M.L. with Augmentation for Boosting the discovery of di-Higgs search

Kayoung Ban (반가영)

(Yonsei University)

In collaboration with

Won Sang Cho(SNU), Dongsub Lee(SNU), Chan Beom Park(IBS-CTPU),
Jae-hyeon Park(KIAS), Seong Chan Park(Yonsei Univ.)

ban94gy@yonsei.ac.kr

2018.12.07
KEK-PH2018 winter

서울대학교 SEOUL NATIONAL UNIVERSITY

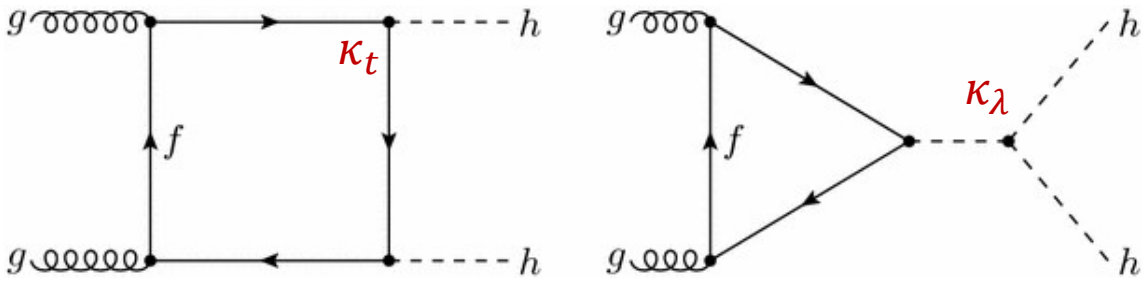
Institute for Basic Science 기초과학연구원

imagine the impossible KIAS KOREA INSTITUTE FOR ADVANCED STUDY

연세대학교 YONSEI UNIVERSITY

Introduction

- **Higgs pair production**



- **Higgs self-coupling**

$$\mathcal{L}_h = \frac{1}{2}\partial^\mu h \partial_\mu h - \frac{1}{2}m_h^2 h^2 - \kappa_\lambda \lambda_{SM} v h^3 - \frac{m_t}{v}(v + \kappa_t h)(\bar{t}_L t_R + h.c) + \cdots$$

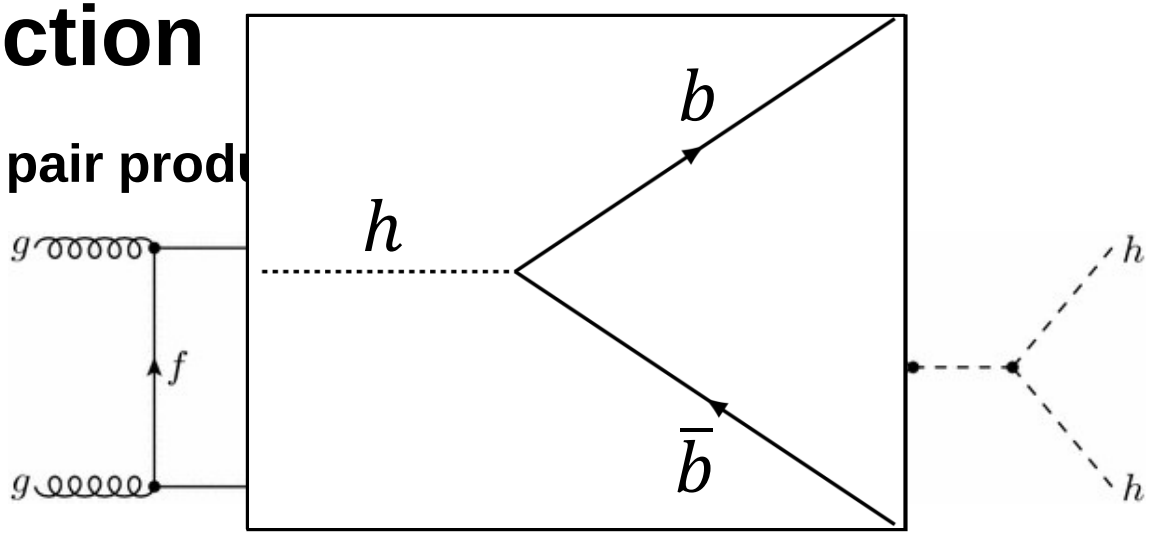
✓ Di-Higgs search has a meaning to measure the Higgs self coupling and to understand the Higgs potential

✓ The topology is determined by each Higgs decay chain

Figure ref : arXiv:1512.08928

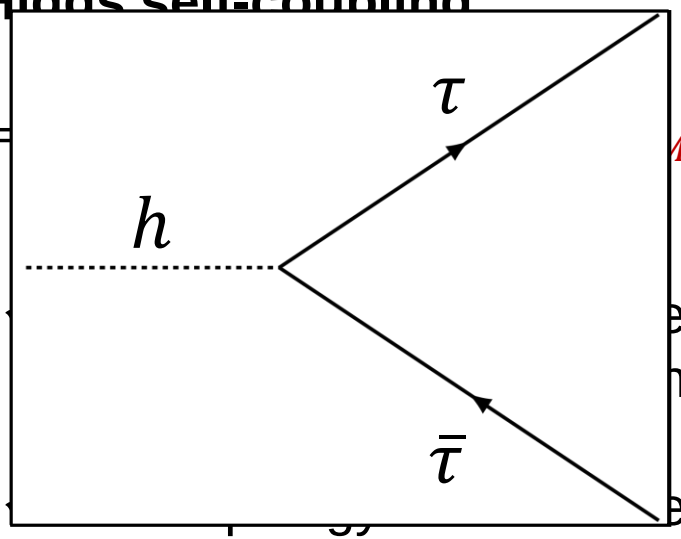
Introduction

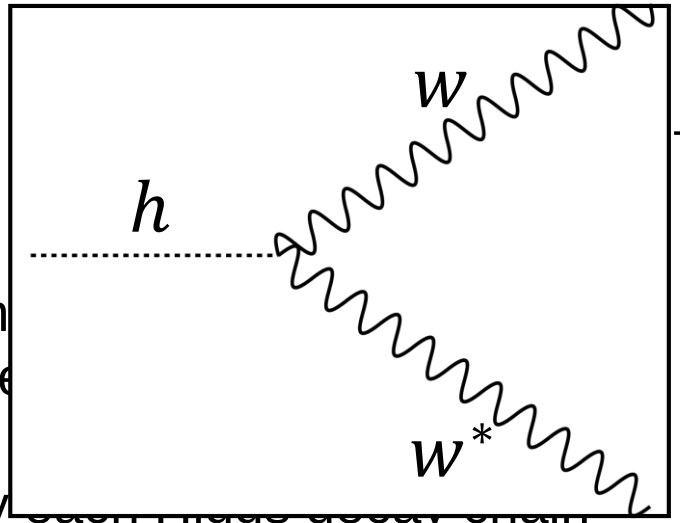
- **Higgs pair produ**



- **Higgs self-coupling**

$\mathcal{L}_h =$ $_4 v h^3$ $+ \cdots$





eanin
nd the
ed by

Figure ref : arXiv:1512.08928

Signal vs Background

@ 14TeV, 39.64 fb (HH), 953.6 fb (TT)

Channel	Leptons	X section	Topology $* \ l = e \ or \ \mu$
HH2Tau	0	~1.2	$hh \rightarrow b\ b\ \tau\ \tau \rightarrow \boldsymbol{b\ b\ \tau_h\ \tau_h + met}$
TT2Tau	0	~5097.2	$\bar{t}t \rightarrow b\ w\ b\ w \rightarrow b\ b\ \tau\ \tau + met \rightarrow \boldsymbol{b\ b\ \tau_h\ \tau_h + met}$
HH2Tau	1	~1.3	$hh \rightarrow b\ b\ \tau\ \tau \rightarrow \boldsymbol{b\ b\ \tau_h\ l + met}$
HH2W*W	1	~0.15	$hh \rightarrow b\ b\ w\ w^* \rightarrow b\ b\ \tau\ l + met \rightarrow \boldsymbol{b\ b\ \tau_h\ l + met}$
HH2WW*	1	~0.15	$hh \rightarrow b\ b\ w\ w^* \rightarrow b\ b\ l\ \tau + met \rightarrow \boldsymbol{b\ b\ l\ \tau_h + met}$
TT2Tau	1	~5546.3	$\bar{t}t \rightarrow b\ w\ b\ w \rightarrow b\ b\ \tau\ \tau + met \rightarrow \boldsymbol{b\ b\ \tau_h\ l + met}$
TT1Tau	1	~29700.2	$\bar{t}t \rightarrow b\ w\ b\ w \rightarrow b\ b\ \tau\ l + met \rightarrow \boldsymbol{b\ b\ \tau_h\ l + met}$
HH2Tau	2	~0.36	$hh \rightarrow b\ b\ \tau\ \tau \rightarrow \boldsymbol{b\ b\ l\ l + met}$
HH2W*W1Tau	2	~0.08	$hh \rightarrow b\ b\ w\ w^* \rightarrow b\ b\ \tau\ l + met \rightarrow \boldsymbol{b\ b\ l\ l + met}$
HH2WW*1Tau	2	~0.08	$hh \rightarrow b\ b\ w\ w^* \rightarrow b\ b\ l\ \tau + met \rightarrow \boldsymbol{b\ b\ l\ l + met}$
HH2WW0Tau	2	~0.47	$hh \rightarrow b\ b\ w\ w^* \rightarrow \boldsymbol{b\ b\ l\ l + met}$
TT2Tau	2	~1508.7	$\bar{t}t \rightarrow b\ w\ b\ w \rightarrow b\ b\ \tau\ \tau + met \rightarrow \boldsymbol{b\ b\ l\ l + met}$
TT1Tau	2	~16158.3	$\bar{t}t \rightarrow b\ w\ b\ w \rightarrow b\ b\ \tau\ l + met \rightarrow \boldsymbol{b\ b\ l\ l + met}$
TT0Tau	2	~43263.9	$\bar{t}t \rightarrow b\ w\ b\ w \rightarrow \boldsymbol{b\ b\ l\ l + met}$

