



Isolating semi-leptonic $H \rightarrow WW^*$ decays for Bell inequality tests

Federica Fabbri^{1,2,3,a} , James Howarth^{1,b} , Théo Maurin^{1,c}

¹ SUPA - School of Physics and Astronomy, University of Glasgow, Glasgow, United Kingdom

² Dipartimento di Fisica e Astronomia A. Righi, Università di Bologna, Bologna, Italy

³ INFN Sezione di Bologna, Bologna, Italy

Received: 9 August 2023 / Accepted: 19 December 2023
© The Author(s) 2024

Abstract We present a method for identifying $H \rightarrow WW^* \rightarrow \ell\nu jj$ events in the presence of large Standard Model backgrounds and illustrate how this decay mode may be applied to the study of Bell-type Inequalities. Our findings reveal the feasibility of complete reconstruction of such Higgs decays and the efficacy of our suggested combination of selection criteria in effectively mitigating the otherwise overwhelming backgrounds. Our approach is based on a combination of bottom and charm tagging, alongside innovative reconstruction techniques. A realistic simulation based on publicly available object identification, reconstruction, and tagging efficiencies from the ATLAS experiment is used to explore the potential sensitivity to violations of the Collins-Gisin-Linden-Massar-Popescu (CGLMP) inequality in existing and expected future data collected at the Large Hadron Collider (LHC). It is found that, the proposed method provides a powerful means of distinguishing the Higgs decay mode from the background, allowing us to achieve an expectation of approximately 3σ significance in detecting violations of these inequalities with 300 fb^{-1} of data, soon-to-be collected by the LHC.

1 Introduction

The production of Higgs bosons decaying to a pair of W bosons has been observed by the ATLAS and CMS collaborations at $\sqrt{s} = 8 \text{ TeV}$, and more recently has been probed at $\sqrt{s} = 13 \text{ TeV}$ by both collaborations. The total cross-section is measured to a precision of 12%/12% in the gluon-gluon fusion production mode and to 25%/39% in the associated vector boson production modes by ATLAS and CMS, respectively [1, 2]. However, in each of these results, only the

fully leptonic decay modes ($H \rightarrow WW^* \rightarrow \ell\nu\ell\nu$, where $\ell = e, \mu$ ¹) were considered due the significant Standard Model (SM) background processes that mimic this final state when considering cases where one or both W 's may decay hadronically. Due to the relatively low leptonic branching fraction of the W , these results have large statistical uncertainties, at the same level of the experimental and theoretical systematic uncertainties. Combining the currents results with events collected from final states not restricted to be fully leptonic would more than double the statistics available for the $H \rightarrow WW^*$ measurements. Another severe limitation of the dileptonic final states is the presence of two neutrinos, that makes high resolution reconstruction of the Higgs boson kinematics almost impossible. This limits the potential to measure phenomena unique to the Higgs boson with LHC data, such as a proposed measurement of quantum entanglement in a qutrit system [3]. In this article, we propose a method to identify and fully reconstruct the more prevalent semi-leptonic decay mode ($H \rightarrow WW^* \rightarrow \ell\nu jj$, where $j = q = c, s$) whilst also reducing the SM background to a manageable level. Furthermore, we demonstrate the utility of these techniques by investigating the potential to measure Bell Inequality violation in Higgs boson decays using a combination of data already collected at the LHC experiments during Run2 ($\sqrt{s} = 13 \text{ TeV}$) and data that are being collected in the emergent Run3 ($\sqrt{s} = 13.6 \text{ TeV}$) and the future High Luminosity LHC (HL-LHC).

2 Background suppression and event reconstruction

In the $H \rightarrow WW^*$ decay mode, one of the W bosons is necessarily off its mass-shell (indicated by an asterisk) due to

^a e-mail: federica.fabbri@cern.ch

^b e-mail: james.howarth@glasgow.ac.uk (corresponding author)

^c e-mail: t.maurin.1@research.gla.ac.uk

¹ For the duration of this article, charge conjugation is assumed unless explicitly stated otherwise. That is, W^+ and W^- will both be referred to as W , similarly for leptons and quarks.

the mass of the Higgs boson (~ 125 GeV) being significantly lower than twice the on-shell W mass (~ 160 GeV). All studies of the $H \rightarrow WW^*$ process have so-far focused on the fully leptonic final state ($\ell\nu\ell\nu$) since this is the only decay mode that is not overwhelmed by background events from other SM processes. In particular, the semi-leptonic decay mode ($H \rightarrow WW^* \rightarrow \ell\nu qq'$) is not normally considered as it is dwarfed by the W +jets process. The cross-section of semi-leptonic $H \rightarrow WW^*$ production is 4.3 pb ($m_H = 125.0$ GeV) [4, 5] but the cross-section of the $W(\rightarrow l\nu)$ +jets process is $\sim 42,000$ pb at next-to-leading-order accuracy in perturbative QCD [6]. In order to successfully isolate the Higgs signal, the W +jets process must be suppressed by 4 orders of magnitude.

The study of this specific final state offers advantages beyond the mere addition of a new channel to Higgs measurements, as it facilitates the full reconstruction of the Higgs system, a task that proves challenging in fully-leptonic $H \rightarrow WW^*$ decays.

In hadron collider experiments, neutrinos are not measured directly but their presence is inferred by requiring momentum conservation in the plane transverse to the beam direction. It is possible to determine the total p_x and p_y components of the sum of all neutrinos present in the event via missing transverse momentum ($P_{x/y}^{miss}$), but not their p_z or, in the case of multiple neutrinos, their individual p_x or p_y components. This presents a problem for the measurement of Bell inequalities proposed by Barr [3], because these require that the Higgs itself, and all of its decay particles, be fully reconstructed. At the time of writing, no dedicated studies exist on how to solve this problem for the fully-leptonic $H \rightarrow WW^*$ case. In the semi-leptonic case, if the leptonically-decaying W were assumed to be on-shell (and therefore, the hadronic W to be off-shell), the on-shell W mass could be used as an additional constraint, along with the $P_{x/y}^{miss}$, to reconstruct the leptonic W . Two additional hadronic jets in an event would then need to be paired under some criterion (such as $m(jj) < 80$ GeV) to reconstruct the off-shell hadronic W and both would need to combine to reconstruct a Higgs boson with a mass of 125 GeV. This approach would be similar to the techniques employed in the reconstruction of semi-leptonic $t\bar{t}$ decays [7]. However, whilst this approach would have a high efficiency on the $H \rightarrow WW^*$ signal, it would also select a large fraction of W +jets, which also contain an on-shell leptonic W and could easily have additional hadronic jets that satisfy whatever criterion were chosen for the Higgs reconstruction. The problem of isolating the signal Higgs with respect to the W +jets background would remain.

