

CMS Draft Analysis Note

The content of this note is intended for CMS internal use and distribution only

2011/08/15

Head Id: 73281

Archive Id: 75367:75375

Archive Date: 2011/08/04

Archive Tag: trunk

Search for the Standard Model Higgs boson in the decay channel $H \rightarrow ZZ \rightarrow 2\ell 2b$

L. Borrello⁵, D. Bortoletto⁵, R. Castello⁴, G. Codispoti⁶, A. De Cosa³, F. Fabozzi³, J.P. Fernandez², J. Fernandez de Troconiz⁶, P. Garcia-Abia², O. Gonzalez², J. Hernandez², M. Kress⁵, L. Lista³, M. Narain¹, M. Segala¹, and M. Vidal⁵

¹ Brown University, Providence, RI, USA

² CIEMAT, Madrid, Spain

³ INFN Sezioni di Napoli, Napoli, Italy

⁴ INFN Sezioni di Torino, Torino, Italy

⁵ Purdue University, West Lafayette, IN, USA

⁶ Universidad Autonoma de Madrid, Madrid, Spain

Abstract

We present a search for the standard model Higgs boson through its decay to two Z bosons, subsequently decaying to two leptons (electrons or muons) and two b-quark jets. The analysis exploits the kinematics of the final state, consisting on two leptons and two b-jets, to select the Higgs boson signal in the mass range from 200 GeV to 600 GeV. The full analysis is run using 1 fb^{-1} of data collected by CMS in 2011.

This box is only visible in draft mode. Please make sure the values below make sense.

PDFAuthor: To be filled in

PDFTitle: Search for the Standard Model Higgs boson in the decay channel H to ZZ to 2l2b

PDFSubject: CMS Higgs boson search

PDFKeywords: CMS, LHC, physics, data analysis, Higgs boson

Please also verify that the abstract does not use any user defined symbols

1 Contents

2	1	Introduction	2
3	2	Data and Monte Carlo samples	2
4	2.1	Background Monte Carlo	2
5	2.2	Signal Monte Carlo	3
6	2.3	Data Samples	3
7	3	Event Selection	4
8	3.1	Preselection	5
9	3.2	B-tagging requirements	6
10	3.3	Missing Transverse Energy	6
11	3.4	Kinematic Selection	6
12	3.5	Angular Selection	8
13	4	Data-MC Comparison	8
14	5	Background Determination from Data	12
15	5.1	$t\bar{t}$ Background	12
16	5.2	Mistags	16
17	6	Correlated and Uncorrelated Systematic Uncertainties	17
18	6.1	Leptons selection and Trigger efficiency	17
19	6.2	Leptons energy scale and resolution	20
20	6.3	Evaluation of systematic uncertainties	20
21	6.4	b -tagging Systematics	21
22	7	Signal Optimization	23
23	7.1	Angular Optimization	23
24	8	Statistical Analysis and Results	26
25	9	Conclusions	27
26	10	Update for Lepton-Photon-2011	27
27	10.1	b -tagging Systematics	27
28	10.2	Leptons selection and Trigger efficiency	28
29	10.3	Leptons energy scale and resolution	30
30	10.4	Evaluation of systematic uncertainties	30
31	A	B-tagging optimization	35
32	B	Z+HF Background Checks	37
33	B.1	Z+HF Background from data	37
34	B.2	Z+HF Shape Systematics from MC	38
35	C	Kinematic Optimization	40
36	D	Mistag Efficiencies and Scale Factors	43

37 1 Introduction

38 The Higgs boson is an essential element of the Standard Model (SM) of particles and their inter-
 39 actions explaining the origin of mass and playing a key role in the mechanism of electroweak
 40 symmetry breaking. Precision measurements of the W boson and top quark masses together
 41 with the Z pole variables constrain the mass of the Higgs boson which is a free parameter of
 42 the theory: currently the 95% one-sided confidence level upper limit on m_H is 158 GeV. Direct
 43 searches for the Higgs at LEP-II [1] set a 95% C.L. lower limit on m_H of 114.4 GeV, while the re-
 44 gion between 158 GeV and 175 GeV has been excluded by the Tevatron experiments [2, 3]. The
 45 95% C.L. upper limit on the Higgs mass increases from 158 GeV to 185 GeV when the results
 46 from the direct searches are included [4].

47 The CMS collaboration is performing searches for the Higgs in several decay modes. This
 48 comprehensive effort aims at gaining sensitivity over a large range of Higgs masses by com-
 49 bining many different analyses. We expect that this effort will finally explore the region with
 50 $m_H \geq 2m_Z$ which is not probed at the Tevatron. In this note we report a detailed study of the
 51 prospect for the search for the Higgs boson in $H \rightarrow ZZ$ when one of Z decays as $Z \rightarrow \ell^-\ell^+$
 52 and the other as $Z \rightarrow b\bar{b}$. The dominant background is $Z+jets$ production. Other minor back-
 53 grounds are $t\bar{t}$ and diboson production. The data and Monte Carlo sample used in this analysis
 54 are described in Section 2. The analysis strategy including the signal to background optimiza-
 55 tion are discussed in Sections 3 and 7. Contamination is estimated from background-enriched
 56 control samples. Methods to provide a data-driven derivation of the background is presented
 57 in Section 5. A detailed discussion of the systematic errors, including the impact of the pile-up
 58 events expected in the 2011 LHC run, as well as the modeling of the b -tagging algorithms in
 59 CMS, is described in Section 6. The strategy and the experimental methods are validated using
 60 1 fb^{-1} data collected by the CMS experiment in 2011 as shown in Section 8.

61 2 Data and Monte Carlo samples

62 For this round of the analysis we are using 1 fb^{-1} of data collected in 2011. collected in 2011.
 63 The prompt reconstruction pass is used in order to compare with the MC samples using the
 64 same reconstruction software release.

65 The data and Monte Carlo (MC) sample used in this analysis are described herein.

66 The MC samples, described in detail in the next sections, have been produced superimposing
 67 a number of minimum bias interactions to simulate the pile-up interactions in the event data.

68 2.1 Background Monte Carlo

69 The dominant background in this analysis is the inclusive Z production together with jets,
 70 in particular with a pair of b -jets. We used the ALPGEN generator [5], which is expected to
 71 reproduce well the multiplicity of jets in the events. The inclusive ALPGEN $Z+jets$ samples
 72 (Table 2) do not include the production of hard b quarks. Therefore, dedicated samples with Z
 73 + $b\bar{b}$ pair production have been generated (Table 1). We have developed a method to properly
 74 combine the samples avoiding event double-counting.

75 Other minor backgrounds are $t\bar{t}$, ZZ , WW , and WZ production. The samples used are listed in
 76 Table 3.

Name	σ LO(NLO) [pb]	Lumi LO(NLO) [fb^{-1}]
ZBB0JetsToLNu_TuneZ2_7TeV-alpgen-tauola	1.703(2.912)	204.0(119.3)
ZBB1JetsToLNu_TuneZ2_7TeV-alpgen-tauola	0.962(1.645)	210.6(123.1)
ZBB2JetsToLNu_TuneZ2_7TeV-alpgen-tauola	0.3639(0.6223)	29.79(17.4)
ZBB3JetsToLNu_TuneZ2_7TeV-alpgen-tauola	0.1598(0.2733)	68.20(39.9)
ZCC0JetsToLNu_TuneZ2_7TeV-alpgen-tauola	1.707(2.919)	256.6(150.1)
ZCC1JetsToLNu_TuneZ2_7TeV-alpgen-tauola	0.9526(1.6289)	193.5(113.2)
ZCC2JetsToLNu_TuneZ2_7TeV-alpgen-tauola	0.3659(0.6257)	29.34(17.16)
ZCC3JetsToLNu_TuneZ2_7TeV-alpgen-tauola	0.1643(0.2810)	61.94(36.2)

Table 1: Z+Heavy Flavour MC samples.

Name	σ LO(NLO) [pb]	Lumi LO(NLO) [fb^{-1}]
Z0Jets_TuneZ2_7TeV-alpgen-tauola	1929(2450)	0.74(0.59)
Z1Jets_ptZ-0to100_TuneZ2_7TeV-alpgen-tauola	380(483)	39.1(30.9)
Z1Jets_ptZ-100to300_TuneZ2_7TeV-alpgen-tauola	8.72(11.1)	1375.4(1080.6)
Z1Jets_ptZ-300to800_TuneZ2_7TeV-alpgen-tauola	7.386×10^{-2} (9.38e-02)	76945(60588)
Z1Jets_ptZ-800to1600_TuneZ2_7TeV-alpgen-tauola	1.374×10^{-4} (1.745x10 $^{-4}$)	3657328(2879753)
Z2Jets_ptZ-0to100_TuneZ2_7TeV-alpgen-tauola	104(132)	12.9(10.1)
Z2Jets_ptZ-100to300_TuneZ2_7TeV-alpgen-tauola	8.53(10.8)	131.7(104.7)
Z2Jets_ptZ-300to800_TuneZ2_7TeV-alpgen-tauola	0.1151(0.1462)	5693(4482)
Z2Jets_ptZ-800to1600_TuneZ2_7TeV-alpgen-tauola	3.023×10^{-4} (3.839x10 $^{-4}$)	205901(162135)
Z3Jets_ptZ-0to100_TuneZ2_7TeV-alpgen-tauola	22.9(29.1)	14.1(11.1)
Z3Jets_ptZ-100to300_TuneZ2_7TeV-alpgen-tauola	3.95(5.02)	74.4(58.5)
Z3Jets_ptZ-300to800_TuneZ2_7TeV-alpgen-tauola	8.344×10^{-2} (0.106)	2198(1730)
Z3Jets_ptZ-800to1600_TuneZ2_7TeV-alpgen-tauola	2.480×10^{-4} (3.15x10 $^{-4}$)	98544(77584)
Z4Jets_ptZ-0to100_TuneZ2_7TeV-alpgen-tauola	4.6(5.84)	13.6(18.7)
Z4Jets_ptZ-100to300_TuneZ2_7TeV-alpgen-tauola	1.30(1.65)	93.7(73.7)
Z4Jets_ptZ-300to800_TuneZ2_7TeV-alpgen-tauola	3.935×10^{-2} (4.997x10 $^{-2}$)	1163(915)
Z4Jets_ptZ-800to1600_TuneZ2_7TeV-alpgen-tauola	1.394×10^{-4} (1.77x10 $^{-4}$)	93859(73921)
Z5Jets_ptZ-0to100_TuneZ2_7TeV-alpgen-tauola	1.135(1.441)	27.4(21.6)
Z5Jets_ptZ-100to300_TuneZ2_7TeV-alpgen-tauola	0.4758(0.6043)	64.4(50.7)
Z5Jets_ptZ-300to800_TuneZ2_7TeV-alpgen-tauola	1.946×10^{-4} (2.471x10 $^{-4}$)	1278(1006)
Z5Jets_ptZ-800to1600_TuneZ2_7TeV-alpgen-tauola	7.195×10^{-5} (9.138x10 $^{-5}$)	440111(346531)

Table 2: Z+jets MC samples.

2.2 Signal Monte Carlo

The signal MC samples used in this analysis are summarized in Table 4. A dedicated simulation program [6] has been used to generate signal events which generates spin correlations through the $2 \rightarrow 1 \rightarrow 2 \rightarrow 4$ process. PYTHIA [7] is used for hadronization and further interface to the CMS software (CMSSW) framework.

2.3 Data Samples

We are using the double lepton primary datasets corresponding to 1 fb^{-1} of data collected in 2011. We also use the Muon-Electron primary dataset for the determination of the $t\bar{t}$ background from data. The data samples are listed in Table 5.

Name	σ LO(NLO) [pb]	Lumi LO(NLO) [fb^{-1}]
TTTo2L2Nu2B_7TeV-powheg-pythia6	15.86(16.7)	63.04(59.87)
ZZtoAnything_TuneZ2_7TeV-pythia6-tauola	4.30(5.9)	490.7(357.4)
WZtoAnything_TuneZ2_7TeV-pythia6-tauola	10.4(18.3)	202.7(115.2)
WWtoAnything_TuneZ2_7TeV-pythia6-tauola	27.8(42.9)	74.2(48.1)

Table 3: $t\bar{t}$, ZZ, WW, and WZ MC samples.

Sample	$\sigma \times \text{BR}$ at NLO [pb]	Lumi [fb^{-1}]
SMHiggsToZZTo2L2Q_M-200_7TeV-jhu-pythia6	0.1441	757.2
SMHiggsToZZTo2L2Q_M-250_7TeV-jhu-pythia6	0.1063	1034.6
SMHiggsToZZTo2L2Q_M-300_7TeV-jhu-pythia6	0.0802	1360.4
SMHiggsToZZTo2L2Q_M-350_7TeV-jhu-pythia6	0.0769	1418.4
SMHiggsToZZTo2L2Q_M-400_7TeV-jhu-pythia6	0.0553	1987.8
SMHiggsToZZTo2L2Q_M-450_7TeV-jhu-pythia6	0.0353	3113.5
SMHiggsToZZTo2L2Q_M-500_7TeV-jhu-pythia6	0.0224	4591.8
SMHiggsToZZTo2L2Q_M-1000_7TeV-jhu-pythia6		

Table 4: Signal Monte Carlo samples used in the analysis. Only muon and electron final states in the leptonic Z decay are considered, with a total branching ratio $B(ZZ \rightarrow 2\ell 2j) = 2 \times 0.067 \times 0.699 = 0.0937$.

Dataset name
/DoubleMu/Run2011A-May10ReReco-v1/AOD
/DoubleElectron/Run2011A-May10ReReco-v1/AOD
/MuEG/Run2011A-May10ReReco-v1/AOD
/DoubleMu/Run2011A-PromptReco-v4/AOD
/DoubleElectron/Run2011A-PromptReco-v4/AOD
/MuEG/Run2011A-PromptReco-v4/AOD

Table 5: 2011 data samples.

3 Event Selection

- 86 The main signature of $H \rightarrow ZZ \rightarrow 2\ell 2j$ is the presence of a lepton pair and a jet pair having
 87 an invariant mass around the Z mass. The width of the invariant mass is dominated by the Z
 88 intrinsic width for the di-lepton pair, and by the jet resolution for the jet pair. The Higgs boson
 89 mass can be reconstructed as the invariant mass of the four objects in the final state, $m_{\ell\ell jj}$.
 90
- 91 The main background to this channel is the SM production of Z in association with jets. This
 92 background can be reduced requiring the identification (tagging) of the two final jets as b jets,
 93 hence restricting the Higgs decay to $H \rightarrow ZZ \rightarrow 2\ell 2b$. Production of Z boson pairs and $t\bar{t}$ also
 94 contribute to the total background. ZZ production can be reduced exploiting the significant
 95 boost of the Z bosons coming from the Higgs decay, when the Higgs mass is significantly larger
 96 than $2m_Z$. $t\bar{t}$ events have a significant missing transverse energy (\cancel{E}_T) due to the presence of
 97 the neutrinos in the final state that is not present in the Higgs signal.
- 98 In the following sub-sections we describe the selection of leptons and b-jets, then in a dedicated
 99 section we present the kinematic selection adopted to reduce the SM backgrounds.

3.1 Preselection

We select events with 2 leptons and 2 jets according to the lepton and jet selection described in the following Sections. We don't apply b-tagging requirements at the pre-selection stage. We also select events with only one jet for background control purposes. All events are required to have a di-lepton invariant mass $m_{\ell\ell}$ greater than 50 GeV.

Isolation is an important variable to distinguish leptons from EW processes from the ones from QCD background and b -decays. We define isolation variables for the tracker, electromagnetic and hadronic calorimeters as follows: $I_{\text{trk}} = \sum_{\text{tracks}} p_T$, $I_{\text{ECAL}} = \sum_{\text{ECAL}} E_T$, $I_{\text{HCAL}} = \sum_{\text{HCAL}} E_T$, where the sums are performed on all objects falling within a cone $\Delta R = \sqrt{\Delta\eta^2 + \Delta\phi^2} = 0.3$ around the lepton candidate momentum direction. The energy deposits and the track associated to the lepton candidate are excluded from the sums. We ensure lepton isolation applying cuts on the relative combined isolation, defined as:

$$R = (I_{\text{ECAL}} + I_{\text{HCAL}} + I_{\text{trk}}) / p_T.$$

3.1.1 Muons

We select global muon pairs with p_T greater than 20 GeV and 40 GeV, for the lower and greater momentum lepton respectively, and $|\eta| < 2.4$. We require the following quality selection recommended in previous studies [8]:

- Normalized χ^2 of the global track < 10
- At least 10 tracker hits
- At least 1 pixel hit
- At least 1 hit in the muon subsystem
- At least 2 matched muon stations
- Transverse impact parameter with respect to the primary vertex, $d_{xy} < 0.02$ cm
- Combined relative isolation $R < 0.15$

The transverse impact parameter cut has been tightened w.r.t. Ref. [8] in order to reduce the possible background coming from muons produced in the decays of heavy flavour mesons.

3.1.2 Electrons

We select GSF electrons with p_T greater than 20 GeV and 40 GeV, for the lower and greater momentum lepton respectively, and $|\eta| < 2.5$ that fulfill the following conditions [9]:

- Electron rejection if $1.4442 < |\eta| < 1.566$ (no electrons in the EB/EE gap)
- WP95 requirements with combined isolation $R < 0.15$
- At least one electron passing WP80 with combined isolation $R < 0.15$

WP95 and WP80 are defined in Ref. [9], and refer to two sets of cuts on electron's discriminating variables and isolation variables I_{trk}/E_T , I_{ECAL}/E_T and I_{HCAL}/E_T , that have a selection efficiency of approximately 95% and 80% respectively.

3.1.3 Jets

Particle Flow jets [10] reconstructed with the anti- k_T [11] algorithm with radius $R = 0.5$ are considered in this analysis. Selected jets must have $p_T > 30$ GeV and $|\eta| < 2.4$.

130 Jet energy corrections are applied at L2 and L3. Corrections at L1, taking into account pile-up
 131 effects, are also included as default in the 2011 data.

132 The selected jets (j) are cleaned-up vetoing the presence of selected leptons (ℓ) within a cone
 133 around the jet by applying the cut $\Delta R(\ell, j) > 0.5$.

134 3.2 B-tagging requirements

135 After the pre-selection described in Section 3.1, heavy flavor tagging requirements are applied.
 136 Since our signal has two b -jets in the final state, the double-tag selection was found to be opti-
 137 mal improving the signal to background ratio. Therefore, we define a double tag region requir-
 138 ing at least two b -tagged jets in the event. An exclusive tagged region is defined by requiring
 139 that one of the two jets is tag as a b -jet.

140 These two regions are the baseline selection for the analysis.

141 The tagging algorithm used as default in the analysis is the Track Counting High Efficiency (TCHE) [12].
 142 For the double tag region a combination of two working points is used, requiring one of the
 143 jets being tag with the medium working point (TCHEM), and the other jet with the loose work-
 144 ing point (TCHEL). For the single tag region one of the jets is required to be TCHEL, and the
 145 other jet is required to have a b -tag discriminator value $<$ TCHEM. These working points were
 146 chosen based on the tagging studies described in Appendix A.

147 Other algorithms, like Simple Secondary Vertex (SSV), performance variations (High Efficiency
 148 or High Purity), and working points (Loose, Medium, and Tight) are also checked but they are
 149 not used for the baseline results.

150 The b -tagging efficiencies, scale factors data/Monte Carlo, and uncertainties can be found in
 151 the following reference [13] provided by the BTV POG. The numbers for our default working
 152 points are summarized in table 6. Scale factors data/Monte Carlo are also needed for light jets
 153 misidentified as b -jets. These factors are computed as a function of p_T and η of the jets and are
 154 listed in Appendix D.

Tagger	ϵ_b^{data}	ϵ_b^{MC}	SF_b
TCHEM	$\sim 0.7 \pm -$	$\sim 0.7 \pm -$	0.9 ± 0.13
TCHEL	$\sim 0.8 \pm -$	$\sim 0.8 \pm -$	0.9 ± 0.13

Table 6: Tagging efficiency for b jets in data and, Monte Carlo, and their ratio from the BTV POG.(Work ongoing in the BTV POG)

155 3.3 Missing Transverse Energy

156 At the first stage of the analysis, requiring 2 b -tagged jets and no further selection, one of the
 157 main sources of background is the $t\bar{t}$ production. These $t\bar{t}$ events contain two real b -jets and a
 158 significant amount of \cancel{E}_T originated from neutrinos.

159 We reduce substantially the presence of $t\bar{t}$, while keeping the signal, requiring particle flow
 160 MET-Significance < 10 . The MET-Significance distribution after double b -tagging requirements
 161 is shown in Figure 1.

