

A Collider Search for Dark Matter Produced in Association
with a Higgs Boson with the CMS Detector at the 13 TeV LHC

By

DUSTIN RAY BURNS

B.S. (Georgia Institute of Technology) 2011

M.S. (University of California, Davis) 2012

DISSERTATION

Submitted in partial satisfaction of the requirements for the degree of

DOCTOR OF PHILOSOPHY

in

Physics

in the

OFFICE OF GRADUATE STUDIES

of the

UNIVERSITY OF CALIFORNIA

DAVIS

Approved:

Chair Michael Mulhearn

Robin Erbacher

Albert De Roeck

Committee in Charge

2017

Copyright © 2017 by
Dustin Ray Burns
All rights reserved.

CONTENTS

LIST OF FIGURES

LIST OF TABLES

Chapter 1

Results

1.1 Statistical analysis

1.1.1 Systematic uncertainties

This section covers the systematic, including statistical, errors associated with the analysis. The general strategy is to duplicate the systematics applied in the SM H to 4l search, including errors on the background estimation from data, and add the errors associated with MET modelling, guided by the mono-H to $\gamma\gamma$ strategy.

1.1.1.1 Uncertainties on Reducible Background estimation

Sources of systematic uncertainties for the methods presented above potentially arises from the different composition of background processes (DY , $t\bar{t}$, WZ , $Z\gamma$) in the region where we measure and where we apply the fake ratios. OS method corrects for the resulting bias via the “3P1F component” of its prediction. SS method corrects explicitly the electron fake rates for the correct fraction of photon conversions, but no attempt is made to correct the muon fake rates. The closure tests presented here are used to assess a possible residual bias in the methods.

Statistics in 4l Control Sample The limited size of the samples in the control regions where we measure and where we apply the fake ratios is the source of the statistical uncertainties of the method. The dominating statistical uncertainty is driven by the number of events in the control region and is typically in the range of 1-10%.

Sensitivity of Fake Ratios to Background Composition Compositions of reducible background processes (DY , $t\bar{t}$, WZ , $Z\gamma^{(*)}$) in the region where we measure and where we apply the fake ratio method are typically not the same. This is the main source of the systematic uncertainty of the fake ratio method.

This uncertainty can be estimated by measuring the fake ratios for individual background processes in the $Z + 1L$ region in simulation. The weighted average of these individual fake ratios is the fake ratio that we measure in this sample (in simulation). The exact composition of the background processes in the 2P+2F region where we plan to apply the fake ratios can be determined from simulation, and one can reweigh the individual fake ratios according to the 2P+2F composition. The difference between the reweighed fake ratio and the average one can be used as a measure of the uncertainty on the measurement of the fake ratios. The fake ratios for individual processes, the average fake ratio and the reweighed fake ratio determined by simulation are shown in table below. The effect of this systematic uncertainty is propagated to the final estimates, and it amounts to about 32% for $4e$, 33% for $2e2\mu$ and 35% for 4μ final state.

Shape Uncertainty In order to estimate the uncertainty on the m_{4l} shape we have looked at the differences between the shapes of predicted background distributions for all three channels, and between both of the two methods. The envelope of differences between these shapes of distributions is used as an estimate of the shape uncertainty. The uncertainty is estimated to be roughly in the range of 5% - 15%. Since the difference of the shapes slowly varies with m_{4l} , it is taken as a constant versus m_{4l} and is absorbed in the much larger uncertainty on the predicted yield of backgrounds.

1.1.1.2 MET systematics

There are two types of systematic uncertainties related to the modeling of MET, those from the measurement of real MET, as from the signal samples or backgrounds with neutrinos, and those from fake MET, due to the mismeasurement of jets and other objects. The fake MET systematics apply to the Higgs signals with no associated W production and the non-resonant backgrounds.

The uncertainties from the modeling of real met are measured by varying several

corrections used to calculate MET, propogating these variations to the efficiency of MC samples to pass the MET selection [?]. The list of corrections used in this calculation are given below. Each correction is varied up and down by one standard deviation of the input distribution, with the systematic uncertainty taken as the maximim difference in efficiencies accross all correction variations.

- Jet energy
- Jet resolution
- Muon energy
- Electron energy
- Tau energy
- Photon energy
- Unclustered jet energy

The efficiencies for the V+H and a benchmark signal sample to pass the MET selection of $E_T^{\text{MISS}} > 60$ GeV after each correccion is varied up and down are given in Table ?? and Table ??, respectively.

Correction	Efficiency Up	Efficiency Down
Original PFMET		
Jet energy		
Jet resolution		
Muon energy		
Electron energy		
Tau energy		
Photon energy		
Unclustered jet energy		

Correction	Efficiency Up	Efficiency Down
Original PFMET		
Jet energy		
Jet resolution		
Muon energy		
Electron energy		
Tau energy		
Photon energy		
Unclustered jet energy		

The second systematic uncertainty results from the modeling of fake MET, primarily due to the mismeasurement of jets (see Section ??). This systematic is measured in the the sideband CR as the percent difference between the efficiency for the data and total background sample to pass the MET selection. The efficiency for data is XX, MC is XX, giving a difference of around 40% which is taken as the systematic.

