Jets, Substructure, and Searching for Dark Matter 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

August 2018

© Massachusetts Institute of Technology 2018. All rights reserved.

Author .....
Department of Physics
August 32, 2018

Certified by .....
Christoph M. E. Paus
Professor of Physics
Thesis Supervisor

Accepted by .....
Somebody
Chairman, Department Committee on Graduate Theses

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

by

Siddharth Madhavan Narayanan

Submitted to the Department of Physics
on August 32, 2018, 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 (DM). Many theories propose and experiments test the hypothesis that DM may have a particle physics origin, but this remains unproven. One such experiment is the Compact Muon Solenoid (CMS) at the Large Hadron Collider (LHC). If DM couples to particles present in protons, it is possible that DM is produced in collisions at the LHC. Because DM is effectively invisible to CMS, we must look for collisions in which DM is produced in association with one or more Standard Model (SM) particles. This thesis describes three different scenarios for the SM particle hypothesis: a single top quark, a single Higgs boson, or two light quarks. All three cases result in complicated detector signatures due to the hadronization of final-state quarks. Improved jet substructure techniques to identify these unique signatures are presented. Since the observed data is consistent with SM backgrounds in all three searches, we translate this result into the most stringent constraints to date on the relevant beyond-SM models.

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

Acknowledgments

Thank you to Professor Markus Klute for nice barbecues.

Contents

1 The Standard Model and Beyond **9**

1.0.1 The Standard Model of Particle Physics 9

2 The CMS experiment at the LHC **11**

2.1 The Large Hadron Collider 12

2.2 The Compact Muon Solenoid 12

2.2.1 Detector systems 12

2.2.2 Particle reconstruction algorithms 12

2.3 Simulation of collisions 12

2.3.1 Physics simulation 12

2.3.2 Detector simulation 12

3 Hadronic Resonance Identification **13**

3.1 Reconstruction 13

3.2 Identification 16

3.2.1 Substructure 18

3.2.2 Heavy flavor identification 20

3.3 Data validation 20

4 The Search for $t + p_{\mathbf{T}}^{\mathbf{miss}}$ **21**

4.1 Signal selection 25

4.1.1 Online trigger selection 26

4.1.2 Offline signal selection 27

4.2 Background estimation 30
4.2.1 Visible final states to constrain invisible final states 30
4.2.2 Theoretically-limited extrapolations 44
4.3 Results 50
4.3.1 Constraints on mono-top models 54

5 The Search for Invisible Decays of the Higgs Boson 59
5.1 Signal selection 60
5.1.1 Online trigger selection 60
5.1.2 EW and QCD production of electroweak bosons 68
5.1.3 Sensitivity optimization 71
5.2 Background estimation 72
5.3 Results 74

Chapter 1

The Standard Model and Beyond

1.0.1 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 (the quarks and leptons), respecting

Chapter 2

The CMS experiment at the LHC

2.1 The Large Hadron Collider

2.2 The Compact Muon Solenoid

2.2.1 Detector systems

Silicon tracker

Electromagnetic calorimetry

Hadronic calorimetry

Muon chambers

Online trigger system

2.2.2 Particle reconstruction algorithms

Tracks and vertices

Electrons and photons

Jets

Muons

Particle flow algorithm

Missing momentum

2.3 Simulation of collisions

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 ($W, Z \to q\bar{q}'$), the Higgs boson (typically $H \to b\bar{b}$), and the top quark ($t \to bW(\to q\bar{q}')$). These quarks hadronize into jets (described in Chapter 1), which are typically reconstructed at the LHC using the anti-k_{T} algorithm (described in Chapter 2). 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 jets. In preparation for Chapter 4, we will take the top quark as a concrete example. In principle, 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 approximate angular separation between the quarks from a heavy resonance decay is[???]:

$$\Delta R \sim \frac{2M}{p_{\mathrm{T}}} \tag{3.1}$$

where M is the resonance mass and p_{T} is the resonance transverse momentum. Setting $M = m_t$ and $\Delta R = 1.2$ (i.e. the radius at which three $R = 0.4$ jets start to

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}.

overlap), we extract a "merging scale" of 300 GeV. This can be verified by checking the distribution of the "decay radius" in top quark simulation. Here, we define decay radius as:

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

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}, where we restrict the resonance to satisfy $|\eta| < 2.5$. If we are interested in top quarks with $p_{\mathrm{T}} > 250$ GeV (motivated by the trigger selection (Section 4.1)), then over half of top quarks will be fully contained within a jet of radius 1.5. That is, at $p_{\mathrm{T}} \approx 250$ GeV, it is equally likely that a top quark's decay products will fall within a single large-radius 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 $p_{\mathrm{T}} > 250$ 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 [???].

to pick circular jets, whereas the Cambridge-Aachen (CA) algorithm allows for 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 that sits closer to m_t than anti-k_{T}. Because of this, and the general improvement in S/B near the top mass peak, 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 can accidentally be clustered into a jet from the primary vertex. Fundamental quantities (like top quark momentum) are uncorrelated with the number of primary vertices (N_{PV}), but reconstructed quantities can acquire such a dependence due to the extra radiation. These additional particles can bias the energy scale of the jet (e.g. the mass) as well as geometric observables (described in Section 3.2). 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 using all particles (without PUPPI filtering) have a jet mass and τ_{32}^{SD} distributions (Figure 3.4) in which both the mean and variance have an N_{PV}-dependence. Adding PUPPI stabilizes the mean and ensures that the

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.

variance does not grow at large N_{PV}.