162 3.4 Kinematic Selection

163 A cut-based analysis, as described in Section 7, is performed as a cross check of the main pro-
 164 cedure based on an angular Likelihood Discriminant. The kinematic selection described in this

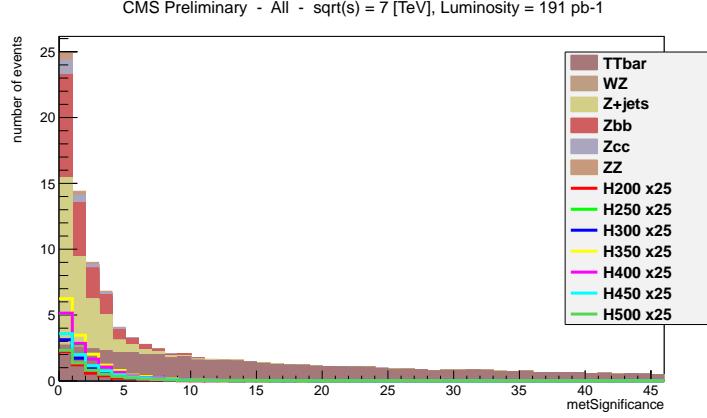


Figure 1: MET-Significance distribution for several Higgs masses (lines) and backgrounds (stacked histograms).

section is used in the cut-based approach.

The signal kinematic selection depends on the unknown Higgs boson mass. Since we search the whole accessible Higgs mass range, in order to avoid the complications due to the use of different selections for different Higgs mass hypotheses, we apply kinematic cuts that depend on the reconstructed Higgs mass, $m_{\ell\ell jj}$. In most of the cases, the cuts are linear functions of $m_{\ell\ell jj}$.

Figure 2 shows the distribution of the measured p_T of the di-lepton pair for three signal mass hypotheses and for the Z+jets background from b and c quarks. The p_T tends to be larger for larger values of m_H . A linear cut is applied, according to:

$$p_T(\ell\ell) > \left(150 - \frac{6}{15}(500 - m_{\ell\ell jj}) \right). \quad (1)$$

This means that a cut of $p_T(\ell\ell) > 50$ GeV is applied when $m_{\ell\ell jj} = 250$ GeV, and $p_T(\ell\ell) > 150$ GeV when $m_{\ell\ell jj} = 500$ GeV.

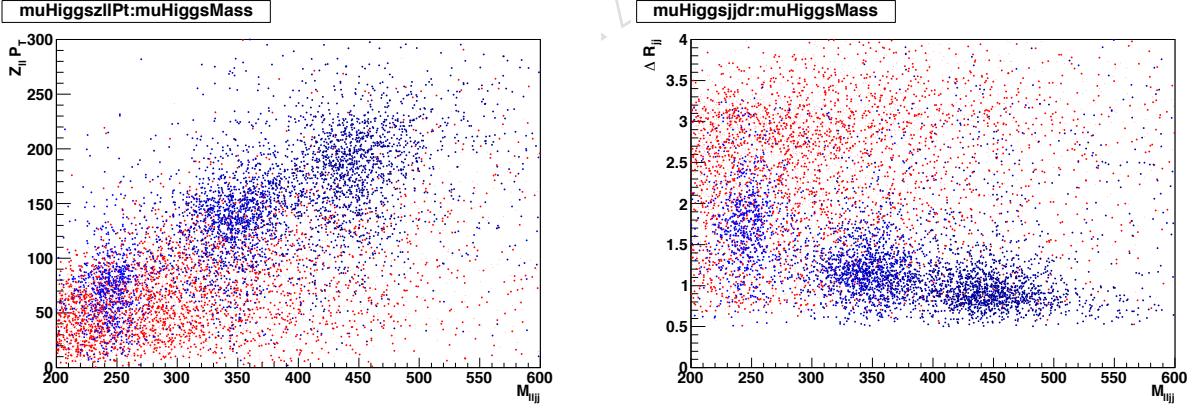


Figure 2: Distribution of the di-lepton pair transverse momentum (left) and di-jet pair angular separation (right), for signal (blue, $m_H = 250, 350$ and 450 GeV) and for the Z+jets background from c and b quarks (red).

The angular separation of jets, $\Delta R(jj)$, tends to be larger if the Z which decays in two b quarks is boosted. Figure 2 shows the distribution of $\Delta R(jj)$ for three signal mass hypotheses and for the Z+jets background from b and c quarks. High values of m_H yield more collinear jets. In

order to reduce the Z+jets background, we apply the linear cut:

$$\Delta R(jj) < \left(1.9 + \frac{1}{500}(250 - m_{\ell\ell jj}) \right). \quad (2)$$

- 172 The applied cut is $\Delta R(jj) < 1.9$ for $m_{\ell\ell jj} = 250$ GeV and $\Delta R(jj) < 1.4$ for $m_{\ell\ell jj} = 500$ GeV
 173 Figure 3 shows the distribution of the invariant mass $m_{\ell\ell jj}$ for the candidates surviving the
 174 final selection. The plot is normalized to signal and background expectations for an integrated
 175 luminosity of 1 fb^{-1} .

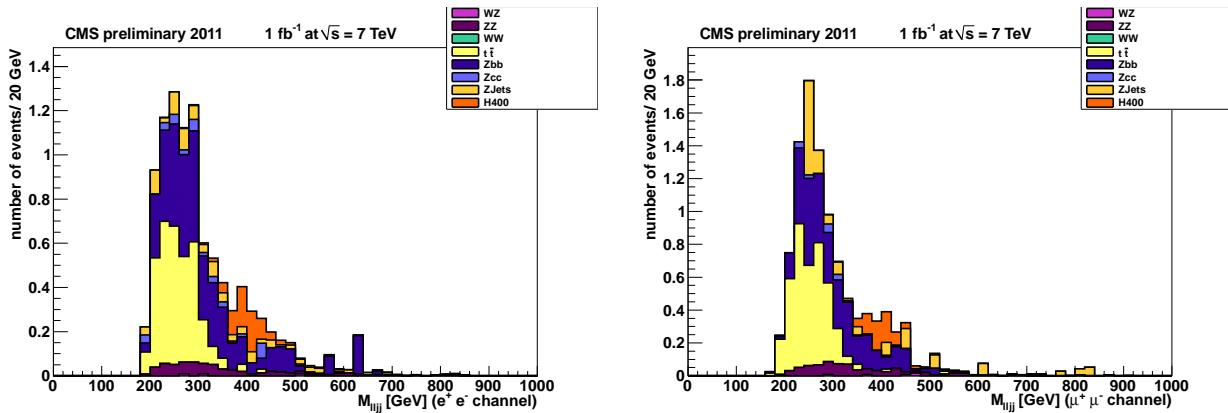


Figure 3: Distribution of the invariant mass $m_{\ell\ell jj}$ for the events passing the final selection. Higgs signal with a mass of 400 GeV is shown in orange, superimposed to the SM backgrounds: WW, WZ, $t\bar{t}$, Z+jets: heavy, and light quarks and ZZ.

176 3.5 Angular Selection

- 177 Taking advantage of the angular information in the production and decay chain of the pro-
 178 cesses contributing in this analysis, an helicity Likelihood Discriminant (LD) was built to sep-
 179 arate signal and background. The method and its validation are described in Reference [14].

180 4 Data-MC Comparison

- 181 Every time new data is available we check if the CMS simulation reproduce it in the pre-
 182 selection region described in Section 3.1. As an example, the m_{jj} , m_{ll} , and MET-Significance are
 183 shown in Figures 4 to 6 for a 1 fb^{-1} of data.
 184 Distributions of the angular variables used in the helicity LD, and the helicity LD itself, are also
 185 checked as shown in Figures 7 to 12. These distributions are checked after preselection and
 186 tagging requirements.

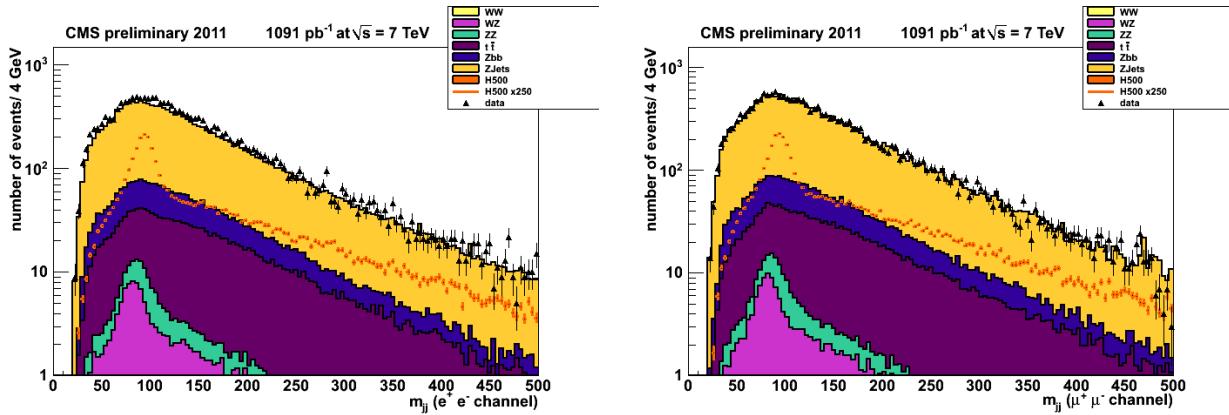


Figure 4: Distribution of the m_{jj} for electron (left) and muon (right) events in $1 fb^{-1}$.

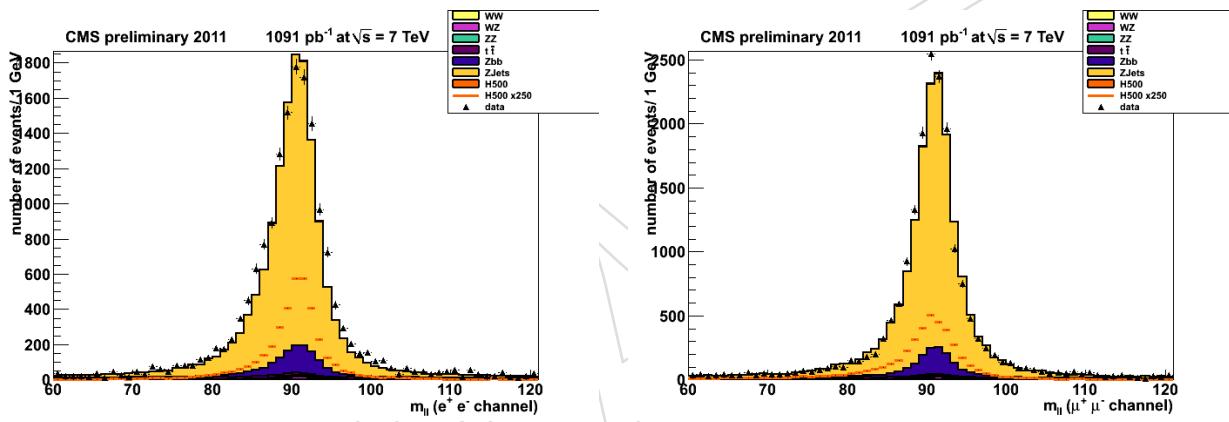


Figure 5: Distribution of the m_{ll} for electron (left) and muon (right) events in $1 fb^{-1}$.

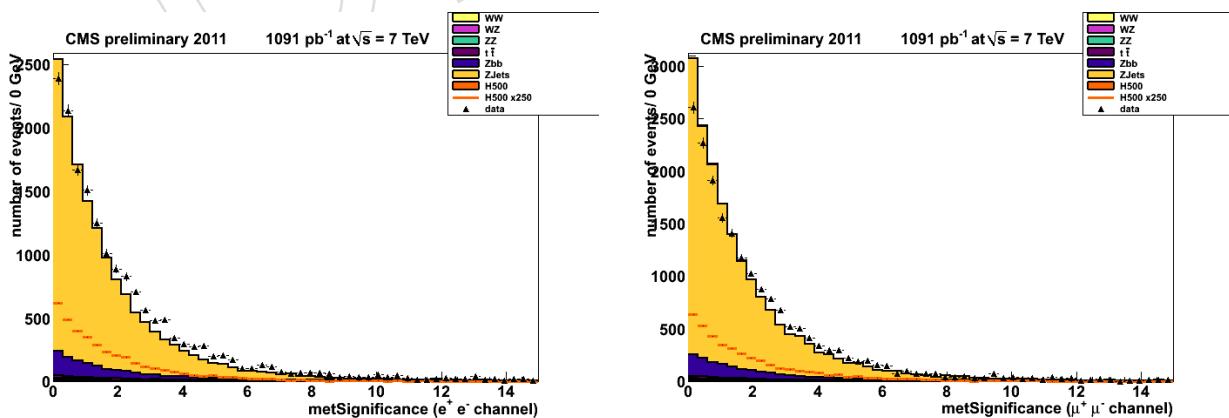


Figure 6: Distribution of the MET-Significance for electron (left) and muon (right) events in $1 fb^{-1}$.

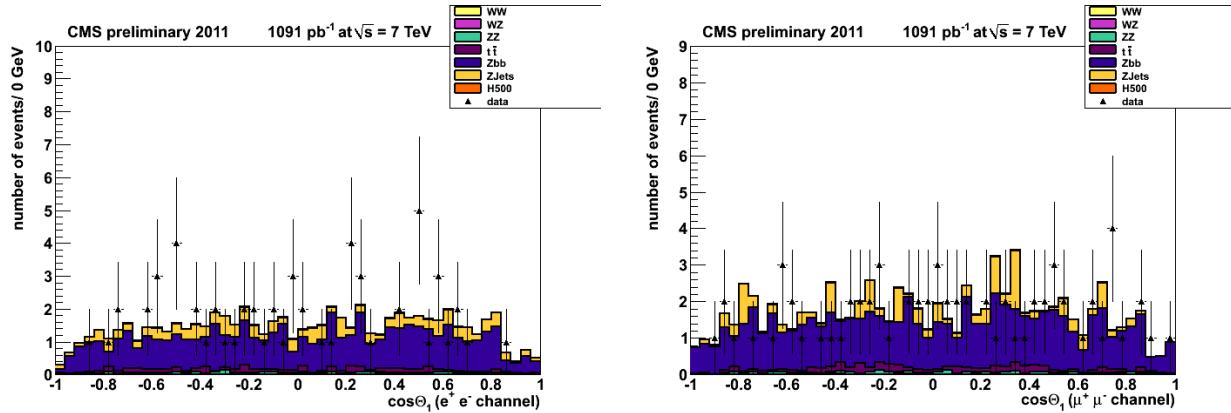


Figure 7: Distribution of the $\cos\theta_1$ for electron (left) and muon (right) events in $1 fb^{-1}$.

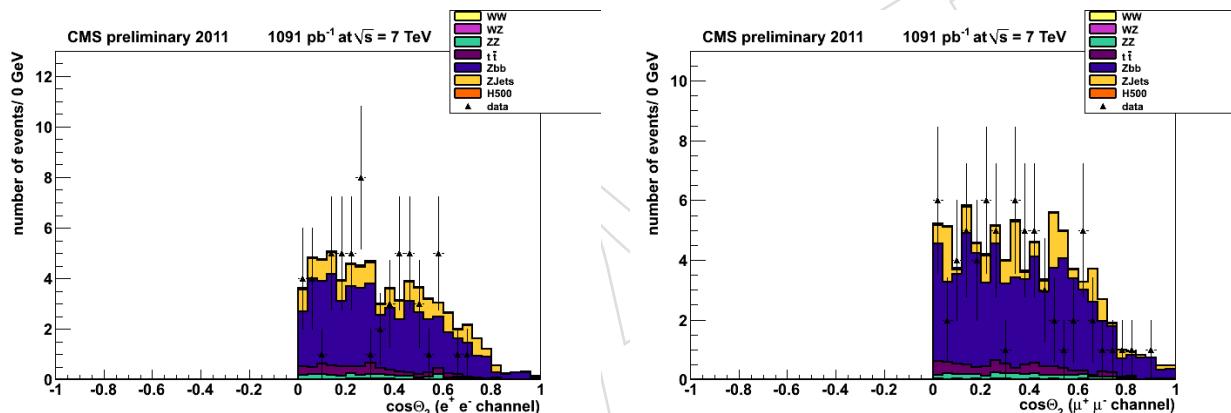


Figure 8: Distribution of the $\cos\theta_2$ for electron (left) and muon (right) events in $1 fb^{-1}$.

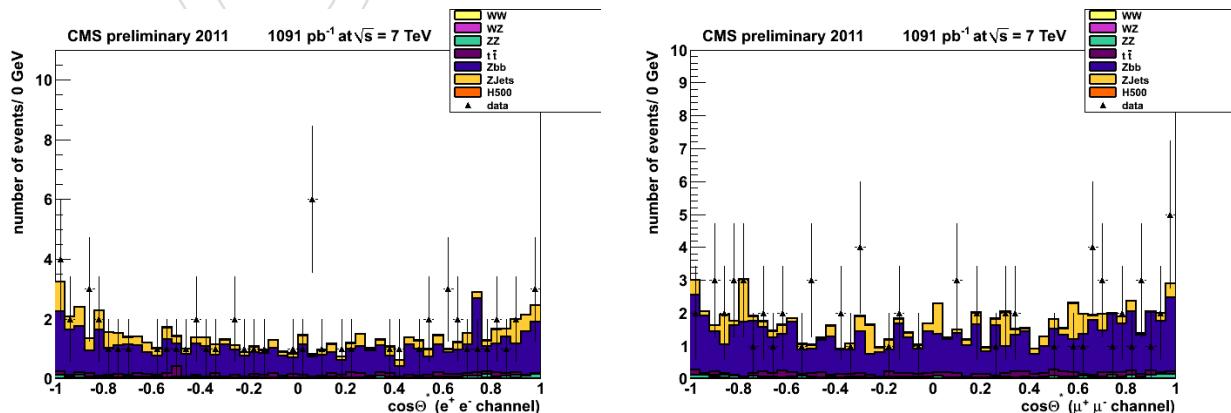


Figure 9: Distribution of the $\cos\theta^*$ for electron (left) and muon (right) events in $1 fb^{-1}$.

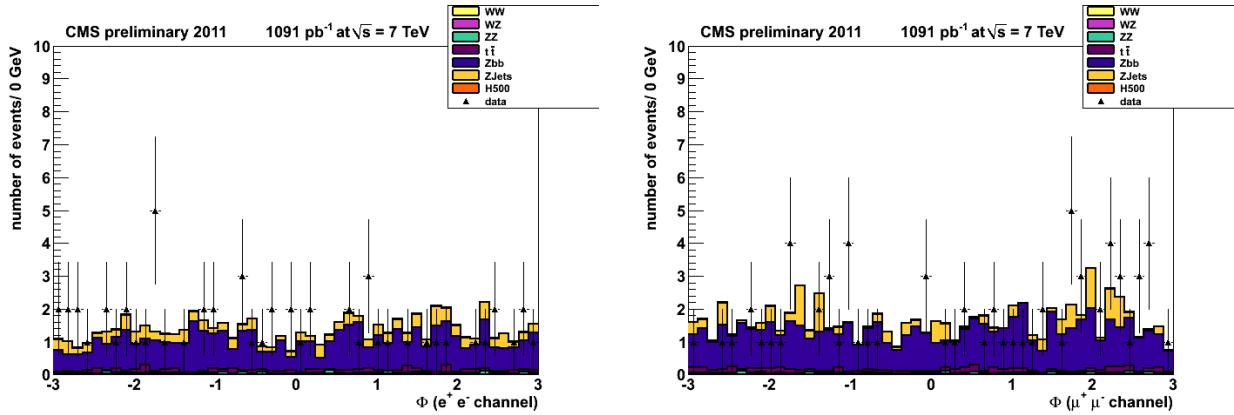


Figure 10: Distribution of the ϕ for electron (left) and muon (right) events in $1 fb^{-1}$.

