

# Supervising Deep Neural Networks with Topological Augmentation in search for di-Higgs Production at the LHC

---

**Won Sang Cho\***

*Department of Physics and Astronomy, Seoul National University, Seoul 151-747, Korea*  
E-mail: [wscho@snu.ac.kr](mailto:wscho@snu.ac.kr)

**Dongsub Lee**

*Department of Physics and Astronomy, Seoul National University, Seoul 151-747, Korea*  
E-mail: [dongsub93@snu.ac.kr](mailto:dongsub93@snu.ac.kr)

**Chan Beom Park**

*Center for Theoretical Physics of the Universe, Institute for Basic Science (IBS), Daejeon, 34126, Korea*  
E-mail: [cbpark@ibs.re.kr](mailto:cbpark@ibs.re.kr)

**Jae-hyeon Park**

*Korea Institute for Advanced Study, Seoul 130-722, Korea*  
E-mail: [jhpark@kias.re.kr](mailto:jhpark@kias.re.kr)

Augmentation of invisible information with respect to many hypothetical models of background and signal processes, can highly improve the performance of the machine learning classifiers for HEP event discrimination. In this regard, di-Higgs search in the channels with multiple invisible final states, is one of the most important applications. Focusing on the di-Higgs channels with 2 bottom quarks + 0/1/2-leptons/taus + MET from  $bbWW$  and  $bb\tau\tau$  productions, we introduce various augmentation schemes and ways to build better multi-class classifiers using deep neural networks. We conclude our study with demonstration how much the new deep learning classifiers supervised by physical augmentation, can improve the discovery potential of di-Higgs production at the LHC, and discuss on the implications for future collider study.

*The 39th International Conference on High Energy Physics (ICHEP2018)*  
4-11 July, 2018  
Seoul, Korea

---

\*Speaker.

## 1. Topological Augmentation for di-Higgs Production and Decays & DNN Classifier

**Motivation** As demonstrated in many studies for machine learning HEP event classifiers nowadays, descent deep learning models which are trained even not-necessarily with high level features variables induced from physical invariances, can have highly improved performances in experimental data analysis, once provided by enough data [1]. However, still there exist obvious ways for human physics models to contribute more crucially, and one of them would be involved with the ways of feeding relevant data to the machines, under the supervision of physics models in purpose. In this regard, in search for di-Higgs productions using the  $2b + 2l + \text{MET}$  signatures from  $hh \rightarrow bbWW$  and  $hh \rightarrow bb\tau\tau$  production channels (as signal), against the  $t\bar{t}$  production and decay channels (as main background), we try to augment the invisible and missing momentum degrees of freedom, with respect to every relevant hypothetical decay topologies in the signal and background processes, and train deep neural networks under the supervision including the augmented missing features, in order to improve the performance of machine learning event classifiers searching for di-Higgs production at the LHC.

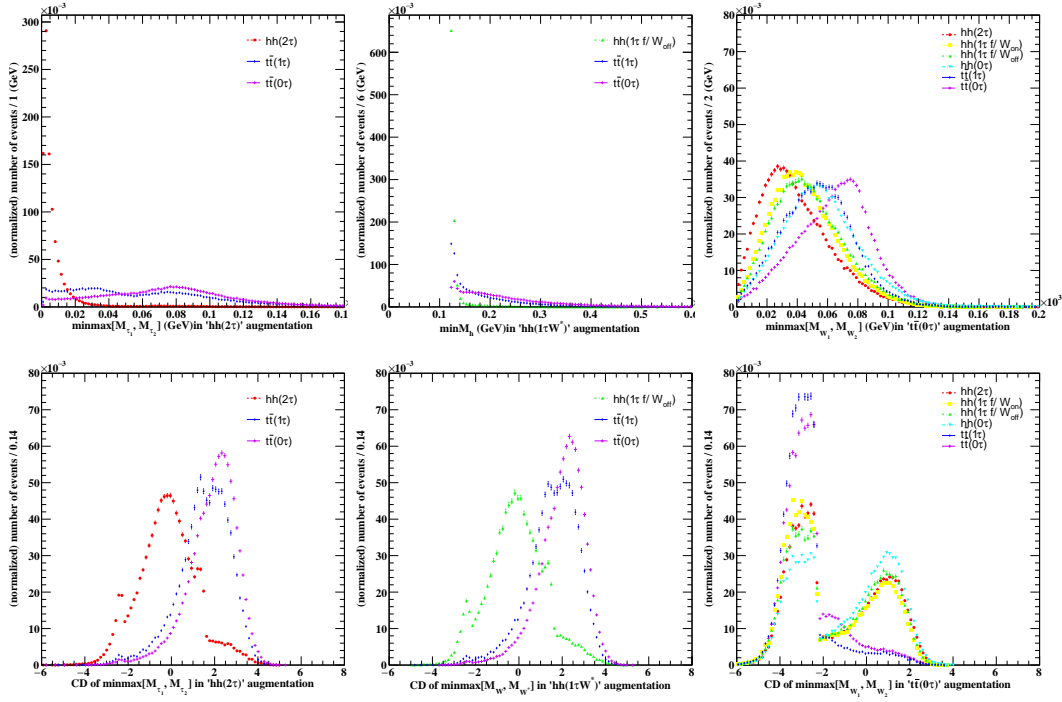
**Augmentation** Considering the final states including  $(2b+2l+\text{MET})$ , we categorized the signal and background processes into 7-classes, by the number of leptonically decaying taus which can make kinematically distinctive features with more complicated decay topology. As significant portion of energy-momentum flow defining the decay topology, is carried by invisible neutrinos in each topological classes, sizable portion of topological information is blind to machines when training for classifiers with conventional visible features.

