Systematics in charged Higgs searches in CMS

Lauri A. Wendland[∗] **on behalf of the CMS collaboration**
Helsinki Institute of Physics, Finland
E-mail: `lauri.wendland@helsinki.fi`

Data-driven methods for measuring the backgrounds for charged Higgs boson searches in the CMS experiment at the CERN LHC are presented. First results from the LHC collision data for systematic uncertainties and background estimation are shown. Plans for the estimation of systematic uncertainties for the signal are discussed.

[∗]Speaker.

1. Introduction

The charged Higgs boson decay $H^{\pm} \to \tau^{\pm}\nu_{\tau}$ is very interesting, since it provides one of the earliest measurements of the charged Higgs boson at the CMS detector at the LHC [1]. Particularly in the case where the charged Higgs boson is lighter than the top quark, the charged Higgs production cross-section via $gg/q\bar{q} \to t\bar{t}$ and the coupling $t \to H^{\pm}b$ are boosted. Furthermore, the branching ratio of $H^{\pm} \to \tau^{\pm}\nu_{\tau}$ is almost one for a large part of the allowed $\tan\beta$ region.

The decay signature of such a decay is classified into

- fully hadronic $t\bar{t} \to (H^{\pm}b)(W^{\mp}\bar{b}) \to (\tau\nu_{\tau}b)(\bar{b}q\bar{q}) \to (\text{hadr.} + 2\nu_{\tau}b)(\bar{b}q\bar{q})$,

- semi-leptonic $t\bar{t} \to (H^{\pm}b)(W^{\mp}\bar{b}) \to (\tau\nu_{\tau}b)(\bar{b}\ell^{\mp}\nu_{\ell}) \to (\text{hadr.} + 2\nu_{\tau}b)(\bar{b}\ell^{\mp}\nu_{\ell})$ or
  $t\bar{t} \to (H^{\pm}b)(W^{\mp}\bar{b}) \to (\tau\nu_{\tau}b)(\bar{b}q\bar{q}) \to (\ell^{\mp}\nu_{\ell}2\nu_{\tau}b)(\bar{b}q\bar{q})$,

- and dileptonic $t\bar{t} \to (H^{\pm}b)(W^{\mp}\bar{b}) \to (\tau\nu_{\tau}b)(\bar{b}\ell^{\mp}\nu_{\ell}) \to (\ell^{\mp}\nu_{\ell}2\nu_{\tau}b)(\bar{b}\ell^{\mp}\nu_{\ell})$

final states, where $\ell =$ e or μ. The physical signature involves thus the measurement of hadronically decaying tau leptons, isolated electrons and muons, b-hadronic jets, hadronic jets of lighter flavors, and missing energy from the neutrinos. In essence, the successful measurement of the charged Higgs boson in this decay mode requires information of every subdetector of the CMS detector. Consequently, in terms of systematic uncertainties, one has to address the whole systematics menu: the isolated electron/muon reconstruction, identification and fake rate; the hadronically decaying tau lepton energy scale, identification and fake-rate; b-jet tagging and mistagging; and hadronic jet and missing transverse energy energy scale. The current understanding of the systematic uncertainties of these identities is reviewed in Section 2.

The aforementioned systematics menu needs to be applied for the measurement of the charged Higgs boson signal from data. The systematic uncertainties of the main backgrounds, the QCD multi-jet events, $t\bar{t}$ production, W+jets, and Z/γ^*, can be reduced by measuring them separately from the data. Strategies for such measurements are presented in Section 3.

2. Systematic uncertainties of the physics objects

2.1 Systematic uncertainty of isolated electrons

The systematic uncertainty of the reconstruction and identification of isolated electrons with $p_T > 20$ GeV/c has been determined from $Z \to$ ee events with the tag-and-probe technique and from $W \to e\nu_e$ events with a maximum likelihood fit of the transverse mass [2]. Both cut-based and category-based (amount of bremsstrahlung) electron identification methods were considered with working points of 80% and 95% for the electron identification efficiency. For the 95% working point of the category-based electron identification, a difference of 3% (5%) between MC and data was observed for $|\eta| < 1.4$ ($1.4 < |\eta| < 3.0$) when combining the W-tagging and Z tag-and-probe measurements.

To measure the probability for a fake isolated electron, events were selected with the single jet trigger with uncorrected $E_T > 15$ GeV. The events were required to have missing $E_T < 30$ GeV and the electron candidates were required to be outside the jet that was triggered. The fake rate

was measured to be 0.7% to 4% for the electron p_T range of 20-60 GeV/c for the category based electron identification with the 95% working point. A reasonable agreement was found between data and simulation results.

