PROCEEDINGS OF SCIENCE

Portraying Double Higgs at the Large Hadron Collider

Jeong Han Kim
Department of Physics, University of Notre Dame, 225 Nieuwland Hall Notre Dame, IN 46556, USA
E-mail: jkim73@nd.edu

Minho Kim
Department of Physics, POSTECH, 77 Cheongam-ro, Nam-gu, Pohang, 37673, Korea
E-mail: kmhmon@postech.ac.kr

Kyoungchul Kong*
Department of Physics and Astronomy, University of Kansas, Lawrence, KS 66045, USA
E-mail: kckong@ku.edu

Konstantin T. Matchev
Institute for Fundamental Theory, Physics Department, University of Florida, Gainesville, FL 32611, USA
E-mail: matchev@phys.ufl.edu

Myeonghun Park
Institute of Convergence Fundamental Studies and School of Liberal Arts, Seoultech, 232 Gongneung- ro, Nowon-gu, Seoul, 01811, Korea
E-mail: seoultech@seoultech.ac.kr

We examine the discovery potential for double Higgs production at the high luminosity LHC in the final state with two b-tagged jets, two leptons and missing transverse momentum. Although this dilepton final state has been considered a difficult channel due to the large backgrounds, we argue that it is possible to obtain a sizable signal significance, by adopting a deep learning framework making full use of the relevant kinematics along with the jet images from the Higgs decay. The proposed method can be easily generalized to the semi-leptonic channel of double Higgs production, as well as to other processes with similar final states.

Artificial Intelligence for Science, Industry and Society, AISIS2019
October 21-25, 2019
Universidad Nacional AutÃşnoma de MÃl'xico, Mexico City, MÃl'xico

*Speaker.

PoS(AISIS2019)013

1. Introduction

The discovery of the Higgs boson (h) with a mass $m_h = 125$ GeV [1,2] jumpstarted a comprehensive program of the precision measurements of all Higgs couplings [3]. The current results for the couplings to fermions and gauge bosons [4] appear to be in agreement with the SM predictions. However, probing the triple and quartic Higgs self-couplings is notoriously difficult [5–12]. Yet, the knowledge of those couplings is crucial for understanding the exact mechanism of electroweak symmetry breaking and the origin of mass in our universe. The measurement of double Higgs production and the associated triple Higgs coupling measurement is a guaranteed physics at the next run of the LHC program [3,13] among many possibilities [14,15].

The Higgs self-interaction is parameterized as $V = \frac{m_h^2}{2} h^2 + \kappa_3 \lambda_3^{\rm SM} v h^3 + \frac{1}{4} \kappa_4 \lambda_4^{\rm SM} h^4$, where $\lambda_3^{\rm SM} = \lambda_4^{\rm SM} = \frac{m_h^2}{2v^2}$ are the SM values, κ_3 and κ_4 parameterize deviations from those, and $v \approx 256$ GeV is the Higgs vacuum expectation value. In order to access κ_3 (κ_4), one has to measure the double (triple) Higgs boson production at the LHC with high luminosity (HL) or future colliders.

Due to the small signal cross-section (σ_{hh}), *it is necessary to combine as many different channels as possible to discover the double Higgs production and study the triple Higgs coupling* [13]. Among all possible channels, one specific process, $hh \to (b\bar{b})(W^\pm W^\mp)$, has so far been relatively overlooked, although it has the second largest branching fraction. This is mainly due to the large SM background cross-section $\sigma_{bknd} \sim 10^5 \sigma_{hh}$ (at the 14 TeV LHC), which is predominantly due to top quark pair production ($t\bar{t}$). In particular, there have been very few studies on the resulting dilepton final state [9–12,16]. The existing analyses employ sophisticated algorithms (neural networks (NN) [9], deep neural networks (DNN) [12,17], boosted decision tree (BDT) [10,16], etc.) to increase the signal sensitivity, but show somewhat pessimistic results, with a significance no better than 1σ at the HL-LHC with 3 ab^{-1} luminosity [9,10,12,16].

