



Searching for new Higgs particles in resonant top quark pair production

Dominic Stafford, Durham University

September 10, 2018

Abstract

Searches for new scalar and pseudoscalar heavy Higgs bosons are considered. The interference $M_{t\bar{t}}$ distribution is simulated for a benchmark scenario using an EFT approach in MG5_aMC@NLO using reweighting, which is found to give a significant improvement to the lineshape resolution compared with previous approaches which did not use reweighting. Using W and top tagging to improve limits on the couplings of heavy Higgs candidates at different masses was then considered, and it was found that these approaches notably improved the limits which could be set on both 400GeV and 750GeV candidates.

Contents

1	Introduction	3
2	Simulating the signal lineshape	3
2.1	Theoretical Background	3
2.2	Technical setup	5
2.3	Results	5
3	Data Analysis	6
3.1	Search method	6
3.2	Event reconstruction without HOT	7
3.3	Event reconstruction with HOT	8
3.4	Search variables	10
3.5	Results	11
4	Conclusion	12
5	Acknowledgements	12
6	Appendix: Details of simulation of interference lineshape	16

1 Introduction

Many current Beyond the Standard Model (BSM) Theories include an extended Higgs Sector. The simplest of these theories- for instance the two Higgs Doublet Model (2HDM) [1], the Minimally Supersymmetric Standard Model [2], and composite Higgs Theories [3]- include additional scalar (H) or pseudoscalar (A) neutral Higgs Bosons, which may be detectable at the LHC. These particles are not expected to couple strongly to the Weak gauge Bosons due to various theoretical constraints, but are expected to couple strongly to the top quark due to its very high mass of 173GeV. The Heavy Higgs Bosons are therefore expected to decay predominantly to two top quarks (Assuming $m_{H/A} > 2m_t$), and so this is considered a very profitable channel in which to search for these new particles.

A likely mode of production for heavy Higgs bosons is via a top loop coupling quarks to gluons (Figure 1). This loop introduces a non-trivial phase, which results in interference between the signal and the Standard Model (SM) background, which can be larger than the resonant component of the signal. As a result, the signal may not give a peak in the $M_{t\bar{t}}$ as one would expect, but instead a peak-dip, dip-peak or pure dip structure [4, 5], depending on the precise parameters of the model.

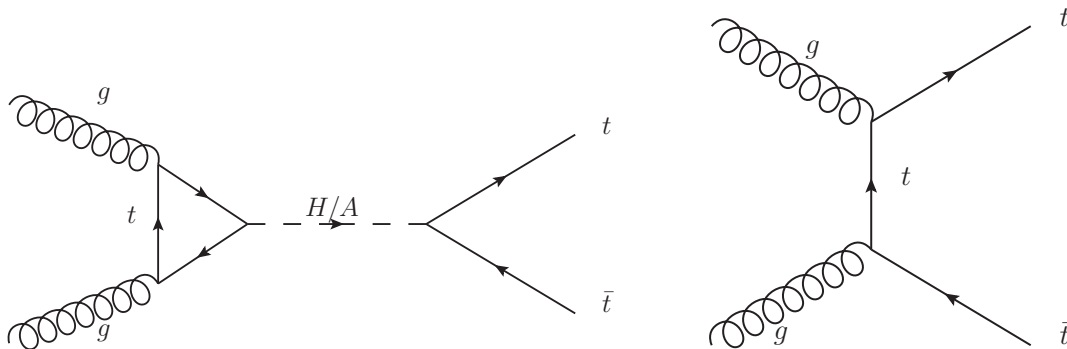


Figure 1: Example LO graphs of signal (left) and background (right) $t\bar{t}$ production

This report is divided into two sections: the first looks at producing accurate predictions of the lineshape of the interference component of the signal using an effective field theory approach, whilst the second looks at using heavy object tagging (HOT) to improve detection of these signals.

2 Simulating the signal lineshape

2.1 Theoretical Background

As has already been mentioned, accurate predictions of the signal lineshape are believed to be very important in searches for the new resonances in the $t\bar{t}$ decays. In order to achieve this accuracy, it is desirable to simulate the process at next-to leading-order

(NLO) accuracy. However, this is difficult since the LO production diagrams already include a topquark loop, which results in a very large number of NLO diagrams (see figure 2.1). A solution is to use effective field theory (EFT) to model this loop as a contact interaction, as described in reference [6]. This EFT approach is only normally valid in the case where the particle in the loop is too heavy relative to the other particles to be resolved, i.e. $2m_t > m_{H/A}$, however since a decay to two tops is being investigated, one has the opposite assumption: $m_{H/A} > 2m_t$. The solution proposed in reference [6] is to multiply the coupling constants of the Higgs to the gluons C_{HG} and C_{AG} by a complex factor $a + bi$, which mimics the phase factor introduced by a resolved top loop. Unfortunately this factor was not included here due to time constraints, but could easily be added in to further analyses.

