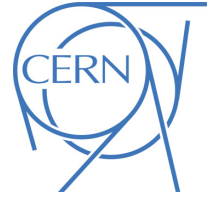**ATLAS Note**

ATL-COM-PHYS-2017-629

18th May 2017

Draft version 0.1

Improved signal–background discrimination in searches for di-Higgs production in the $bb\tau_\text{had}\tau_\text{lep}$ channel at 13 TeV using a novel approach to kinematic reconstruction

Agni Bethani, Jonathon Langford, Katie Spencer and Terry Wyatt.

University of Manchester, UK.

Events containing two b-jets and two candidate τ-lepton decays are tested for their kinematic consistency with the hypothesis $hh \to bb\tau_h\tau_l$. In this method the measured b-jet energies are re-scaled, subject to constraints on the resulting corrected masses of the bb and $\tau_h\tau_l$ systems, on the magnitude and direction of the corrected $\not{E}_T$, and on the factors by which the b-jet energies are re-scaled. An optimimum choice of b-jet energy re-scaling is obtained event by event by minimising a measure, ψ, of the kinematic consistency of an event with the signal hypothesis. The resulting value of ψ provides an event-by-event discriminant against background processes, since these are unlikely to yield events kinematically consistent with the signal hypothesis, even after re-scaling of the b-jet energies. We combine this novel discriminant with the output from the BDT that has been designed for the ongoing ATLAS search for $hh \to bb\tau_h\tau_l$; we find that this approach yields a significant reduction in the level of selected background at constant signal selection efficiency, and thus a higher a priori sensitivity to the presence of a signal, when compared to a selection that uses the BDT output alone. This is particularly the case for low higgs boson pair masses, where the kinematic discrimination of the BDT is modest. For example, in the case of a graviton resonance of mass 300 GeV a reduction by roughly a factor of two in background can be achieved for the same signal efficiency. Monte Carlo simulations are found to describe well the distributions of relevant kinematic variables in the 36.1 fb^{-1} data set collected by ATLAS at 13 TeV (blinded in regions of high signal sensitivity). We propose when the 36.1 fb^{-1} data set is unblinded that the distribution of ψ in a signal-enhanced region would, at bare minimum, provide a useful cross check of the analysis that uses the BDT output alone.

Contents

1 Introduction and Overview

It is widely accepted that new physics must exist beyond the Standard Model (BSM), and the discovery of a 125 GeV resonance by both ATLAS and CMS at the Large Hadron Collider (LHC) in 2012 opened up promising new channels for BSM searches [1, 2]. The properties of this resonance, including its dominant production mechanisms and branching fractions, are so far found to be consistent with Standard Model (SM) predictions for the Higgs boson, h [3]. However, for the properties of the Higgs potential, and consequently the mechanism of electroweak symmetry breaking to be tested in full, measurements must be made of the strength of the self-coupling, λ_{hhh}, of the Higgs boson.

In the SM the production of events containing a pair of Higgs bosons can occur via the processes illustrated in figure 1. Di-Higgs searches offer the best direct probe of λ_{hhh} and hence present an important

(a) Production via the Higgs trilinear self-coupling.   (b) Production via Higgs-fermion coupling.
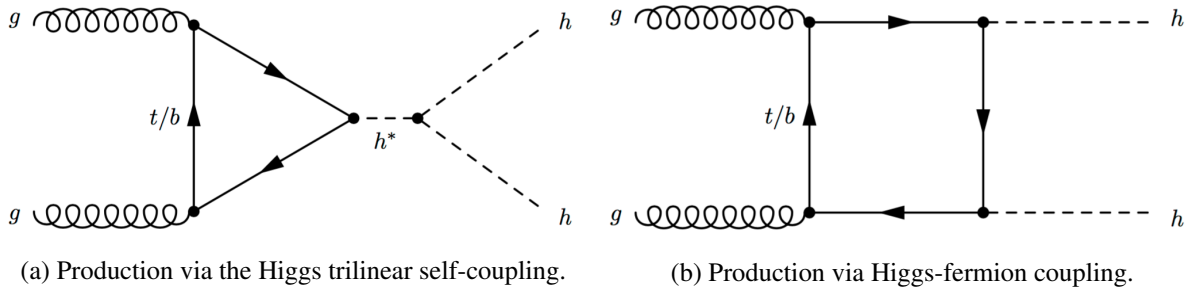
Figure 1: Leading order Feynman diagrams for non-resonant SM di-Higgs production at the LHC [4].
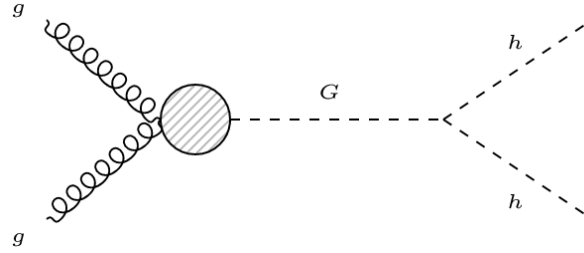
Figure 2: Leading order Feynman diagrams for resonant di-Higgs production via a graviton at the LHC.

opportunity for testing the SM at the LHC and beyond [5]. Searches for SM production have been conducted by ATLAS in Run 1 using 20.3 fb^{-1} of pp collision data at a centre-of-mass energy $\sqrt{s}$ = 8 TeV in the $bb\gamma\gamma$, $bb\tau\tau$, $\gamma\gamma WW^*$ and $bbbb$ channels [4]. A combined analysis allowed a 95% confidence level (CL) upper-limit to be set on the observed cross-section at 70 times the SM prediction. Studies have continued in Run 2 at $\sqrt{s}$ = 13 TeV with 13.2 fb^{-1} of collision data in the $\gamma\gamma WW^*$ and $bbbb$ channels and 3.3 fb^{-1} in the $bb\gamma\gamma$ channel[6–8]. Despite the current progress, the small cross-section for production means that sensitivity to a SM di-Higgs signal will not be possible until the High-Luminosity stage of the LHC (HL-LHC) from 2026 onwards [9].

Nevertheless, searches for di-Higgs production present an opportunity to discover new physics before the HL-LHC due to the several potential BSM enhancements to the SM cross-section. These could arise, for example, through the resonant decay of a new TeV-scale particle such as a Kaluza-Klein (KK) graviton G (as illustrated in figure 2) in the Randall-Sundrum (RS) model or a heavy neutral Higgs H boson in the Minimal Supersymmetric Standard Model (MMSM) [10, 11]. In light of this, alongside searches for SM di-Higgs production in the above mentioned ATLAS analyses, resonant pair production has also been explored.

This note presents a novel approach to event reconstruction in the di-Higgs $bb\tau_h\tau_l$ final state, where the subscripts h and l refer to hadronically and leptonically decaying τs respectively. In this method the measured b-jet energies are re-scaled, subject to constraints on the resulting corrected masses of the bb and $\tau_h\tau_l$ systems, on the magnitude and direction of the corrected $\not{E}_T$, and on the factors by which the b-jet energies are re-scaled.

The initial development of the analysis is described in section 3. This is performed using data and Monte Carlo (MC) corresponding to 13.2 fb^{-1} of integrated luminosity at proton-proton centre-of-mass energy $\sqrt{s}$ = 13 TeV collected by the ATLAS experiment in 2015 and the early part of 2016. The initial development is presented in the context of BSM resonant production of a Higgs boson pair, but the derived method is equally valid in the search for processes that lead to a continuum distribution for the mass of the Higgs boson pair (as is the case in the SM). An optimimum choice of b-jet energy re-scaling is obtained event by event by minimising a measure, ψ, of the kinematic consistency of an event with the signal hypothesis. The resulting value of ψ provides an event-by-event discriminant against background processes, since these are unlikely to yield events kinematically consistent with the signal hypothesis, even after re-scaling of the b-jet energies.

The results of applying this new analysis method are presented in section 4. An initial demonstration of the performance of the new discriminant using the 13.2 fb^{-1} data set is presented in section 4.1. The analysis is then extended to the full currently available (2015–2016) run 2 ATLAS data and Monte Carlo samples corresponding to an integrated luminosity of 36.1 fb^{-1}. In section 4.2 results are presented

that demonstrate that even after a tight requirement on the $bb\tau_h\tau_l$ BDT score [12] is imposed the ψ variable provides considerable additional discriminating power against background. In section 4.3 the applicability of the new approach using the variable ψ to searches for SM continuum production of di-Higgs events is explored.

A summary and conclusions are presented in section 5, including a concrete proposal for how the new discriminant ψ could, at bare minimum, provide a useful cross check of the analysis that uses the BDT output alone once the 36.1 fb^{-1} ATLAS data set is unblinded.

2 ATLAS Data and Monte Carlo Samples

The ATLAS detector is described in [13].

This analysis makes use of the data and MC ntuples produced for the $bb\tau_h\tau_l$ search described in [12]. A summary of the MC generators and PDFs used to create the MC samples for background and signal processes is given in table 1.

Process	Event Generator	PDF
Background		
$t\bar{t}$, single-top s, Wt channels	Powheg-Box [14] & Pythia 6.428 [15]	CT 10 [16] & CTEQ6L1 [17]
single-top t channel	Powheg-Box & Pythia 6.428	CT10f4 [16] & CTEQ6L1
$W/Z + jets$	MG5_aMC [18] & Pythia 8.186 [19]	NNPDF2.3LO [20]
di-boson	Sherpa 2.1.1 [21]	CT10
Vh	Powheg-Box v2 & Pythia 8.186	CTEQ6L1
Signal		
Resonant $G \to hh$	MG5_aMC & Pythia 8.186	NNPDF2.3LO
SM di-Higgs production	MG5_aMC & Herwig [22]	CTEQ6L1

Table 1: A summary of the event generators and PDFs used to create the MC samples for both background and signal processes. The signal samples assume a Higgs boson mass of $m_h = 125.09$ GeV.

Full details of the employed object selections and corrections, overlap removal, event selections, etc, are given in [12]. In brief, events are required to contain exactly two jets with transverse momentum $p_T > 20$ GeV and pseudorapidity $|\eta| < 2.4$ that are b-tagged according to the MV2c10 algorithm at its '70% efficiency' working point. At least one of the jets must satisfy $p_T > 45$ GeV. Events must contain exactly one candidate τ_h with $p_T > 20$ GeV and $|\eta| < 2.5$ that is associated with one or three tracks in the inner detector and satisfies the 'medium' working point of a BDT-based identification algorithm. Exactly one electron or muon candidate with opposite charge to the τ_h and passing a 'loose' working point requirement on isolation is required. Events are required to pass a single electron or single muon trigger and the offline requirement on the lepton p_T is set to be 1 GeV above that required in the trigger (which is typically in the range 24–27 GeV, depending on lepton type and data taking period). Matching is required between the lepton candidate and the object that fired the single lepton trigger. Electrons are required to satisfy $|\eta| < 2.47$ and pass a 'medium' or 'tight' working point requirement on identification (depending on data taking period) and muons are required to satisfy $|\eta| < 2.7$ and pass a 'loose' working point requirement on identification. Electron and τ_h candidates in the barrel-endcap transition region of the calorimeter ($1.37 < |\eta| < 1.52$) are vetoed.

3 Methodology

The method to improve the signal–background discrimination in the search for di-Higgs boson production in the $bb\tau_h\tau_l$ final state is presented in this section. Firstly, the expected characteristics of a signal event are outlined (section 3.1). Secondly, the initial event reconstruction is defined and the observed kinematic characteristics are compared to those expected (section 3.2). The results motivate a novel reconstruction method for analysing the underconstrained system of particles, which is subsequently described in detail (section 3.3). Following this, the motivation for an event-by-event overall optimisation is presented with a demonstration that this provides significantly improved signal–background discrimination in di-Higgs boson searches (section 3.4). Note, all figures and results in this section are produced with the data and Monte Carlo samples corresponding to a 13.2 fb^{-1} ATLAS data set for which the Higgs boson pair search has been unblinded. The term *raw* is used to define variables before the novel reconstruction is applied, and *corrected* indicates variables calculated after the reconstruction. This section outlines the development of the analysis technique with respect to signal samples corresponding to the resonant production of gravitons (G) of various masses that decay to a pair of Higgs bosons[1].

3.1 The expected topology of an 'ideal' signal event

The novel approach to signal–background discrimination makes use of the expected kinematic characteristics of a Higgs boson pair event decaying into the $bb\tau_h\tau_l$ final state. Therefore, it is necessary to outline the expected topology of a signal event, which is later used to assist in background rejection. Figure 3 presents the expected characteristics of a $hh \rightarrow bb\tau_h\tau_l$ event, in the transverse plane. The variable $\not{E}_T$ represents the missing transverse energy in the system. For resonances with mass: $m_G \gg 2m_h$, the final state particles are expected to be boosted along their parent Higgs boson direction due to the initial momenta of the Higgs bosons. Thus, the bb di-jet system and the $\tau_h\tau_l$ system are expected to be narrowly collimated. This effect becomes less pronounced with decreasing resonance mass as the Higgs pair are produced with smaller relative momentum. As a result, the expected opening angles of the bb and $\tau_h\tau_l$ systems increase. Ultimately, in the $m_G = 260$ GeV signal sample, the Higgs pair are produced approximately at rest, hence there is no preferred direction for the final state particles.

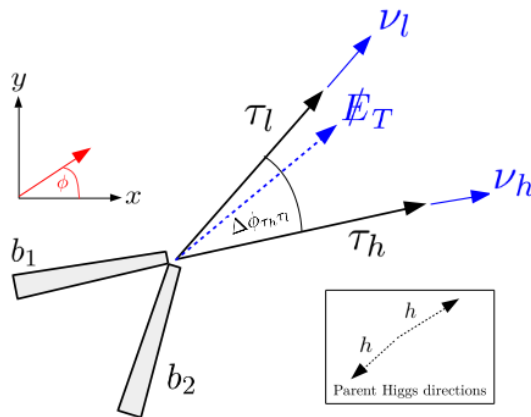Figure 3: A diagram showing the expected topology of a signal event in the transverse plane.

---

[1] The application of the analysis to the search for non-resonant SM production of Higgs boson pairs is presented in section 4.3.

Initially, the $\not{E}_T$ is assumed to be exclusively due to neutrinos in the subsequent τ decays. The neutrinos, represented by $\nu_{h/l}$ in figure 3, are highly boosted along their parent τ direction. The approximation is made that the neutrinos are assumed to be collinear with their respective τ. Also, the τ_l decay involves two neutrinos, whereas the τ_h involves only one. As a result, a larger missing energy component is expected along τ_l. Assuming perfect reconstruction of all physics objects, the 'ideal' signal event is defined by the following constraints:

- Reconstructing the invariant mass of the $\tau_h\tau_l$ system, $M_{\tau\tau}$, including an appropriate decomposition of the observed $\not{E}_T$ into ν_h and ν_l, gives m_h.

- Reconstructing the invariant mass of the bb system, M_{bb}, gives m_h.

- The missing energy, $\not{E}_T$, is directed inside the azimuthal opening angle of the $\tau_h\tau_l$ system, with a preference for lying closer to τ_l.

- Also, the invariant mass of the four-body system, $M_{bb\tau\tau}$, incorporating the missing energy into the reconstruction, should be consistent with a peak at the resonance mass.

3.2 Initial reconstruction and observed event characteristics

As mentioned in section 3.1, the $\not{E}_T$ is initially assumed to be due to the neutrinos produced in the τ decays. Thus, the $\not{E}_T$ vector is expected to lie within the $\tau_h\tau_l$ opening angle. To investigate this property, a parameter Ω is defined to characterize the direction of the $\not{E}_T$ with respect to the $\tau_h\tau_l$ system. Firstly, an intermediate parameter ω is calculated using equation 1

$$\omega = \frac{\Delta\phi_{\tau_{close},\not{E}_T}}{\Delta\phi_{\tau_h,\tau_l}}. \tag{1}$$

The variable $\Delta\phi_{\tau_{close},\not{E}_T}$ refers to the difference in ϕ angle between the $\not{E}_T$ vector and its closest τ, and $\Delta\phi_{\tau_h,\tau_l}$ represents the azimuthal opening angle between τ_h and τ_l in the transverse plane, as illustrated in figure 4.

The final parameter, Ω, is then computed as follows:

- For events in which the $\not{E}_T$ lies inside of the $\tau_h\tau_l$ opening angle

$$\Omega = 0 + \omega, \tag{2}$$

when the $\not{E}_T$ is closest to τ_h, and

$$\Omega = 1 - \omega, \tag{3}$$

when closest to τ_l. In summary, for the subset of events with $\not{E}_T$ inside $\tau_h\tau_l$, $\Omega \in [0,1]$, where $\Omega = 0$ and $\Omega = 1$ correspond to the $\not{E}_T$ lying along the τ_h and τ_l directions respectively.