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 (Chapter 2), a substantial fraction originates from extra radiation being absorbed into the jet. These extra particles can arise from pile-up (although this is accounted for by PUPPI), initial state radiation, and underlying event. Many algorithms exist to "groom" such particles from a jet after it has been clustered; here, we will discuss and use the soft drop (SD) method [???]. SD functions by traversing the CA clustering history in reverse and removing branches of the clustering tree that are deemed to be too soft or wide-angled. More formally, at each node in the clustering tree, the softer branch of the node will be removed if

(a) Jet mass, LQG

(b) Jet mass, Top

(c) τ_{32}^{SD}, LQG

(d) τ_{32}^{SD}, Top

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.

it satisfies the condition:

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

where $p_{\mathrm{T,i}}$ refers to the p_{T} of the i-th branch of the node; ΔR_{12} is the ΔR-distance between the two branches; and R and β are tunable parameters. This condition is satisfied if the two branches are very far apart (assuming $\beta \geq 0$) or if the splitting is very asymmetric in momentum. 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 mass and heavy flavor mesons (Section 3.2.2), 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 [???]. It is defined as:

$$\tau_N = \frac{\sum_{i \in \mathrm{jet}} p_{\mathrm{T},i} \min\{\Delta R_{ia} | a \in A\}}{\sum_{i \in \mathrm{jet}} p_{\mathrm{T},i} R} \tag{3.4}$$

(a) Ungroomed, LQG

(b) Ungroomed, Top

(c) SD, LQG

(d) SD, Top

Figure 3.5: Distribution 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 parton that initiates the jet.

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 "subjets." Since the k_{T} distance metric is proportional to ${}^{\Delta R^2}/_{R^2}$, this approximates the ideal minimization. The set of axes A is taken to be the directions of the N subjets. A small τ_N indicates a high degree of compatibility with the N-axis hypothesis. Therefore, we expect a 3-pronged (e.g. top) jet to satisfy $\tau_3 \ll \tau_2$, whereas a 1-pronged (e.g. LQG) jet should satisfy $\tau_3 \lesssim \tau_2$ (for optimal choice of A, $\tau_N \leq \tau_M$ if $N > M$ for any jet). Correspondingly, we take $\tau_{32} \equiv {}^{\tau_3}/_{\tau_2}$ to be the tagging observable.

A combined tagger

3.2.2 Heavy flavor identification

3.3 Data validation

Chapter 4

The Search for $t + p_{\mathrm{T}}^{\mathrm{miss}}$

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.

To illustrate how beyond-SM physics can produce this final state, we introduce two DM models: 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 motivitated and designed agnostically, 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 FCNC V is assumed to couple to a fermionic DM candidate χ. A partial Lagrangian of the interaction terms is given by:

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

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 $p_{\mathrm{T}}^{\mathrm{miss}}$.

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 $\mathrm{SU}(2)_\mathrm{L}$ 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 (i.e. strongest couplings are within generations), 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 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 $\mathrm{SU}(2)_\mathrm{L}$ symmetry, we make the assumption that $g_u^V = g_d^V$ and $g_u^A = g_d^A$.

(a) $(g_u^{V,A})_{ij} \approx \delta_{ij}$ (b) $(g_u^{V,A})_{ij} \approx \delta_{i1}\delta_{j3} + \delta_{i3}\delta_{j1}$

Figure 4.2: Possible DM production at the LHC, assuming a simplified spin-1 extension to the SM.

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:

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

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.

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 $p_{\text{T}}^{\text{miss}}$ distribution, whereas the scalar resonance produces a $p_{\text{T}}^{\text{miss}}$ distribution peaking at approximately $m_\phi/2$.

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.

(a) FCNC

(b) Scalar resonance

Figure 4.4: Spectra of DM (missing) momentum under various signal hypothesis.

(a) Missing momentum

(b) Jet p_{T}

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

4.1 Signal selection

When looking at events that pass a simple set of criteria (moderate $p_{\mathrm{T}}^{\mathrm{miss}}$ and one CA15 jet), it is clear (Figure 4.5) that the highest signal sensitivity is found in regions of high $p_{\mathrm{T}}^{\mathrm{miss}}$ and jet p_{T}. The three primary background processes are:

- $Z \to \nu\nu$. When the Z is produced in association with one or more jets, the jet system can (with low probability) pass the criteria used to select a top jet. The neutrinos manifest as $p_{\mathrm{T}}^{\mathrm{miss}}$.

- $W \to \ell\nu$. As in the case of the Z, additional jets can 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.

- $t\bar{t} \to bq\bar{q}' + \bar{b}\ell\nu$. 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 $p_{\mathrm{T,L1}}^{\mathrm{miss}} > 70$ GeV, where:

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

Events that pass this selection are sent to the HLT system, where we place requirements on the both the missing momentum ($p_{\mathrm{T,HLT}}^{\mathrm{miss}}$) and the missing hadronic momentum ($H_{\mathrm{T,HLT}}^{\mathrm{miss}}$). These are defined as:

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

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

The HLT decides to keep an event if $\min(p_{\mathrm{T,HLT}}^{\mathrm{miss}}, H_{\mathrm{T,HLT}}^{\mathrm{miss}})$ 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.

Note that in all of the trigger decisions, muons are implicitly (in the L1) or explicitly (in the HLT) 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.

