Boosting the $H \rightarrow \text{invisibles}$ searches with $Z$ boson polarization

Dorival Gonçalves\textsuperscript{1}

PITT PACC, Department of Physics and Astronomy,
University of Pittsburgh, 3941 O’Hara St., Pittsburgh, PA 15260, USA

Junya Nakamura\textsuperscript{2}

Institut für Theoretische Physik, Universität Tübingen, 72076 Tübingen, Germany

It is argued that, in $H \rightarrow \text{invisibles}$ searches with $Z(\ell\ell)H$ associated production at the LHC, the signal efficiency can be sensibly improved via a detailed study of the $Z$ boson polarization, discriminating between the signal and the dominant-irreducible $Z(\ell\ell)Z(\nu\nu)$ background. We first present a comprehensive polarization study, obtaining the complete set of angular coefficients $A_i$ in the Collins-Soper reference frame and identifying the dominant phenomenological effects. Then, we show the results for a realistic Monte Carlo study to $H \rightarrow \text{invisibles}$, taking the polarization analysis into account. We obtain about 20% improvement in the upper bound for the branching ratio of the Higgs boson to invisible particles, assuming 300 fb$^{-1}$ of data at the 13 TeV LHC.

I. INTRODUCTION

Bounding the invisible decay rate of the observed Higgs boson is one of the major targets of the LHC programme \cite{1,2}. While the Standard Model (SM) predicts a very small rate $\text{BR}_{H \rightarrow \text{inv}} \simeq 0.1\%$ \cite{3}, there are many extensions of the SM, referred to as Higgs Portal models \cite{4,5,6,7}, that predict a significantly larger $\text{BR}_{H \rightarrow \text{inv}}$. Therefore, the observation of invisible Higgs boson decays above the small SM rate would be a smoking gun signature for physics beyond the SM and could be the first direct evidence for the underlying microphysics of the Dark Sector.

Direct searches for invisible decays of the Higgs boson have been actively conducted at the large hadron collider (LHC) by the ATLAS and CMS Collaborations in several Higgs production channels \cite{8,9}. From combinations of these searches, the current upper bounds are $\text{BR}_{H \rightarrow \text{inv}} < 25\%$ by ATLAS \cite{10} and $\text{BR}_{H \rightarrow \text{inv}} < 24\%$ by CMS \cite{11} at the 95\% confidence level, where the SM Higgs production cross sections are assumed. The $ZH$ associated production, in which the $Z$ boson decays to a charged lepton pair, either electron or muon, provides significant constraints to $\text{BR}_{H \rightarrow \text{inv}}$ on its own \cite{12,13}. The signature is characterized by a large missing transverse momentum recoiling against a charged lepton pair that reconstructs the $Z$ boson mass. The dominant background after signal extraction is $ZZ$ \cite{11,14,15}, where one of the $Z$ bosons decays to a charged lepton pair and the other decays to neutrinos, hence it is an irreducible background. Yet, it is possible to measure the polarization of the $Z$ boson from angular distribution of the charged lepton pair for both the $ZH$ signal and the $ZZ$ background.

In this paper, we study in detail the possibility of enhancing the $ZH$ signal significance by making use of the difference in $Z$ boson polarization between the signal and the dominant $ZZ$ background, following the approach presented in Ref. \cite{20}. Although this information is disregarded in the present experimental analyses \cite{8,9}, we show that the proposed method can be a key ingredient to pin down the $H \rightarrow \text{invisibles}$ signal in the $Z(\ell\ell)H$ channel, separating the signal and background underlying production dynamics more accurately.

The paper is organized as follows. In Section II we show that the signal and the $ZZ$ background predict different states of $Z$ polarization and how we use this information for our purpose. In Section III the effects of $Z$ polarization on observables that are constructed with the charged leptons are discussed. In Section IV we present the results of our analyses. In Section V we conclude.