Using the program package OPTIMASS [2] which had been developed for kinematic mass function minimization with constraints, we augmented the missing momentum degrees of freedom by imposing the sets of full kinematic constraints, each of which is distinctively defined in a given topological class.

In result, as main augmented features from OPTIMASS, we can obtain 1) a minimized subsystem mass -  $\hat{m}(a)$  dubbed as an *optimass* of a subsystem -  $a$ , and 2) an optimized chi-square variable -  $\hat{\chi}^2$  with which we call  $\hat{\delta} \equiv 2\log(\hat{\chi})$ , a *compatibility-distance* of an input event to be consistent with the imposed set of kinematic constraints. Actually, these two augmented variables are not independently calculated, but they are closely related and computed concurrently in the process of constrained minimization of the subsystem mass w.r.t the invisible d.o.f defined in a given hypothetical decay topology, so that the solution of OPTIMASS minimizes a subsystem mass in the space which also mimizes the compatibility distance for the imposed constraints.

The solution of the OPTIMASS can also be considered as generalized MAOS momenta [3] for under-constrained event topologies, just with more flexibility of choosing 1) the type of hypothetical decay topology, 2) the form of target mass function to be minimized in the chosen topology, and 3) the kinds of kinematic constraints which are consistent with the topology, so that the specified missing d.o.f are totally fixed while the mass minimization subject to the kinematic constraints.

**Augmented Features** Figure 1 shows some optimasses,  $\hat{m}$  in the [top row] with their compatibility-distances,  $\hat{\delta}$  in the [bottom row], for various event processes - ( $hh$  with 0/1/2-taus relics, and  $t\bar{t}$  with 0/1-tau relics), in the three augmentation models - 1) [left column] ' $hh(2\tau)$ '-augmentation:  $hh \rightarrow 2b + 2\tau \rightarrow 2b + 2l + \text{MET}(4\nu)$ , 2) [central column] ' $hh(1\tau W^*)$ '-augmentation:  $hh \rightarrow 2b + W(\rightarrow l + \nu) + W^*(\rightarrow \nu + \tau) \rightarrow 2b + 2l + \text{MET}(4\nu)$ , and 3) [right column] ' $t\bar{t}(0\tau)$ '-augmentation:



**Figure 1:** 1) optimass  $\hat{m}$  [top-row] & 2) compatibility-distance  $\hat{\delta}$  [bottom-row], for various signal and background processes listed in each legend, for three distinctive augmentation models for each column - a)  $hh(2\tau)$  [left], b)  $hh(1\tau W^*)$  [center], and c)  $t\bar{t}(0\tau)$  [right]. More details in the paragraph - ‘Augmented Features’.

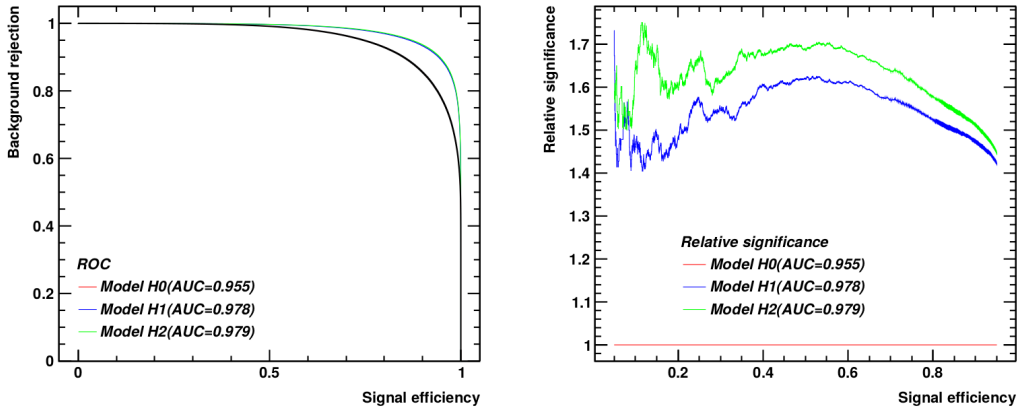
$t\bar{t} \rightarrow bbWW \rightarrow 2b + 2l + \text{MET}(2\nu)$ .