2. Improvement in the signal signal significance via kinematics and color-flow

There have been a lot of studies on the double Higgs but none of them emphasizes the potential role of the dilepton final state and it was considered to be one of worst channels. It was only very recent that this channel has been revisited. Ref. [18] proposes a novel kinematic method, which relies on two new kinematic functions, *Topness* and *Higgsness* [18]. They characterize features of the major ($t\bar{t}$) background and the signal (hh) events, respectively, rendering a way to disentangle two different event topologies. The method also utilizes two less commonly used variables, the subsystem M_{T2} (or subsystem M_2) [19–21] for $t\bar{t}$ production and the subsystem $\sqrt{\hat{s}}_{min}$ (or subsystem M_1) [21–23] for hh production. The basic idea behind this approach is very general but works the best with the double Higgs production in the dilepton channel, as described below.

For any given event, *Topness* [18,24] quantifies the degree of consistency to the dileptonic $t\bar{t}$ production, with 6 unknowns (the three-momenta of the two neutrinos, $\vec{p}_\nu$ and $\vec{p}_{\bar{\nu}}$) and four on-shell constraints, m_t, $m_{\bar{t}}$, m_{W^+} and m_{W^-}. An ansatz for the neutrino momenta can be obtained by minimizing the quantity

$$\chi^2_{ij} \equiv \min_{\vec{p}_T = \vec{p}_{\nu T} + \vec{p}_{\bar{\nu} T}} \left[\frac{\left(m^2_{b_i \ell^+ \nu} - m_t^2\right)^2}{\sigma_t^4} + \frac{\left(m^2_{\ell^+ \nu} - m_W^2\right)^2}{\sigma_W^4} + \frac{\left(m^2_{b_j \ell^- \bar{\nu}} - m_t^2\right)^2}{\sigma_t^4} + \frac{\left(m^2_{\ell^- \bar{\nu}} - m_W^2\right)^2}{\sigma_W^4} \right],$$

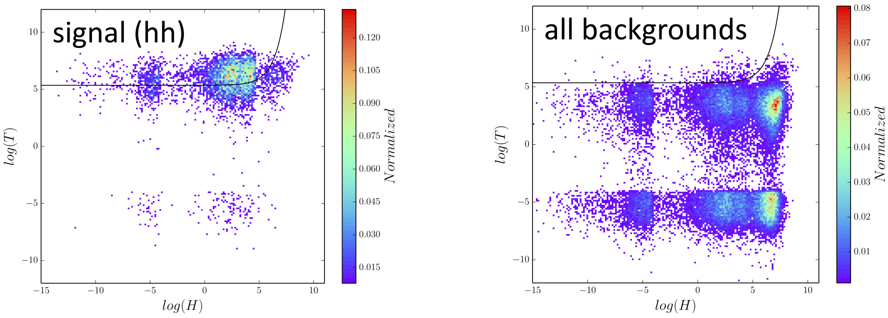

Figure 1: Two-dimensional correlation plots for Higgsness and Topness for signal (left) and all backgrounds (right). The solid curve represents a suitable cut to maximize the signal significance.

subjected to the missing transverse momentum constraint, $\vec{\not{p}}_T = \vec{p}_{\nu T} + \vec{p}_{\bar{\nu} T}$. Since there is a twofold ambiguity in the paring of a b-quark and a lepton, we define *Topness* as the smaller of the two χ^2s,

$$T \equiv \min\left(\chi^2_{12}, \chi^2_{21}\right) . \tag{2.1}$$

In double Higgs production, the invariant mass, m_{bb}, cut has been used to identify the Higgs decay ($h \to b\bar{b}$) and suppress the SM backgrounds. To characterize the decay of the other Higgs boson, $h \to WW^* \to \ell^+\ell^-\nu\bar{\nu}$, we introduce *Higgsness* [18] as follows:

$$H \equiv \min\left[\frac{\left(m^2_{\ell^+\ell^-\nu\bar{\nu}} - m^2_h\right)^2}{\sigma^4_{h_\ell}} + \frac{\left(m^2_{\nu\bar{\nu}} - m^2_{\nu\bar{\nu},peak}\right)^2}{\sigma^4_\nu} \right. \tag{2.2}$$

$$\left. + \, min\left(\frac{\left(m^2_{\ell^+\nu} - m^2_W\right)^2}{\sigma^4_W} + \frac{\left(m^2_{\ell^-\bar{\nu}} - m^2_{W^*,peak}\right)^2}{\sigma^4_{W_*}} , \; \frac{\left(m^2_{\ell^-\bar{\nu}} - m^2_W\right)^2}{\sigma^4_W} + \frac{\left(m^2_{\ell^+\nu} - m^2_{W^*,peak}\right)^2}{\sigma^4_{W_*}}\right)\right] ,$$