II. $Z$ BOSON POLARIZATION

In general, the $Z \rightarrow \ell^+\ell^-$ decay angular distribution for the $pp \rightarrow Z(\ell^+\ell^-)+X$ process can be described as

$$\frac{1}{\sigma} \frac{d\sigma}{d\cos\theta d\phi} = 1 + \cos^2\theta + A_2(1 - 3\cos^2\theta) + A_3 \sin 2\theta \cos\phi + A_4 \sin^2\theta \cos 2\phi + A_5 \cos \theta + A_6 \sin \theta \cos \phi + A_7 \sin \theta \sin \phi + A_8 \sin 2\theta \sin \phi + A_9 \sin^2\theta \sin 2\phi,$$  \hspace{1cm} (1)

where $\theta (0 \leq \theta \leq \pi)$ and $\phi (0 \leq \phi \leq 2\pi)$ are the polar and azimuthal angles of the lepton ($\ell^-$) in the $Z$ boson rest frame. We follow the notation of Ref. \cite{20}. The eight coefficients $A_i$ ($i = 2 \text{ to } 9$) correspond, in the most general case, to the number of degrees of freedom for the polarization of a spin 1 particle and uniquely parameterize $Z$ boson polarization. In this context, the lepton angular distribution works as an analyser, probing the underlying production dynamics encoded in the $A_i$ coefficients.

The value of $A_i$ depends on a reference frame. We choose it following Collins and Soper (Collins-Soper

\textsuperscript{1}dorival@pitt.edu
\textsuperscript{2}junya.nakamura@itp.uni-tuebingen.de
Table I. Angular coefficients at percent unit, for \(Z(\ell^- \ell^+)H\) and \(Z(\ell^- \ell^+)Z\) at LO and NLO QCD, after the selection in Eq. 2. The statistical uncertainty at the one standard deviation for the last digit is shown in parentheses.

|        | \(A_2\) | \(A_3\) | \(A_4\) | \(A_5\) | \(A_6\) | \(A_7\) | \(A_8\) | \(A_9\) |
|--------|---------|---------|---------|---------|---------|---------|---------|---------|
| \(ZH\) (LO) | 0.03(6) | 0.2(1)  | -80.0(1)| -0.08(8)| -0.01(8)| 0.04(8) | 0.1(1)  | 0.1(1)  |
| \(ZH\) (NLO)| 1.7(1)  | 0.0(3)  | -75.0(3)| -0.1(2) | 0.6(2)  | -0.2(2) | -0.0(3) | 0.1(3)  |
| \(ZZ\) (LO) | 48.12(9)| 0.3(1)  | 41.0(1) | 0.0(1)  | 0.2(1)  | 0.1(1)  | -0.1(1) | -0.1(1) |
| \(ZZ\) (NLO)| 49.1(2) | 0.0(4)  | 40.8(4) | -0.3(3) | 4.8(3)  | 0.2(3)  | 1.0(4)  | 0.1(4)  |

This frame is well recognized and the angular coefficients for the Drell-Yan \(Z\) boson production have been measured by ATLAS \cite{22,23} and CMS \cite{24}.

To access the potential of a polarization analysis to boost the signal \(Z(\ell\ell)H\) (inv) discrimination against its leading backgrounds, in particular the dominant \(Z(\ell\ell)Z(\nu\nu)\) contributions, we calculate the coefficients \(A_i\) at the fixed leading-order (LO) and next-to-leading-order (NLO) QCD with MadGraph5_aMC@NLO \cite{25} for the 13 TeV LHC, applying the signal selections

\[
75 \text{ GeV} < m_{\ell\ell} < 105 \text{ GeV}, \quad p_{T\ell\ell} > 200 \text{ GeV}. \quad (2)
\]