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:

$$\epsilon_{\mathrm{trig}}(U) = \frac{N_{\text{pass trig}}(U)}{N(U)} \tag{4.6}$$

This is measured using $W \to \mu\nu$ events containing one or more high-p_{T} jets. The events are triggered using single-μ triggers (which have lower thresholds and efficiencies ~ 1 in this phase space). We then require events have exactly one well-identified

Figure 4.6: Efficiency of the $p_{\mathrm{T}}^{\mathrm{miss}}$ trigger measured in single-muon events.

muon and at least one jet with $p_{\mathrm{T}} > 100$ GeV. Figure 4.6 shows the efficiency as a function of U. To avoid the steeply-rising part of the curve, we choose a threshold of $U > 250$ GeV in the analysis.

4.1.2 Offline signal selection

Events in the signal regions (SRs) are then selected according to Table 4.1, chosen to be consistent with the signal topology while mitigating the aforementioned SM backgrounds. As described in Section 3.2.1, 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 $p_{\mathrm{T}}^{\mathrm{miss}}$ distributions, as predicted by MC and as observed in collected data, in the two signal regions.

(a) Loose SR

(b) Tight SR

Figure 4.7: $p_{\mathrm{T}}^{\mathrm{miss}}$ 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 fluctations in data (black bars) and MC (grey band).

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
$p_{\mathrm{T}}^{\mathrm{miss}} > 250$ GeV	Signal events should have large missing momentum. Exact threshold is chosen to maximize online trigger efficiency.
1 CA15 jet with $p_{\mathrm{T}} > 250$ GeV	Top quark candidate. Recoils against $p_{\mathrm{T}}^{\mathrm{miss}}$, so threshold is set at 250 GeV.
CA15 jet $110 < m_{\mathrm{SD}} < 210$ 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 $e, \mu, \tau_{\mathrm{h}}$	Suppress W+jet and $t\bar{t}$ processes.
No identified γ	Suppress γ+jet processes.
$\min_{\mathrm{jets}} \Delta\phi(\mathrm{jet}, p_{\mathrm{T}}^{\mathrm{miss}}) > 0.5$	Remove events with large $p_{\mathrm{T}}^{\mathrm{miss}}$ caused by mismeasured jets.
CA15 jet BDT	Identifying top decay structure. If the jet passes the tight WP, it is placed in the "tight" SR. Otherwise, if it only passes the loose WP, it is placed in the "loose" SR.

4.2 Background estimation

Searching for DM amounts to looking for an excess of data events over the SM prediction at large values of $p_{\mathrm{T}}^{\mathrm{miss}}$. Therefore, the $p_{\mathrm{T}}^{\mathrm{miss}}$ 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 further, a "data-driven" approach is used to estimate the SM processes in the SR. In this context, "data-driven" refers to the use of control data (i.e. data that cannot contain signal) to directly estimate or supplement the estimation of SM processes in the SR.

4.2.1 Visible final states to constrain invisible final states

As a starting point, let us tackle the estimation of $Z \to \nu\nu$ in the SR. Since the momentum imbalance (up to experimental effects) in a $Z \to \nu\nu$ 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 $p_{\mathrm{T}}^{\mathrm{miss}}(Z \to \nu\nu)$ by measuring $p_{\mathrm{T}}^{Z}(Z \to \mu\mu)$, 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 momentum imbalance. Effects like jet energy scale and acceptance can impact $p_{\mathrm{T}}^{\mathrm{miss}}$, but not $p_{\mathrm{T}}^{\mu\mu}$. Therefore, instead of directly measuring $p_{\mathrm{T}}^{\mu\mu}$ in $\mu\mu$ events, we define and use the hadronic recoil U:

$$\vec{U} = \vec{p}_{\mathrm{T}}^{\,\mathrm{miss}} + \sum_{\mu} \vec{p}_{\mathrm{T}}^{\,\mu} + \sum_{e} \vec{p}_{\mathrm{T}}^{\,e} + \sum_{\gamma} \vec{p}_{\mathrm{T}}^{\,\gamma} \tag{4.7}$$

In the SR (where there are no e, μ, γ), $U = p_{\mathrm{T}}^{\mathrm{miss}}$. In $Z \to \mu\mu$ events, U mimics the momentum imbalance, if we had pretended the identified muons did not exist when computing $p_{\mathrm{T}}^{\mathrm{miss}}$. Therefore, U is an exact analogy for $p_{\mathrm{T}}^{\mathrm{miss}}$ in the SR. Figure 4.8

(a) $Z \to \nu\nu$ (b) $Z \to \mu\mu$

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 $p_{\mathrm{T}}^{\mu\mu}$ is largely independent of the jet.

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:

$$T_{Z,i}^{\mu\mu} = \frac{N_i^{\mathrm{SR}}(Z \to \nu\nu)}{N_i^{\mu\mu}(Z \to \mu\mu)} \tag{4.8}$$

The transfer factors are estimated using MC simulation. To encode the effects of

(a) U in loose CR

(b) U in tight CR

(c) $m_{\mu\mu}$

(d) Leading muon p_{T}

Figure 4.9: Various kinematic distributions in the two mono-top $\mu\mu$ CRs. Note the clearly discernable peak in the $m_{\mu\mu}$ distribution near m_Z

