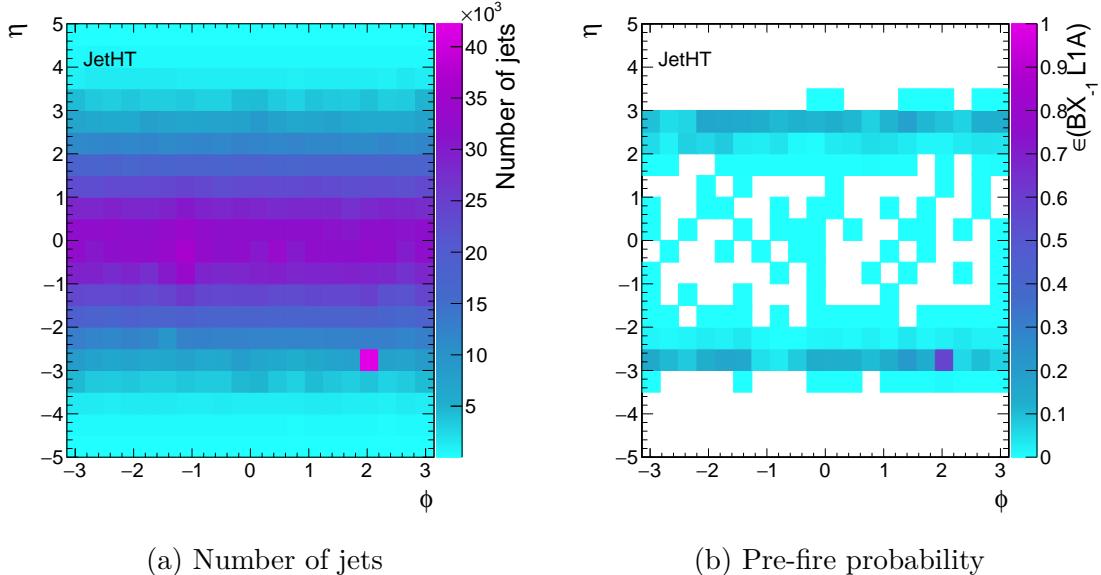


Figure 5.7: Distribution of jets and pre-fire events as a function of the jet location in the detector. Note the spike near $(\eta, \phi) = (-2.8, 2)$.

triggers (SingleMuon) are largely unaffected by the ECAL system, and therefore this measurement of $\epsilon_{\text{pre-fire}}$ is the least biased.

The probability of at least one jet pre-firing in an event is:

$$\epsilon_{\text{pre-fire}}^{\text{event}} = 1 - \prod_{j \in \text{jets}} \left(1 - \epsilon_{\text{pre-fire}}(p_T^j, \eta^j)\right) \quad (5.4)$$

The ϕ -dependence has been dropped, since it is clear from Figure 5.7 that the effect can be averaged over ϕ (once the spike is removed). Figure 5.10 shows $\epsilon_{\text{pre-fire}}(p_T, \eta)$ using muon-triggered and jet-triggered events. In the former case, statistical fluctuations make the region with $p_T > 250$ GeV unusable. Fortunately, this is the region in which the trigger bias is smallest, and so we switch to the jet triggered measurement above this threshold. A 20% uncertainty is assessed on the efficiency, which is derived from the difference between the SingleMuon and JetHT measurements.

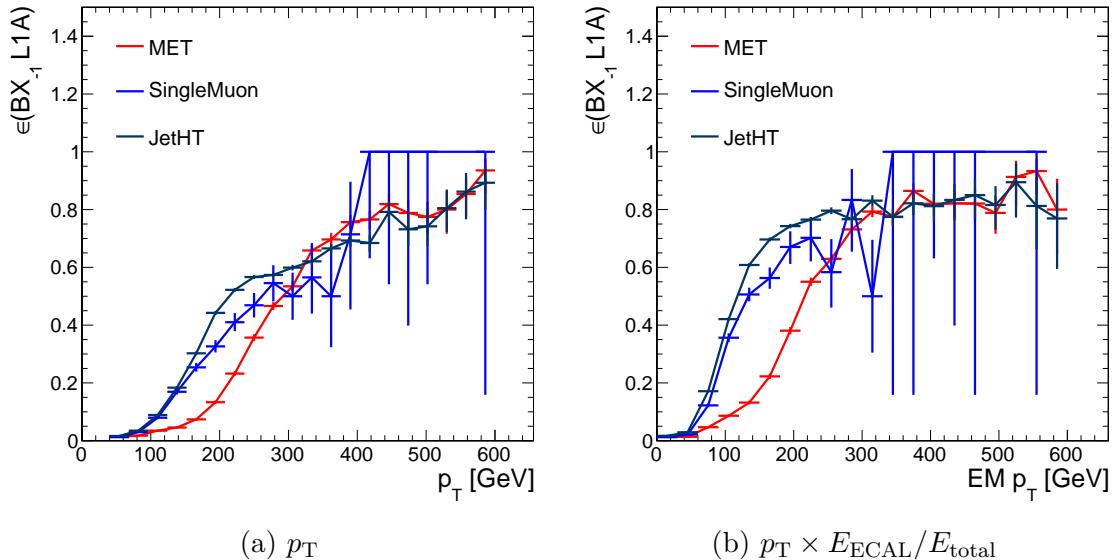


Figure 5.8: Probability that a given jet with $2.25 < |\eta| < 3$ causes a pre-fire in the L1 trigger due to ECAL mistiming. Two parameterizations are used: jet p_T and EM p_T . The three curves refer to which set of triggers are used to select BX_0 .

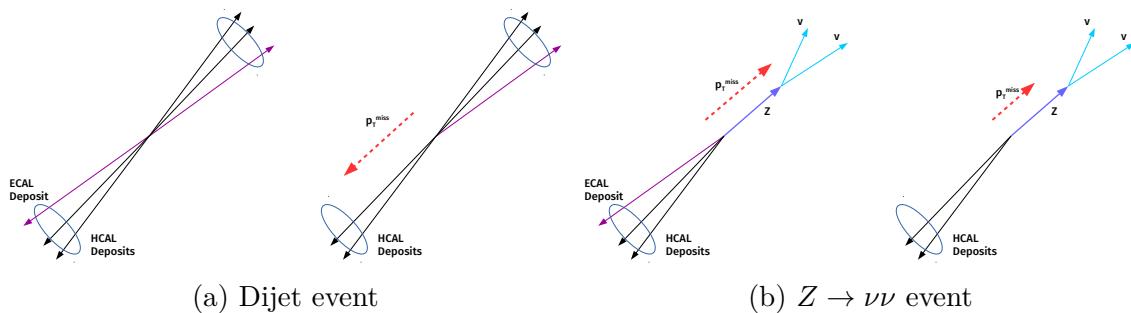


Figure 5.9: How zero suppression in the ECAL due to pre-firing can bias certain events. Subfigure (a) shows a dijet event, in which the loss of an ECAL deposit reduces the total H_T of the event, thereby lowering the probability of a jet-based trigger to fire. Subfigure (b) shows a $Z \rightarrow \nu\nu$ event, in which the loss of an ECAL deposit reduces the total p_T^{miss} of the event.

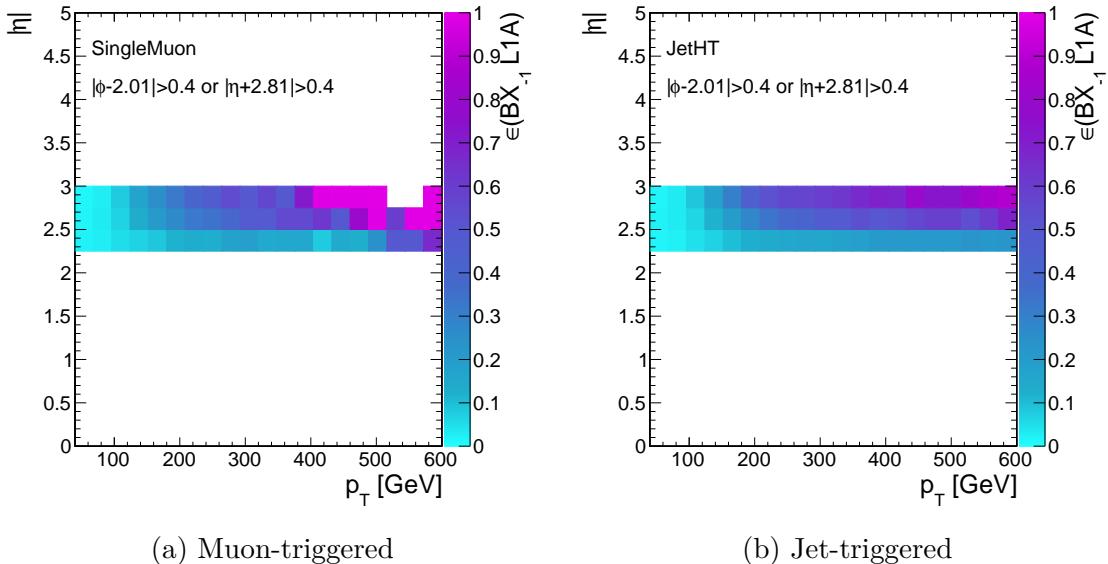


Figure 5.10: $\epsilon_{\text{pre-fire}}(p_{\text{T}}, \eta)$ with two different sets of reference triggers used to select BX_0 .

