SEARCH FOR THE $H \to W^+W^-$ PROCESS AT THE LHeC EXPERIMENT*

Hollis Williams

School of Engineering, University of Warwick, Coventry CV4 7AL, UK

(Received May 19, 2020; accepted August 12, 2020)

We consider the decay of the Higgs boson to W^+ W^- at a proposed Large Hadron Electron Collider and determine the likelihood of detecting a signal for the Higgs mass from its decay product Wjets by imposing cuts to select candidate jet pairs and optimizing the value of the angular separation ΔR. It was found that at the LHeC experiment (CM energy $\sqrt{s} = 1.3$ TeV and luminosity of 100 fb^{-1} per year), the highest efficiency is obtained with $\Delta R = 0.4$, along with a selection scheme of $10 < m < 85$ GeV, $|\Delta\eta| < 1$, p_{T} of jets 1 and 2 between 10–20 GeV and p_{T} of jets 3 and $4 > 10$ GeV: this led to an efficiency between 7.1–7.5% for finding the invariant 4-jet mass in a mass region < 140 GeV. Under signal-to-background comparison, the signal showed a 3.8σ excess compared to the charged current W-jet background.

DOI:10.5506/APhysPolB.51.1811

1. Introduction

A Higgs boson is an excitation of the Higgs field, the field from which fundamental fermions and massive gauge bosons acquire their masses (with the fermions getting their masses from the Yukawa coupling to this field, and the gauge boson obtaining it via the Higgs mechanism). The major puzzle which the Higgs mechanism solves is the mass of the electroweak gauge bosons [1–4]. In order to experimentally study the Higgs boson in more detail and determine more of its properties, it would be necessary to produce and detect a large number of them on a reliable basis. The LHC is not suitable for this purpose as it has a large QCD background, making it difficult to separate the signal associated with a Higgs decay channel from other interactions which involve multi-jet final states. Although the discovery of the Higgs boson was a key discovery of the LHC, we still lack detailed understanding of its properties and couplings (including whether it

really is a fundamental scalar in the sense that leptons are fundamental, or whether it is a composite particle) [5, 6]. The branching ratios also need to be measured rigorously and checked against the predictions of the Standard Model, and it remains to be seen whether the measured total Higgs decay cross section can be accounted for using only the particles of the Standard Model.

By studying how it couples to its decay products, we may also uncover properties which are unexpected or not explained in the SM. It is known that the Higgs has even parity and zero spin, and hence it represents an (apparently) fundamental scalar, unlike scalar mesons, which are hadronic composites. Its mass is a free parameter of the SM given by $m_H = 2\lambda v^2$. It is a neutral particle and as a consequence of its role in generating mass, it couples to mass. In theory, the Higgs can decay to any other particle in the SM, but the coupling is proportional to mass, so the largest branching ratio should be to the most massive particle which is kinematically accessible. It follows that the decay mode with the largest branching ratio is $b\bar{b}$. In this work, we study the possibility of using an *ep* collider to search for Higgs boson production via Higgs decay to W-boson pairs.

The LHeC is one of several options for a post-LHC collider. At the time of writing, it is probably the least popular of the available possibilities, and a very strong technical case would need to be made for it to replace the preference for a new hadron collider. A case also needs to be made that the LHeC could function as a Higgs boson factory at a level competitive with that of an electron–positron collider, and we take first steps in this direction in our article. As might be expected due to this relative lack of popularity, there are much fewer studies of the physics and technological feasibility of the LHeC and a correspondingly small number of studies of Higgs boson production at the LHeC. Almost all of the existing studies of the LHeC of this kind focus on exotic Higgs decays, BSM physics and BSM Higgs boson signatures [7–11]. In our study, we are instead concerned with probing the limits of the Standard Model and the SM Higgs boson, with the belief that this will give us the best possible clues of the kind of physics which lies beyond the Standard Model. A recent study suggests detecting the Higgs boson at the LHeC via photoproduction (a very different process), but ultimately concludes that photoproduction is not a feasible way of detecting the Higgs boson signal at the LHeC [12]. On the other hand, we demonstrate that the process which we study is feasible for signal detection and that it merits further study with more sophisticated techniques.

The main existing study of the design and conception of the LHeC can be found in [13]. This article dedicates very few pages to Higgs boson production and suggests detection of the Higgs by reconstruction of its dominant decay channel $H \to \bar{b}b$. However, the suggestion that one could use this

decay channel to find the Higgs boson is not novel, and it was only one of a selection of complementary decay channels considered in the LHC discovery of the Higgs [5, 6]. In any case, the most significant evidence for the Higgs boson at the LHC came from two rather different processes, $H \to \gamma\gamma$ and $H \to ZZ^* \to 4l$. In fact, the possibility of the LHeC as a Higgs boson factory where the Higgs boson is produced and studied on a reliable basis is simply not mentioned at all in [13] and no reference is made to a Higgs factory.

Searching for the Higgs boson via the decay channel mentioned in [13] is unlikely to work. In general, it is shown that reconstruction from the $\bar{b}b$ process suffers from difficulties which our approach does not have, since it is necessary to develop methods which distinguish the b jets from other quark-jets which can easily be misidentified as b-tagged jets (especially due to top production). Reference [13] does not consider b-quarks produced during the parton showering and so overestimates the signal-to-background expectation by a factor of five. The argument relies heavily on the energy resolution and b-tagging capabilities of the detector in question: these details are impossible to know without more realistic detector simulations and rely on engineering developments in detector technology. Along with this, the estimation of the background rejection remains subject to large uncertainties. In summary, the method of signal detection which is mentioned in [13] is likely not feasible and would require substantial developments in b-tagging techniques which are not guaranteed.