Signal vs Background

@ 14TeV, 39.64 fb (HH), 953.6 fb (TT)

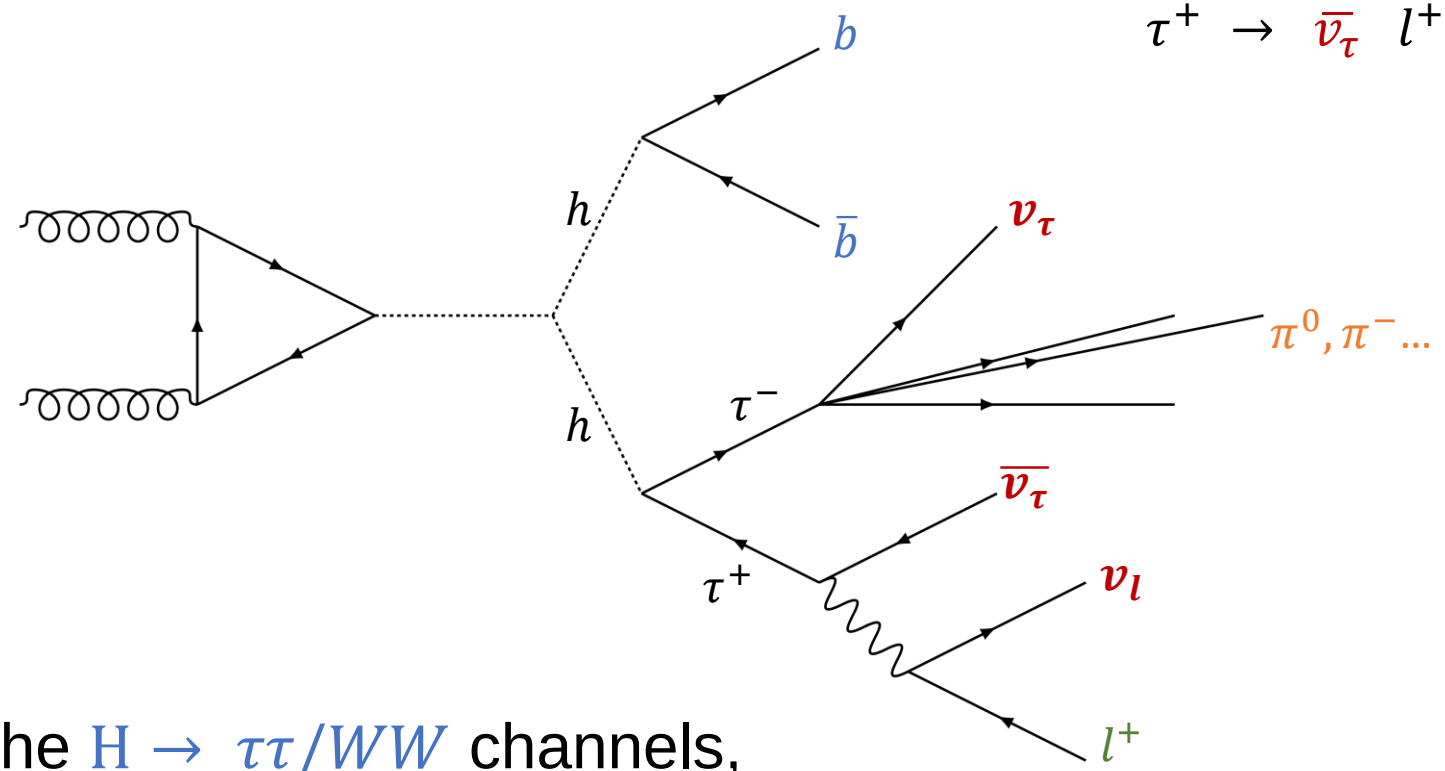
Channel	Leptons	X section	Topology * $l = e\ or\ \mu$
HH2Tau	0	~1.2	$hh \rightarrow b\,b\,\tau\,\tau \rightarrow \boldsymbol{b\,b\,\tau_h\,\tau_h + met}$
TT2Tau	0	~5097.2	$\bar{t}t \rightarrow b\,w\,b\,w \rightarrow b\,b\,\tau\,\tau + met \rightarrow \boldsymbol{b\,b\,\tau_h\,\tau_h + met}$
HH2Tau	1	~1.3	$hh \rightarrow b\,b\,\tau\,\tau \rightarrow \boldsymbol{b\,b\,\tau_h\,l + met}$
HH2W*W	1	~0.15	$hh \rightarrow b\,b\,w\,w^* \rightarrow b\,b\,\tau\,l + met \rightarrow \boldsymbol{b\,b\,\tau_h\,l + met}$
HH2WW*	1	~0.15	$hh \rightarrow b\,b\,w\,w^* \rightarrow b\,b\,l\,\tau + met \rightarrow \boldsymbol{b\,b\,l\,\tau_h + met}$
TT2Tau	1	~5546.2	$\bar{t}t \rightarrow b\,w\,b\,w \rightarrow b\,b\,\tau\,\tau + met \rightarrow \boldsymbol{b\,b\,\tau_h\,l + met}$
TT1Tau	[illegible]	[illegible]	[illegible] $\boldsymbol{b\,\tau_h\,l + met}$
HH2Tau	2	[illegible]	[illegible] $\boldsymbol{b\,b\,l\,l + met}$
HH2W*W1Tau	2	~0.08	$hh \rightarrow b\,b\,w\,w^* \rightarrow b\,b\,\tau\,l + met \rightarrow \boldsymbol{b\,b\,l\,l + met}$
HH2WW*1Tau	2	~0.08	$hh \rightarrow b\,b\,w\,w^* \rightarrow b\,b\,l\,\tau + met \rightarrow \boldsymbol{b\,b\,l\,l + met}$
HH2WW0Tau	2	~0.47	$hh \rightarrow b\,b\,w\,w^* \rightarrow \boldsymbol{b\,b\,l\,l + met}$
TT2Tau	2	~1508.7	$\bar{t}t \rightarrow b\,w\,b\,w \rightarrow b\,b\,\tau\,\tau + met \rightarrow \boldsymbol{b\,b\,l\,l + met}$
TT1Tau	2	~16158.3	$\bar{t}t \rightarrow b\,w\,b\,w \rightarrow b\,b\,\tau\,l + met \rightarrow \boldsymbol{b\,b\,l\,l + met}$
TT0Tau	2	~43263.9	$\bar{t}t \rightarrow b\,w\,b\,w \rightarrow \boldsymbol{b\,b\,l\,l + met}$

Too Many $t\bar{t}$ Background

How to reconstruct the missing information?

Ex) HH2tau : $b\ b\ l\ \tau_h\ + met$

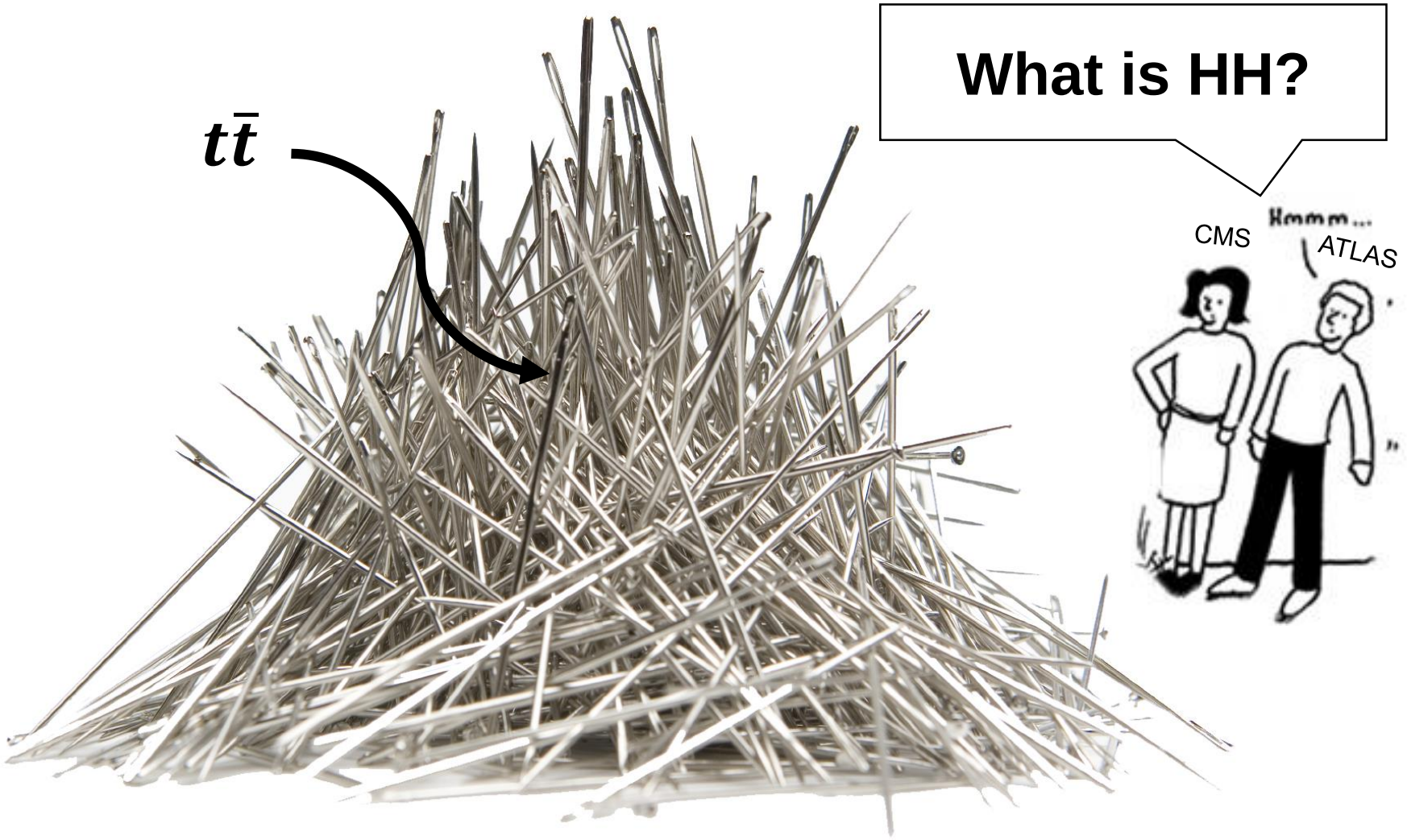
$$hh \rightarrow b\ b\ \tau^-\tau^+$$
$$\tau^- \rightarrow \nu_\tau\ \tau_h$$
$$\tau^+ \rightarrow \bar{\nu}_\tau\ l^+\ \tau_l$$