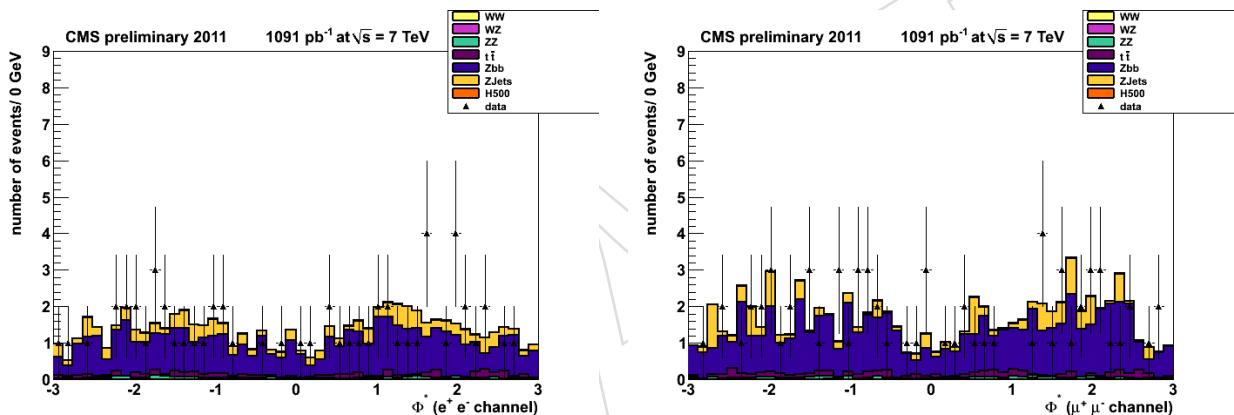


Figure 11: Distribution of the ϕ^* for electron (left) and muon (right) events in $1 fb^{-1}$.

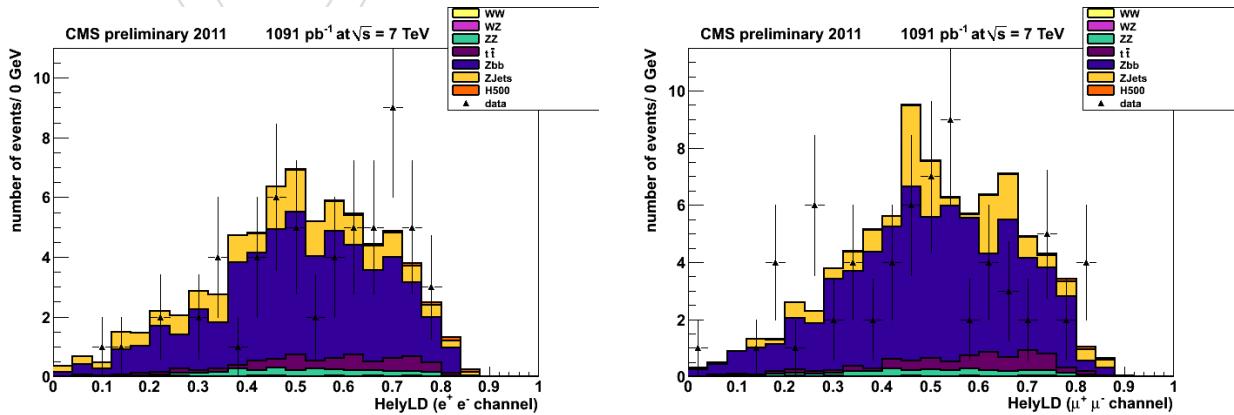


Figure 12: Distribution of the helicity LD for electron (left) and muon (right) events in $1 fb^{-1}$.

187 5 Background Determination from Data

188 The MC prediction of yields and shapes of the various relevant backgrounds are affected by
 189 theoretical uncertainties. We define several data-driven methods to determine backgrounds
 190 from background-enriched regions obtained by inverting some of the selection criteria.

191 The main procedure for background estimation from side bands is extensively described in the
 192 complementary Analysis Note [14].

193 5.1 $t\bar{t}$ Background

194 The $t\bar{t}$ background is estimated from the data using $e\mu$ events passing the same cuts as the
 195 signal. This method accounts for other small backgrounds (as $WW + \text{jets}$, $Z \rightarrow \tau\tau + \text{jets}$, single
 196 top, fakes) where the lepton flavour symmetry can be invoked as well.

197 Top-pair Monte Carlo studies show that the $e\mu$ vs. $ee + \mu\mu$ symmetry works very well at the
 198 level of the shapes of the distributions of all considered variables. Also, the relative event
 199 normalization is consistent with one, within the MC finite statistical errors. For instance, the
 200 $e\mu/(ee + \mu\mu)$ relative event normalizations are 0.997 ± 0.019 after selection and kinematical
 201 cuts, and 1.000 ± 0.034 after b-tagging.

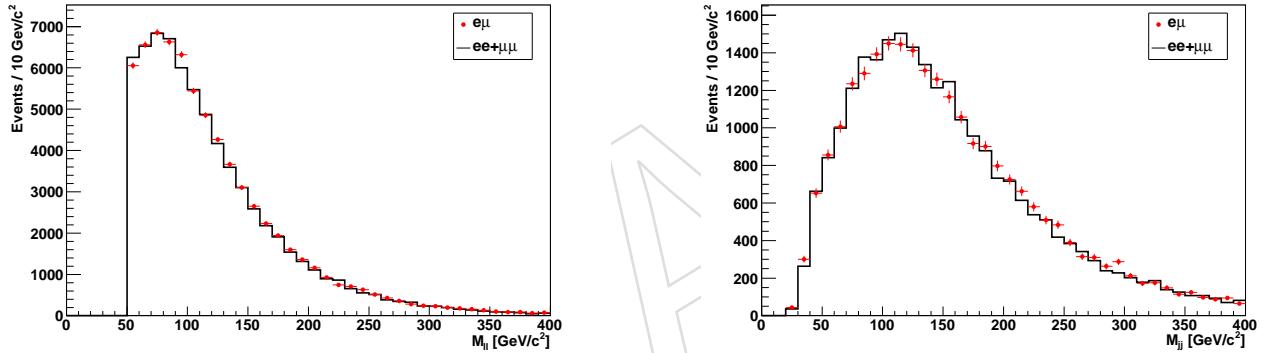


Figure 13: (Left) Top MC dilepton invariant mass for events satisfying $M_{ll} > 50 \text{ GeV}/c^2$. (Right) Top MC dijet invariant mass for events satisfying $71 \text{ GeV}/c^2 < M_{ll} < 111 \text{ GeV}/c^2$.

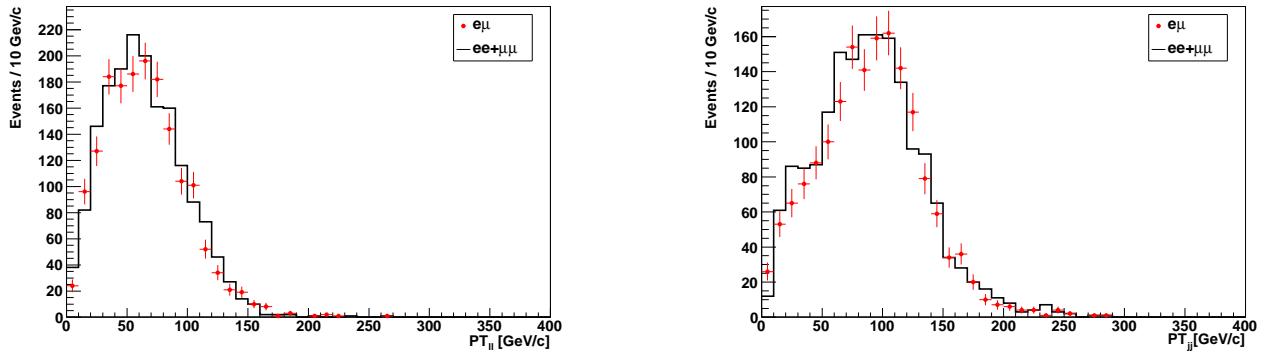


Figure 14: Top MC dilepton transverse momentum (left) and dijet transverse momentum (right) for events satisfying $71 \text{ GeV}/c^2 < M_{ll} < 111 \text{ GeV}/c^2$ and $71 \text{ GeV}/c^2 < M_{jj} < 111 \text{ GeV}/c^2$. Both jets have been b-tagged using the TCHEM prescription.

202 Figures 13-15 show a comparison of the $ee + \mu\mu$ and $e\mu$ top MC distributions of several relevant variables, for events with at least two leptons and two jets passing selection cuts. Only

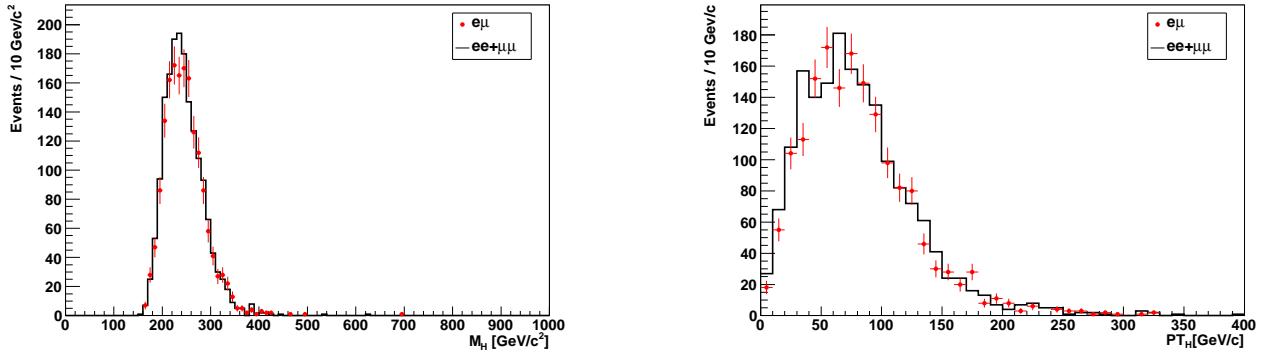


Figure 15: Top MC dilepton+dijet “Higgs” invariant mass (left) and transverse momentum (right) for events satisfying $71 \text{ GeV}/c^2 < M_{ll} < 111 \text{ GeV}/c^2$ and $71 \text{ GeV}/c^2 < M_{jj} < 111 \text{ GeV}/c^2$. Both jets are b-tagged using the TCHEM selection.

204 the hardest- $\sum P_T$ dilepton combination, and the dijet combinations with largest TCHEM dis-
 205 criminator values are considered. Extra cuts are detailed in the captions. The normalization is
 206 arbitrary.

Cuts	Pythia Top MC	Total Pythia MC	$e\mu$ data
$M_{ll} > 50 \text{ GeV}/c^2$	960	1043	1167
$70 \text{ GeV}/c^2 < M_{ll} < 110 \text{ GeV}/c^2$	339	362	415
$70 \text{ GeV}/c^2 < M_{jj} < 110 \text{ GeV}/c^2$	73.1	78.5	107
≥ 1 TCHEM b-tagged jet	59.9	62.1	86
≥ 1 TCHEL & ≥ 1 TCHEM	34.5	35.2	42
≥ 2 TCHEM b-tagged jet	24.4	24.8	32
MET Significance < 10	6.3	6.5	8

Table 7: Comparison of 2011 $e\mu$ data to $e\mu$ Pythia top MC event yields, corresponding to an integrated luminosity of 868 pb^{-1} . “Total MC” contains the top, WW, $Z \rightarrow \tau\tau$, and single top MC contributions. Every cut in a line assumes all cuts in lines above it.

207 The 2011 $e\mu$ data compares reasonable well to the sum of the $e\mu$ Herwig top MC prediction
 208 and other small backgrounds (WW, $Z \rightarrow \tau\tau$, single top). Table 7 compares the event yields for
 209 Pythia MC. Table 8 compares the event yields for Herwig MC, and Figures 16-20 compare sev-
 210 eral distributions, for events with at least two leptons and and two jets passing selection cuts.
 211 Only the hardest- $\sum P_T$ dilepton combination, and the dijet combination with largest TCHEM
 212 discriminator values are considered. Again other extra cuts are detailed where appropriate.

213 After kinematical selection, the contribution of the small backgrounds is always within one
 214 sigma of the data statistical error. One also notes that differences between Pythia and Herwig
 215 top MC predictions are important. Estimation of the top background using the data is needed.

Cuts	Herwig Top MC	Total Herwig MC	$e\mu$ data
$M_{ll} > 50 \text{ GeV}/c^2$	1022	1105	1167
$70 \text{ GeV}/c^2 < M_{ll} < 110 \text{ GeV}/c^2$	361	383	415
$70 \text{ GeV}/c^2 < M_{jj} < 110 \text{ GeV}/c^2$	79.3	84.6	107
≥ 1 TCHEM b-tagged jet	66.6	68.9	86
≥ 1 TCHEL & ≥ 1 TCHEM	37.6	38.2	42
≥ 2 TCHEM b-tagged jet	28.0	28.5	32
MET Significance < 10	7.1	7.3	8

Table 8: Comparison of 2011 $e\mu$ data to $e\mu$ Herwig top MC event yields, corresponding to an integrated luminosity of 868 pb^{-1} . “Total MC” contains the top, WW , $Z \rightarrow \tau\tau$, and single top MC contributions. Every cut in a line assumes all cuts in lines above it.

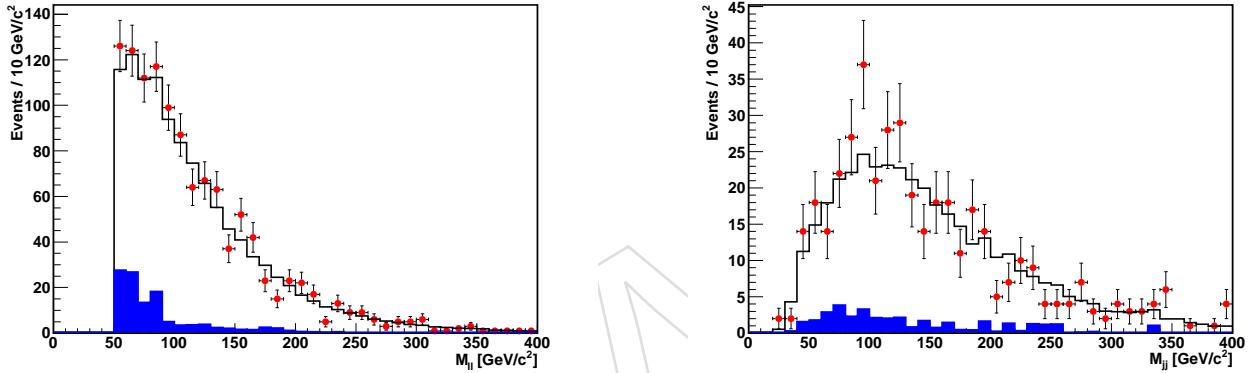


Figure 16: Comparison of 2011 $e\mu$ data (points) to to Herwig top MC (white; other small backgrounds, blue), corresponding to an integrated luminosity of 868 pb^{-1} . (Left) Dilepton invariant mass for events satisfying $M_{ll} > 50 \text{ GeV}/c^2$. (Right) Dijet invariant mass for events satisfying $71 \text{ GeV}/c^2 < M_{ll} < 111 \text{ GeV}/c^2$.

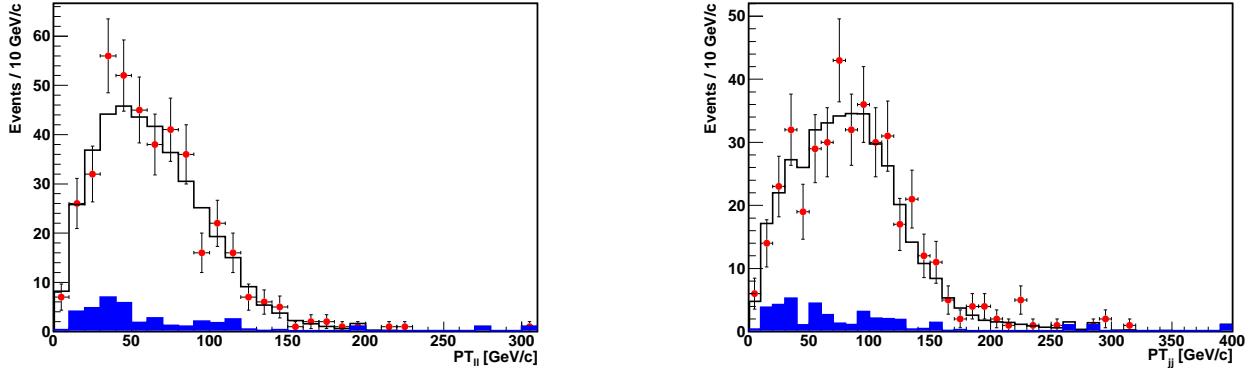


Figure 17: Comparison of 2011 $e\mu$ data (points) to to Herwig top MC (white; other small backgrounds, blue), corresponding to an integrated luminosity of 868 pb^{-1} . Dilepton transverse momentum (left) and dijet transverse momentum (right) for events satisfying $71 \text{ GeV}/c^2 < M_{ll} < 111 \text{ GeV}/c^2$.

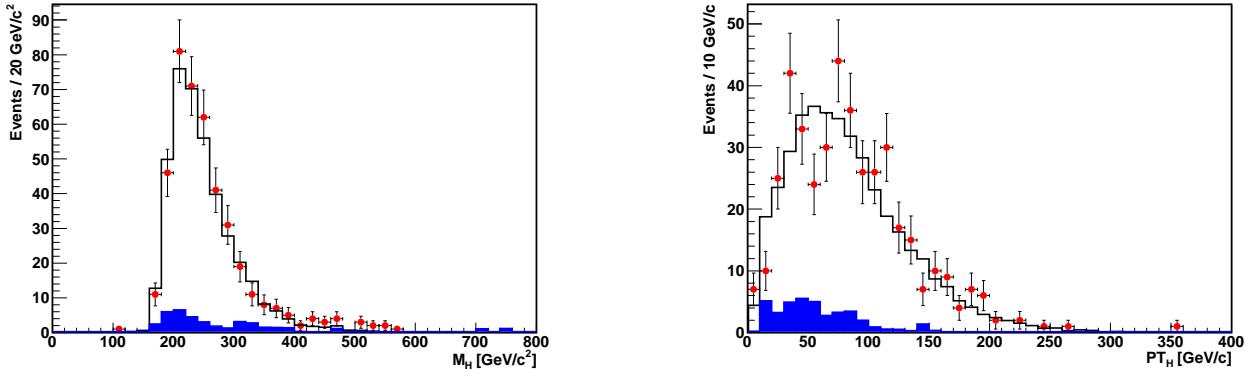


Figure 18: Comparison of 2011 $e\mu$ data (points) to Herwig top MC (white; other small backgrounds, blue), corresponding to an integrated luminosity of 868 pb^{-1} . Dilepton+ dijet “Higgs” invariant mass (left) and transverse momentum (right) for events satisfying $71 \text{ GeV}/c^2 < M_{ll} < 111 \text{ GeV}/c^2$.

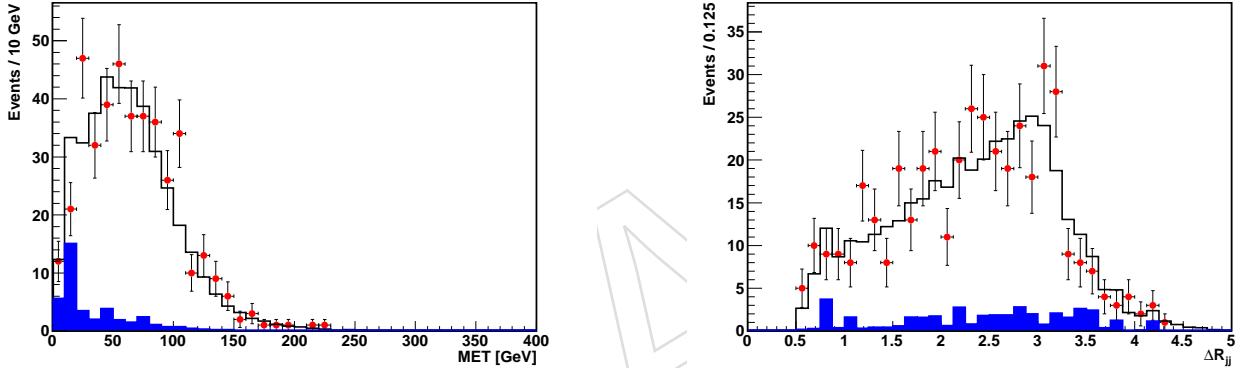


Figure 19: Comparison of 2011 $e\mu$ data (points) to Herwig top MC (white; other small backgrounds, blue), corresponding to an integrated luminosity of 868 pb^{-1} . Missing transverse energy (left) and dijet ΔR_{jj} separation (right) for events satisfying $71 \text{ GeV}/c^2 < M_{ll} < 111 \text{ GeV}/c^2$.