where m_{W^*} is the invariant mass of the lepton-neutrino pair resulting from the off-shell W. The m_{W^*} distribution has an end-point at around $m_h - m_W$, and its peak is located at

$$m^{peak}_{W^*} = \frac{1}{\sqrt{3}}\sqrt{2\left(m^2_h + m^2_W\right) - \sqrt{m^4_h + 14 m^2_h m^2_W + m^4_W}} . \tag{2.3}$$

Note also that $m^{peak}_{\nu\bar{\nu}} = m^{peak}_{\ell\ell} \approx 30$ GeV is the location of the peak in the $\frac{d\sigma}{dm_{\nu\bar{\nu}}}$ or $\frac{d\sigma}{dm_{\ell\ell}}$ distribution [18, 25]. The σ parameters in Eqs. (2.1) and (2.3) stand for the experimental uncertainties and intrinsic particle widths. In principle, they can be treated as free parameters, and tuned by a neutral network (NN) or a boosted decision tree (BDT) algorithms. In our numerical study, we choose $\sigma_t = 5$ GeV, $\sigma_W = 5$ GeV, $\sigma_{W^*} = 5$ GeV, $\sigma_{h_\ell} = 2$ GeV, and $\sigma_\nu = 10$ GeV. For detailed discussion in the rest of this article, we follow the same procedure described in Ref. [18] for event generation of signal and background, parton-shower / hadronization and semi-realistic detector effects.

Scatter distributions of *Higgsness* and *Topness* are shown in Fig. 1 for (left) the signal and (right) all the backgrounds ($t\bar{t}$, $t\bar{t}h$, $t\bar{t}V$, $\ell\ell bj$, $\tau\tau bb$ and others). The dominant $t\bar{t}$ events are expected to be on the lower-right corner with smaller *Topness* and larger *Higgsness*. The hh events are, on the other hand, expected to have smaller *Higgsness* and larger *Topness*. This motivates the use of a curve in the $(\log H, \log T)$ space as a cut in order to separate signal and backgrounds.

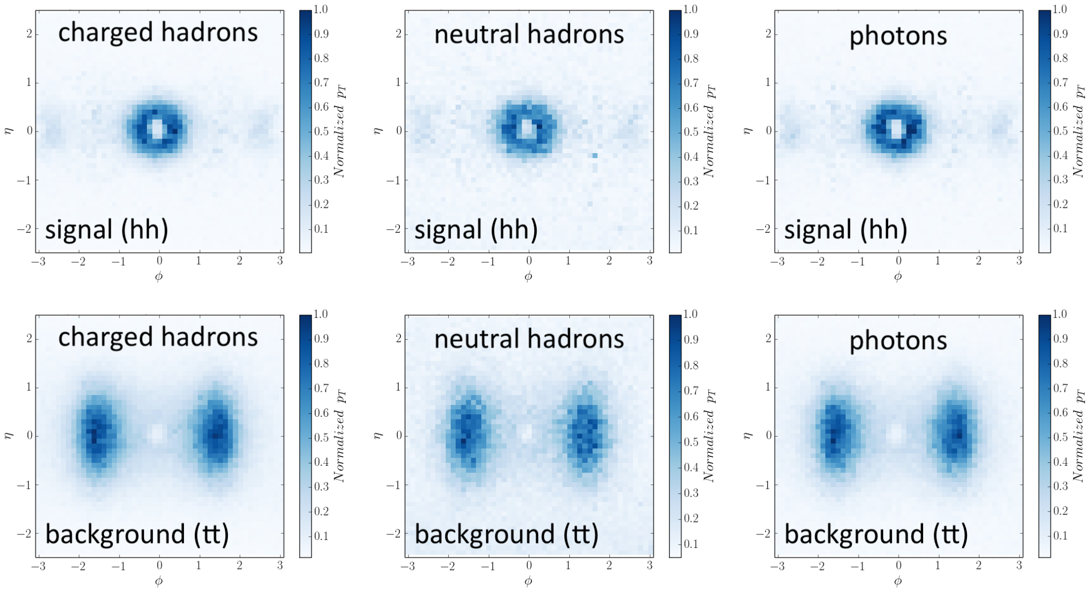

Figure 2: The (preliminary) cumulative average of the images for the signal (top) and the $t\bar{t}$ background (bottom). The origin of the (ϕ, η) plane is taken to be the center of the b quark pair and the density indicates the total p_T in each pixel. Images from the left to the right are obtained from charged hadrons (1st column), neutral hadrons (2nd), and photons (3rd).