1.1.1.3 Additional systematics

The main experimental uncertainties which affect both signal and background are the uncertainty on the integrated luminosity (2.6%) and the uncertainty on the lepton identification and reconstruction efficiency (ranging from 2.5–9% on the overall event yield for the 4μ and $4e$ channels, respectively). Experimental uncertainties for the reducible background estimation, described in Section ??, vary between 36% (4μ) and 43% ($4e$). The uncertainty on the lepton energy scale is determined by considering the $Z \rightarrow \ell\ell$ mass distributions in data and simulation. Events are separated into categories based on the p_T and η of one of the two leptons, determined randomly, and integrating over the other. The dilepton mass distributions are then fit to a Breit-Wigner parameterization convolved with a double-sided Crystal Ball function. The offset in the measured peak position with respect to the nominal Z boson mass in data and simulation are extracted, and the results are shown in Fig. 1.1. The relative difference between data and simulation is propagated

to the reconstructed four-lepton mass from simulated Higgs boson events. The results of the propagation can be seen in Fig. 1.2. In the case of electrons, since the same data set is used to derive and validate the momentum scale corrections, the size of the corrections are taken into account for the final value of the uncertainty. The uncertainty is determined to be 0.04% (0.3%) for the 4μ ($4e$) channels, respectively. The uncertainty on the 4ℓ mass resolution coming from the uncertainty on the per-lepton energy resolution is 20%, as described in Section ??.

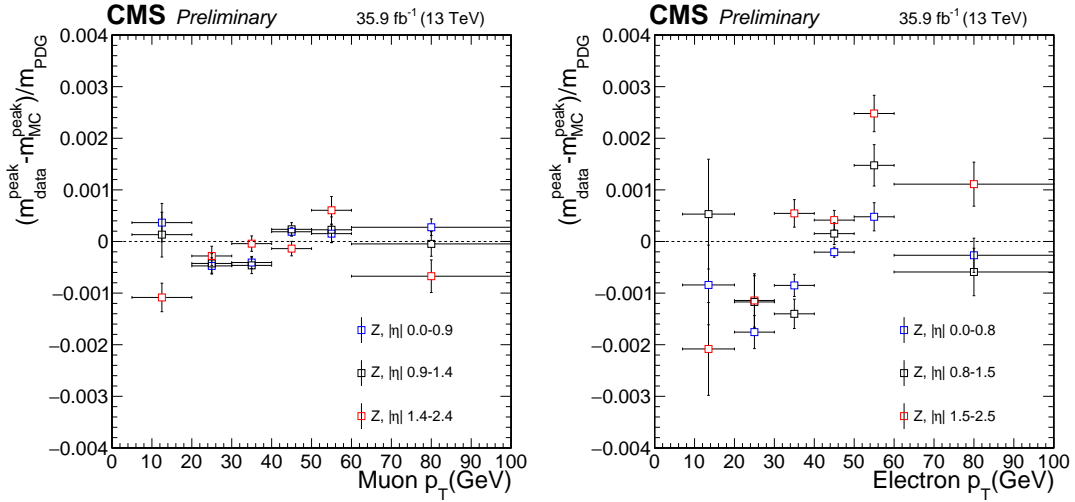


Figure 1.1: Difference between the $Z \rightarrow \ell\ell$ mass peak positions in data and simulation normalized by the nominal Z boson mass obtained as a function of the p_T and $|\eta|$ of one of the leptons regardless of the second for muons (left) and electrons (right).

Theoretical uncertainties which affect both the background signal and background estimation include uncertainties from the renormalization and factorization scale and choice of PDF set. The uncertainty from the renormalization and factorization scale is determined by varying these scales between 0.5 and 2 times their nominal value while keeping their ratio between 0.5 and 2. The uncertainty from the PDF set is determined by taking the root mean square of the variation when using different replicas of the default NNPDF set. An additional uncertainty of the 10% on the K factor used for the $gg \rightarrow ZZ$ prediction is applied as described in Section ?. A systematic uncertainty of 2% on the branching ratio of $H \rightarrow ZZ \rightarrow 4\ell$ only affects the signal yield. In the case of event categorization, experimental and theoretical uncertainties which account