Reference [13] does mention the possibility of detecting the Higgs at the LHeC via final states such as WW or via photoproduction. The former is exactly what we study in this article, and in [12] it is shown that the latter is not feasible. Our study is novel in the process which we consider for detection and in our demonstration that this process is a feasible way to detect a signal. Although our study is not the first to suggest that the LHeC could function as a Higgs boson factory, our suggestion that the SM Higgs could be produced and detected on a reliable basis at the LHeC via the $H \to W^+W^-$ appears to be novel. Finally, we should mention that the Higgs boson coupling to the b quark has already been well-studied, whereas the approach which we suggest would also allow for many opportunities to study the much less well-understood $H \to WW$ coupling [14].

At the LHeC, the most probable mechanism for Higgs boson production is a charged current or neutral current interaction via W- or Z-boson fusion, resulting in a Higgs boson, a jet, and an electron neutrino [15]. In Fig. 1, we give some possible tree-level Feynman diagrams which contribute to Higgs boson production via W-boson fusion: the left-hand Feynman diagram contributes to Higgs boson production at the LHeC, and the right-hand diagram contributes to Higgs boson production at the LHC. In general, identification

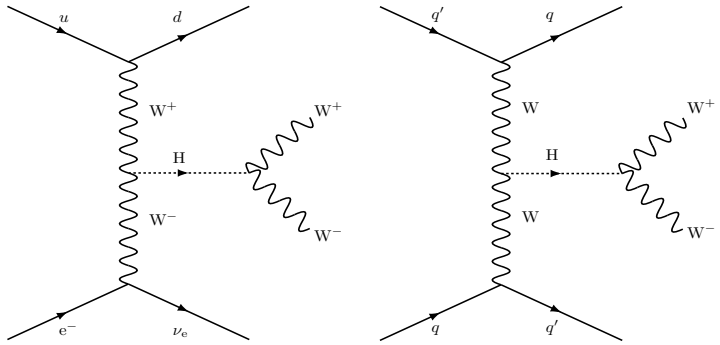

Fig. 1. Tree-level Feynman diagrams contributing to Higgs boson production via W boson fusion. The left-hand diagram contributes to Higgs production at the LHeC, and the right-hand diagram contributes to Higgs production at the LHC.

of the Higgs at the LHC via hadronic decay of the W boson is not considered viable because it is difficult to distinguish final-state jet decays of the Higgs from the huge number of other events at the LHC which involve multi-jet final states. At the LHeC, the QCD background is cleaner by a factor of 100 and the DIS final state is clean. It is possible to detect the W boson through its leptonic decays (such as $e^+\nu$) but these have relatively small branching ratios. The branching ratio for the W boson decay to two jets is dominant, with the ratio of $W^+ \to q\bar{q}$ being $(67.41 \pm 0.27)\%$ [16]. Given the abundance of this branching fraction, there is the potential to not only study $H \to WW^*$ in more detail at the LHeC but also to use it to measure Higgs boson production, so creating a precision Higgs factory where the Higgs can be created on a reliable basis. There is interest in studying the HWW coupling because, unlike at the LHC where the HZZ and HWW couplings cannot be separated, the two are separate at the LHeC and the latter could have contributions which are not explained by the SM. A more detailed knowledge of Higgs couplings is also necessary to determine if the fundamental fermions do in fact obtain their masses via Yukawa coupling to the Higgs field.

Studying the $H \to W^+W^-$ process in detail relies on the possibility of tagging on W jets (as has already been done at the LHC). The production of a W boson and another virtual W boson is followed by hadronic decays of both Ws, resulting in four or five W jets in the final state (we will leave aside the leptonic decays). When the W boson is very energetic, the two jets will be close together and will merge, meaning that we actually end up reconstructing one single jet characterised by a two-prong structure. The analysis of this channel relies on the possibility of being able to distinguish the W jets from jets due to quarks and gluons produced in strong interactions.

2. Large Hadron Electron Collider

Physicists are currently exploring options for a next-generation collider at the energy frontier: two possibilities are a new electron–positron collider or an LHeC (Large Hadron–Electron Collider). The former would be similar to the LEP (Large Electron–Positron Collider) and the latter would be similar to HERA at DESY, but with a greater centre-of-mass energy compared to CM energy of 318 GeV at HERA. The advantage of an *ep* collider is that it offers the opportunity to observe phenomena which would be observed in a *pp* collider with a cleaner decay environment and reduced contamination from unwanted multi-jet final states. One important aspect of the LHeC which separates it from the LEP is that it complements the LHC: the LHeC would provide an electron beam between 60 and 150 GeV (compared to 27.5 GeV for the lepton beam at HERA) which would be collided with the intense hadron beams already provided by the LHC. This would increase the kinematic range by a factor of twenty for Q^2 and inverse x, and there would be an increase over the integrated luminosity of HERA of two orders of magnitude, with a luminosity of 10^{33} cm^{-2} s^{-1}.