2.2 Systematic uncertainty of isolated muons

The systematic uncertainty of the reconstruction and identification of muons has been studied with inclusive muons of $p_T > 15$ GeV/c with standard muon identification methods [3]. The statistical uncertainty of this method was found to be 3.0% and it was found to dominate the total uncertainty. Studies done with $Z \rightarrow \mu\mu$ events and the tag-and-probe technique agree with the obtained 3.0% uncertainty.

The probability of charged pions, kaons, or protons to fake a muon were measured from tracks coming from identified K_S, ϕ, and Λ resonances. With such method, the fake rate was determined to be $(1.0 \pm 0.2) \times 10^{-4}$ with good agreement of data and simulation results.

2.3 Systematic uncertainty of electromagnetic energy scale

The systematic uncertainty of the electromagnetic energy scale has been evaluated from $\pi^0/\eta \rightarrow \gamma\gamma$ decays by comparing the reconstructed diphoton mass peaks between data and simulated results [4]. With such method, a systematic uncertainty of 0.9% (2.2%) was obtained for $|\eta| < 1.4$ ($1.4 < |\eta| < 3.0$) with the E_T cut identified as the leading source of systematic uncertainty. The absolute energy scale was obtained from test-beam results.

2.4 Systematic uncertainty of jet energy scale

The jet energy scale corrections at CMS use a factorized approach [5]. First an offset correction is made followed by relative η correction and absolute jet p_T corrections. The correction factors can be obtained either from simulation or in-situ with the di-jet p_T balance method [7, 8]. The data vs. simulation comparison has been studied for three anti-k_T-based jet algorithms: calorimeter jets, jet-plus-track (JPT) jets, and particle-flow (PF) jets [6]. The scale uncertainty has been estimated from simulation-based correction factors to be 10%$\oplus$2%/η (5%$\oplus$2%/η) for calorimeter jets (JPT and PF jets).

The absolute jet energy scale is planned to be obtained from γ+jets events once enough data is available. First such studies hint that the current scale uncertainty is quite conservative.

2.5 Systematic uncertainty of missing transverse energy scale

The missing E_T (MET) scale has been estimated from a sample with at least two jets with $p_T > 25$ GeV/c and $|\eta| < 3$ [9]. The scale uncertainty has been studied for the three used MET algorithms: type-II corrected calorimeter MET, track-corrected MET, and particle-flow MET. Comparing the data and simulation results yields 10% as a conservative estimate of the MET scale uncertainty for all three MET algorithms.

2.6 Systematic uncertainty of b-jet tagging

The systematic uncertainty of b-jet tagging efficiency is evaluated from b-jets containing muons [10]. After selecting events with at least one jet of $p_T > 30$ GeV/c and containing one muon

in the event with $p_T > 5$ GeV/c, a template method based on p_T^{rel} shape, i.e. the momentum of the muon transverse to the jet axis, is used to determine the b-tag efficiency. The template shapes for b- and c-jets and for other flavor jets are determined from simulations. The systematic uncertainty averaged over the impact parameter and secondary vertex based tagging algorithms with different working points was found to be 19% for a sample with average p_T of the jets 31 GeV/c.

The b-mistag rate is evaluated from tracks with negative impact parameter or from secondary vertices with negative decay lengths with $\varepsilon_{data}^{mistag} = \varepsilon_{data}^{-} \frac{\varepsilon_{MC(uds)}^{mistag}}{\varepsilon_{MC(all)}^{-}}$, where the minus sign denotes the negative tag rate and uds refers to the light flavors. The systematic uncertainty for the b-mistag rate was found to be 3% (6-12%, 40-60%) for the operating point of 10% (1%, 0.1%) of light flavors passing the b-tag. The dominating sources for the systematic uncertainty were established as the templates b- and c-flavor fractions (20% rel.), parton distribution function used for the gluon fraction (20% rel.), presence of long lived K_S^0 and Λ decays (10-20% rel.), and mismeasured tracks (50% rel.).