- To ensure Ω is a continuous variable, events with $\not{E}_T$ outside of the $\tau_h\tau_l$ system but closest to τ_h are chosen to have

$$\Omega = 0 - \omega, \tag{4}$$

and when the $\not{E}_T$ is outside but closest to τ_l
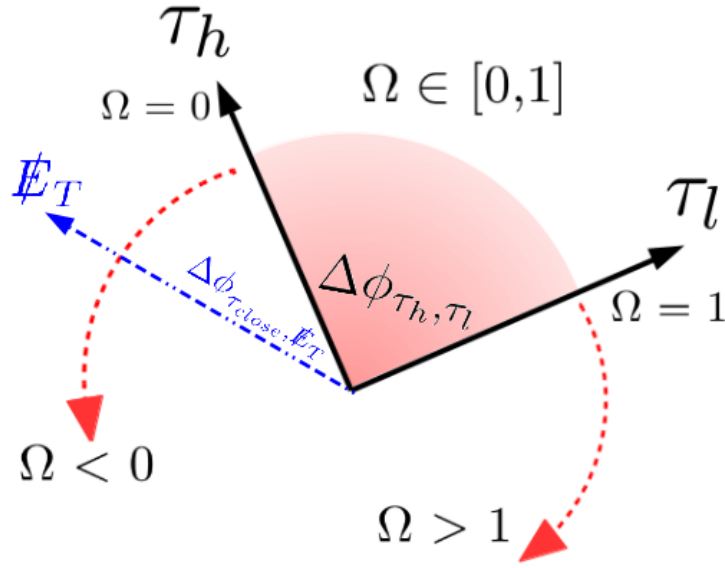
$$\Omega = 1 + \omega. \tag{5}$$

Figure 4: A diagram to summarise the Ω parameter, where an example $\not{E}_T$ vector is drawn along with the required angles.

The different regions in the transverse plane are labelled by their respective Ω values in figure 4. An example $\not{E}_T$ vector is also shown, where the relevant angles have been labelled. Ω is designed to scale with the opening angle of the $\tau_h\tau_l$ system and therefore should be independent of the resonance mass. Figure 5 shows a comparison between data and SM Monte Carlo for the raw Ω distribution.

Distributions corresponding to resonant production of gravitons with masses m_G in the range 260–500 GeV are also shown in figure 5. The distribution of Ω shown in figure 5 contradicts the naive assumption that the $\not{E}_T$ lies within the $\tau_h\tau_l$ system: showing that approximately 40% of signal events[2] are defined by $\Omega \notin [0,1]$. As a result, separate reconstruction methods are implemented to incorporate $\not{E}_T$ into the event for the cases that $\not{E}_T$ lies inside and outside the $\tau_h\tau_l$ system. However, figure 5 does demonstrate a number of the expected characteristics. Importantly, a greater fraction of events satisfy $\Omega \in [0,1]$ for di-Higgs signal samples compared to the background, in which the distribution extends out to much larger $|\Omega|$ values. Also, for events with $\not{E}_T$ inside $\tau_h\tau_l$, the signal distribution peaks mostly towards τ_l ($\Omega \approx 1$), whereas background favours $\not{E}_T$ closer to τ_h ($\Omega \approx 0$). This offers further discriminating power for events with $\Omega \in [0,1]$. Another key observation is the *scaling* behaviour of Ω. As expected, all signal distributions follow extremely similar shapes, regardless of the graviton mass. Hence, Ω offers a scale-invariant parameter to aid in signal–background discrimination.

To reconstruct an event with the $\not{E}_T$ vector directed within the transverse opening angle of the $\tau_h\tau_l$ system, two effective massless particles, ν_h and ν_l, are incorporated into the reconstruction. The particles correspond to the neutrinos associated with the visible τ-leptons and, assuming collinearity, inherit the (η,ϕ) values of their respective visible τ decay products. The transverse momenta, $p^T_{\nu_{h/l}}$, of $\nu_{h/l}$ are calculated by assuming that their vector sum is equal to the observed $\not{E}_T$. The following simultaneous

---

[2] It can be noted that the fraction of signal events in which the $\not{E}_T$ does not lie within the $\tau_h\tau_l$ system is constant at around 40%, independent of the value of m_G.
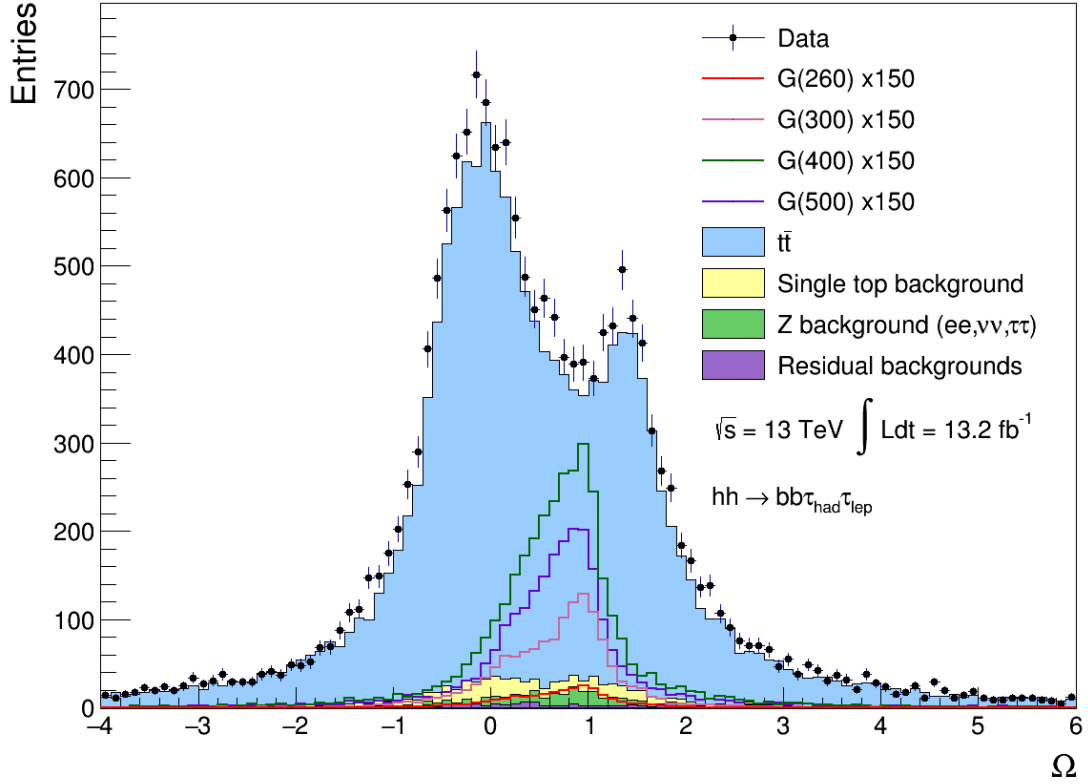
Figure 5: Distribution of the raw Ω parameter. Data are shown as the points with error bars, which are purely statistical. The SM Monte Carlo backgrounds are stacked to compare with data. Systematic uncertainties have not been evaluated. Distributions corresponding to resonant production of gravitons with masses m_G in the range 260–500 GeV are also shown. Note that the signal distributions have been multiplied by a factor of 150 to enable comparison with dominant backgrounds.

equations define $p^T_{\nu_{h/l}}$, where $\not{E}_x$ and $\not{E}_y$ are the x and y components of the missing transverse energy

$$\begin{aligned} p^T_{\nu_h} \cos \phi_{\tau_h} + p^T_{\nu_l} \cos \phi_{\tau_l} &= \not{E}_x, \\ p^T_{\nu_h} \sin \phi_{\tau_h} + p^T_{\nu_l} \sin \phi_{\tau_l} &= \not{E}_y. \end{aligned} \tag{6}$$

The variable ϕ_{τ_h} (ϕ_{τ_l}) corresponds to the azimuthal angle of the visible decay products of the hadronically (leptonically) decaying τ. This leads to the following expressions for $p^T_{\nu_{h/l}}$

$$\begin{aligned} p^T_{\nu_h} &= \frac{\not{E}_x - \not{E}_y \cot \phi_{\tau_l}}{\cos \phi_{\tau_h} - \sin \phi_{\tau_h} \cot \phi_{\tau_l}}, \\ p^T_{\nu_l} &= \frac{\not{E}_x - \not{E}_y \cot \phi_{\tau_h}}{\cos \phi_{\tau_l} - \sin \phi_{\tau_l} \cot \phi_{\tau_h}}. \end{aligned} \tag{7}$$

Subsequently, a full 3D reconstruction of the $\not{E}_T$ in the system is possible, assuming ν_h and ν_l inherit the η values of their parent τ. The z-component of the missing energy enters the system via the z-component of the effective particles' momenta, $p^z_{\nu_{h/l}}$, defined by equation 8

$$p^z_{\nu_i} = p^T_{\nu_i} \cot \theta_{\tau_i}, \tag{8}$$

where θ_{τ_i} represents the polar angle of the visible τ_i and $i = h, l$. Equation 9 is used to reconstruct the invariant mass of the $\tau_h \tau_l$ system,

$$M_{\tau\tau}{}^2 = (q_{\tau_h} + q_{\tau_l} + q_{\nu_h} + q_{\nu_l})^2, \tag{9}$$

where q_{τ_h}, q_{τ_l}, q_{ν_h} and q_{ν_l} are the four-momenta of the τ_h, τ_l, ν_h and ν_l respectively. To clarify, the four-momenta of the effective particles are defined as:

$$q_{\nu_i}^{\mu} = (E_{\nu_i}, p_{\nu_i}^T \cos \phi_{\tau_i}, p_{\nu_i}^T \sin \phi_{\tau_i}, p_{\nu_i}^z)^T, \tag{10}$$

where $E_{\nu_i} = \sqrt{p_{\nu_i}^{T}{}^2 + p_{\nu_i}^{z}{}^2}$ is the energy of the effective particle, and again, subscript $i = h, l$.

For events in which the missing energy lies outside of the $\tau_h \tau_l$ system ($\Omega \notin [0,1]$) and within $90°$ of the closest τ, a second reconstruction method is used. In the invariant mass reconstruction, only the component of the $\not{E}_T$ vector along the direction of the closest τ is included. Thus, in contrast to the method described above, only one effective massless particle, ν_{close}, is introduced. The effective particle is defined to be collinear with the closest τ, inheriting the τ's η and ϕ values. Equations 11 and 12 define the effective particle's transverse momentum and z-component of the momentum respectively

$$p_{\nu_{close}}^T = \sqrt{(\not{E}_x \cos \phi_{\tau_{close}})^2 + (\not{E}_y \sin \phi_{\tau_{close}})^2}, \tag{11}$$

$$p_{\nu_{close}}^z = p_{\nu_{close}}^T \cot \theta_{\tau_{close}}. \tag{12}$$

The parameter $\phi_{\tau_{close}}$ ($\theta_{\tau_{close}}$) is the azimuthal (polar) angle of the closest τ. Equation 13 is used to calculate $M_{\tau\tau}$ for this subset of events, where $q_{\nu_{close}}$ is the four-momentum of ν_{close}

$$M_{\tau\tau}{}^2 = (q_{\tau_h} + q_{\tau_l} + q_{\nu_{close}})^2. \tag{13}$$

Note that, for events in which the $\not{E}_T$ lies outside the $\tau_h \tau_l$ system and **further** than $90°$ from the closest τ, there is no sensible reconstruction method. These events are unlikely to be signal and are thus removed. Figure 6 summarises the missing energy reconstruction method implemented for the different situations.

For all events, the invariant mass of the b-tagged jets is calculated using equation 14, where the four-momenta of the leading (sub-leading) b-jet is given by q_{b_1} (q_{b_2})

$$M_{bb}{}^2 = (q_{b_1} + q_{b_2})^2. \tag{14}$$

The total invariant mass of the $bb\tau_h\tau_l$ system is reconstructed by summing the four momenta of all final state particles including the missing energy contribution. For events in which the $\not{E}_T$ vector is directed inside the $\tau_h \tau_l$ opening angle, equation 15 is used

$$M_{bb\tau\tau}{}^2 = (q_{b_1} + q_{b_2} + q_{\tau_h} + q_{\tau_l} + q_{\nu_h} + q_{\nu_l})^2. \tag{15}$$

This incorporates the two effective particles, ν_h and ν_l, defined to be collinear with the visible τ candidates. Equation 16 calculates the total invariant mass for events in which the $\not{E}_T$ lies outside the $\tau_h \tau_l$ system, but within $90°$ of the closest τ

$$M_{bb\tau\tau}{}^2 = (q_{b_1} + q_{b_2} + q_{\tau_h} + q_{\tau_l} + q_{\nu_{close}})^2. \tag{16}$$
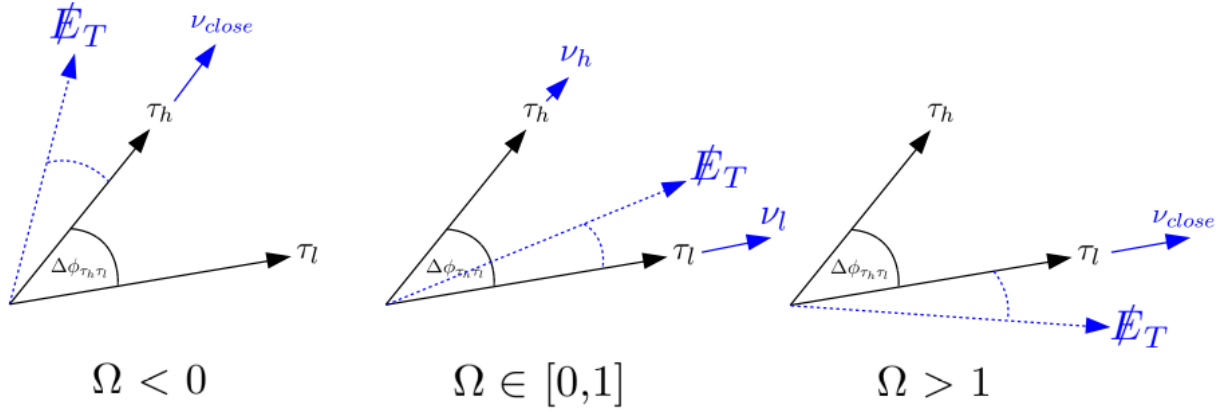
Figure 6: A series of diagrams to show the missing energy reconstruction for the different situations. On the left, represents $\Omega < 0$, when the $\not{E}_T$ lies outside on the τ_h side. The middle diagram shows the reconstruction technique for $\not{E}_T$ inside of the $\tau_h\tau_l$ opening angle ($\Omega \in [0,1]$). Finally, the diagram on the right is for $\Omega > 1$, where the missing energy lies outside but closer to τ_l.

This incorporates ν_{close} into the reconstruction.

The raw M_{bb}, $M_{\tau\tau}$ and $M_{bb\tau\tau}$ distributions are presented in figure 7. Note that events in which $\not{E}_T$ lies outside $\tau_h\tau_l$ and further than 90° from the closest τ have been omitted from the distributions. As before, the signal distributions have been multiplied by a factor of 150 to allow comparison with a much larger background. The M_{bb} spectra for the signal samples peak to the left of the Higgs mass (m_h = 125 GeV) in figure 7(a) and are characterized by a long tail on the low mass side, suggesting jet momenta are often underestimated. This property is somewhat accounted for by the uncertainty on jet reconstruction, including the presence of neutrinos produced in the semileptonic decays of b-flavoured hadrons. The signal distributions for $M_{\tau\tau}$ in figure 7(b) also peak to the left of m_h. The signal distributions in figure 7(c) peak to the left of the relevant graviton mass. Also, the signal $M_{bb\tau\tau}$ spectra are relatively broad. Combining these observations with the Ω distribution in figure 5, emphasises that a significant proportion of signal events do not conform to the ideal $hh \rightarrow bb\tau_h\tau_l$ kinematics, as outlined in section 3.1. Poor reconstruction of physics objects and detector resolution effects are modifing the true kinematics of the event, such as the $\not{E}_T$ direction and the invariant masses. Therefore, some way of incorporating these effects must be included in the analysis.

3.3 Novel reconstruction method: scaling the b-jets

The discrepancy between the event kinematic characteristics observed in figures 5 and 7 and the naively expected 'ideal' kinematic characteristics motivates the novel reconstruction method. The reconstruction technique subtly adjusts the final state kinematics to check if an event can be made to look consistent with di-Higgs boson production. In turn, this may offer potential discriminating power between signal and background by demanding that the new, scaled event conforms to the 'ideal' characteristics.