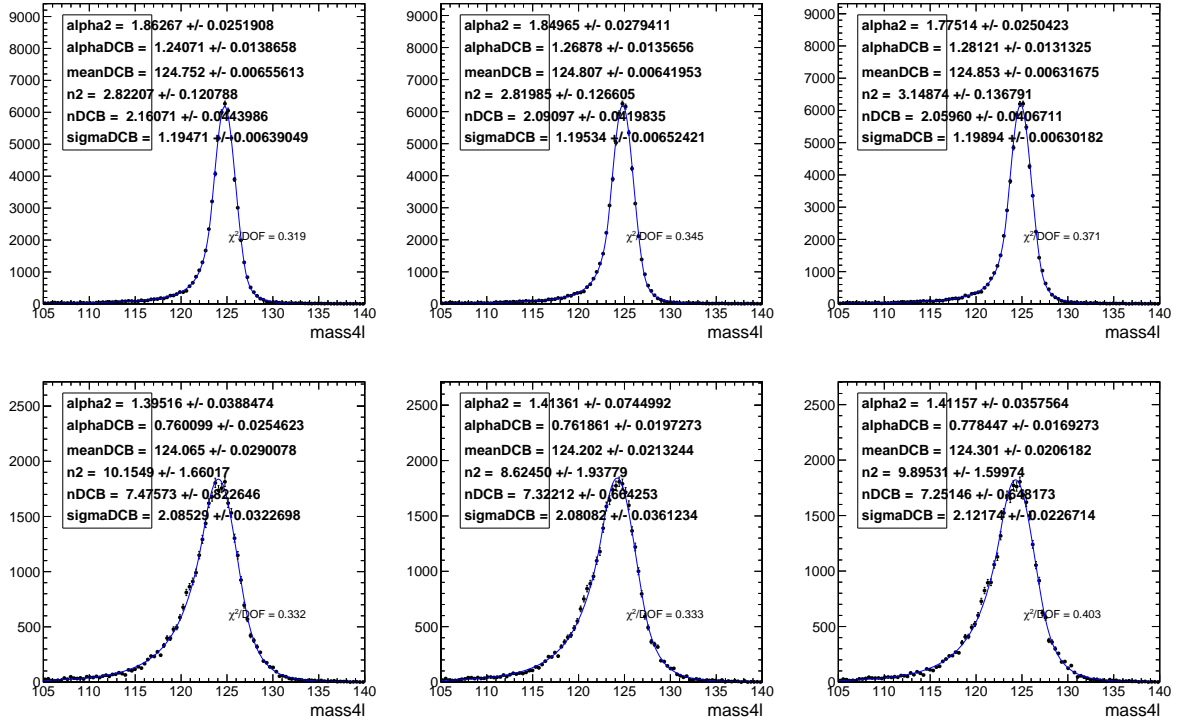


Figure 1.2: Different $m_{4\ell}$ distributions after propagating the biases in Fig. 1.1 to Higgs boson events. The change in the mean of the double crystal ball is used to determine the systematic uncertainty due to the lepton momentum scale. The middle plot shows the nominal distribution, while the left (right) plots show the down (up) systematic variations. The 4μ channel is shown in the top row and the $4e$ channel is shown in the bottom row.

for possible migration of signal and background events between categories are included. The main sources of uncertainty on the event categorization include the QCD scale, PDF set, and the modeling of hadronization and the underlying event. These uncertainties amount to between 4–20% for the signal and 3–20% for the background depending on the category. The lower range corresponds to the VBF and VH processes and the upper range corresponds to the $gg \rightarrow H$ process yield in the VBF-2jet-tagged category. Additional uncertainties come from the imprecise knowledge of the jet energy scale (from 2% for the $gg \rightarrow H$ yield in the untagged category to 15% for $gg \rightarrow H$ yield in the VBF-2jet-tagged category) and b-tagging efficiency and mistag rate (up to 6% in the -tagged category).

Table 1.3: Summary of the experimental systematic uncertainties in the $H \rightarrow 4\ell$ measurements.

Summary of relative systematic uncertainties	
Common experimental uncertainties	
Luminosity	2.6 %
Lepton identification/reconstruction efficiencies	2.5 – 9 %
Background related uncertainties	
Reducible background (Z+X)	36 – 43 %
Signal related uncertainties	
Lepton energy scale	0.04 – 0.3 %
Lepton energy resolution	20 %

1.1.2 Limit setting

The primary tool used to interpret the analysis described above in the context of the signal models is the HiggsAnalysis-CombinedLimit package [?], a collection of RooStats-based software [?] used within the Higgs PAG [?], hereafter called the combine tool, or simply combine. Specifically, limits (one-sided Bayesian credible intervals) are set on the expected and observed signal production cross section times branching ratio BR of H to four leptons for the various benchmarks using the asymptotic CL_s method [?], an approach to calculating a profile likelihood ratio using an approximation of the LHC test-statistic distributions. The upper limit on the cross section gives the maximum number of events that can be attributed to the signal process, consistent with the data that is observed. Combine finds the upper limit as the numerical solution to Equation 1.1, which sets the integral of the posterior probability $p(\sigma|D)d\sigma$ equal to the desired confidence level for the measurement, typically 95%. The posterior probability gives the degree of belief that σ lies in the interval $[\sigma, \sigma + d\sigma]$, and is formed by inverting a multi-Poisson model, $p(D|\sigma, \boldsymbol{\theta})$ using Bayes' theorem (Equation 1.2), after numerically marginalizing priors describing

Table 1.4: Summary of the theory systematic uncertainties in the $H \rightarrow 4\ell$ measurements for the inclusive analysis

Summary of inclusive theory uncertainties	
QCD scale (gg)	$\pm 3.9 \%$
PDF set (gg)	$\pm 3.2 \%$
Bkg K factor (gg)	$\pm 10 \%$
QCD scale (VBF)	$+0.4/-0.3 \%$
PDF set (VBF)	$\pm 2.1 \%$
QCD scale (WH)	$+0.5/-0.7 \%$
PDF set (WH)	$\pm 1.9 \%$
QCD scale (ZH)	$+3.8/-3.1 \%$
PDF set (ZH)	$\pm 1.6 \%$
QCD scale ($t\bar{t}H$)	$+5.8/-9.2 \%$
PDF set ($t\bar{t}H$)	$\pm 3.6 \%$
BR($H \rightarrow ZZ \rightarrow 4\ell$)	2%
QCD scale ($q\bar{q} \rightarrow ZZ$)	$+3.2/-4.2 \%$
PDF set ($q\bar{q} \rightarrow ZZ$)	$+3.1/-3.4 \%$
Electroweak corrections ($q\bar{q} \rightarrow ZZ$)	$\pm 0.1 \%$

the uncertainties $\pi(\boldsymbol{\theta})$ Equation 1.3.