In the $H \rightarrow \tau\tau/WW$ channels,
Higgs decay has more than 2 neutrinos

→ **We need to reconstruct the missing information**

A Needle(HH) in a stack of Needles($t\bar{t}$)



OptiMass introduction

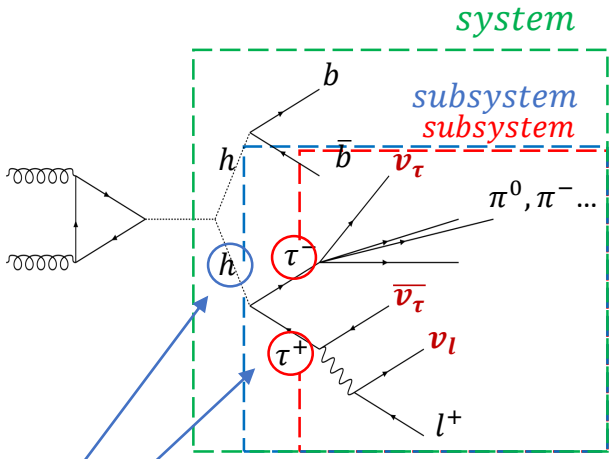
M = *minimized mass variable function*
λ_a = *augmented Lagrange parameter*
μ = *penalty parameter*
$\vec{x}$ = ***missing momentum***
c_a = ***physical constraints***

- Augmented Lagrangian method

$$\mathcal{L}_k(\vec{x} \mid \lambda, \mu) = \boldsymbol{M}(\vec{x}) - \sum_{a=1}^{m} \lambda_{ak}\, c_a(\vec{x}) + \frac{1}{2\mu_k} \sum_a c_a^2(\vec{x}).$$

$$\lambda_a^{k+1} = \lambda_a^k - \frac{c_a(\vec{x}_k)}{\mu_k} \rightarrow \text{iteratively updated}$$

$$\boldsymbol{CD} = ||c(\vec{x}_k)|| < \eta^* \quad with\ ||c(\vec{x}_k)|| = \sqrt{\sum_a |c_a(\vec{x}_k)|^2}$$



For each (sub)system, we have **Physical hypothesis** models

✓ **OM (OptiMass)** ~ TARGET MASS

✓ **CD (Compatibility Distance)** ~ **0** for TRUE SYSTEM

✓ $\vec{x}$: Reconstructed **missing momentum** (next page)

OPTIMASS: A Package for the Minimization of Kinematic Mass Functions with Constraints
- Cho, Won Sang *et al.* JHEP 1601 (2016) 026 arXiv:1508.00589

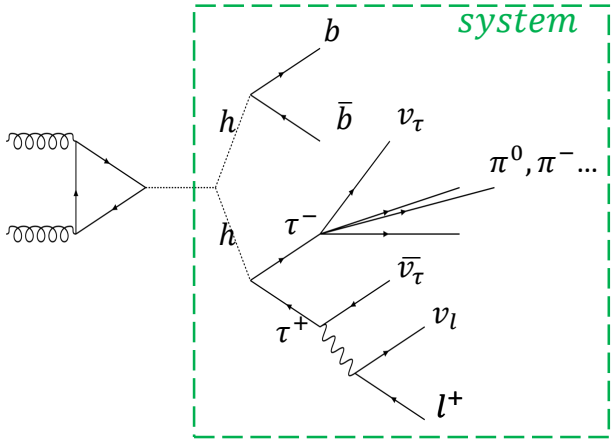
OptiMass introduction

High Level Feature

- ✓ **OM (OptiMass)** ~ TARGET MASS
- ✓ **CD (Compatibility Distance)** ~ **0** for TRUE SYSTEM

$\vec{x}$

- ✓ We can also get **Raw Level Features** by **OPTIMASS** which are reconstructed **Kinematic Variables!**



- ✓ Like P_{T,v_l} *or* η_{v_τ} *or* $\Delta R_{v_{\tau 1} v_{\tau 2}}$ *or* $\Delta\phi_{v1}$ …

	P_T	ΔR	Δϕ	m	η	m_T
h2ta	⋮	⋮	⋮	⋮	⋮	⋮
h1tawos			⋮	⋮	⋮	⋮
h2tawoff	⋮			⋮	⋮	⋮
h0ta	⋮	⋮	⋮			⋮
t2ta	⋮	⋮	⋮	⋮		
t1ta	⋮	⋮	⋮	⋮	⋮	⋮
t0ta	⋮	⋮	⋮	⋮	⋮	⋮

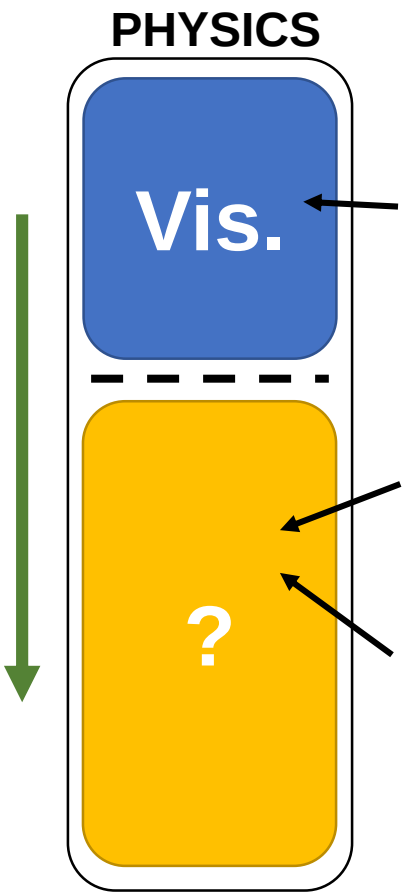
For Each Reconst. ν and ν combinations

OPTIMASS: A Package for the Minimization of Kinematic Mass Functions with Constraints
- Cho, Won Sang *et al.* JHEP 1601 (2016) 026 arXiv:1508.00589

Machine Learning with Topological Augmentation

- Distinctive decay topologies categorized by tau decay kinematics

  ⇒ Characterized by each constraints, c_a , systems and subsystem



❖ **Visible : CMS/ATLAS common features ~ $p_T, \Delta R, \Delta\phi$ …**

>> CMS-HIG 17-002

❖ **H.L. : (Augmented) High Level Features by OPTIMASS**

❖ **R.L. : (Augmented) Raw Level Features by OPTIMASS**

Event Selection

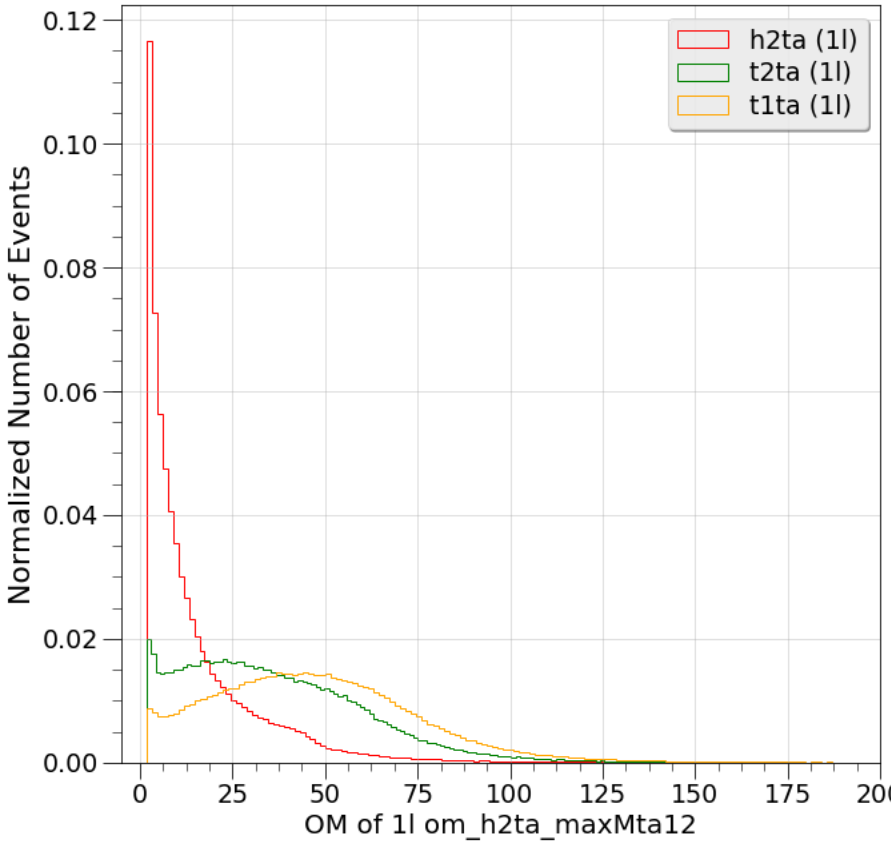
- **Using …**
  - MC : MG5_aMC v2.5.5 + MADSPIN (Decay)
  - Showering : PYTHIA8
  - Detector Simulation : Delphes
  - CT10nlo for all Channel