Along with *Higgsness* and *Topness*, we employ the M_{T2} variables for the $b\bar{b}$ subsystem ($M_{T2}^{(b)}$) and the lepton subsystem ($M_{T2}^{(\ell)}$) [20], and the subsystem $\hat{s}_{min}^{(\ell\ell)}$ variable for $h \to W^{\pm}W^{*\mp} \to \ell^+\ell^-\nu\bar{\nu}$ [22, 23]. The M_{T2} is defined as

$$M_{T2}(\tilde{m}) \equiv \min\left\{\max\left[M_{TP_1}(\vec{p}_{\nu T}, \tilde{m}),\, M_{TP_2}(\vec{p}_{\bar{\nu}T}, \tilde{m})\right]\right\}, \tag{2.4}$$

where $\tilde{m}$ is the test mass for the daughter particle. The minimization among the transverse masses of the parent particles M_{TP_i} $(i = 1, 2)$ is performed over the transverse neutrino momenta $\vec{p}_{\nu T}$ and $\vec{p}_{\bar{\nu}T}$ subjected to the $\not{\vec{p}}_T$ constraint [19–21, 26–29]. In the case of $M_{T2}^{(b)}$, the two Ws play a role of two missing neutrinos. The vertical lines at $M_{T2}^{(b)} = 190$ GeV and $M_{T2}^{(\ell)} = 6$ GeV represent the optimized cuts, suppressing $t\bar{t}$ and $\tau\tau bb$ (Drell-Yan) backgrounds respectively.

The $\hat{s}_{min}^{(\mathrm{v})}$ variable [21–23] is defined as

$$\hat{s}_{min}^{(\mathrm{v})} = m_{\mathrm{v}}^2 + 2\left(\sqrt{|\vec{P}_T^{\mathrm{v}}|^2 + m_{\mathrm{v}}^2}\,|\not{\vec{p}}_T| - \vec{P}_T^{\mathrm{v}} \cdot \not{\vec{p}}_T\right), \tag{2.5}$$

where the script (v) represents a set of visible particles under consideration. The m_{v} and $\vec{P}_T^{\mathrm{v}}$ denote their invariant mass and transverse momentum, respectively. The $\hat{s}_{min}^{(\mathrm{v})}$ variable provides the minimum value of the Mandelstam invariant mass $\hat{s}$ which is consistent with the observed visible 4-momentum vector. The $\sqrt{\hat{s}_{min}^{(\ell\ell)}}$ distribution has an endpoint at around m_h for hh events. All other backgrounds, however, extend above this point. This justifies the use $\sqrt{\hat{s}_{min}^{(\ell\ell)}} < 130$ GeV as a cut to reduce the backgrounds. We observe that $\sqrt{\hat{s}_{min}^{(bb\ell\ell)}}(hh)$ provides a good measure of the true $\sqrt{\hat{s}}(hh)$, while $\sqrt{\hat{s}_{min}^{(bb\ell\ell)}}(t\bar{t})$ peaks lower, near the $2m_t$ threshold. Secondly, both $\sqrt{\hat{s}}(hh)$ and $\sqrt{\hat{s}}(t\bar{t})$ peak at