We propose instead to sacrifice using the $P_{x/y}^{miss}$ directly and the requirement that the leptonic W be on-shell and impose a stricter requirement on the other side of the event, i.e. that the hadronically decaying W boson be on-shell. We

then use a technique adapted from $t\bar{t}$ dileptonic final state reconstruction called Neutrino Weighting [8] to reconstruct the final state and, additionally, to also suppress the SM W +jets background.

2.1 Monte Carlo processes and detector effects

To determine the feasibility of differentiating signal $H \rightarrow WW^*$ events from background SM processes and of observing Bell's inequality violation in Higgs events under real experimental conditions, signal and background events are simulated using the POWHEG generator [9–12]. One million signal events are generated for semi-leptonically decaying $H \rightarrow WW^*$ events via the gluon-gluon fusion production mode [13]. Similarly, one million background events each are generated for $t\bar{t}$, W +jets and tW processes. In all cases, PYTHIA [14, 15] is used to perform the parton shower and hadronisation of the events. The only exception are the Diboson processes, generated using SHERPA [16]. Only processes that generate a single lepton in the final state are considered.

RIVET [17] is used to process the HEPMC [18, 19] events and reconstruct physics objects. Physics objects are constructed from stable particles in the event record (particles with a mean lifetime greater than 30 ps). Electrons and muons are identified as originating from a W boson decay (including via intermediate τ leptons) and all photon radiation within a cone of $\Delta R < 0.1$ (where $\Delta R^2 = \Delta\eta^2 + \Delta\phi^2$) is summed into the lepton four-momentum and removed from the event record (excluding photons from hadrons decays). Jets are reconstructed using the anti- k_t algorithm with a distance parameter of $R = 0.4$ and are tagged as having b or c flavour via ghost matching [20]. $P_{x/y}^{miss}$ is defined using only neutrinos originating from a W boson decay (including via intermediate τ leptons). Neutrinos and leptons deriving from hadron decays are included in the input to the jet clustering. These object definitions follow those typically used by the ATLAS and CMS collaborations during LHC Run2.

In order to simulate the effect of the detector response, publicly available acceptance and efficiency values from the ATLAS collaboration are applied to the physics objects. The trigger efficiencies for electrons [21] and muons [22] are based on the assumption of using single lepton unrescaled triggers (with a p_T threshold of 20 GeV). The major effect of these efficiencies is to effectively set the minimum p_T requirement for the leptons employed in the analysis. The lepton identification and reconstruction efficiencies for electrons and muons are taken to be 90% and 96.1%, respectively, based on the *medium* definitions used by ATLAS [23, 24]. We assume a jet reconstruction efficiency of unity and neglect the effects due to limited resolution. Jet smearing based on [25] was tested and was found to have a limited impact on the results described in the following sections. Furthermore, recent experimental results that demonstrate the significant

potential for improving jet resolution and response through the utilization of machine learning techniques [26] imply that these effects will be subdominant in analyses on real data. Moreover, in real-experiment analyses, the impact of jet resolution and response can be further mitigated by implementing dedicated regions and observables designed to constrain the jet calibration in-situ.

2.2 Neutrino weighting

Neutrino Weighting (NW) is a technique used to reconstruct leptonically decaying W bosons. It has been used to great effect in the dileptonic $t\bar{t}$ signature, most recently in measurements of spin correlation in $t\bar{t}$ events by ATLAS [27,28], but it requires some adaptation to be applicable to the $H \rightarrow WW^* \rightarrow \ell\nu jj$ signature. NW solves the problem of reconstructing W bosons that decay leptonically by introducing assumptions on the unconstrained parameters. In the case of dileptonic $t\bar{t}$, these assumptions are the values for the pseudo-rapidity of each of the two neutrinos in the final state. In our modification, these assumptions are the pseudo-rapidity of the neutrino (η_ν) and the off-shell mass of the leptonically decaying W boson ($m_{W(lep)}$). Since neither of these values are known a-priori, the full system must be reconstructed under many possible assumptions for $m_{W(lep)}$ and η_ν . With these two assumptions, as well as an assumption on the mass of the Higgs (which we take to always be 125 GeV) and the reconstructed mass of the hadronically decaying W , the system is fully constrained and can be solved for the neutrino’s momentum x and y components.² There are zero, one, or two potential real solutions (corresponding to the roots of the polynomial) for each assumption of $m_{W(lep)}$ and η_ν . For each assumption, a weight is generated based on the difference between the observed $P_{x/y}^{miss}$ and the reconstructed neutrino kinematics:

$$w = \exp\left(\frac{(v_x - P_x^{miss})^2}{\sigma_x^2}\right) \cdot \exp\left(\frac{(v_y - P_y^{miss})^2}{\sigma_y^2}\right), \quad (1)$$

where $v_{x/y}$ is the x/y component of the reconstructed neutrino’s momentum, $P_{x/y}^{miss}$ is the observed $P_{x/y}^{miss}$ in the experiment in the x/y direction, and $\sigma_{x/y}$ is the experimental resolution of $P_{x/y}^{miss}$ (the values for these resolutions only serve to scale the weight and have no physical impact on the result if taken equal in x and y). The resultant weight is high when the reconstructed neutrino matches the observed $P_{x/y}^{miss}$ and low when it does not and reaches a maximum when the chosen values for η_ν and m_W are closest to the true ones. An example of the distribution of weights for a single example

² The full details of this calculation and a python implementation are available from the authors upon request.