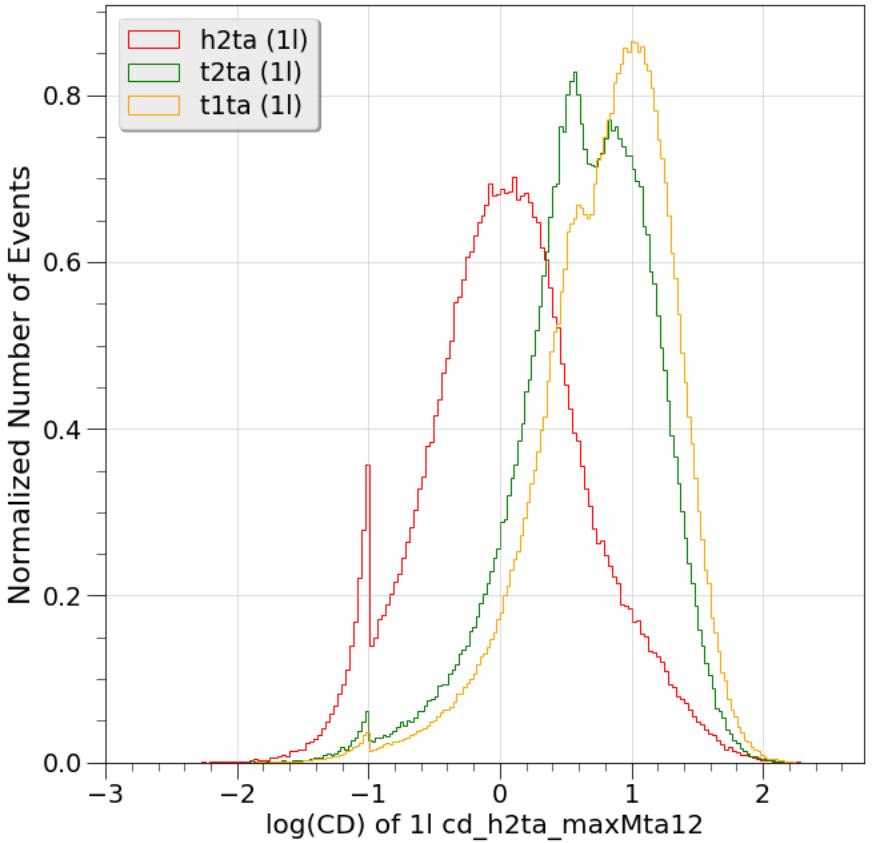
- **Cut …**
  - Delphes : Using delphes_card_ATLAS.dat
  - Tau Tagging : $\Delta R < 0.4$, $\Delta R_{track} < 0.2$, $P_T^{Track} > 1.0$, $P_T^{\tau} > 2.0$, $\eta_{\tau} < 2.5$
    1-prong Eff = 70% , N(>1)-prong eff = 60%
  - Others are same with delphes_card_ATLAS.dat default