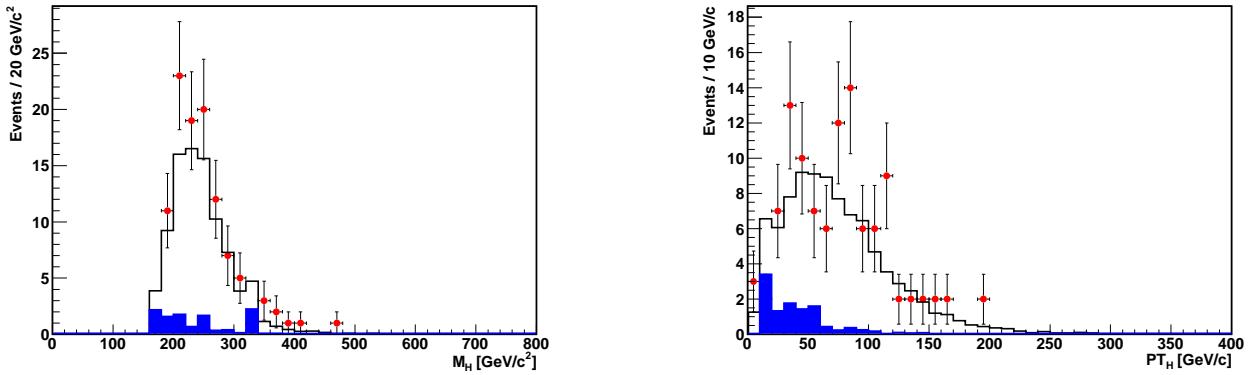


Figure 20: Comparison of 2011 $e\mu$ data (points) to Herwig top MC (white; other small backgrounds, blue), corresponding to an integrated luminosity of 868 pb^{-1} . Dilepton+ dijet “Higgs” invariant mass (left) and transverse momentum (right) for events satisfying $71 \text{ GeV}/c^2 < M_{ll} < 111 \text{ GeV}/c^2$ and $71 \text{ GeV}/c^2 < M_{jj} < 111 \text{ GeV}/c^2$.

216 **5.2 Mistags**

217 Another source of background consists of jets that, although initiated by u,d,s quarks or gluons¹, pass all selection cuts, including b -tagging requirement. The contribution from this back-
218 ground is estimated as the efficiency for tagging light jets, denoted as *mistag rate* [15]. In simu-
219 lated events, it can be directly measured as:

$$\epsilon_{MC}^{\text{mistag}} = \frac{N^{\text{TAGGED light}}}{N^{\text{ALL light}}} \quad (3)$$

221 Jet flavour identification considers all partons generated within a cone $\Delta R < 0.3$ around the jet
222 axis. If neither a b nor a c quark is found, the flavor *light* is assigned to the jet, if the highest
223 energetic parton is either an u,d or s quark or a gluon.

224 However, mistag rates estimation must be cross-checked using data-driven techniques, so that
225 it does not rely completely on the Monte Carlo modelling of jet-flavor content. At CMS, the
226 negative taggers method [15] is used. It uses the fact that impact-parameter (IP) of b -jets is
227 mainly positive, whereas for light jets is symmetric around 0 due to uncertainties in the track
228 measurements. The tagger discriminant, given by the IP significance, can also be defined as
229 positive or negative accordingly. Given its symmetry around zero, the number of tagged jets
230 with negative tagger is a good estimate of the number of light jets tagged on the positive side.
231 Further details are found in [15].

232 The working points for b -taggers at CMS are defined for different b -tagging efficiency and
233 mistag rates. Correction factors $\epsilon_{DATA}^{\text{mistag}} / \epsilon_{MC}^{\text{mistag}}$ are available as functions of jet E_T , η in the
234 database, and should be applied to simulated events before the comparison with the data.

235 Here, we use the same negative taggers method to investigate if the Higgs selection imposes
236 any bias on the mistag rates, in which case the correction factors provided by the b Tag POG
237 group would not be appropriate. The data consists on 17 pb^{-1} of 2011 run period, and for
238 Monte Carlo, the sample of Higgs signal for mass 400 GeV (see Table 4) was used. Only muons
239 were considered. Because of the small statistics – to be updated as 2011 data becomes available
240 – neither the cut on the invariant mass of the muon pair nor the muon-jet separation was im-
241 posed. Only taggeable jets are considered, both in the numerator and the denominator of Eq. 3.
242 All jets are required to have neutral EM energy fraction < 1 , neutral HAD energy fraction < 1 ,
243 charged HAD energy fraction > 0 , charged EM energy fraction < 1 ; number of charged particle
244 tracks > 0 and number of charged+neutral particles > 1 . A charged particle track is defined
245 as a track fulfilling the cuts: Number of pixel hits > 1 , number of silicon strip+pixel hits > 7 ,
246 $p_T > 1 \text{ GeV}$, normalized $\chi^2 < 5$, transverse impact parameter $< 0.2 \text{ cm}$ and longitudinal impact
247 parameter $< 17 \text{ cm}$.

248 Results are presented in Tables 9 and 10. In spite of the large uncertainties, due to low statis-
249 tics, the results are in very good agreement with the official b Tag POG numbers, given in [15].
250 Therefore, the correction factors available in CMSSW database can be directly used in our anal-
251 ysis.

252 The contribution from charm-initiated jets, in addition to the mistags from light jets, is also
253 shown.

254 In summary, the mistag estimation in our analysis region using the negative tag method is in
255 good agreement with the estimations provided by the BTV POG. During the analysis official

¹In this note, we refer to jets initiated by either u,d,s quarks or gluons as *light jets*

Jet E_T (GeV)	Mistag	Mistag+charm
Data (TCHEM)		
>30	0.009 ± 0.005	0.019 ± 0.010
30-45	0.012 ± 0.018	0.028 ± 0.038
45-80	0.014 ± 0.016	0.032 ± 0.032
>80	0.012 ± 0.019	0.024 ± 0.035
Monte Carlo (TCHEM)		
>30	0.012 ± 0.004	0.026 ± 0.006
30-45	0.007 ± 0.006	0.015 ± 0.008
45-80	0.009 ± 0.006	0.021 ± 0.008
>80	0.015 ± 0.015	0.031 ± 0.020

Table 9: Mistag rates and rates of light+c flavor jets using TCHEM tagging, in bins of jet E_T .

Jet E_T (GeV)	Mistag	Mistag+charm
Monte Carlo (TCHEL)		
>30	0.086 ± 0.011	0.110 ± 0.012
30-45	0.059 ± 0.018	0.075 ± 0.019
45-80	0.070 ± 0.016	0.091 ± 0.017
>80	0.102 ± 0.039	0.127 ± 0.041
Data (TCHEL)		
>30	0.051 ± 0.012	0.065 ± 0.014
30-45	0.060 ± 0.031	0.076 ± 0.038
45-80	0.027 ± 0.012	0.035 ± 0.015
>80	0.133 ± 0.080	0.166 ± 0.094

Table 10: Mistag rates and rates of light+c flavor jets using TCHEL tagging, in bins of jet E_T .

256 numbers from the BTV POG shown in Appendix D are used.

257 6 Correlated and Uncorrelated Systematic Uncertainties

258 Luminosity uncertainty has been estimated to be 6% [16, 17]. Sources of systematic uncertainties
 259 are the knowledge of lepton selection and b-tagging efficiency, the lepton and jets resolution
 260 and scales, the \cancel{E}_T modeling, and knowledge of the signal and background shapes and cross
 261 sections.

262 Studies about jet resolution and efficiency, pile-up effects, and signal cross section uncertainties
 263 are described in Reference [14]. In the following subsections we describe lepton selection
 264 efficiencies, b-tagging, and background systematics.

265 6.1 Leptons selection and Trigger efficiency

266 The efficiency of selecting a *lepton object* has been evaluated both on data 2011 and on simulation,
 267 selecting electron and muon objects closer as possible to the ones requested in the analysis
 268 from a single lepton trigger dataset. Therefore each event is required to contain at least two leptons
 269 belonging to Z decay and one jet, in order to match the analysis event topology. The total
 270 efficiency measurement can be factorized into five sequential relative measurements. The five
 271 steps are: tracking efficiency, reconstruction efficiency, identification efficiency, isolation effi-

272 cency and the online or total trigger efficiency. The total is thus given by the product:

$$\epsilon_{lepton} = \epsilon_{tracking} * \epsilon_{RECO/Tracking} * \epsilon_{ID/RECO} * \epsilon_{ISO/ID} * \epsilon_{Trigger/ISO} \quad (4)$$

273 Each term is considered separately in the following and the values are provided by the official
 274 “tag-and-probe” method, a generic tool for measuring efficiencies using known resonances in
 275 p_T and η bins ($\epsilon_{tracking}$ is assumed to be very close to 100%) [18]. Given the intermediate p_T
 276 range ([20-100] GeV) of the lepton required for this analysis the Z resonance and its leptonic
 277 decay products are chosen to measure the efficiency with this method. It requires a mass con-
 278 straint from a pair of same flavor lepton objects where on the two is tightly selected, the “tag”,
 279 to ensure enough purity, and the other leg, the “probe” is used to measure efficiencies for a
 280 given identification criterium: efficiency is defined as the ratio of the number of probes passing
 281 the set of cuts to the total number of probes before the cuts.

282 **Reconstruction and ID.** Efficiencies computed for the lepton reconstruction and identification
 283 on data and simulation are evaluated as a function of the different η regions in a unique pT
 284 range. Tight cuts on the lepton Tag are requested, asking for a matching with the single trigger
 285 object and a very high reconstruction quality (VBTF) for both electrons and muons. For muons,
 286 identification criteria used for selecting the passing probes matches the ones of a VBTF muons
 287 with further requirements, like transverse and longitudinal impact parameter. For electrons the
 288 $\epsilon_{RECO/Tracking}$ is evaluated, meaning the efficiency of reconstructing an electron with GSF al-
 289 gorithm given the original super cluster in the calorimeter. Results are reported in Table 23.
 290 Instead, the electron identification efficiency is factorized together with isolation, by asking
 291 how many electrons having a WP95 (or WP80) are reconstructed, given a GSF electron, and
 292 values are reported in Table 24.

$\epsilon_{RECO} * \epsilon_{ID}$ for muons			
η coverage	p_T range (GeV)	Efficiency [%] (Data)	Data/MC ratio
$ \eta < 1.20$ $1.20 < \eta < 2.40$	20-100	96.0 ± 0.1	0.996 ± 0.001
	20-100	96.0 ± 0.1	0.986 ± 0.001
ϵ_{RECO} for electrons			
η coverage	p_T range (GeV)	Efficiency [%] (Data)	Data/MC ratio
$ \eta < 0.80$ $0.80 < \eta < 1.44$ $1.44 < \eta < 1.57$ $1.57 < \eta < 2.0$ $2.0 < \eta < 2.5$	20-100	97.7 ± 0.3	0.999 ± 0.005
	20-100	94.2 ± 0.3	0.964 ± 0.003
	20-100	96.0 ± 0.6	0.99 ± 0.04
	20-100	95.1 ± 0.5	0.992 ± 0.006
	20-100	93.6 ± 0.4	1.001 ± 0.006

Table 11: Reco/ ID efficiency values from 2011 data and simulation using Tag and Probe method.

293 **Isolation.** Efficiencies for isolation have been computed requiring the passing probe to satisfy

294 the criteria adopted for isolating leptons in this analysis. For the moment, the jet energy corrections
 295 due the pile-up are not taken into account in the definition of the isolation variable with
 296 respect to which the efficiency is computed, but studies have been demonstrated as changes
 297 are minor. Results are shown in Table 24. As explained, the electron identification efficiency is
 298 factorized together with isolation and a unique number is reported in the table, for the WP95.

ϵ_{ISO} for muons			
η coverage	p_T range (GeV)	Efficiency [%] (Data)	Data/MC ratio
$ \eta < 0.9$	20-40	94.5 ± 0.5	0.987 ± 0.006
$0.9 < \eta < 2.4$	20-40	96.5 ± 0.4	0.995 ± 0.005
$ \eta < 0.9$	40-100	98.7 ± 0.2	0.994 ± 0.002
$0.9 < \eta < 2.4$	40-100	99.2 ± 0.2	0.996 ± 0.002
$\epsilon_{ID} * \epsilon_{ISO}$ for electrons (WP95)			
η coverage	p_T range (GeV)	Efficiency [%] (Data)	Data/MC ratio
$ \eta < 1.5$	20-40	91.6 ± 0.5	0.988 ± 0.006
$1.5 < \eta < 2.5$	20-40	83.1 ± 0.8	0.998 ± 0.011
$ \eta < 1.5$	40-100	94.8 ± 0.2	0.988 ± 0.003
$1.5 < \eta < 2.5$	40-100	94.1 ± 0.3	1.016 ± 0.064

Table 12: Isolation (together with identification, for electrons) efficiency values from 2011 data and simulation with Tag and Probe method.

299 **Trigger.** Concerning the trigger, efficiency that a lepton identified with the analysis criteria
 300 fired on the trigger used in the analysis has been evaluated. Using again the “tag-and-probe”
 301 method, a tight requirement on the tag is done, keeping the probe unbiased. Then, probe is
 302 asked to geometrically match with the offline trigger object and a value for the efficiency is
 303 extracted. Matching efficiency is of the order of permille, therefore negligible. For muons, the
 304 HLT_DoubleMu7 unprescaled trigger used in the analysis have been estimated on data. The
 305 HLT path is emulated on MC, since it is not simulated in the sample, by applying the proper
 306 threshold to trigger objects reconstructed on-line. The efficiency for this trigger has been com-
 307 puted for the single leg and has been found to vary as a function of η of the muon, therefore
 308 efficiencies for three distinct regions have been evaluated (These results will be updated with
 309 the new unprescaled trigger HLT_Mu13_Mu8, currently in use for the HLT menu). A simi-
 310 lar study has been performed using single muon trigger, by choosing the lowest unprescaled
 311 available one in both data and simulation, the HLT_IsoMu17. Here the efficiencies measured on
 312 data are lower than for the double trigger case, in particular for the outermost η region, likely
 313 due to tighter requirements on L1 quality level for the single trigger with respect to double. For
 314 electrons, efficiency for the HLT_Ele17Calo_Ele8Calo was measured on data to be greater than
 315 99% for electrons with p_T larger than 20 GeV through the entire ECAL fiducial region. Results
 316 are reported in Table 25.

$\epsilon_{Trigger}$ for single leg of HLT_DoubleMu7			
η coverage	p_T range (GeV)	Efficiency [%] (Data)	Data/MC ratio
$ \eta < 0.9$	20-100	96.7 ± 0.2	0.987 ± 0.002
$0.9 < \eta < 2.1$	20-100	95.0 ± 0.2	0.97 ± 0.003
$2.1 < \eta < 2.4$	20-100	91.3 ± 0.9	0.99 ± 0.01
$\epsilon_{Trigger}$ for HLT_IsoMu17			
η coverage	p_T range (GeV)	Efficiency [%] (Data)	Data/MC ratio
$ \eta < 0.9$	20-100	87.7 ± 0.4	0.975 ± 0.005
$0.9 < \eta < 2.1$	20-100	81.6 ± 0.5	0.954 ± 0.006
$2.1 < \eta < 2.4$	20-100	63.15 ± 1.6	1.72 ± 0.05

Table 13: Trigger efficiency values from 2011 data and simulation with Tag and Probe method.

6.2 Leptons energy scale and resolution

Systematic uncertainties can arise also from the lepton energy/momentum scale and resolution. Muon transverse momentum bias has been measured from 2010 data [19] as a function of the muon kinematic variables, and a clear dependence from the azimuthal angle of the muon was found. Nevertheless the maximum variation was found to be $\Delta p_T / p_T < 1\%$. For the electron energy scale variation of 2% for the central region and 5% for the endcap is applied. Numbers have been extracted from prompt reco, where laser corrections for the ECAL transparency were not yet applied. For a more consistent estimation the study will be repeated with re-reco sample.

6.3 Evaluation of systematic uncertainties

For each range a value for the ratio between efficiency on data and on simulation given by tag-and-probe method and the relative uncertainty is provided. The uncertainty on this ratio (σ_R) is computed by propagating single data and MC efficiency uncertainty. The final one considered includes:

$$\sigma_R^2 = \sigma_{stats+fit}^2 + \sigma_{syst}^2 \quad (5)$$

where σ_{syst} is intrinsic from tag and probe method (up to 0.4% for muons, 0.5% for electrons) and $\sigma_{stats+fit}$ is the statistical plus fit uncertainty. Each data/MC ratio has been used as correction factor to properly re-weight the MC events. In most of the cases the corrections factors to be applied have found to be larger than the calculated uncertainty on them. Therefore, in order to correctly evaluate the systematic uncertainty on the signal, the corrected MC has been varied by assigning $\pm \sigma_R$ of the uncertainty computed on the correction factor. Then, using a cut and count approach, the variation of the final signal yields in the window $M_{Higgs} * 0.94 < M_{lljj} < M_{Higgs} * 1.1$ with respect to the corrected one, for different mass values, has been evaluated. No significative mass-dependence has been observed. The selection

340 applied to compute the final yields is the one for 2b-tag category, but no variation is expected if
 341 considering the other categories, since the lepton selection cuts do not change. The uncertainty
 342 due to the muon trigger has been evaluated by varying in the same ways as done for the recon-
 343 struction and identification and in the case of double trigger the single leg efficiency has been
 344 squared to correctly compute the contribution of both legs. The dielectron trigger is assumed
 345 to be fully efficient on data and therefore a 1% systematic is assigned to it. Uncertainty on
 346 the muon momentum scale is propagated through the final yield by varying the muon p_T ac-
 347 cording to the function described in [19]. Same method has been adopted for quoting electron
 energy scale uncertainty. Final resulting values for the systematics are reported in Table 26.

category	$H \rightarrow ZZ \rightarrow \mu\mu jj$	$H \rightarrow ZZ \rightarrow ee jj$
Reco-ID-isolation	0.8%	3.4%
Trigger (DoubleMu7)	1%	/
Trigger (IsoMu17)	2.4%	/
Trigger (Ele17CaloEle8Calo)	/	1.0%
momentum/energy scale	1.0%	3.0%

Table 14: Systematic uncertainties estimated for muon and for the electron channel.

348

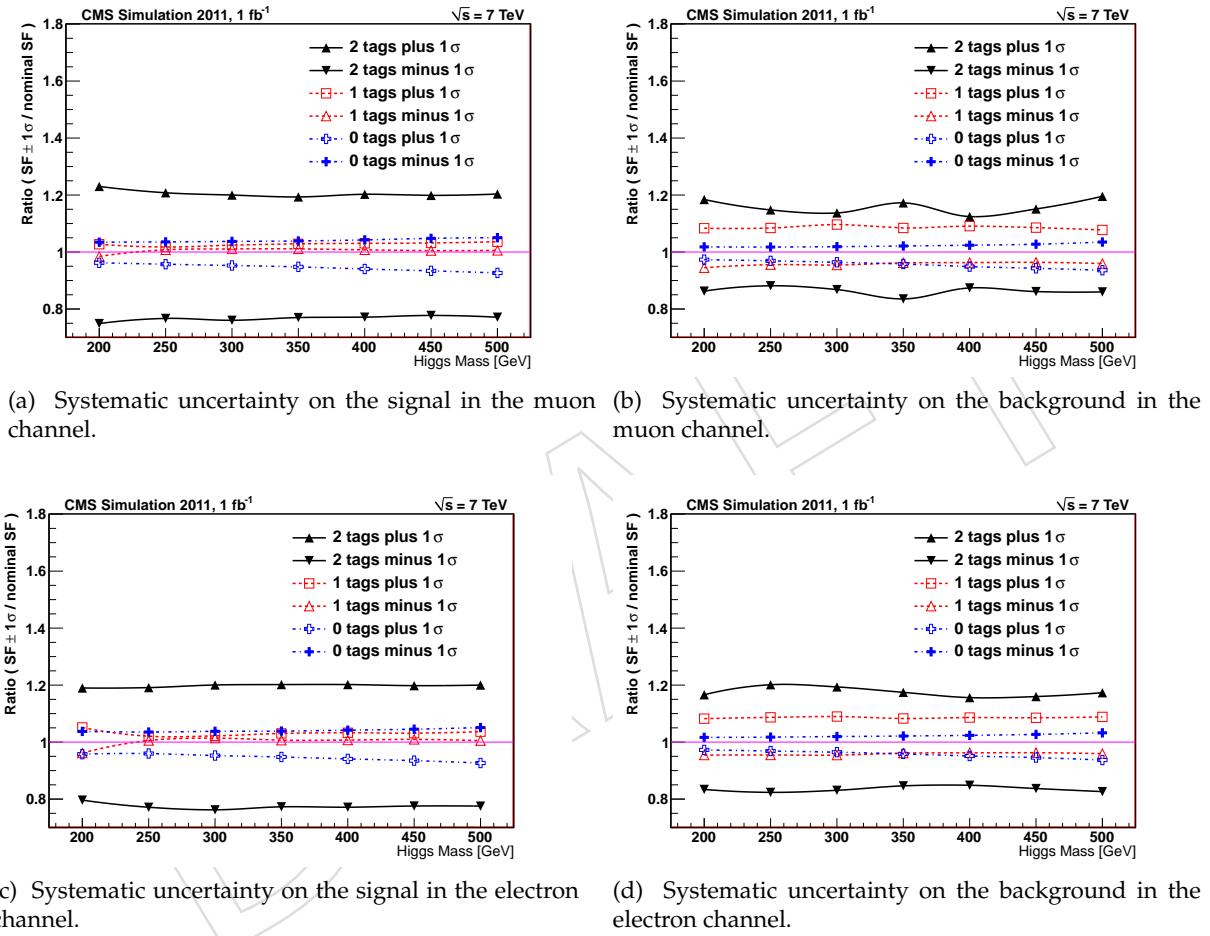
349 6.4 *b*-tagging Systematics

350 A data-to-Monte Carlo scale factor (SF_b) of 0.9 ± 0.15 has been measured for events containing
 351 b -jets. This SF_b corrects for the 10% more efficient identification of b -jets in Monte Carlo com-
 352 pared to data. Likewise, a mistag rate scale factor (SF_{mistag}) for light quarks mismeasured
 353 as b -jets has been measured over a range of p_T and η for the jets. To study the systematic effects
 354 of b -tagging, both the SF_b and SF_{mistag} were simultaneously varied up and down by 15%.