5.1.2 EW and QCD production of electroweak bosons

The primary backgrounds to the VBF production of invisibly-decaying Higgs bosons are $Z(\rightarrow \nu\nu) + 2 \text{ jet}$ and $W(\rightarrow \ell\nu) + 2 \text{ jet}$ production. At leading order, the relevant Feynman diagrams are either of the order $\alpha_{\text{EW}}^2 \alpha_{\text{QCD}}^4$ or α_{EW}^6 . We refer to the former as the QCD production mode and the latter as the EW mode. Examples Feynman diagrams are shown in Figure 5.11. The EW mode is essentially vector boson fusion, and so the terms EW and VBF will be used interchangeably.

As the vector boson is not directly detectable, the only experimental signatures are the jets. The jet kinematics are sensitive to the production mode (vector boson fusion vs QCD), as well as the spin of the produced boson. Some conclusions can be drawn from the kinematic distributions (Figure 5.12):

1. The yield ($\sigma \times A$) of the three VBF processes are relatively close in the relevant phase space (assuming $\mathcal{B}(H \rightarrow \chi\bar{\chi}) = 1$), but the QCD processes are 1-2 orders of magnitude higher.
2. The jet p_{T} and $p_{\text{T}}^{\text{miss}}$ distributions in the signal are comparable to or softer than

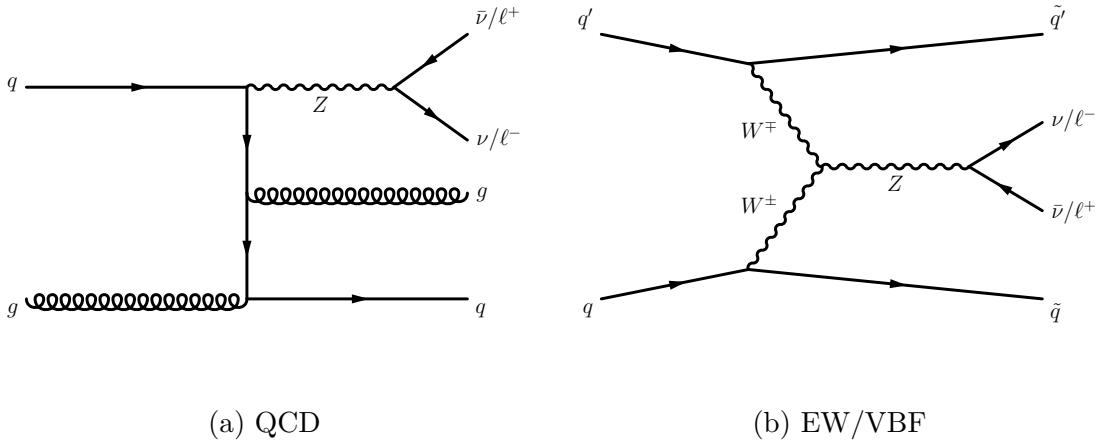


Figure 5.11: Examples of the two modes of producing Z bosons in association with 2 jets. Similar diagrams exist for W boson production.

the background processes. This is in contrast to other DM searches, in which the signal p_T^{miss} distribution is much harder than SM predictions.

3. VBF $H \rightarrow \chi\bar{\chi}$ produces fewer jets than SM processes.
 4. VBF $H \rightarrow \chi\bar{\chi}$ produces relatively forward jets. QCD $V + \text{jets}$ produces mostly central jets. VBF $V + \text{jets}$ produces jets that are somewhere between these distributions.

To fully exploit these kinematic distributions, we look at *VBF-tag* observables, which are functions of the two leading jets. These are defined as:

m_{jj} : Invariant mass of the dijet system.

$\Delta\eta_{jj}$: Absolute value of the difference in pseudorapidity of the two jets.

$\Delta\phi_{jj}$: Absolute value of the difference in azimuthal angle of the two jets.

These distributions are shown in Figure 5.13. The first two distributions look different in QCD and VBF processes and are therefore useful to reduce QCD backgrounds. On the other hand, $\Delta\phi_{jj}$ is sensitive to the spin of the boson produced in a VBF process, and therefore can distinguish between Higgs and electroweak boson production.

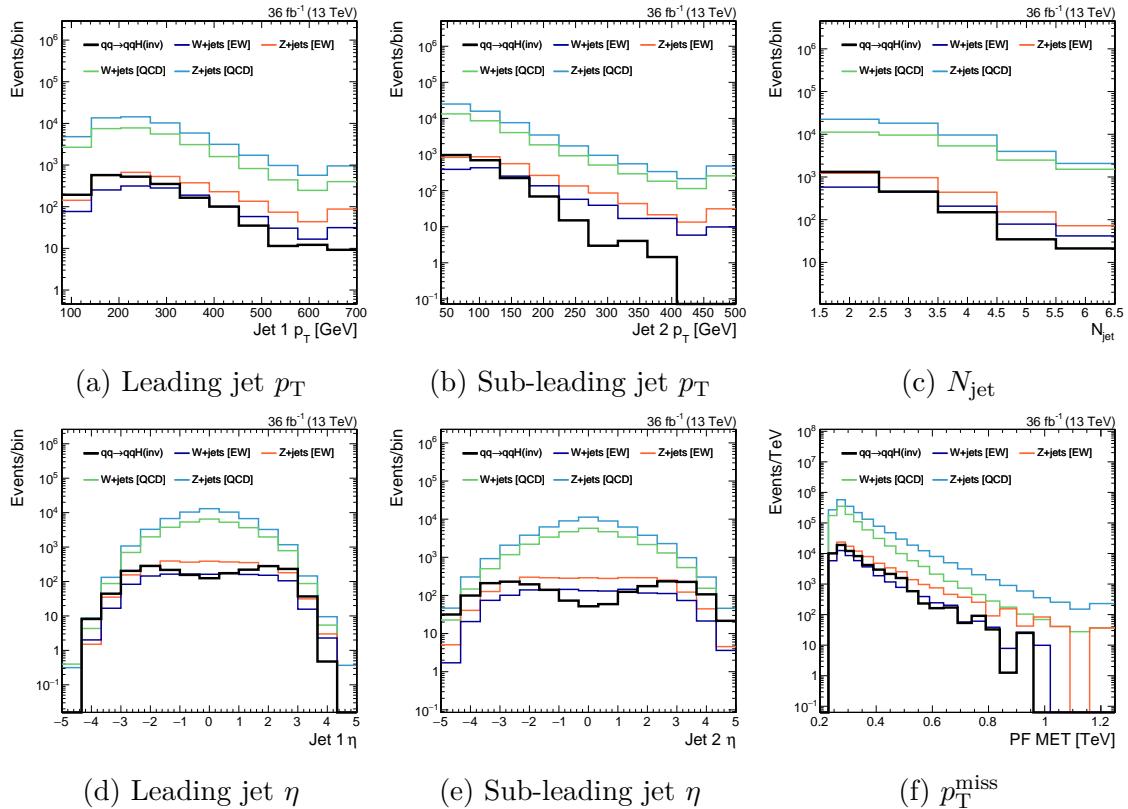


Figure 5.12: Event kinematic distributions, as compared between H vs Z vs W production, and VBF vs QCD modes.

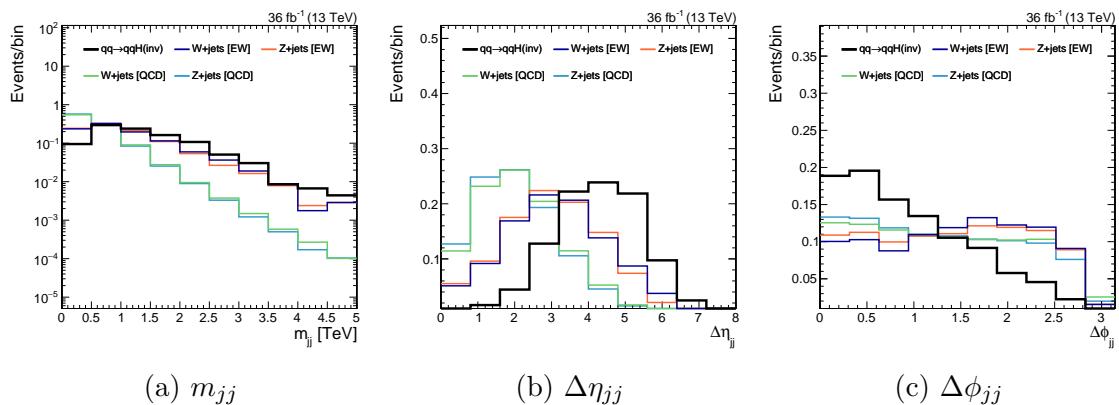


Figure 5.13: VBF tag observable distributions, as compared between H vs Z vs W production, and VBF vs QCD modes.