In figure 7(a), it is evident there exists a large uncertainty on the b-jet object reconstruction. Poor energy reconstruction of the b-jets can subsequently modify the measured $\not{E}_T$, thus affecting all of the variables defined in section 3.2 (Ω, M_{bb}, $M_{\tau\tau}$ and $M_{bb\tau\tau}$). Therefore, by scaling the transverse momenta of the
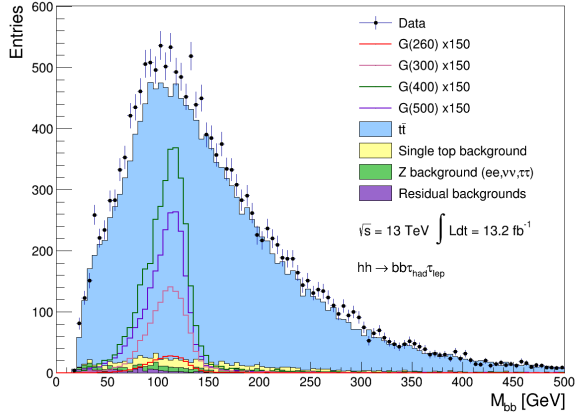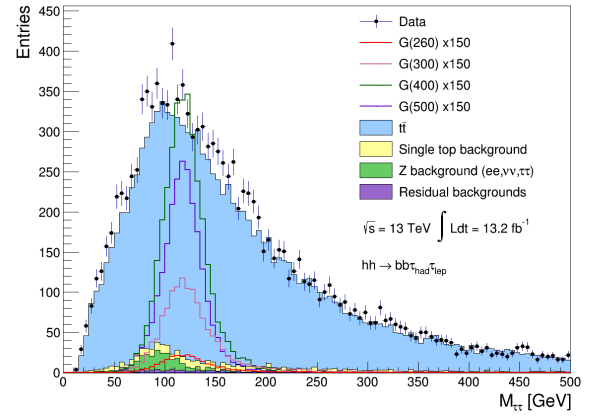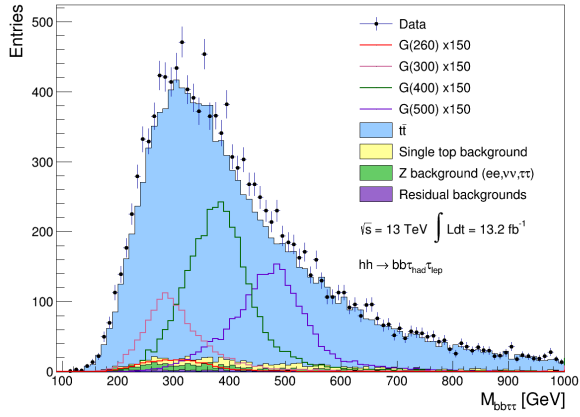
(a) Raw M_{bb}.

(b) Raw $M_{\tau\tau}$.

(c) Raw $M_{bb\tau\tau}$.

Figure 7: Raw invariant mass distributions for different systems of final state particles. The signal distributions have been multiplied by a factor of 150.

b-jets, it may be possible to correct the mismeasurement of the bb di-jet system and, in turn, rectify the kinematics of the event. For example, the corrected $\not{E}_T$ may be pushed inside of the $\tau_h\tau_l$ system and $M_{bb}/M_{\tau\tau}$ may be driven towards m_h. Ultimately, the scaled event is classified to be signal-like if the corrected kinematics adhere to the characteristics outlined in section 3.1.

Event by event, the algorithm iterates over p_T scale factors, χ_1 and χ_2, for the leading and sub-leading b-jets respectively, such that the corrected transverse momenta, $p_{b_i}^T{}'$, are given by equation 17

$$p_{b_i}^T{}' = \chi_i\, p_{b_i}^T, \tag{17}$$

where the subscript $i = 1,2$, labels the leading and sub-leading b-jets and $p_{b_i}^T$ represents the un-scaled (*raw*) jet transverse momenta. Initially, χ_1 and χ_2 were set to run over one hundred and fifty iterations between 0.5 and 2. These χ limits span a wide kinematic phase space, increasing the likelihood that an event can be scaled to look signal-like. It is unlikely that the raw b-jet momenta are mismeasured by a factor of 2 from their true values. However the χ limits of 0.5 and 2 offer a good starting point and the range is restricted later in the analysis. For each (χ_1,χ_2) combination, the corrected kinematic variables: $\not{E}_T'$, Ω', M_{bb}', $M_{\tau\tau}'$ and $M_{bb\tau\tau}'$ are computed. First, the corrected missing transverse energy is derived by imposing conservation of momentum in the x-y plane. Equation 18 is used to calculate the x and y components of the missing energy solely due to the poor energy reconstruction of the b-jets, $\not{E}_x(b_i)$ and $\not{E}_y(b_i)$,

$$\not{E}_x(b_i) = (\chi_i - 1)p_{b_i}^T \cos\phi_{b_i},$$
$$\not{E}_y(b_i) = (\chi_i - 1)p_{b_i}^T \sin\phi_{b_i}, \tag{18}$$

where ϕ_{b_i} is defined as the azimuthal angle of the ith jet. Equation 19 is then implemented to extract the corrected x and y components of the missing energy

$$\not{E}_x{}^{corr} = \not{E}_x{}^{raw} - \sum_{i=1}^{2} \not{E}_x(b_i),$$
$$\not{E}_y{}^{corr} = \not{E}_y{}^{raw} - \sum_{i=1}^{2} \not{E}_y(b_i). \tag{19}$$

After correcting the missing transverse energy, the other variables are determined using the reconstruction methods outlined in section 3.2, but replacing the raw $\not{E}_T$ with $\not{E}_T' = (\not{E}_x{}^{corr}, \not{E}_y{}^{corr})$.

This analysis is performed for each (χ_1,χ_2) combination, scanning the kinematic phase space to search for a scaling that adjusts the final system towards the characteristics of a $hh \to bb\tau_h\tau_l$ process. A schematic portraying the method is presented in figure 8. In qualitative terms the approach can be described as follows: if a scale factor combination exists which fine-tunes the kinematics to conform to the 'ideal' characteristics then it is accepted as signal; otherwise the event lies too far in the kinematic phase space and it is classified as background.

As a first attempt at a more quantitative investigation, the distribution of scale factor combinations in the χ_1-χ_2 plane was investigated by demanding that the kinematic properties of the scaled events satisfy three constraints which correspond to the 'ideal' signal characteristics. The initial constraints were defined as follows:

    1. The corrected invariant mass of the b-jet system, M_{bb}' is within 1 GeV of m_h.
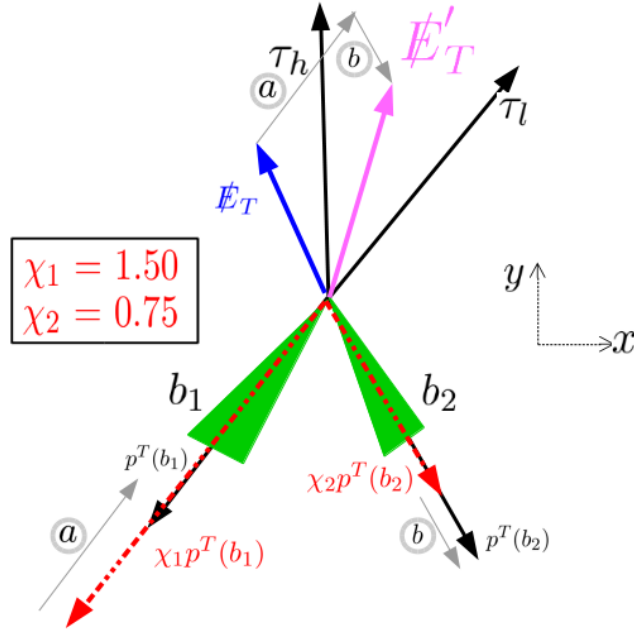
Figure 8: A schematic of an example event in the plane transverse to the beam direction to explain the novel reconstruction method. The originally measured raw p_Ts of the b-jets are indicated as black solid arrows. The p_Ts of the b-jets after re-scaling are indicated as red dot-dashed arrows. Example scale factors: $(\chi_1, \chi_2) = (1.50, 0.75)$ are indicated. The measured raw $\not{E}_T$ is indicated as the blue solid arrow. The corrections to $\not{E}_T$ resulting from the re-scaling of the b-jet p_Ts are indicated by the light grey solid arrows. The final corrected missing energy, $\not{E}'_T$, is indicated by the magenta solid arrow.

2. The corrected missing transverse energy, $\not{E}_T$, lies inside of the $\tau_h \tau_l$ opening angle or within an angle $10°$ to the closest τ.

3. The corrected invariant mass of the $\tau_h \tau_l$ system, $M'_{\tau\tau}$ is within 3 GeV of m_h.

If no (χ_1, χ_2) combination can satisfy all three constraints simultaneously, then the event is removed. If multiple combinations pass for a given event, then the scale factor combination is chosen as the (χ_1, χ_2) that takes $M'_{\tau\tau}$ closest to m_h. The optimum scale factor combinations passing the constraints are shown as contour plots in the χ_1-χ_2 plane for the G(500) signal in figure 9(a) and combined background samples in figure 9(b). In the signal distribution, a dense elliptical region centered close to $(\chi_1, \chi_2) = (1,1)$ is observed. This emphasises that the majority of signal events passing the constraints need little to no scaling to satisfy the 'ideal' kinematic constraints. This property is not limited to the G(500) signal; the samples corresponding to all other graviton mass values produce distributions similar to the ellipse in figure 9(a). Conversely, the background contour plot shows a more uniform spread of scale factor combinations, indicating there is no particularly preferred region of scale factor phase space.

Equation 20 defines an ellipse in the χ_1-χ_2 plane

$$\mathcal{E} = \frac{((\chi_1 - h)\cos A + (\chi_2 - k)\sin A)^2}{a^2} + \frac{((\chi_1 - h)\sin A - (\chi_2 - k)\cos A)^2}{b^2} = 1. \tag{20}$$

The parameters a and b are the semi-major and semi-minor axes respectively, (k,h) represent the (χ_1, χ_2) co-ordinates of the ellipse centre and A is the tilt angle of the ellipse with respect to the vertical axis. To match the shape of the ellipse in figure 9(a), the parameters were determined empirically to be: $a = 0.49$,
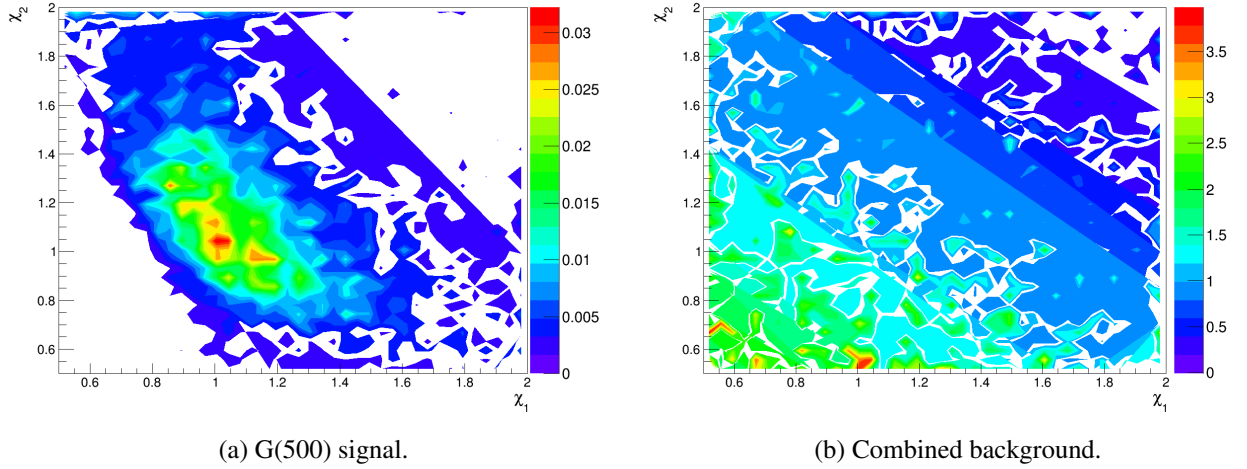
(a) G(500) signal.

(b) Combined background.

Figure 9: Contour plots showing the distribution of optimum (χ_1, χ_2) in scale factor space for events satisfying the kinematic constraints given in the text.

$b = 0.22$, $(k,h) = (1.05, 1.10)$ and $A = 0.65\pi$. Thus, for a given (χ_1, χ_2) combination, $\mathcal{E}$ quantifies a weighted distance from the 'ideal' signal position in scale factor phase space, the ellipse centre. This can be used as a parameter measuring the magnitude of scaling an event needs to look signal-like. In addition, looking at the background distribution in figure 9(b), the region towards the origin of the plot contains a significant proportion of events, with the probabiltiy density decreasing towards the upper right half of the plot. In contrast, the probabiltiy density distribution for the signal in figure 9(a) is such that events are less likely to lie on the same side of the ellipse centre to the origin than on the opposite side. Thus to increase the discriminating power, an additional sign is applied to $\mathcal{E}$, which is negative if the optimum (χ_1, χ_2) combination lies on the origin side of the ellipse major axis and is otherwise positive.

For any given event, there is typically a range of scale factor combinations χ_1, χ_2 that satisfy the set of constraints given above on the corrected variables. The choice of scale factor combination that takes $M'_{\tau\tau}$ closest to m_h is adequate for a first exploration, but is somewhat arbitrary. The next step is to define a method to choose the optimum combination of χ_1, χ_2 b-jet scaling factors based on a combined assessment of the corrected kinematic properties of an event.

3.4 Global optimisation: ψ

The algorithm was adapted to iterate only over (χ_1, χ_2) combinations with an ellipse parameter, $|\mathcal{E}| < 2$. Reducing the scale factor phase space in this manner better matches the expected uncertainty on jet energy reconstruction and should favour signal events given the emergence of the ellipse in figure 9(a). For each (χ_1, χ_2) with $|\mathcal{E}| < 2$, a penalty function ψ is calculated, which effectively quantifies the 'distance' a scaled event is from the 'ideal' characteristics outlined in section 3.1. To perform a global optimisation, the ψ function takes a combination of the corrected variables as input. The optimum scale factor combination for a given event is then chosen as the (χ_1, χ_2) corresponding to the *lowest* ψ score, effectively the scaling which most resembles an 'ideal' signal event. This approach discriminates between signal and background events as it is more likely to find a scaling scenario that can take a raw signal event closer to

the 'ideal' characteristics in all corrected variables, and thus signal events are expected to have smaller minimum ψ values than is the case for background events.

The ψ penalty function is split into four components: $\psi_{\tau\tau}$, ψ_{bb}, $\psi_{\not{E}_T}$ and $\psi_{\mathcal{E}}$, which quantify a distance from the 'ideal' in a number of different corrected parameters. All components include additional weights, σ, which enable each to contribute equally in the total ψ score. The weights were chosen 'by hand' such that the individual ψ component distributions are all similar in shape for a given signal sample (see figure 14). The individual ψ components are defined in the following subsections, where their design incorporates some of the main features in the raw variable distributions.

3.4.1 Invariant mass of the $\tau_h\tau_l$ system: $\psi_{\tau\tau}$

The component $\psi_{\tau\tau}$ calculates the squared distance between the corrected invariant mass of the $\tau_h\tau_l$ system and the Higgs mass. It is defined in equation 21

$$\psi_{\tau\tau} = \frac{(M'_{\tau\tau} - m_h)^2}{\sigma^2_{\tau\tau}},\tag{21}$$

where $\sigma^2_{\tau\tau}$ is chosen to give a score $\psi_{\tau\tau} = 10$ for $|M'_{\tau\tau} - m_h| = 25$ GeV: $\sigma^2_{\tau\tau} = 62.5$ GeV^{-2}.

3.4.2 Invariant mass of the b-tagged jet system: ψ_{bb}

The component ψ_{bb} is the squared distance between the corrected invariant mass of the b-jet system and m_h. It is defined in equation 22

$$\psi_{bb} = \frac{(M'_{bb} - m_h)^2}{\sigma^2_{bb}}.\tag{22}$$

The signal M_{bb} distributions in figure 7(a) are asymmetric about the Higgs mass, with a longer tail on the low mass side. To account for this, the ψ_{bb} score is defined to have a steeper gradient on the high mass side as it is less likely to find signal events in this region. Thus, for $M'_{bb} < m_h$, σ^2_{bb} is chosen to give $\psi_{bb} = 10$ at $M'_{bb} \approx 80$ GeV: $\sigma^2_{bb} = 200$ GeV^{-2}. For $M'_{bb} > m_h$, a score of $\psi_{bb} = 10$ corresponds to $M'_{bb} = 150$ GeV: $\sigma^2_{bb} = 62.5$ GeV^{-2}.