355 The study was performed separately for the muon and electron channels. The effect was calcu-
 356 lated independently for signal and the combined background channels excluding $t\bar{t}$. We expect
 357 real b -jets in the $t\bar{t}$ sample, but due to the lack of sufficient number of events surviving the full
 358 event selection requirements, the systematic uncertainty on the $t\bar{t}$ sample was assumed to be
 359 the same as that computed for the signal samples.

360 Table 22 gives the b -tagging systematic uncertainty for the signal, $t\bar{t}$, and background channels
 361 excluding $t\bar{t}$, for muons and electrons. The systematic effect is computed as the ratio of the
 362 number of tagged jets with a SF varied by plus and minus 15% to the number of tagged jets
 363 with the nominal SF . The uncertainty is reported for the cases where both jets are tagged, at
 364 least 1 jet is tagged, and no jets are tagged.

365 Figure 21 shows the b -tagging systematic uncertainty for the signal and background channels
 366 excluding $t\bar{t}$ in the muon and electron samples for the different cases of b -category where both
 367 jets are tagged, at least 1 jet is tagged, and no jets are tagged.



(c) Systematic uncertainty on the signal in the electron channel.

(d) Systematic uncertainty on the background in the electron channel.

Figure 21: b -tagging systematic uncertainty for the signal and background channels excluding $t\bar{t}$ in the muon and electron channel.

	2 Tags	1 Tag	0 Tags
Muons			
Signal	20%	1%	3%
$t\bar{t}$	20%	1%	3%
Background	20%	8%	2%
Electrons			
Signal	20%	1%	3%
$t\bar{t}$	20%	1%	3%
Background	20%	8%	2%

Table 15: b -tagging systematic uncertainty

368 7 Signal Optimization

369 In order to increase the signal over background ratio in the analysis, we perform an optimization
 370 process taking as benchmark the object definitions described in section 3.1, a cut on MET
 371 significance smaller than 10 described in section 3.3, plus a cut on the Z boson mass window
 372 as expected from our signal requiring $75 < m_{jj} < 105$ and $70 < m_{ll} < 110$.

373 For the double tag category the baseline tagging selections defined in subsection 3.2 is applied.
 374 The final step in the optimization is done based on a Likelihood Discriminant based on helicity
 375 angles, described in the following section. A cut-based kinematic selection is used as a
 376 consistency cross check as shown in Appendix C.

377 7.1 Angular Optimization

378 The angular optimization is based on the LD introduced in section 3.5. We select events with a
 379 LD value greater than 0.5.

380 We check the performance of the analysis applying the described optimization to MC simulation
 381 normalized to $1 fb^{-1}$. The m_{jjll} distribution in the final region is shown in Figure 22.
 382 Cutting on different mass ranges we obtain different optimizations as a function of the Higgs
 383 mass, as shown in Tables 16 to 21.

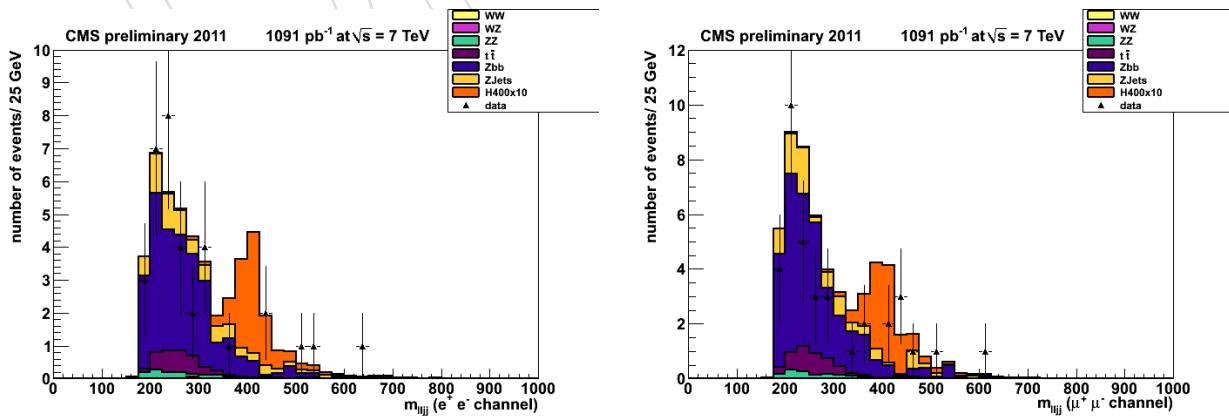


Figure 22: Distribution of the m_{jjll} for electron (left) and muon (right) events in $1 fb^{-1}$. Higgs signal multiplied by 10 with a mass of 400 GeV is shown in orange, superimposed to the SM backgrounds.

Process	Muons	Electrons
$M_H = 250$ [235.5-275.5]	0.82	0.73
ZJets	1.10	1.34
ZZ	0.25	0.30
WZ	0.04	0.02
WW	0.00	0.00
Zbb	8.42	5.85
$t\bar{t}$	1.47	1.09
Total Bkg	11.27	8.60
Data	5	7

Table 16: Number of events expected in the final region within a Higgs mass range of [235.5-275.5].

Process	Muons	Electrons
$M_H = 300$ [282.5-330.5]	0.52	0.54
ZJets	1.17	0.93
ZZ	0.26	0.19
WZ	0.03	0.05
WW	0.00	0.00
Zbb	3.73	4.71
$t\bar{t}$	0.71	0.64
Total Bkg	5.91	6.51
Data	2	6

Table 17: Number of events expected in the final region within a Higgs mass range of [282.5-330.5].

Process	Muons	Electrons
$M_H = 350$ [329.5-385.5]	1.08	1.01
ZJets	0.77	0.84
ZZ	0.18	0.16
WZ	0.01	0.01
WW	0.00	0.00
Zbb	3.25	1.98
$t\bar{t}$	0.15	0.16
Total Bkg	4.36	3.16
Data	3	1

Table 18: Number of events expected in the final region within a Higgs mass range of [329.5-385.5].

Process	Muons	Electrons
$M_H = 400 [376.5-440.5]$	0.78	0.75
ZJets	0.52	0.67
ZZ	0.10	0.09
WZ	0.03	0.04
WW	0.00	0.00
Zbb	1.08	1.12
$t\bar{t}$	0.02	
Total Bkg	1.74	1.94
Data	2	2

Table 19: Number of events expected in the final region within a Higgs mass range of [376.5-440.5].

Process	Muons	Electrons
$M_H = 450 [423.5-495.5]$	0.53	0.49
ZJets	0.78	0.52
ZZ	0.10	0.06
WZ	0.04	0.04
WW	0.00	0.00
Zbb	0.76	0.50
$t\bar{t}$	0.04	0.00
Total Bkg	1.72	1.13
Data	4	2

Table 20: Number of events expected in the final region within a Higgs mass range of [423.5-495.5].

Process	Muons	Electrons
$M_H = 500 [470.5-550.5]$	0.32	0.33
ZJets	0.32	0.37
ZZ	0.06	0.07
WZ	0.02	0.02
WW	0.00	0.00
Zbb	0.91	0.69
$t\bar{t}$	0.04	0.01
Total Bkg	1.35	1.16
Data	2	2

Table 21: Number of events expected in the final region within a Higgs mass range of [470.5-550.5].

384 8 Statistical Analysis and Results

385 We determine the expected upper limits to the Standard Model Higgs production cross section
 386 as a function of the Higgs boson mass assuming as reference an integrated luminosity of 1 fb^{-1} .
 387 We use the official tool developed by the CMS Higgs combination group [20] that supports
 388 different methods, including Bayesian approach, frequentist profile likelihood, and Feldman-
 389 Cousins [21] methods, and the modified frequentist CLs method [22] with Cousins-Highland
 390 integration of nuisance parameters for the treatment of systematic uncertainties [23]. The tool
 391 uses the RooStats [24] engine from ROOT as internal implementation. We determine the
 392 expected upper limit to the Higgs boson production cross section times branching fraction to
 393 $\ell\ell b\bar{b}$. The limit is expressed as ratio r of the determined upper limit to the cross section times
 394 branching ratio divided by its standard model expectation. A value of the Higgs boson mass
 395 m_H is excluded if, for that mass hypothesis, r is less than one.

396 There are two main possible general approaches to determine the limit to the Higgs cross sec-
 397 tions.

398 The simplest approach, referred to as “cut and count” analysis, uses only the number of events
 399 selected within a given window in the reconstructed Higgs mass, $m_{\ell\ell jj}$, around an assumed
 400 value m_H of the Higgs boson mass. The Higgs cross section limit is determined from the ex-
 401 pected number of signal and background events passing the selections s and b respectively. We
 402 combine using Poissonian statistics the counting information from the two channels with elec-
 403 trons and muons. The sources of systematic uncertainties on s and b are taken into account in
 404 the determination of the limit assuming log-normal distributions of the nuisance parameters.

405 Another possible approach, referred to as “shape analysis”, takes into account the measured
 406 distribution of variables that can discriminate Higgs signal events against background events.
 407 In the case of the $H \rightarrow ZZ \rightarrow \ell^+\ell^- b\bar{b}$ analysis, the main discriminant variable is the recon-
 408 structed Higgs boson mass, $m_{\ell\ell jj}$, which is peaked around the true Higgs boson mass, m_H , for
 409 the signal and has a broader distribution for background processes. We can develop an anal-
 410 ysis that does not select signal events cutting on $m_{\ell\ell jj}$, and we can use the distribution of $m_{\ell\ell jj}$
 411 for selected events, comparing to the expected distribution from signal and background. The
 412 possible sources of systematic uncertainties on both the expected signal and background yields
 413 and the shape of signal and background distributions are considered in the limit extraction
 414 procedure. The shape analysis can be implemented in the Higgs combination tool using an
 415 unbinned approach with an extended likelihood function defined by:

$$416 -\ln \mathcal{L} = \sum_{j=1}^k \left[s_j + b_j - \sum_{i=1}^{n_j} \left(s_i \mathcal{P}_s^{(j)}(m_i^{(j)}) + b_i \mathcal{P}_b^{(j)}(m_i^{(j)}) \right) \right], \quad (6)$$

417 where k is the number of channels (in our case $k = 2$, and $j = 1, 2$ correspond to the electron and
 418 muon channels), n_j is the number of selected candidate events in the channel j , s_j and b_j are
 419 the expected signal and background events, $\mathcal{P}_s^{(j)}$ and $\mathcal{P}_b^{(j)}$ are the probability density functions
 420 for signal and background, and $m_i^{(j)}$ are the n_j values of $m_{\ell\ell jj}$ of the selected candidates in the
 421 channel j . As alternative, the distribution of $m_{\ell\ell jj}$ can be sub-divided into bins, and in that
 422 cases the number of selected events in each bin is considered with the corresponding expected
 423 number of events from signal and background in that bin, and the information from all bins is
 424 combined using a Poissonian likelihood, similarly to the “cut and count” case, but with more
 425 channels.

426 Results combining all the $H \rightarrow ZZ \rightarrow 2\ell 2q$ categories (with and without tagging) are described

⁴²⁶ in Reference [14].

⁴²⁷ 9 Conclusions

⁴²⁸ We present a search for a Standard Model Higgs boson decay to two Z bosons with a subsequent decay to a semileptonic final state with two leptons and two b-jets, $H \rightarrow ZZ \rightarrow 2l2b$.

⁴³⁰ We performed an optimization of the selection cuts based on kinematic and topological quantities. A study of systematics errors have been done focusing mainly on lepton selection efficiency and b-tagging. Full analysis is validated using 1 fb^{-1} data collected in 2011.

⁴³³ 10 Update for Lepton-Photon-2011

⁴³⁴ The results showed in the previous chapters of this document were targeted for the EPS-2011 conference. Since then, a significant amount of data was taken and the results are updated for Lepton-Photon-2011.

⁴³⁷ Taking advantage of the new performance studies [25] provided by the BTV POG, the b -tagging systematics are recalculated as shown in the following section.

⁴³⁹ The efficiency of selecting a *lepton object* and its systematic uncertainty have been evaluated both on 2011 data and MC simulation.

⁴⁴¹ 10.1 b -tagging Systematics

⁴⁴² A data-to-Monte Carlo scale factor (SF_b) has been measured for events containing b -jets as a function of p_T and η for the jets. This SF_b corrects for the more efficient identification of b -jets in Monte Carlo compared to data. Likewise, a mistag rate scale factor (SF_{mistag}) for light quarks misreconstructed as b -jets has been measured over a range of p_T and η for the jets. To study the systematic effects of b -tagging, both the SF_b and SF_{mistag} were simultaneously varied up and down by 10%.

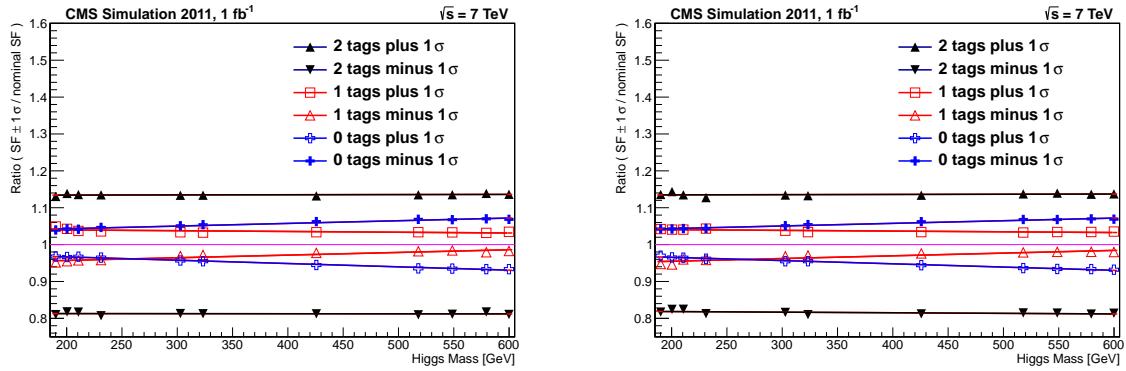
⁴⁴⁸ The study was performed separately for the muon and electron channels, calculating the effect for signal MC.

Signal	2 Tags	1 Tag	0 Tags
Muons	+13/-18%	3%	3-7%
Electrons	+13/-18%	3%	3-7%

Table 22: b -tagging systematic uncertainty

⁴⁵⁰ Table 22 gives the b -tagging systematic uncertainty for the signal, for muons and electrons. The systematic effect is computed as the ratio of the number of tagged jets with a SF varied by plus and minus 10% to the number of tagged jets with the nominal SF . The uncertainty is reported for the cases where both jets are tagged, at least 1 jet is tagged, and no jets are tagged. In the analysis, the exact systematic uncertainty as a function of the Higgs mass is applied.

⁴⁵⁵ Figure 23 shows the b -tagging systematic uncertainty for the signal in the muon and electron samples for the different cases of b -category where both jets are tagged, at least 1 jet is tagged, and no jets are tagged.



(a) Systematic uncertainty on the signal in the muon channel.
(b) Systematic uncertainty on the signal in the electron channel.

Figure 23: b -tagging systematic uncertainty for the signal in the muon and electron channel.

458 10.2 Leptons selection and Trigger efficiency

459 The efficiency of selecting a *lepton object* has been evaluated both on data 2011 and on simulation,
460 selecting electron and muon objects closer as possible to the ones requested in the analysis
461 from a single lepton trigger dataset. Therefore each event is required to contain at least two lep-
462 tons belonging to Z decay and one jet, in order to match the analysis event topology.

463 The total efficiency measurement can be factorized into five sequential relative measurements.
464 The five steps are: tracking efficiency, reconstruction efficiency, identification efficiency, iso-
465 lation efficiency and the online or total trigger efficiency. The total is thus given by the product:

$$\epsilon_{\text{lepton}} = \epsilon_{\text{tracking}} * \epsilon_{\text{RECO/Tracking}} * \epsilon_{\text{ID/RECO}} * \epsilon_{\text{ISO/ID}} * \epsilon_{\text{Trigger/ISO}} \quad (7)$$

466 Each term is considered separately in the following and the values are provided by the official
467 “tag-and-probe” method, a generic tool for measuring efficiencies using known resonances in
468 p_T and η bins ($\epsilon_{\text{tracking}}$ is assumed to be very close to 100%) [18].

469 Given the intermediate p_T range ([20-100] GeV) of the lepton required for this analysis the
470 Z resonance and its leptonic decay products are chosen to measure the efficiency with this
471 method. It requires a mass constraint from a pair of same flavor lepton objects where on the
472 two is tightly selected, the “tag”, to ensure enough purity, and the other leg, the “probe” is
473 used to measure efficiencies for a given identification criteria: efficiency is defined as the ratio
474 of the number of probes passing the set of cuts to the total number of probes before the cuts.

475 **Reconstruction and ID.** Efficiencies computed for the lepton reconstruction and identification
476 on data and simulation are evaluated as a function of the different η regions in a unique p_T
477 range.

478 Tight cuts on the lepton Tag are requested, asking for a matching with the single trigger ob-
479 ject and a very high reconstruction quality (VBTF) for both electrons and muons. For muons,
480 identification criteria used for selecting the passing probes matches the ones of a VBTF muons
481 with further requirements, like transverse and longitudinal impact parameter. For electrons
482 the $\epsilon_{\text{RECO/Tracking}}$ is evaluated, meaning the efficiency of reconstructing an electron with GSF
483 algorithm given the original super cluster in the calorimeter. Results are reported in Table 23.
484 Instead, the electron identification efficiency is factorized together with isolation, by asking
485 how many electrons having a WP95 (or WP80) are reconstructed, given a GSF electron, and

⁴⁸⁶ values are reported in Table 24.

$\epsilon_{RECO} * \epsilon_{ID}$ for muons			
η coverage	p_T range (GeV)	Efficiency [%] (Data)	Data/MC ratio
$ \eta < 1.20$ $1.20 < \eta < 2.40$	20-100	96.0 ± 0.1	0.996 ± 0.001
	20-100	96.0 ± 0.1	0.986 ± 0.001
ϵ_{RECO} for electrons			
η coverage	p_T range (GeV)	Efficiency [%] (Data)	Data/MC ratio
$ \eta < 0.80$ $0.80 < \eta < 1.44$ $1.44 < \eta < 1.57$ $1.57 < \eta < 2.0$ $2.0 < \eta < 2.5$	20-100	97.7 ± 0.3	0.999 ± 0.005
	20-100	94.2 ± 0.3	0.964 ± 0.003
	20-100	96.0 ± 0.6	0.99 ± 0.04
	20-100	95.1 ± 0.5	0.992 ± 0.006
	20-100	93.6 ± 0.4	1.001 ± 0.006

Table 23: *Reco/ID efficiency values from 2011 data and simulation using Tag and Probe method.*

⁴⁸⁷ **Isolation.** Efficiencies for isolation have been computed requiring the passing probe to satisfy
⁴⁸⁸ the criteria adopted for isolating leptons in this analysis. For the moment, the jet energy correc-
⁴⁸⁹ tions due the pile-up are not taken into account in the definition of the isolation variable with
⁴⁹⁰ respect to which the efficiency is computed, but studies have been demonstrated as changes
⁴⁹¹ are minor. Results are shown in Table 24. As explained, the electron identification efficiency is
⁴⁹² factorized together with isolation and a unique number is reported in the table, for the WP95.