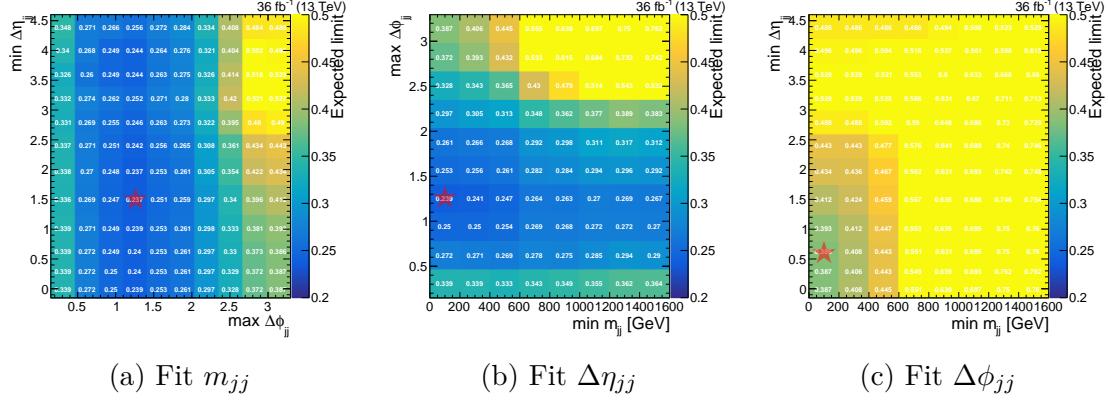


Figure 5.14: Optimization of the dijet kinematic selection, in three different fitting distribution scenarios.

5.1.3 Sensitivity optimization

A *baseline* selection is defined as:

- $p_T^{\text{miss}} > 250$ GeV: driven by trigger efficiency, as discussed in Section 5.1.1.
- $p_T^{\text{jet}} > 80, 40$ GeV: require two VBF jets, lower p_T thresholds set by trigger efficiency
- $N_{e,\mu,\tau,\gamma} = 0$: veto leptonic decays of Z and W , $t\bar{t}$, diboson production, γ +jet, etc.
- $\min \Delta\phi(\text{jet}, p_T^{\text{miss}}) > 0.4$: remove QCD multijet events.
- $|p_{T,\text{calo}}^{\text{miss}} - p_T^{\text{miss}}| < p_T^{\text{miss}}/2$: remove miscalibrated events.

As the tag variables each show some level of separation between signal and backgrounds, we can choose to either fit the distributions or use them to select events. To find the optimal choice, we fit each of the distributions in turn, and scan the other two observables. The details of this fit and the background estimation are described in Section 5.2. The metric is chosen to be the expected 95% CLs upper limit on $\mathcal{B}(H \rightarrow \chi\bar{\chi})$. Figure 5.14 shows the result of this optimization. The dijet mass is found to be the best distribution to fit, while requiring $\Delta\eta_{jj} > 1$ and $\Delta\phi_{jj} < 1.5$.

5.2 Background estimation

To estimate the combined m_{jj} spectra of the EW and QCD V +jet backgrounds, we employ a similar visible-to-invisible strategy as described in Section 4.2. In this case, the transfer factors \mathbf{T} are a function of m_{jj} . Control regions are defined using dilepton (single-lepton) selections to estimate the Z (W) contributions. Again, p_T^{miss} is replaced by U (Equation 4.6) to mimic the signal region selection. Because there are *two* components to estimate in each CR (QCD and EW), we slightly modify the likelihood. Adding only the $\mu\mu$ CR to constrain the Z +jet component for now:

$$\begin{aligned} \mathcal{L}(\mathbf{d} | \mu, \boldsymbol{\mu}_{\text{SR}}^{Z \rightarrow \nu\nu}, \boldsymbol{\theta}) = \prod_{i \in \text{bins}} & \left[\text{Pois} \left(d_i^{\text{SR}} \mid \mu S_i^{\text{SR}}(\boldsymbol{\theta}) + \mu_{\text{SR},i}^{Z \rightarrow \nu\nu} + \frac{\mu_{\text{SR},i}^{Z \rightarrow \nu\nu}}{T_{Z,i}^{\text{QE}}(\boldsymbol{\theta})} + B_i^{\text{SR}}(\boldsymbol{\theta}) \right) \right. \\ & \times \text{Pois} \left(d_i^{\mu\mu} \mid \frac{\mu_{\text{SR},i}^{Z \rightarrow \nu\nu}}{T_{Z,i}^{\mu\mu}(\boldsymbol{\theta})} + \frac{\mu_{\text{SR},i}^{Z \rightarrow \nu\nu}}{T_{Z,i}^{\mu\mu}(\boldsymbol{\theta}) T_{Z,i}^{\text{QE}}(\boldsymbol{\theta})} + B_i^{\mu\mu}(\boldsymbol{\theta}) \right) \left. \right] \\ & \times \prod_{j=0}^{n_\theta} p_j(\theta_j) \end{aligned} \quad (5.5)$$

While the notation largely follows that used in Equation 4.11, one additional term has been introduced. This is a *transfer factor* linking the QCD and EW components in the signal region, so that the only free parameter is $\boldsymbol{\mu}_{\text{SR}}^{Z \rightarrow \nu\nu}$:

$$T_{Z,i}^{\text{QE}} = \frac{N_i^{\text{SR}}(\text{QCD } Z \rightarrow \nu\nu)}{N_i^{\text{SR}}(\text{EW } Z \rightarrow \nu\nu)} \quad (5.6)$$

where as always, the yields N are predicted using MC. Kinematic distributions from the two dilepton CRs are shown in Figure 5.15.

In the region $m_{jj} > 2.5$ TeV, the statistical power of the dilepton regions is extremely limited. For this reason, and to estimate the W +jets contribution in the SR, we add two single-lepton CRs in analogy to what is done in Section 4.2. Figure 5.16 shows the level of agreement between the data and MC in these CRs. The likelihood

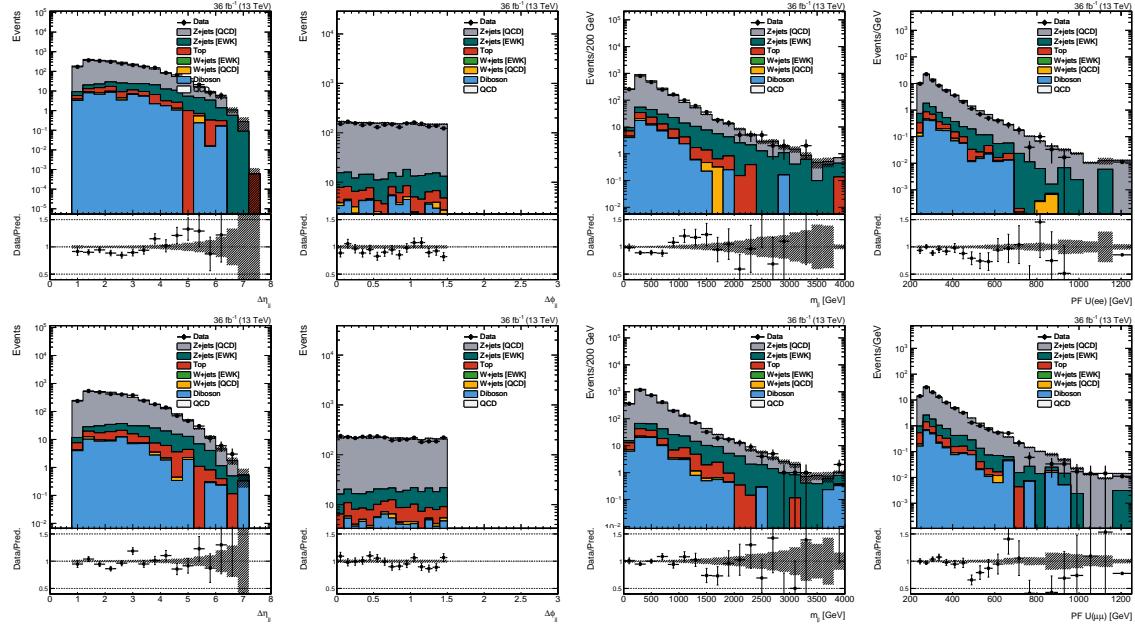


Figure 5.15: Dijet and recoil distributions in the dielectron (top) and dimuon (bottom) CRs.