The LHeC could potentially be realized as a ring–ring or a ring–linac configuration. In the ring–ring configuration, the same geometry is used for both components and the technology of the ring setup has been extensively studied at HERA and LEP. The electrons are accelerated in a ring, whereas in the ring–linac configuration, the electrons are accelerated to the required energies in a linear accelerator before being collided with the protons travelling around the LHC. The process of generating intense lepton beams in a storage ring is well-understood. However, the ring–linac configuration has the advantage that the infrastructure of the linac only meets with the ring in the vicinity of the interaction vertex, minimising interference due to hadron beams [13]. An initial 500 MeV electron bunch originating at the injector is accelerated to 10 GeV in each linac, leading to a final energy of 60 GeV at the interaction vertex after passing through the entire setup three times. The 60 GeV beam is then collided with the proton beam from the LHC.

3. Simulated samples

The search was performed using a simulated LHeC experiment with $\sqrt{s} = 1.3$ TeV *ep* collisions and luminosity of 100 fb^{-1} per year. In almost all of the analysis of the jet kinematics, there was a p_{T} cut on the jets which would assist in jet reconstruction. Low p_{T} jets are typical when the jets are formed via hadronization of QCD radiation. These background jets can be due to quarks or gluons emitted by particles inside the signal jet which then fragment and hadronize to form new jets [17].

These false jets typically have smaller values of p_{T} and so can be removed with a cut on transverse momentum of the jets. When reconstructing the jets, we must also consider the separation of the jets, their size and the algorithm used. An algorithm which is in use at the LHC (and in simulations of LHeC for consistency purposes) is the anti-k_{T} algorithm, along with a distance parameter. A typical way of trying to categorize signal jets which emerge from a decay is to study how many of them have merged or separated configurations. For example, we could consider a dijet with $\Delta R = 0.4$ as being separated, whereas a pair of jets with lower separation merge to form one jet. ΔR is defined as follows:

$$\Delta R = \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2}\,, \tag{1}$$

where η is the pseudorapidity and ϕ is the usual angular coordinate. The anti-k_{T} algorithm is an example of a sequential recombination jet algorithm for jet reconstruction. These algorithms normally have as their parameter the power of the energy scale in the distance measure, and the 'anti' comes from the fact that the power is negative, as opposed to the ordinary k_{T} algorithm. The distance measure between entities i and j is defined as follows:

$$d_{ij} = \min\left(k_{\mathrm{t}i}^{2p}, k_{\mathrm{t}j}^{2p}\right)\frac{\Delta_{ij}^2}{R^2}\,, \tag{2}$$

where $\Delta_{ij}^2 = (y_i - y_j)^2 + (\phi_i + \phi_j)^2$ and $k_{\mathrm{t}i}$, y_i and ϕ_i are the transverse momentum, rapidity, and the azimuth of the particle i, respectively. As mentioned, the parameter p controls the relative power of the energy *versus* geometric scales [18].

Generation of events and cross sections (and numerical evaluation of relevant matrix elements) was carried out via MadGraph5, which was used to generate both signal and background events [17]. The partons produced by MadGraph5 were assigned 4-vectors and were then showered by PYTHIA, an event generator which simulates the hadronisation and decay of the showers. PYTHIA describes the parton distributions, the initial- and final-state parton showers, and the interactions between the partons in the showers [19]. CTEQ5L was used as the parton distribution set for the proton: CTEQ5L is parametrised by a leading order fit to experimental data which constrains the Q^2 dependence to agree with the Altarelli–Parisi QCD equations [20]. The resulting events were assigned 4-vectors and interfaced to Delphes, a detector package which includes simulations of systems to trigger on tracks and simulations of calorimeters and muon detectors. Delphes analyses events generated by PYTHIA and creates a dataset as an output which can be used for reconstruction [21].

4. Event selection

4.1. Parton level plots

We begin by checking that our samples are for the WW^* decay where one of the W bosons is off mass shell, and not WW, where both W bosons are on-shell. The latter case makes the Higgs decay to two W bosons effectively impossible since $m_H < 2m_W$. The partonic mass distributions were plotted for W^+ and W^- and shown to be almost the same. In both cases, a sharp peak was observed around the mass of the W boson (approximately 80 GeV), along with a smaller peak between 10–50 GeV, reaching a highest point between 30–40 GeV. Since the off-shell W boson has a smaller mass, this confirmed that the samples generated WW^*. It was also confirmed that a plot of the Higgs mass at parton level led to a peak at the Higgs mass as expected (around 125 GeV). This is shown in Fig. 2 where the mass at parton level is plotted (the units of mass always being GeV).

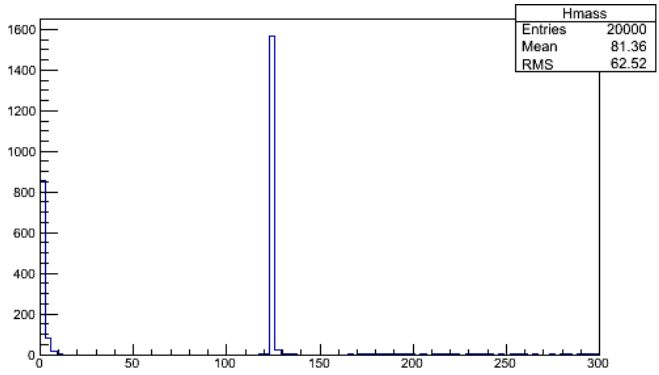

Fig. 2. Parton level plot of Higgs mass.