PoS(AISIS2019)013

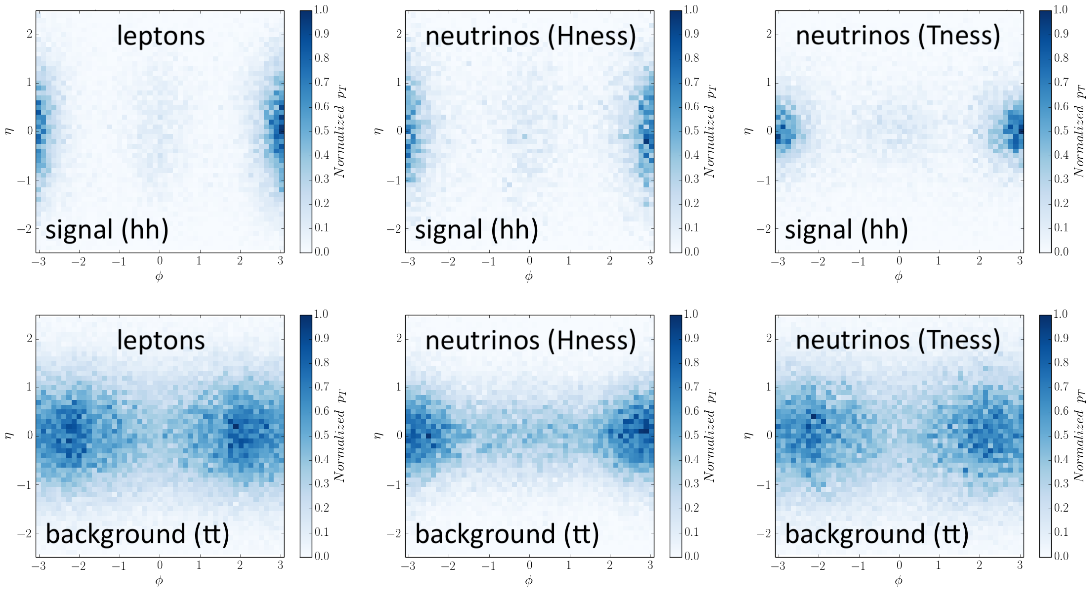

Figure 3: The same as in Fig. 2 but for leptons (1st column) and neutrinos with approximate momentum reconstruction using Higgsness (2nd) and Topness (3rd).

~ 400 GeV. This implies that while the two top quarks are produced near threshold ($2m_t$), the two Higgs bosons are produced well above the corresponding $2m_h$ threshold. Consequently, the two top quarks are more or less at rest, while the two Higgs bosons are relatively boosted and their decay products tend to be more collimated. This observation motivates the use of simple kinematic variables such as $\Delta R_{\ell\ell}$, ΔR_{bb}, $m_{\ell\ell}$ and m_{bb} as a part of basic selection cuts introduced in Ref. [18].

Another difference between the signal ($hh \to WW^* \to b\bar{b}\ell\bar{\ell}\nu\bar{\nu}$) and the dominant background ($t\bar{t} \to b\bar{b}\ell\bar{\ell}\nu\bar{\nu}$) is that two b-quarks in the signal arise from the color-singlet (h) and those in the $t\bar{t}$ from color-octet (gg in the initial state). Therefore hadrons from the color singlet tend to be closer to each other [30–33]. Ref. [34] utilized the color-flow (first time for double Higgs) and showed the significant increase in the final signal significance.

Fig. 2 shows the cumulative average of the jet images for the signal (top panel) and the $t\bar{t}$ background (bottom panel). The origin of the (ϕ, η) plane is taken to be the center of the b quark pair and the density indicates the total p_T in each pixel. Images from the left to the right are obtained from charged hadrons (1st column), neutral hadrons (2nd), and photons (3rd). In Ref. [34], we used the convolutional neural networks (CNN) with the first three images (charged hadrons, neutral hadrons, photons) along with kinematic variables. These two recent studies show that one can enhance the signal sensitivity significantly via interplay of kinematics and machine learning algorithm. We regard the $hh \to b\bar{b}WW^*$ channel as important as other channels such as $bb\gamma\gamma$, $bbbb$ and $bb\tau\tau$.

3. Playing with neural networks

The first improvement that we are targeting is the use of momenta of leptons and the reconstructed neutrino momenta. Lepton momenta enter in kinematic variables such as $m_{\ell\ell}$, $\Delta R_{\ell\ell}$ etc.