is modified to include the constraints of the single-lepton CRs:

$$\begin{aligned} \mathcal{L}(\mathbf{d} | \mu, \boldsymbol{\mu}_{\text{SR}}^{Z \rightarrow \nu\nu}, \boldsymbol{\theta}) = & \\ \prod_{i \in \text{bins}} \left[\text{Pois} \left\{ d_i^{\text{SR}} \mid \mu S_i^{\text{SR}}(\boldsymbol{\theta}) + \left(1 + \frac{1}{T_{Z,i}^{\text{QE}}(\boldsymbol{\theta})} \right) \left(1 + \frac{1}{T_{Z/W,i}^{\text{SR}}(\boldsymbol{\theta})} \right) \mu_{\text{SR},i}^{Z \rightarrow \nu\nu} + B_i^{\text{SR}}(\boldsymbol{\theta}) \right\} \right. & \\ \times \prod_{X=\mu,e} \text{Pois} \left\{ d_i^X \mid \left(1 + \frac{1}{T_{Z,i}^{\text{QE}}(\boldsymbol{\theta})} \right) \frac{\mu_{\text{SR},i}^{Z \rightarrow \nu\nu}}{T_{W,i}^X(\boldsymbol{\theta}) T_{Z/W,i}^{\text{SR}}(\boldsymbol{\theta})} + B_i^X(\boldsymbol{\theta}) \right\} & \\ \times \prod_{X=\mu\mu,ee} \text{Pois} \left\{ d_i^X \mid \left(1 + \frac{1}{T_{Z,i}^{\text{QE}}(\boldsymbol{\theta})} \right) \frac{\mu_{\text{SR},i}^{Z \rightarrow \nu\nu}}{T_{Z,i}^X(\boldsymbol{\theta})} + B_i^X(\boldsymbol{\theta}) \right\} \right] \times \prod_{j=0}^{n_\theta} p_j(\theta_j) & \quad (5.7) \end{aligned}$$

To validate that the transfer factors are reasonably well-simulated (within the assigned uncertainties), Figure 5.17 uses the following ratios of CRs as proxies for transfer

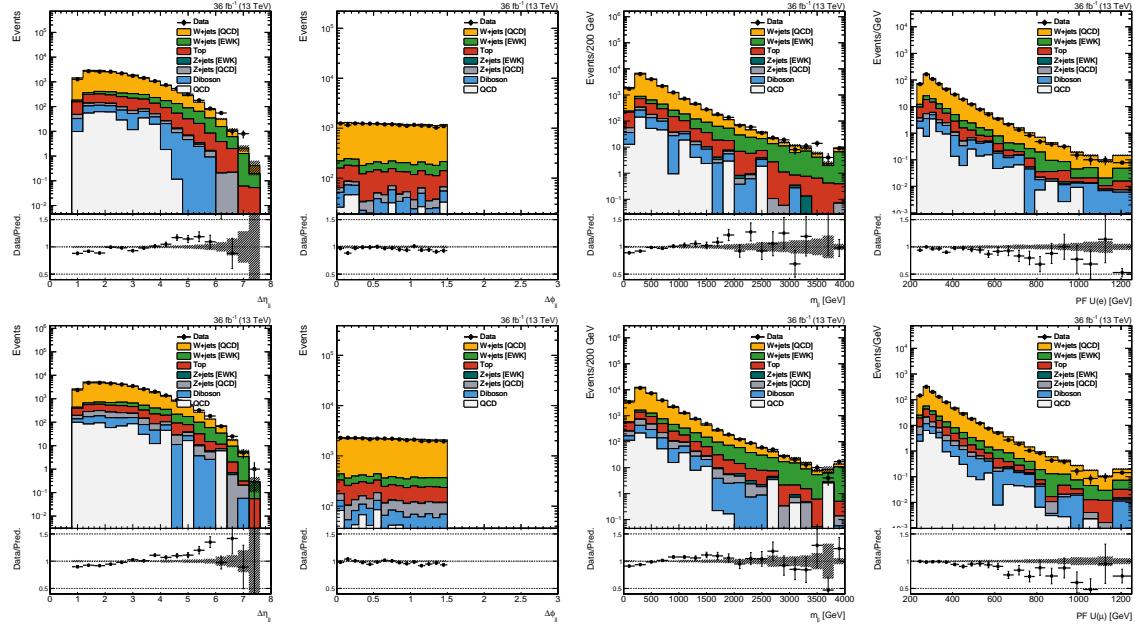


Figure 5.16: Dijet and recoil distributions in the single-electron (top) and single-muon (bottom) CRs.

factors:

$$\begin{aligned}
 \mathbf{T}_Z^{\mu\mu}, \mathbf{T}_Z^{ee} &\sim \frac{N_{\mu\mu}(Z \rightarrow \mu\mu)}{N_{ee}(Z \rightarrow ee)} \\
 \mathbf{T}_W^{\mu}, \mathbf{T}_W^e &\sim \frac{N_\mu(W \rightarrow \mu\nu)}{N_e(W \rightarrow e\nu)} \\
 \mathbf{T}_{Z/W}^{\text{SR}} &\sim \frac{N_\mu(W \rightarrow \mu\nu) + N_e(W \rightarrow e\nu)}{N_{\mu\mu}(Z \rightarrow \mu\mu) + N_{ee}(Z \rightarrow ee)}
 \end{aligned} \tag{5.8}$$

5.3 Results

The dijet mass distribution in data is fit in all signal and control regions, the results of which are shown in Figure 5.18. As no statistically significant excess over the Standard Model is observed, we translate the results into upper limits on the branching ratio of $H \rightarrow \chi\bar{\chi}$. As the signal hypothesis, both the VBF and gluon fusion (with 2 extra jets) Higgs production modes are considered; the latter contaminates the SR due to its relatively large cross section. After the signal region selection criteria, the

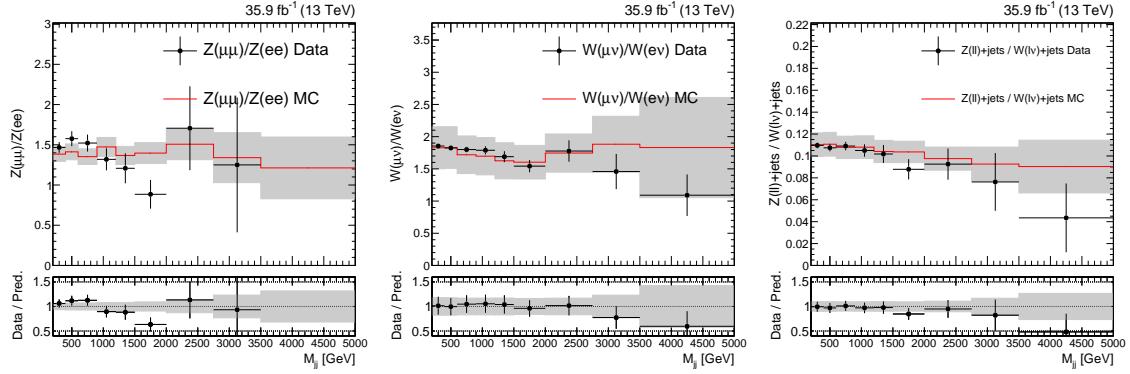


Figure 5.17: Validation of the VBF transfer factors using control region data. The transfer factor proxies are found to agree quite well with the data within the post-fit uncertainties

two modes contribute approximately equal yields. Assuming $m_H = 125$ GeV, the observed 95% CL upper limit is 0.33. Assuming a background-only hypothesis, the expected distribution of upper limits has median 0.33, with the 1 standard deviation band covering [0.18, 0.35]; the observation therefore represents an upwards fluctuation slightly under 1σ .

We further scan m_H and set upper limits on $\sigma(qq \rightarrow qqH)\mathcal{B}(H \rightarrow \chi\bar{\chi})$. The upper limits are shown in Figure 5.19, and the observed (expected) limits exclude $m_H < 540$ GeV (635 GeV) assuming a branching ratio of 100%.