2.7 Systematic uncertainty of hadronically decaying tau leptons

No estimate for the hadronically decaying tau lepton (tau jet) reconstruction and identification uncertainty or for the tau-jet energy scale has been so far published. For the tau-jet reconstruction and identification uncertainty, a crude estimate of 10% can be used. For the tau-jet energy scale, a conservatively estimated uncertainty of 10% can be taken based on the jet energy scale uncertainty. The jet$\rightarrow\tau$ fake-rate has been evaluated for the four tau-jet reconstruction algorithms (track-corrected tau, particle-flow tau, hadron-plus-strips tau, and tau neural classifier) [11]. The fake-rate was determined from a sample taken with the single jet trigger with uncorrected jet $E_T > 15$ GeV/c. Such sample is expected to be dominated by QCD multi-jet events. The fake-rate is defined as the number of tau-jet candidates passing the tau identification divided by the number of all tau-jet candidates in a given bin of tau-jet p_T or η. For tau-jets of $p_T > 30$ GeV/c, the data and simulation were found to disagree by 20-30% independent of the tau-jet algorithm and working point. The exact cause of the discrepancy is under further study.

3. Strategies for dedicated background measurements

With a cross-section exceeding 6×10^7 pb for $30 < \hat{p}_T < 300$ GeV/c, where $\hat{p}_T$ is the p_T of the jets in the rest frame of the simulated hard interaction, QCD multi-jet background is by far the largest background that has to be considered for the light charged Higgs boson searches from $t\bar{t}$ production. The subleading backgrounds are the W+jets (2.5×10^4 pb) and $t\bar{t}$ ($165 \times (1 - \mathrm{Br}(t \rightarrow bH^{\pm}))^2$ pb). For the di-leptonic final state, also the Z/γ^* background has to be considered. With dedicated background measurements it is possible to bypass some of the sources of systematic uncertainty such as theoretical uncertainty of background cross-section.

3.1 QCD multi-jet background measurement for the fully hadronic final state

The measurement of the QCD multi-jet background from data for the fully hadronic final state is planned to be done in two parts: measurement of the jet$\rightarrow\tau$ fake-rate and the full event

selection with tau-jet identification replaced by the jet→τ fake-rate. This strategy relies on the well established fact that tau-jet identification is practically uncorrelated with any other event selections.

The QCD multi-jet sample can be selected either with a single jet or with a single tau trigger. The single tau trigger seems to be a more promising option, since the jet triggers are heavily prescaled, and since the QCD multi-jet events still dominate the sample even after the single tau trigger. To reduce possible contamination of the sample by W+jets events, the sample can be further cleaned by requiring an upper limit on the missing E_T and by requiring at least N_{jets} central hadronic jets in the event with $E_T > 30$ GeV. After the cleaning of the sample, tau-jet candidates that have been matched to the trigger object, that are central, and that have jet $E_T > 30$ GeV are considered. The jet→τ fake-rate is obtained by dividing the number of tau-jet candidates that have passed the standard tau-jet identification algorithm by the total number of tau-jet candidates in bins of tau-jet candidate E_T and η. With the careful cleaning of the W+jets events, the statistical uncertainty of the fake-rate measurement with early data is expected to dominate the bias coming from the presence of electroweak events in the data sample.

The overall number of QCD multi-jet events is planned to be measured from data by applying the same event selections as for the signal selection with the tau-jet identification part factorized out by the jet→τ fake-rate measurement. To ensure that QCD multi-jet events dominate the selected sample, it might be necessary to further factorize or parametrize the missing E_T event selection.

3.2 QCD multi-jet background measurement for the semi-leptonic final state

Data-driven methods have been successfully implemented to measure the QCD multi-jet background in semi-leptonic $t\bar{t}$ decays [12]. These methods are also applicable to the charged Higgs boson searches in the semi-leptonic final state. For the muonic final state, the number of QCD multi-jet events is measured with the ABCD method. In this method, the relative isolation variable and the impact parameter of the muon are used to define four regions, in which QCD multi-jet events are dominating in three regions and the last region is the signal region. Since the chosen variables are independent, the number of QCD multi-jet events can be evaluated from the measurement of the three other regions. Otherwise, the standard set of event selections is applied.

A complementary method is to measure the shape of the relative isolation variable distribution in the region dominated by QCD multi-jet events and to extrapolate the curve to the signal region. A third method is to reverse the relative isolation cut to select a QCD-enriched sample to obtain kinematic distributions of this background. The overall normalization can be obtained also here from the ABCD method. This method is applicable to both muonic and electronic final states.

Additionally, for the electronic final state, the QCD multi-jet background can be measured successfully with a template fit method [12].

3.3 Electroweak background measurement for the fully hadronic final state

The W+jets and $t\bar{t}$ backgrounds are planned to be measured from data with the embedding technique. In this technique μ+jets events are selected from data and the muon is replaced with a generator level tau lepton, forced to decay hadronically, and reconstructed. After the embedding of the tau lepton, the standard set of event selections is applied. Since no mass requirements are planned to be made, the W+jets and $t\bar{t}$ do not need to be separated.