3.4.3 Direction of the missing transverse energy: $\psi_{\not{E}_T}$

$\psi_{\not{E}_T}$ incorporates the direction of the corrected missing energy into a component of ψ. The corrected parameter Ω' could be used to define the $\psi_{\not{E}_T}$ score for all situations. However, an asymmetry parameter A_1, defined by equation 23

$$A_1 = \frac{1}{2}\left[\frac{f_{\tau_l} - f_{\tau_h}}{f_{\tau_l} + f_{\tau_h}} + 1\right],\tag{23}$$

is found to have greater signal–background discrimination for the subset of events with $\not{E}'_T$ directed inside the $\tau_h\tau_l$ opening angle. The definition of A_1 maps onto the limits of Ω, such that $A_1 \in [0,1]$ where $A_1 = 0$ (1) corresponds to the $\not{E}_T$ directed along τ_h (τ_l). The parameter $f_{\tau_i} = p^T_{\nu_i}/p^T_{\tau_i}$ is the assigned ν transverse momentum along τ_i divided by the measured τ transverse momentum. A new variable, the '$\not{E}_T$ parameter' is constructed out of both Ω and A_1; it takes the value of Ω for events with the missing

energy outside of the $\tau_h\tau_l$ system and takes the value of A_1 when the missing energy is inside the $\tau_h\tau_l$ system. Figure 10 compares the raw Ω and raw $\not{E}_T$ parameter distributions for $\not{E}_T$ inside the $\tau_h\tau_l$ system; the greater signal–background discrimination for the $\not{E}_T$ parameter for this sub-set of events is clearly evident.
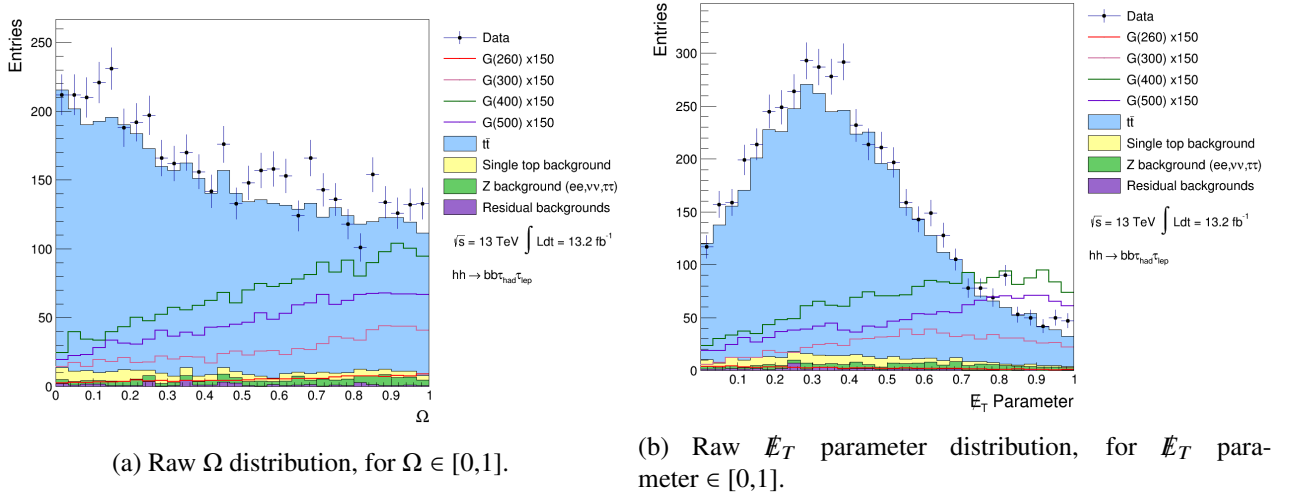

(a) Raw Ω distribution, for $\Omega \in [0,1]$.

(b) Raw $\not{E}_T$ parameter distribution, for $\not{E}_T$ parameter $\in [0,1]$.

Figure 10: Comparing the raw Ω and raw $\not{E}_T$ parameter distributions. It can be seen that the $\not{E}_T$ parameter has greater discriminating power for the subset of events with $\not{E}_T$ inside the $\tau_h\tau_l$ system.

The distribution of the $\not{E}_T$ parameter over the entire allowed range is shown in figure 11. The contribution, $\psi_{\not{E}_T}$, of the $\not{E}_T$ to the penalty function ψ is defined separately in different regions of the $\not{E}_T$ parameter:

- The most likely region to find the missing energy in signal events is directed inside the $\tau_h\tau_l$ opening angle and towards the τ_l. Thus, for corrected $\not{E}_T$ parameter $\in [0.8,1]$, $\psi_{\not{E}_T} = 0$.

- $\psi_{\not{E}_T}$ is defined to inherit a score of 8 for when the corrected missing energy is directed along τ_h. Equation 24 is used to calculate $\psi_{\not{E}_T}$ for corrected $\not{E}_T$ parameter $\in [0,0.8]$, such that $\sigma^2_{\not{E}_T,1} = 0.08$

$$\psi_{\not{E}_T} = \frac{(A_1' - 0.8)^2}{\sigma^2_{\not{E}_T,1}}. \tag{24}$$

- Signal events are not expected to have the corrected missing energy outside of the $\tau_h\tau_l$ system, and thus a high score must be given to this situation. Figure 11 shows that for signal events with the missing energy outside $\tau_h\tau_l$, it is more likely that the $\not{E}_T$ is closer to τ_l. Therefore, the $\psi_{\not{E}_T}$ component requires a shallower gradient on the τ_l side. For corrected $\not{E}_T$ parameter > 1 (on the τ_l side), $\psi_{\not{E}_T}$ is defined by equation 25

$$\psi_{\not{E}_T} = \frac{(\Omega' - 1)^2}{\sigma^2_{\not{E}_T,2}}. \tag{25}$$

The weighting $\sigma^2_{\not{E}_T,2} = 0.005$ corresponds to $\psi_{\not{E}_T} = 8$ at corrected $\not{E}_T$ parameter $= 1.2$.
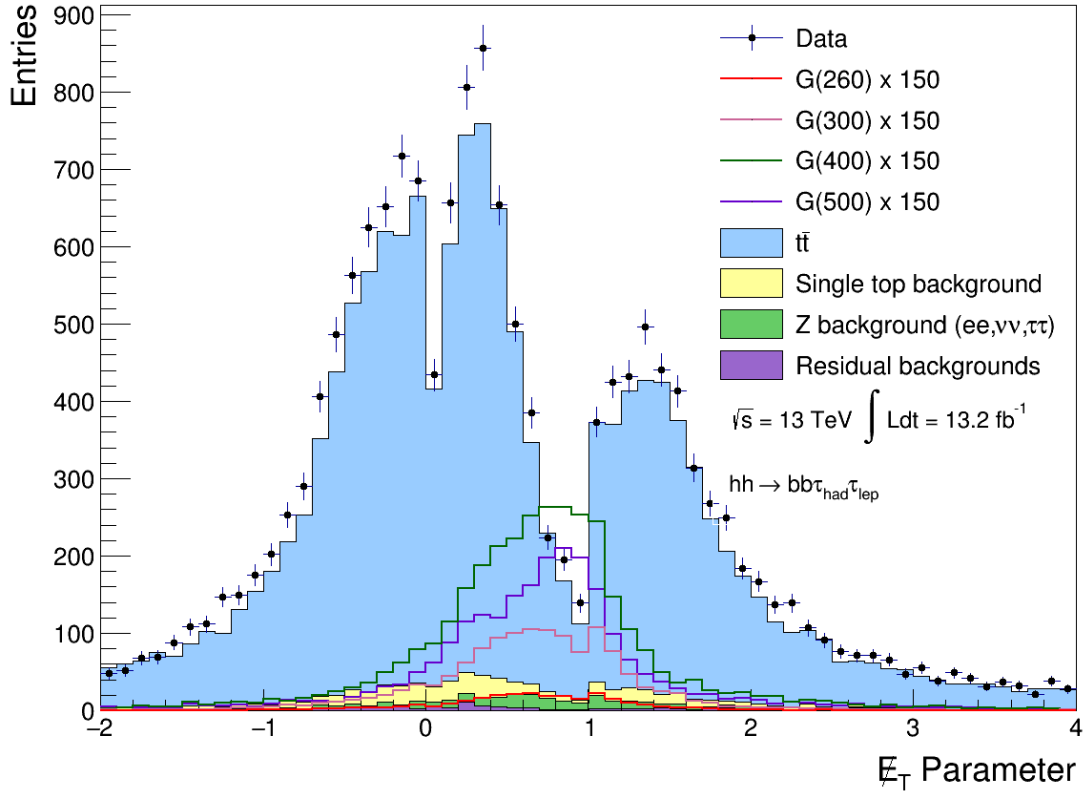
Figure 11: Raw $\not{E}_T$ parameter distribution.

- Equation 26 is used to calculate $\psi_{\not{E}_T}$ for a corrected $\not{E}_T$ parameter < 0 (on the τ_h side)

$$\psi_{\not{E}_T} = 8 + \frac{|\Omega'|^2}{\sigma^2_{\not{E}_T,3}}. \qquad (26)$$

This is the least likely signal scenario and thus corresponds to the highest score. The definition ensures $\psi_{\not{E}_T}$ is a continuous variable across the whole $\not{E}_T$ parameter range. A weighting of $\sigma^2_{\not{E}_T,3} = 0.00125$ is chosen to give a steeper gradient for this scenario.

3.4.4 Distance of (χ_1,χ_2) from the ellipse centre: $\psi_{\mathcal{E}}$

The parameter $\psi_{\mathcal{E}}$ introduces a component into the penalty function ψ that quantifies the distance a scale factor combination χ_1,χ_2 is from the ellipse centre ($\mathcal{E} = 0$). Figure 12 presents the $\mathcal{E}$ distribution for both signal and background samples obtained after the optimization procedure described in section 3.4.5. Note, values are limited to the range $\mathcal{E} \in$ [-2,2], since the algorithm iterates only over scale factors inside the ellipse, $|\mathcal{E}| < 2$. As expected, the signal spectra peak closer to the ellipse centre, whereas the combined background distribution is more uniform across the whole range. In addition, the signal spectra are defined by a longer tail on the $\mathcal{E} > 0$ side, echoing the finding in section 3.3 based on figure 9.

379   Therefore, a steeper scoring gradient is defined for $\mathcal{E} < 0$. Equation 27 is used to define $\psi_\mathcal{E}$

$$\psi_\mathcal{E} = \frac{\mathcal{E}^2}{\sigma_\mathcal{E}^{2\,\pm}}, \tag{27}$$

380   where $\sigma_\mathcal{E}^{2\,+} = 0.2$ and $\sigma_\mathcal{E}^{2\,-} = 0.0845$ represent the weightings for $\mathcal{E} > 0$ and $\mathcal{E} < 0$ respectively.

381   In figure 12, spikes are seen at the values $\mathcal{E} = \pm 2$, corresponding to events whose optimum scaling
382   combination (χ_1, χ_2) lies at the ellipse edge. The effect is much more pronounced in the background
383   distribution, although a small number of signal events behave also in this way. This reflects the fact that,
384   within the limits placed on the factors by which the b-jet energies are allowed to be re-scaled, background
385   processes yield events that are rather unlikely to be kinematically consistent with the signal hypothesis.
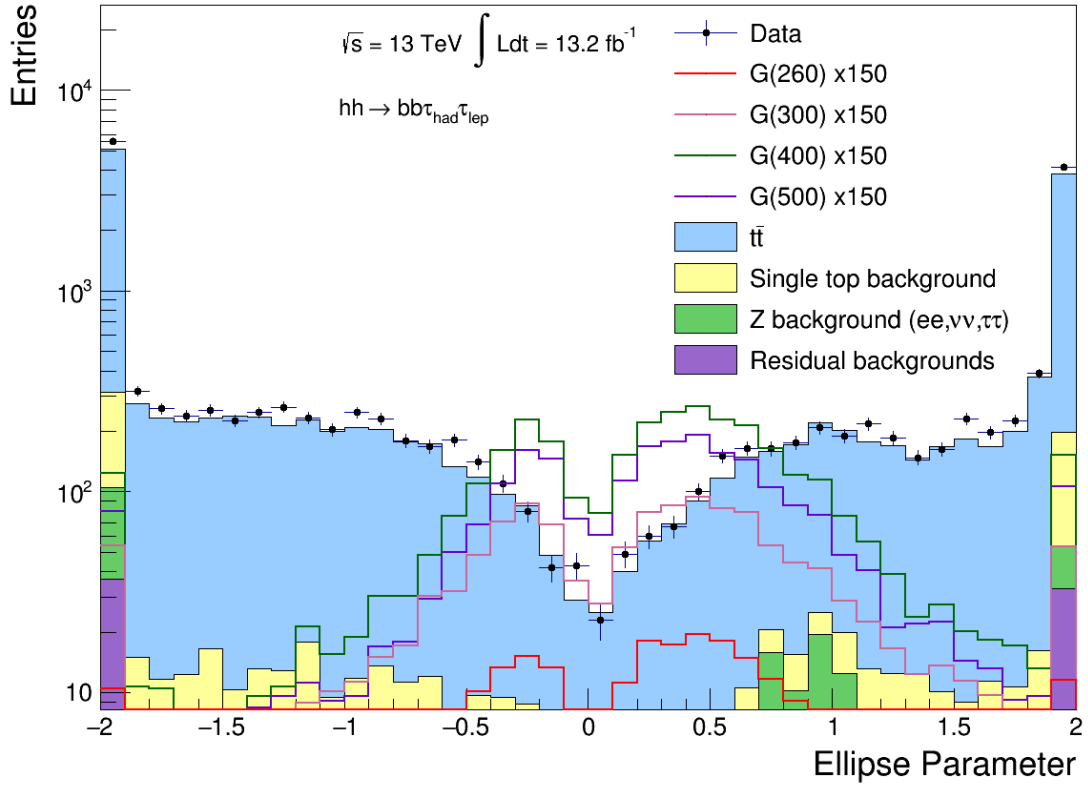
Figure 12: $\mathcal{E}$ parameter distribution after global optimisation.

386   **3.4.5 Total ψ**

387   The total penalty function is defined as the quadrature sum of the individual components, shown in equa-
388   tion 28

$$\psi = \sqrt{\psi_{\tau\tau}^2 + \psi_{bb}^2 + \psi_{\not{E}_T}^2 + \psi_\mathcal{E}^2}. \tag{28}$$

389 After calculating a ψ value for each (χ_1, χ_2) combination within the ellipse, the algorithm chooses the
390 optimum scaling for a given event as the χ_1, χ_2 combination that corresponds to the lowest possible value
391 of ψ. Note that a score of $\psi = 0$ corresponds to the 'ideal' signal kinematics. If no (χ_1, χ_2) combination
392 with $|\mathcal{E}| < 2$ exists which can take the corrected missing energy within 90° of the closest τ, then the
393 reconstruction methods outlined in 3.2 are invalid and the event is removed as background.

The distribution of the minimised ψ values is presented in figure 13. The shapes of the signal distribution
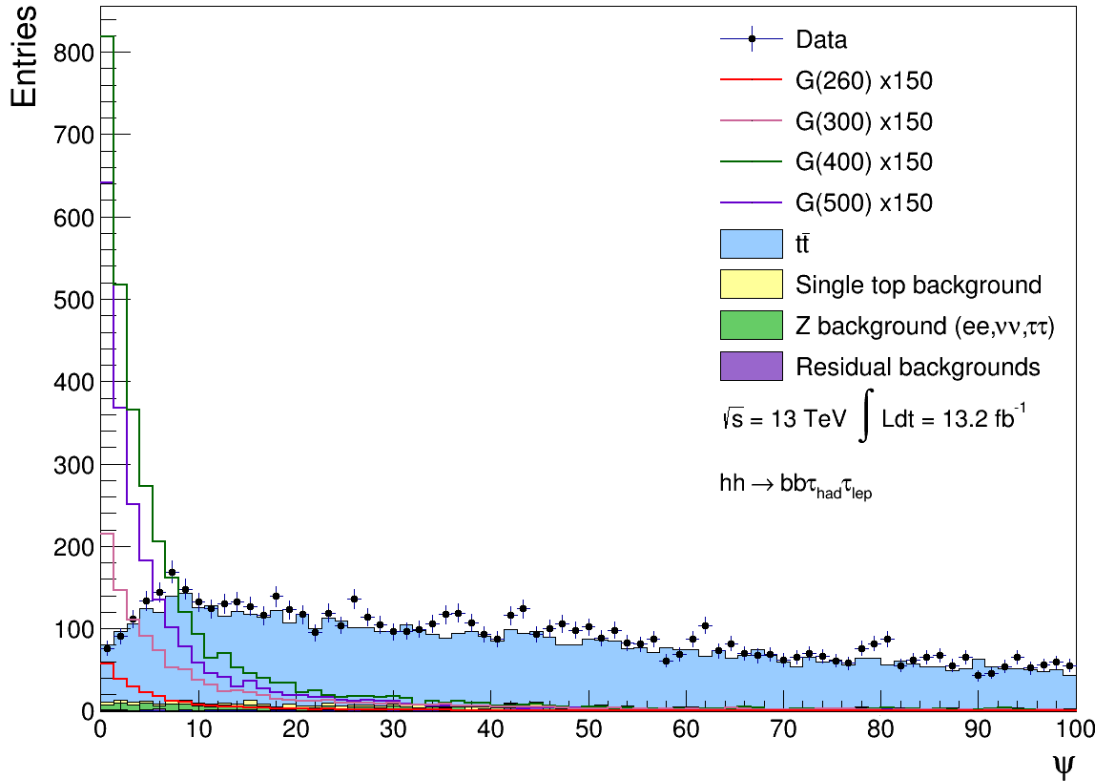
Figure 13: Total ψ distribution. All parameters are calculated using the optimum (χ_1, χ_2) combination for a given event.