The \(Z(\ell\ell)Z\) background takes into account its interference with \(\gamma(\ell\ell)Z\). The results are summarized, at percent unit, in Tab. I. The \(ZH\) signal and \(ZZ\) background lepton angular distributions are governed by \(A_4\) and \(A_{2,4}\), respectively. The QCD NLO corrections are visibly large only in \(A_4\) for \(ZH\) and in \(A_6\) for \(ZZ\). The non-zero \(A_6\) in \(ZZ\) introduces an asymmetry between \(\ell^-\) and \(\ell^+\), which can be confirmed by observing the sign change of the \(A_6\) term in Eq. 1 after interchanging \(\ell^-\) and \(\ell^+\) (i.e. \(\theta \rightarrow \pi - \theta\) and \(\phi \rightarrow \phi + \pi\)). This indicates that \(A_6\) does not contribute when we do not distinguish \(\ell^-\) and \(\ell^+\). We take this approach, in order to make analysis simpler. Because the difference in \(A_6\) between the signal and the background is not large, the loss of information on \(A_6\) represents a sub-leading effect. Consequently, both for the \(ZH\) signal and for the \(ZZ\) background, Eq. 1 effectively simplifies to

\[
\frac{1}{\sigma \ d \cos \theta \ d \phi} \frac{d\sigma}{d\cos \theta} = 1 + \cos^2 \theta + A_2(1 - 3 \cos^2 \theta) + A_4 \sin^2 \theta \cos 2\phi, \quad (3)
\]

where the angles \(\theta\) and \(\phi\) are defined in the restricted ranges \(0 \leq \theta \leq \pi/2\) and \(0 \leq \phi \leq \pi/2\) as a result of not distinguishing \(\ell^-\) and \(\ell^+\). They can be obtained from

\[
\begin{align*}
\cos \theta &= \frac{2[q^0 p_T^3 - q^3 p_T^0]}{Q \sqrt{Q^2 + |q_T|^2}}, \\
\cos \phi &= \frac{2 \sin \theta}{Q^2 |q_T|^2 \sqrt{Q^2 + |q_T|^2}},
\end{align*}
\]

where \(q^\mu = (q^0, q_T, q^3)\) and \(p_T^\mu = (p_T^0, \vec{p}_T, p_T^3)\) are four-momenta of the reconstructed \(Z\) boson and either of the leptons, respectively, in the laboratory frame and \(Q\) is the reconstructed \(Z\) invariant mass \(m_{\ell\ell}\) \cite{20}.

In Fig. 1, we show the \(p_{T\ell\ell}\) distributions for the coefficients in Eq. 3 calculated at LO, imposing the invariant mass cut in Eq. 2. The difference in \(A_4\) between \(ZH\) and
ZZ increases as the \( p_{T\ell\ell} \) grows up. Hence, the signal and background \( Z \rightarrow \ell^+\ell^- \) angular distributions become more distinct at the boosted regime, where they acquire, in particular, an extra characteristic \( \phi \) modulation. In this way, the polarization study dovetails nicely with the usual boosted strategy for the \( H \rightarrow \text{invisibles} \) search in the \( Z(\ell\ell)H \) channel.

In Fig. 3 we show the ratio of the normalized \( \langle \cos \theta, \phi \rangle \) distribution for the \( ZH \) process to that for the \( ZZ \) process at the LO, imposing the selections in Eq. 2. The large differences in \( A_2 \) and \( A_4 \) between signal and background result in phenomenologically relevant kinematic profiles in the two-dimensional \( \langle \cos \theta, \phi \rangle \) distribution, where the signal to background ratio is sensibly enhanced for \( \langle \cos \theta \rightarrow 1, \phi \rightarrow \pi/2 \rangle \) and suppressed for \( \langle \cos \theta \rightarrow 0, \phi \rightarrow 0 \rangle \). In Sec. IV we will show that this distribution can be a key element in boosting the signal from background discrimination.