$$\int_0^{\sigma_{\text{Up}}} p(\sigma|D)d\sigma = 1 - \alpha \quad (1.1)$$

$$p(\sigma|D) = p(D|\sigma)\pi(\sigma)/p(D) \quad (1.2)$$

$$p(D|\sigma) = \int p(D|\sigma, \boldsymbol{\theta})\pi(\boldsymbol{\theta})d\boldsymbol{\theta} \quad (1.3)$$

Combine takes as input specially formatted text files called data cards, which contain the signal, background, and data yields, along with associated systematic uncertainties, and point to the location of the file containing the shape distributions for each sample. When the shape distributions are included, combine computes the limits accross all bins, then combines the results. So, a non-shape based limit is equivalent to a shape-based limit with one bin. This can be used as a cross check that the shape-based limit is behaving correctly, and to analyze the improvement of using the shape-based approach over non-shape based. The event yields and systematic uncertainties are discussed later in this section.

Combine returns the 95% confidence level (CL) expected and observed limits on the signal strength parameter, μ , as well as one and two sigma deviations of the expected limit, where μ is the signal scale factor in the i^{th} bin in the mean count, $n_i = \mu * s_i + b_i$, where s_i and b_i are the signal and background yields, respectively.

Limits are set for the three decay channels, $4e$, 4μ , and $2e2\mu$ individually, then combined for the $4l$ limit. The signal strength parameter gives the ratio of the 95% CL expected signal yield to the theoretical yield, or more generally,

$$\mu[\sigma_{\text{norm}} * BR] = \frac{\sigma_{95\%CL} * BR}{\sigma_{\text{norm}} * BR}, \quad (1.4)$$

where σ_{norm} is the cross section used to normalize the signal yields given as input to combine, typically the theoretical production cross section, but also set to 1 pb for comparison to other H decay channels. This formula is used to calculate the 95% CL cross section times branching ratio, which is the desired output variable. This limit can then be compared to the theoretical cross section times BR. The points where the upper limit is lower than the theory value are interpreted as exclusions.

1.2 Results

After the event selection is optimized and all of the uncertainties are accounted for, limits are set on the signal strength parameter, μ , and presented in two ways. For each model, the limits are shown for μ , to be easily compared with the limits obtained by the other H decay channels, and as cross section limit versus mass plots. These results are found for both the cut-and-count and MVA based strategies. Until the selection is frozen and the data unblinded, only expected limits will be plotted in the SR.

1.2.1 Cut-and-count based

1.2.1.1 Zp2HDM

The 1D slice of mass points fixing $m_{A0} = 300$ GeV is selected for the limits plots since this region has the highest cross section and the largest branching fraction to DM production. The 1D limits are presented for μ in Figure ?? and $\sigma_{95\%CL} * BR$ in Figure ??, obtained using the optimized selection given in Section ??.

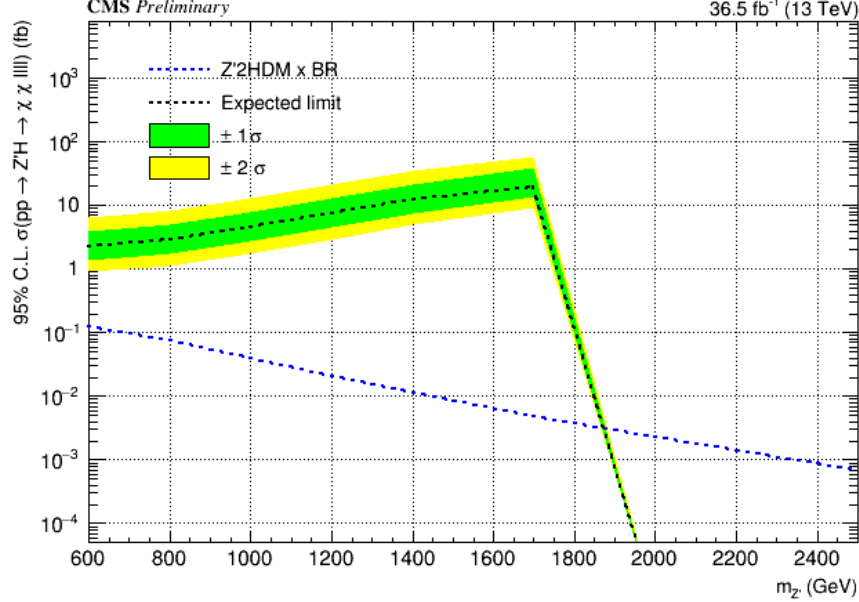


Figure 1.3: PLACEHOLDER 1D μ limits for $m_{A0} = 300$ GeV for Zp2HDM.