4.2. W-jets selection

A selection criterion on the transverse momentum of the jets was required to remove a large number of the false jets which are present due to background QCD radiation. To determine this criterion, the invariant mass of the two leading jets in each event (jets 1 and 2) was plotted. Kinematically, it was expected that the W jets in an event could be detected via the jet pairs composed of the leading jets, and so the jet pair composed of leading jets 1 and 2 should reconstruct to have the invariant mass of an on-shell W boson, whereas the pair composed of leading jets 3 and 4 should reconstruct to the invariant mass of an off-shell W boson. A cut of $p > 30$ GeV was found to remove a large number of events with background jets whilst

retaining the structure of the W peak, hence it was determined that this cut would be used as a maximum for removing background events, being lowered as necessary to detect the off-shell W^* (since a cut of $p_T > 30$ GeV is rather high to see a signal from a W^* between 10 and 50 GeV). In fact, a cut of $p_T > 30$ on all jets is very high for charged-current DIS and parton level observation showed that the transverse momentum p_T of the generated Higgs boson peaked at 50 GeV, whereas the p_T of the on-shell W boson would be expected to peak at 40 GeV.

4.3. Signal selection

To reconstruct the Higgs mass, a cut of $|\Delta\eta| < 1$ (where η is the pseudo-rapidity) was imposed for the difference between two jets in a jet pair. The four W-tagged jets being used to reconstruct the Higgs mass can obviously appear in multiple combinations (for example, we could have the jet pairs (1,3) and (2,4) both in the necessary mass window, but not (1,2) and (3,4)). Ideally, one would like to consider the η differences for all the jets in an event and not just the 4 leading jets, so the masses of all possible permuations of jet pairs in an event was considered where each permutation contains 4 jets. Analysis confirmed that the signal for the Higgs boson could be improved by taking cuts of $p_T > 20$ GeV on jets 1 and 2, $p_T > 10$ GeV on jets 3 and 4, $|\Delta\eta| < 1$ and invariant mass m of jets 3 and 4 larger than 10 GeV. The histogram (shown in Fig. 3) shows a fairly high number of entries around the Higgs mass region with these selection criteria.

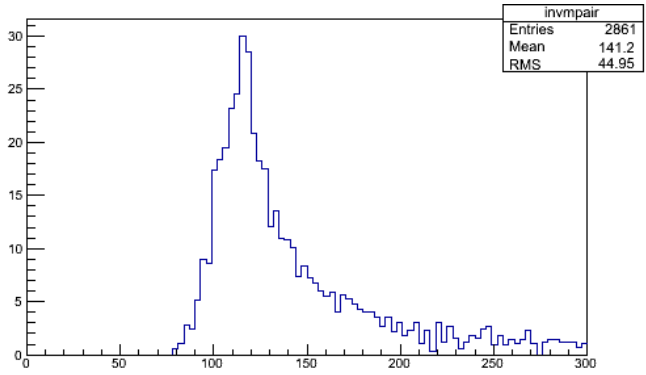

Fig. 3. Mass distribution for the 4-jet system (cuts of $p_T > 10$ on (3,4), $m > 10$). Units are GeV/c and GeV, respectively.

Up to this point, all samples used for the analysis assume a value of angular separation between jets of $\Delta R = 0.4$. This is likely optimal for the types of jet which are being studied. This assumption was confirmed by

repeating the analysis for other values of ΔR and finding that other values apart from $\Delta R = 0.4$ or $\Delta R = 0.5$ reduce the number of events falling into the Higgs mass region. In Figs. 4 and 5, we show the resulting mass distribution when the analysis is repeated for $\Delta R = 0.4$ and $\Delta R = 0.5$, respectively. The actual optimal value was confirmed later via selection efficiencies.

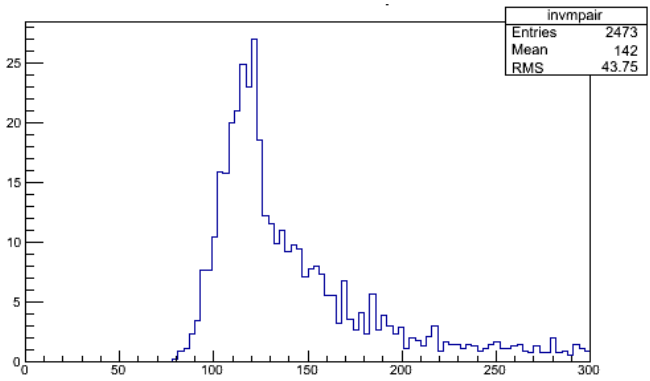

Fig. 4. Mass distribution for the 4-jet system ($\Delta R = 0.5$).

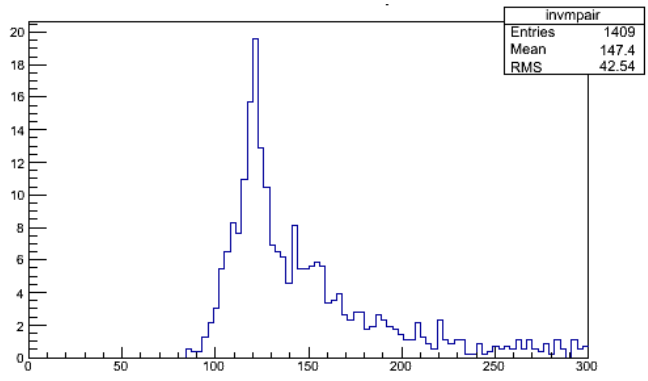

Fig. 5. Mass distribution for the 4-jet system ($\Delta R = 0.7$).