### III. EFFECTS OF Z POLARIZATION ON LEPTON OBSERVABLES

Observables that are constructed by the leptons from the \( Z \) boson decay can be, in general, largely affected by the \( Z \) boson polarization. Here we illustrate it with three phenomenologically relevant observables: the transverse lepton momentum \( p_{T\ell} \), the rapidity separation \( \Delta y_{\ell\ell} \) (\( \geq 0 \)) and the azimuthal angle separation \( \Delta \phi_{\ell\ell} \) (\( 0 \leq \Delta \phi_{\ell\ell} \leq \pi \)), all in the laboratory frame. These observables are used in the signal selection of the ATLAS and CMS analyses [8, 9]. The transverse momentum \( p_{T} \) of the harder lepton (\( \ell_1 \)) and the softer lepton (\( \ell_2 \)) are written in terms of the angles \( \theta \) and \( \phi \) defined in the Collins-Soper frame as [20]

\[
p_{T\ell_{1(2)}} = \frac{1}{2} \left[ q_T^2 + Q^2 \sin^2 \theta + q_T^2 \sin^2 \theta \cos^2 \phi \pm 2 q_T \sqrt{Q^2 + q_T^2} \sin \theta \cos \phi \right],
\]

where \( q_T \) is the \( Z \) transverse momentum \( (q_T = |q_T|) \). In the same manner, \( \Delta \phi_{\ell\ell} \) and \( \Delta y_{\ell\ell} \) are given by

\[
e^{2\Delta y_{\ell\ell}} = \frac{(\sqrt{q_T^2 + Q^2} + Q \cos \theta)^2 - q_T^2 \sin^2 \theta \cos^2 \phi}{(\sqrt{q_T^2 + Q^2} - Q \cos \theta)^2 - q_T^2 \sin^2 \theta \cos^2 \phi},
\]

\[
\cos \Delta \phi_{\ell\ell} = \frac{q_T^2 (1 - \sin^2 \theta \cos^2 \phi) - Q^2 \sin^2 \theta}{q_T^2 (1 - \sin^2 \theta \cos^2 \phi) - Q^2 \sin^2 \theta} \left[ 1 + \frac{4q_T^2 Q^2 \sin^2 \theta \sin^2 \phi}{(q_T^2 (1 - \sin^2 \theta \cos^2 \phi) - Q^2 \sin^2 \theta)^2} \right]^{-\frac{1}{2}}.
\]

Their \( \langle \theta, \phi \rangle \) dependence in fact shows that these observables are sensitive to the \( Z \) polarization. Searches for \( H \rightarrow \text{invisibles} \) with the \( ZH \) production at the LHC are

* More precisely, \( \Delta R_{\ell\ell}^2 = \Delta \phi_{\ell\ell}^2 + \Delta y_{\ell\ell}^2 \) is used in the ATLAS analysis.
performed in high $q_T$ boosted regions \cite{8, 10}. In boosted regions $Q/q_T < 1$, the above observables can be expanded as
\begin{equation}
 p_{T\ell_{1}(2)} = \frac{1}{2} q_T \left( 1 \pm \sin \theta \cos \phi + O(Q^2/q_T^2) \right), \tag{6a}
\end{equation}

\begin{equation}
 e^{2\Delta y_{\ell\ell}} = 1 + \frac{4 \cos \theta}{1 - \sin^2 \theta \cos^2 \phi} \frac{Q}{q_T} + O(Q^2/q_T^2), \tag{6b}
\end{equation}

\begin{equation}
 \cos \Delta \phi_{\ell\ell} = 1 + O(Q^2/q_T^2). \tag{6c}
\end{equation}

The $(\theta, \phi)$ dependence vanishes in $\Delta y_{\ell\ell}$ and $\Delta \phi_{\ell\ell}$ in the limit $Q/q_T \to 0$, while it still survives in $p_{T\ell_{1}(2)}$. Therefore, among these three observables, only $p_{T\ell_{1}(2)}$ is sensitive to $Z$ polarization for highly boosted events. This can be confirmed in Fig. 3 in which we show the normalized distributions for $p_{T\ell_{1}}$ (left), $\Delta y_{\ell\ell}$ (middle) and $\cos \Delta \phi_{\ell\ell}$ (right), for $ZH$ and $ZZ$ at the LO, imposing the selections from Eq. 2. We observe a large difference only in the $p_{T\ell_{1}}$ distribution, which originates from the sizable difference in $Z$ polarization shown in Tab. 1 or in Figs. 1 and 2. To summarize, in high $q_T$ regions, only limited observables can be sensitive to $Z$ polarization and the $p_{T\ell_{1}}$ is one of them.