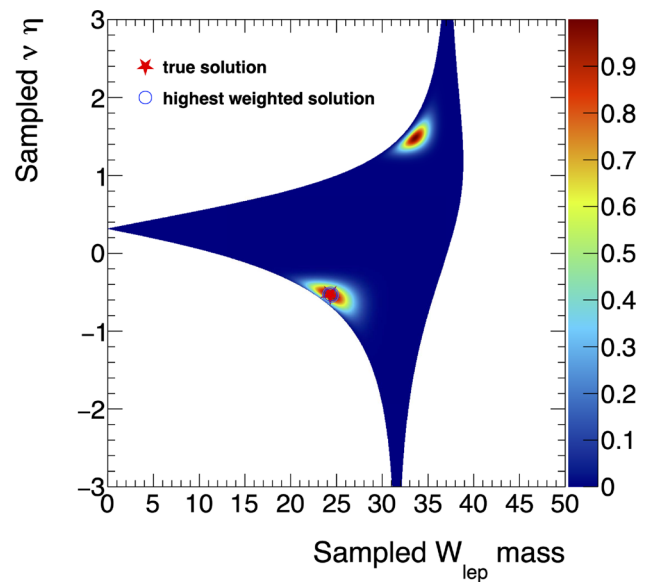


Fig. 1 The distribution of weights for each assumption of leptonic W boson mass and neutrino pseudo-rapidity, with markers indicating the location of the true values of these parameters and the scan point with the highest weight

collision event is shown in Fig. 1. As can be seen from Eq. (1), the weight is defined between 0 and 1, and the distribution peaks at the same values as the true $m_{W(lep)}$ and η_ν . One can also observe a second peak, unrelated to the true values of $m_{W(lep)}$ and η_ν corresponding to the second root. In the example event shown in Fig. 1, the solution with the highest weight lies at the same values as the true quantities. A discussion of the performance of the NW algorithm is reported in Sect. 2.4.

The weight from NW functions is more than just a method of selecting the most-likely values of $m_{W(lep)}$ and η_ν , it can also be used to suppress background processes with remarkable efficiency. This is discussed further in Sect. 2.4.

2.3 Charm tagging

Thanks to the enormous statistics provided by the LHC, it is possible to restrict the decay signature to events where the hadronic W boson decays to a charm and strange³ quark. As we have already assumed one on-shell W bosons decaying hadronically to perform the NW reconstruction, the major backgrounds to the $H \rightarrow WW^*$ signal arise from the production of a single leptonically decaying W boson with additional jets that arise from gluon radiation, from pairs of W bosons where only one boson decays hadronically and from

³ To the aim of the analysis reported here the exact flavour of the quark produced in association with the c is not relevant, given that all down-type quark have the same spin analyzing power.

Table 1 c and b -tagging efficiencies used in the ATLAS experiment [29,30]. The numbers refer to the probability for the tagging algorithms to assign the c or b flavour to jets containing a B -hadron (b -jet), a c -hadron and no B -hadrons (c -jet) or no heavy-flavour hadrons (l -jet). These numbers are directly used in the analysis presented in this paper to emulate realistic c -tagging performances

true flavour	c -tagging efficiency	b -tagging efficiency
b -jet	0.14	0.77
c -jet	0.4	0.2
l -jet	0.016	0.008

$t\bar{t}$ events, that contain two W bosons originating by the top quark decay.

Both the ATLAS and CMS experiments heavily utilise sophisticated algorithms to identify (tag) hadronic jets containing B -hadrons, which experiments use to infer the presence of b quarks in an event. The predominant feature used by these algorithms is the long lifetime of B -hadrons, leading to so called *displaced vertices* in the experiment inner trackers, a feature conspicuously absent from jets originating from so-called *light jets* (i.e. those initiated by gluons or u, d, s quarks). Jets originating from charm quarks occupy an interesting grey area as hadrons with charm flavour also have sufficiently long lifetimes to create displaced vertices and both ATLAS and CMS have created dedicated c tagging algorithms to identify them. This can be useful for reducing the W +jets background to our signal. The additional jets are more likely to be light jets from gluons or the light flavour quarks (such as $u\bar{u}$ or $d\bar{d}$) compared to heavier quarks pairs (such as $c\bar{c}$), whereas in our signal, the on-shell hadronic W decays to a ud pair or a cs pair with similar branching fractions. By requiring the presence of exactly one c -tagged jet, the SM background (which contains either none or two) can be reduced.

In this study, we simulate charm tagging using a technique called ‘ghost-matching’, where jets in Monte Carlo events are determined to have either b or c flavour by setting unstable B and D hadrons energies to zero and allowing them to be clustered into stable particle jets. Any jet with a clustered B hadron is determined to be a b -jet and any with a clustered D hadron and no B hadrons is tagged as a charm jet. Only hadrons with p_T above 5 GeV are considered. We then apply efficiency and mis-tag factors based on recent ATLAS values that are summarised in Table 1 [29,30].

For the $H \rightarrow WW^* \rightarrow cj\ell\nu$ we are particularly interested in identifying the following features:

- Exactly one c -tagged jet.
- One or more light jets (additional light jets could be radiated in the event).
- Exactly zero b -tagged jets.

The first two points are derived from the condition that one of the W bosons decays hadronically to a cs pair. The third point is necessary to reject the $t\bar{t}$ background process, which contains two b quarks in the final state. The second point requires the least consideration; all algorithms employed by LHC experiments are very efficient at rejecting light jets, on the other hand the interplay between b -tagging and c -tagging is non-trivial. The efficiency to mis-tag a b -jet for a c -jet, and vice-versa, can be relevant, as shown in Table 1.

2.4 Event selection

Based on the discussions in Sects. 2.3, the following criteria are used to select events:

- pre-selection:
 - Exactly 1 lepton with $p_T > 20$ GeV
 - Exactly 0 b -tagged jets
- c -tagging selection:
 - 2 or more jets, exactly one of which must be c -tagged.
 - At least 1 (c -jet, l -jet) pair with $|m_{cl} - 80.6| < 10$ GeV
- A reconstructed leptonic W boson from NW with $w > 0.7$.
- Maximum 2 light jets.
- Invariant mass of the lepton and the c -tagged jet $m(\ell c) < 80$ GeV.