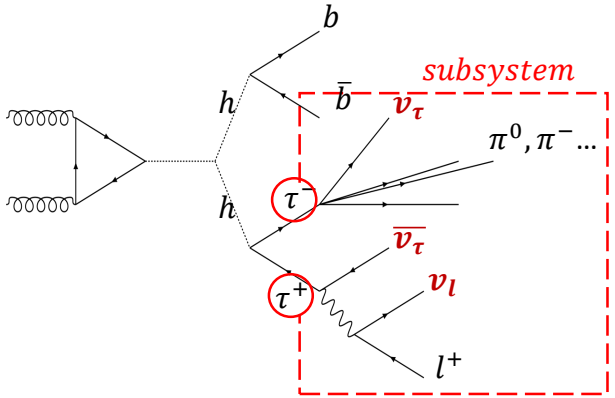
- Hadronic Tau decay using ALL N-prongs

H.L. Result of 1 Lepton

※ H.L. = High Level Features
= OM / CD



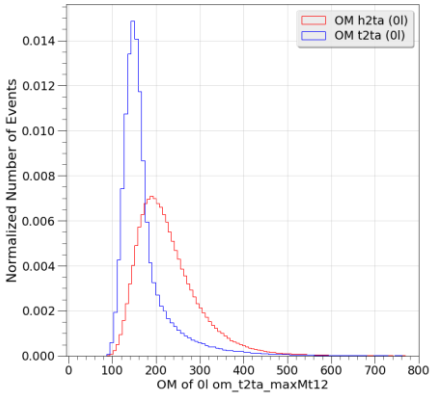


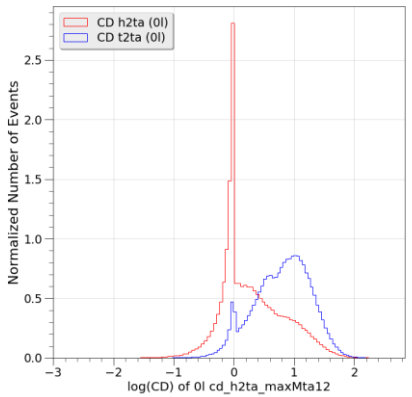


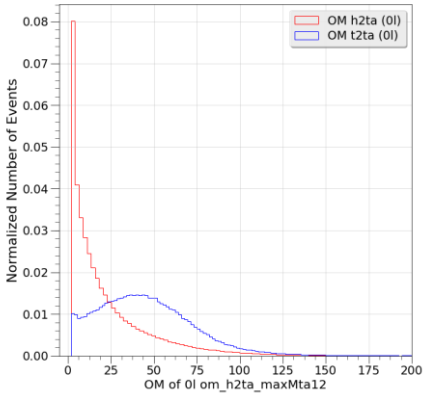
Constraints of Mass

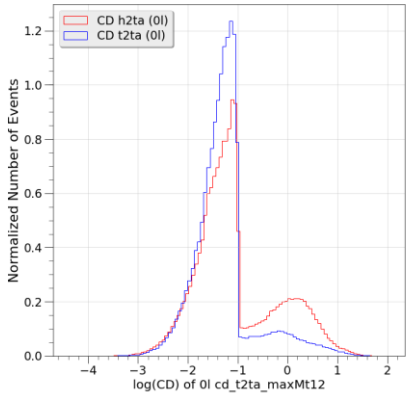
- H(tautau) = 125.
- Tau1 = Tau2

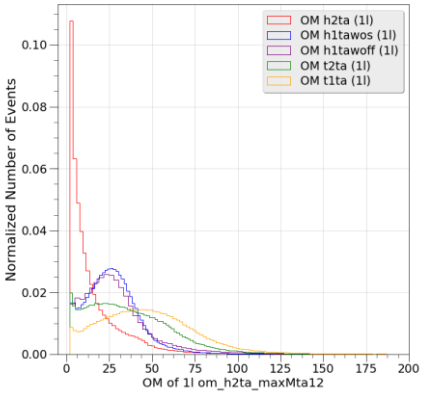
H.L. Result of 0/1 Leptons

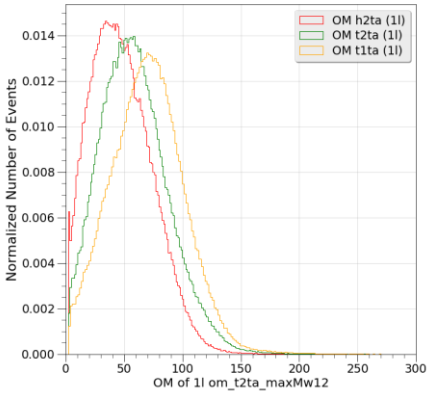


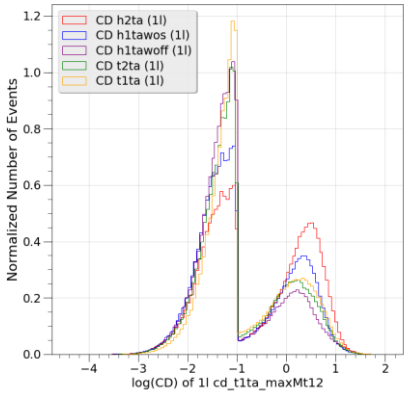


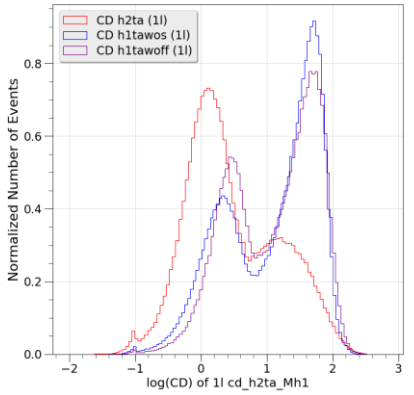


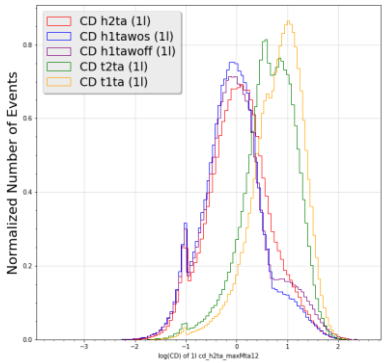


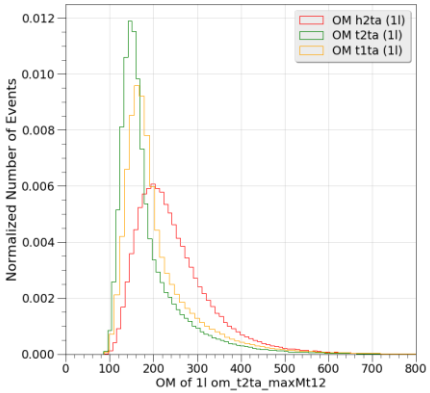


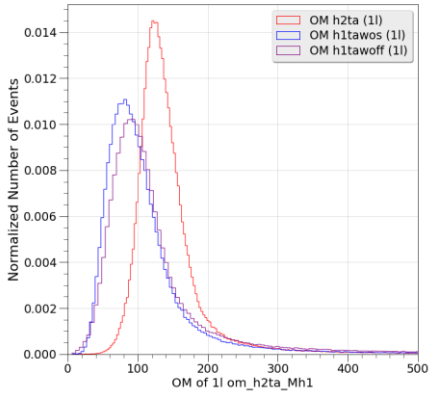


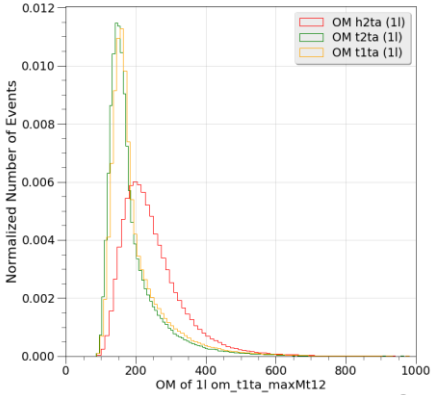






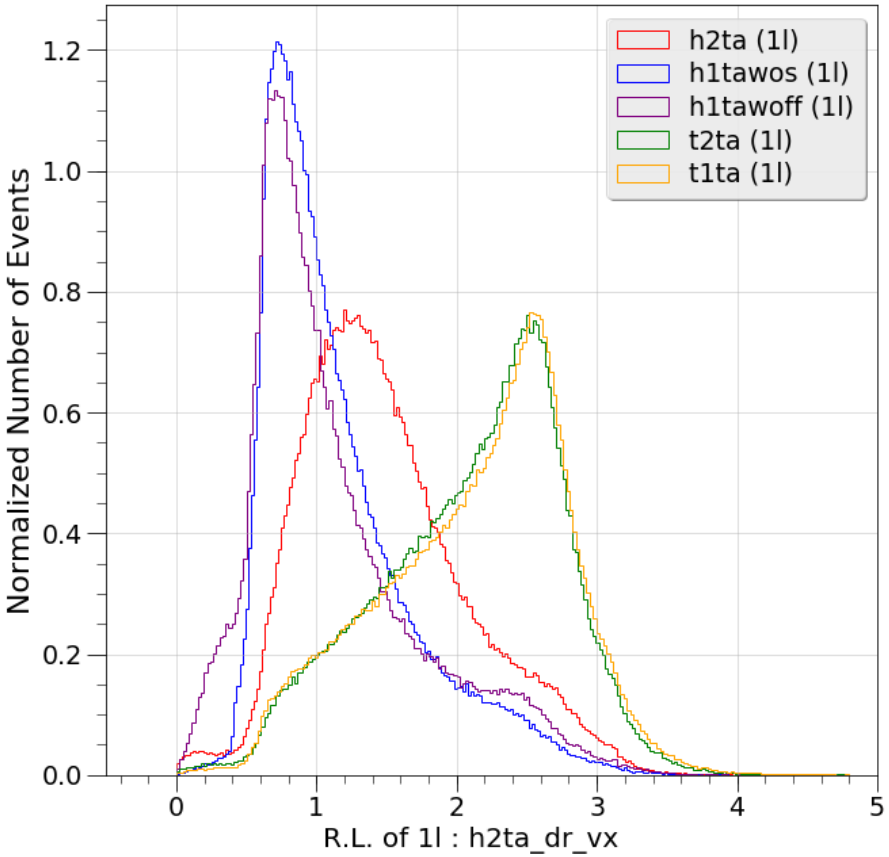


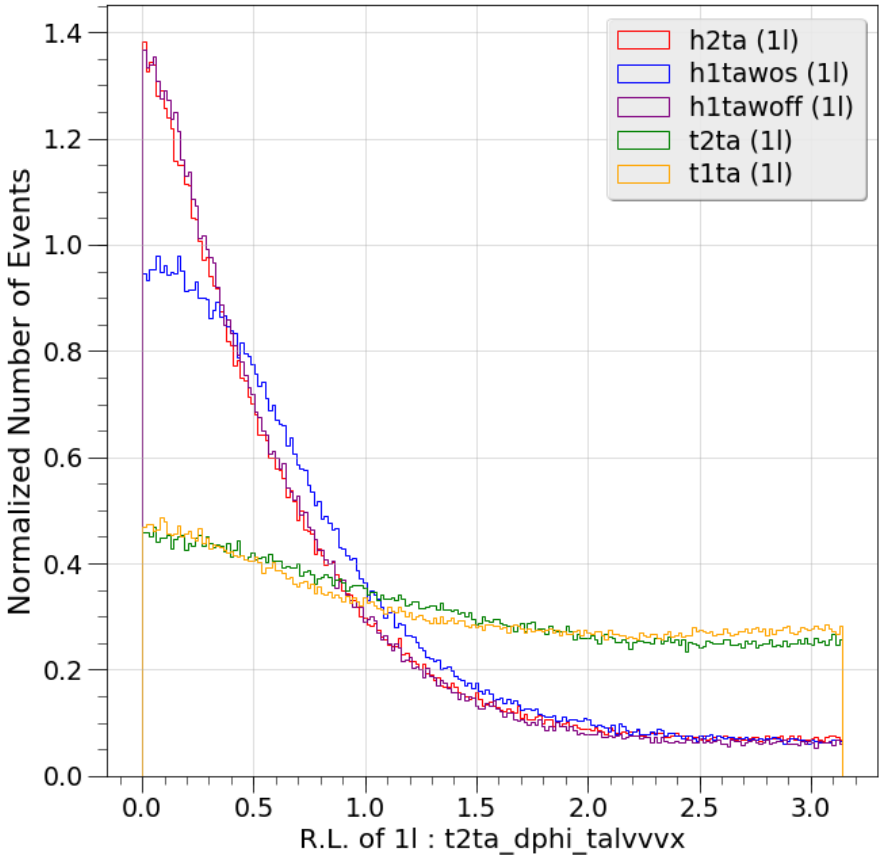


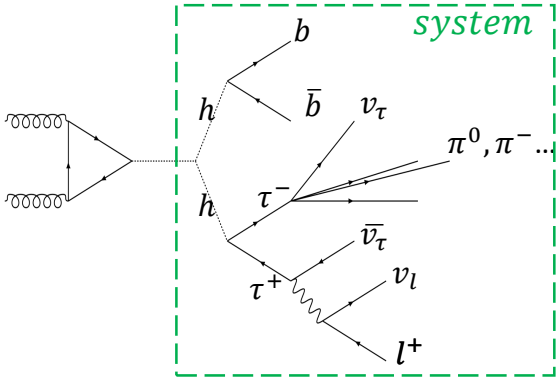


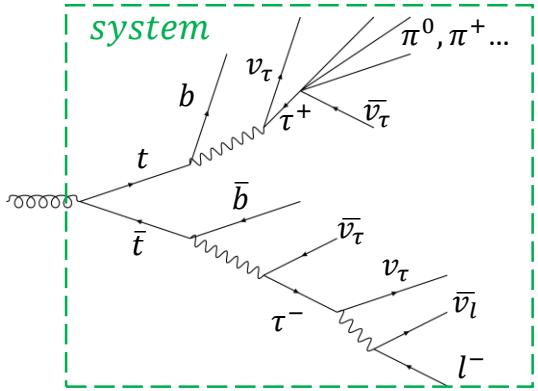
R.L. Result of 1 Leptons

※ R.L. = Raw Level Features
= P_T / M / η / ΔR …



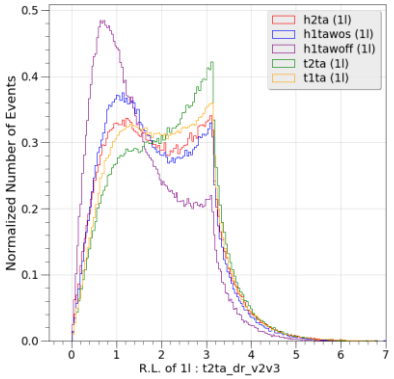


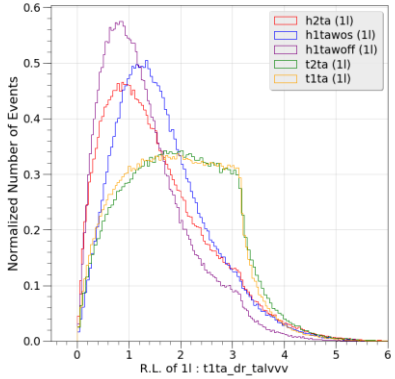


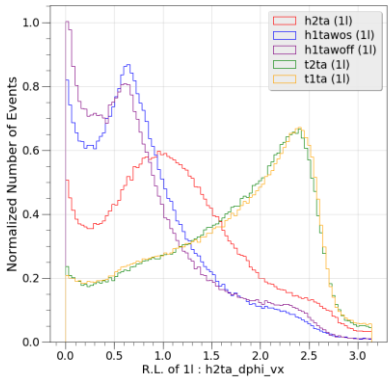


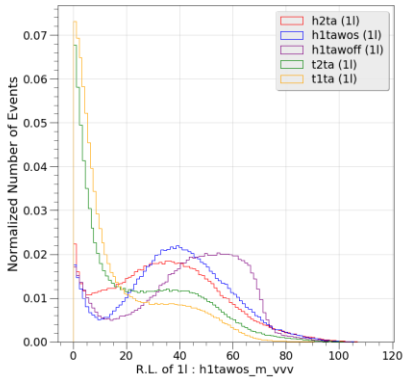
R.L. Result of 0/1 Leptons

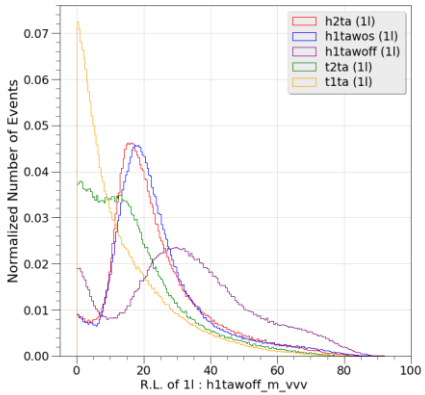
※ R.L. = Raw Level Features
= P_T / M / η / ΔR …