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

Criterion	Notes
$U > 250$ GeV	Mimicking the selection in the SR; also constrained by trigger thresholds.
1 CA15 jet with $p_{\mathrm{T}} > 250$ GeV	Same as SR
CA15 jet $110 < m_{\mathrm{SD}} < 210$ GeV	Same as SR
Well-identified μ^-, μ^+ pair, with $\lvert m_{\mu\mu} - m_Z \rvert < 30$ GeV	Identifying the $Z \to \mu\mu$ resonance.
No identified e, τ_{h}	Same as SR.
No identified γ	Same as SR
$\min_{\mathrm{jets}} \Delta\phi(\mathrm{jet}, U) > 0.5$	Same as SR
CA15 jet BDT	Same as SR

various uncertainties, we introduce nuisance parameters $\boldsymbol{\theta}$. That is:

$$T^X_{Y,i} \to T^X_{Y,i}(\boldsymbol{\theta}) \equiv T^X_{Y,i} \times \prod_{j=0}^{n_\theta} (1 + \theta_j) \tag{4.9}$$

$$\theta_j \sim p_j(\theta_j) \tag{4.10}$$

where n_θ is the number of nuisance parameters and $p_j(\theta_j)$ is some prior distribution for each nuisance (see below for how the priors are used). The priors are typically chosen to have a central value (e.g. mean, median) 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 $\mathrm{Pois}(d|\lambda)$ 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:

$$\begin{aligned}\mathcal{L}(\boldsymbol{d} \mid \mu, \boldsymbol{\mu}^{Z\to\nu\nu}_{\mathrm{SR}}, \boldsymbol{\theta}) = \prod_{i \in \mathrm{bins}} \Bigg[&\mathrm{Pois}\left(d^{\mathrm{SR}}_i \mid \mu S^{\mathrm{SR}}_i(\boldsymbol{\theta}) + \mu^{Z\to\nu\nu}_{\mathrm{SR},i} + B^{\mathrm{SR}}_i(\boldsymbol{\theta})\right) \\ &\times \mathrm{Pois}\left(d^{\mu\mu}_i \;\middle|\; \frac{\mu^{Z\to\nu\nu}_{\mathrm{SR},i}}{T^{\mu\mu}_{Z,i}(\boldsymbol{\theta})} + B^{\mu\mu}_i(\boldsymbol{\theta})\right)\Bigg] \times \prod_{j=0}^{n_\theta} p_j(\theta_j)\end{aligned} \tag{4.11}$$

where the following notation is used:

d_i^X : The data observed in bin i of region X. For now, $X = \mathrm{SR}, \mu\mu$.

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.

$\mu_{\mathrm{SR},i}^P$: 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. Here, "minor" refers to all SM processes that are not the signal and are not estimated using a data-driven method.

The signal and background yields $\boldsymbol{S}$ and $\boldsymbol{B}$ are estimated using MC. Note that the inclusion of the priors in the likelihood constrains the nuisance parameters to be close to their "nominal" values; moving a θ_j to fit the data incurs a large cost from the prior.

If we set $B_i = \mu = 0$ (the null hypothesis, ignoring small minor backgrounds), a simple picture emerges of how the likelihood is maximized. The parameters $\boldsymbol{\mu}_{\mathrm{SR}}^{Z\to\nu\nu}$ float freely to satisfy $d_{\mathrm{SR,i}} \sim \mu_{\mathrm{SR},i}^{Z\to\nu\nu}$ and $d_{\mu\mu,\mathrm{i}} \sim \mu_{\mathrm{SR},i}^{Z\to\nu\nu}/T_{Z,i}^{\mu\mu}(\boldsymbol{\theta})$. If both constraints cannot be satisfied simultaneously by scaling $\boldsymbol{\mu}_{\mathrm{SR}}^{Z\to\nu\nu}$, the (constrained) nuisance parameters $\boldsymbol{\theta}$ modify the transfer factor $\boldsymbol{T}_Z^{\mu\mu}$. Table 4.6 shows the relevant uncertainties for $\boldsymbol{T}_Z^{\mu\mu}$, 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 $\boldsymbol{T}_Z^{\mu\mu}(\boldsymbol{\theta})$ have two salient features:

1. The values are strictly greater than one. This is due to (a) $\mathcal{B}(Z \to \nu\nu) > \mathcal{B}(Z \to \mu\mu)$ and (b) a non-100% efficiency in reconstructing and identifying muons. This implies that the constraining power of the $\mu\mu$ CR is less than that of the SR, especially at high U (i.e. the Poisson uncertainties are larger).

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

(a) Loose

(b) Tight

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

(a) Loose (b) Tight

Figure 4.11: The transfer factors $T^{\mu\mu}_{Z,i}$ 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.

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

To account for point (1), we can simply add more control data by also looking at $Z \to ee$ 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 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 $Z \to \ell\ell$ CRs, we can use the hadronic recoil U in CRs that measure visible final states of W and $t\bar{t}$.

(a) U in loose CR (b) U in tight CR (c) m_{ee}

Figure 4.12: Various kinematic distributions in the two mono-top ee CRs.

(a) Loose (b) Tight

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.

(a) $W \to \ell\nu$ (b) $t\bar{t} \to bqq' + b\ell\nu$

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

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 targetted process in the SR: $\boldsymbol{T}_{t\bar{t}}^{b\mu}$ and $\boldsymbol{T}_{W}^{\mu}$. 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 $\boldsymbol{T}_{t\bar{t}}^{\mu}$. 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.

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.

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

Criterion	Notes
$U > 250$ GeV	Mimicking the selection in the SR; also constrained by trigger thresholds.
1 CA15 jet with $p_{\mathrm{T}} > 250$ GeV	Same as SR
CA15 jet $110 < m_{\mathrm{SD}} < 210$ 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 $W \to \mu\nu$
No identified e, τ_{h}	Same as SR.
No identified γ	Same as SR
$\min_{\mathrm{jets}} \Delta\phi(\mathrm{jet}, U) > 0.5$	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 split based on the jet BDT score.