If there are multiple combinations of light-jets and c -jet, the pair with mass closest to the $m_W = 80.4$ GeV is chosen as reconstructed hadronic W and employed in the NW. Using this selection we consider two charm tagging possibilities; an idealised case (with perfect identification efficiency and perfect b and light jet rejection) and a realistic case with $\epsilon_c = 40\%$, shown in Table 1. The latter is representative of the possibility with state-of-the-art charm tagging, whereas the former is provided as an indication of the best-case-scenario if significant improvements could be made. In all of the fol-

Table 2 The expected yields for an integrated luminosity of 300 fb^{-1} in a selection with idealised charm tagging and Run2-like charm tagging. The signal over the signal plus background ($S/(S+B)$) is also shown. The uncertainties are statistical only

Process	Idealised	$\epsilon_c = 40\%$
W + jets	13131 ± 785	10444 ± 664
WW	2298 ± 31	1137 ± 22
$t\bar{t}$	601 ± 76	1453 ± 119
tW	217 ± 8	350 ± 11
Higgs	5967 ± 76	2843 ± 56
$S/(S+B)$	0.27	0.18

lowing tables and figures the realistic charm tagging case is used unless otherwise specified.

Given that the mass of the hadronic W is an ingredient of the NW, the charm tagging also slightly influences the performance of the NW. The algorithm has indeed a 78% probability of finding a solution in signal events in the idealised case, while the probability is 72% in the realistic case. The resolution of m_W mass reconstruction is slightly below 10% in the idealised case and slightly above in the realistic one. The performance of the NW can be further optimised by fine tuning the number of points scanned in the η_ν and m_W plane.

The expected fraction of the signal (S) compared to the full simulation, including backgrounds (B), as a function of the weight from NW is presented in Fig. 2. Here the NW requirement is applied on top of the pre-selection and c -tagging requirements. Applying only the pre-selection, the fraction of signal over the entire simulated sample is ~ 0.0001 . The c -tagging requirements help in selecting the signal, increasing the $\frac{S}{S+B}$ by a factor ten, but it is only the constraint on the NW score that allows to really isolate the signal. The requirement at 0.7 on the NW score has indeed a 0.45 efficiency on the signal, against a 0.005 efficiency on the backgrounds. On top of the NW score request two additional selection requirements are applied, mainly targeting the reduction of $t\bar{t}$ background. To enter the selection the two b -jets included in the $t\bar{t}$ process must be mis-reconstructed as light jets or c -tagged jets. The limit on the number of jets remove events in the first category while selecting on $m(\ell c)$ removes events from the second category. The $m(\ell c)$ distribution is indeed related to the mass of the parent particle, so it shows a very distinct distribution in the $t\bar{t}$ case, where the lepton and the c -tagged jet (mis-reconstructed b -jet) derive from a top-quark decay. The expected composition of the data after the pre-selection and c -tagging requirements and after the whole selection is presented in Fig. 3. Even with the requirements on the hadronic W mass and the veto on b -tagged jets, the W +jets and $t\bar{t}$ processes significantly dominate over the signal before the requirement on the weight from NW is applied. After the whole selection, the contribution from $t\bar{t}$ and the W +jets processes is drastically reduced and the signal is well visible.

The expected number of signal and background events for 300 fb^{-1} of $\sqrt{s} = 13 \text{ TeV}$ data (the total expected Run2 + Run3 luminosities for ATLAS and CMS) are presented in Table 2. 2843 signal events are expected under realistic charm-tagging conditions, with a signal purity of 18%. This is a remarkably high purity given that the cross-section for the background process (W +jets in particular) are orders of magnitude higher than the signal process. Under perfect charm-tagging performance, these numbers increase to 5967 signal events and a purity of 27%. Though such performance is unlikely to be achievable, these values serve to act as an upper bound on the performance of the techniques

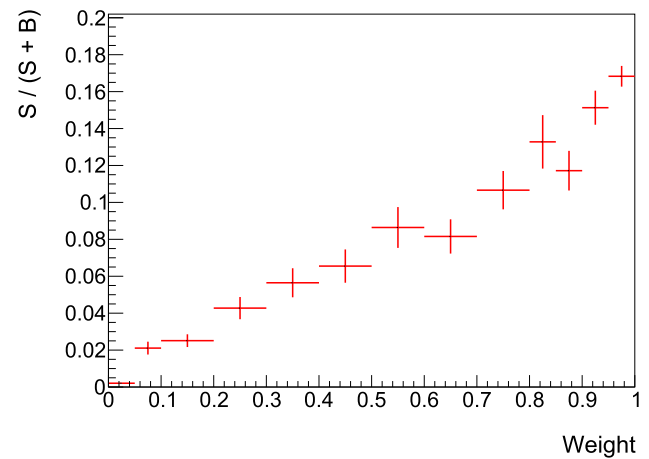


Fig. 2 The signal over signal plus background as a function of the requirement applied on the score from the Neutrino Weighting. Only events passing the pre-selection and the c -tagging selection are included in the plot

described in this paper. When including realistic jet smearing the purity degrades to 13% for the realistic charm tagging case and this is likely a fair representation of what is achievable under LHC Run2 conditions. However, advanced analysis techniques, such as multivariate approaches, could be employed to maximise the separation between signal and background exploiting other observables that manifest a different behavior between the signal and the main backgrounds, e.g. $\Delta\phi(\ell, s)$ shown in Fig. 3 (bottom).

3 Observing bell-inequality violation in Higgs events

Many Bell-type inequalities exist for a variety of different scenarios. For the $H \rightarrow WW^*$ process, which is a pair of three-outcome spin states (qutrit), the Collins–Gisin–Linden–Massar–Popescu (CGLMP) inequality has been shown to be the optimal choice [31,32]. The applicability of this inequality to fully-leptonic $H \rightarrow WW^*$ events has already been discussed in previous works. We follow the strategy proposed by Barr [3] with modifications only to account for the semi-leptonic final state. This is possible because the spin analysing power of the down-type quarks deriving by the W decay is 1 as in the charged leptons case. As a consequence the technique introduced in [3] can be applied also to the case presented here using the lepton and the down-type quark instead of the charged lepton pair to reconstruct the observables of interest. The down-type quark employed in the following is identified with the light jet used to reconstruct the hadronic W .