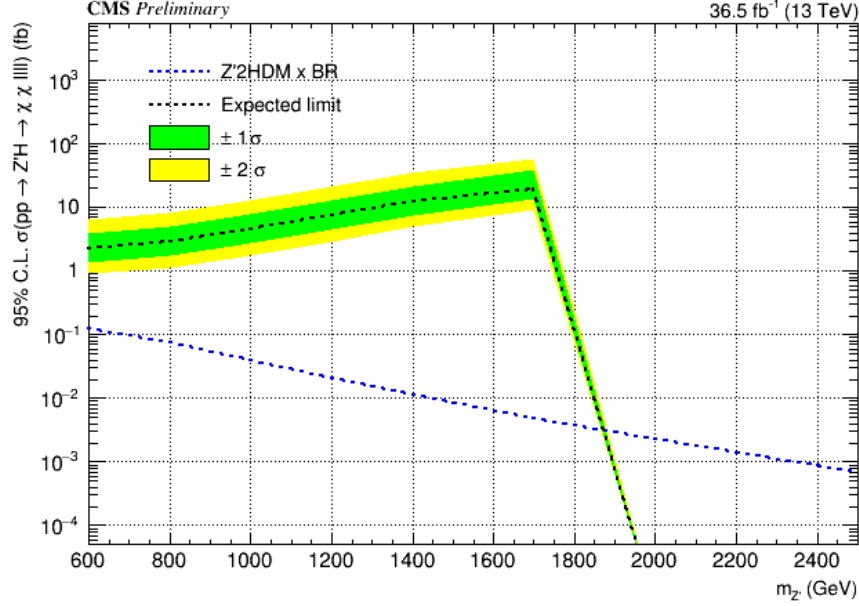


Figure 1.4: PLACEHOLDER 1D cross section times BR limits for $m_{A0} = 300$ GeV for Zp2HDM.

1.2.1.2 ZpBaryonic

The 1D slice of mass points fixing $m_\chi = 1$ GeV is selected for the limits plots since this region has the highest cross section. The 1D limits are presented for μ in Figure ?? and $\sigma_{95\%CL} * BR$ in Figure ??, obtained using the optimized selection given in Section ??.

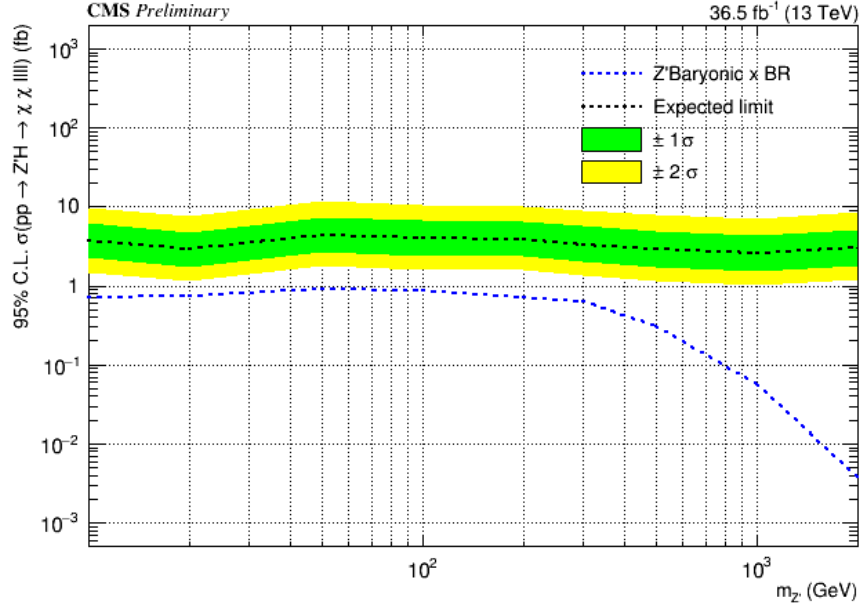
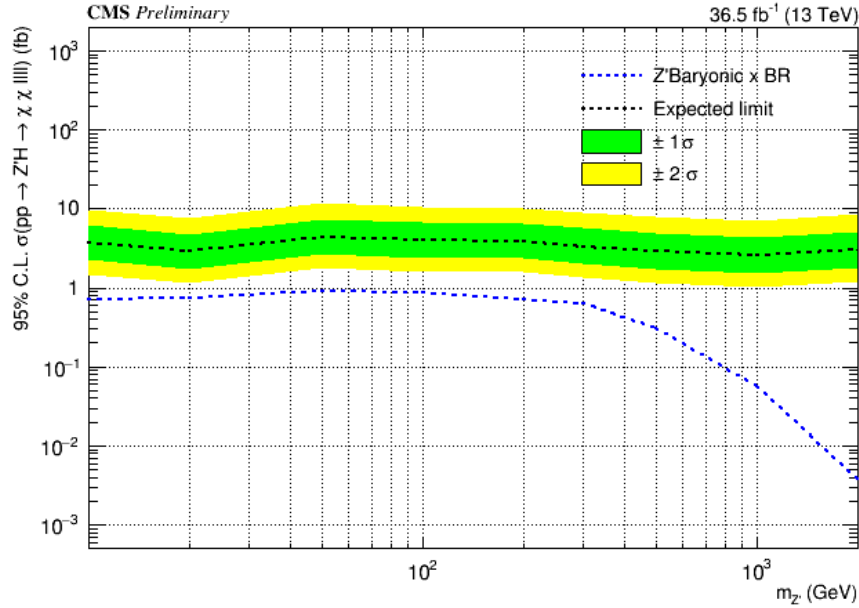
1.2.2 MVA based