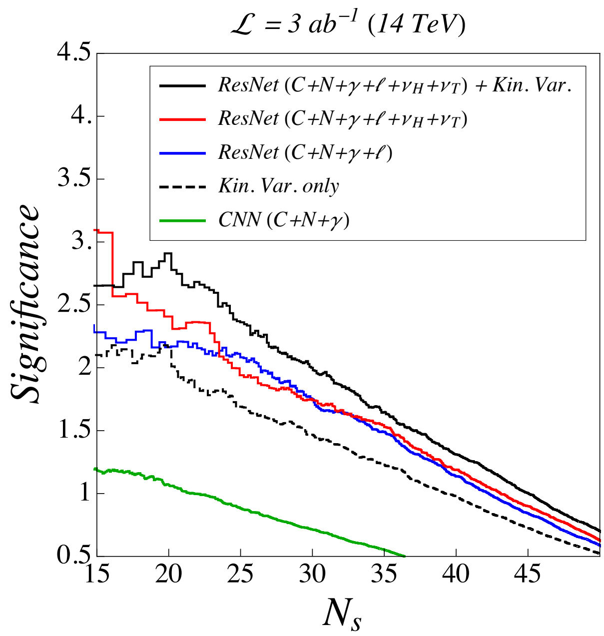

Figure 4: The significance to observe the double Higgs production at the 14 TeV LHC. The acronyms C, N, γ, ℓ, ν_H, ν_T denote charged, neutral, photon, lepton, neutrino (Higgsness), and neutrino (Topness) images respectively. The acronym *Kin. Var.* stands for sixteen kinematic variables, $p_{T\ell_1}$, $p_{T\ell_2}$, $\not{P}_T$, $m_{\ell\ell}$, m_{bb}, $\Delta R_{\ell\ell}$, ΔR_{bb}, p_{Tbb}, $p_{T\ell\ell}$, $min[\Delta R_{b\ell}]$, Topness, Higgsness, $M_{T2}^{(b)}$, $M_{T2}^{(\ell)}$, $\hat{s}_{min}^{(\ell\ell)}$ and $\hat{s}_{min}^{(bb\ell\ell)}$.

However, since jet images are used (for color flow), it would make sense to study images of leptons and neutrinos. Correlation among the b-tagged jet, leptons and neutrinos will be learned naturally via images in neural networks. Fig. 3 show the cumulative average of the lepton images (1st column) and neutrino images (2nd and 3rd) for the signal (top) and the $t\bar{t}$ background (bottom) before the baseline cuts. Two neutrino images are obtained using Higgsness (2nd) and Topness (3rd). Although they are only approximate, they do exhibit noticeable difference. As expected, neutrino images are supposed to the same as leptons images.

In Fig. 4. we show our final result on the signal significance considering the dominant background ($t\bar{t}$) only. The green curve is obtained using CNN described in Ref. [18] with charged, neutral and photon images only. When adding an additional lepton image, we employ a deeper neural network based on ResNet [35] topology as shown in Fig. 5. The blue curve shows that the lepton image significantly improves the result, by capturing an orthogonal information in that the dilepton system of double Higgs production is back-to-back with respect to the $b\bar{b}$-system. The improvement from the neutrino images reconstructed from the Higgsness and Topness turns out to be mild as shown in the red curve. We find that there is a marginal improvement in a small signal efficiency region. Although the neutrino images of double Higgs production and $t\bar{t}$ are manifestly different, the fact that they are highly correlated to lepton images, as shown in Fig. 3, reduces its effectiveness. When we combine the sixteen kinematic variables together with all images, the overall significance can be substantially improved as shown by solid black curve. The kinematic variables themselves (dashed black line), however, are outperformed by ResNet, indicating that the image-based neural network has a potential to improve a conventional analysis which utilizes

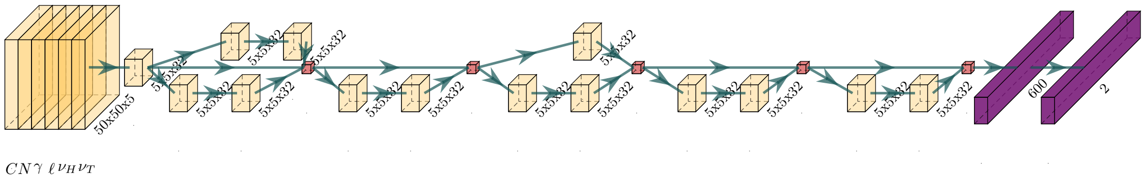

Figure 5: The charged, neutral, photon, lepton, neutrino (Higgsness) and neutrino (Topness) images are fed to a deeper neural network based on ResNet [35] topology.

high-level kinematic variables only.