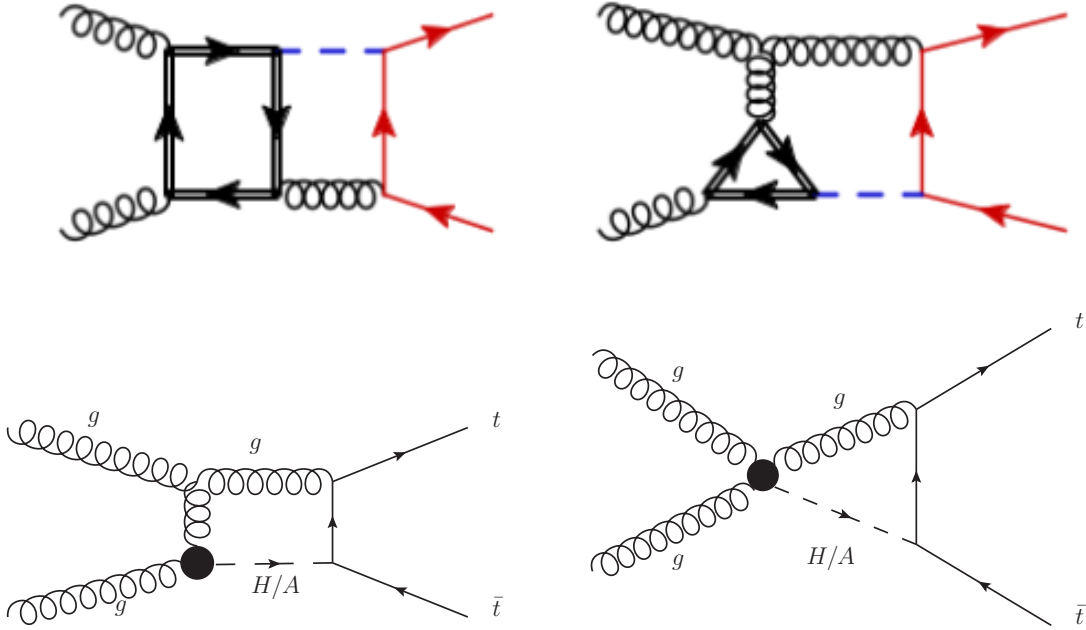


Figure 2: Top: Example two-loop NLO diagrams for the signal process (source: [6])
Bottom: The corresponding diagrams with the top loop replaced by an eft vertex

The total amplitude for the signal plus background process is:

$$|\mathcal{M}_{SM} + \mathcal{M}_{BSM}|^2 = |\mathcal{M}_{SM}|^2 + |\mathcal{M}_{BSM}|^2 + 2Re(\mathcal{M}_{SM}\mathcal{M}_{BSM}^*), \quad (1)$$

where the first term on the right-hand side of the equation is the Standard model background, the second term is the resonant component of the BSM signal, and the third term is the interference. The first and second of these terms can be easily calculated in isolation, but the interference, which is the component of greatest interest because it is the component most affected by the accuracy of the simulation and can change the shape of the signal, can only be found as a part of the overall amplitude. It could be calculated by subtracting the pure resonant and background components from the total

amplitude, however this tends to lead to poor statistics, so the preferred method is to also calculate the the total amplitude when the signal has been multiplied by -1 , and then calculate the interference as:

$$2\text{Re}(\mathcal{M}_{SM}\mathcal{M}_{BSM}^*) = \frac{1}{2}(|\mathcal{M}_{SM} + \mathcal{M}_{BSM}|^2 - |\mathcal{M}_{SM} - \mathcal{M}_{BSM}|^2). \quad (2)$$

2.2 Technical setup

The benchmark used here was selected for consistency with that used by F. Hagemann in [10], which was originally based on benchmark C1 in [6]. The values of the parameters are summarised in table 2.2. This benchmark has $m_H < 2m_t < m_A$, and so only the pseudoscalar heavy Higgs can decay to a top quark pair, and so one would expect a single peak-dip structure in the interference lineshape centered on m_A . Unlike the scenarios discussed previously, this benchmark allows for an additional BSM fermion in the loop connecting the gluons to the heavy Higgs and so there are additional couplings y_F and \tilde{y}_F . The previous analyses assumed a non zero complex phase multiplying the signal to simulate the phase coming from the eft vertex, however here this complex number was taken to be $c = a + ib = 1$ for simplicity.