5. Signal-to-background comparison

A signal-to-background comparison was difficult for a study of this kind, as the main background was due to multi-jet final states from charged-current DIS or from photoproduction of multi-jets. Since it is non-trivial to produce high statistics of multi-jets, it would be normal in this situation

to make cuts at parton level and then optimize the ΔR parameter for the separation between candidate W jets. However, it was still desirable to have a preliminary background sample to show that it was suppressed by the proposed cuts on the signal sample (for example, a sample of background W jets). Such a sample should also take account of the asymmetry in the WW^* decay. The analysis was repeated with the BG sample over the same number of events and a comparison made with the previous results. This comparison is shown in Fig. 6. For reference, note that the total production cross section in fb of the SM Higgs boson via CC interactions in electron–proton collisions with electron beam energy of 50 GeV is 81 fb [13].

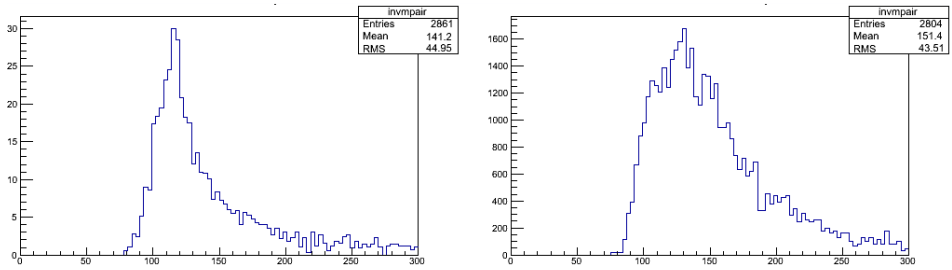

Fig. 6. Signal-to-background comparison. The left-hand figure shows the shape of the mass distribution for the signal sample and the right-hand figure gives the same mass distribution when the analysis is repeated with the background sample.

Even accounting for normalization, the suppression of background events does not look promising, so a stricter mass cut is required on the 4-jet of $m < 130$ GeV or $m < 140$. The result of the same analysis with this cut is shown in Fig. 7.

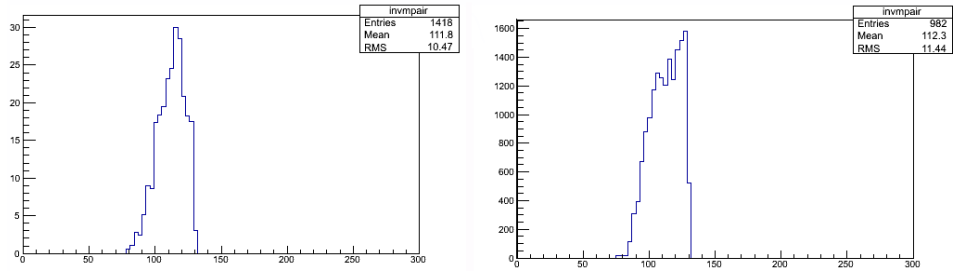

Fig. 7. Signal-to-background comparison with mass cut $m < 130$ GeV on 4-jets. The left-hand figure corresponds to the signal sample and the right-hand figure corresponds to the same analysis with the background sample.

A strict cut is required on the mass of the 4-jet in order to see a good signal-to-background ratio: 1.44 for $m < 130$ and 1.28 for $m < 140$ compared to 1.02 for no mass cut.

5.1. Selection efficiencies

Efficiencies were calculated for various cuts on p_{T} and m at $\Delta R = 0.4$. The selection efficiency was defined as the ratio between the number of events passing the selection criteria with 4-jet masses below 140 GeV and the total number of events. $p_{\mathrm{T}ij}$ denotes the p_{T} cut on the jet pair composed of jets i and j, and m is the lower mass cut on the 4-jet. In Table I, we show the signal-to-background ratio for various sets of selection criteria.

TABLE I

Number of events passing selection criteria ($m < 140$ GeV).

Cuts	Signal	Background	S/B ratio
$p_{\mathrm{T}12} > 20,\ p_{\mathrm{T}34} > 20,\ m > 10$	186	89	2.09
$p_{\mathrm{T}12} > 20,\ p_{\mathrm{T}34} > 20,\ m > 20$	161	78	2.06
$p_{\mathrm{T}12} > 20,\ p_{\mathrm{T}34} > 15,\ m > 20$	472	302	1.56
$p_{\mathrm{T}12} > 20,\ p_{\mathrm{T}34} > 15,\ m > 10$	594	391	1.52
$p_{\mathrm{T}12} > 20,\ p_{\mathrm{T}34} > 10,\ m > 10$	1644	1282	1.28
$p_{\mathrm{T}12} > 20,\ p_{\mathrm{T}34} > 10,\ m > 20$	1140	845	1.35
$p_{\mathrm{T}12} > 15,\ p_{\mathrm{T}34} > 10,\ m > 20$	1192	885	1.35
$p_{\mathrm{T}12} > 15,\ p_{\mathrm{T}34} > 10,\ m > 10$	1731	1353	1.28
$p_{\mathrm{T}12} > 10,\ p_{\mathrm{T}34} > 10,\ m > 20$	1195	887	1.35
$p_{\mathrm{T}12} > 10,\ p_{\mathrm{T}34} > 10,\ m > 10$	1738	1363	1.28