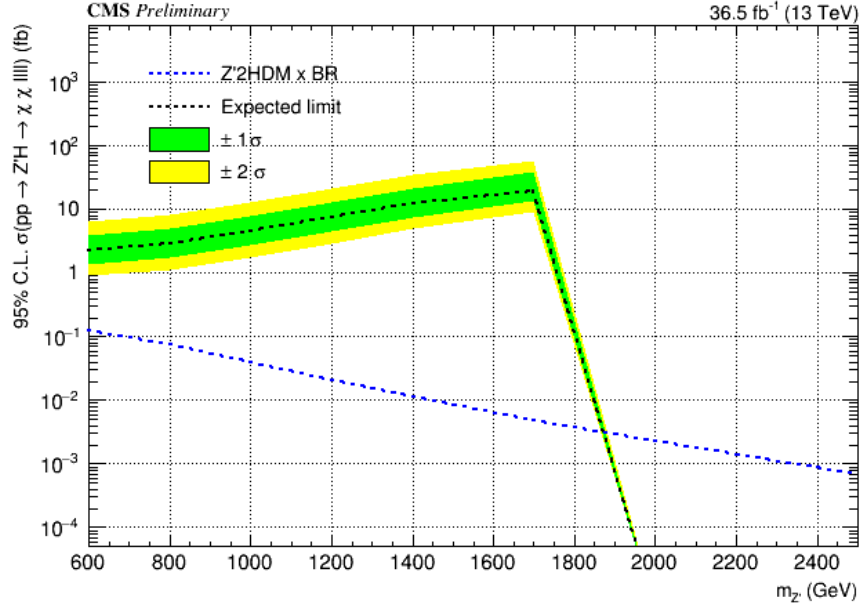
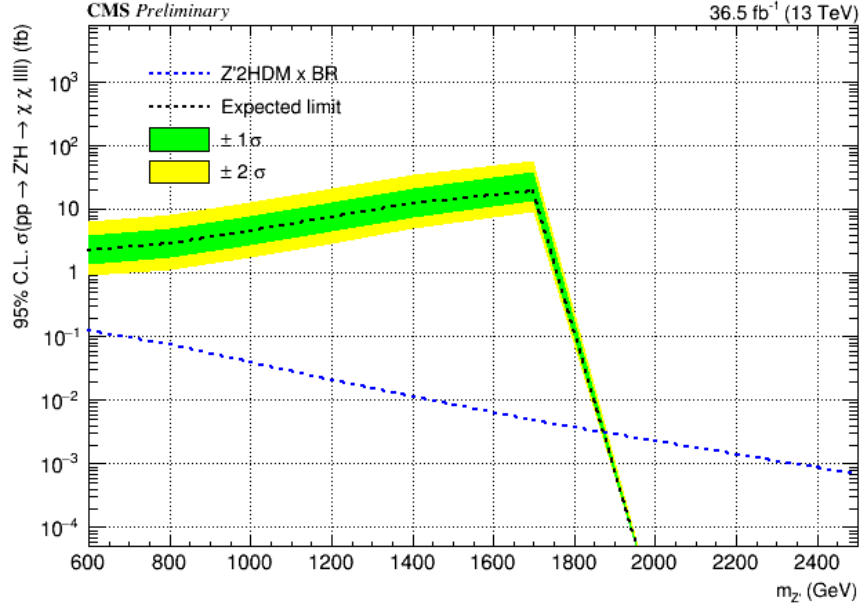
1.2.2.1 Zp2HDM

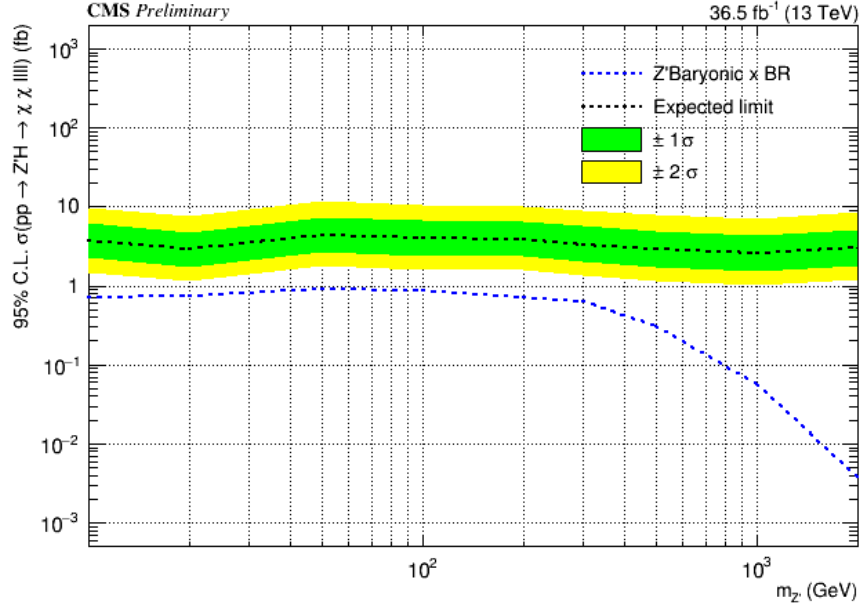
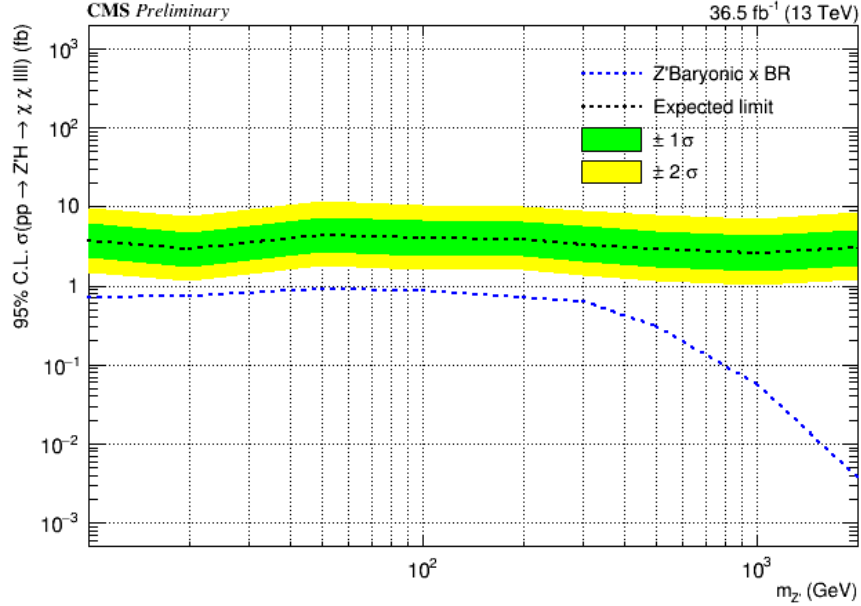
The 1D limits are presented for μ in Figure ?? and $\sigma_{95\%CL} * BR$ in Figure ??, obtained using the optimized selection given in Section ??.