$y_t = 0.05$	$\tilde{y}_t = 0.5$	$y_F = -0.0354$	$\tilde{y}_F = 0.354$
$C_{tG} = 9.56 \times 10^{-6}$	$C_{HG} = 4.73 \times 10^{-3}$	$C_{AG} = -6.49 \times 10^{-3}$	$\Lambda = 1.0 \times 10^3$
$m_{u^l} = 172.5$	$m_{u_{yukawa}} = 172.5$	$m_H = 300$	$m_A = 450$

Table 1: The parameters used in the simulation. All values are in GeV

The events were simulated at NLO using MadGraph5_aMCNLO [7] using the AHTtbar model [8] and showered using Pythia8 [9], using the NNPDF30_nnlo_as_0118_hessian pdf set. The analysis in [10] followed this method but encountered problems since the statistical uncertainties from the two large values on the right-hand side of equation 2 tended to dominate the signal. The solution used here was to use the event-by-event NLO reweighting capabilities of MadGraph5_aMC@NLO [11] to calculate the interference for each event. The results were then analysed using MadAnalysis5 [12], and a graph of the interference was produced. Full details of the method for a reader wishing to reproduce the results are included in the appendix.

2.3 Results

A graph is shown in figure 2.3 of the interference component of the signal, calculated as described above, alongside a previous result from [10], which did not use reweighting, to give a comparison. Both graphs show the expected peak-dip structure about $m_A = 450\text{GeV}$, however it is clear that reweighting dramatically improves the resolution of the interference lineshape.

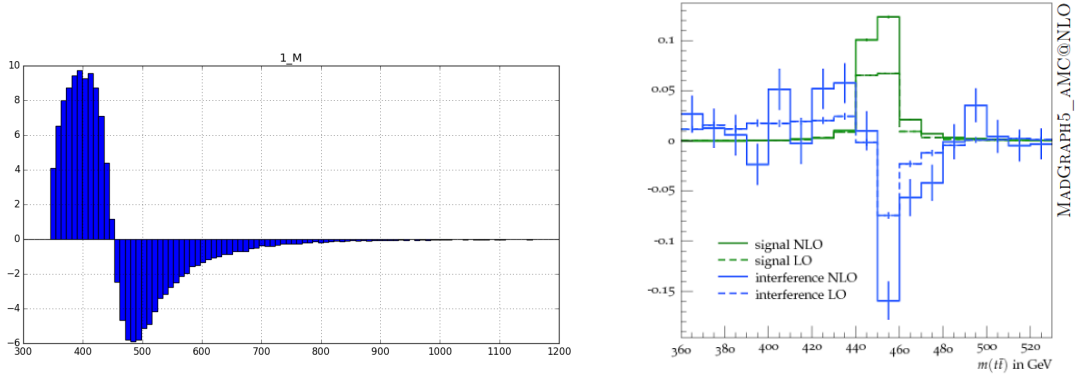


Figure 3: Left: The graph of the interference lineshape. Right: Graphs of the resonant and interference components of the signal produced without reweighting (Source: [10])

3 Data Analysis

3.1 Search method

There are ongoing searches to attempt to identify the signals discussed above in $t\bar{t}$ production at the LHC [13, 14]. The difficulty these searches face is that the signal is likely to be small compared to the large standard model $t\bar{t}$ background, especially for higher heavy Higgs masses. One potential way to improve the sensitivity of these searches is to make use of heavy object tagging (HOT) to directly tag merged jets coming from the top decays, rather than attempting to identify the top quarks from a number of decay products. Here we focussed on replicating the expected limits from simulation data found in the semileptonic channel by [14] using a very similar methodology, and then investigated whether improvements could be made using HOT.

The semileptonic decay channel of the $t\bar{t}$ system is shown in figure 3.1- both the top and the antitop quarks decay quickly to a bottom (anti)quark and a $W^{+(-)}$ boson. One of the W bosons then decays to a lepton-neutrino pair, the other to a quark-antiquark pair. If the lepton is an electron or a muon it should be registered in the detector, and one should also be able to detect tau leptons which decay leptonically (some momentum will go into the neutrinos produced in this decay, but these momenta will simply be added to that of the neutrino produced in the W decay in the missing energy). Hadronically decaying tau leptons are not considered here as they do not give a sufficiently distinctive signal to differentiate $t\bar{t}$ events from the QCD background. The two b quarks and the two light quarks from the W decay form jets and can be identified using a jet clustering algorithm. The neutrino is not observed, but its transverse momentum can be inferred to be the negative of the sum of the transverse momenta of all the detected particles coming from the interaction (the “missing” transverse momentum).