The CGLMP inequality (\mathcal{I}_3) may be constructed from the expectation values of operators (\mathcal{B}) constructed from angular observables based upon a choice of three orthonormal axes

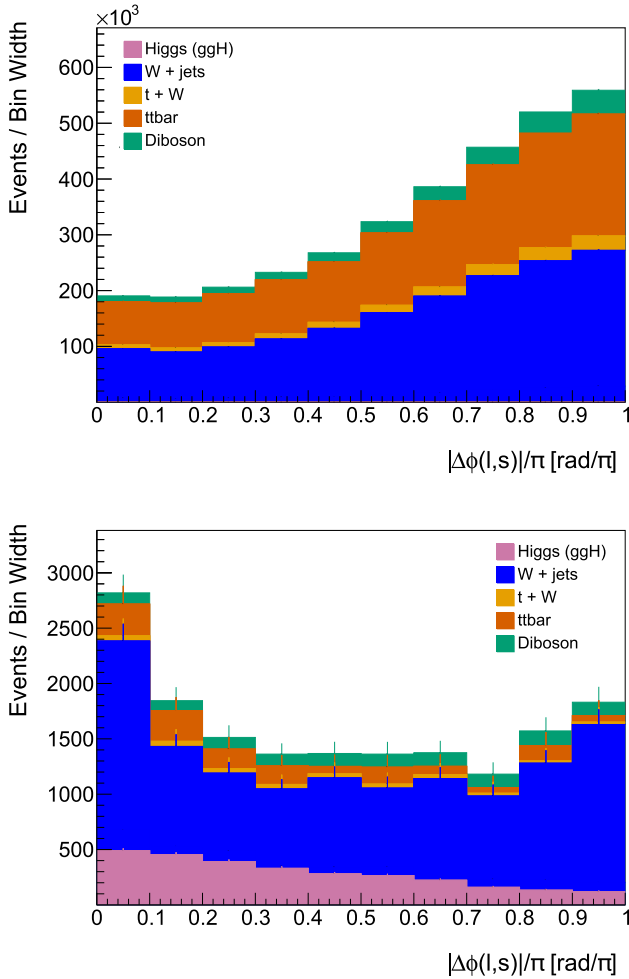


Fig. 3 Number of events as a function of the azimuthal angle between the lepton and the down-type quark ($\Delta\phi(l, s)$) applying only the pre-selection and c -tagging requirements (top) and applying the whole selection (bottom)

($\hat{x}\hat{y}\hat{z}$). These axes are defined based on the direction of the W^+ boson momenta (\hat{x}) in the W^+W^- rest frame, the direction perpendicular to the plane formed by \hat{x} and the direction of the incoming protons (\hat{y}), and the right-handed remaining orthogonal direction (\hat{z}) (a more in-depth discussion on these definitions and their similarity to those used in top spin measurements is provided in [3]). The CHLMP inequality is therefore defined as:

$$\mathcal{I}_3^{xyz} = \max(\langle \mathcal{B}_{\text{CGLMP}}^{xy} \rangle, \langle \mathcal{B}_{\text{CGLMP}}^{yz} \rangle, \langle \mathcal{B}_{\text{CGLMP}}^{zx} \rangle), \tag{2}$$

and this inequality is violated when:

$$\mathcal{I}_3^{xyz} > 2. \tag{3}$$

The operators are themselves constructed from the products of cosines of spin analysers of the parent W boson (charged leptons and strange-type jets in this topology). For example,

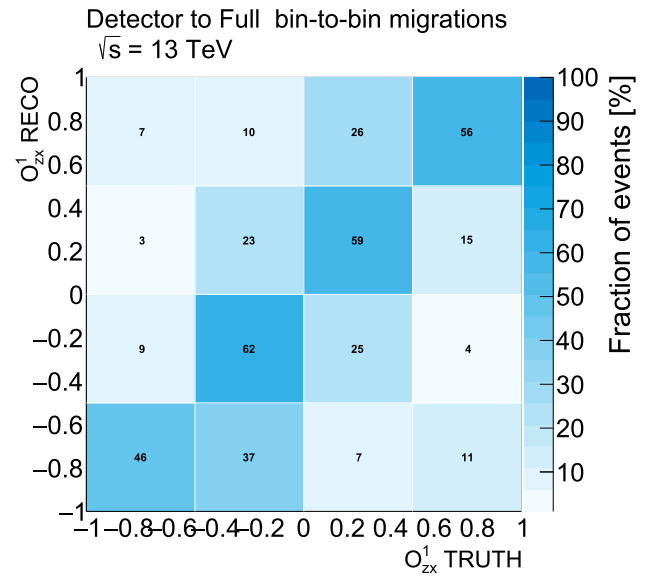


Fig. 4 Migration matrices for one of the three observables composing the expectation value of the zx Bell Operator

$\mathcal{B}_{\text{CGLMP}}^{xy}$, following the formalism in [3] is formed from the summation of the averages of three observables:

$$\mathcal{O}_{xy}^1 = \frac{8}{\sqrt{3}} \langle O_{xy}^1 \rangle$$

$$O_{xy}^1 = \xi_x^+ \xi_x^- + \xi_y^+ \xi_y^-$$

$$\mathcal{O}_{xy}^2 = 25 \langle O_{xy}^2 \rangle$$

$$O_{xy}^2 = ((\xi_x^+)^2 - (\xi_y^+)^2)((\xi_x^-)^2 - (\xi_y^-)^2)$$

$$\mathcal{O}_{xy}^3 = 100 \langle O_{xy}^3 \rangle,$$

$$O_{xy}^3 = \xi_x^+ \xi_y^+ \xi_x^- \xi_y^-,$$

where ξ_{xyz}^\pm is the cosine of the angle between the spin analyser from the parent W^\pm boson and the spin analysing axis $\hat{x}\hat{y}\hat{z}$. Similar sets of three observables exist for the $\mathcal{B}_{\text{CGLMP}}^{yz}$ and $\mathcal{B}_{\text{CGLMP}}^{zx}$ operators. In each event the sign of the two W is determined by the sign of the charged lepton. Each of the observables is constructed event-by-event and used to fill histograms from which the means are later determined. The means are then used to construct the relevant operator and, hence, determine the value for the inequality.