One obvious conclusion from the table is that lowering the mass cut from $m > 20$ to $m > 10$ increases the selection efficiency but lowers the signal-to-background ratio. The difference in signal-to-background ratio in this case is negligible (between 0.03 and 0.07) and so the focus was placed on the candidate cut sets which produced the greatest number of events, especially since signal-to-background was not a major part of the study. The cuts which produced the largest efficiencies were $p_{\mathrm{T}12} > 20,\ p_{\mathrm{T}34} > 10,\ m > 10$, $p_{\mathrm{T}12} > 15,\ p_{\mathrm{T}34} > 10,\ m > 10$ and $p_{\mathrm{T}12} > 10,\ p_{\mathrm{T}34} > 10,\ m > 10$. The analysis was repeated for the above optimal cut sets over samples with ΔR ranging from 0.3 to 0.7: the results are shown in Table II for signal only.

The analysis was repeated for a mass cut of < 130 GeV on the 4-jets. Table III repeats the original analysis of Table II with a mass cut of less than 130 GeV on the mass of the 4-jet, and Table IV repeats the analysis for ΔR ranging from 0.3 to 0.7 with the same mass cut.

TABLE II

Number of events passing selection criteria for $\Delta R = 0.3$–0.7.

Cuts	0.3	0.4	0.5	0.6	0.7
$p_{\mathrm{T}12} > 20,\ p_{\mathrm{T}34} > 10,\ m > 10$	715	1644	1416	1027	720
$p_{\mathrm{T}12} > 15,\ p_{\mathrm{T}34} > 10,\ m > 10$	757	1731	1488	1098	765
$p_{\mathrm{T}12} > 10,\ p_{\mathrm{T}34} > 10,\ m > 10$	758	1738	1494	1110	777

TABLE III

Number of events passing selection criteria ($m < 130$ GeV).

Cuts	Signal	Background	S/B ratio
$p_{\mathrm{T}12} > 20,\ p_{\mathrm{T}34} > 20,\ m > 10$	132	43	3.07
$p_{\mathrm{T}12} > 20,\ p_{\mathrm{T}34} > 20,\ m > 20$	113	34	3.32
$p_{\mathrm{T}12} > 20,\ p_{\mathrm{T}34} > 15,\ m > 20$	378	195	1.94
$p_{\mathrm{T}12} > 20,\ p_{\mathrm{T}34} > 15,\ m > 10$	482	260	1.85
$p_{\mathrm{T}12} > 20,\ p_{\mathrm{T}34} > 10,\ m > 10$	1418	982	1.44
$p_{\mathrm{T}12} > 20,\ p_{\mathrm{T}34} > 10,\ m > 20$	966	633	1.53
$p_{\mathrm{T}12} > 15,\ p_{\mathrm{T}34} > 10,\ m > 20$	1014	673	1.51
$p_{\mathrm{T}12} > 15,\ p_{\mathrm{T}34} > 10,\ m > 10$	1497	1048	1.43
$p_{\mathrm{T}12} > 10,\ p_{\mathrm{T}34} > 10,\ m > 20$	1017	675	1.51
$p_{\mathrm{T}12} > 10,\ p_{\mathrm{T}34} > 10,\ m > 10$	1503	1056	1.42

TABLE IV

Number of events passing selection criteria for $\Delta R = 0.3$–0.7.

Cuts	0.3	0.4	0.5	0.6	0.7
$p_{\mathrm{T}12} > 20,\ p_{\mathrm{T}34} > 10,\ m > 10$	625	1418	1220	869	606
$p_{\mathrm{T}12} > 15,\ p_{\mathrm{T}34} > 10,\ m > 10$	666	1497	1285	930	647
$p_{\mathrm{T}12} > 10,\ p_{\mathrm{T}34} > 10,\ m > 10$	667	1503	1291	941	656

This lowered the selection efficiency but increased the signal-to-background ratio. Using calculated values for the Higgs cross section in an ep collider, a value for the background cross section and the numbers of events computed in Tables II and IV, values were calculated for expected numbers of signal and background events along with the expected significance, and the results shown in Tables V and VI for mass cuts of $m < 130$ GeV and $m < 140$ GeV, respectively. As a result, the actual signal-to-background ratios were much smallerthan the ones calculated initially, since the background cross section is larger than the signal cross section.

TABLE V

Expected numbers of events ($m < 130$ GeV).

Cuts	Signal	Background	Significance
$p_{\mathrm{T}12} > 20,\ p_{\mathrm{T}34} > 20,\ m > 10$	568	120400	1.6
$p_{\mathrm{T}12} > 20,\ p_{\mathrm{T}34} > 20,\ m > 20$	486	95200	1.6
$p_{\mathrm{T}12} > 20,\ p_{\mathrm{T}34} > 15,\ m > 20$	1625	546000	2.2
$p_{\mathrm{T}12} > 20,\ p_{\mathrm{T}34} > 15,\ m > 10$	2073	728000	2.4
$p_{\mathrm{T}12} > 20,\ p_{\mathrm{T}34} > 10,\ m > 10$	6097	2749600	3.7
$p_{\mathrm{T}12} > 20,\ p_{\mathrm{T}34} > 10,\ m > 20$	4154	1772400	3.1
$p_{\mathrm{T}12} > 15,\ p_{\mathrm{T}34} > 10,\ m > 20$	4360	1884400	3.21
$p_{\mathrm{T}12} > 15,\ p_{\mathrm{T}34} > 10,\ m > 10$	6437	2934400	3.8
$p_{\mathrm{T}12} > 10,\ p_{\mathrm{T}34} > 10,\ m > 20$	4373	1890000	3.2
$p_{\mathrm{T}12} > 10,\ p_{\mathrm{T}34} > 10,\ m > 10$	6463	2956800	3.8

TABLE VI

Expected number of events ($m < 140$ GeV).