In Ref. \cite{20}, it is found that a higher lepton $p_T$ cut can improve the signal significance as we may expect from the $p_{T\ell_{2}}$ distribution in Fig. 3. However, it is also found that the highest signal significance can be achieved by setting the lepton $p_T$ cut as small as possible and analyzing the $(\cos \theta, \phi)$ distribution directly. For our polarization analysis based on the $(\cos \theta, \phi)$ distribution, it is best to soften the lepton selections, as they can generally disturb the $(\cos \theta, \phi)$ distribution \cite{25}. Thus, in our hadron level study presented in Sec. IV, we lower the lepton transverse momentum selection to $p_{T\ell} > 5 \text{ GeV}$.

In the next section, we will quantify the impact of the $Z$ polarization to the $Z(\ell\ell)H(\text{inv})$ analysis via a realistic Monte Carlo study. Instead of probing the $Z$ polarization indirectly via the observables $\Delta \phi_{\ell\ell}$, $\Delta y_{\ell\ell}$ or $p_{T\ell_{1}(2)}$, we will directly explore the lepton angular distribution $(\cos \theta, \phi)$.

\section{IV. RESULTS}

We now scrutinise the potential improvements from the polarization study to $pp \to Z(\ell\ell)H(\text{inv})$ analysis. This search is characterised by a boosted leptonic $Z$ boson decay, recoiling against large transverse missing energy from $H \to \text{inv}$ \cite{8, 10, 19}. The dominant backgrounds for this signature are $t\bar{t}$+jets, $Z/\gamma^*+jets$, and diboson pairs ($ZZ, WW \to \ell\ell\nu\nu$ and $ZW \to \ell\ell\nu\nu$).

Our signal and background samples are simulated with SHHERPA+OPENLOOPS \cite{27, 29}. The $ZH_{DY}$ Drell-Yan signal component, $t\bar{t}$ and diboson pair samples are generated with the MC@NLO algorithm \cite{30, 31}, and the $Z/\gamma^*+jets$ is generated up to two extra jet emissions at NLO with the MEPS@NLO algorithm \cite{32}. We also account for the loop-induced gluon fusion $ZH_{GF}$ signal component at leading order accuracy merged up to one extra jet emission via the CKKW algorithm \cite{19, 33, 35}. Spin correlations and finite width effects from vector bosons are accounted for in our simulation. Hadronization and underlying event effects are simulated \cite{36}.

We start the analysis requiring two same-flavour opposite sign leptons ($\ell = e, \mu$) with $p_{T\ell} > 5 \text{ GeV}$ and $|\eta_\ell| < 2.5$, within the $Z$-boson invariant mass window $m_{\ell\ell} - m_Z < 15 \text{ GeV}$. Events with extra leptons are vetoed. Since most of the signal sensitivity resides in the boosted kinematics, we require $E_T > 200 \text{ GeV}$.

Jets are defined with the anti-$k_T$ jet algorithm with radius $R = 0.4 \cite{37}$, $p_{Tj} > 30 \text{ GeV}$, and $|\eta_j| < 5$. To