⁴⁹³ **Trigger.** Concerning the trigger, efficiency that a lepton identified with the analysis criteria
⁴⁹⁴ fired on the trigger used in the analysis has been evaluated. Using again the “tag-and-probe”
⁴⁹⁵ method, a tight requirement on the tag is done, keeping the probe unbiased. Then, probe is
⁴⁹⁶ asked to geometrically match with the offline trigger object and a value for the efficiency is
⁴⁹⁷ extracted. Matching efficiency is of the order of permille, therefore negligible. For muons,
⁴⁹⁸ the HLT_DoubleMu7 and HLT_Mu13_Mu8 unprescaled trigger used in this analysis have been
⁴⁹⁹ estimated on data. Since the HLT path is not simulated in the MC sample, the trigger is not
⁵⁰⁰ applied on MC and this last is directly rescaled with the efficiency measured on data. The
⁵⁰¹ efficiency for this trigger have been found to vary as a function of η of the muon, therefore
⁵⁰² efficiencies for three distinct regions have been evaluated. In order to maximize the efficiency
⁵⁰³ on the signal, the lowest unprescaled available single muon trigger, HLT_IsoMu24, has been
⁵⁰⁴ used in the analysis as well. Here the efficiencies measured on data are lower than for the
⁵⁰⁵ single leg of double trigger case, in particular for the outermost η region, likely due to tighter
⁵⁰⁶ requirements on L1 quality level for the single trigger with respect to double. The efficiencies
⁵⁰⁷ have been also evaluated in three different period, corresponding to the 2011 technical stops
⁵⁰⁸ where, due to some changes at L1 seed, different values for the efficiencies are expected. Values
⁵⁰⁹ are properly combined by weighting the final correction factor to be applied on MC event
⁵¹⁰ according to the integrated luminosity in where efficiencies are found be constant. Results are

ϵ_{ISO} for muons			
η coverage	p_T range (GeV)	Efficiency [%] (Data)	Data/MC ratio
$ \eta < 0.9$	20-40	94.5 ± 0.5	0.987 ± 0.006
$0.9 < \eta < 2.4$	20-40	96.5 ± 0.4	0.995 ± 0.005
$ \eta < 0.9$	40-100	98.7 ± 0.2	0.994 ± 0.002
$0.9 < \eta < 2.4$	40-100	99.2 ± 0.2	0.996 ± 0.002
$\epsilon_{ID} * \epsilon_{ISO}$ for electrons (WP95)			
η coverage	p_T range (GeV)	Efficiency [%] (Data)	Data/MC ratio
$ \eta < 1.5$	20-40	91.6 ± 0.5	0.988 ± 0.006
$1.5 < \eta < 2.5$	20-40	83.1 ± 0.8	0.998 ± 0.011
$ \eta < 1.5$	40-100	94.8 ± 0.2	0.988 ± 0.003
$1.5 < \eta < 2.5$	40-100	94.1 ± 0.3	1.016 ± 0.064

Table 24: Isolation (together with identification, for electrons) efficiency values from 2011 data and simulation with Tag and Probe method.

511 reported in Table 25. For electrons, efficiency for the HLT_Ele17Calo_Ele8Calo was measured
 512 on data to be greater than 99% for electrons with p_T larger than 20 GeV through the entire
 513 ECAL fiducial region.

514 10.3 Leptons energy scale and resolution

515 Systematic uncertainties can arise also from the lepton energy/momentum scale and resolu-
 516 tion. Muon transverse momentum bias has been measured from 2010 data [19] as a function of
 517 the muon kinematic variables, and a clear dependence from the azimuthal angle of the muon
 518 was found. Nevertheless the maximum variation was found to be $\Delta p_T / p_T < 1\%$. For the
 519 electron energy scale variation of 2% for the central region and 5% for the endcap is applied.
 520 Numbers have been extracted from prompt reco, where laser corrections for the ECAL trans-
 521 parency were not yet applied. For a more consistent estimation the study will be repeated with
 522 re-reco sample.

523 10.4 Evaluation of systematic uncertainties

524 For each range a value for the ratio between efficiency on data and on simulation given by tag-
 525 and-probe method and the relative uncertainty is provided. The uncertainty on this ratio (σ_R) is
 526 computed by propagating single data and MC efficiency uncertainty. The final one considered
 527 includes:

$$\sigma_R^2 = \sigma_{stats+fit}^2 + \sigma_{syst}^2 \quad (8)$$

528 where σ_{syst} is intrinsic from tag and probe method (up to 0.4% for muons, 0.5% for electrons)
 529 and $\sigma_{stats+fit}$ is the statistical plus fit uncertainty. Each data/MC ratio has been used as cor-

$\epsilon_{Trigger}$ for HLT_Mu13_Mu8 (only single leg Mu13 reported)				
period (Integrated luminosity)	$ \eta < 0.8$	$0.8 < \eta < 1.2$	$1.2 < \eta < 2.4$	
up to EPS ($1.1 fb^{-1}$)	$97.7 \pm 0.1\%$	$95.8 \pm 0.1\%$	$91.5 \pm 0.2\%$	
after EPS ($219 pb^{-1}$)	$98.0 \pm 0.1\%$	$95.8 \pm 0.1\%$	$90.1 \pm 0.2\%$	

$\epsilon_{Trigger}$ for HLT_IsoMu24				
period (Integrated luminosity)	$ \eta < 0.8$	$0.8 < \eta < 1.47$	$1.47 < \eta < 2.10$	$ \eta > 2.1$
up to MayTS ($217 pb^{-1}$)	$98.0 \pm 0.1\%$	$92.5 \pm 0.1\%$	$89.0 \pm 0.1\%$	$75.8 \pm 0.2\%$
MayTS to EPS ($920 pb^{-1}$)	$97.9 \pm 0.1\%$	$94.5 \pm 0.1\%$	$94.1 \pm 0.1\%$	$88.3 \pm 0.2\%$
After EPS ($219 pb^{-1}$)	$97.0 \pm 0.1\%$	$94.2 \pm 0.1\%$	$92.3 \pm 0.1\%$	$81.4 \pm 0.2\%$

Table 25: Trigger efficiency values from 2011 data and simulation with Tag and Probe method. In the upper table results for double muon trigger efficiency are reported: almost identical efficiency for Mu13 and Mu8 leg are found (differences $< 0.5\%$). In the lower table results for the single muon trigger are reported, related to different periods in which changes are expected.

reaction factor to properly re-weight the MC events. In most of the cases the corrections factors to be applied have found to be larger than the calculated uncertainty on them. Therefore, in order to correctly evaluate the systematic uncertainty on the signal, the corrected MC has been varied by assigning $\pm \sigma_R$ of the uncertainty computed on the correction factor.

Then, using a cut and count approach, the variation of the final signal yields in the window $M_{Higgs} * 0.94 < M_{lljj} < M_{Higgs} * 1.1$ with respect to the corrected one, for different mass values, has been evaluated. No significative mass-dependence has been observed. The selection applied to compute the final yields is the one for 2b-tag category, but no variation is expected if considering the other categories, since the lepton selection cuts do not change.

The dielectron trigger is assumed to be fully efficient on data and therefore a 1% systematic is assigned to it.

Uncertainty on the muon momentum scale is propagated through the final yield by varying the muon p_T according to the function described in [19]. Same method has been adopted for quoting electron energy scale uncertainty. Final resulting values for the systematics are reported in Table 26.

category	$H \rightarrow ZZ \rightarrow \mu\mu jj$	$H \rightarrow ZZ \rightarrow ee jj$
Reco-ID-isolation	0.8%	3.4%
Trigger (HLT_Mu13_Mu8)	>2%	/
Trigger (HLT_IsoMu24)	>2.%	/
Trigger (Ele17CaloEle8Calo)	/	1.0%
momentum/energy scale	1.0%	3.0%

Table 26: *Systematic uncertainties estimated for muon and for the electron channel.*

545 References

- [1] G. Abbiendi et al., The ALEPH, DELPHI, L3 and OPAL Collaborations, The LEP Working Group for Higgs Boson Searches, "Search for the Standard Model Higgs Boson at LEP", *Phys. Lett.* **B565** (2003) 61–75. doi:10.1016/S0370-2693(03)00614-2.
- [2] T. Aaltonen et al, The CDF and D0 Collaborations, "Combination of Tevatron searches for the standard model Higgs boson in the W+W- decay mode", *Phys. Rev. Lett.* **104** (2010) 061802. doi:10.1103/PhysRevLett.104.061802.
- [3] T. Aaltonen et al, The CDF, D0 Collaborations, the TEVNPHWG Working Group, "Combined CDF and D0 Upper Limits on Standard Model Higgs Boson Production with up to 8.2 fb-1 of Data", *FERMILAB-CONF-11-044-E* (2011) arXiv:1103.3233.
- [4] The ALEPH, CDF, D0, DELPHI, L3, OPAL, SLD Collaborations, the LEP Electroweak Working Group, the Tevatron Electroweak Working Group, and the SLD electroweak and heavy flavour groups, "Precision Electroweak Measurements and Constraints on the Standard Model", *LEPEWWG 2010-01* (2010) arXiv:1012.2367.
- [5] ALPGEN generator reference.
- [6] Yanyan Gao, Andrei V. Gritsan, Zijin Guo, Kirill Melnikov, Markus Schulze and Nhan V. Tran, "Spin-Determination of Single-Produced Resonances at Hadron Colliders",.
- [7] T. Sjöstrand, L. Lönnblad and S. Mrenna, "Title: PYTHIA 6.2 Physics and Manual", arXiv:0108264.
- [8] CMS Collaboration, "Performance of CMS muon identification in pp collisions at $\sqrt{s} = 7$ TeV", *CMS PAS MUO-2010-002* (2010).
- [9] CMS Collaboration, "Electron Reconstruction and Identification at $\sqrt{s} = 7$ TeV", *CMS Detector Performance Summary DP-2010-032* (2010).
- [10] CMS Collaboration, "Particle-flow commissioning with muons and electrons from J/Psi, and W events at 7 TeV", *CMS PAS PFT-2010-003* (2010).
- [11] M. Cacciari, G. P. Salam, and G. Soyez, "The anti- k_t jet clustering algorithm", *JHEP* **04** (2008) 063. doi:10.1088/1126-6708/2008/04/063.
- [12] A. Rizzi, F. Palla, and G. Segneri, "Track impact parameter based b-tagging with CMS", *CMS Note CMS-AN-2006/019* (2006).
- [13] CMS Collaboration, "Commissioning of b-jet identification with pp collisions at $\sqrt{s}=7$ TeV", *CMS Physics Analysis Summary CMS-PAS-BTV-10-001* (2010).
- [14] S. Bolognesi, A. Bonato, A. D. Del Re, A. and Gritsan et al., "Search for a Semileptonic Decay of a SM Higgs or BSM Boson $H \rightarrow ZZ \rightarrow 2l2j$ ", *CMS Note CMS-AN-2011/100*.
- [15] "Mistag note reference should go here",.
- [16] CMS Collaboration, "Measurement of CMS luminosity", *CMS PAS EWK-2010-004* (2010).
- [17] CMS Collaboration, "Absolute luminosity normalization", *CMS DPS 2011-002* (2011).

- 582 [18] C. Collaboration, "Measurements of Inclusive W and Z Cross Sections in pp Collisions at
583 $\sqrt{s} = 7$ TeV", *JHEP* **1101** (2011) 080. doi:10.1007/JHEP01(2011)080.
- 584 [19] CMS Collaboration Collaboration, "Performance of muon reconstruction and
585 identification in pp collisions at $\sqrt{s}=7$ TeV", *CMS PAS CMS-MUO-10-004under
586 approval* (2010).
- 587 [20] C. H. C. Group, "Documentation of the RooStats-based statistics tools for Higgs PAG",
588 *CMS TWiki SWGuideHiggsAnalysisCombinedLimit* (2011).
- 589 [21] G. J. Feldman and R. D. Cousins, "Unified approach to the classical statistical analysis of
590 small signals", *Phys. Rev. D* **57** (1998) 3873–3889. doi:10.1103/PhysRevD.57.3873.
- 591 [22] A. L. Read, "Modified Frequentist Analysis of Search Results (The CLs Method)", *CERN
592 OPEN* **2000-205** (2000).
- 593 [23] R. D. Cousins and V. L. Highland, "Incorporating systematic uncertainties into an upper
594 limit", *Nucl. Instr. Meth. A* **320** (1992) 331–335.
595 doi:doi:10.1016/0168-9002(92)90794-5.
- 596 [24] L. Moneta, K. Belasco, K. Cranmer et al., "The RooStats Project", arXiv:1009.1003.
- 597 [25] CMS Collaboration, "Performance of b -jet identification in CMS", *CMS Physics Analysis
598 Summary CMS-PAS-BTV-11-001* (2011).

DRAFT

599 A B-tagging optimization

600 Six different cases for defining b -tag categories were studied for the Track Counting High Efficiency (TCHE) and Track Counting High Purity (TCHP) B-taggers. For each case, the 95% C.L.
 601 upper limit on expected $\sigma \times \text{BR}$ for the process $H \rightarrow ZZ \rightarrow \ell\ell + 2\text{jets}$ was computed. In order
 602 to arrive at the optimal choice of the b -tag category and the algorithm, the one which led to the
 603 lowest 95% C.L. upper limit on the $\sigma \times \text{BR}$ was chosen.

605 We determined which B-tagging algorithm gives the lowest value for the upper 95% expected
 606 limit on the ratio of the number of signal events to the number of expected signal events. For
 607 each case the event was categorized to have two, one or zero tagged jets by checking if the
 608 jets passed or failed a series of discriminating values for the given b -tagging algorithm. If both
 609 the jets passed the tagging requirements the event was considered to have two tagged jets. If
 610 this failed, we checked whether there was exactly one jet that passed the tagging requirement;
 611 in this case, the event was categorized as one tagged jet event. If this also failed, the event
 612 was considered to have zero tagged jets. The list below describes the six cases which were
 613 considered for each algorithm.

- 614 • Case 1: 2 jets pass loose, 1 jet passes loose, No jets pass
- 615 • Case 2: 2 jets pass loose, 1 jet passes medium, No jets pass
- 616 • Case 3: 1 jet passes loose and 1 jet passes medium, 1 jet passes loose, No jets pass
- 617 • Case 4: 1 jet passes loose and 1 jet passes medium, 1 jet passes medium, No jets pass
- 618 • Case 5: 2 jets pass medium, 1 jet passes loose, No jets pass
- 619 • Case 6: 2 jets pass medium, 1 jet passes medium, No jets pass

	Loose	Medium
TCHE	1.7	3.3
TCHP	1.19	1.93

Table 27: Loose and medium working points for the Track Counting High Efficiency and Track Counting High Purity B-tagging algorithms

620 In the list above “loose” and “medium” refer to the loose and medium working points for each
 621 algorithm, their values can be found in Table 27.

622 Figure 24 shows the 95 % C.L. upper limit on expected $\sigma \times \text{BR}$ for the process $H \rightarrow ZZ \rightarrow \ell\ell +$
 623 2jets for each tagging case and for various masses of the Higgs Boson. Although a single case
 624 does not always give the lowest upper limit, on average case 3 for the Track Counting High
 625 Efficiency gives the best overall results and should thus be used for the tagging requirement.

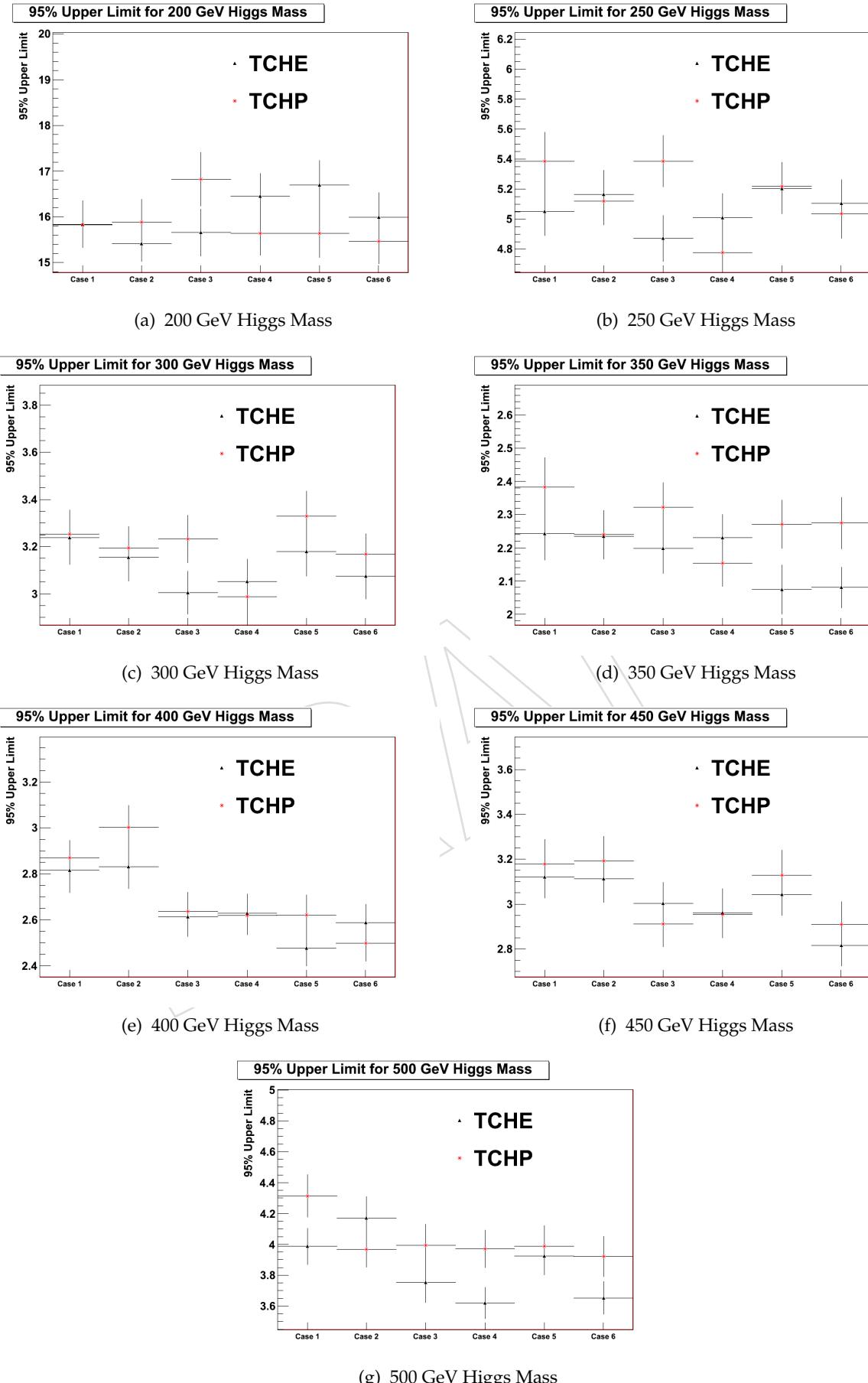


Figure 24: 95 % upper Limit for each Higgs Mass. Each plot shows how the tagging algorithms perform for each case.

626 B Z+HF Background Checks

627 B.1 Z+HF Background from data

628 A selection of a relatively pure control sample of Zbb events, removing the possible presence
 629 of the Higgs signal, is achieved by changing the signal selection by imposing a veto on the
 630 presence of a Z decaying into a $b\bar{b}$ pair.

631 We select events passing the Higgs preselection described in Section 3.1, we require the pres-
 632 ence of two muons with $p_T > 20$ GeV, $d_{xy} < 0.02$ cm, a di-lepton invariant mass within 10 GeV
 633 of the Z mass, a positive b-tag of both jets and a \cancel{E}_T smaller than 35 GeV, in order to reduce
 634 the $t\bar{t}$ background. The resulting distribution of the mass difference $m_{\ell\ell jj} - m_{jj}$ as a function of
 635 the di-jet invariant mass is shown in Figure 25. In order to reduce the possible signal contam-
 636 ination, we require the di-jet invariant mass to be greater than 120 GeV. This defines the Zbb
 637 enriched region.