394
395 in ψ are significantly different to those of the combined background, emphasising the strong discrimin-
396 ating power available in this variable. The signal distributions show a sharp peak at low ψ, decaying to
397 very few events with $\psi > 20$. In constrast, the background distribution is almost flat across the whole ψ
398 range plotted in figure 13.

399 The distributions for the individual components of ψ are shown in figure 14. It can be noted that for any
400 given signal sample the distributions for each individual component of the ψ function look similar. This
401 enables each component to contribute significantly to ψ and has been a factor in the choice of the values
402 of the weights 'σ_i^2' in section 3.4.

403 In the $\psi_{\not{E}_T}$ distribution of figure 14(c), a spike at around $\psi_{\not{E}_T} \leq 8$ is apparent in the distribution of
404 background events. This occurs due to the discontinuity in the gradient of $\psi_{\not{E}_T}$ with Ω' at the point
405 $\Omega' = 0$. Negative values of Ω' are particularly heavily penalised in the definition of $\psi_{\not{E}_T}$ (equation 26),
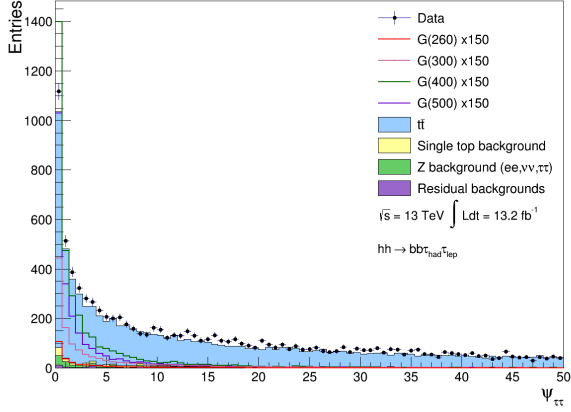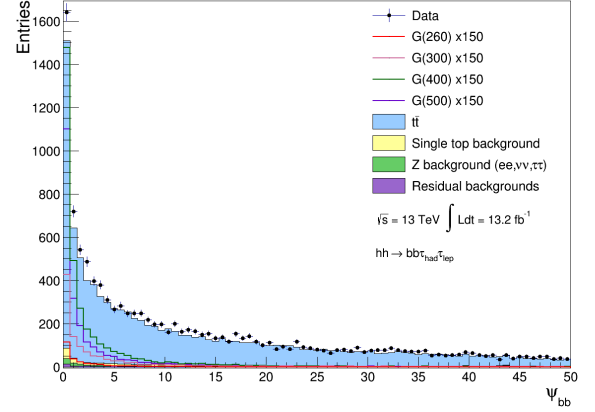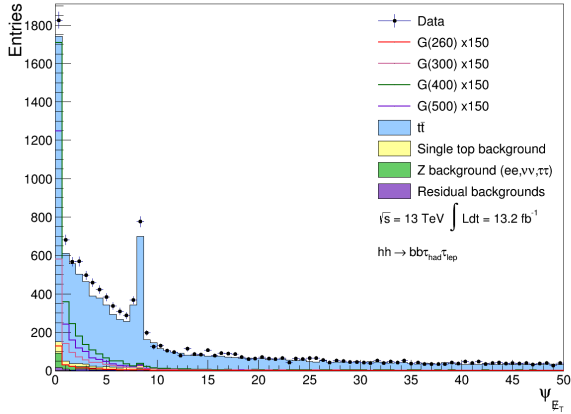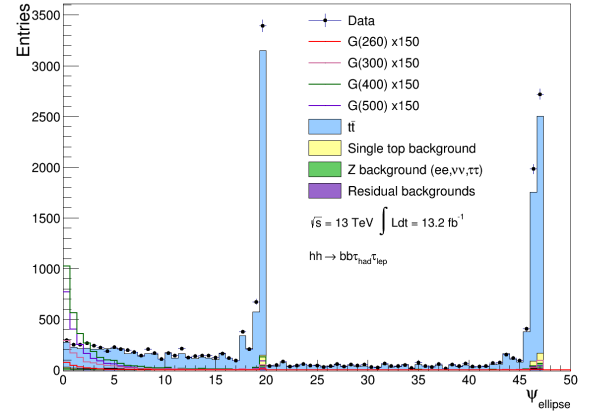
(a) $\psi_{\tau\tau}$ distribution.

(b) ψ_{bb} distribution.

(c) $\psi_{\not{E}_T}$ distribution.

(d) $\psi_{\mathcal{E}}$ distribution.

Figure 14: Distributions of the individual ψ components. All parameters are calculated using the optimum (χ_1, χ_2) combination for a given event.

and this results in something of a bias in the optimisation for events to be assigned values of Ω' at zero or slightly above. In the $\psi_{\mathcal{E}}$ distribution of figure 14(d), two spikes appear at $\psi_{\mathcal{E}} \approx 20$ and $\psi_{\mathcal{E}} \approx 50$. As can be verified by re-arranging equation 27, the positions of these spikes correspond to the spikes in events with optimized values of χ_1, χ_2 on the ellipse boundary at $\mathcal{E} = 2$ and $\mathcal{E} = -2$ respectively, shown in figure 12. It can be seen in figure 13 that any spikes in the individual components of ψ are largely washed out in the overall distribution of ψ. In addition, looking at the overall ψ distribution in figure 13, it can be anticipated that the optimum cut on this variable would be somewhere in the range $5 < \psi < 20$. Therefore, the artifact at the ellipse boundary will not affect the final signal–background discrimination.

One of the main features of this analysis is that the input parameters of ψ: $M'_{\tau\tau}$, M'_{bb}, the corrected $\not{E}_T$ parameter and the ellipse parameter $\mathcal{E}$ have little to no dependence on the Higgs boson pair (resonance) mass. Therefore, ψ itself is independent of the energy scale of the Higgs pair production search. Figure 15 presents the fraction of events in percent remaining for each Monte Carlo sample as a function of a cut on ψ. All signal spectra follow roughly the same curve, highlighting the **scale invariant** nature of ψ. This result is useful in allowing the global optimisation analysis to be applied for all graviton masses and
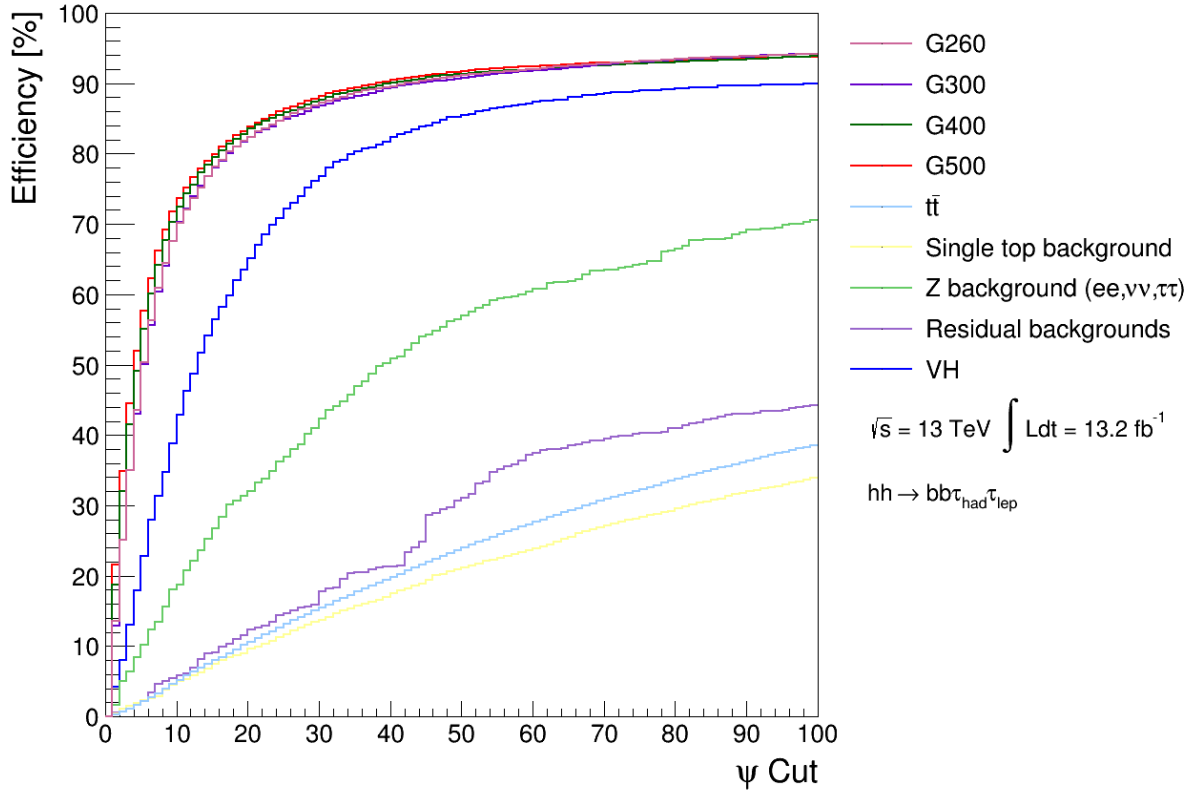
Figure 15: Percentage of events remaining in each sample as a function of a cut on the ψ variable.

for standard model di-Higgs events. It is important at this stage to stress that the $M'_{bb\tau\tau}$ variable has not been used in the ψ analysis. Thus, there is significantly more discriminating power available if we wish to localise our search to a specific range of $M'_{bb\tau\tau}$ values.

4 Results

This section demonstrates the performance of the ψ analysis through its impact on the 2-body and 4-body masses along with the corrected orientation of the missing energy (shown via the $\not{E}_T$ parameter). An initial demonstration of the performance using the 13.2 fb^{-1} data set is presented in section 4.1. The analysis is then extended to the full currently available (2015–2016) run 2 ATLAS data and Monte Carlo samples corresponding to an integrated luminosity of 36.1 fb^{-1}. In section 4.2 a comparison of the current ATLAS BDT performance with and without the ψ analysis is presented. In light of the excellent performance of ψ as a scaling variable, its application to SM di-Higgs searches is explored in section 4.3.

4.1 The discriminating power of ψ using the 13.2 fb^{-1} samples

Figure 16 shows the corrected variable distributions after global optimisation is performed. Note, the distributions are produced with respect to the optimum (χ_1, χ_2) combination for each event, which gives the lowest ψ score. The $M'_{\tau\tau}$ and M'_{bb} signal spectra in figure 16 present sharper peaks compared to the respective raw distributions in figure 7, demonstrating the effect of constraining the system to the Higgs mass. The corrected four-body invariant mass signal spectra in figure 16d are also characterised by narrower peaks than in the raw $M_{bb\tau\tau}$ distribution, and the peak positions are shifted closer towards the graviton mass used to generate the different samples. This behaviour confirms the physical argument used to motivate the analysis method, as without constraining the four-body mass, the distributions have been amended towards their expected characteristics. The corrected $M'_{bb\tau\tau}$ distribution with a cut on ψ at: $\psi < 5$, is presented in figure 17. Comparing to figure 16d, significant background rejection is observed whilst maintaining a strong signal efficiency. At this point it is possible to localise the search to a specific mass window by cutting on the $M'_{bb\tau\tau}$ variable.
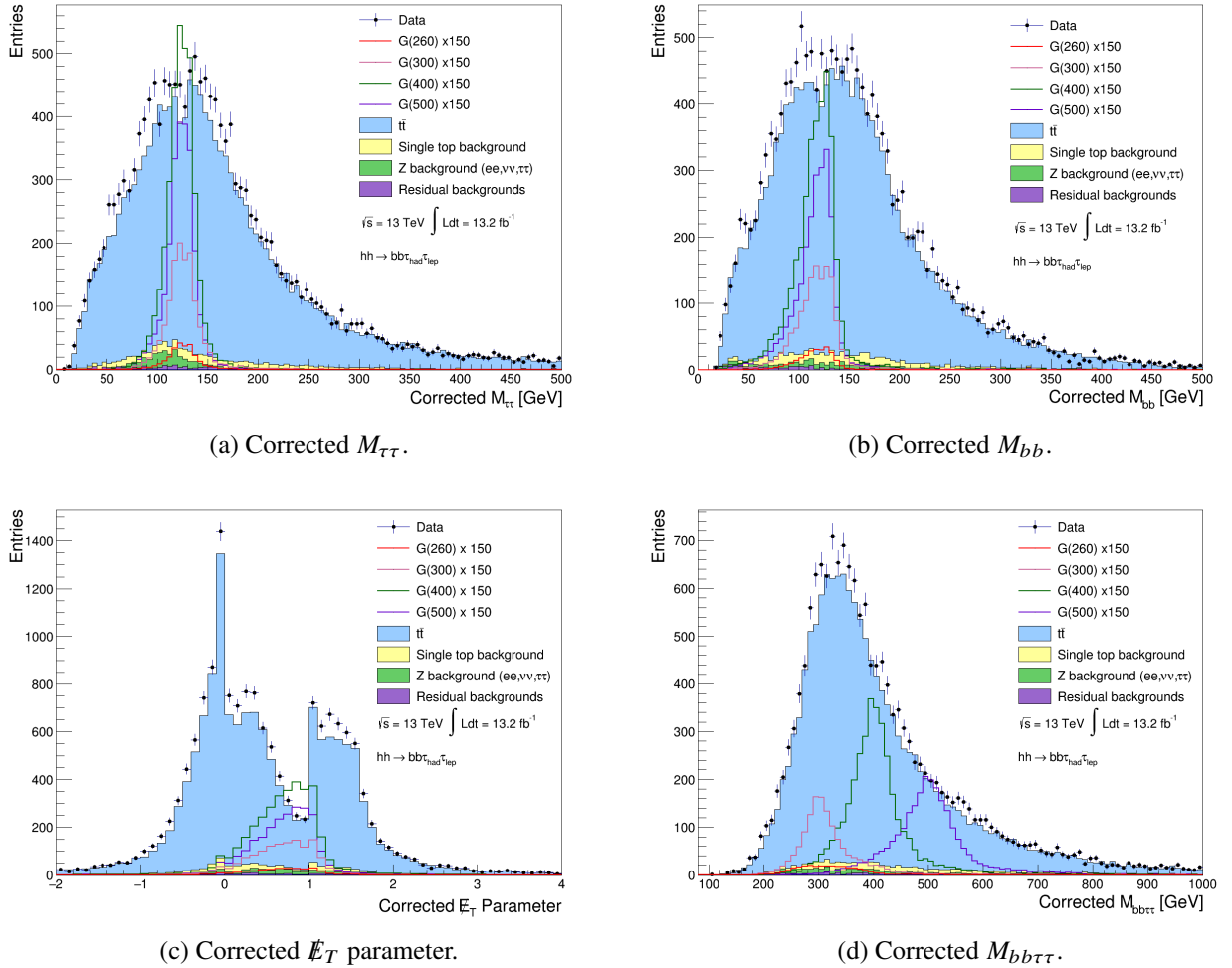

(a) Corrected $M_{\tau\tau}$.

(b) Corrected M_{bb}.

(c) Corrected $\not{E}_T$ parameter.

(d) Corrected $M_{bb\tau\tau}$.