Type	Uncertainty	Reference and $\int \mathcal{L}$
Isolated electrons, reconstruction and ID	3%	[2], 198 nb^{-1}
Isolated electrons, fake rate	5%	[2], 78 nb^{-1}
Isolated muons, reconstruction and ID	3%	[3], 198 nb^{-1}
Isolated muons, fake rate	negligible	[3], 0.47 nb^{-1}
Electromagnetic energy scale	0.9% / 2.2%	[4], 123 nb^{-1}
Jet energy scale	5/10%$\oplus$2%/η	[6], 123 nb^{-1}
Missing E$_T$ energy scale	10%	[9], 11.7 nb^{-1}
b-tagging, efficiency	19%	[10], 8 nb^{-1}
b-tagging, mistag rate	3-60%	[10], 12 nb^{-1}
Hadronic tau decays, energy scale	10%	estimate
Hadronic tau decays, reconstruction and ID	10%	estimate
Hadronic tau decays, jet$\rightarrow\tau$ fake-rate	20-30%	[11], 8.4 nb^{-1}
Luminosity measurement	11%	[13]
Underlying event	10%	[14]

Table 1: Summary of systematic uncertainties.

3.4 Drell-Yan background measurement for the di-leptonic final state

The $Z/\gamma^* \rightarrow ee/\mu\mu$ background can be measured from the data by counting the events where the di-lepton mass is within 15 GeV/c^2 of the Z mass [12]. The number of Z/γ^* with invariant di-lepton mass outside this mass window is determined from simulation and the measured number of events is corrected correspondingly. A conservative estimate of the systematic uncertainty of this method is 50%.

4. Summary

The current understanding of the systematic uncertainties for the light charged Higgs boson search via $t\bar{t}$ production is summarized in table 1. Estimates of the luminosity measurement [13] and underlying event uncertainty [14] are also shown in the table. Since the results shown are based on a very small amount of integrated luminosity, many measurements of the systematic uncertainties have been affected by statistical uncertainties, whose influence will decrease as more data is collected. The dedicated background measurements are still very much work in progress and will be evaluated when more data become available.

Acknowledgements

The author wishes to thank the Waldemar von Frenckell foundation for travel support.

References

[1] The CMS Collaboration, *The CMS experiment at the CERN LHC*, JINST 3:S08004 (2008).

[2] The CMS Collaboration, *Electron reconstruction and identification at $\sqrt{s}$ = 7 TeV*, CMS PAS EGM-10-004.

[3] The CMS Collaboration, *Measurements of Inclusive W and Z Cross Sections in pp Collisions at $\sqrt{s}$=7 TeV*, CMS PAS EWK-10-002.

[4] The CMS Collaboration, *Electromagnetic calorimeter calibration with 7 TeV data*, CMS PAS EGM-10-003.

[5] The CMS Collaboration, *Jet Performance in pp Collisions at $\sqrt{s}$=7 TeV*, CMS PAS JME-07-002.

[6] The CMS Collaboration, *Jet Performance in pp Collisions at $\sqrt{s}$=7 TeV*, CMS PAS JME-10-003.

[7] The D0 Collaboration, *Determination of the absolute jet energy scale in the D0 calorimeters*, Nucl. Inst. Meth. A 424 (1999) 323-394.

[8] The CDF Collaboration, *Determination of the jet energy scale at the Collider Detector at Fermilab*, Nucl. Inst. Meth. A 566 (2006) 375-412 [`hep-ex/0510047`].

[9] The CMS Collaboration, *Missing Transverse Energy Performance in Minimum-Bias and Jet Events from pp Collisions at $\sqrt{s}$=7 TeV*, CMS PAS JME-10-004.

[10] The CMS Collaboration, *Commissioning of b-jet identification with pp collisions at $\sqrt{s}$ = 7TeV*, CMS PAS BTV-10-001.

[11] The CMS Collaboration, *Tau Commissioning with 7-TeV data*, CMS PAS PFT-10-004.

[12] The CMS Collaboration, *Selection of Top-Like Events in the Dilepton and Lepton-plus-Jets Channels in Early 7 TeV Data*, CMS PAS TOP-10-004.

[13] The CMS Collaboration, *Measurement of CMS Luminosity*, CMS PAS EWK-10-004.

[14] The CMS Collaboration, *Measurement of the Underlying Event Activity at the LHC with $\sqrt{s}$=7 TeV*, CMS PAS QCD-10-010.