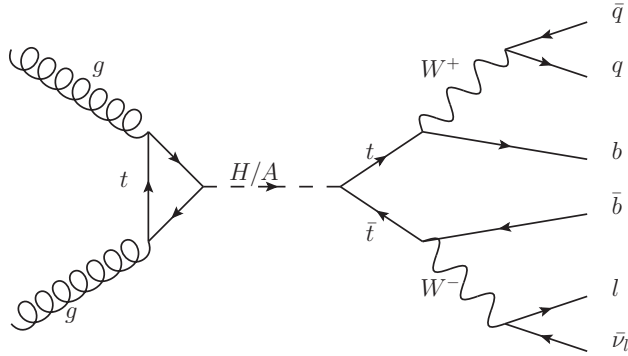


Figure 4: Full Feynmann diagram for a semi-leptonically decaying signal event

3.2 Event reconstruction without HOT

The simulated data was in the form of ROOT data files containing particle-flow candidates. These candidates were then required to pass certain quality requirements, which were an attempt to replicate those in [14] to give comparibility of results (the quality requirements in that note were generally motivated by the specifications of the CMS detector [15]):

Leptons

For use in later event reconstruction, all leptons were tagged in two categories "loose" and "tight". For a muon to be tagged as loose it was required to have a transverse momentum $p_t > 10\text{GeV}$, a pseudorapidity $|\eta| < 2.5$ and a particle flow isolation $\mathcal{I} < 0.25$. To be tagged as tight, a muon was required to have $p_t > 26\text{GeV}$, $|\eta| < 2.4$ and $\mathcal{I} < 0.15$. A loose electron was required to have $p_t > 20\text{GeV}$ and $|\eta| < 2.5$, whilst a tight electron was required to have $p_t > 30\text{GeV}$ and $|\eta| < 2.5$, with the region $1.4442 < |\eta| < 1.5660$, which corresponds the transition region between the barrel and the end-caps of the electronic calorimeter, excluded.

Jets

All jets were required to have $p_t > 20\text{GeV}$ and $|\eta| < 2.4$. Jets were then b-tagged using the medium working point of the cMVA_{v2} algorithm [16], which uses a boosted decision tree taking the outputs of several individual algorithms as inputs. The CSV algorithm, which uses a neural network based on track and vertex information, was also trialed, but this was found to give slightly weaker limits than the cMVA_{v2} algorithm. Jets which were not b-tagged were identified as light jets.

Event Selection

In order to be considered in the analysis, each event was required to contain at least one tight muon, two b-tagged jets and two light jets which pass the above quality requirements. Since the light jets from a W decay are often charm jets, which are relatively likely to be mistagged as b jets, it might be profitable to also allow additional b jets to be considered as light jets, however this was not done here for simplicity. In order to suppress the large backgrounds from Drell-Yan and other multi-lepton processes, the process was only permitted to contain one loose lepton (which from the previous

requirement is in fact a tight lepton). Furthermore, each process was required to have $m_T^W > 50\text{GeV}$, where m_T^W is defined as:

$$m_T^W = \sqrt{2p_T^l p_T^{miss} (1 - \cos\Delta\phi(\vec{p}_T^l, \vec{p}_T^{miss}))} \quad (3)$$

This event selection give efficiencies of about 11% for the 400GeV hypothesis and 14% for the 750GeV hypothesis, whilst the $t\bar{t}$ background efficiency was 4%. These efficiencies are slightly higher than those achieved in reference [14], largely because some of the selection requirements used there, notably the triggering requirements used to give comparability with experimental data, were not implemented here.

Reconstructing the neutrino

As has already been mentioned, neutrinos cannot be detected in the CMS detector, and must be reconstructed using the missing transverse momentum of the event and other constraints (the z component of the neutrino momentum cannot be identified from the missing momentum of the system in the z direction since the momentum of particles which escape along the beam-pipe is not known). Here the requirement was made that $(p(l) + p(\nu))^2 = m_W^2$ (which follows from the fact the lepton and neutrino were produced in a W decay), and that the neutrino was massless and on-shell (i.e. $p_x(\nu)^2 + p_y(\nu)^2 + p_z(\nu)^2 = E(\nu)^2$). One could then simply take the x and y components of the neutrino momentum to be those of the missing transverse momentum, and the z component and the energy could then be computed using these constraints. One problem was that the conditions gave a quadratic equation for p_z , which could have a pair of either real or complex solutions. For real solutions the solution with the smallest $|p_z|$ was taken, whilst for complex solutions the real part was taken, in accordance with the recommendations of [17].