As described in the beginning of this chapter, each Higgs production mode corresponds to a potential invisible Higgs search channel. While VBF is the most sensitive, the total sensitivity can be improved by statistically combining all channels. Other results from CMS cover searches for associated production of a Higgs boson, either with a leptonically-decaying Z boson ([?]) or a hadronically-decaying weak boson ([?]); and for gluon fusion production, with at least one jet originating from the initial state or heavy quark loop ([?]). The details of these searches are left to the referenced literature, but a summary of their results is provided in Figure 5.20. When statistically combining the results, most experimental nuisances are treated as correlated between the searches, with the exception of the VBF jet energy scale dependence. This is because the VBF category selects forward jets, whereas other searches generally probe

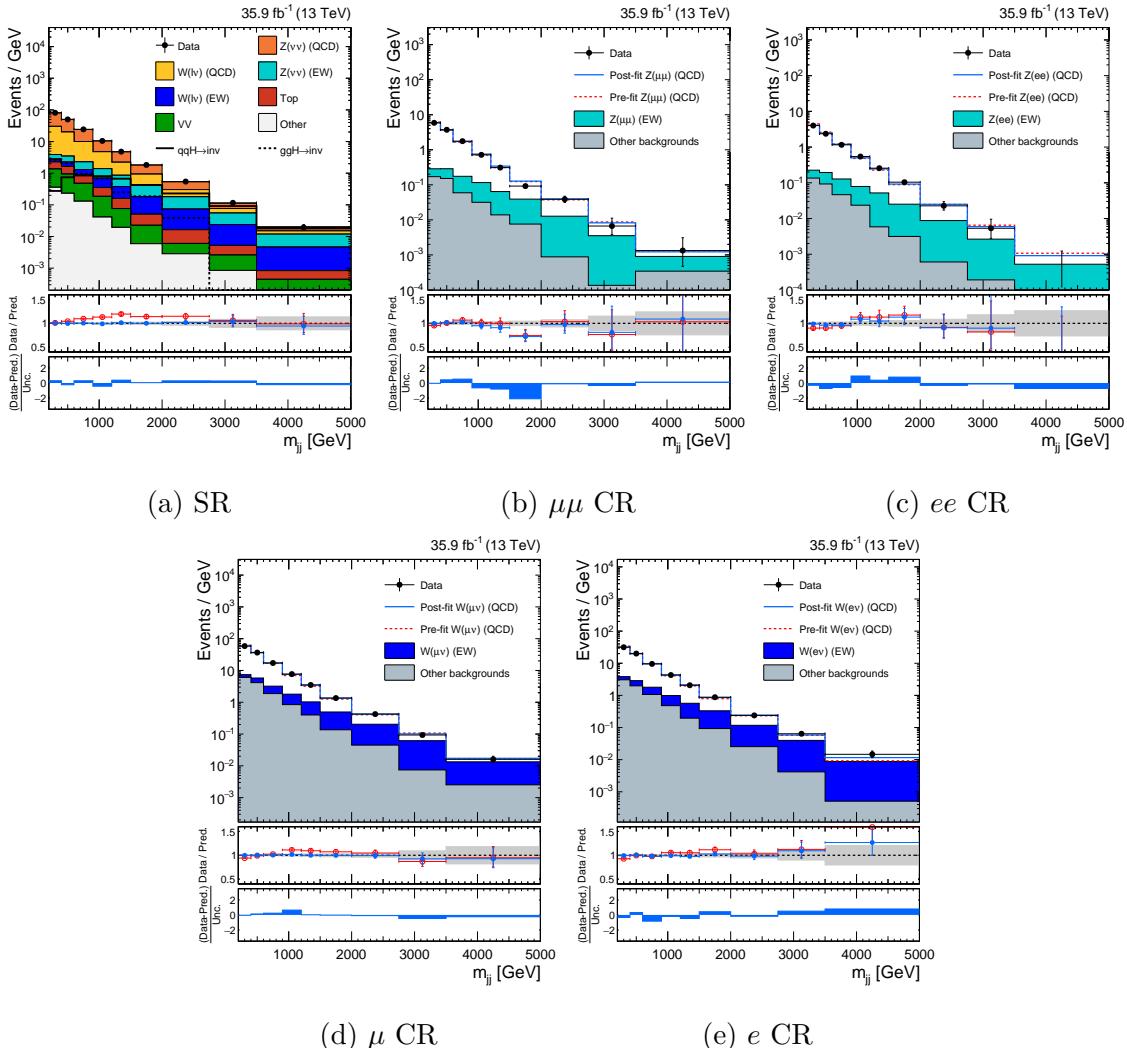


Figure 5.18: Post-fit m_{jj} distributions in the various signal and control regions. The uncertainties (gray bands) and bin pulls (blue bands) are defined by varying the nuisances by one standard deviation around the maximum likelihood estimate.

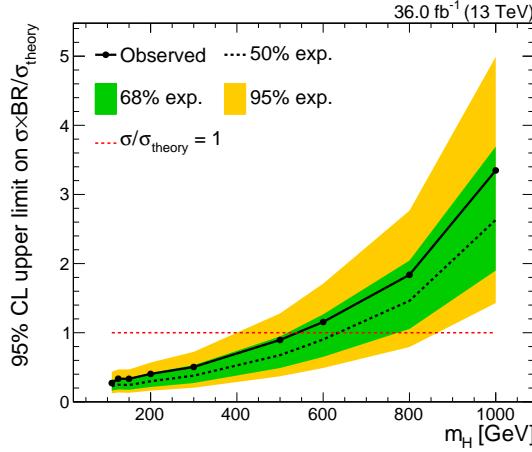


Figure 5.19: Upper limits on $\sigma \times \mathcal{B}/\sigma_{\text{theory}}$ as a function of m_H , where σ refers to the total production cross section of the Higgs boson with mass m_H . If one assumes that $\sigma = \sigma_{\text{theory}}$, then the upper limits can be interpreted directly as constraints on $\mathcal{B}(H \rightarrow \chi\bar{\chi})$.

central jets. Theoretical nuisances (e.g. those affecting W/Z or ZZ/WZ ratios) are left uncorrelated between all searches. The combined result constrains $\mathcal{B}(H \rightarrow \chi\bar{\chi})$ to be less than 0.26 at 95% CL, which approximately corresponds to a 1 standard deviation fluctuation upward relative to the median expected limit of 0.20.

Higgs-mediated DM can also be probed by direct detection (DD) experiments. We interpret the combined 90% CL upper limit (for consistency with DD conventions) on $\mathcal{B}(H \rightarrow \chi\bar{\chi})$ as an upper limit on the spin-independent cross section of DM-nucleon scattering. First, we convert the branching ratio into a partial width:

$$\Gamma_{H \rightarrow \chi\bar{\chi}} = \frac{\mathcal{B}(H \rightarrow \chi\bar{\chi}) \cdot \Gamma_{\text{SM}}}{1 - \mathcal{B}(H \rightarrow \chi\bar{\chi})}, \text{ where } \Gamma_{\text{SM}} = 4 \text{ GeV} \quad (5.9)$$

Then, using the results described in Reference [?], $\Gamma_{H \rightarrow \chi\bar{\chi}}$ is translated into $\sigma_{\chi N}^{\text{SI}}(m_\chi)$, where the only free parameter is the DM mass. Figure 5.21 compares the CMS exclusions with those from DD experiments [?]. Also shown for comparison is the *neutrino floor*, which is the cross section of coherent scattering of solar neutrinos, and a limiting factor for DD experiments. At low m_χ , CMS is able to significantly extend the DD constraints, reaching well below the neutrino floor.

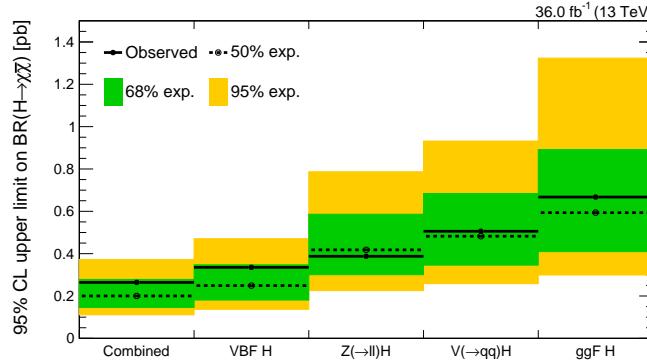


Figure 5.20: Upper limits on $\mathcal{B}(H \rightarrow \chi\bar{\chi})$ after statistically combining all of the CMS searches for $H \rightarrow \chi\bar{\chi}$ conducted on 36 fb^{-1} of data collected in 2016. For comparison, the upper limits of each of the individual categories are also shown.

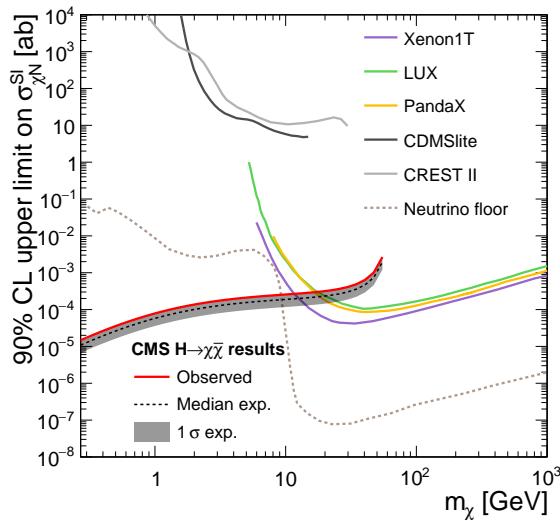


Figure 5.21: Upper limits on $\sigma_{\chi N}^{\text{SI}}$ as a function of m_χ . Shown are the combined results from the CMS invisible Higgs searches, as well as various direct detection experiments.