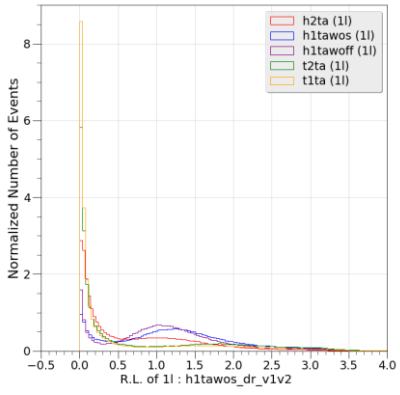


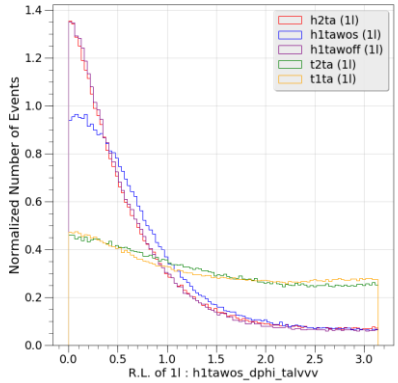


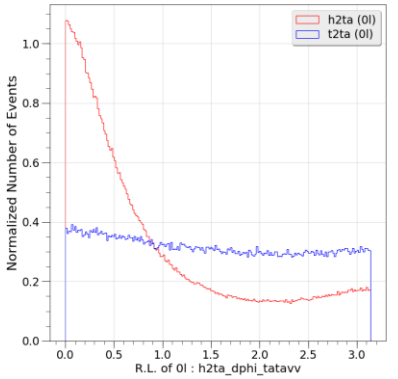


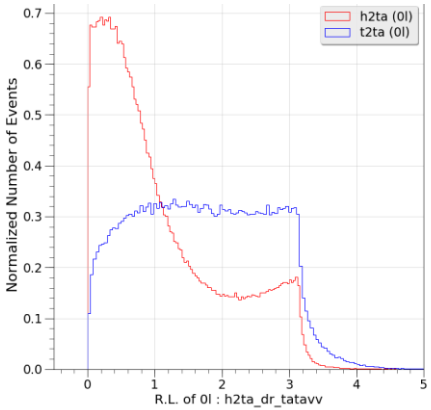


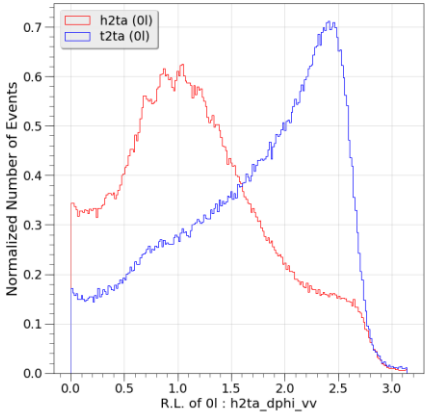


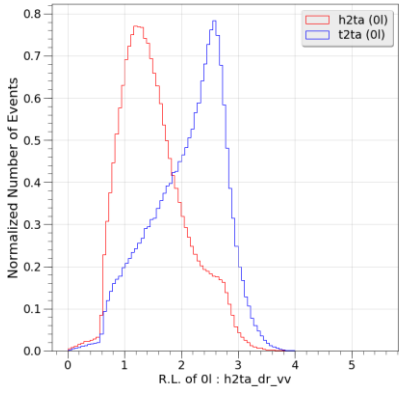


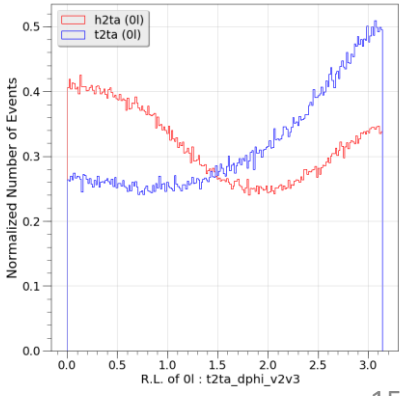






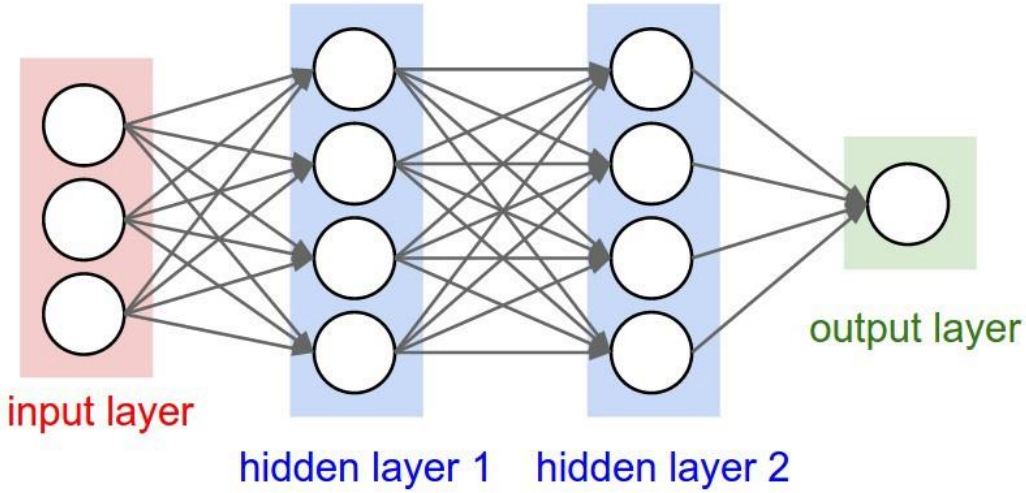