Figure 16: Corrected variable distributions after global optimisation, produced with respect to the optimum (χ_1, χ_2) combination for each event.
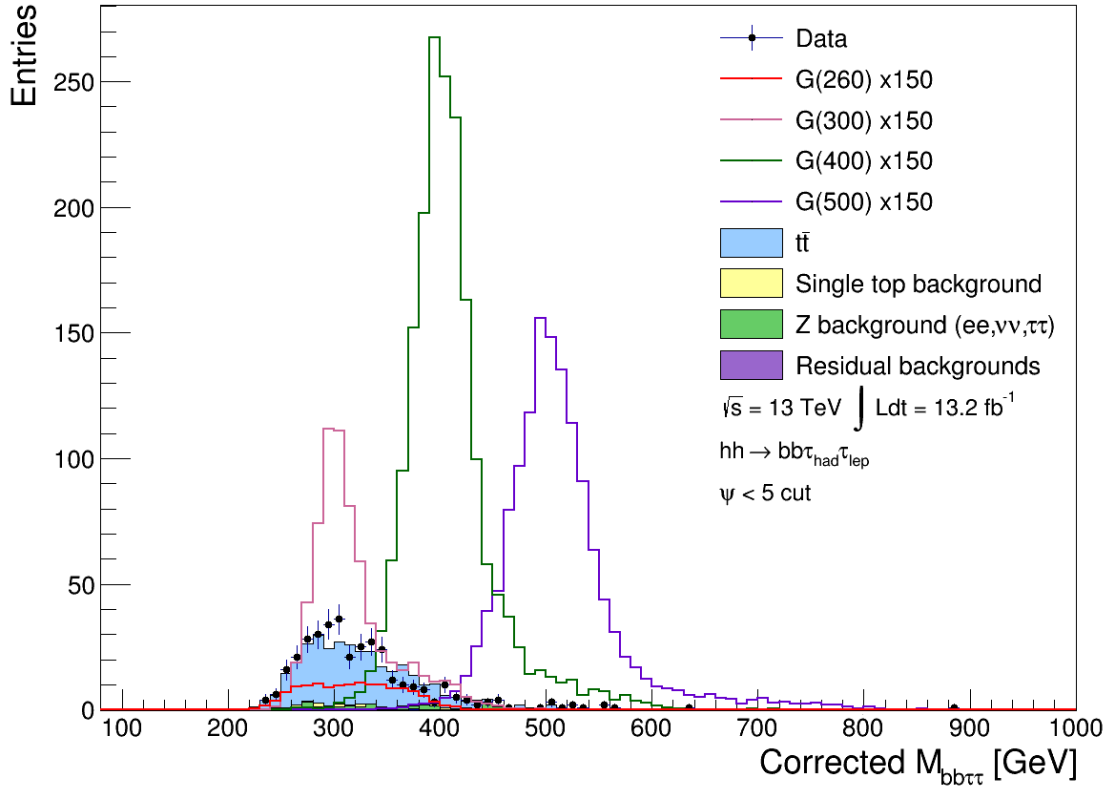
Figure 17: Corrected $M_{bb\tau\tau}$ distribution after a cut on the ψ scoring function is made at: $\psi < 5$.

Cut-and-count significance plots were produced for each signal sample as a function of a cut on $M'_{bb\tau\tau}$. This involves incrementing a cut on the corrected four-body mass between 0 GeV and 1 TeV in 100 iterations, such that events with $M'_{bb\tau\tau}$ less than the cut are removed. For each mass cut value the number of remaining signal and background events is tallied and the significance is evaluated using equation 29

$$\sigma = \frac{N_S}{\sqrt{N_S + N_B}}, \tag{29}$$

where N_S and N_B are the remaining number of signal and background events respectively. Note, this is performed separately for each mass sample. Figure 18 presents the cut-and-count significance plots for the G(400) and G(500) signal samples. Although the plots show ψ to have excellent signal–background discrimination, it is imperative to show that the variable holds additional discriminating power, which will build upon the current ATLAS search method. The following section details a comparison to the ATLAS BDT analysis, in which the effectiveness of ψ after the multi-variate approach is scrutinised.

4.2 Comparison to the ATLAS BDT analysis

In this section, results are presented that demonstrate that even after a tight requirement on the $bb\tau_h\tau_l$ BDT score [12] is imposed the ψ variable provides considerable additional discriminating power against background.

(a) Produced with respect to the G(400) signal sample.

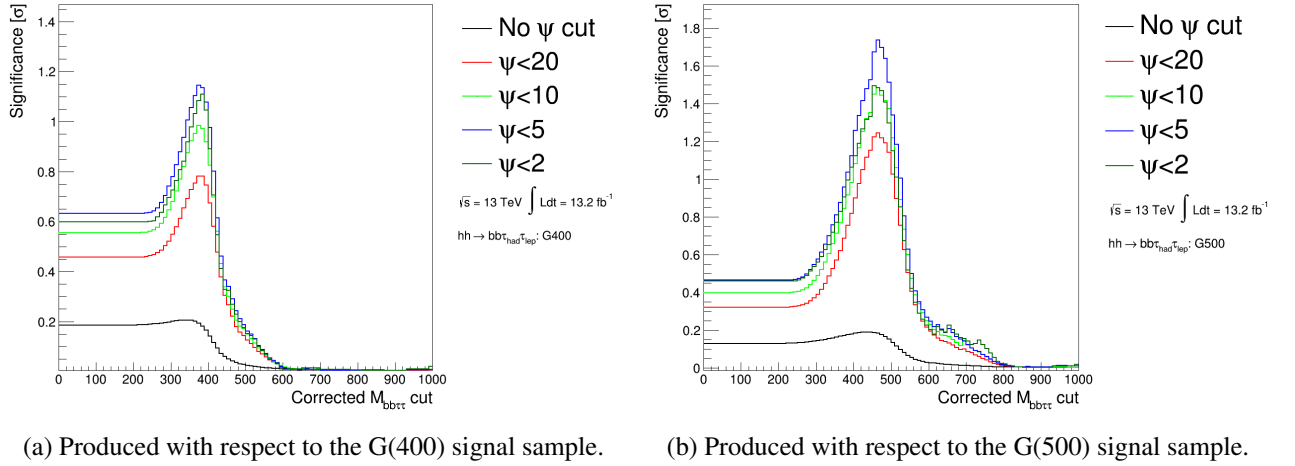(b) Produced with respect to the G(500) signal sample.

Figure 18: Cut-and-count significance plots as a function of a cut on the corrected four-body mass, $M'_{bb\tau\tau}$.

At this point, the analysis moved to the full ATLAS run 2 36.1 fb^{-1} data and Monte Carlo samples. Figure 19 shows the ψ variable distribution for the larger dataset. The data is not blinded in this distribution since without any restriction on the corrected four-body mass, $M'_{bb\tau\tau}$, the number of background events is significantly greater than those predicted by the Monte Carlo signal samples, and thus it would be unlikely to see any hint of a signal in the data. Excellent agreement between data and SM background Monte Carlo is observed over the entire range of ψ.

The distributions with the 36 fb^{-1} dataset of the individual ψ components: $\psi_{\tau\tau}$, ψ_{bb}, $\psi_{\not{E}_T}$ and $\psi_{\mathcal{E}}$, are presented in figure 20. Excellent agreement between data and SM background Monte Carlo is again observed over the entire range of each component of ψ.

Distributions of the BDT score, trained on the G(300) and G(400) signal samples, are given in figures 21(a) and 21(c) respectively. Note, the discrimination between signal and background decreases for lower masses due to increasing overlap between signal and background-like characteristics in the input parameters. That is, it is more difficult to define a set of cuts which can distinguish between signal and background for lower graviton masses. Figures 21(b) and 21(d) show the ψ distributions for the G(300) and G(400) signal on top of background, where a cut on the respective BDT score is imposed. The BDT cuts are highlighted by the dashed lines in the BDT score distributions. It can be seen that a significant number of background events that are given a signal-like BDT score by the ATLAS analysis, are characterized by a high ψ score. Thus, imposing an additional requirement on ψ can significantly improve background rejection. Note, data points cannot be shown on these distributions as the ATLAS group are keeping the full dataset blinded with respect to the BDT score distributions. Hence, any other variable distributions with similiar signal–background discrimination must also be kept blinded. The distributions of BDT score and ψ after a cut on the BDT corresponding to G(260) and G(500) are shown in Appendix 1.

Cut-and-count significance plots were again produced for each signal sample, but now as a function of a cut on the respective BDT score. The BDT score cut was incremented between -1 and +1 in 100 iterations, and a significance was calculated for each cut value using equation 29. The analysis is then repeated with different requirements on the ψ variable. Cut-and-count plots for the G(300) and G(400) samples are
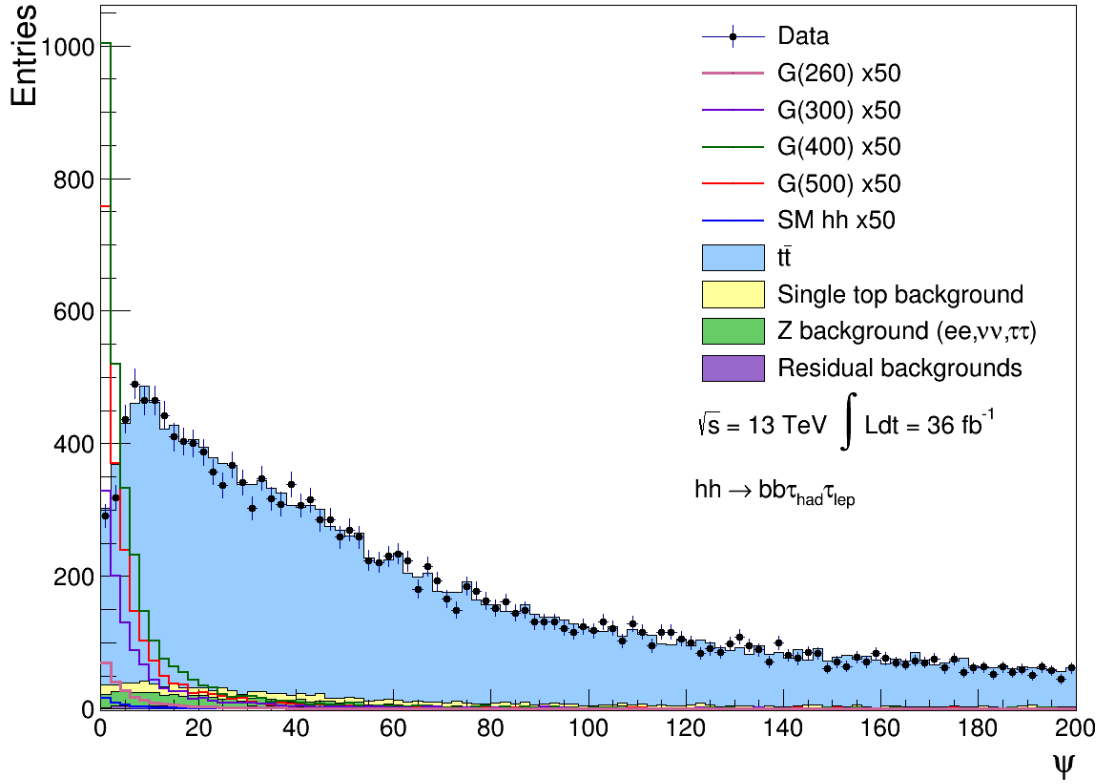
Figure 19: ψ distribution with the full 36 fb^{-1} dataset.

presented in figures 22(a) and 22(b) respectively. Comparing the (black) curves with no ψ requirement, to the optimum ψ cut of $\psi < 5$ (dark green), a huge improvement in terms of peak significance is observed. With an additional ψ requirement, the G(300) peak significance increases by approximately a factor of two, and the G(400) peak significance is taken from 2σ to approximately 2.7σ.

A standard method for presenting discrimination power is in the form of ROC curves. For a given BDT score requirement, the signal efficiency, ϵ, and background rejection, R, are calculated using equations 30 and 31 respectively,

$$\epsilon = \frac{S_{remaining}}{S_{total}}, \tag{30}$$

$$R = 1 - \frac{B_{remaining}}{B_{total}}, \tag{31}$$

where the variables $S_{remaining}$ ($B_{remaining}$) and S_{total} (B_{total}) are the remaining number and total number of signal (background) events respectively. Curves were then produced by plotting signal efficiency versus background rejection for 500 different BDT score requirements between -1 and +1. The analysis was repeated with additional ψ requirements imposed. ROC curves for the G(300) and G(400)
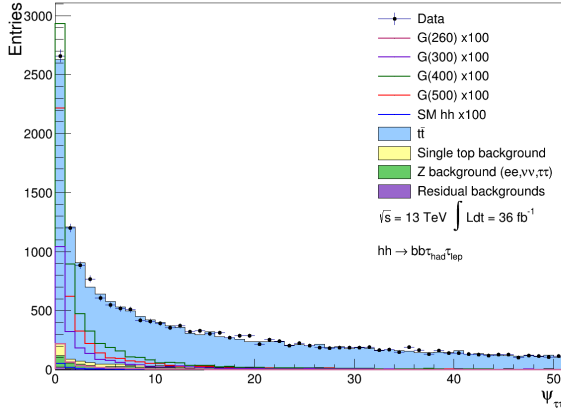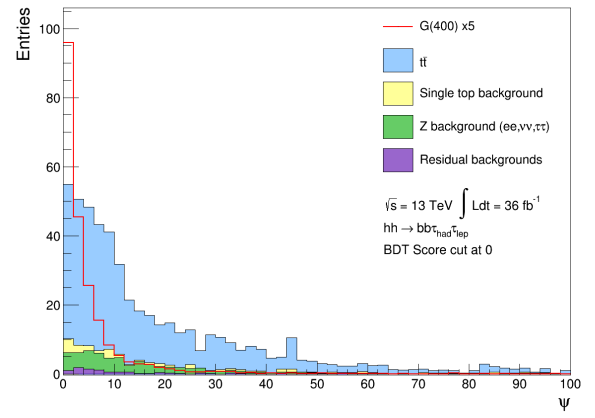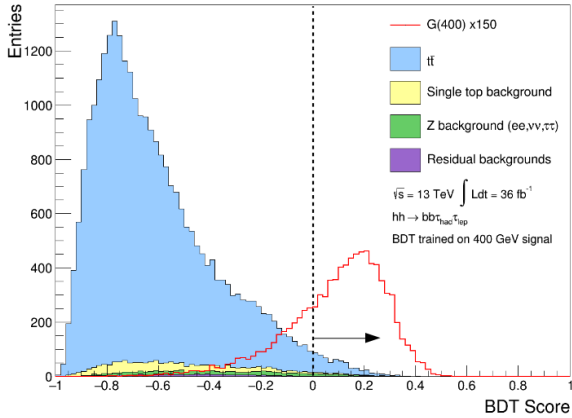
(a) $\psi_{\tau\tau}$ distribution.
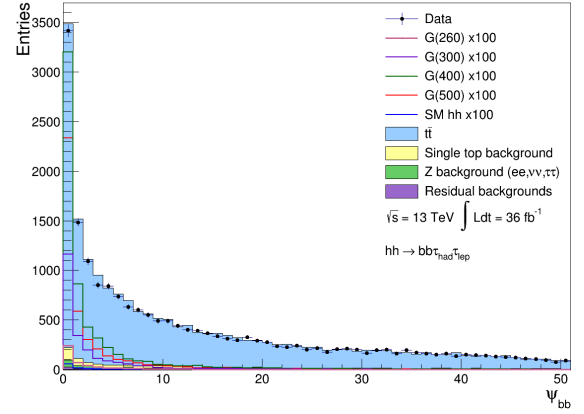
(b) ψ_{bb} distribution.
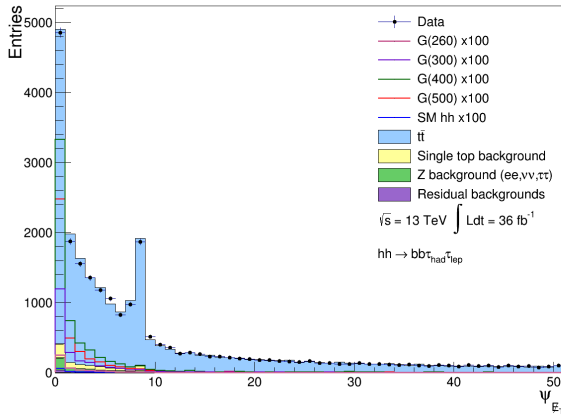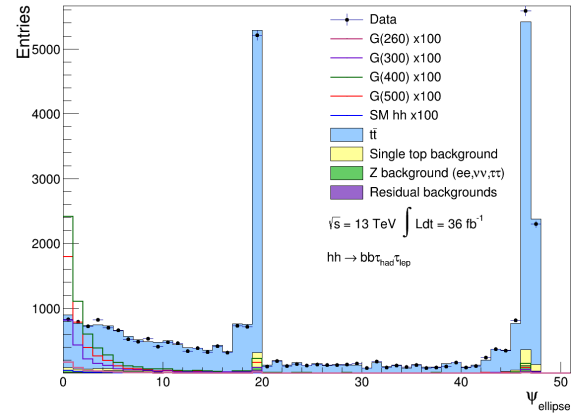
(c) $\psi_{\not{E}_T}$ distribution.

(d) $\psi_{\mathcal{E}}$ distribution.