Reconstructing the ttbar system

Two light jets were chosen from the list of all light jets which passed the quality requirements such that the sum of their 4-momenta gave the closest invariant mass to that of the W boson. For simplicity, only the two leading b jets (highest p_T) were considered (though since most of the selected events only had two b-tagged jets this is probably not a significant issue. The one of these b jets which, when its four momentum was added to that of the lepton and the neutrino gave the closest invariant mass to the top quark was identified as coming from the leptonically decaying top quark, whilst the other was identified as coming from the hadronically decaying top quark. This then allowed the 4-momenta of the top quarks to be calculated as:

$$\begin{aligned} p_{t_{lep}} &= p_{b_{lep}} + p_l + p_\nu, \\ p_{t_{had}} &= p_{b_{had}} + p_q + p_{\bar{q}}. \end{aligned}$$

3.3 Event reconstruction with HOT

So far the decay products have been assumed to be distinct jets, however if a decaying object is heavy enough its decay products may all be merged into a single ‘‘fat jet’’. In this process it is quite likely that this will occur, especially for the 750GeV heavy Higgs

mass hypothesis, where the top quark transverse momentum is quite high, and so the two light jets may be merged into a single W jet, or they may be merged together with the b jet to give a top jet (Figure 3.3).

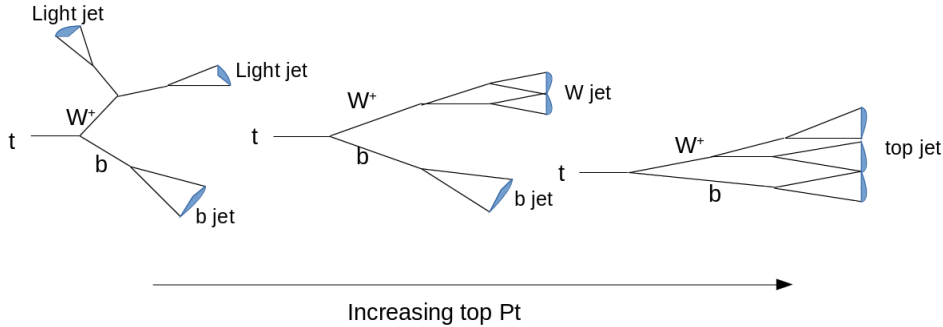


Figure 5: The sequence of merging as the top p_T is increased. Left: Unmerged jets, Center: A W jet and a b jet, Right: A top jet

Tagging fat jets

Fat jets were provided in the AK8 format, with top and W tagging discriminants provided by the Deep AK8 tagger, which utilises a deep neural network to exploit the full raw data from the jet (as opposed to simpler taggers which focus only on relevant derived variables such as the jet mass and the jet n-subjettiness). Since there are not yet any well-established working points for this tagger, simple loose, medium and tight working points were constructed from data by selecting working points that gave 10%, 1% and 0.1% mistag probabilities. Since the data also did not contain generator-level information about the flavour of the fat jets, cuts were made on the generator level mass, with jets with $m_{jet} > 160\text{GeV}$ taken as top jets, and those with $60\text{GeV} < m_{jet} < 120\text{GeV}$ taken as W jets. This gave loose/medium/tight efficiencies of 86%/59%/17.7% for top tagging and 51%/14.5%/2.7% for W tagging (though note these numbers are not necessarily that meaningful since the generator values were only taken from masses rather than flavour information). Due to relatively low numbers of events containing fat jet candidates, the loose working points were used for both types of tagging.

Leptons and b jets

In order to find the bjets and leptons necessary for reconstruction, the normal jets and leptons were treated in the same way as in section 3.2, with the exception that any jets within $\Delta R = \sqrt{\Delta\eta^2 + \Delta\phi^2} < 0.8$ of a fat jet which was tagged as either a top jet or a W jet were removed to prevent double counting.

Reconstructing the $t\bar{t}$ system

As with the case for no HOT, all events were required to have a single tight lepton and no additional loose leptons, and the neutrino could be reconstructed as in section 3.2. If a top jet was tagged, it was assumed to be the hadronically decaying top jet, and so the only additional requirement was to have one b jet (if there was more than one the leading b jet was taken), from which one could reconstruct the leptonically decaying top quark by adding the lepton and the neutrino. If a W jet was tagged, two b-tagged jets

were required, and the one giving the best leptonically decaying top mass was taken to come from that decay, and the other from the hadronic decay, as in section 3.2. The hadronically decaying top quark momentum can then just be found as the sum of the momenta of the W jet and the b jet coming from the hadronic decay.

The event selection efficiency for top tagging is about 0.5% for the 400GeV resonance and 5% for the 750GeV resonance, with a tagging efficiency for the $t\bar{t}$ background of 0.26%. For W tagging the selection efficiency is about 0.4% for the 400GeV mass and 2% for the 750GeV mass, with a background efficiency of 0.25%.