Cuts	Signal	Background	Significance
$p_{\mathrm{T}12} > 20,\ p_{\mathrm{T}34} > 20,\ m > 10$	800	249200	1.6
$p_{\mathrm{T}12} > 20,\ p_{\mathrm{T}34} > 20,\ m > 20$	692	218400	1.5
$p_{\mathrm{T}12} > 20,\ p_{\mathrm{T}34} > 15,\ m > 20$	2030	845600	2.2
$p_{\mathrm{T}12} > 20,\ p_{\mathrm{T}34} > 15,\ m > 10$	2554	1094800	2.4
$p_{\mathrm{T}12} > 20,\ p_{\mathrm{T}34} > 10,\ m > 10$	7069	3589600	3.7
$p_{\mathrm{T}12} > 20,\ p_{\mathrm{T}34} > 10,\ m > 20$	4902	2366000	3.2
$p_{\mathrm{T}12} > 15,\ p_{\mathrm{T}34} > 10,\ m > 20$	5126	2478000	3.3
$p_{\mathrm{T}12} > 15,\ p_{\mathrm{T}34} > 10,\ m > 10$	7443	3788400	3.8
$p_{\mathrm{T}12} > 10,\ p_{\mathrm{T}34} > 10,\ m > 20$	5139	2483600	3.3
$p_{\mathrm{T}12} > 10,\ p_{\mathrm{T}34} > 10,\ m > 10$	7473	3816499	3.8

The background sample used at this point was for background due to W jets and so the analysis was also run over a sample for more general QCD background. We repeated the analysis for events with invariant 4-jet masses below 130 GeV and added in the effect of the second background. The numbers of events were calculated and then scaled up appropriately: the results are shown in Table VII.

The number of background events is now quite high: choosing the signal for the first set of cuts in Table VI, the expected background events would have an uncertainty of $\sqrt{120400} = 347$ such that the signal would be seen at 1.5σ. An improvement to 3σ could likely be achieved with improvements to the analysis and use of a more refined selection strategy and improved technique for signal-to-background comparison.

TABLE VII

Expected number of events with QCD background ($m < 130$ GeV).

Cuts	Signal	Background	Significance
$p_{\mathrm{T}12} > 20,\ p_{\mathrm{T}34} > 20,\ m > 10$	800	453600	1.2
$p_{\mathrm{T}12} > 20,\ p_{\mathrm{T}34} > 20,\ m > 20$	692	364400	1.1
$p_{\mathrm{T}12} > 20,\ p_{\mathrm{T}34} > 15,\ m > 20$	2030	1838400	1.5
$p_{\mathrm{T}12} > 20,\ p_{\mathrm{T}34} > 15,\ m > 10$	2554	2438000	1.6
$p_{\mathrm{T}12} > 20,\ p_{\mathrm{T}34} > 10,\ m > 10$	7069	10218000	2.2
$p_{\mathrm{T}12} > 20,\ p_{\mathrm{T}34} > 10,\ m > 20$	4902	6570800	1.9
$p_{\mathrm{T}12} > 15,\ p_{\mathrm{T}34} > 10,\ m > 20$	5126	6945600	1.9
$p_{\mathrm{T}12} > 15,\ p_{\mathrm{T}34} > 10,\ m > 10$	7443	11015400	2.2
$p_{\mathrm{T}12} > 10,\ p_{\mathrm{T}34} > 10,\ m > 20$	5139	7038800	1.9
$p_{\mathrm{T}12} > 10,\ p_{\mathrm{T}34} > 10,\ m > 10$	7473	11174899	2.2

6. Conclusion

After cuts and analysis had been carried out, it was found that a good selection scheme for the search which we are studying is $|\Delta\eta| < 1, 10 < m < 85$ GeV, p_{T} of jets 1 and 2 between 10–20% and p_{T} of jets 3 and 4 > 10 GeV, and that the selection efficiency is highest for $\Delta R = 0.4$, leading to an efficiency between 7.1–7.5% for finding the invariant 4-jet mass in a mass region < 140 GeV. In fact, $\Delta R = 0.4$ was optimal for all cut sets where it was varied, but this could be unique to the particular decay we were studying, as other Higgs decays often show a strong dependence on ΔR. It was also found that we could begin to incorporate a background sample due to W jets with a signal which would be observed with 3.8σ, but attempts to make this figure more realistic by adding QCD background resulted in a very high number of background events.