3.1 Unfolding

In order to correct for detector effects, many collider measurements use a technique called unfolding. We replicate this technique to remove the smearing effects imposed by our identification, efficiency, and reconstruction process using a type of unfolding called Iterative Bayesian Unfolding [33]. In particular, we use the procedure described in several ATLAS top quark measurements [7,34] and implemented

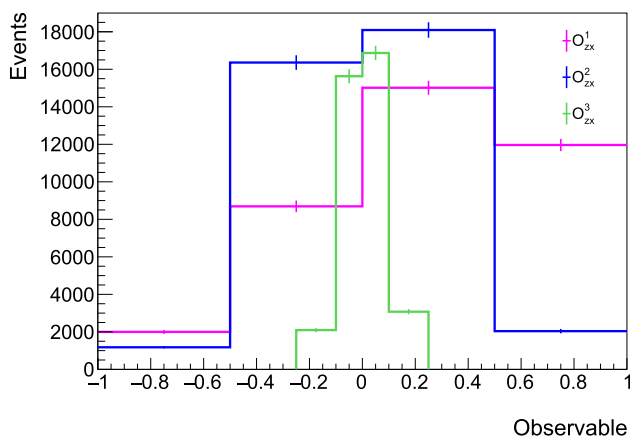


Fig. 5 Unfolded histograms for the three observables composing the expectation value of the z_x Bell Operator. The uncertainties on each bin are statistical

Table 3 The expected sensitivity for different values of integrated luminosity for a realistic and idealised case. The uncertainties are statistical only

Luminosity [fb^{-1}]	$\langle \mathcal{B}_{\text{CGLMP}}^{z_x} \rangle$ (idealised)	Significance (idealised)
139	2.45 ± 0.25 (0.18)	1.8 (2.5)
300	2.45 ± 0.17 (0.12)	2.65 (3.75)
3000	2.45 ± 0.05 (0.04)	9.0 (11.25)

in the ROOUNFOLD package [35]. The nine observables of interest are unfolded using two iterations and extrapolated to the full semi-leptonic phase space, at which point the mean of each observable is extracted to determine $\langle \mathcal{B}_{\text{CGLMP}}^{xy} \rangle$, $\langle \mathcal{B}_{\text{CGLMP}}^{yz} \rangle$, and $\langle \mathcal{B}_{\text{CGLMP}}^{zx} \rangle$. With the choice of axes and selection described in Sect. 2.4, $\langle \mathcal{B}_{\text{CGLMP}}^{zx} \rangle$ is always the largest and thus is the value taken to determine \mathcal{I}_3^{xyz} . The normalised migration matrices for one of the observables used to build this operator is shown in Fig. 4. The chosen binning results in diagonal matrices and consequently a stable unfolding. The limitations of the reconstruction are also apparent and manifest as a smearing of the reconstructed quantity towards zero.

The unfolded observables employed to determine $\langle \mathcal{B}_{\text{CGLMP}}^{z_x} \rangle$ are shown in Fig. 5.

3.2 Expected precision

Using the event selection, reconstruction, and unfolding described in the previous sections, the expected measurement and sensitivity for \mathcal{I}_3 is determined for three different values of integrated luminosity at a centre of mass energy of 13 TeV: 139 fb^{-1} (corresponding to the amount of data collected by both ATLAS and CMS during LHC Run2), 300 fb^{-1} (corresponding to the expected data that will be collected by the combination of Run2 and Run3), and 3000 fb^{-1} (corresponding to the expected total integrated lumi-

nosity for HL-LHC). In all cases, the centre of mass energy is 13 TeV though, in practice, the LHC Run3 centre-of-mass energy is slightly higher and the energy for HL-LHC may reach 14 TeV. This small changes in the center of mass energy does not largely effect the numbers projected here, as the signal and background cross-sections are only mildly affected. Moreover it is possible to combine the results on the significance across the different center of mass energies. The uncertainties considered are only statistical, these are evaluated taking into account the unfolding procedure. The statistical uncertainty is estimated using pseudo-experiments, derived smearing every bin of the input distributions with a Poisson distribution. The standard deviation of the unfolded pseudo-experiments in each bin is taken as statistical uncertainty and then propagated to the histograms mean and to \mathcal{I}_3^{xyz} .

The expected measured values and significances for each of these luminosity scenarios are presented in Table 3. The central value is 2.45, indicating violation of the \mathcal{I}_3 inequality. The expected uncertainties and significances of these values are presented for the charm tagging efficiency described in Sect. 2.3 and for an idealised case. Though such a case is not realistic for the existing LHC experiments, it provides a useful upper-bound on the results. For 300 fb^{-1} the expected significance is slightly below 3 sigmas in the realistic case, and above it in the idealised case, suggesting that the evidence could be reached by the LHC experiments already at the end of Run3. In the HL-LHC scenario the expected significance rises to 9 sigmas in the realistic scenario and to 11 in the idealised one. The impact of the assumption of perfect jet reconstruction was tested here using performances derived from ATLAS public jet performance results [25] for the case with non-idealised charm tagging, where the significances are slightly diluted to 1.6, 2.4, and 7.5 sigma for 130 fb^{-1} , 300 fb^{-1} , and 3000 fb^{-1} , respectively.