In summary, we discussed a novel method incorporated in deep learning framework, which could bring a significant increase in the signal sensitivity for *hh* production in the dilepton channel compared to previous analyses [9, 10, 16]. The discussed method is quite general and can be easily applied to other processes such as the semi-leptonic final state, resonant *hh* production, non-resonant production with more than one Higgs boson, etc. It is straightforward to generalize the idea to different topologies in searches for other BSM particles as well.

Acknowledgments

This work is supported in part by the U.S. Department of Energy under grant No. DE-SC0019474.

References

[1] ATLAS collaboration, G. Aad et al., *Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC*, Phys. Lett. **B716** (2012) 1–29, [1207.7214].

[2] CMS collaboration, S. Chatrchyan et al., *Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC*, Phys. Lett. **B716** (2012) 30–61, [1207.7235].

[3] HL/HE WG2 GROUP collaboration, M. Cepeda et al., *Higgs Physics at the HL-LHC and HE-LHC*, 1902.00134.

[4] ATLAS, CMS collaboration, G. Aad et al., *Measurements of the Higgs boson production and decay rates and constraints on its couplings from a combined ATLAS and CMS analysis of the LHC pp collision data at $\sqrt{s} = 7$ and 8 TeV*, JHEP **08** (2016) 045, [1606.02266].

[5] ATLAS collaboration, *Study of the double Higgs production channel $H(\to b\bar{b})H(\to \gamma\gamma)$ with the ATLAS experiment at the HL-LHC*, .

[6] ATLAS collaboration, *Projected sensitivity to non-resonant Higgs boson pair production in the $b\bar{b}b\bar{b}$ final state using proton-proton collisions at HL-LHC with the ATLAS detector*, .

[7] J. H. Kim, Y. Sakaki and M. Son, *Combined analysis of double Higgs production via gluon fusion at the HL-LHC in the effective field theory approach*, Phys. Rev. **D98** (2018) 015016, [1801.06093].

[8] CMS collaboration, C. Collaboration, *Combination of searches for Higgs boson pair production in proton-proton collisions at $\sqrt{s} = 13$ TeV*, .

[9] CMS collaboration, *Higgs pair production at the High Luminosity LHC*, .

PoS(AISIS2019)013

[10] CMS collaboration, *Projected performance of Higgs analyses at the HL-LHC for ECFA 2016*, .

[11] J. Baglio, A. Djouadi, R. Gröber, M. M. Mühlleitner, J. Quevillon and M. Spira, *The measurement of the Higgs self-coupling at the LHC: theoretical status*, JHEP **04** (2013) 151, [1212.5581].

[12] CMS collaboration, A. M. Sirunyan et al., *Search for resonant and nonresonant Higgs boson pair production in the* $b\bar{b}\ell\nu\ell\nu$ *final state in proton-proton collisions at* $\sqrt{s} = 13$ *TeV*, JHEP **01** (2018) 054, [1708.04188].

[13] J. Alison et al., *Higgs Boson Pair Production at Colliders: Status and Perspectives*, in *Double Higgs Production at Colliders Batavia, IL, USA, September 4, 2018-9, 2019* (B. Di Micco, M. Gouzevitch, J. Mazzitelli and C. Vernieri, eds.), 2019, 1910.00012, https://lss.fnal.gov/archive/2019/conf/fermilab-conf-19-468-e-t.pdf.

[14] J. H. Kim, K. Kong, B. Nachman and D. Whiteson, *The motivation and status of two-body resonance decays after the LHC Run 2 and beyond*, 1907.06659.

[15] N. Craig, P. Draper, K. Kong, Y. Ng and D. Whiteson, *The unexplored landscape of two-body resonances*, Acta Phys. Polon. **B50** (2019) 837, [1610.09392].

[16] A. Adhikary, S. Banerjee, R. K. Barman, B. Bhattacherjee and S. Niyogi, *Revisiting the non-resonant Higgs pair production at the HL-LHC*, JHEP **07** (2018) 116, [1712.05346].

[17] ATLAS collaboration, G. Aad et al., *Search for non-resonant Higgs boson pair production in the* $bb\ell\nu\ell\nu$ *final state with the ATLAS detector in pp collisions at* $\sqrt{s} = 13$ *TeV*, 1908.06765.

