Study of kinematic observables sensitive to the Higgs boson production channel in $pp \rightarrow Hjj$ process

N Belyaev$^1$, R Konoplich$^2$ and K Prokofiev$^3$

$^1$ National Research Nuclear University MEPhI (Moscow Engineering Physics Institute), Kashirskoe highway 31, Moscow, 115409, Russia
$^2$ Department of Physics, New York University, 4 Washington Place, New York, NY 10003, USA
$^3$ Department of Physics and Institute for Advanced Study, the Hong Kong University of Science and Technology

E-mail: NLBelyayev@mephi.ru, RK60@nyu.edu, Kirill.Prokofiev@cern.ch

Abstract. After the Higgs boson discovery at the LHC, a lot of additional measurements should be performed to understand in details the properties of the observed particle. These measurements include cross sections measurements, couplings measurements, studies of the interaction vertex structures etc. One of the most perspective subjects to study is the kinematics of the production jets, associated with the Higgs boson. It is demonstrated, that the kinematic correlations of such jets can be used to distinguish different production channels of Higgs boson: gluon-gluon fusion (ggF) and vector boson fusion (VBF). Such separation plays an important role because possible beyond the Standard Model contributions in ggF and VBF channels lead to different effects, which should be taken into account in searches for BSM physics.

1. Introduction

After the Higgs boson was discovered by the ATLAS and CMS collaborations at the LHC [1,2], the particle content of Standard Model (SM) became complete. Searches for physics beyond the Standard Model (BSM) are currently ongoing. Several impressive discoveries based on the LHC data were done in last years, namely the pentaquark observation [3], the observation of CP violation in rare baryon decays [4] and measurement of the Higgs boson mass [5].

Separation of gluon-fusion (ggF) and vector boson fusion (VBF) Higgs boson production channels was the main subject of study in many works [6–11]. However, due to the progress in neural networks and boosted decision tree (BDT) classifiers [12], there are plenty of new possibilities to improve the ggF-VBF separation even more. The separation of these channels is important because many BSM theories predict different effects for ggF and VBF vertices [13–16]. The separate investigation of both ggF and VBF production mechanisms will thus help to reveal the properties of discovered Higgs boson.

4 Also at Physics Department, Manhattan College, 4513 Manhattan College Parkway, Riverdale, New York, NY 10471, USA
2. Higgs model
One of the efficient ways to study possible BSM effects is to use the Effective field theory approach (EFT). This approach is based on the effective Lagrangian of particular model. Effective Lagrangian usually includes high dimensional operators which contribute to different vertices. In this study, ggF and VBF Higgs boson production channels are considered in the framework of the widely used Higgs Characterisation model [8] for the aMC@NLO Monte Carlo generator [17]. The tree-level HVV Lagrangian of this model can be written as following:

\[
\mathcal{L}_0^V = \left\{ c_\alpha \kappa_{SM} \left[ \frac{1}{2} g_{HZZ} Z_\mu Z^\mu + g_{HWW} W_\mu^+ W^-\mu \right] - \frac{1}{2} c_\alpha \kappa_{H\gamma\gamma} g_{H\gamma\gamma} A_{\mu\nu} A^{\mu\nu} + s_\alpha \kappa_{A\gamma\gamma} g_{A\gamma\gamma} A_{\mu\nu} \tilde{A}^{\mu\nu} \right\}
\]

where \( \Lambda \) is the new physics energy scale and the field strength tensors are defined as follows:

\[
V_{\mu\nu} = \partial_\mu V_\nu - \partial_\nu V_\mu \quad (V = A, Z, W^\pm),
\]

\[
G_{\mu\nu} = \partial_\mu G^a_{\nu} - \partial_\nu G^a_{\mu} + g_s f^{abc} G^b_{\mu} G^c_{\nu}.
\]

The dual tensor \( \tilde{V}_{\mu\nu} \) is defined as:

\[
\tilde{V}_{\mu\nu} = \frac{1}{2} \epsilon_{\mu\nu\rho\sigma} V^{\rho\sigma}.
\]

Lagrangian [1] contains CP-even and CP-odd Higgs boson states. Case \( c_\alpha = 1 \) corresponds to a pure Standard Model CP-even Higgs boson state while \( c_\alpha = 0 \) corresponds to a pure BSM CP-odd Higgs boson state. Higgs boson states with different CP-parities can be mixed by setting the \( c_\alpha \) parameter to some value inside the interval \((0, 1)\).