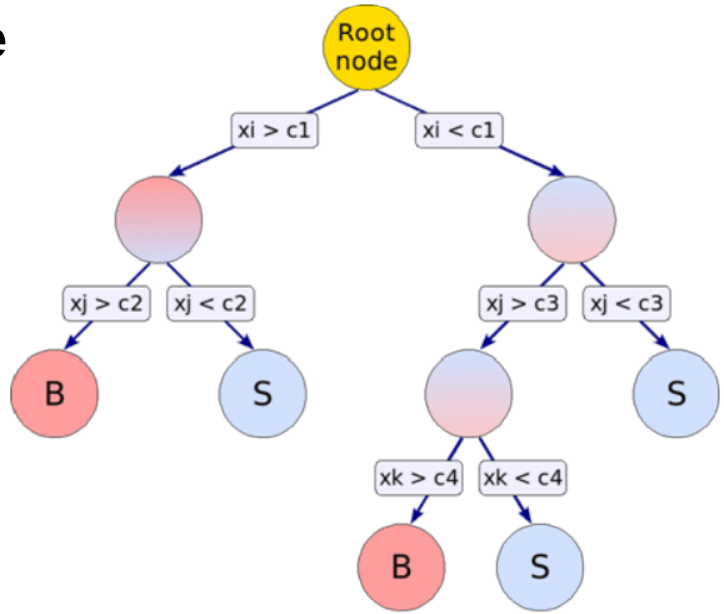


Classifier Model : DNN vs BDT

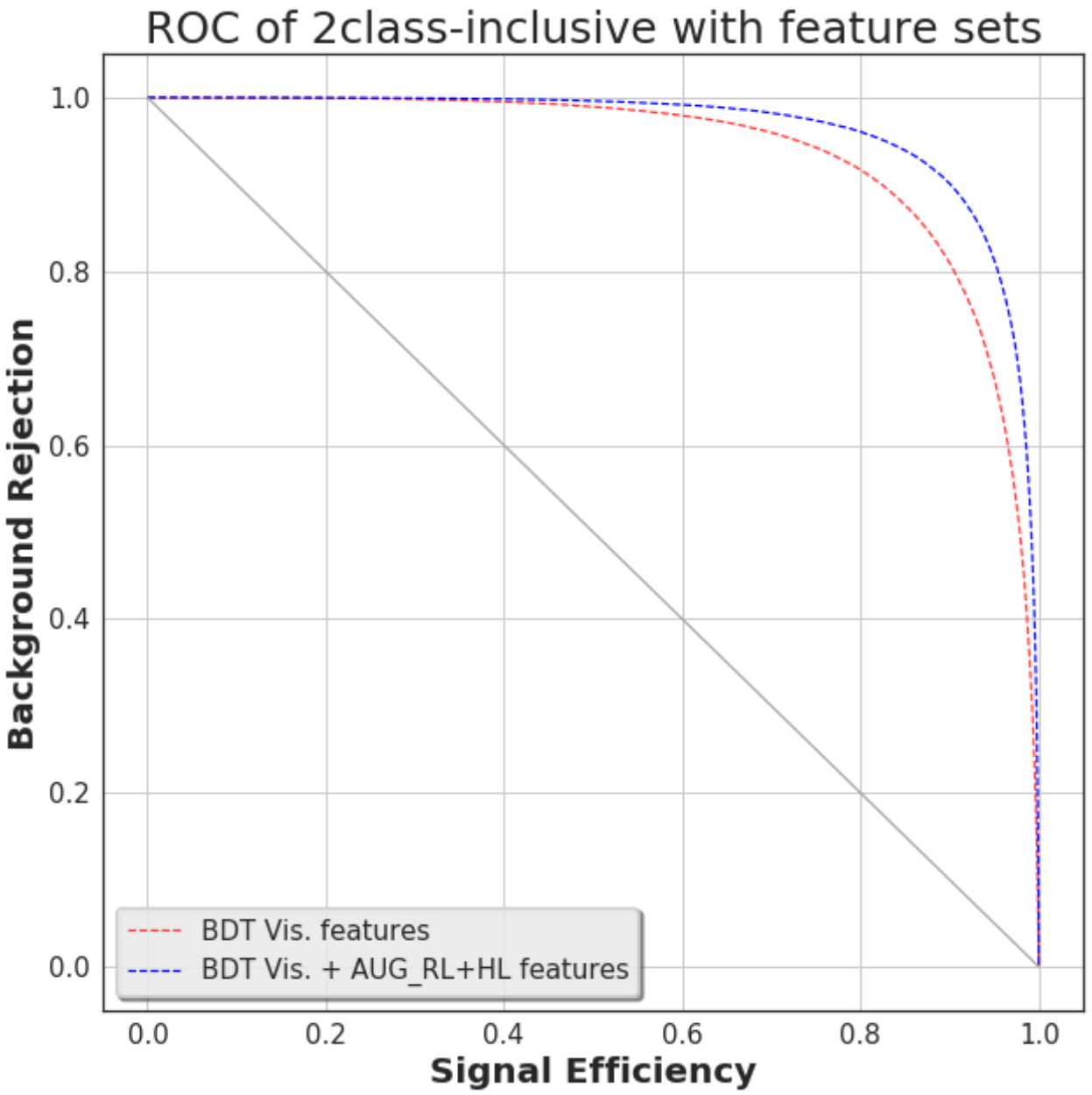
- **Deep Neural Network**



- **Boost Decision Tree**

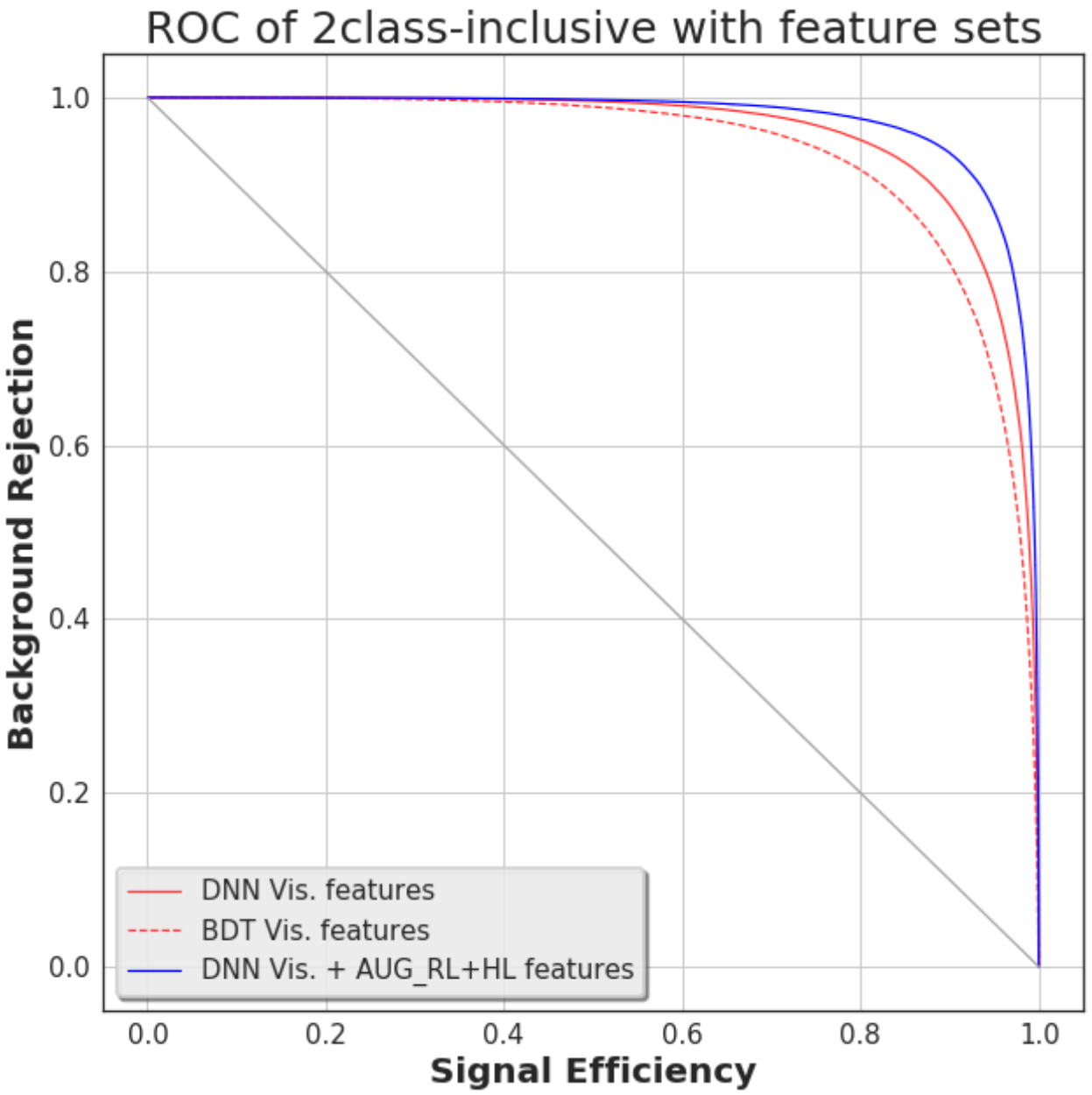


ROC Result for 0 lep (2b+2tau_h)



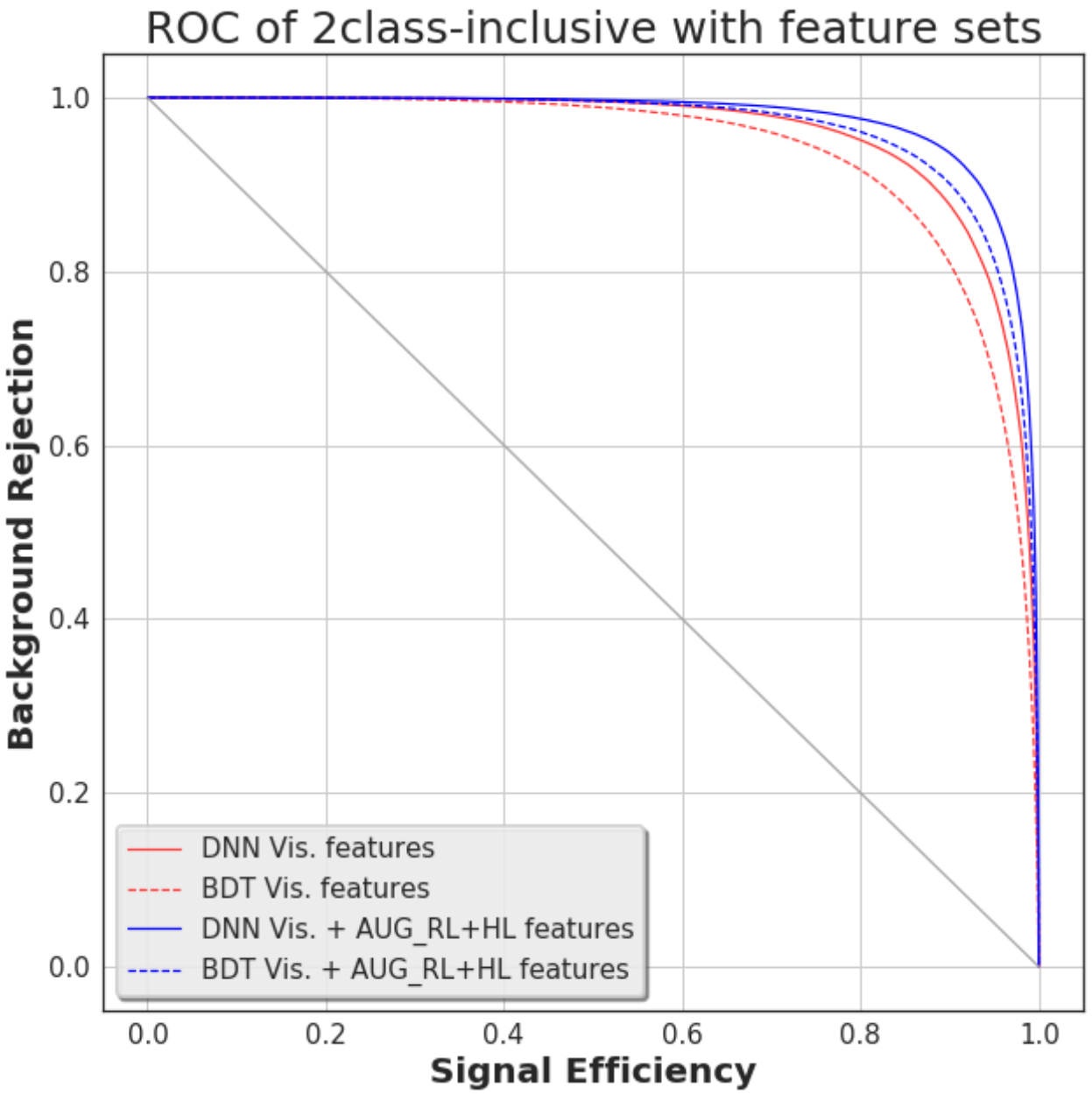
Model	AUC
BDT + Vis	0.9375
BDT + Vis + **R.L. + H.L**	**0.9642**