4 Conclusions

In this paper, we have demonstrated a method for isolating a previously inaccessible Higgs decay mode. The adaptation of Neutrino Weighting to semi-leptonic $H \rightarrow WW^*$ final state provides a new tool to separate the signal from the overwhelming background allowing to reach a signal purity of 18%, in combination with c -tagging and b -tagging. We have implemented a simple selection on the score of the NW, but the separation between signal and background could be further enhanced by introducing a multivariate analysis technique that combines the NW with other variables sensitive to the differences between $H \rightarrow WW^*$ and W +jets. Using the NW and the c -tagging combined with the W and Higgs mass constraints, we have fully reconstructed the final state and prototyped a measurement of the CGLMP Bell-type inequal-

ity in this topology. The results are promising and a significance of almost 3 sigma is expected with the soon-to-be collected luminosity of 300 fb^{-1} at the LHC, whilst observation could be reached using the data sample expected for HL-LHC. We found that the reconstruction of the off-shell leptonic W boson and the identification of the c -quark (and hence, s -quark) type jets are limiting factors for the measurement. Innovative charm tagging, with increased efficiency and an improved mis-tag rate for b -jets would represent a large improvement on the result both in terms of reduction of the statistical uncertainty and in the increase of the signal purity. This kind of improvement seems feasible in the near future, given recent advancements in flavour tagging algorithms [36,37]. We focused only on the CGLMP inequalities as a probe to study the quantum nature of the Higgs boson but the same approach could be applied to measure other entanglement witnesses valid for qutrit systems introduced in recent phenomenological studies [38]. This could facilitate observation of entanglement in $H \rightarrow WW^*$ events before reaching the statistic required to observe the violation of Bell inequalities.

Funding Funding was provided by Royal Society (URFR1191524).

Data Availability Statement This manuscript has no associated data or the data will not be deposited. [Authors' comment: This is a theoretical study and has no experimental data.]

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

Funded by SCOAP³.

References

1. ATLAS Collaboration, Measurements of Higgs boson production by gluon-gluon fusion and vector-boson fusion using $H \rightarrow WW^* \rightarrow e\nu\mu\nu$ decays in pp collisions at $\sqrt{s}=13$ TeV with the ATLAS detector. *Phys. Rev. D* **108**, 032005 (2023). <https://doi.org/10.1103/PhysRevD.108.032005>. arXiv:2207.00338 [hep-ex]
2. CMS Collaboration, Measurements of the Higgs boson production cross section and couplings in the W boson pair decay channel in proton-proton collisions at $\sqrt{s}=13$ TeV. (2022). arXiv:2206.09466 [hep-ex]
3. A.J. Barr, Testing Bell inequalities in Higgs boson decays. *Phys. Lett. B* **825**, 136866 (2022) <https://doi.org/10.1016/j.physletb.2021.136866>. arXiv:2106.01377 [hep-ph]
4. Particle Data Group, Review of Particle Physics. *Prog. Theor. Exp. Phys.* **8**, 083C01 (2020). <https://doi.org/10.1093/ptep/ptaa104> (ISSN: 2050-3911)
5. LHC Higgs Working Group Yellow Report. <https://twiki.cern.ch/twiki/bin/view/LHCPhysics/LHCHWG>. Accessed: 30 Sept 2022
6. 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**, 079 (2014). [https://doi.org/10.1007/JHEP07\(2014\)079](https://doi.org/10.1007/JHEP07(2014)079). arXiv:1405.0301 [hep-ph]
7. ATLAS Collaboration, Measurement of the $t\bar{t}$ production cross-section in the lepton+jets channel at $\sqrt{s}=13$ TeV with the ATLAS experiment. *Phys. Lett. B* **810**, 135797 (2020). <https://doi.org/10.1016/j.physletb.2020.135797>. arXiv:2006.13076 [hep-ex]
8. D0 Collaboration, Measurement of the top quark mass using dilepton events. *Phys. Rev. Lett.* **80**, 2063–2068 (1998). <https://doi.org/10.1103/PhysRevLett.80.2063>. arXiv:hep-ex/9706014
9. P. Nason, A new method for combining NLO QCD with shower Monte Carlo algorithms. *JHEP* **11**, 040. (2004). <https://doi.org/10.1088/1126-6708/2004/11/040>. arXiv:hep-ph/0409146
10. S. Frixione, P. Nason, C. Oleari, Matching NLO QCD computations with Parton Shower simulations: the POWHEG method. *JHEP* **11**, 070 (2007). <https://doi.org/10.1088/1126-6708/2007/11/070>. arXiv:0709.2092 [hep-ph]
11. S. Alioli et al., A general framework for implementing NLO calculations in shower Monte Carlo programs: the POWHEG BOX. *JHEP* **06**, 043 (2010). [https://doi.org/10.1007/JHEP06\(2010\)043](https://doi.org/10.1007/JHEP06(2010)043). arXiv:1002.2581 [hep-ph]
12. T. Ježo, P. Nason, On the treatment of resonances in next-to-leading order calculations matched to a parton shower. *JHEP* **12**, 065 (2015). [https://doi.org/10.1007/JHEP12\(2015\)065](https://doi.org/10.1007/JHEP12(2015)065). arXiv:1509.09071 [hep-ph]
13. S. Alioli et al., NLO Higgs boson production via gluon fusion matched with shower in POWHEG. *JHEP* **04**, 002 (2009). <https://doi.org/10.1088/1126-6708/2009/04/002>. arXiv:0812.0578 [hep-ph]
14. T. Sjöstrand et al., An introduction to PYTHIA 8.2. *Comput. Phys. Commun.* **191**, 159–177 (2015). <https://doi.org/10.1016/j.cpc.2015.01.024>. arXiv:1410.3012 [hep-ph]
15. C. Bierlich, et al., A comprehensive guide to the physics and usage of PYTHIA 8.3. (2022). <https://doi.org/10.21468/SciPostPhysCodeb.8>. arXiv:2203.11601 [hep-ph]
16. Sherpa Collaboration, Event Generation with Sherpa 2.2. *SciPost Phys.* **7**(3), 034 (2019). <https://doi.org/10.21468/SciPostPhys.7.3.034>. arXiv:1905.09127 [hep-ph]
17. A. Buckley et al., Rivet user manual. *Comput. Phys. Commun.* **184**, 2803–2819 (2013). <https://doi.org/10.1016/j.cpc.2013.05.021>. arXiv:1003.0694 [hep-ph]
18. M. Dobbs, J.B. Hansen, The HepMC C++ Monte Carlo event record for High Energy Physics Available via the following web-address: <http://home.cern.ch/mdobbs/HepMC/>. *Comput. Phys. Commun.* **134**(1), 41–46 (2001). [https://doi.org/10.1016/S0010-4655\(00\)00189-2](https://doi.org/10.1016/S0010-4655(00)00189-2) (ISSN: 0010-4655)
19. A. Buckley et al., The HepMC3 event record library for Monte Carlo event generators. *Comput. Phys. Commun.* **260**, 107310 (2021). <https://doi.org/10.1016/j.cpc.2020.107310>. arXiv:1912.08005 [hep-ph]
20. M. Cacciari, G.P. Salam, G. Soyez, The anti-kt jet clustering algorithm. *JHEP* **04**, 063 (2008). <https://doi.org/10.1088/1126-6708/2008/04/063>. arXiv:0802.1189 [hep-ph]
21. ATLAS Collaboration, Performance of electron and photon triggers in ATLAS during LHC Run 2. *Eur. Phys. J. C* **80**(1), 47 (2020). <https://doi.org/10.1140/epjc/s10052-019-7500-2>. arXiv:1909.00761 [hep-ex]
22. ATLAS Collaboration, Performance of the ATLAS muon triggers in Run 2. *JINST* **15**(09), P09015 (2020). <https://doi.org/10.1088/1748-0221/15/09/p09015>. arXiv:2004.13447 [physics.ins-det]