3.4 Search variables

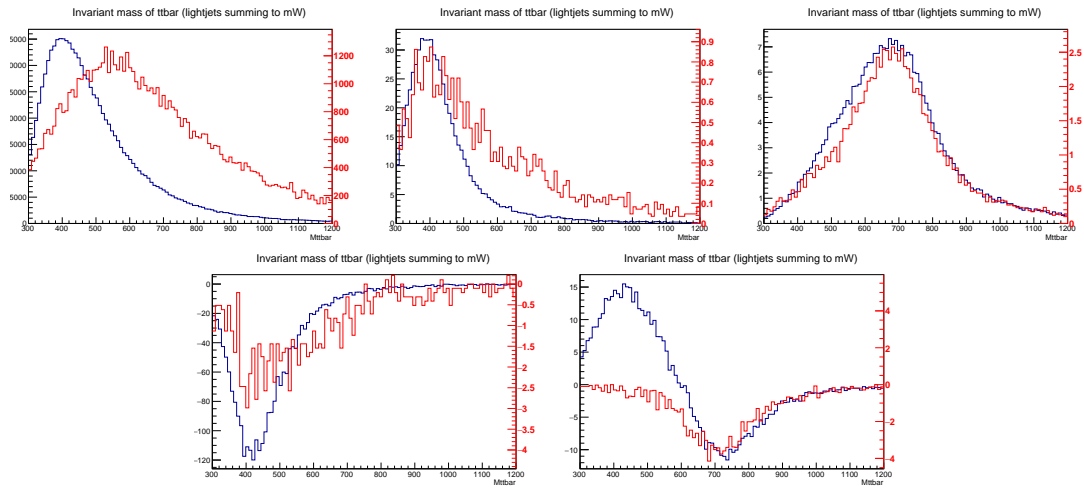


Figure 6: The various mass distributions for the $t\bar{t}$ system. Distributions obtained without HOT are in blue, those obtained with top tagging are in red. Top left: SM background $t\bar{t}$ production. Top center: Resonant distribution for a scalar 400 GeV heavy Higgs. Top right: Resonant distribution for a scalar 750 GeV heavy Higgs. Bottom left: Interference distribution for a scalar 400 GeV heavy Higgs. Bottom right: Interference distribution for a scalar 750 GeV heavy Higgs.

For use in later limit setting the events considered were divided into resonant and positive and negative interference components. The first search variable considered was the invariant mass of the combined $t\bar{t}$ system. As expected, the resonant peak appeared around the invariant mass of the new heavy Higgs boson in the signal samples (slightly shifted to the left as some energy is lost in the reconstruction), with the combined positive and negative interference components giving a peak-dip structure about this mass. The background roughly follows a Poisson distribution, with relatively few events below $2m_t = 340\text{GeV}$. Since top tagging tends to detect higher p_T and hence higher energy events, the background distribution is shifted to the right in this case, the 400GeV resonance has a large positive tail, and the interference graphs have the lower energy

peaks suppressed relative to the higher-energy dips (to the extent that the peak in the 750 GeV distribution completely disappears when using HOT).

The other search variable used was $\cos(\theta)$, where θ is the angle between the momentum of the leptonically decaying top jet in the rest frame of the $t\bar{t}$ system, and the momentum of the $t\bar{t}$ system in the detector frame of reference. This variable is believed to be sensitive to spin correlations, and hence should be different for the signal and background processes. The distribution of this observable is not symmetric, however the events were binned symmetrically for consistency with [14], though since these distributions are especially asymmetric for top tagging, it may be profitable to change this in future searches. The bins chosen were $|\cos(\theta)| < 0.4$, $0.4 < |\cos(\theta)| < 0.6$, $0.6 < |\cos(\theta)| < 0.75$, $0.75 < |\cos(\theta)| < 0.9$ and $|\cos(\theta)| > 0.9$, which reflect the fact that the $\cos(\theta)$ distribution changes more rapidly at high value than low ones. The $m_{t\bar{t}}$ distributions were then provided to the limit setter in a different channel for each of these bins.

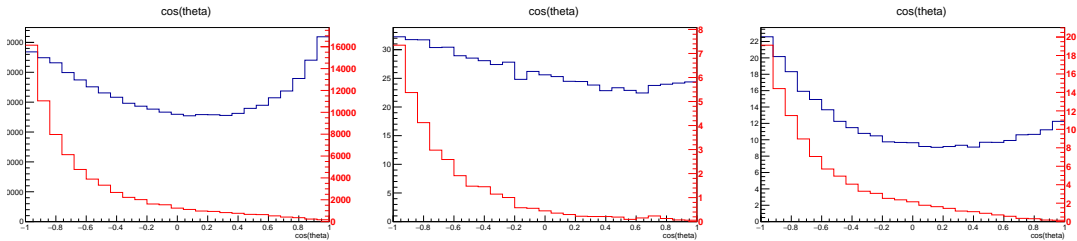


Figure 7: The distributions of the $\cos(\theta)$ variable, again with distributions obtained without HOT in blue and those with top tagging in red. Left: SM background $t\bar{t}$ production. Center: Resonant distribution for a scalar 400 GeV heavy Higgs. Right: Resonant distribution for a scalar 750 GeV heavy Higgs.