Appendix A

The CL_s Method

In the absence of an excess compatible with a signal hypothesis, we define 95% confidence level (CL) exclusion limits on model parameters such as cross sections and particle masses. The method used here was first used for the combination of CMS and ATLAS results for the Higgs search [102] and is based on earlier approaches from many other particle physics experiments [103, 104]. Given an analysis-specific likelihood \mathcal{L} , we define the test statistic as:

$$\tilde{q}_\mu^{\text{obs}} = -2 \ln \frac{\mathcal{L}(\mathbf{d}_{\text{obs}} | \mu, \hat{\boldsymbol{\theta}}_\mu)}{\mathcal{L}(\mathbf{d}_{\text{obs}} | \hat{\mu}, \hat{\boldsymbol{\theta}})} \quad (\text{A.1})$$

where \mathbf{d}_{obs} is a vector of observations in data; μ is a free parameter corresponding to the signal strength; and the nuisance parameter estimates are:

$$\begin{aligned} \hat{\boldsymbol{\theta}}_\mu &= \arg \max_{\boldsymbol{\theta}} \mathcal{L}(\mathbf{d}_{\text{obs}} | \mu, \boldsymbol{\theta}) \\ \hat{\mu}, \hat{\boldsymbol{\theta}}_\mu &= \arg \max_{\boldsymbol{\theta}} \mathcal{L}(\mathbf{d}_{\text{obs}} | \mu, \boldsymbol{\theta}) \end{aligned} \quad (\text{A.2})$$

We further require that $0 \leq \hat{\mu} \leq \mu$, as the goal of \tilde{q}_μ is to measure the compatibility of the data with a non-negative signal strength smaller than the hypothesis μ .

To estimate the significance of the measured $\tilde{q}_\mu^{\text{obs}}$, we generate *pseudo-data* by

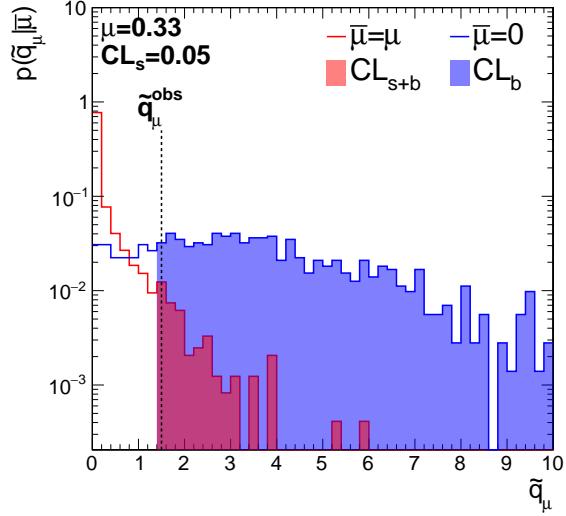


Figure A.1: Distribution of $p(\tilde{q}_\mu | \mu)$ and $p(\tilde{q}_\mu | 0)$ at $\mu = 0.33$. This value is chosen because it gives $CL_s(\mu) = 0.05$, corresponding to the 95% upper limit. The likelihood in this case is the VBF $H \rightarrow \chi\bar{\chi}$ analysis discussed in Chapter 5.

using the likelihood as a generative model. The pseudo-data follows the distribution:

$$\mathbf{d}_{\bar{\mu}} \sim \mathcal{L}(\mathbf{d} | \bar{\mu}, \hat{\boldsymbol{\theta}}_{\bar{\mu}}) \quad (\text{A.3})$$

This leads to a natural definition of the conditional test statistic:

$$\tilde{q}_\mu | \bar{\mu}, \hat{\boldsymbol{\theta}}_{\bar{\mu}} = -2 \ln \frac{\mathcal{L}(\mathbf{d}_{\bar{\mu}} | \mu, \hat{\boldsymbol{\theta}}_\mu)}{\mathcal{L}(\mathbf{d}_{\bar{\mu}} | \hat{\mu}, \hat{\boldsymbol{\theta}})} \quad (\text{A.4})$$

With a sufficiently large pseudo-dataset $\{\mathbf{d}_{\bar{\mu}}\}$, we estimate the probability distribution function $p(\tilde{q}_\mu | \bar{\mu}, \hat{\boldsymbol{\theta}}_{\bar{\mu}})$. Finally, we define the CL_s statistic as:

$$CL_s(\mu) = \frac{CL_{s+b}(\mu)}{CL_b(\mu)}, \text{ where:} \\ CL_{s+b}(\mu) = \int_{\tilde{q}_\mu^{\text{obs}}}^{\infty} d\tilde{q}_\mu p(\tilde{q}_\mu | \mu, \hat{\boldsymbol{\theta}}_\mu), \quad CL_b(\mu) = \int_{\tilde{q}_\mu^{\text{obs}}}^{\infty} d\tilde{q}_\mu p(\tilde{q}_\mu | 0, \hat{\boldsymbol{\theta}}_0) \quad (\text{A.5})$$

These PDFs and the corresponding integrals are shown in Figure A.1.

If $CL_s(\mu) = 1 - \alpha$, then we say the signal with strength μ has been excluded at α confidence level (μ_α^{obs}). One can also identify CL_{s+b} as a p -value, but it does not take

into account the sensitivity of the data to the differences between the background-only and signal+background models. By construction, $\text{CL}_b \leq 1$, which acts as a penalty as $p(\tilde{q}_\mu|0)$ approaches $p(\tilde{q}_\mu|\mu)$.

To estimate the significance of μ_α^{obs} , the above procedure is repeated many times using more pseudo-datasets. Each pseudo-dataset is generated using the maximum likelihood estimate for the background-only hypothesis $\mathcal{L}(\mathbf{d}|0, \hat{\boldsymbol{\theta}}_0)$. The distribution of μ_α assuming no signal can be compared to the observed value. A measure of the significance is $(\mu_\alpha^{\text{obs}} - \langle \mu_\alpha \rangle)/\sigma(\mu_\alpha)$. These numbers are usually quoted as the expected (median) μ_α , with one and two standard deviation coverage bands.

The computation of the CL_s test statistic with many toys can be prohibitive. If the determination of the PDFs p requires N toys, and the evaluation of the distribution of μ_α requires M toys, then at least $N + M$ maximizations of the likelihood are needed. The asymptotic method

Bibliography

- [1] D. J. Gross and F. Wilczek, “Ultraviolet Behavior of Nonabelian Gauge Theories,” *Phys. Rev. Lett.*, vol. 30, pp. 1343–1346, 1973. [,271(1973)].
- [2] H. D. Politzer, “Reliable Perturbative Results for Strong Interactions?,” *Phys. Rev. Lett.*, vol. 30, pp. 1346–1349, 1973. [,274(1973)].
- [3] S. Weinberg, “A Model of Leptons,” *Phys. Rev. Lett.*, vol. 19, pp. 1264–1266, 1967.
- [4] A. Salam, “Weak and Electromagnetic Interactions,” *Conf. Proc.*, vol. C680519, pp. 367–377, 1968.
- [5] S. L. Glashow, “Partial Symmetries of Weak Interactions,” *Nucl. Phys.*, vol. 22, pp. 579–588, 1961.
- [6] B. Graner, Y. Chen, E. G. Lindahl, and B. R. Heckel, “Reduced Limit on the Permanent Electric Dipole Moment of Hg^{199} ;,” *Phys. Rev. Lett.*, vol. 116, p. 161601, Apr. 2016.
- [7] “Review of particle physics,” *Phys. Rev. D*, vol. 98, p. 030001, Aug 2018.
- [8] D. J. Gross and F. Wilczek, “Asymptotically Free Gauge Theories - I,” *Phys. Rev.*, vol. D8, pp. 3633–3652, 1973.
- [9] S. Weinberg, “New approach to the renormalization group,” *Phys. Rev.*, vol. D8, pp. 3497–3509, 1973.
- [10] R. Hagedorn, “Statistical thermodynamics of strong interactions at high-energies,” *Nuovo Cim. Suppl.*, vol. 3, pp. 147–186, 1965.
- [11] M. Gell-Mann, “A Schematic Model of Baryons and Mesons,” *Phys. Lett.*, vol. 8, pp. 214–215, 1964.
- [12] F. Englert and R. Brout, “Broken Symmetry and the Mass of Gauge Vector Mesons,” *Phys. Rev. Lett.*, vol. 13, pp. 321–323, 1964. [,157(1964)].
- [13] P. W. Higgs, “Broken symmetries, massless particles and gauge fields,” *Phys. Lett.*, vol. 12, pp. 132–133, 1964.