23. ATLAS Collaboration, Electron and photon performance measurements with the ATLAS detector using the 2015-2017 LHC proton-proton collision data. *JINST* **14**(12), P12006 (2019). <https://doi.org/10.1088/1748-0221/14/12/P12006>. arXiv:1908.00005 [hep-ex]
24. ATLAS Collaboration, Muon reconstruction performance of the ATLAS detector in proton-proton collision data at $\sqrt{s}=13$ TeV. *Eur. Phys. J. C* **76**(5), 292 (2016). <https://doi.org/10.1140/epjc/s10052-016-4120-y>. arXiv:1603.05598 [hep-ex]
25. ATLAS Collaboration, Jet energy scale and resolution measured in proton-proton collisions at $\sqrt{s}=13$ TeV with the ATLAS detector. *Eur. Phys. J. C* **81**(8), 689 (2021). <https://doi.org/10.1140/epjc/s10052-021-09402-3>. arXiv:2007.02645 [hep-ex]
26. ATLAS Collaboration, New techniques for jet calibration with the ATLAS detector (2023). arXiv:2303.17312 [hep-ex]
27. ATLAS Collaboration, Measurements of top-quark pair spin correlations in the $e\mu$ channel at $\sqrt{s}=13$ TeV using pp collisions in the ATLAS detector. *Eur. Phys. J. C* **80**(8), 754 (2020). <https://doi.org/10.1140/epjc/s10052-020-8181-6>. arXiv:1903.07570 [hep-ex]
28. D0 Collaboration, Measurement of the top quark mass using dilepton events. *Phys. Rev. Lett.* **80**, 2063–2068 (1998). <https://doi.org/10.1103/PhysRevLett.80.2063>. arXiv:hep-ex/9706014
29. ATLAS Collaboration, ATLAS b-jet identification performance and efficiency measurement with $t\bar{t}$ events in pp collisions at $\sqrt{s}=13$ TeV. *Eur. Phys. J. C* **79**(11), 970 (2019). <https://doi.org/10.1140/epjc/s10052-019-7450-8>. arXiv:1907.05120 [hep-ex]
30. ATLAS Collaboration, ATLAS flavourtagging algorithms for the LHC Run 2 pp collision dataset (2022). arXiv:2211.16345 [physics.data-an]
31. D. Kaszlikowski et al., Clauser-Horne inequality for three-state systems. *Phys. Rev. A* **65**, 032118 (2002). <https://doi.org/10.1103/PhysRevA.65.032118>.
32. D. Collins et al., Bell inequalities for arbitrarily high-dimensional systems. *Phys. Rev. Lett.* **88**, 040404 (2002). <https://doi.org/10.1103/PhysRevLett.88.040404>.
33. G. D'Agostini, A multidimensional unfolding method based on Bayes' theorem. *Nucl. Instrum. Methods Phys. Res. Sect. A Accel. Spectrom. Detect. Assoc. Equip.* **362**(2), 487–498 (1995). [https://doi.org/10.1016/0168-9002\(95\)00274-X](https://doi.org/10.1016/0168-9002(95)00274-X) (ISSN: 0168-9002)
34. ATLAS Collaboration, Measurements of top-quark pair spin correlations in the $e\mu$ channel at $\sqrt{s}=13$ TeV using pp collisions in the ATLAS detector. *Eur. Phys. J. C* **80**(8), 754 (2020). <https://doi.org/10.1140/epjc/s10052-020-8181-6>. arXiv:1903.07570 [hep-ex]
35. L. Brenner et al., Comparison of unfolding methods using RooFitUnfold. *Int. J. Mod. Phys. A* **35**(24), 2050145 (2020). <https://doi.org/10.1142/S0217751X20501456>. arXiv:1910.14654 [physics.data-an]
36. CMS Collaboration, A new calibration method for charm jet identification validated with proton-proton collision events at $\sqrt{s}=13$ TeV. *JINST* **17**(03), P03014 (2022). <https://doi.org/10.1088/1748-0221/17/03/P03014>. arXiv:2111.03027 [hep-ex]
37. ATLAS Collaboration, Neural Network Jet Flavour Tagging with the Upgraded ATLAS Inner Tracker Detector at the High-Luminosity LHC. Tech. rep. Geneva: CERN (2022). <http://cds.cern.ch/record/2839913>
38. J. A. Aguilar-Saavedra, et al., Testing entanglement and Bell inequalities in $H \rightarrow ZZ$. *Phys. Rev. D* **107**(1), 016012 (2023). <https://doi.org/10.1103/PhysRevD.107.016012>. arXiv:2209.13441 [hep-ph]