Criterion	Notes
$U > 250$ GeV	Mimicking the selection in the SR; also constrained by trigger thresholds.
1 CA15 jet with $p_{\mathrm{T}} > 250$ GeV	Same as SR
CA15 jet $110 < m_{\mathrm{SD}} < 210$ 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 $W \to \mu\nu$
No identified e, τ_{h}	Same as SR.
No identified γ	Same as SR
$\min_{\mathrm{jets}} \Delta\phi(\mathrm{jet}, U) > 0.5$	Same as SR
CA15 jet BDT	Same as SR

(a) U in loose CR (b) U in tight CR (c) m_{SD}

Figure 4.15: Various kinematic distributions in the two mono-top $b\mu$ CRs.

(a) U in loose CR (b) U in tight CR (c) m_{SD}

Figure 4.16: Various kinematic distributions in the two mono-top μ CRs.

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. ($\boldsymbol{T}^{b\mu}_{t\bar{t}}$)	1 s.d. ($\boldsymbol{T}^{\mu}_{W}$)	1 s.d. ($\boldsymbol{T}^{\mu}_{t\bar{t}}$)	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

(a) Loose $\boldsymbol{T}_{t\bar{t}}^{b\mu}$ (b) Tight $\boldsymbol{T}_{t\bar{t}}^{b\mu}$

(c) Loose $\boldsymbol{T}_{W}^{\mu}$ (d) Tight $\boldsymbol{T}_{W}^{\mu}$

(e) Loose $\boldsymbol{T}_{t\bar{t}}^{\mu}$ (f) Tight $\boldsymbol{T}_{t\bar{t}}^{\mu}$

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

(a) U in loose be CR

(b) U in tight be CR

(c) m_{SD} in tight be CR

(d) U in loose e CR

(e) U in tight e CR

(f) m_{SD} in tight e CR

Figure 4.18: Various kinematic distributions in the mono-top be CRs (top) and e CRs (bottom).

(a) Loose $\boldsymbol{T}^{be}_{t\bar{t}}$

(b) Loose $\boldsymbol{T}^{e}_{W}$

(c) Loose $\boldsymbol{T}^{e}_{t\bar{t}}$

(d) Tight $\boldsymbol{T}^{be}_{t\bar{t}}$

(e) Tight $\boldsymbol{T}^{e}_{W}$

(f) Tight $\boldsymbol{T}^{e}_{t\bar{t}}$

Figure 4.19: The transfer factors $\boldsymbol{T}^{be}_{t\bar{t}}$, $\boldsymbol{T}^{e}_{W}$, and $\boldsymbol{T}^{e}_{t\bar{t}}$

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

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

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 $Z \to \nu\nu$ 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 $\sigma(pp \to Z \to \nu\nu) > \sigma(pp \to Z \to \ell^+\ell^-)$; accordingly, to alleviate this limitation, we look to a process with a much bigger cross-section. In similar regions of final-state phase space, $\sigma(pp \to \gamma + \text{jets}) \sim 30 \times \sigma(pp \to Z(\to \nu\nu) + \text{jets})$. Therefore, it is natural to use the γ+jet production spectrum as a way to estimate the Z+jet spectrum. As before, let us define another transfer factor:

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

However, unlike the $\boldsymbol{T}_Y^X$ we have discussed so far (which correlate similar processes, e.g. $Z \to \nu\nu / Z \to \mu\mu$), $\boldsymbol{T}_\gamma^\gamma$ is highly sensitive to the theoretical predictions of the Z and γ spectra.

To reduce the impact of higher-order effects on $\boldsymbol{T}_\gamma^\gamma$, 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 simulation. While producing sufficient NLO simulation for the analysis is prohibitive, we can compute certain inclusive distributions at NLO. Since $U \approx p_{\mathrm{T}}^{V}$ 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 large corrections, both at low and high p_{T}^{V}. 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 clearly is insufficient to cover NLO effects. Therefore, we compute a simple correction for the NLO QCD effects, known as a k-factor:

$$k_{Z,\mathrm{QCD}}(p_{\mathrm{T}}^{Z}) = \frac{d\sigma_{\mathrm{NLO\ QCD}}(Z)/dp_{\mathrm{T}}^{Z}}{d\sigma_{\mathrm{LO}}(Z)/dp_{\mathrm{T}}^{Z}} \tag{4.14}$$

We include another$k$-factor, k_{EWK}, to correct for higher-order EWK effects. Unlike k_{QCD}, k_{EWK} is derived using a theoretical calculation [7, 8, 11] instead of NLO simulation. k_{EWK} covers NLO EWK 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 for each process. Therefore, transfer factors like $\boldsymbol{T}_{\gamma}^{\gamma}$ are strongly sensitive to NLO effects, i.e.:

$$T_{\gamma,i}^{\gamma} = \frac{N_i^{\mathrm{SR}}(Z \to \nu\nu)}{N_i^{\gamma}(\gamma)} \neq \frac{N_i^{\mathrm{SR,LO}}(Z \to \nu\nu)}{N_i^{\gamma,\mathrm{LO}}(\gamma)} \tag{4.15}$$

The distributions in Sections 4.1-4.2 are all corrected using these k-factors. Figure 4.22 shows the equivalent for CRs that target γ+jet events, and Table 4.7 describes the selection used to define these CRs.