- [14] P. W. Higgs, “Broken Symmetries and the Masses of Gauge Bosons,” *Phys. Rev. Lett.*, vol. 13, pp. 508–509, 1964. [,160(1964)].
- [15] P. W. Higgs, “Spontaneous Symmetry Breakdown without Massless Bosons,” *Phys. Rev.*, vol. 145, pp. 1156–1163, 1966.
- [16] G. S. Guralnik, C. R. Hagen, and T. W. B. Kibble, “Global Conservation Laws and Massless Particles,” *Phys. Rev. Lett.*, vol. 13, pp. 585–587, 1964. [,162(1964)].
- [17] T. W. B. Kibble, “Symmetry breaking in nonAbelian gauge theories,” *Phys. Rev.*, vol. 155, pp. 1554–1561, 1967. [,165(1967)].
- [18] Y. Nambu, “Quasi-Particles and Gauge Invariance in the Theory of Superconductivity,” *Physical Review*, vol. 117, pp. 648–663, Feb. 1960.
- [19] J. Goldstone, “Field theories with Superconductor solutions,” *Il Nuovo Cimento*, vol. 19, pp. 154–164, Jan. 1961.
- [20] G. Aad *et al.*, “Combined Measurement of the Higgs Boson Mass in pp Collisions at $\sqrt{s} = 7$ and 8 TeV with the ATLAS and CMS Experiments,” *Phys. Rev. Lett.*, vol. 114, p. 191803, 2015.
- [21] N. Cabibbo, “Unitary Symmetry and Leptonic Decays,” *Physical Review Letters*, vol. 10, pp. 531–533, June 1963.
- [22] M. Kobayashi and T. Maskawa, “CP-Violation in the Renormalizable Theory of Weak Interaction,” *Progress of Theoretical Physics*, vol. 49, pp. 652–657, Feb. 1973.
- [23] Z. Maki, M. Nakagawa, and S. Sakata, “Remarks on the Unified Model of Elementary Particles,” *Progress of Theoretical Physics*, vol. 28, pp. 870–880, Nov. 1962.
- [24] Y. Fukuda *et al.*, “Evidence for Oscillation of Atmospheric Neutrinos,” *Physical Review Letters*, vol. 81, pp. 1562–1567, Aug. 1998.
- [25] D. Hanneke, S. Fogwell, and G. Gabrielse, “New Measurement of the Electron Magnetic Moment and the Fine Structure Constant,” *Phys. Rev. Lett.*, vol. 100, p. 120801, Mar. 2008.
- [26] V. Khachatryan *et al.*, “Precise determination of the mass of the Higgs boson and tests of compatibility of its couplings with the standard model predictions using proton collisions at 7 and 8 TeV,” *Eur. Phys. J.*, vol. C75, no. 5, p. 212, 2015.
- [27] A. Einstein, “Die Feldgleichungen der Gravitation. (German) [The field equations of gravitation],” *Ständiger Beobachter der Preussischen Akademie der Wissenschaften, Part 2*, pp. 844–847, 1915.

- [28] P. J. E. Peebles and B. Ratra, “The Cosmological constant and dark energy,” *Rev. Mod. Phys.*, vol. 75, pp. 559–606, 2003. [,592(2002)].
- [29] G. Bertone, D. Hooper, and J. Silk, “Particle dark matter: evidence, candidates and constraints,” *Physics Reports*, vol. 405, p. 279, 2005.
- [30] J. L. Feng, “Dark matter candidates from particle physics and methods of detection,” *Ann. Rev. Astron. Astrophys.*, vol. 48, p. 495, 2010.
- [31] T. A. Porter, R. P. Johnson, and P. W. Graham, “Dark matter searches with astroparticle data,” *Ann. Rev. Astron. Astrophys.*, vol. 49, p. 15, 2011.
- [32] K. G. Begeman, A. H. Broeils, and R. H. Sanders, “Extended rotation curves of spiral galaxies - Dark haloes and modified dynamics,” *Monthly Notices of the Royal Astronomical Society*, vol. 249, pp. 523–537, Apr. 1991.
- [33] N. Aghanim *et al.*, “Planck 2018 results. VI. Cosmological parameters,” 2018.
- [34] L. Accardo *et al.*, “High Statistics Measurement of the Positron Fraction in Primary Cosmic Rays of 0.5–500 GeV with the Alpha Magnetic Spectrometer on the International Space Station,” *Phys. Rev. Lett.*, vol. 113, p. 121101, 2014.
- [35] M. Ackermann *et al.*, “Searching for Dark Matter Annihilation from Milky Way Dwarf Spheroidal Galaxies with Six Years of Fermi Large Area Telescope Data,” *Phys. Rev. Lett.*, vol. 115, no. 23, p. 231301, 2015.
- [36] A. M. Sirunyan *et al.*, “Search for new physics in final states with an energetic jet or a hadronically decaying W or Z boson and transverse momentum imbalance at $\sqrt{s} = 13$ TeV,” *Phys. Rev.*, vol. D97, no. 9, p. 092005, 2018.
- [37] D. S. Akerib *et al.*, “Results from a search for dark matter in the complete LUX exposure,” *Phys. Rev. Lett.*, vol. 118, no. 2, p. 021303, 2017.
- [38] M. Perelstein, “Introduction to Collider Physics,” pp. 421–486, 2011.
- [39] J. C. Collins, D. E. Soper, and G. F. Sterman, “Factorization of Hard Processes in QCD,” *Adv. Ser. Direct. High Energy Phys.*, vol. 5, pp. 1–91, 1989.
- [40] Y. L. Dokshitzer, “Calculation of the Structure Functions for Deep Inelastic Scattering and e+ e- Annihilation by Perturbation Theory in Quantum Chromodynamics.,” *Sov. Phys. JETP*, vol. 46, pp. 641–653, 1977. [Zh. Eksp. Teor. Fiz.73,1216(1977)].
- [41] G. Altarelli and G. Parisi, “Asymptotic Freedom in Parton Language,” *Nucl. Phys.*, vol. B126, pp. 298–318, 1977.
- [42] V. N. Gribov and L. N. Lipatov, “Deep inelastic e p scattering in perturbation theory,” *Sov. J. Nucl. Phys.*, vol. 15, pp. 438–450, 1972. [Yad. Fiz.15,781(1972)].

- [43] R. D. Ball *et al.*, “Parton distributions for the LHC Run II,” *JHEP*, vol. 04, p. 040, 2015.
- [44] J. Alwall, R. Frederix, S. Frixione, V. Hirschi, F. Maltoni, O. Mattelaer, H. S. Shao, T. Stelzer, P. Torrielli, and M. Zaro, “The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations,” *JHEP*, vol. 07, p. 079, 2014.
- [45] R. Frederix and S. Frixione, “Merging meets matching in MC@NLO,” *JHEP*, vol. 12, p. 061, 2012.
- [46] C. Oleari, “The POWHEG BOX,” *Nuclear Physics B Proceedings Supplements*, vol. 205, pp. 36–41, Aug. 2010.
- [47] T. Sjöstrand, S. Ask, J. R. Christiansen, R. Corke, N. Desai, P. Ilten, S. Mrenna, S. Prestel, C. O. Rasmussen, and P. Z. Skands, “An Introduction to PYTHIA 8.2,” *Comput. Phys. Commun.*, vol. 191, pp. 159–177, 2015.
- [48] B. Andersson, G. Gustafson, G. Ingelman, and T. Sjostrand, “Parton Fragmentation and String Dynamics,” *Phys. Rept.*, vol. 97, pp. 31–145, 1983.
- [49] L. Evans and P. Bryant, “LHC Machine,” *JINST*, vol. 3, p. S08001, 2008.
- [50] C. De Melis, “The CERN accelerator complex. Complexe des accélérateurs du CERN,” Jan 2016. General Photo.
- [51] M. Benedikt, P. Collier, V. Mertens, J. Poole, and K. Schindl, “LHC Design Report. 3. The LHC injector chain,” 2004.
- [52] C. Collaboration. https://twiki.cern.ch/twiki/bin/view/CMSPublic/SWGuideGlobalHLT#Trigger_development_for_Run_2.
- [53] S. Chatrchyan *et al.*, “The CMS experiment at the CERN LHC,” *Journal of Instrumentation*, vol. 3, no. 08, p. S08004, 2008.
- [54] G. Aad *et al.*, “The atlas experiment at the cern large hadron collider,” *Journal of Instrumentation*, vol. 3, no. 08, p. S08003, 2008.
- [55] S. Chatrchyan *et al.*, “Description and performance of track and primary-vertex reconstruction with the CMS tracker,” *JINST*, vol. 9, no. 10, p. P10009, 2014.
- [56] P. Billoir and S. Qian, “Simultaneous pattern recognition and track fitting by the kalman filtering method,” *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, vol. 294, no. 1, pp. 219 – 228, 1990.
- [57] R. Fruehwirth, “Application of Kalman filtering to track and vertex fitting,” *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, vol. 262, no. 2, pp. 444 – 450, 1987.