As shown in the [bottom-left] panel, the di-Higgs with  $2\tau$  relic events tend to have smaller  $\hat{\delta}$  values than other background processes, when augmented by its own ‘ $hh(2\tau)$ ’-model. For the same reason, as shown in the [bottom-right] panel, the  $t\bar{t}$  production with  $(0/1\tau)$  relic events tend to have smaller  $\hat{\delta}$  values than the signal di-Higgs processes, if augmented by their own ‘ $t\bar{t}(0/1\tau)$ ’-model. This phenomenon can also be checked in the [bottom-central] panel.

For the three topological augmentations for the various signal and background processes, a lot of optimass  $\hat{m}$  variables can be defined and obtained. Actually there exist lots of interesting optimass variables which can be consistently defined using the sub-decay systems in each distinctive augmentation topology, and we shows some of them in the top-row as the following - 1) [top-left]  $\min\{\max[M_{\tau_1}, M_{\tau_2}]\}$  in ‘ $hh(2\tau)$ ’-augmentation, 2) [top-center]  $\min\{M_{h \rightarrow WW^*}\}$  in ‘ $hh(1\tau W^*)$ ’-augmentation, 3) [top-right]  $\min\{\max[M_{t_1}, M_{t_2}]\}$  in ‘ $t\bar{t}(0\tau)$ ’-augmentation.

Interestingly, the value of such  $\hat{m}$  variables of a process, if their objective mass function (to be minimized) can be consistently defined in a given augmentation model, tend to be ranged well below the very value of the objective mass function *for the true invisible momenta, if exist* (like as  $m_{T/T2}(W) \leq M(W)$ ), while some other processes which are not consistent with the given augmentation models, tend to violate and be squeezed out of the  $\hat{m}^{max}$  boundary. We can check this behavior quite well with the top panel plots.

**Training DNN Classifiers** For the signal and background processes categorized by the 7 topological classes, we augmented each process event data by the all 7 topological models, and utilized



**Figure 2:** a) ROC (Receiver Operating Characteristic) curves, and b) relative significances of di-Higgs event discrimination, using 3 different DNN classifiers, trained with augmented features ( $H2$  and  $H1$ ) over only visible features  $H0$ . More details in the paragraph - ‘Training DNN Classifiers’

various optimasses,  $\hat{m}$  and compatibility distances,  $\hat{\delta}$ , for the training of multi-class classifiers using deep neural networks in addition to the conventional feature variables used in [4].

Figure 2-a) [left] shows ROC curves of the (binarized by sig vs bg) DNN classifier models -  $H0$ ,  $H1$ ,  $H2$ , which are trained with feature variable sets - {‘visible only’, ‘augmented only’, ‘visible + augmented’}, respectively, and Figure 2-b) [right] shows the relative significances when using  $H1$  and  $H2$  with respect to the  $H0$  model. We checked that significant gains can be obtained, when we employ various augmented missing features for building machine learning HEP event classifiers, especially with the demonstration of the performance in search for the rare (non-resonant SM) di-Higgs productions at the LHC. More extensive studies including more broad signatures will also be addressed in the work in progress [5].

## References

- [1] P. Baldi, P. Sadowski and D. Whiteson, “Enhanced Higgs Boson to  $\tau^+\tau^-$  Search with Deep Learning,” Phys. Rev. Lett. **114**, no. 11, 111801 (2015) [arXiv:1410.3469 \[hep-ph\]](#).
- [2] W. S. Cho, J. S. Gainer, D. Kim, S. H. Lim, K. T. Matchev, F. Moortgat, L. Pape and M. Park, “OPTIMASS: A Package for the Minimization of Kinematic Mass Functions with Constraints,” JHEP 1601 (2016) 026 [arXiv:1508.00589 \[hep-ph\]](#).
- [3] W. S. Cho, K. Choi, Y. G. Kim and C. B. Park, “M(T2)-assisted on-shell reconstruction of missing momenta and its application to spin measurement at the LHC,” Phys. Rev. D **79**, 031701 (2009) [arXiv:0810.4853 \[hep-ph\]](#).
- [4] CMS Collaboration [CMS Collaboration], “Search for resonant and non-resonant Higgs boson pair production in the  $b\bar{b}l\nu l\nu$  final state at  $\sqrt{s} = 13$  TeV,” CMS-PAS-HIG-17-006, JHEP 01 (2018) 054 [arXiv:1708.04188 \[hep-ph\]](#).
- [5] K. Y. Ban, et. al in preparation.