Figure 20: A plot of the individual ψ components with the full 36 fb^{-1} samples.

samples are shown in figures 23(a) and 23(b) respectively. In figure 23(a) it can be seen for the G(300) signal that over a wide range of signal efficiencies an additional ψ requirement can reduce the level of selected background by a factor of approximately two at constant signal efficiency. For example, at a signal efficiency of around 62% the background rejection can be improved from approximately 0.95 to 0.975. Alternatively, at a signal efficiency of around 75% the background rejection can be improved from approximately 0.90 to nearly 0.95. Similarly, a ψ requirement on the G(400) sample improves background rejection significantly for a given signal efficiency. The cut-and-count significance plots and the ROC curves both emphasise the success of the global optimisation technique, where ψ is shown to have additional discriminating power with respect to the BDT score variable. Equivalent cut-and-count significance plots and ROC curves for the G(260) and G(500) signal samples are shown in Appendix 2. Note, the relative effectiveness of the ψ variable increases with lower graviton masses since the signal–background discrimination in the BDT score decreases.
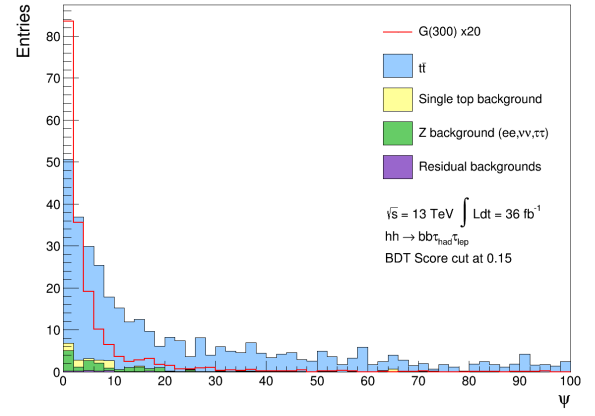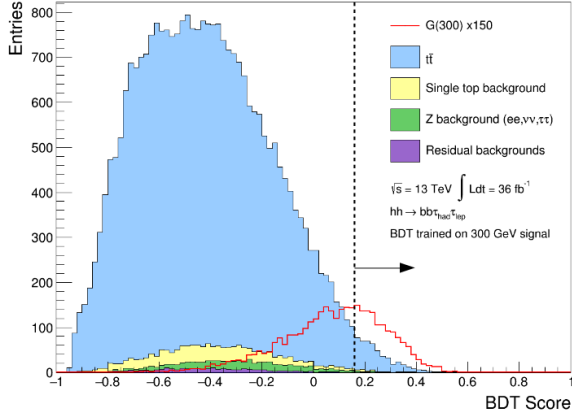
(a) BDT score distribution for training on the G(300) sample. The dashed line respesents a cut at 0.15, where the arrow signifies events which are kept.

(b) G(300) signal ψ distribution on top of background, where a cut to remove events with BDT score < 0.15 is imposed.


(c) BDT score distribution for training on the G(400) sample. The dashed line respesents a cut at 0, where the arrow signifies events which are kept.

(d) G(400) signal ψ distribution on top of background, where a cut to remove events with BDT score < 0 is imposed.

Figure 21: BDT score distributions for training on the G(300) and G(400) samples separately. ψ distributions are presented next to the BDT score plots, for the respective signal samples. A cut on the BDT score is imposed to produce the ψ distributions, identified by the dashed lines. Data is blinded from these plots due to high signal–background discrimination in the variables.
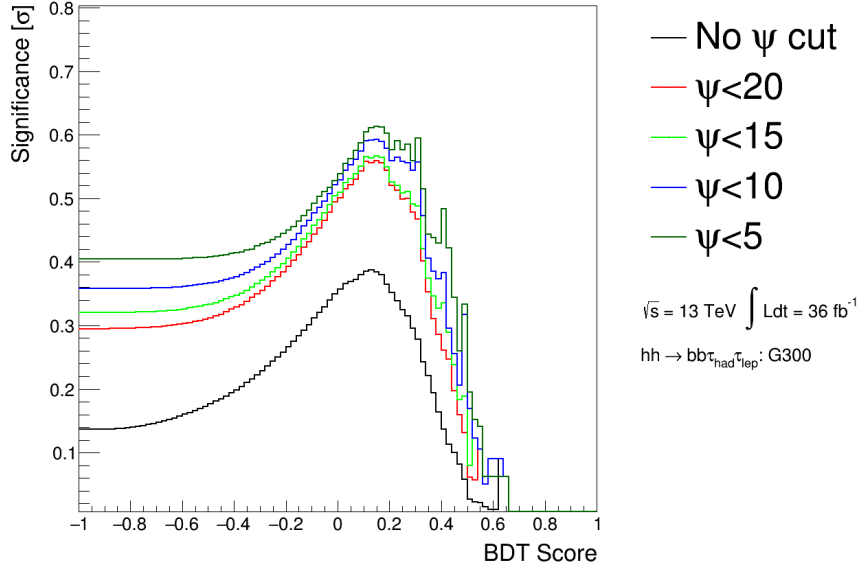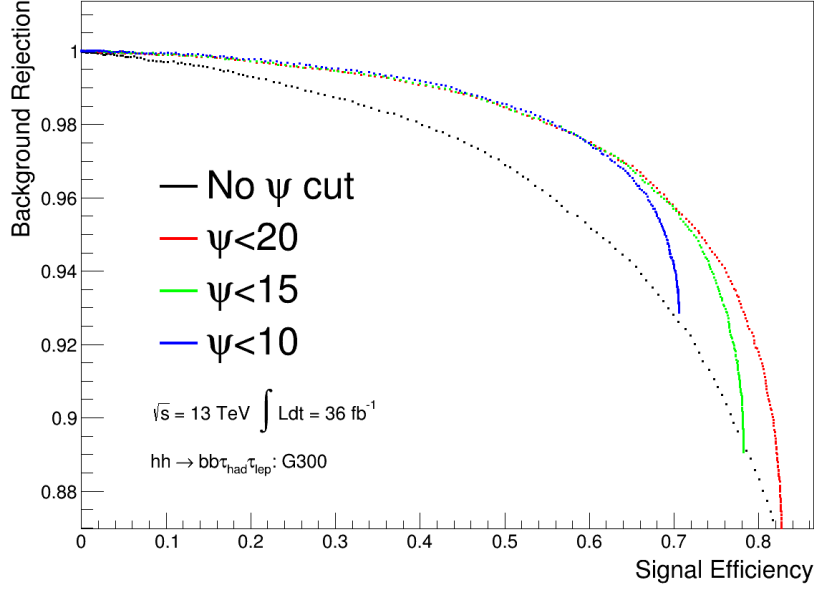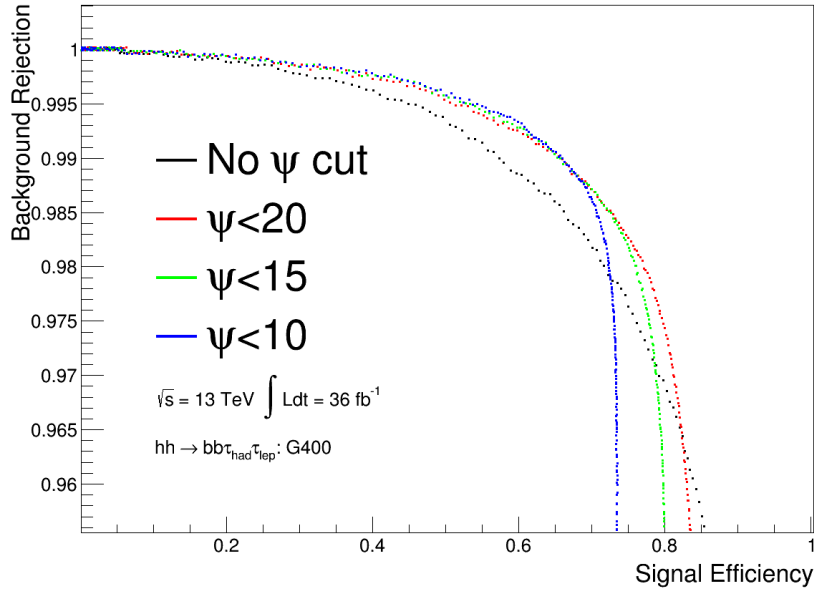
(a) G(300) distribution.
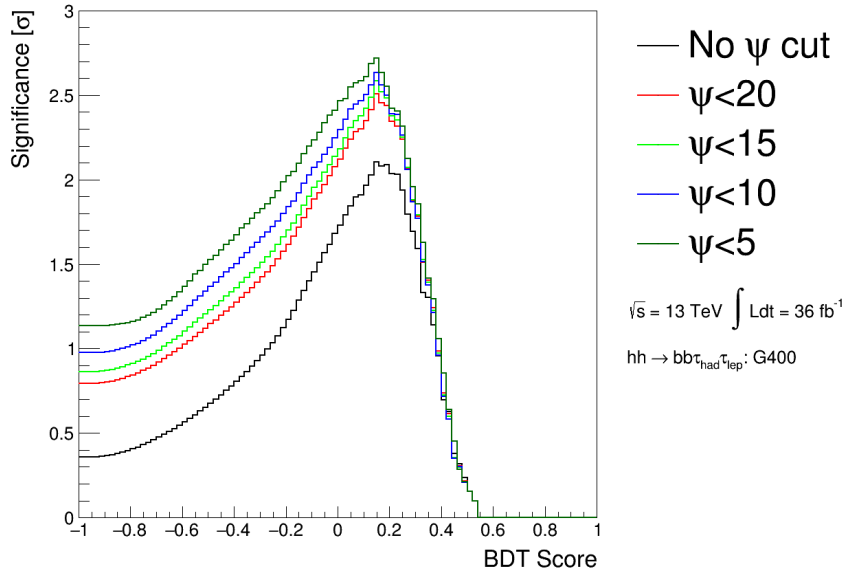
(b) G(400) distribution.

Figure 22: Cut-and-count significance plots as a function of a cut on the BDT score, trained on the G(300) and G(400) samples. The coloured curves have been created with an additional requirement on ψ.

(a) BDT score trained on the G(300) sample.

(b) BDT score trained on the G(400) sample.

Figure 23: ROC curves for the G(300) and G(400) samples. Created by plotting signal efficiency versus background rejection for 500 different BDT requirements. The coloured curves were produced with additional requirements on ψ.

4.3 Application of ψ to non-resonant di-Higgs production searches

Non-resonant Higgs pair production has no well defined four-body mass spectrum and therefore the search cannot be localised to a specific mass window. Therefore, as ψ is a scale invariant parameter, it has huge potential for application in SM di-Higgs production searches. Figure 24(a) shows the BDT score distribution for training on the SM $hh \rightarrow bb\tau_h\tau_l$ Monte Carlo sample. The overlap between signal and background in this plot stresses the difficulty in distinguishing between the two, even with a multi-variate analysis approach. It is important to note the large signal multiplier, which highlights the extremely small SM di-Higgs production cross-section. The dashed line represents a cut to remove events with BDT score < 0.15. This cut is imposed on the SM signal ψ distribution, shown in figure 24(b). As before, a large proportion of background passing the BDT cut inherits a high ψ score.

(a) SM BDT score distribution. (b) SM ψ distribution with a cut at BDT score < 0.15.
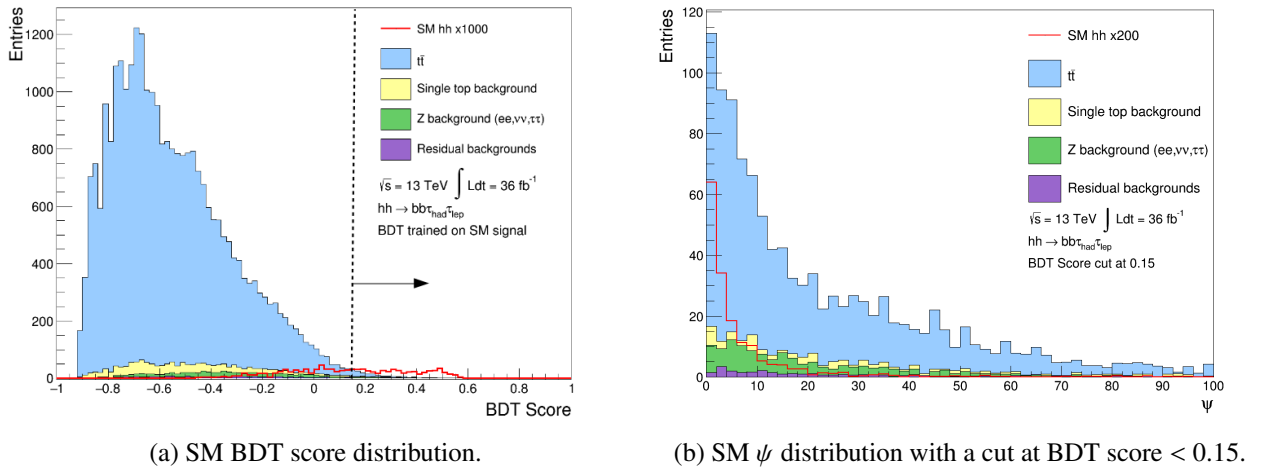
Figure 24: BDT score distribution for the BDT trained on the SM signal sample. The ψ distribution for this signal is also shown, where a cut to remove events with BDT score < 0.15 is imposed. The cut is highlighted by the dashed line in the score distribution.

Figure 25(a) displays the cut-and-count significance plot and Figure 25(b) the ROC curves produced with respect to the SM signal sample, implementing the same techniques described in section 4.2. Again, vast improvements in peak significance and background rejection are observed in figures 25(a) and 25(b) respectively, when additional ψ requirements are imposed. However, the absolute peak significance is much smaller compared to the resonant di-Higgs production samples because of the small SM cross-section. All in all, the application of ψ to SM di-Higgs production can significantly improve upon background rejection in a notoriously difficult search channel.

5 Summary and Conclusions

In this note we have presented a novel approach to the search for the process $hh \rightarrow bb\tau_h\tau_l$. Events containing two b-jets and two candidate τ-lepton decays are tested for their kinematic consistency with the hypothesis $hh \rightarrow bb\tau_h\tau_l$. In this method the measured b-jet energies are re-scaled, subject to constraints on the resulting corrected masses of the bb and $\tau_h\tau_l$ systems, on the magnitude and direction of the corrected $\not{E}_T$, and on the factors by which the b-jet energies are re-scaled. Having re-scaled the b-jet
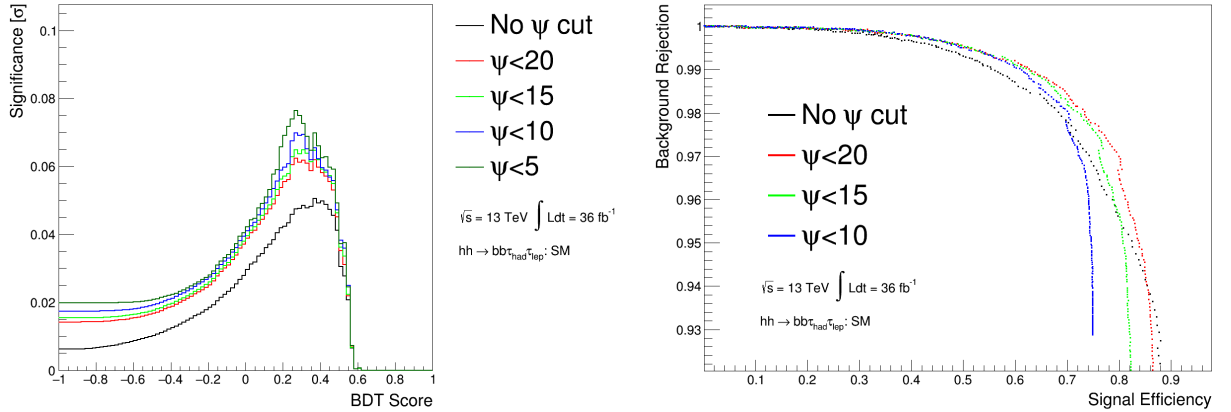
(a) Cut-and-count significance plot as a function of BDT score, when trained on the SM signal sample.

(b) ROC curves for the SM signal sample, with and without additional ψ requirements.

Figure 25: Distributions to show the effectiveness of the global optimisation parameter, ψ, to SM Higgs pair production searches.

energies the corrected $\not{E}_T$ is ascribed to the system of neutrinos produced in the decay of the $\tau_h\tau_l$ system and is taken into account (using the colinear approximation) in calculating the mass of the $\tau_h\tau_l$ system. An optimimum choice of b-jet energy re-scaling is obtained event by event by minimising a measure, ψ, of the kinematic consistency of an event with the signal hypothesis. The resulting value of ψ provides an event-by-event discriminant against background processes, since these are unlikely to yield events kinematically consistent with the signal hypothesis, even after re-scaling of the b-jet energies.