- [58] K. Rose, “Deterministic annealing for clustering, compression, classification, regression, and related optimization problems,” *Proceedings of the IEEE*, vol. 86, pp. 2210–2239, Nov 1998.
- [59] R. Fruehwirth, W. Waltenberger, and P. Vanlaer, “Adaptive Vertex Fitting,” Tech. Rep. CMS-NOTE-2007-008, CERN, Geneva, Mar 2007.
- [60] A. M. Sirunyan *et al.*, “Identification of heavy-flavour jets with the CMS detector in pp collisions at 13 TeV,” *JINST*, vol. 13, no. 05, p. P05011, 2018.
- [61] A. Benaglia *et al.*, “The CMS ECAL performance with examples,” *Journal of Instrumentation*, vol. 9, no. 02, p. C02008, 2014.
- [62] V. Khachatryan *et al.*, “Performance of Photon Reconstruction and Identification with the CMS Detector in Proton-Proton Collisions at $\sqrt{s} = 8$ TeV,” *JINST*, vol. 10, no. 08, p. P08010, 2015.
- [63] C. Biino, “The CMS Electromagnetic Calorimeter: overview, lessons learned during Run 1 and future projections,” *Journal of Physics: Conference Series*, vol. 587, no. 1, p. 012001, 2015.
- [64] *The CMS hadron calorimeter project: Technical Design Report*. Technical Design Report CMS, Geneva: CERN, 1997.
- [65] A. M. Sirunyan *et al.*, “Particle-flow reconstruction and global event description with the CMS detector,” *JINST*, vol. 12, no. 10, p. P10003, 2017.
- [66] C. Collaboration, “Performance of the CMS hadron calorimeter with cosmic ray muons and LHC beam data,” *Journal of Instrumentation*, vol. 5, p. T03012, Mar. 2010.
- [67] A. M. Sirunyan *et al.*, “Performance of the CMS muon detector and muon reconstruction with proton-proton collisions at $\sqrt{s} = 13$ TeV,” *JINST*, vol. 13, no. 06, p. P06015, 2018.
- [68] V. Khachatryan *et al.*, “The CMS trigger system,” *JINST*, vol. 12, no. 01, p. P01020, 2017.
- [69] J. M. Campbell, J. W. Huston, and W. J. Stirling, “Hard Interactions of Quarks and Gluons: A Primer for LHC Physics,” *Rept. Prog. Phys.*, vol. 70, p. 89, 2007.
- [70] G. L. Bayatian *et al.*, “CMS Physics Technical Design Report, Volume I: Detector Performance and Software,” 2006.
- [71] S. Chatrchyan *et al.*, “Performance of CMS muon reconstruction in pp collision events at $\sqrt{s} = 7$ TeV,” *JINST*, vol. 7, p. P10002, 2012.
- [72] W. Adam, R. Frühwirth, A. Strandlie, and T. Todorov, “Reconstruction of electrons with the gaussian-sum filter in the cms tracker at the lhc,” *Journal of Physics G: Nuclear and Particle Physics*, vol. 31, no. 9, p. N9, 2005.

- [73] M. Cacciari, G. P. Salam, and G. Soyez, “The anti- k_t jet clustering algorithm,” *JHEP*, vol. 04, p. 063, 2008.
- [74] S. D. Ellis and D. E. Soper, “Successive combination jet algorithm for hadron collisions,” *Phys. Rev.*, vol. D48, pp. 3160–3166, 1993.
- [75] Y. L. Dokshitzer, G. D. Leder, S. Moretti, and B. R. Webber, “Better jet clustering algorithms,” *JHEP*, vol. 08, p. 001, 1997.
- [76] “FastJet User Manual,” *Eur. Phys. J.*, vol. C72, p. 1896, 2012.
- [77] V. Khachatryan *et al.*, “Jet energy scale and resolution in the CMS experiment in pp collisions at 8 TeV,” *JINST*, vol. 12, no. 02, p. P02014, 2017.
- [78] V. Khachatryan *et al.*, “Reconstruction and identification of τ lepton decays to hadrons and ν_τ at CMS,” *JINST*, vol. 11, no. 01, p. P01019, 2016.
- [79] D. Bertolini, P. Harris, M. Low, and N. Tran, “Pileup Per Particle Identification,” *JHEP*, vol. 10, p. 059, 2014.
- [80] “CMS Offline Software.” <https://github.com/cms-sw/cmssw>.
- [81] S. Agostinelli *et al.*, “GEANT4: A Simulation toolkit,” *Nucl. Instrum. Meth.*, vol. A506, pp. 250–303, 2003.
- [82] J. Allison *et al.*, “Geant4 developments and applications,” *IEEE Trans. Nucl. Sci.*, vol. 53, p. 270, 2006.
- [83] J. Shelton, “Jet Substructure,” pp. 303–340, 2013.
- [84] A. J. Larkoski, S. Marzani, G. Soyez, and J. Thaler, “Soft Drop,” *JHEP*, vol. 05, p. 146, 2014.
- [85] J. Thaler and K. Van Tilburg, “Identifying Boosted Objects with N-subjettiness,” *JHEP*, vol. 03, p. 015, 2011.
- [86] G. Kasieczka, T. Plehn, T. Schell, T. Strebler, and G. P. Salam, “Resonance Searches with an Updated Top Tagger,” *JHEP*, vol. 06, p. 203, 2015.
- [87] I. Moult, L. Necib, and J. Thaler, “New Angles on Energy Correlation Functions,” *JHEP*, vol. 12, p. 153, 2016.
- [88] J. Therhaag, “TMVA Toolkit for multivariate data analysis in ROOT,” *PoS*, vol. ICHEP2010, p. 510, 2010.
- [89] J. Schmidhuber, “Deep Learning in Neural Networks: An Overview,” *arXiv e-prints*, p. arXiv:1404.7828, Apr. 2014.
- [90] S. Kallweit, J. M. Lindert, P. Maierhöfer, S. Pozzorini, and M. Schönherr, “NLO electroweak automation and precise predictions for W+multijet production at the LHC,” *JHEP*, vol. 04, p. 012, 2015.

- [91] S. Kallweit, J. M. Lindert, S. Pozzorini, M. Schönherr, and P. Maierhöfer, “NLO QCD+EW automation and precise predictions for V+multijet production,” in *50th Rencontres de Moriond on QCD and High Energy Interactions La Thuile, Italy, March 21-28, 2015*, 2015.
- [92] J. H. Kuhn, A. Kulesza, S. Pozzorini, and M. Schulze, “Electroweak corrections to hadronic photon production at large transverse momenta,” *JHEP*, vol. 03, p. 059, 2006.
- [93] A. Denner, S. Dittmaier, T. Kasprzik, and A. Mück, “Electroweak corrections to W + jet hadroproduction including leptonic W-boson decays,” *JHEP*, vol. 08, p. 075, 2009.
- [94] A. Denner, S. Dittmaier, T. Kasprzik, and A. Mück, “Electroweak corrections to dilepton + jet production at hadron colliders,” *JHEP*, vol. 06, p. 069, 2011.
- [95] A. Denner, S. Dittmaier, T. Kasprzik, and A. Mück, “Electroweak corrections to monojet production at the LHC,” *Eur. Phys. J. C*, vol. 73, p. 2297, 2013.
- [96] J. H. Kuhn, A. Kulesza, S. Pozzorini, and M. Schulze, “Logarithmic electroweak corrections to hadronic Z+1 jet production at large transverse momentum,” *Phys. Lett. B*, vol. 609, p. 277, 2005.
- [97] J. H. Kuhn, A. Kulesza, S. Pozzorini, and M. Schulze, “One-loop weak corrections to hadronic production of Z bosons at large transverse momenta,” *Nucl. Phys. B*, vol. 727, p. 368, 2005.
- [98] J. H. Kuhn, A. Kulesza, S. Pozzorini, and M. Schulze, “Electroweak corrections to hadronic production of W bosons at large transverse momenta,” *Nucl. Phys. B*, vol. 797, p. 27, 2008.
- [99] L. Moneta, K. Belasco, K. S. Cranmer, S. Kreiss, A. Lazzaro, D. Piparo, G. Schott, W. Verkerke, and M. Wolf, “The RooStats Project,” *PoS*, vol. ACAT2010, p. 057, 2010.
- [100] D. A. *et al*, “Dark Matter Benchmark Models for Early LHC Run-2 Searches: Report of the ATLAS/CMS Dark Matter Forum,” 2015.
- [101] D. de Florian *et al*, “Handbook of LHC Higgs Cross Sections: 4. Deciphering the Nature of the Higgs Sector,” Tech. Rep. FERMILAB-FN-1025-T, Oct 2016. 869 pages, 295 figures, 248 tables and 1645 citations. Working Group web page: <https://twiki.cern.ch/twiki/bin/view/LHCPhysics/LHCHXSWG>.
- [102] “Procedure for the LHC Higgs boson search combination in Summer 2011,” Aug 2011.
- [103] T. Junk, “Confidence level computation for combining searches with small statistics,” *Nucl. Instrum. Meth.*, vol. A434, pp. 435–443, 1999.

- [104] A. L. Read, “Presentation of search results: The CL(s) technique,” *J. Phys.*, vol. G28, pp. 2693–2704, 2002. [,11(2002)].