[18] J. H. Kim, K. Kong, K. T. Matchev and M. Park, *Measuring the Triple Higgs Self-Interaction at the Large Hadron Collider*, 1807.11498.

[19] C. G. Lester and D. J. Summers, *Measuring masses of semiinvisibly decaying particles pair produced at hadron colliders*, Phys. Lett. **B463** (1999) 99–103, [hep-ph/9906349].

[20] M. Burns, K. Kong, K. T. Matchev and M. Park, *Using Subsystem MT2 for Complete Mass Determinations in Decay Chains with Missing Energy at Hadron Colliders*, JHEP **03** (2009) 143, [0810.5576].

[21] A. J. Barr, T. J. Khoo, P. Konar, K. Kong, C. G. Lester, K. T. Matchev et al., *Guide to transverse projections and mass-constraining variables*, Phys. Rev. **D84** (2011) 095031, [1105.2977].

[22] P. Konar, K. Kong and K. T. Matchev, $\sqrt{\hat{s}}_{min}$ *: A Global inclusive variable for determining the mass scale of new physics in events with missing energy at hadron colliders*, JHEP **03** (2009) 085, [0812.1042].

[23] P. Konar, K. Kong, K. T. Matchev and M. Park, *RECO level* $\sqrt{s}_{min}$ *and subsystem* $\sqrt{s}_{min}$*: Improved global inclusive variables for measuring the new physics mass scale in* $\not{E}_T$ *events at hadron colliders*, JHEP **06** (2011) 041, [1006.0653].

[24] M. L. Graesser and J. Shelton, *Hunting Mixed Top Squark Decays*, Phys. Rev. Lett. **111** (2013) 121802, [1212.4495].

[25] W. S. Cho, D. Kim, K. T. Matchev and M. Park, *Probing Resonance Decays to Two Visible and Multiple Invisible Particles*, Phys. Rev. Lett. **112** (2014) 211801, [1206.1546].

[26] D. Debnath, D. Kim, J. H. Kim, K. Kong and K. T. Matchev, *Resolving Combinatorial Ambiguities in Dilepton* $t\bar{t}$ *Event Topologies with Constrained* M_2 *Variables*, Phys. Rev. **D96** (2017) 076005, [1706.04995].

[27] P. Konar, K. Kong, K. T. Matchev and M. Park, *Superpartner Mass Measurement Technique using 1D Orthogonal Decompositions of the Cambridge Transverse Mass Variable M_{T2}*, *Phys. Rev. Lett.* **105** (2010) 051802, [0910.3679].

[28] P. Konar, K. Kong, K. T. Matchev and M. Park, *Dark Matter Particle Spectroscopy at the LHC: Generalizing M(T2) to Asymmetric Event Topologies*, *JHEP* **04** (2010) 086, [0911.4126].

[29] W. S. Cho, K. Choi, Y. G. Kim and C. B. Park, *Gluino Stransverse Mass*, *Phys. Rev. Lett.* **100** (2008) 171801, [0709.0288].

[30] F. Maltoni, K. Paul, T. Stelzer and S. Willenbrock, *Color Flow Decomposition of QCD Amplitudes*, *Phys. Rev.* **D67** (2003) 014026, [hep-ph/0209271].

[31] J. Gallicchio and M. D. Schwartz, *Seeing in Color: Jet Superstructure*, *Phys. Rev. Lett.* **105** (2010) 022001, [1001.5027].

[32] ATLAS collaboration, M. Aaboud et al., *Measurement of colour flow using jet-pull observables in $t\bar{t}$ events with the ATLAS experiment at $\sqrt{s} = 13\,TeV$*, *Eur. Phys. J.* **C78** (2018) 847, [1805.02935].

[33] J. Cogan, M. Kagan, E. Strauss and A. Schwarztman, *Jet-Images: Computer Vision Inspired Techniques for Jet Tagging*, *JHEP* **02** (2015) 118, [1407.5675].

[34] J. H. Kim, M. Kim, K. Kong, K. T. Matchev and M. Park, *Portraying Double Higgs at the Large Hadron Collider*, *JHEP* **09** (2019) 047, [1904.08549].

[35] S. R. Kaiming He, Xiangyu Zhang and J. Sun, *Deep Residual Learning for Image Recognition*, 1512.03385.