Now that we can describe $\boldsymbol{T}_{\gamma}^{\gamma}$ 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 [1]. By varying μ_F and μ_R by factors of 0.5 and

(a) LO vs. NLO (QCD)

(b) k-factors

Figure 4.20: Theoretical predictions for p_{T}^{Z} in Z+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.

(a) Z

(b) W

(c) γ

Figure 4.21: Differential k-factors for each of the V+jet processes.

(a) U in loose CR (b) U in tight CR (c) p_{T}^{γ}

Figure 4.22: Various kinematic distributions in the two mono-top γ CRs.

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

Criterion	Notes
$U > 250$ GeV	Mimicking the selection in the SR
1 CA15 jet with $p_{\mathrm{T}} > 250$ GeV	Same as SR
CA15 jet $110 < m_{\mathrm{SD}} < 210$ GeV	Same as SR
No identified $\mu, e, \tau_{\mathrm{h}}$	Same as SR.
Well-identified γ with $p_{\mathrm{T}}^{\gamma} > 175$ GeV	High-p_{T} photon, set by trigger threshold
$\min_{\mathrm{jets}} \Delta\phi(\mathrm{jet}, U) > 0.5$	Same as SR
CA15 jet BDT	Same as SR

2, we assess the effect of the integration scale choices on $\boldsymbol{T}$. 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 EWK effects, following what is suggested in References [3, 4, 5, 7, 8, 9, 10, 11, 12] and agreed upon in the LHC Dark Matter Working Group. These EWK uncertainties break down into three categories:

- Unknown Sudakov logarithms in the NLL correction. These uncertainties are correlated across processes (Z, W, γ).
- Missing NNLO EWK 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 processes, they do not affect transfer factors that correlate processes differing only in decay mode or acceptance. This is simply because these uncertainties primarily deal with the initial state or the production of an electroweak boson, which is not related to the description of the decay to leptons or the experimental identification of leptons and b-jets. That is:

$$T_{Z,i}^{\mu\mu} = \frac{N_i^{\mathrm{SR}}(Z \to \nu\nu)}{N_i^{\mu\mu}(Z \to \mu\mu)} \approx \frac{N_i^{\mathrm{SR,LO}}(Z \to \nu\nu)}{N_i^{\mu\mu,\mathrm{LO}}(Z \to \mu\mu)} \tag{4.16}$$

Now that we have tools to construct transfer factors of the form $N(V)/N(V')$ with reasonably small uncertainties (i.e. smaller than the statistical uncertainty of the data), it is natural to add another transfer factor to our toolbox:

$$T_{Z/W,i}^{\mathrm{SR}} = \frac{N_i^{\mathrm{SR}}(Z \to \nu\nu)}{N_i^{\mathrm{SR}}(W \to \ell\nu)} \tag{4.17}$$

This allows us to use the e, μ CRs (which target W+jet production) to further reduce

(a) Loose Z/γ

(b) Tight Z/γ

(c) Loose Z/W

(d) Tight Z/W

Figure 4.23: The transfer factors $\boldsymbol{T}_{\gamma}^{\gamma}$ and $\boldsymbol{T}_{Z/W}^{\mathrm{SR}}$; and corresponding shape uncertainties.

the uncertainty in the estimation of $Z \to \nu\nu$ 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:

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

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 $\boldsymbol{T}_{\gamma}^{\gamma} \ll 1$, and the same holds for the effective transfer factor $\boldsymbol{T}_{Z/W}^{\mathrm{SR}} \times \boldsymbol{T}_{W}^{\mu}$. 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:

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

The discussion in this section has largely relied on arguments from simulation and calculation. We can, however, validate that our estimates of $\boldsymbol{T}$ and the corresponding uncertainties are reasonable by using CR data and appropriate proxies. For example, to test $\boldsymbol{T}_\gamma^\gamma \sim N(Z \to \nu\nu)/N(\gamma)$, we can look at $N(Z \to \mu\mu)/N(\gamma)$. 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 results of a ML fit under a background-only hypothesis (i.e. setting $\mu = 0$) 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.

(a) Loose SR

(b) Tight SR

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

(a) Loose μ CR

(b) Tight μ CR

(c) Loose $b\mu$ CR

(d) Tight $b\mu$ CR

(e) Loose e CR

(f) Tight e CR

(g) Loose be CR

(h) Tight be CR

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 $\mu = 0$.

(a) Loose $\mu\mu$ CR

(b) Tight $\mu\mu$ CR

(c) Loose ee CR

(d) Tight ee CR

(e) Loose γ CR

(f) Tight γ CR

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 $\mu = 0$.

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 $\sigma_{95\%\ \mathrm{CLs}} < \sigma_{\mathrm{theory}}$ 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).

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+backgroundl 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 Section [**?**].

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 $\phi \to t\psi$ is strongly suppressed. The values of the couplings a_q, b_q and a_ψ, b_ψ are similarly fixed to reasonable values. The cross section (and therefore the signal strength 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.

(a) Vector couplings (b) Axial couplings

Figure 4.29: 95% CLs upper limits on the signal strength ($\mu = \sigma/\sigma_{\text{theory}}$) of the flavor-violating process $pp \to t\chi\bar{\chi}$ 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).