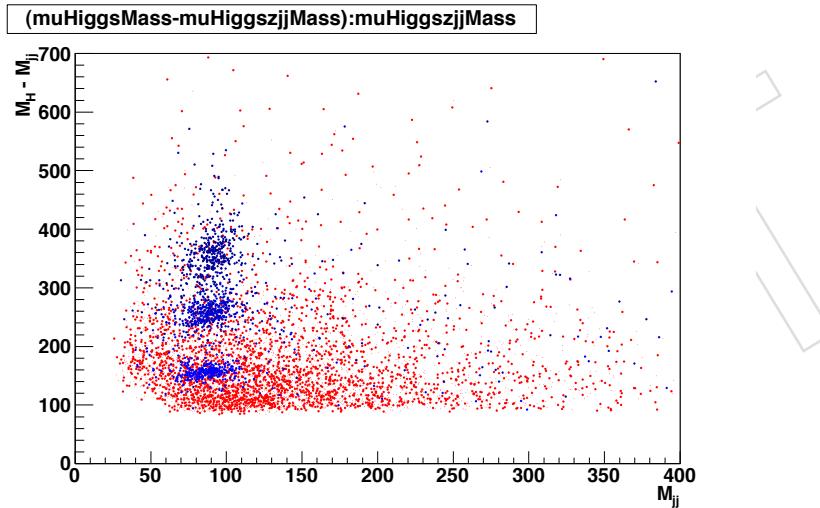


Figure 25: Distribution of the di-jet invariant mass, m_{jj} , versus the difference of the Higgs reconstructed mass, $m_{\ell\ell jj}$ minus m_{jj} . Three Higgs mass hypotheses (250, 350 and 450 GeV) are shown in blue, $Z+jets$ is shown in red.

638 The number of expected candidates in the Zbb -enriched region is shown in Table 28. In this
 639 region there is some contamination from $Z+light$ -quark jets, which are expected to have very
 640 similar kinematic properties. A 24% residual contamination from $t\bar{t}$ processes is also present
 641 in the sample. This contribution is relatively well known, measured by CMS, and can be sub-
 642 tracted using MC.

Process	Expected events (1fb^{-1})
Higgs ($m_H = 350$ GeV)	0.34
Zbb, Zcc	23.42
$Z+jets$ (light quarks)	5.88
$t\bar{t}$	5.53
ZZ	0.33
WZ	0.33

Table 28: Number of expected events surviving the Zbb -enriched selection for an integrated luminosity of 1fb^{-1} .

643 The number of expected events in the signal region is shown in Table 29. This region is obtained
 644 replacing the cut $m_{jj} > 120$ GeV with the request that $|m_{jj} - m_Z| < 15$ GeV.

Process	Expected events (1fb^{-1})
Higgs ($m_H = 350$ GeV)	2.00
Zbb, Zcc	9.09
Z+jets (light quarks)	0.86
t̄t	1.28
ZZ	2.45
WZ	0.02

Table 29: Number of expected events surviving the Zbb-enriched selection with the inversion of the di-jet selection cut to $|m_{jj} - m_Z| < 15$ GeV for an integrated luminosity of 1fb^{-1} .

645 The expected number of Z+heavy-quark jets in the Zbb-enriched region is about 2.56 times the
 646 number of events in the signal region ($|m_{jj} - m_Z| < 15$ GeV). This allows to determine the
 647 amount of background in the signal region with a relative uncertainty of about 20% assuming
 648 an integrated luminosity of 1fb^{-1} .

649 A comparison of the main kinematic distributions of Zbb events in the signal and control re-
 650 gions is shown in Figures 26, 27, showing the distributions of $p_T(\ell\ell)$, $\Delta R(jj)$, m_{jj} and $m_{\ell\ell jj}$ re-
 651 spectively. The agreement of the distributions in the two regions is generally good.

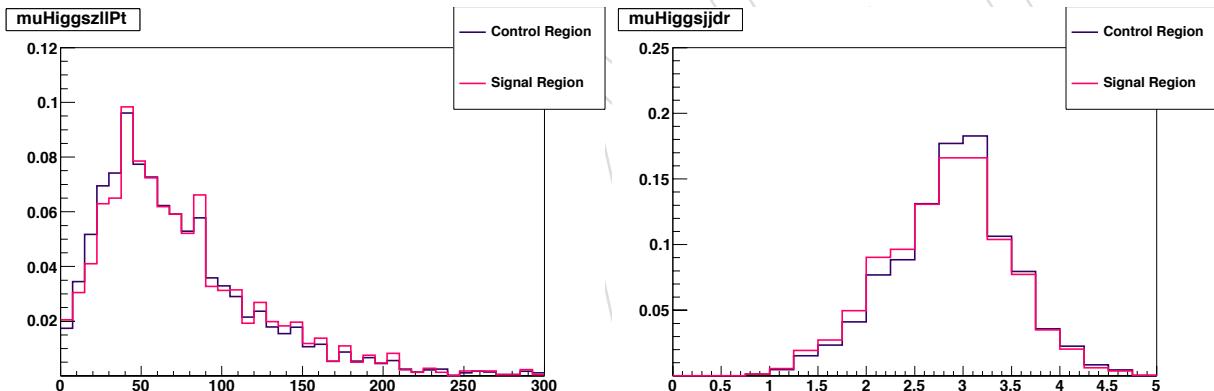


Figure 26: $p_T(\ell\ell)$ (left) and $\Delta R(jj)$ (right) for the Zbb-enriched control region and the signal region ($|m_{jj} - m_Z| < 15$ GeV).

652 B.2 Z+HF Shape Systematics from MC

653 The main background in the analysis is the Z+heavy flavor production.
 654 We compare two different MC generators, Alpgen and Madgraph, as a preliminary estimation
 655 of the uncertainty in this shape. Figure 28 shows the $M_{j_\ell ll}$ invariant mass using Z+jets events
 656 generated with Alpgen and Madgraph MC. Due to the lack of statistics in the Madgraph sample
 657 the comparison is done without requiring b-tagging.
 658 The difference in yields between both samples in final regions is $\sim 20\%$. However, we use a
 659 30% uncertainty being conservative due to the lack of statistics in the available MC samples.

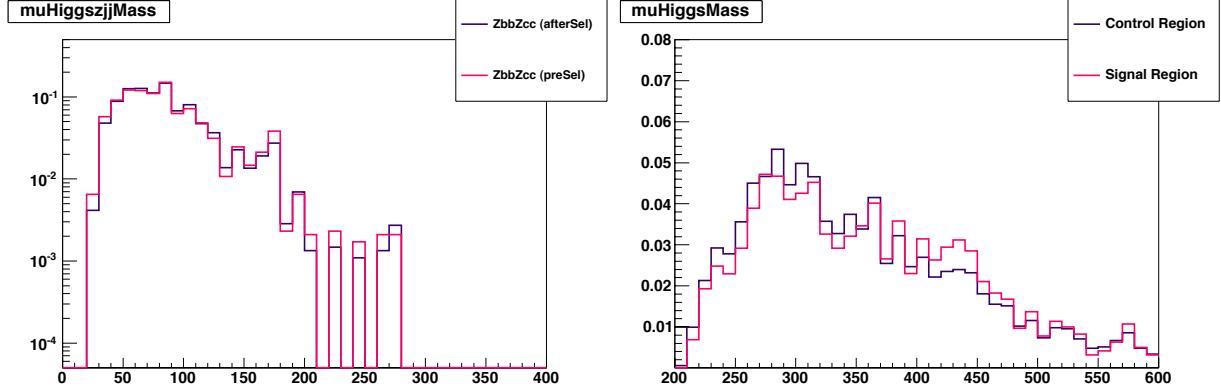


Figure 27: m_{jj} (left) and $m_{\ell\ell jj}$ (right) for the Zbb-enriched control region and the signal region ($|m_{jj} - m_Z| < 15 \text{ GeV}$).

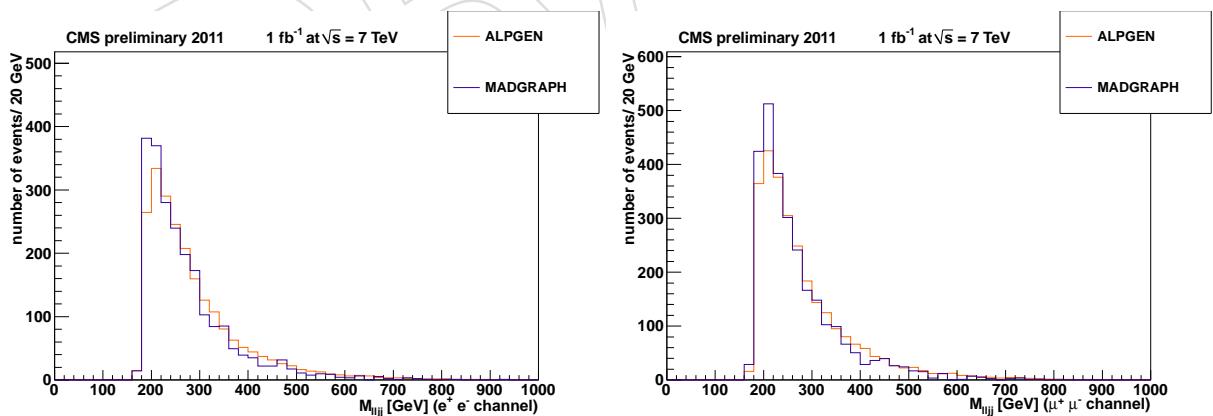


Figure 28: M_{jll} shape comparison using Algen and Madgraph generators to produce $Z+jets$. Electron (left) and muon (right) selections are shown.

660 C Kinematic Optimization

661 The kinematic optimization is a cut-based procedure, as function of the j_{ll} reconstructed mass,
 662 as explained in section 3.4.

663 We check the performance of the analysis applying the described optimization to MC simu-
 664 lation normalized to 1 fb^{-1} . The results are shown in Figure 29 where the Higgs signal with
 665 mass of 400 GeV is shown.

666 Using the events in Figure 29 and cutting on different mass ranges is possible to perform a cut
 667 and count experiment and extract a 95%CL limit as in shown in Table 30 to 35.

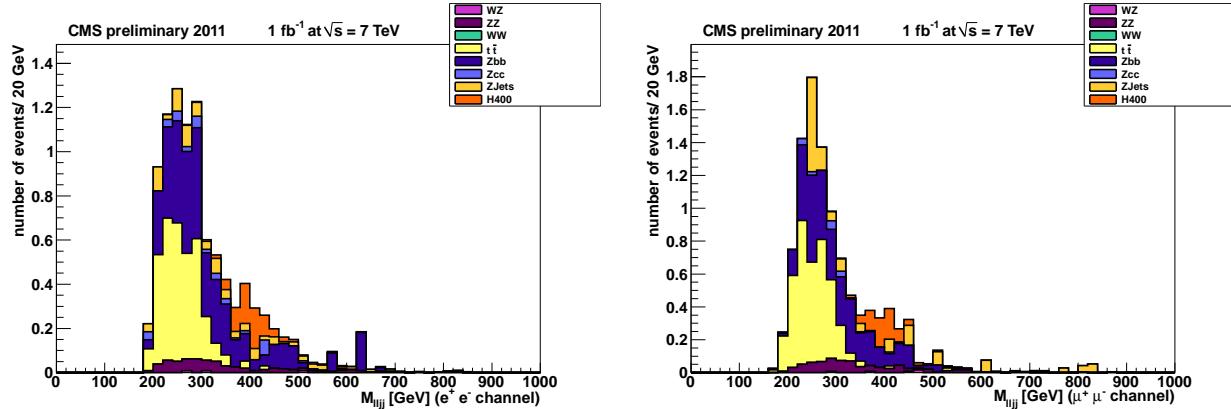


Figure 29: Background distribution of the $m_{j_{ll}}$ for electron (left) and muon (right) events in 1 fb^{-1} . Higgs signal with an invariant mass of 400 GeV is shown.

Process	Electrons	Muons	Limit@95%CL
$M_H = 250$ [237-260]	0.37	0.34	(BayesianSimple, 1000 toys)
Z+Jets	0.57	0.18	$r < 9.38$
ZZ	0.11	0.10	
WZ	0	0.01	
WW	0	0	
Zbb	1.10	0.88	
Zcc	0.05	0.09	
$t\bar{t}$	1.47	1.24	
Total Bkg	3.30	2.50	

Table 30: Number of events expected in the final region within a Higgs mass range of [237-260]. Also the expected limit is shown for the cut and count experiment

Process	Electrons	Muons	Limit@95%CL
$M_H = 300$ [280-232]	0.27	0.29	(BayesianSimple, 1000 toys)
Z+Jets	0.13	0.15	$r < 10.55$
ZZ	0.19	0.14	
WZ	0.01	0.01	
WW	0	0	
Zbb	0.71	0.95	
Zcc	0.09	0.07	
$t\bar{t}$	0.73	0.77	
Total Bkg	1.85	2.09	

Table 31: Number of events expected in the final region within a Higgs mass range of [280-232]. Also the expected limit is shown for the cut and count experiment

Process	Electrons	Muons	Limit@95%CL
$M_H = 350$ [330-380]	0.61	0.58	(BayesianSimple, 1000 toys)
Z+Jets	0.05	0.10	$r < 4.06$
ZZ	0.11	0.08	
WZ	0.01	0	
WW	0	0	
Zbb	0.68	0.60	
Zcc	0.01	0.07	
$t\bar{t}$	0.05	0.11	
Total Bkg	0.91	0.96	

Table 32: Number of events expected in the final region within a Higgs mass range of [330-380]. Also the expected limit is shown for the cut and count experiment

Process	Electrons	Muons	Limit@95%CL
$M_H = 400$ [390-460]	0.41	0.42	(BayesianSimple, 1000 toys)
Z+Jets	0.20	0.13	$r < 4.86$
ZZ	0.11	0.06	
WZ	0.02	0	
WW	0	0	
Zbb	0.41	0.31	
Zcc	0.02	0.07	
$t\bar{t}$	0.02	0.03	
Total Bkg	0.78	0.60	

Table 33: Number of events expected in the final region within a Higgs mass range of [390-460]. Also the expected limit is shown for the cut and count experiment

Process	Electrons	Muons	Limit@95%CL
$M_H = 450$ [420-550]	0.38	0.39	(BayesianSimple, 1000 toys)
Z+Jets	0.21	0.15	$r < 5.13$
ZZ	0.12	0.09	
WZ	0.02	0.009	
WW	0	0	
Zbb	0.37	0.42	
Zcc	0.01	0.07	
$t\bar{t}$	0.02	0.02	
Total Bkg	0.75	0.76	

Table 34: Number of events expected in the final region within a Higgs mass range of [420-550]. Also the expected limit is shown for the cut and count experiment

Process	Electrons	Muons	Limit@95%CL
$M_H = 500$ [470-1000]	0.39	0.26	(BayesianSimple, 1000 toys)
Z+Jets	0.15	0.17	$r < 6.18$
ZZ	0.09	0.11	
WZ	0.009	0.01	
WW	0	0	
Zbb	0.42	0.44	
Zcc	0.07	0.01	
$t\bar{t}$	0.02	0	
Total Bkg	0.75	0.75	

Table 35: Number of events expected in the final region within a Higgs mass range of [470-1000]. Also the expected limit is shown for the cut and count experiment

668 **D Mistag Efficiencies and Scale Factors**

$ \eta_{min} $	$ \eta_{max} $	$ p_{T,min} $	$ p_{T,max} $	lightjet-eff	lightjet-eff-err	SF_{light}	$SF_{light-err}$
0	0.8	20	30	0.003966	0.001208	1.30895	0.28861
0	0.8	30	40	0.00611	0.001681	1.3127	0.27687
0	0.8	40	50	0.008224	0.002146	1.31524	0.26591
0	0.8	50	60	0.010307	0.002603	1.31663	0.25568
0	0.8	60	70	0.01236	0.003053	1.31696	0.24617
0	0.8	70	80	0.014383	0.003494	1.31628	0.23735
0	0.8	80	90	0.016376	0.003928	1.31468	0.2292
0	0.8	90	100	0.018338	0.004354	1.31222	0.22167
0	0.8	100	110	0.02027	0.004773	1.30899	0.21476
0	0.8	110	120	0.022172	0.005183	1.30504	0.20843
0	0.8	120	130	0.024044	0.005586	1.30047	0.20266
0	0.8	130	140	0.025886	0.005981	1.29533	0.19742
0	0.8	140	150	0.027697	0.006368	1.2897	0.19268
0	0.8	150	160	0.029478	0.006748	1.28366	0.18841
0	0.8	160	170	0.031229	0.007119	1.27728	0.1846
0	0.8	170	180	0.03295	0.007483	1.27062	0.18121
0	0.8	180	190	0.03464	0.007839	1.26377	0.17822
0	0.8	190	200	0.0363	0.008187	1.2568	0.17559
0	0.8	200	210	0.03793	0.008528	1.24977	0.17332
0	0.8	210	220	0.03953	0.00886	1.24276	0.17136
0	0.8	220	230	0.041099	0.009185	1.23585	0.16969
0	0.8	230	240	0.042639	0.009502	1.22911	0.16828
0	0.8	240	250	0.044148	0.009811	1.22261	0.16712
0	0.8	250	260	0.045627	0.010113	1.21642	0.16616
0	0.8	260	270	0.047075	0.010406	1.21061	0.16539
0	0.8	270	280	0.048493	0.010692	1.20527	0.16478
0	0.8	280	290	0.049882	0.01097	1.20046	0.1643
0	0.8	290	300	0.05124	0.011241	1.19625	0.16393
0	0.8	300	310	0.052567	0.011503	1.19272	0.16364
0	0.8	310	320	0.053865	0.011758	1.18994	0.16339
0	0.8	320	330	0.055132	0.012005	1.18798	0.16318
0	0.8	330	340	0.056369	0.012244	1.18692	0.16296
0	0.8	340	350	0.057576	0.012475	1.18683	0.16272
0	0.8	350	360	0.058752	0.012698	1.18778	0.16242
0	0.8	360	370	0.059898	0.012914	1.18984	0.16205
0	0.8	370	380	0.061015	0.013122	1.19309	0.16156
0	0.8	380	390	0.0621	0.013322	1.1976	0.16095
0	0.8	390	400	0.063156	0.013514	1.20345	0.16017
0	0.8	400	410	0.064181	0.013699	1.2107	0.15921
0	0.8	410	999	0.078288	0.01671	1.21943	0.15803

Table 36: Efficiencies, Scale Factors and errors for mistags using TCHEM. Part I

$ \eta_{min} $	$ \eta_{max} $	$ p_{T,min} $	$ p_{T,max} $	lightjet-eff	lightjet-eff-err	SF_{light}	$SF_{light-err}$
0.8	1.6	20	30	0.003958	0.00121	1.31897	0.29197
0.8	1.6	30	40	0.006109	0.001647	1.31014	0.27345
0.8	1.6	40	50	0.008261	0.002085	1.30131	0.25645
0.8	1.6	50	60	0.010414	0.002524	1.29253	0.24091
0.8	1.6	60	70	0.012568	0.002963	1.2838	0.22678
0.8	1.6	70	80	0.014724	0.003402	1.27517	0.21398
0.8	1.6	80	90	0.016881	0.003841	1.26666	0.20247
0.8	1.6	90	100	0.019039	0.004281	1.25829	0.19217
0.8	1.6	100	110	0.021199	0.004722	1.25009	0.18303
0.8	1.6	110	120	0.02336	0.005163	1.2421	0.175
0.8	1.6	120	130	0.025522	0.005604	1.23433	0.168
0.8	1.6	130	140	0.027685	0.006046	1.22683	0.16198
0.8	1.6	140	150	0.029849	0.006488	1.2196	0.15688
0.8	1.6	150	160	0.032015	0.00693	1.21269	0.15264
0.8	1.6	160	170	0.034182	0.007373	1.20611	0.14919
0.8	1.6	170	180	0.03635	0.007817	1.1999	0.14648
0.8	1.6	180	190	0.03852	0.008261	1.19409	0.14444
0.8	1.6	190	200	0.04069	0.008705	1.1887	0.14303
0.8	1.6	200	210	0.042862	0.00915	1.18375	0.14216
0.8	1.6	210	220	0.045036	0.009595	1.17929	0.1418
0.8	1.6	220	230	0.04721	0.01004	1.17533	0.14186
0.8	1.6	230	240	0.049386	0.010486	1.1719	0.1423
0.8	1.6	240	250	0.051563	0.010932	1.16903	0.14305
0.8	1.6	250	260	0.053741	0.011379	1.16675	0.14406
0.8	1.6	260	270	0.05592	0.011826	1.16508	0.14526
0.8	1.6	270	280	0.058101	0.012274	1.16406	0.14659
0.8	1.6	280	290	0.060283	0.012722	1.16371	0.148
0.8	1.6	290	300	0.062466	0.01317	1.16405	0.14941
0.8	1.6	300	310	0.06465	0.013619	1.16512	0.15078
0.8	1.6	310	320	0.066836	0.014068	1.16695	0.15203
0.8	1.6	320	330	0.069023	0.014518	1.16955	0.15312
0.8	1.6	330	340	0.071211	0.014968	1.17296	0.15397
0.8	1.6	340	350	0.073401	0.015419	1.17721	0.15454
0.8	1.6	350	360	0.075591	0.01587	1.18232	0.15475
0.8	1.6	360	370	0.077783	0.016321	1.18833	0.15455
0.8	1.6	370	380	0.079976	0.016773	1.19525	0.15388
0.8	1.6	380	390	0.082171	0.017225	1.20312	0.15268
0.8	1.6	390	400	0.084366	0.017677	1.21196	0.15088
0.8	1.6	400	410	0.086563	0.01813	1.2218	0.14844
0.8	1.6	410	999	0.092652	0.019406	1.23267	0.14527