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{Figure_4}
\caption{Signal (red) and background (blue/green) $E_T$ (left) and $\Delta \phi_{\ell\ell}$ (right) distributions. The background (signal) histograms are (non-)stacked. We apply all the selections described in the text but $E_T > 200 \text{ GeV}$ and $\Delta \phi_{\ell\ell} < \pi/2$ (left panel) and $\Delta \phi_{\ell\ell} < \pi/2$ (right panel). We consider the $\sqrt{s} = 13 \text{ TeV}$ LHC.}
\end{figure}
tame the $t\bar{t}$+jets background, we veto events with two or more jets or containing a $b$-jet. Our study assumes 70% $b$-tagging efficiency and 1% miss-tag rate. We further optimise the signal selection, requiring $\Delta\phi(\ell\ell, \not{p}_T^{miss}) > 2.8$, $|E_T - p_{T,\ell\ell}|/p_{T,\ell\ell} < 0.4$, transverse mass $m_T > 200$ GeV and $\Delta\phi_{\ell\ell} < \pi/2$, following the CMS analysis [8]. An additional selection $\Delta\phi(p_T^{miss}, j) > 0.5$ is implemented to the one-jet category to further suppress the $Z/\gamma^*+jets$ background [8]. The resulting signal and background $E_T$ distributions are displayed in Fig. [1](left). The $t\bar{t}$ and $Z/\gamma^*+jets$ backgrounds get rapidly depleted for large $E_T$, and the diboson contributions ($ZZ, WW \rightarrow \ell\ell\nu\nu$ and $ZW \rightarrow \ell\ell\nu\nu$ result as the leading background components.

While the selection $\Delta\phi_{\ell\ell} < \pi/2$ can further suppress some of the backgrounds, such as $t\bar{t}$, see Fig. [2] (right), it will have reduced impact in the dominant background process $Z(\ell\ell)Z(\nu\nu)$ at the boosted regime. The potential sensitivity in the $\Delta\phi_{\ell\ell}$ observable to separate the signal $Z(\ell\ell)H(\text{inv})$ from the background $Z(\ell\ell)Z(\nu\nu)$ channels could arise only from the $Z$ polarization, however this information is suppressed at the boosted kinematics, as discussed in Sec [III]. Conversely, the direct $(\cos\theta, \phi)$ analysis becomes even more powerful for large transverse momentum, see Fig. [1] hence this two-dimensional profile become a key ingredient to achieve more accurate limits.

To quantify the possible gains with the polarization study, we perform a binned log-likelihood analysis, invoking the CL$_S$ method [35] on the rate and compare with the analysis based on the $(\cos\theta, \phi)$ distribution. Our results assume 5% systematic uncertainty on the background rate modeled as a nuisance parameter. In Fig. [5], we show the 95% confidence level bound on the $Z(\ell\ell)H$ production times the invisible Higgs branching ratio $BR(H \rightarrow \text{inv})$ normalized by the SM $Z(\ell\ell)H$ production rate, $\sigma_{SM}$. The polarization study largely improves the $H \rightarrow \text{inv}$s bound and makes it less systematic limited at large collider luminosities. This is because of the larger signal over background ratio $S/B$ for a sizeable portion of the $(\cos\theta, \phi)$ parameter space. As a result, we can improve the bound from $BR(H \rightarrow \text{inv}) \lesssim 21\%$ to $\lesssim 17\%$ by adding the polarization analysis, assuming $L = 300$ fb$^{-1}$.

V. CONCLUSION

In this publication, we present a method to enhance the sensitivity on the $H \rightarrow \text{inv}$s searches with $Z(\ell\ell)H$ associated production at the LHC. The proposal relies on the accurate study of the $Z$ boson polarization to disentangle, with greater precision, the signal and background underlying production dynamics. We first calculate the complete set of angular coefficients $A_i$ in the $Z$ boson Collins-Soper frame at NLO QCD precision. The signal and background present very distinct coefficients, consequently, their $Z \rightarrow \ell\ell$ angular distributions display relevant phenomenologically differences. Performing a realistic Monte Carlo analysis, we show that these polarization effects can significantly improve the $H \rightarrow \text{inv}$ bounds. Assuming an integrated luminosity of 300 fb$^{-1}$ at the 13 TeV LHC, we achieve a Higgs to invisibles limit that is 20% stronger by including the polarization effects into the analysis. As this proposal relies only on lepton reconstruction, it presents small experimental uncertainties and can be promptly included in the ATLAS and CMS studies.

ACKNOWLEDGMENTS

DG was funded by U.S. National Science Foundation under the grant PHY-1519175. JN appreciates the support from the Alexander von Humboldt Foundation.