A particular feature of this approach is that the kinematic quantities considered are almost entirely independent of the total mass of the produced Higgs boson pair system. For example, the azimuthal angular separations between the corrected $\not{E}_T$ and the observed τ decay products are scaled by the azimuthal opening angle between the two observed τs. Similarly, the magnitude of the corrected $\not{E}_T$ is related to the p_Ts of the observed τ decay products. An advantage of this 'scaling' approach is that the discriminant ψ does not need to be optimised separately for each considered Higgs boson pair production mechanism; the shape of the distribution in ψ is independent of the distribution of produced Higgs boson pair masses. (See, e.g., figure 15).

We demonstrate that the variable ψ provides signal–background discriminating power that is complementary to that provided by the family of BDTs that have been tuned separately for each considered production mechanism and resonance mass as part of the ongoing ATLAS search for $hh \to bb\tau_h\tau_l$ [12]. For example, we find that combining selection cuts on ψ and the BDT output yields a significant reduction in the level of selected background at constant signal selection efficiency, and thus a higher a priori sensitivity to the presence of a signal, when compared to a selection that uses the BDT output alone. This is particularly the case for low higgs boson pair masses, where the kinematic discrimination of the BDT is modest. For example, in the case of a graviton resonance of mass 300 GeV a reduction by roughly a factor of two in background can be achieved for the same signal efficiency. (See, e.g., figures 21 and 23).

Monte Carlo simulations are found to describe well the distributions of relevant kinematic variables in the 36.1 fb^{-1} data set collected by ATLAS at 13 TeV (blinded in regions of high signal sensitivity). (See, e.g., figures 19 and 20).

We propose when the 36.1 fb^{-1} ATLAS data set is unblinded that the novel discriminant ψ would, at bare minimum, provide a useful cross check of the analysis that uses the BDT output alone. In concrete terms we propose that for each flavour of BDT the following distributions should be examined

- distribution of ψ, separately in a signal-enhanced and a signal-depleted region, as defined by the BDT output, and

- distribution of BDT output, separately in a signal-enhanced and a signal-depleted region, as defined by the value of ψ (say, $\psi < 10$ and $\psi > 10$).

These distributions would help cross check

- the discrimination power of the BDTs,

- the degree to which the SM background model is able to describe the data in signal enhanced and signal depleted regions of phase space, and

- for the presence/absence of evidence for a BSM signal.

We note that the technique could trivially be adapted to the search for $Zh \rightarrow bb\tau\tau$, by modifiying the constraint on the bb system mass to be consistent with the Z boson mass. A modified version of this technique could be applied to the search in the $hh \rightarrow bb\tau_h\tau_h$ channel with two hadronic τ decays.

Acknowledgements

## Appendix 1: Distributions of BDT score and ψ for the G(260) and G(500) samples

In this appendix the BDT score distributions are presented for the BDT trained on the G260 and G500 signal samples. The ψ distributions for the relevant signal masses are also shown after a cut on the BDT score has been imposed. The dashed line signals the position of the cut, where the arrow points to events which are kept. Again, a significant number of background events passing the cut are attributed high ψ scores. As the graviton mass decreases, the discriminating power of the BDT score decreases since it more difficult to define non-overlapping signal-like and background-like regions in the input parameters. The distributions are presented in figure 26

## Appendix 2: Results for the G(260) and G(500) signal samples

The cut-and-count significance plots and the ROC curves for the G260 and G500 signal are presented here. The plots were produced using the same technique outlined in 4.2. Imposing additional ψ requirements vastly improves signal–background discrimination in the G260 sample. The effect is less pronounced in the G500 distributions because it is comparitively easier to discriminate between signal and background in the BDT input parameters. Therefore, there is less scope for additional discriminating power in ψ variable since the BDT signal–background discrimination is already so strong. Nevertheless, an improvement is still observed. Note, low statistics limit the reliability of the G260 significance plot at the high BDT score end.

References

[1] G. e. a. Aad, *Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC*, Phys. Lett. **B716** (2012) 1, arXiv: `1207.7214 [hep-ex]`.

[2] e. a. Chatrchyan Serguei,
Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC,
Phys. Lett. **B716** (2012) 30, arXiv: `1207.7235 [hep-ex]`.

[3] P. D. Group, *Status of Higgs Boson Physics*, 2016, (visited on 29/12/2016).

[4] A. Collaboration, *Searches for Higgs boson pair production in the $hh \rightarrow bb\tau\tau$, $\gamma\gamma WW^*$, $\gamma\gamma bb$, bbbb channels with the ATLAS detector*, Phys. Rev. D **92** (9 2015) 092004,
URL: `http://link.aps.org/doi/10.1103/PhysRevD.92.092004`.

[5] J. Baglio et al., *The measurement of the Higgs self-coupling at the LHC: theoretical status*,
JHEP **04** (2013) 151, arXiv: `1212.5581 [hep-ph]`.

[6] M. Aaboud et al., *Search for pair production of Higgs bosons in the $b\bar{b}b\bar{b}$ final state using proton–proton collisions at $\sqrt{s}$ = 13 TeV with the ATLAS detector*,
Phys. Rev. **D94** (2016) 052002, arXiv: `1606.04782 [hep-ex]`.

[7] 'Search for Higgs boson pair production in the final state of $\gamma\gamma WW^*(\rightarrow l\nu jj)$ using 13.3 fb^{-1} of pp collision data recorded at $\sqrt{s}$ = 13 TeV with the ATLAS detector',
tech. rep. ATLAS-CONF-2016-071, CERN, 2016,
URL: `http://cds.cern.ch/record/2206222`.
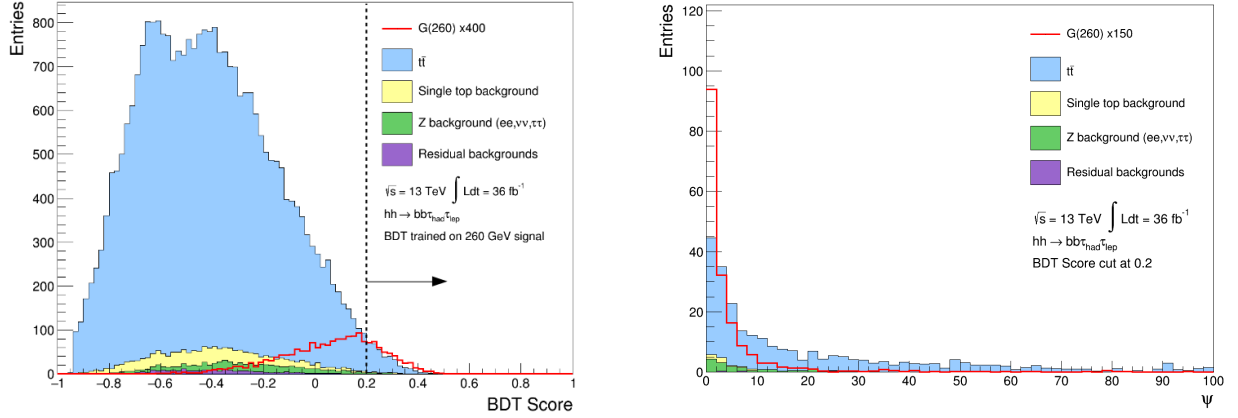
(a) BDT score distribution for training on the G260 sample. The dashed line respesents a cut at 0.2, where the arrow signifies events which are kept.

(b) G260 signal ψ distribution on top of background, where a cut to remove events with BDT score < 0.2 has been imposed.
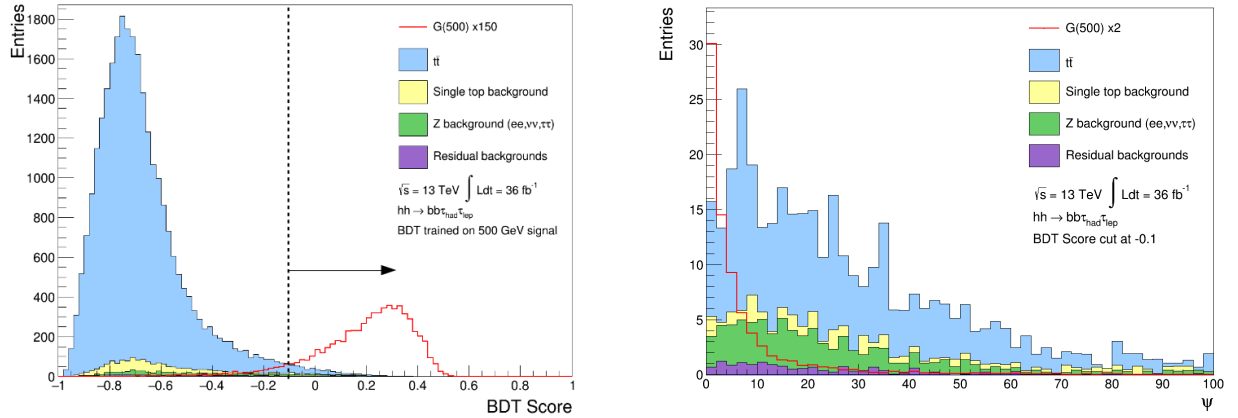
(c) BDT score distribution for training on the G500 sample. The dashed line respesents a cut at -0.1, where the arrow signifies events which are kept.

(d) G500 signal ψ distribution on top of background, where a cut to remove events with BDT score < -0.1 has been imposed.

Figure 26: BDT score distributions for training on the G260 and G500 samples separately. ψ distributions are presented next to the BDT score plots, for the respective signal samples. A cut on the BDT score has been imposed to produce the ψ distributions, identified by the dashed lines. Data has been blinded from these plots due to high signal–background discrimination
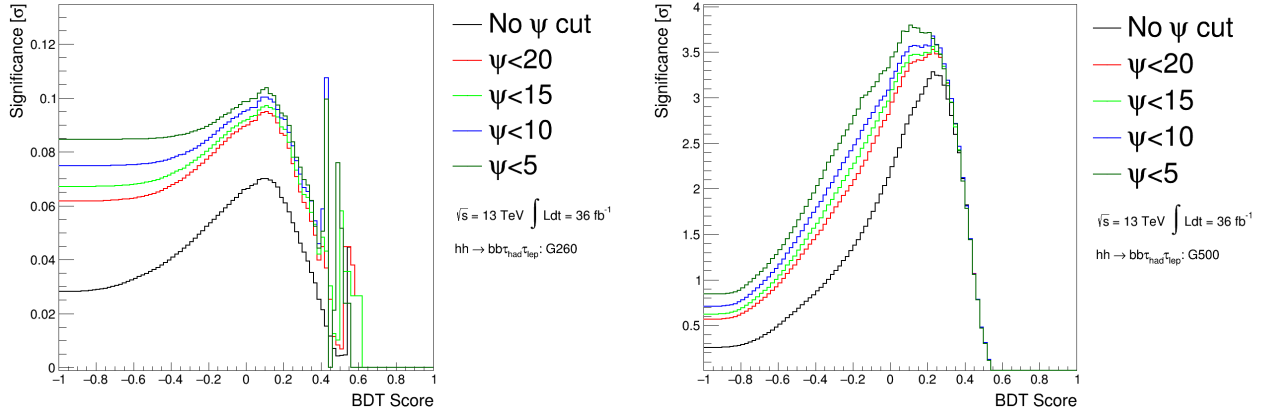
(a) G260 cut-and-count significance plot for the BDT score when trained on the G260 sample.

(b) G500 cut-and-count significance plot for the BDT score when trained on the G500 sample.
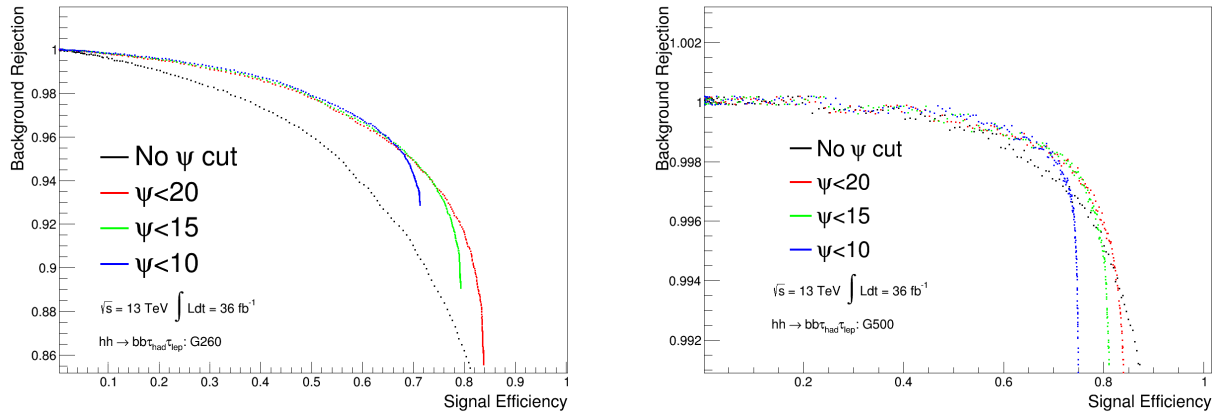
(c) ROC curves for BDT score when trained on the G260 sample.

(d) ROC curves for BDT score when trained on the G500 sample.

Figure 27: Cut-and-count significance plots and ROC curves as a function of a cut on the BDT score, trained on the G260 and G500 samples. The coloured curves have been created with additional requirements on the ψ variable.

[8] 'Search for Higgs boson pair production in the $b\bar{b}\gamma\gamma$ final state using pp collision data at $\sqrt{s} = 13$ TeV with the ATLAS detector', tech. rep. ATLAS-CONF-2016-004, CERN, 2016, URL: https://cds.cern.ch/record/2138949.

[9] F. Goertz, A. Papaefstathiou, L. L. Yang and J. Zurita, *Higgs Boson self-coupling measurements using ratios of cross sections*, JHEP **06** (2013) 016, arXiv: 1301.3492 [hep-ph].

[10] A. L. Fitzpatrick, J. Kaplan, L. Randall and L.-T. Wang, *Searching for the Kaluza-Klein Graviton in Bulk RS Models*, JHEP **09** (2007) 013, arXiv: hep-ph/0701150 [hep-ph].

[11] G. C. Branco et al., *Theory and phenomenology of two-Higgs-doublet models*, Phys. Rept. **516** (2012) 1, arXiv: 1106.0034 [hep-ph].

[12] 'Search for resonant and non-resonant di-Higgs production in the $bb\tau au\tau au$ decay channel with 36.1 fb^{-1} pp collision data at $\sqrt{s} = 13$ TeV with the ATLAS detector', tech. rep. ATLAS-COM-2017-007, CERN, 2017, URL: https://cds.cern.ch/record/2240998?.

[13] ATLAS Collaboration, *The ATLAS Experiment at the CERN Large Hadron Collider*, JINST **3** (2008) S08003.

[14] S. Alioli, P. Nason, C. Oleari and E. Re, *NLO Higgs boson production via gluon fusion matched with shower in POWHEG*, JHEP **04** (2009) 002, arXiv: 0812.0578 [hep-ph].

[15] T. Sjostrand, S. Mrenna and P. Z. Skands, *PYTHIA 6.4 Physics and Manual*, JHEP **05** (2006) 026, arXiv: hep-ph/0603175 [hep-ph].

[16] H.-L. Lai et al., *New parton distributions for collider physics*, Phys. Rev. **D82** (2010) 074024, arXiv: 1007.2241 [hep-ph].

[17] J. Pumplin et al., *New generation of parton distributions with uncertainties from global QCD analysis*, JHEP **07** (2002) 012, arXiv: hep-ph/0201195 [hep-ph].

[18] 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*, JHEP **07** (2014) 079, arXiv: 1405.0301 [hep-ph].

[19] T. Sjostrand, S. Mrenna and P. Z. Skands, *A Brief Introduction to PYTHIA 8.1*, Comput. Phys. Commun. **178** (2008) 852, arXiv: 0710.3820 [hep-ph].

[20] R. D. Ball et al., *Parton distributions with LHC data*, Nucl. Phys. **B867** (2013) 244, arXiv: 1207.1303 [hep-ph].

[21] T. Gleisberg et al., *Event generation with SHERPA 1.1*, JHEP **02** (2009) 007, arXiv: 0811.4622 [hep-ph].

[22] G. Corcella et al., *HERWIG 6: An Event generator for hadron emission reactions with interfering gluons (including supersymmetric processes)*, JHEP **01** (2001) 010, arXiv: hep-ph/0011363 [hep-ph].

List of contributions

This analysis was performed in the context of a final year undergraduate 'MPhys' research project at the University of Manchester, UK, by students **Jonathon Langford** and **Katie Spencer**, who designed, wrote and ran all of the required computer code and produced all of the analysis results presented here. The project was supervised by **Terry Wyatt**, who also suggested some aspects of the final algorithm. Specialist knowledge of the $hh \to bb\tau_h\tau_l$ analysis and datasets, as well as general analysis advice, was provided by **Agni Bethani**.