Table 37: Efficiencies, Scale Factors and errors for mistags using TCHEM. Part II

$ \eta_{min} $	$ \eta_{max} $	$ p_{T,min} $	$ p_{T,max} $	lightjet-eff	lightjet-eff-err	SF_{light}	$SF_{light-err}$
1.6	2.4	20	30	0.003327	0.000902	1.04679	0.26316
1.6	2.4	30	40	0.005602	0.001357	1.05954	0.24352
1.6	2.4	40	50	0.007879	0.001814	1.07138	0.22573
1.6	2.4	50	60	0.010157	0.002271	1.08233	0.20971
1.6	2.4	60	70	0.012435	0.002729	1.09242	0.19537
1.6	2.4	70	80	0.014715	0.003187	1.10169	0.18264
1.6	2.4	80	90	0.016996	0.003647	1.11016	0.17144
1.6	2.4	90	100	0.019277	0.004107	1.11787	0.16168
1.6	2.4	100	110	0.02156	0.004567	1.12485	0.1533
1.6	2.4	110	120	0.023844	0.005029	1.13114	0.14621
1.6	2.4	120	130	0.026128	0.005491	1.13675	0.14033
1.6	2.4	130	140	0.028414	0.005954	1.14173	0.13558
1.6	2.4	140	150	0.030701	0.006418	1.14611	0.13188
1.6	2.4	150	160	0.032988	0.006882	1.14991	0.12916
1.6	2.4	160	170	0.035277	0.007347	1.15318	0.12734
1.6	2.4	170	180	0.037567	0.007813	1.15594	0.12633
1.6	2.4	180	190	0.039858	0.008279	1.15822	0.12605
1.6	2.4	190	200	0.042149	0.008747	1.16005	0.12644
1.6	2.4	200	210	0.044442	0.009214	1.16147	0.1274
1.6	2.4	210	220	0.046736	0.009683	1.16251	0.12886
1.6	2.4	220	230	0.04903	0.010153	1.16321	0.13074
1.6	2.4	230	240	0.051326	0.010623	1.16358	0.13296
1.6	2.4	240	250	0.053623	0.011094	1.16366	0.13544
1.6	2.4	250	260	0.055921	0.011565	1.16349	0.1381
1.6	2.4	260	270	0.058219	0.012037	1.1631	0.14087
1.6	2.4	270	280	0.060519	0.01251	1.16251	0.14366
1.6	2.4	280	290	0.06282	0.012984	1.16177	0.14639
1.6	2.4	290	300	0.065122	0.013459	1.16089	0.14899
1.6	2.4	300	310	0.067424	0.013934	1.15992	0.15138
1.6	2.4	310	320	0.069728	0.01441	1.15889	0.15347
1.6	2.4	320	330	0.072033	0.014886	1.15782	0.15519
1.6	2.4	330	340	0.074339	0.015364	1.15675	0.15646
1.6	2.4	340	350	0.076646	0.015842	1.15571	0.1572
1.6	2.4	350	360	0.078953	0.01632	1.15473	0.15732
1.6	2.4	360	370	0.081262	0.0168	1.15385	0.15676
1.6	2.4	370	380	0.083572	0.01728	1.15309	0.15543
1.6	2.4	380	390	0.085883	0.017761	1.15248	0.15326
1.6	2.4	390	400	0.088194	0.018243	1.15207	0.15015
1.6	2.4	400	410	0.090507	0.018725	1.15187	0.14604
1.6	2.4	410	999	0.083835	0.017345	1.15193	0.14084

Table 38: Efficiencies, Scale Factors and errors for mistags using TCHEM. Part III

$ \eta_{min} $	$ \eta_{max} $	$ p_{T,min} $	$ p_{T,max} $	lightjet-eff	lightjet-eff-err	SF _{light}	SF _{light-err}
0	0.5	20	30	0.027579	0.005726	1.1926	0.14437
0	0.5	30	40	0.04283	0.008816	1.13739	0.13122
0	0.5	40	50	0.057605	0.011807	1.10553	0.12395
0	0.5	50	60	0.071916	0.0147	1.08581	0.11948
0	0.5	60	70	0.085769	0.017498	1.0732	0.11656
0	0.5	70	80	0.099175	0.020202	1.06515	0.11459
0	0.5	80	90	0.112143	0.022815	1.06016	0.11324
0	0.5	90	100	0.124682	0.025339	1.05735	0.11232
0	0.5	100	110	0.136801	0.027776	1.05614	0.1117
0	0.5	110	120	0.14851	0.030129	1.05614	0.11131
0	0.5	120	130	0.159817	0.032399	1.05707	0.1111
0	0.5	130	140	0.170731	0.034588	1.05874	0.11102
0	0.5	140	150	0.181263	0.036699	1.06101	0.11104
0	0.5	150	160	0.19142	0.038733	1.06376	0.11116
0	0.5	160	170	0.201212	0.040694	1.06691	0.11134
0	0.5	170	180	0.210649	0.042582	1.0704	0.11159
0	0.5	180	190	0.219739	0.044401	1.07417	0.11188
0	0.5	190	200	0.228492	0.046152	1.07818	0.11221
0	0.5	200	210	0.236917	0.047838	1.0824	0.11259
0	0.5	210	220	0.245023	0.04946	1.08679	0.11299
0	0.5	220	230	0.252819	0.051021	1.09135	0.11342
0	0.5	230	240	0.260314	0.052523	1.09604	0.11387
0	0.5	240	250	0.267518	0.053968	1.10085	0.11434
0	0.5	250	260	0.27444	0.055358	1.10577	0.11483
0	0.5	260	270	0.281088	0.056695	1.11079	0.11534
0	0.5	270	280	0.287472	0.057982	1.11589	0.11586
0	0.5	280	290	0.293602	0.059221	1.12106	0.11639
0	0.5	290	300	0.299486	0.060413	1.12631	0.11694
0	0.5	300	310	0.305133	0.061561	1.13162	0.11749
0	0.5	310	320	0.310553	0.062667	1.13698	0.11806
0	0.5	320	330	0.315755	0.063734	1.14239	0.11863
0	0.5	330	340	0.320748	0.064762	1.14786	0.11921
0	0.5	340	350	0.325542	0.065755	1.15336	0.1198
0	0.5	350	360	0.330144	0.066715	1.15891	0.12039
0	0.5	360	370	0.334566	0.067644	1.16449	0.12099
0	0.5	370	380	0.338815	0.068543	1.1701	0.1216
0	0.5	380	390	0.3429	0.069415	1.17575	0.12221
0	0.5	390	400	0.346832	0.070263	1.18142	0.12283
0	0.5	400	410	0.350619	0.071087	1.18712	0.12345
0	0.5	410	999	0.354395	0.071853	1.19284	0.12407

Table 39: Efficiencies, Scale Factors and errors for mistags using TCHEL. Part I

$ \eta_{min} $	$ \eta_{max} $	$ p_{T,min} $	$ p_{T,max} $	lightjet-eff	lightjet-eff-err	SF_{light}	$SF_{light-err}$
0.5	1	20	30	0.02837	0.005898	1.16097	0.13972
0.5	1	30	40	0.045219	0.009309	1.11035	0.12791
0.5	1	40	50	0.061432	0.012587	1.08207	0.12127
0.5	1	50	60	0.077021	0.015736	1.06481	0.11711
0.5	1	60	70	0.092003	0.018757	1.05379	0.11434
0.5	1	70	80	0.106389	0.021656	1.04666	0.11242
0.5	1	80	90	0.120195	0.024434	1.04211	0.11106
0.5	1	90	100	0.133434	0.027096	1.03937	0.11009
0.5	1	100	110	0.14612	0.029643	1.03794	0.1094
0.5	1	110	120	0.158268	0.032079	1.0375	0.10892
0.5	1	120	130	0.169891	0.034408	1.03781	0.1086
0.5	1	130	140	0.181003	0.036632	1.03871	0.1084
0.5	1	140	150	0.191619	0.038755	1.04008	0.10831
0.5	1	150	160	0.201751	0.04078	1.04184	0.1083
0.5	1	160	170	0.211415	0.04271	1.04392	0.10835
0.5	1	170	180	0.220625	0.044548	1.04626	0.10846
0.5	1	180	190	0.229393	0.046297	1.04882	0.10862
0.5	1	190	200	0.237735	0.04796	1.05158	0.10882
0.5	1	200	210	0.245664	0.049541	1.05449	0.10905
0.5	1	210	220	0.253194	0.051042	1.05755	0.10931
0.5	1	220	230	0.26034	0.052467	1.06072	0.10959
0.5	1	230	240	0.267115	0.05382	1.064	0.1099
0.5	1	240	250	0.273533	0.055102	1.06738	0.11023
0.5	1	250	260	0.279608	0.056317	1.07084	0.11057
0.5	1	260	270	0.285354	0.057469	1.07437	0.11093
0.5	1	270	280	0.290786	0.058561	1.07797	0.1113
0.5	1	280	290	0.295917	0.059595	1.08163	0.11169
0.5	1	290	300	0.300761	0.060574	1.08534	0.11209
0.5	1	300	310	0.305332	0.061503	1.08909	0.11249
0.5	1	310	320	0.309644	0.062384	1.09289	0.11291
0.5	1	320	330	0.313712	0.06322	1.09674	0.11333
0.5	1	330	340	0.317548	0.064015	1.10061	0.11376
0.5	1	340	350	0.321168	0.064771	1.10452	0.1142
0.5	1	350	360	0.324585	0.065491	1.10846	0.11464
0.5	1	360	370	0.327813	0.06618	1.11243	0.11509
0.5	1	370	380	0.330867	0.066839	1.11642	0.11555
0.5	1	380	390	0.333759	0.067472	1.12044	0.11601
0.5	1	390	400	0.336505	0.068083	1.12448	0.11647
0.5	1	400	410	0.339118	0.068674	1.12853	0.11694
0.5	1	410	999	0.346231	0.070115	1.13261	0.11741

Table 40: Efficiencies, Scale Factors and errors for mistags using TCHEL. Part II

$ \eta_{min} $	$ \eta_{max} $	$ p_{T,min} $	$ p_{T,max} $	lightjet-eff	lightjet-eff-err	SF_{light}	$SF_{light-err}$
1	1.5	20	30	0.026104	0.005427	1.10451	0.12748
1	1.5	30	40	0.045451	0.009323	1.09097	0.12246
1	1.5	40	50	0.064032	0.013062	1.0809	0.11881
1	1.5	50	60	0.081866	0.016649	1.07339	0.11612
1	1.5	60	70	0.09897	0.020087	1.06781	0.11411
1	1.5	70	80	0.11536	0.02338	1.06374	0.1126
1	1.5	80	90	0.131055	0.026531	1.06087	0.11148
1	1.5	90	100	0.146071	0.029546	1.05895	0.11065
1	1.5	100	110	0.160426	0.032426	1.05783	0.11006
1	1.5	110	120	0.174137	0.035177	1.05735	0.10965
1	1.5	120	130	0.187221	0.037802	1.05742	0.1094
1	1.5	130	140	0.199695	0.040305	1.05795	0.10927
1	1.5	140	150	0.211578	0.042689	1.05888	0.10924
1	1.5	150	160	0.222886	0.044958	1.06015	0.1093
1	1.5	160	170	0.233636	0.047116	1.0617	0.10944
1	1.5	170	180	0.243845	0.049167	1.06352	0.10964
1	1.5	180	190	0.253532	0.051115	1.06556	0.10989
1	1.5	190	200	0.262713	0.052963	1.0678	0.11019
1	1.5	200	210	0.271405	0.054715	1.07021	0.11053
1	1.5	210	220	0.279626	0.056374	1.07278	0.11091
1	1.5	220	230	0.287393	0.057946	1.07549	0.11131
1	1.5	230	240	0.294723	0.059433	1.07832	0.11175
1	1.5	240	250	0.301634	0.060839	1.08127	0.11221
1	1.5	250	260	0.308143	0.062169	1.08432	0.1127
1	1.5	260	270	0.314267	0.063425	1.08747	0.1132
1	1.5	270	280	0.320023	0.064611	1.0907	0.11372
1	1.5	280	290	0.325429	0.065732	1.094	0.11426
1	1.5	290	300	0.330502	0.066791	1.09738	0.11482
1	1.5	300	310	0.33526	0.067792	1.10082	0.11538
1	1.5	310	320	0.339718	0.068738	1.10433	0.11596
1	1.5	320	330	0.343896	0.069634	1.10788	0.11655
1	1.5	330	340	0.347809	0.070484	1.11149	0.11715
1	1.5	340	350	0.351476	0.07129	1.11515	0.11776
1	1.5	350	360	0.354913	0.072057	1.11885	0.11838
1	1.5	360	370	0.358139	0.072788	1.12258	0.11901
1	1.5	370	380	0.361169	0.073488	1.12636	0.11964
1	1.5	380	390	0.364022	0.074159	1.13018	0.12028
1	1.5	390	400	0.366715	0.074807	1.13402	0.12093
1	1.5	400	410	0.369264	0.075434	1.1379	0.12158
1	1.5	410	999	0.378831	0.077389	1.1418	0.12224

Table 41: Efficiencies, Scale Factors and errors for mistags using TCHEL. Part III

$ \eta_{min} $	$ \eta_{max} $	$ p_{T,min} $	$ p_{T,max} $	lightjet-eff	lightjet-eff-err	SF_{light}	$SF_{light-err}$
1.5	2	20	30	0.030334	0.00628	1.09064	0.11867
1.5	2	30	40	0.051453	0.010532	1.09114	0.11784
1.5	2	40	50	0.071417	0.01455	1.09173	0.1171
1.5	2	50	60	0.090268	0.018342	1.09241	0.11644
1.5	2	60	70	0.108043	0.021916	1.09319	0.11587
1.5	2	70	80	0.124782	0.025281	1.09406	0.11538
1.5	2	80	90	0.140526	0.028444	1.09501	0.11497
1.5	2	90	100	0.155313	0.031415	1.09604	0.11463
1.5	2	100	110	0.169182	0.034201	1.09716	0.11437
1.5	2	110	120	0.182174	0.03681	1.09834	0.11418
1.5	2	120	130	0.194328	0.039251	1.0996	0.11405
1.5	2	130	140	0.205683	0.041532	1.10093	0.11399
1.5	2	140	150	0.216278	0.043662	1.10233	0.11399
1.5	2	150	160	0.226153	0.045648	1.10379	0.11405
1.5	2	160	170	0.235348	0.047499	1.1053	0.11416
1.5	2	170	180	0.243902	0.049223	1.10688	0.11433
1.5	2	180	190	0.251854	0.050829	1.10851	0.11455
1.5	2	190	200	0.259245	0.052324	1.11019	0.11482
1.5	2	200	210	0.266112	0.053718	1.11191	0.11513
1.5	2	210	220	0.272496	0.055017	1.11368	0.11549
1.5	2	220	230	0.278437	0.056231	1.11549	0.11589
1.5	2	230	240	0.283973	0.057367	1.11734	0.11632
1.5	2	240	250	0.289145	0.058435	1.11923	0.11679
1.5	2	250	260	0.293991	0.059442	1.12114	0.11728
1.5	2	260	270	0.298551	0.060396	1.12309	0.11781
1.5	2	270	280	0.302865	0.061306	1.12505	0.11836
1.5	2	280	290	0.306971	0.06218	1.12704	0.11894
1.5	2	290	300	0.31091	0.063026	1.12905	0.11954
1.5	2	300	310	0.314721	0.063853	1.13108	0.12015
1.5	2	310	320	0.318443	0.064669	1.13311	0.12078
1.5	2	320	330	0.322117	0.065482	1.13516	0.12142
1.5	2	330	340	0.32578	0.0663	1.13721	0.12208
1.5	2	340	350	0.329473	0.067132	1.13927	0.12273
1.5	2	350	360	0.333235	0.067985	1.14132	0.1234
1.5	2	360	370	0.337106	0.068869	1.14338	0.12406
1.5	2	370	380	0.341125	0.069791	1.14542	0.12472
1.5	2	380	390	0.345331	0.07076	1.14746	0.12538
1.5	2	390	400	0.349765	0.071784	1.14948	0.12603
1.5	2	400	410	0.354465	0.072871	1.15149	0.12667
1.5	2	410	999	0.381495	0.078428	1.15347	0.1273

Table 42: Efficiencies, Scale Factors and errors for mistags using TCHEL. Part IV

$ \eta_{min} $	$ \eta_{max} $	$ p_{T,min} $	$ p_{T,max} $	lightjet-eff	lightjet-eff-err	SF _{light}	SF _{light-err}
2	2.4	20	30	0.029576	0.006119	1.06582	0.11628
2	2.4	30	40	0.047334	0.009739	1.06828	0.11596
2	2.4	40	50	0.063891	0.013101	1.07105	0.11567
2	2.4	50	60	0.079299	0.016218	1.07408	0.11541
2	2.4	60	70	0.093609	0.0191	1.07734	0.11517
2	2.4	70	80	0.106873	0.021762	1.08078	0.11497
2	2.4	80	90	0.119141	0.024214	1.08438	0.1148
2	2.4	90	100	0.130467	0.026469	1.08809	0.11466
2	2.4	100	110	0.1409	0.028539	1.09187	0.11457
2	2.4	110	120	0.150493	0.030438	1.09569	0.11451
2	2.4	120	130	0.159297	0.032176	1.0995	0.11449
2	2.4	130	140	0.167363	0.033766	1.10326	0.11451
2	2.4	140	150	0.174743	0.035221	1.10695	0.11458
2	2.4	150	160	0.181489	0.036552	1.11051	0.11469
2	2.4	160	170	0.187651	0.037773	1.11392	0.11485
2	2.4	170	180	0.193281	0.038895	1.11713	0.11507
2	2.4	180	190	0.198432	0.03993	1.1201	0.11533
2	2.4	190	200	0.203153	0.040891	1.1228	0.11564
2	2.4	200	210	0.207497	0.04179	1.12518	0.11601
2	2.4	210	220	0.211515	0.04264	1.12721	0.11644
2	2.4	220	230	0.215259	0.043452	1.12885	0.11692
2	2.4	230	240	0.218779	0.044239	1.13006	0.11746
2	2.4	240	250	0.222129	0.045013	1.13081	0.11807
2	2.4	250	260	0.225358	0.045787	1.13104	0.11873
2	2.4	260	270	0.228518	0.046572	1.13073	0.11947
2	2.4	270	280	0.231662	0.047381	1.12984	0.12027
2	2.4	280	290	0.234839	0.048226	1.12832	0.12114
2	2.4	290	300	0.238103	0.04912	1.12614	0.12207
2	2.4	300	310	0.241504	0.050075	1.12326	0.12309
2	2.4	310	320	0.245093	0.051102	1.11964	0.12417
2	2.4	320	330	0.248923	0.052215	1.11525	0.12533
2	2.4	330	340	0.253044	0.053426	1.11003	0.12657
2	2.4	340	350	0.257508	0.054746	1.10397	0.12789
2	2.4	350	360	0.262367	0.056188	1.09701	0.12928
2	2.4	360	370	0.267672	0.057765	1.08911	0.13076
2	2.4	370	380	0.273475	0.059488	1.08025	0.13233
2	2.4	380	390	0.279826	0.06137	1.07038	0.13398
2	2.4	390	400	0.286778	0.063424	1.05946	0.13572
2	2.4	400	410	0.294381	0.06566	1.04745	0.13755
2	2.4	410	999	0.187409	0.0418	1.03431	0.13947

Table 43: Efficiencies, Scale Factors and errors for mistags using TCHEL. Part V