We should also point out that this is a preliminary study and that further work would be required before our conclusion could be stated more firmly. The main reason is that we had not always been running on full signal statistics, only using 10 000 events in the interest of efficiency when running samples with many modified versions of the analysis code to find effective cuts (something like a quarter of the full signal statistics). We also did not account for the hadronic branching fraction of the W boson. It is likely that both of these factors led to a small signal cross section of around 0.1 pb. In our calculations, we have assumed a value of 1 ab^{-1} of luminosity, whereas a revised value of 2 ab^{-1} would double the signal cross section. This would, however, increase the significance by a factor of $\sqrt{2}$ to compensate. A proper comparison of signal-to-background would be difficult in this study and it might be considered satisfactory to have imposed a selection scheme

and optimized ΔR, but it would obviously be desirable to perform a more sophisticated signal-to-background analysis using boosted decision trees [22]. BDT is especially useful when the signal is 'drowned out' by similar-looking background events, since it can be used to identify if events are signal-like or background-like by using Monte Carlo simulations to train the decision tree. This then enables a final determination of signal strength compared to background [23].

It would also be desirable to adjust or refine cuts to increase the efficiency of signal detection, since the mass cuts employed appeared to be drowning out the signal at the region of interest. The next step besides BDT would be to run the analysis again on the full Monte Carlo signal statistics, as this could have an effect on the signal cross section. The effect of other variables could be studied for both signal and background: in particular, adjustments of $\Delta\eta$ were not investigated for background topologies in this study. The study only considered charged current and QCD background. In a more thorough study, different types of neutral current background could be incorporated.

I would like to thank Mario Campanelli and Uta Klein for useful discussions and for providing data samples.

REFERENCES

[1] F. Englert, R. Brout, «Broken Symmetry and the Mass of Gauge Vector Mesons», *Phys. Rev. Lett.* **13**, 321 (1964).

[2] P.W. Higgs, «Broken Symmetries and the Masses of Gauge Bosons», *Phys. Rev. Lett.* **13**, 508 (1964).

[3] A. Salam, J.C. Ward, «Weak and electromagnetic interactions», *Nuovo Cim.* **11**, 568 (1959).

[4] S. Weinberg, «A Model of Leptons», *Phys. Rev. Lett.* **19**, 1264 (1967).

[5] ATLAS Collaboration, «Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC», *Phys. Lett. B* **716**, 1 (2012).

[6] CMS Collaboration, «Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC», *Phys. Lett. B* **716**, 30 (2012).

[7] S. Liu, Y. Tang, C. Zhang, S. Zhu, «Exotic Higgs decay $h \to \phi\phi \to 4b$ at the LHeC», *Eur. Phys. J. C* **77**, 457 (2017).

[8] S. Das, M. Nowakowski, «Light neutral CP-even Higgs boson within the next-to-minimal supersymmetric standard model at the Large Hadron Electron Collider», *Phys. Rev. D* **96**, 055014 (2017).

[9] C. Mosomane, M. Kumar, A. Cornell, B. Mellado, «Exploring CP-even scalars of a Two Higgs-doublet model in future $e^- p$ colliders», *J. Phys.: Conf. Ser.* **889**, 012004 (2017).

[10] L. Rose, O. Fischer, A. Hammad, «Prospects for heavy scalar searches at the LHeC», *Int. J. Mod. Phys. A* **34**, 1950127 (2019).

[11] O. Flores Sánchez *et al.*, «Light charged Higgs boson production at *ep* colliders», in: Conference Proceedings of XXVII International Workshop on Deep-Inelastic Scattering and Related Subjects, Torino, Italy, 2019.

[12] R. Li *et al.*, «Photo-production of the Higgs boson at the LHeC», *Nucl. Phys. B* **958**, 115134 (2020).

[13] HeC Study Group, «A Large Hadron Electron Collider at CERN: Report on the Physics and Design Concepts for Machine and Detector», *J. Phys. G: Nucl. Part. Phys.* **39**, 075001 (2012).

[14] ATLAS Collaboration, «Evidence for the $H \to b\bar{b}$ decay with the ATLAS dector», *J. High Energy Phys.* **2017**, 24 (2017).

[15] J. Esteves, «Understanding the Higgs boson with the Large Hadron Electron Collider», *J. Phys.: Conf. Ser.* **645**, 012009 (2015).

[16] C. Patrignani *et al.*, «Review of Particle Physics», *Chin. Phys. C* **40**, 100001 (2016).

[17] 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», *J. High Energy Phys.* **1407**, 079 (2014), arXiv:1405.0301 [hep-ph].

[18] G. Salam, M. Cacciari, G. Soyez, «The anti-k_T jet clustering algorithm», *J. High Energy Phys.* **0804**, 063 (2008), arXiv:0802.1189 [hep-ph].

[19] S. Mrenna, T. Sjöstrand, P. Skands, «PYTHIA 6.4 physics and manual», *J. High Energy Phys.* **0605**, 026 (2006), arXiv:hep-ph/0603175.

[20] H.L. Lai *et al.*, «Global QCD analysis of parton structure of the nucleon: CTEQ5 parton distributions», *Eur. Phys. J. C* **12**, 375 (2000).

[21] A. Giammanco *et al.*, «DELPHES 3, A modular framework for fast simulation of a generic collider experiment», *J. High Energy Phys.* **1402**, 057 (2014), arXiv:1307.6346 [hep-ex].

[22] A. Hoecker *et al.*, «TMVA: Toolkit for Multivariate Data Analysis», *PoS* **(ACAT)**, 040 (2007), arXiv:physics/0703039.

[23] G. McGregor *et al.*, «Boosted decision trees as an alternative to artificial neural networks for particle identification», *Nucl. Instrum. Methods Phys. Res. A* **543**, 577 (2005), arXiv:physics/0408124.