3.5 Results

Expected limits on the coupling constant of the new Higgs bosons to the top quark, denoted here as g , were set using the combine tool [18], using a custom interference model to handle the negative weight of that part of the interference. Limits were set using events which were only top tagged (green), only W tagged (red), and reconstructed without HOT (blue). Limits were additionally set using top and W tagging together (cyan), and all three methods combined (purple). When using multiple methods combined it was important to put the events tagged with the different methods into different channels, so that the larger backgrounds didn't drown out the signals. It was also important to ensure that the same event couldn't be tagged multiple times in the different channels. The limits are generally as one would expect: the case with top tagging performs better than the HOT for the 750 GeV heavy Higgs boson, and worse for the 400 GeV mass, whilst the W mass gives intermediate limits for both cases. Combining the types of tagging further improves on these limits. In general this improvement is about what one

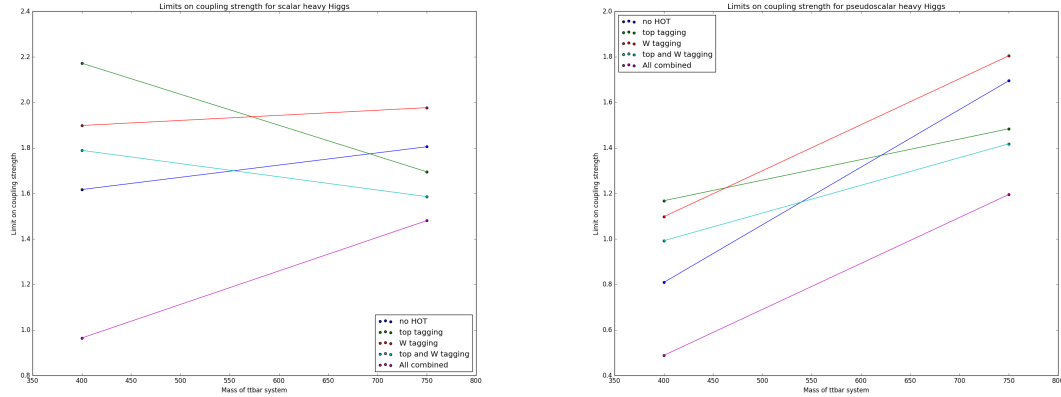


Figure 8: Limits from the different types of tagging. Left: Limits on the scalar couplings C_{HG} , Right: Limits on the pseudoscalar couplings C_{AG} .

would expect, however the limit on the scalar 400GeV coupling seems disproportionately large- the reasons for this are not well understood.

4 Conclusion

This project has demonstrated two important components of searches for new Higgs bosons: improving simulations and improving the detection of any signals produced at the LHC. The importance of reweighting for producing accurate predictions of the interference lineshape has been demonstrated, and this should be useful in future simulations of different theoretical scenarios. For signal detection, both top and W tagging have been shown to give notable improvements to the limits on coupling strengths which can be set, not just for higher heavy Higgs masses as expected but also for lower ones, so these techniques should be considered for further searches using LHC data. This is especially interesting as HOT could also be used in the fully hadronic $t\bar{t}$ decay channel, which has the largest branching ratio but is often neglected due to large QCD backgrounds, since one could potentially observe events with two top tags for very heavy Higgs bosons, which should give quite a distinctive signal. Using both improved simulations and these new tagging methods, one would hope that extended Higgs sector physics may be well searched for masses less than 1TeV and typical coupling constants over the coming years at the LHC.

5 Acknowledgements

I would like to thank my supervisor, Alexander Grohsjean, for providing me with two very interesting projects, and supporting me in undertaking these. I would also like to thank Gerrit van Onsem for providing me with all the ROOT files for my data analysis,

and being very helpful with the entire project. Finally, I would like to thank Olaf Behnke and the entire organising team for giving me and many other students the opportunity to come and experience the very interesting working environment here at DESY, and ensuring the program ran smoothly.