The FCNC model contains many more non-trivial parameters, and so we do not simply fix all but one. First, fixing the couplings (at values agreed upon by the LHC Dark Matter Working Group in Reference [6]), 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 $t^{(*)} \to Vu$ 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 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

(a) Vector, m_V vs g_χ^V

(b) Vector, m_V vs g_q^V

(c) Axial, m_V vs g_χ^A

(d) Axial, m_V vs g_q^A

Figure 4.30: 95% CLs upper limits on the signal strength ($\mu = \sigma/\sigma_{\text{theory}}$) of the flavor-violating process $pp \to t\chi\bar{\chi}$ 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.

sufficiently low masses (still corresponding to viable hypotheses), 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

The discovery of the SM Higgs boson [???] 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) [2]. While gluon fusion is the most frequent mode, the unique detector signatures of the other production modes can be combined the various Higgs decay signatures to define a signal topology with few 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 $H \to \chi\bar{\chi}$. From measurements of the visible branching fractions, we can indirectly place an upper bound of $\mathcal{B}(H \to \chi\bar{\chi}) < 0.2$ [???]. In this chapter, we describe a direct search for $H \to \chi\bar{\chi}$ decays.

As with the case of the mono-top search, the $H \to \chi\bar{\chi}$ process manifests as $p_{\mathrm{T}}^{\mathrm{miss}}$. Each of the aforementioned Higgs production modes translates into a $p_{\mathrm{T}}^{\mathrm{miss}} + X$ 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, as the unique final state topology provides the best sensitivity to $H \to \chi\bar{\chi}$.

(a) $gg \to H(\to \chi\bar{\chi})$+jet(s) (b) $qq' \to H(\to \chi\bar{\chi})$+jet(s) (c) $qq' \to VH(\to \chi\bar{\chi})$

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 $p_{\mathrm{T}}^{\mathrm{miss}}$ through the presence of one or more SM particles in the final state.

5.1 Signal selection

VBF $H \to \chi\bar{\chi}$ events are characterized by large $p_{\mathrm{T}}^{\mathrm{miss}}$ 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 $H \to \chi\bar{\chi}$ 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 $H \to \chi\bar{\chi}$ event with two energetic forward jets ($p_{\mathrm{T}} =$ 180, 107 GeV) and large $p_{\mathrm{T}}^{\mathrm{miss}}$ (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 $p_{\mathrm{T}}^{\mathrm{miss}}$.

(a) Recoil

(b) $H^{\text{miss}}_{\text{T,barrel}}$

Figure 5.3: Trigger efficiency of events with a VBF-like topology (two jets with $p_{\text{T}} > 80, 40$ GeV) as a function of two different observables. Events are split into two categories: those where both jets have $|\eta| < 3$ (BB) and those were 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":

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

As shown in Figure 5.3b, the three categories (BB, BF, All) have similar behavior as a function of $H^{\text{miss}}_{\text{T,barrel}}$. 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. BX_0 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 BX_0.

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 BX_0 (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 (BX_0) 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:

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

However, by definition, pre-fired events cannot be recorded, and therefore $N_{\text{pre-fire}}(p_{\text{T}}, \eta, \phi)$ 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:

$$\epsilon_{\text{pre-fire}}(p_{\mathrm{T}}, \eta, \phi) = \frac{N_{\text{pre-fire } \mathrm{BX}_0|\mathrm{BX}_{-3}}(p_{\mathrm{T}}, \eta, \phi)}{N_{\mathrm{BX}_0|\mathrm{BX}_{-3}}(p_{\mathrm{T}}, \eta, \phi)} \tag{5.3}$$

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 $p_{\mathrm{T}}^{\mathrm{miss}}$-based triggers (MET). This is a consequence of zero suppression biasing the BX_0 triggers, as shown diagrammatically in Figure 5.9. Muon-based

(a) Number of jets

(b) Pre-fire probability

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:

$$\epsilon^{\text{event}}_{\text{pre-fire}} = 1 - \prod_{j \in \text{jets}} \left(1 - \epsilon_{\text{pre-fire}}(p^j_{\text{T}}, \eta^j)\right) \tag{5.4}$$

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 $\epsilon_{\text{pre-fire}}(p_{\text{T}}, \eta)$ using muon-triggered and jet-triggered events. In the former case, statistical fluctuations make the region with $p_{\text{T}} > 250$ 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.

(a) p_{T} (b) $p_{\mathrm{T}} \times E_{\mathrm{ECAL}}/E_{\mathrm{total}}$

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.

(a) Dijet event (b) $Z \to \nu\nu$ event

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 $Z \to \nu\nu$ event, in which the loss of an ECAL deposit reduces the total $p_{\mathrm{T}}^{\mathrm{miss}}$ of the event.

(a) Muon-triggered

(b) Jet-triggered

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 jet and $W(\rightarrow \ell\nu)$+2 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

(a) QCD (b) EW/VBF

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 $p_{\mathrm{T}}^{\mathrm{miss}}$ distribution is much harder than SM predictions.

3. VBF $H \to \chi\bar{\chi}$ produces fewer jets than SM processes.

4. VBF $H \to \chi\bar{\chi}$ produces relatively forward jets. QCD V+jets produces mostly central jets. VBF V+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.