1.2.2.2 ZpBaryonic

The 1D limits are presented for μ in Figure ?? and $\sigma_{95\%CL} * BR$ in Figure ??, obtained using the optimized selection given in Section ??.

Figure 1.5: PLACEHOLDER 1D μ limits for $m_\chi = 1$ GeV for ZpBaryonic.Figure 1.6: PLACEHOLDER 1D cross section times BR limits for $m_\chi = 1$ GeV for ZpBaryonic.

Figure 1.7: PLACEHOLDER 1D μ limits for $m_{A0} = 300$ GeV for Zp2HDM.Figure 1.8: PLACEHOLDER 1D cross section times BR limits for $m_{A0} = 300$ GeV for Zp2HDM.

Figure 1.9: PLACEHOLDER 1D μ limits for $m_\chi = 1$ GeV for ZpBaryonic.Figure 1.10: PLACEHOLDER 1D cross section times BR limits for $m_\chi = 1$ GeV for ZpBaryonic.

Appendices

Appendix A

Production Cross Sections for Benchmark Signal Models

m_χ [GeV]	1	10	50	65	100	200	400	800	1000	1300
EFT_HHxx_scalar	0.10071E+01	0.99793E+00	0.60671E+00	0.48291E-04	0.22725E-05	0.11059E-06	0.36569E-08	0.40762E-10	0.64956E-11	0.51740E-12
EFT_HHxx_combined	0.15731E+01	0.15194E+01	0.34134E+00	0.41039E-04	0.10581E-04	0.16553E-05	0.14628E-06	0.40608E-08	0.85950E-09	0.96480E-10
EFT_HHxg5x	0.15735E+03	0.15594E+03	0.94804E+02	0.12990E-01	0.23075E-02	0.41820E-03	0.45743E-04	0.16734E-05	0.39327E-06	0.49769E-07
EFT_xdxHdHc	0.29530E+00	0.29067E+00	0.10540E+00	0.89849E-01	0.64959E-01	0.30639E-01	0.88644E-02	0.97986E-03	0.33847E-03	0.68674E-04
$\Lambda = 100$ GeV	0.16306E-04	0.15508E-04	0.12088E-05	0.88288E-06	0.53312E-06	0.18046E-06	0.34918E-07	0.27514E-08	0.90662E-09	0.19313E-09
$\Lambda = 1000$ GeV	0.57027E+00	0.57001E+00	0.56025E+00	0.55337E+00	0.53270E+00	0.45792E+00	0.29777E+00	0.10288E+00	0.57444E-01	0.23260E-01
EFT_xgxFHDH										

Table A.1: EFT production cross sections [pb]

m_χ [GeV]	m_S [GeV]							
1	0.21915E+01	0.20798E+01	0.19192E+01	0.18118E+01	0.16735E+01	0.52244E+01	0.41877E+01	0.28732E+01
10	0.17416E+01	0.17420E+01	0.18581E+01	0.17510E+01				0.17398E+01
50	0.39053E+00		0.38877E+00	0.38409E+00	0.37097E+00	0.12861E+01		0.39096E+00
150	0.24136E-05				0.38372E-05	0.21922E-04	0.42337E-03	0.57124E-04
500	0.34099E-08						0.49399E-08	0.25206E-06
1000	0.17012E-10							0.55260E-10
								0.11067E-07

Table A.2: Scalar production cross sections [pb] corresponding to mass points in Table ??

m_χ [GeV]	$m_{Z'}$ [GeV]									
1	0.26615E+01	0.27802E+01	0.33248E+01	0.32341E+01	0.26566E+01	0.23191E+01	0.10842E+01	0.18700E+00	0.11728E-01	0.17399E-07
10	0.21182E-01	0.74027E-01	0.32732E+01	0.32250E+01						0.17380E-07
50	0.36342E-03		0.12726E-01	0.31337E+00	0.21226E+01	0.20120E+01				0.17340E-07
150	0.55972E-05				0.56526E-02	0.18000E+00	0.67266E+00	0.18111E+00		0.16918E-07
500	0.80295E-08						0.36591E-04	0.10368E-01	0.10375E-01	0.13179E-07
1000	0.49387E-10							0.98079E-06	0.57596E-03	0.80146E-08