References

- [1] G. C. Branco et al., “Theory and phenomenology of two-Higgs-doublet models”, *Phys. Rept.* **516** (2012) 1-102, doi:10.1016/j.physrep.2012.02.002, arXiv:1106.0034.
- [2] A. Djouadi, “The anatomy of electro-weak symmetry breaking. II. The Higgs bosons in the minimal supersymmetric standard model”, *Phys. Rept.* **459** (2008) 1-241, doi:10.1016/j.physrep.2007.10.005, arXiv:hep-ph/0503173.
- [3] K. Agashe, R. Contino and A. Pomarol, “The minimal composite Higgs model”, *Nucl. Phys. B* **719** (2005) 165-187, doi:10.1016/j.nuclphysb.2005.04.035, arXiv:hep-ph/0412089
- [4] K. J. F. Gaemers and F. Hoogeveen, “Higgs Production and Decay Into Heavy Flavors With the Gluon Fusion Mechanism”, *Phys. Lett. B* **146** (1984) 347-349, doi:10.1016/0370-2693(84)91711-8.
- [5] D. Dicus, A. Stange and S. Willenbrock, “Higgs decay to top quarks at hadron colliders”, *Phys. Lett. B* **333** (1994) 126-131, doi:10.1016/0370-2693(94)91017-0, arXiv:hep-ph/9404359.
- [6] D. Franzosi, E. Vryonidou and C. Zhang, “Scalar production and decay to top quarks including interference effects at NLO in QCD in an EFT approach”, *J. High Energ. Phys.* (2017), doi:10.1007/JHEP10(2017)096, arXiv:hep-ph/1707.06760.
- [7] J. Alwall et al, “The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations”, arXiv:hep-ph/1405.0301
- [8] C. Zhang, D. Franzosi and E. Vryonidou, <http://hepmdb.soton.ac.uk/hepmdb:0717.0259>.
- [9] T. Sjöstrand et al., “An Introduction to PYTHIA 8.2”, *Computer Physics Communications* **191** (2015) 159-177, doi:10.1016/j.cpc.2015.01.024, arXiv:hep-ph/1410.3012.
- [10] F. Hagemann, “Testing resonant scalar top quark pair production at next-to-leading order in QCD”, DESY summer student project (2017).
- [11] Olivier Mattelaer, “On the maximal use of Monte Carlo samples: re-weighting events at NLO accuracy”, doi:10.1140/epjc/s10052-016-4533-7, arXiv:hep-ph/1607.00763
- [12] E. Conte, B. Fuks and G. Serret, “MadAnalysis 5, a user-friendly framework for collider phenomenology” *Comput. Phys. Commun.* **184** (2013) 222-256 doi:10.1016/j.cpc.2012.09.009 arXiv:hep-ph/1206.1599.
- [13] A. Anuar et al., “Search for a heavy Higgs boson decaying into dilepton $t\bar{t}$ final states at 13 TeV”, CMS Analysis Note AN-16-164, 2017

- [14] M.Gul et al., “Search for a heavy Higgs boson decaying into a pair of top quarks in the semileptonic final state with the 2016 pp dataset”, (Paper awaiting publication)
- [15] CMS Collaboration, “The CMS experiment at the CERN LHC”, *JINST* **3** (2008) S08004, doi:10.1088/1748-0221/3/08/S08004
- [16] CMS Collaboration, “Identification of b quark jets at the CMS experiment in the LHC Run 2”, CMS Physics Analysis Summary BTV-15-001, 2016
- [17] J. Kvita, “Study of methods of resolved top quark reconstruction in semileptonic $t\bar{t}$ decay”, *Nucl Instrum Methods Phys Res A* **900** (2018) 84-100, doi:10.1016/j.nima.2018.05.059, arXiv:hep-ex/1806.05463.
- [18] <https://twiki.cern.ch/twiki/bin/view/CMS/SWGuideHiggsAnalysisCombinedLimit>

6 Appendix: Details of simulation of interference lineshape

In order to simulate a sufficiently large number of events, the simulation was run as 100 batches of 5000 events each on a computer cluster using HTCondor. Since MG5_aMC@NLO does not yet have a formal “gridpack” system for NLO simulations, the first response to the question at <https://answers.launchpad.net/mg5amcnlo/+question/243268> was used. The events were run with reweighting switched on (the reweighting card just changing the parameter `cc` from 1 to -1) and showering switched off, since these options did not appear to work together, and the events were then showered using the command “./bin/shower run_XX”, where XX is the number of the run (if no other trial runs are carried out this will be run_02 since run_01 is produced when setting up the “gridpack”). Since MadAnalysis5 does not support reweighting information in “normal” mode (which uses the interactive python-based shell), it was necessary to use “expert” mode (where one codes directly in c++). To do this a program producing a graph of the M_T distribution using only the positive weight was produced in normal mode, and then the `Build/SampleAnalyzer/User/Analyzer/user.cpp` file inside the folder thus created was edited, and the line `__event_weight__ = event.mc()->weight()` was changed to:

```
WeightCollection reweightcont=event.mc()->multiweights();
std::map<MAuint32,MAfloat64> reweights=reweightcont.GetWeights();
__event_weight__ = (event.mc()->weight() - reweights.find(2)->second)*0.5;
```

This produced a graph in the .saf format, which was converted into a .png file using the saf2png module from <https://github.com/crokkon/saf2png>.