3. Higgs boson production
Kinematic distributions were studied at Monte Carlo level with the Next-to-leading order with quantum chromodynamics corrections (NLO QCD). Two initial data samples in the .lhe file format were generated with the aMC@NLO Monte Carlo generator for \( \sqrt{s} = 13 \) TeV proton-proton collisions. The first data set contains \( 10^5 \) SM ggF events with the following couplings configuration: \( c_\alpha = 1; k_{Hgg} = 1; k_{SM} = 1 \) and \( \Lambda = 10^3 \) GeV. The second data set contains \( 10^5 \) SM VBF events with the following couplings configuration: \( c_\alpha = 1; k_{SM} = 1; k_{H\gamma\gamma} = 1; k_{HZZ} = 1; k_{HWW} = 1 \) and \( \Lambda = 10^3 \) GeV. Feynman diagrams of Leading-order ggF and VBF processes are shown at figure [1]. After the .lhe files were generated, parton showering was applied by PYTHIA8 generator [18] with .stdhep files as the output.
At the last step, stdhep files were processed by DELPHES3 simulation package [19], which took into account simple detector effects of averaged Large hadron collider (LHC) detector. Data files produced by DELPHES3 simulation package were converted into .root format and then processed with the scientific framework ROOT [20].

4. Jet distributions
For this study, only reconstructed jets with transverse momentum $p_{T,jet} > 30$ GeV and pseudorapidity $\eta_j < 5$ were taken into account. Jet correlations were studied for the following kinematic variables:

- The invariant mass of the final state jets: $m_{jj}$;
- The transverse momentum of $X_0$: $p_{T,H}$;
- The number of jets in the event $N_{jets}$;
- The absolute value of difference between pseudorapidities of two jets: $|\Delta \eta_{jj}| = |\eta_{j1} - \eta_{j2}|$;
- The sine of the absolute value of half the difference between phi angle of two jets: $\sin\left(\frac{\phi_{j1} - \phi_{j2}}{2}\right)$;
- The angular separation between two jets: $\Delta R_{jj} = \sqrt{\Delta \eta_{jj}^2 + \Delta \phi_{jj}^2}$;
- The product of pseudorapidities of two jets: $\eta_{j1} \cdot \eta_{j2}$;
- The $\eta$-centrality of $H+2j$ system: $\eta$-centrality $= \frac{\eta_H - \eta_{j1} + \eta_{j2}}{\Delta \eta_{jj}}$;
- The $\vec{p}_T$-centrality of $H+2j$ system: $\vec{p}_T$-centrality $= \frac{|\vec{p}_{T,H} - \vec{p}_{T,j1} + \vec{p}_{T,j2}|}{|\Delta \vec{p}_{T,jj}|}$;
- The balance of transverse momentum of $H+2j$ system: $p_T$-balance $= \frac{|\vec{p}_{T,H} + \vec{p}_{T,j1} + \vec{p}_{T,j2}|}{|\vec{p}_{T,H}| + |\vec{p}_{T,j1}| + |\vec{p}_{T,j2}|}$;
- The normalized sum of jet momenta: $\xi_{jets} = \frac{p_{T,j1} + p_{T,j2}}{E_{j1} + E_{j2}}$;
- The thrust variable: $T = \max\left(\sum_{N_{jets}}^{F \cdot \vec{p}_{i,jet}} |\vec{p}| \right)$;

The thrust axis of a jet, $\hat{T}$, is defined as the direction which maximizes the sum of the longitudinal momenta of the energy clusters. Calculated distributions of the above-defined variables are shown at figure 2 and 3. The observables which are based on the transverse momentum of the
Higgs boson and outgoing jets have better separation power compared with the observables which are based on the angular variables.

![Jet distributions over different kinematic variables](image)

Figure 2. Jet distributions over different kinematic variables: $m_{jj}$ (a), $p_{T,H}$ (b), $|\Delta \eta_{jj}|$ (c), $\eta_1 \cdot \eta_2$ (d), $\Delta R_{jj}$ (e) and $\sin(|\Delta \Phi|)$ (f). Blue lines correspond to ggF channel, while the red ones correspond to VBF.