(a) Leading jet p_{T} (b) Sub-leading jet p_{T} (c) N_{jet}

(d) Leading jet η (e) Sub-leading jet η (f) $p_{\mathrm{T}}^{\mathrm{miss}}$

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

(a) m_{jj} (b) $\Delta\eta_{jj}$ (c) $\Delta\phi_{jj}$

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

(a) Fit m_{jj} (b) Fit $\Delta\eta_{jj}$ (c) Fit $\Delta\phi_{jj}$

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:

- $p_{\mathrm{T}}^{\mathrm{miss}} > 250$ GeV: driven by trigger efficiency, as discussed in Section 5.1.1.
- $p_{\mathrm{T}}^{\mathrm{jet}} > 80, 40$ 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.
- $\min \Delta\phi(\mathrm{jet}, p_{\mathrm{T}}^{\mathrm{miss}}) > 0.4$: remove QCD multijet events.
- $|p_{\mathrm{T,calo}}^{\mathrm{miss}} - p_{\mathrm{T}}^{\mathrm{miss}}| < p_{\mathrm{T}}^{\mathrm{miss}}/2$: 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 $\mathcal{B}(H \to \chi\bar{\chi})$. 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 $\boldsymbol{T}$ are a function of m_{jj}. Control regions are defined using dilepton (single-lepton) selections to estimate the Z (W) contributions. Again, $p_{\mathrm{T}}^{\mathrm{miss}}$ is replaced by U (Equation 4.7) 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:

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

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 $\boldsymbol{\mu}_{\mathrm{SR}}^{Z\to\nu\nu}$:

$$T_{Z,i}^{\mathrm{QE}} = \frac{N_i^{\mathrm{SR}}(\mathrm{QCD}\, Z\to\nu\nu)}{N_i^{\mathrm{SR}}(\mathrm{EW}\, Z\to\nu\nu)} \tag{5.6}$$

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:

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

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:

$$\boldsymbol{T}_Z^{\mu\mu}, \boldsymbol{T}_Z^{ee} \sim \frac{N_{\mu\mu}(Z \to \mu\mu)}{N_{ee}(Z \to ee)}$$
$$\boldsymbol{T}_W^{\mu}, \boldsymbol{T}_W^{e} \sim \frac{N_{\mu}(W \to \mu\nu)}{N_{e}(W \to e\nu)}$$
$$\boldsymbol{T}_{Z/W}^{\mathrm{SR}} \sim \frac{N_{\mu}(W \to \mu\nu) + N_{e}(W \to e\nu)}{N_{\mu\mu}(Z \to \mu\mu) + N_{ee}(Z \to ee)} \tag{5.8}$$

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 $H \to \chi\bar{\chi}$. 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 $\sigma(qq \to qqH)\mathcal{B}(H \to \chi\bar{\chi})$. 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

(a) SR (b) $\mu\mu$ CR (c) ee CR

(d) μ CR (e) e CR

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 $\mathcal{B}(H \to \chi\bar{\chi})$.

searches generally probe central jets. Theoretical nuisances (e.g. those affecting W/Z or ZZ/WZ ratios) are left uncorrelated between all searches. The combined result constrains $\mathcal{B}(H \to \chi\bar{\chi})$ 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 $\mathcal{B}(H \to \chi\bar{\chi})$ as an upper limit on the spin-independent cross section of DM-nucleon scattering. First, we convert the branching ratio into a partial width:

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

Then, using the results described in Reference [???], $\Gamma_{H\to\chi\bar{\chi}}$ is translated into $\sigma^{\text{SI}}_{\chi N}(m_\chi)$, 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

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

extend the DD constraints, reaching well below the neutrino floor.

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

Bibliography

[1] Richard D. Ball, Valerio Bertone, Stefano Carrazza, Christopher S. Deans, Luigi Del Debbio, Stefano Forte, Alberto Guffanti, Nathan P. Hartland, José I. Latorre, Juan Rojo, and Maria Ubiali. Parton distributions for the LHC Run II. *JHEP*, 04:040, 2015.

[2] D. de Florian *et al.* Handbook of LHC Higgs Cross Sections: 4. Deciphering the Nature of the Higgs Sector. Technical Report 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.

[3] Ansgar Denner, Stefan Dittmaier, Tobias Kasprzik, and Alexander Muck. Electroweak corrections to W + jet hadroproduction including leptonic W-boson decays. *JHEP*, 08:075, 2009.

[4] Ansgar Denner, Stefan Dittmaier, Tobias Kasprzik, and Alexander Muck. Electroweak corrections to dilepton + jet production at hadron colliders. *JHEP*, 06:069, 2011.

[5] Ansgar Denner, Stefan Dittmaier, Tobias Kasprzik, and Alexander Mück. Electroweak corrections to monojet production at the LHC. *Eur. Phys. J. C*, 73:2297, 2013.

[6] Daniel Abercrombie *et al.* Dark Matter Benchmark Models for Early LHC Run-2 Searches: Report of the ATLAS/CMS Dark Matter Forum. 2015.

[7] Stefan Kallweit, Jonas M. Lindert, Philipp Maierhöfer, Stefano Pozzorini, and Marek Schönherr. NLO electroweak automation and precise predictions for W+multijet production at the LHC. *JHEP*, 04:012, 2015.

[8] Stefan Kallweit, Jonas M. Lindert, Stefano Pozzorini, Marek Schönherr, and Philipp 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.

[9] Johann 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*, 609:277, 2005.

[10] Johann 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*, 727:368, 2005.

[11] Johann H. Kuhn, A. Kulesza, S. Pozzorini, and M. Schulze. Electroweak corrections to hadronic photon production at large transverse momenta. *JHEP*, 03:059, 2006.

[12] Johann H. Kuhn, A. Kulesza, S. Pozzorini, and M. Schulze. Electroweak corrections to hadronic production of W bosons at large transverse momenta. *Nucl. Phys. B*, 797:27, 2008.