Table A.3: ZpBaryonic production cross sections [pb] corresponding to mass points in Table ??

m_χ [GeV]	$m_{Z'}$ [GeV]									
1	0.61935E-02	0.63192E-02	0.82991E-02	0.11942E-01	0.19171E-01	0.21560E-01	0.16010E-01	0.64416E-02	0.56526E-02	0.58902E-02
10	0.58781E-02	0.58938E-02	0.82944E-02	0.11937E-01						0.58805E-02
50	0.10294E-03		0.77820E-04	0.15258E-04	0.12066E-01	0.12105E-01				0.10387E-03
150	0.28382E-06				0.83917E-05	0.72889E-03	0.65401E-02	0.68337E-03		0.29033E-06
500	0.34689E-09						0.43355E-06	0.87799E-04	0.28292E-05	0.36327E-09
1000	0.20703E-11							0.44782E-07	0.19974E-06	0.27001E-11

Table A.4: ZpHS production cross sections [pb] corresponding to mass points in Table ??

m_{A^0} [GeV]	$m_{Z'}$ [GeV]							
300	42.386	45.097	35.444	26.07	18.942	11.778	7.4456	3.6446
400	5.8513	14.847	14.534	11.792	9.029	5.851	3.7819	1.8758
500		5.9605	8.4961	7.9575	6.5515	4.5063	3.0028	1.5235
600		1.5853	4.6972	5.4808	4.9946	3.7044	2.5694	1.3447
700			2.1092	3.4848	3.6766	3.0253	2.2023	1.1984
800			0.65378	1.9638	2.5511	2.4077	1.8689	1.0692

Table A.5: Zp2HDM production cross sections [pb] corresponding to mass points in Table ??

REFERENCES

- [1] Massimiliano Grazzini, Stefan Kallweit, and Dirk Rathlev. ZZ production at the LHC: Fiducial cross sections and distributions in NNLO QCD. *Phys. Lett. B*, 750:407 – 410, 2015. URL: <http://www.sciencedirect.com/science/article/pii/S0370269315007303>, arXiv:1507.06257, doi:<http://dx.doi.org/10.1016/j.physletb.2015.09.055>.
- [2] John M. Campbell and R. K. Ellis. MCFM for the Tevatron and the LHC. *Nucl. Phys. Proc. Suppl.*, 205:10, 2010. arXiv:1007.3492, doi:10.1016/j.nuclphysbps.2010.08.011.
- [3] John M. Campbell, R. Keith Ellis, and Ciaran Williams. Vector boson pair production at the LHC. *JHEP*, 07:018, 2011. arXiv:1105.0020, doi:10.1007/JHEP07(2011)018.
- [4] John M. Campbell, R. Keith Ellis, and Ciaran Williams. Bounding the Higgs width at the LHC using full analytic results for $gg \rightarrow e^-e^+\mu^-\mu^+$. *JHEP*, 04:060, 2014. arXiv:1311.3589, doi:10.1007/JHEP04(2014)060.
- [5] Marco Bonvini, Fabrizio Caola, Stefano Forte, Kirill Melnikov, and Giovanni Ridolfi. Signal-background interference effects in $gg \rightarrow H \rightarrow WW$ beyond leading order. *Phys. Rev. D*, 88:034032, 2013. arXiv:1304.3053, doi:10.1103/PhysRevD.88.034032.
- [6] Kirill Melnikov and Matthew Dowling. Production of two Z-bosons in gluon fusion in the heavy top quark approximation. *Phys. Lett. B*, 744:43, 2015. arXiv:1503.01274, doi:10.1016/j.physletb.2015.03.030.
- [7] Chong Sheng Li, Hai Tao Li, Ding Yu Shao, and Jian Wang. Soft gluon resummation in the signal-background interference process of $gg(\rightarrow h^*) \rightarrow ZZ$. 2015. arXiv:1504.02388.
- [8] Giampiero Passarino. Higgs CAT. *Eur. Phys. J. C*, 74:2866, 2014. arXiv:1312.2397, doi:10.1140/epjc/s10052-014-2866-7.
- [9] Stefano Catani and Massimiliano Grazzini. An NNLO subtraction formalism in hadron collisions and its application to Higgs boson production at the LHC. *Phys. Rev. Lett.*, 98:222002, 2007. arXiv:hep-ph/0703012, doi:10.1103/PhysRevLett.98.222002.
- [10] Massimiliano Grazzini. NNLO predictions for the Higgs boson signal in the $H \rightarrow WW \rightarrow l\nu l\nu$ and $H \rightarrow ZZ \rightarrow 4l$ decay channels. *JHEP*, 02:043, 2008. arXiv:0801.3232, doi:10.1088/1126-6708/2008/02/043.
- [11] Massimiliano Grazzini and Hayk Sargsyan. Heavy-quark mass effects in Higgs boson production at the LHC. *JHEP*, 09:129, 2013. arXiv:1306.4581, doi:10.1007/JHEP09(2013)129.

- [12] LHC Higgs Cross Section Working Group. Handbook of LHC Higgs Cross Sections: 3. Higgs Properties. CERN Report CERN-2013-004, 2013. [arXiv:1307.1347](#), [doi:10.5170/CERN-2013-004](#).