Variables $m_{jj}$ and $p_{T,H}$ are expected to have different shapes due to kinematics distinctions of parton fusion (ggF) and scattering (VBF) processes, but the effect is moderate. Variables $|\Delta \eta_{jj}|$ and $|\Delta R_{jj}|$ are more useful for the ggF-VBF separation due to greater difference in shapes. Variables $\sin\left(|\Delta \Phi|\right)$ and $\eta_1 \cdot \eta_2$ have limited sensitivity to the Higgs boson production channel.
and their impact on the separation performance should be tested with the BDT discriminators.

![Jet distributions over different kinematic variables: $\eta$-centrality (a), $p_T$-centrality (b), $N_{jets}$ (c), $T$ (d), $p_T$-balance (e) and $\xi_{jets}$ (f). Blue lines correspond to ggF channel, while the red ones corresponds to VBF.](image)

**Figure 3.** Jet distributions over different kinematic variables: $\eta$-centrality (a), $p_T$-centrality (b), $N_{jets}$ (c), $T$ (d), $p_T$-balance (e) and $\xi_{jets}$ (f). Blue lines correspond to ggF channel, while the red ones corresponds to VBF.

Variables $\eta$-centrality, $p_T$-centrality, $N_{jets}$, $\xi_{jets}$ and $p_T$-balance demonstrated in figure 3 have much more sensitivity than the ones shown in figure 2. Observables $p_T$-centrality and $p_T$-balance have qualitatively different shapes which makes them powerful discriminators. Tight cuts can be also set at $\eta$-centrality, $\xi_{jets}$ and $N_{jets}$ observables, while thrust variable $T$ have limited sensitivity. All the reviewed variables can be directly implemented into BDT classification algorithm, which will definitely improve the ggF-VBF separation.
5. Conclusion
In this paper, the Effective field theory approach was used in order to probe the variables sensitive to the Higgs boson production channel. Kinematic distributions were calculated for both ggF and VBF channels in Standard Model case. The obtained results demonstrate high potential for precise ggF-VBF separation based on the BDT algorithms. Reviewed kinematic variables can be also used to distinguish possible CP violation effects in the Higgs sector. The most important variables in terms of ggF-VBF separation are $\eta$-centrality, $\vec{p}_T$-centrality, $p_T$-balance and $\xi_{jets}$.

6. Acknowledgements
The work of N. Belyaev on the selection of discriminating variables was performed within the financial support from Russian Science Foundation - grant No.17-72-10021. The work of R. Konoplich is partially supported by the US National Science Foundation under Grant No.PHY-1402964. The work of K. Prokofiev is partially supported by a grant from the Research Grant Council of the Hong Kong Special Administrative Region, China (Project Nos. CUHK4/CRF/13G).

References
[1] Aad G et al. 2012 Phys. Lett. B 716 1–29
[2] Chatrchyan S et al. 2012 Phys. Lett. B 716 30–61
[3] Aad R et al. 2015 Phys. Rev. Lett. 115 072001
[4] Aad R et al. 2017 Nature Physics 13 391
[5] Aaboud M et al. 2017 Preprint ATLAS-CONF-2017-046
[6] Kruse A et al. 2015 Phys. Rev. D 91 053009
[7] Aaboud M et al. 2016 JHEP 11 112
[8] Maltoni F et al. 2014 Eur. Phys. J. C 74 3065
[9] Klamke G and Zeppenfeld D 2007 JHEP 04 052
[10] Aaboud M et al. 2016 Preprint ATL-PHYS-PUB-2016-008
[11] Khachatryan V et al. 2015 Phys. Rev. D 92 032008
[12] Bramer M 2007 Principles of Data Mining (London: Springer)
[13] Hagiwara K, Li Q and Mawatari K 2009 JHEP 07 101
[14] Hankele V et al. 2006 Phys. Rev. D 74 095001
[15] Agashe K, Contino R and Pomarol A 2005 Nucl. Phys. B 719 165–187
[16] Dolan M J et al. 2014 Phys. Rev. D 90 073008
[17] Alwall J et al. 2011 JHEP 06 128
[18] Ask S et al. 2014 Comp. Phys. Comm. 191 159–177
[19] Delaere C et al. 2014 JHEP 02 057
[20] Brun R et al. 2009 Comp. Phys. Comm. 180 2499–2512