ROC Result for 0 lep (2b+2tau_h)



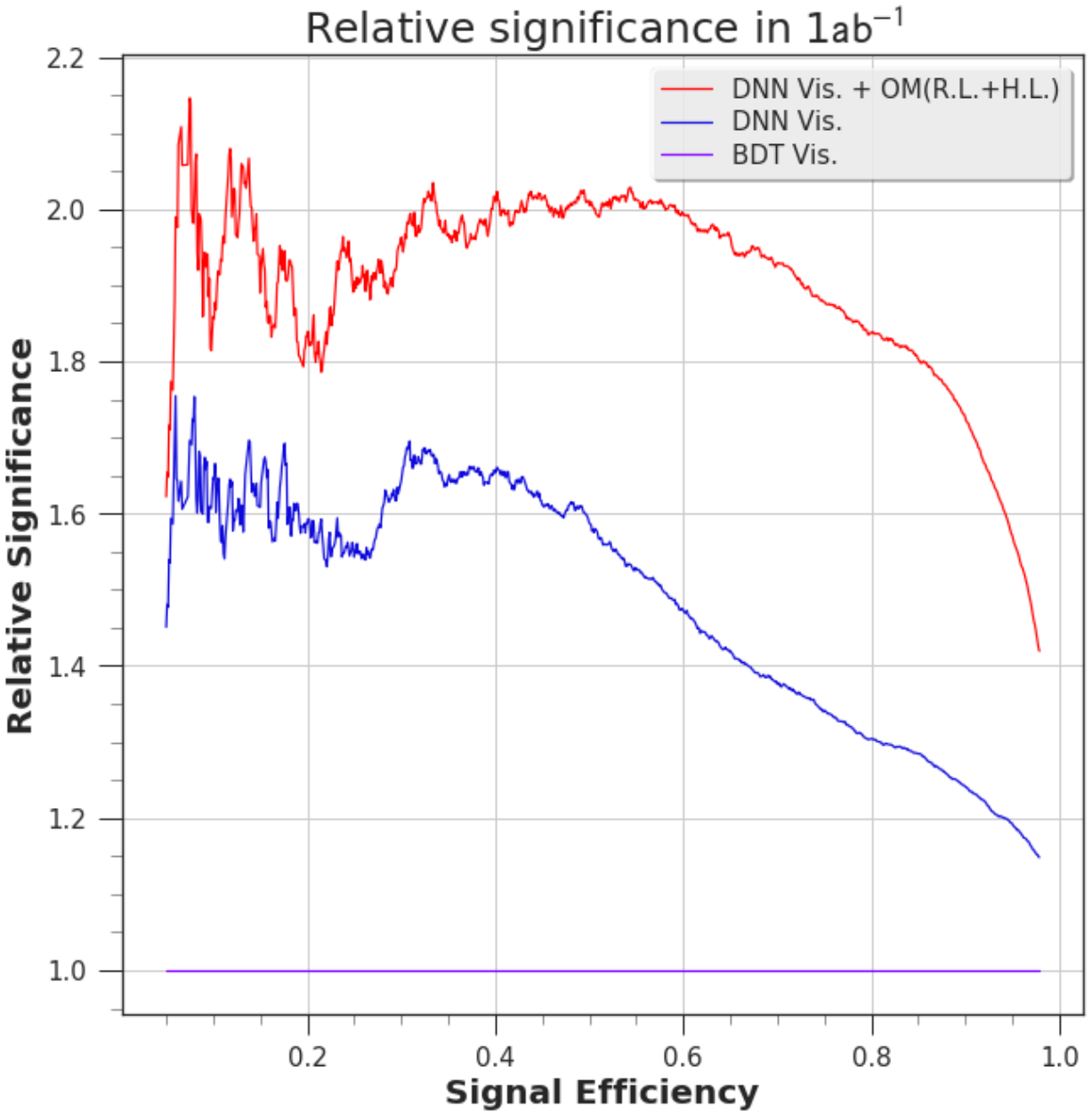
Model	AUC
BDT + Vis	0.9375
DNN + Vis	0.9578
DNN + Vis + **R.L. + H.L**	**0.9744**

ROC Result for 0 lep (2b+2tau_h)

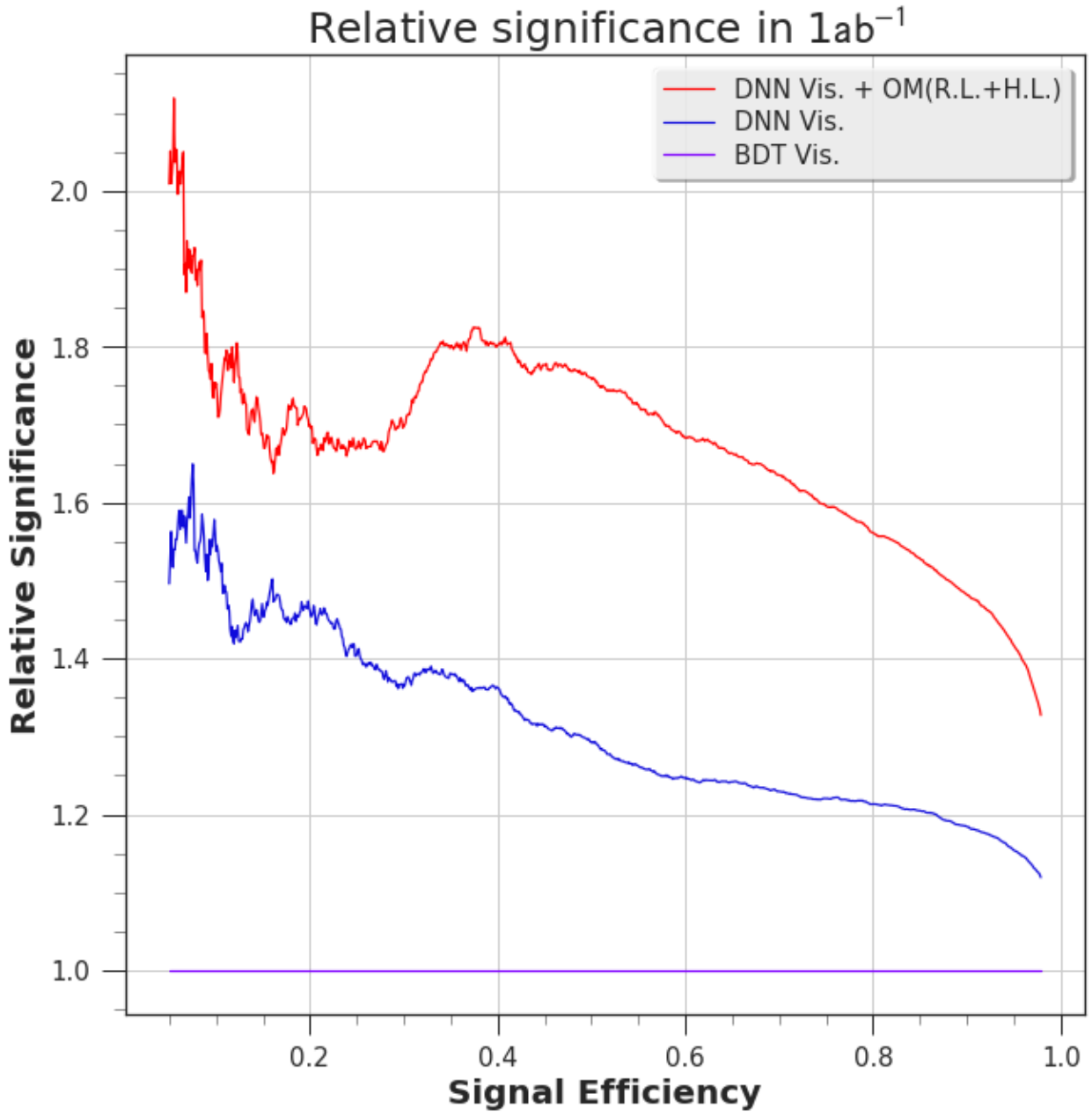


Model	AUC
BDT + Vis	0.9375
DNN + Vis	0.9577
BDT + Vis + **R.L. + H.L**	**0.9642**
DNN + Vis + **R.L. + H.L**	**0.9744**

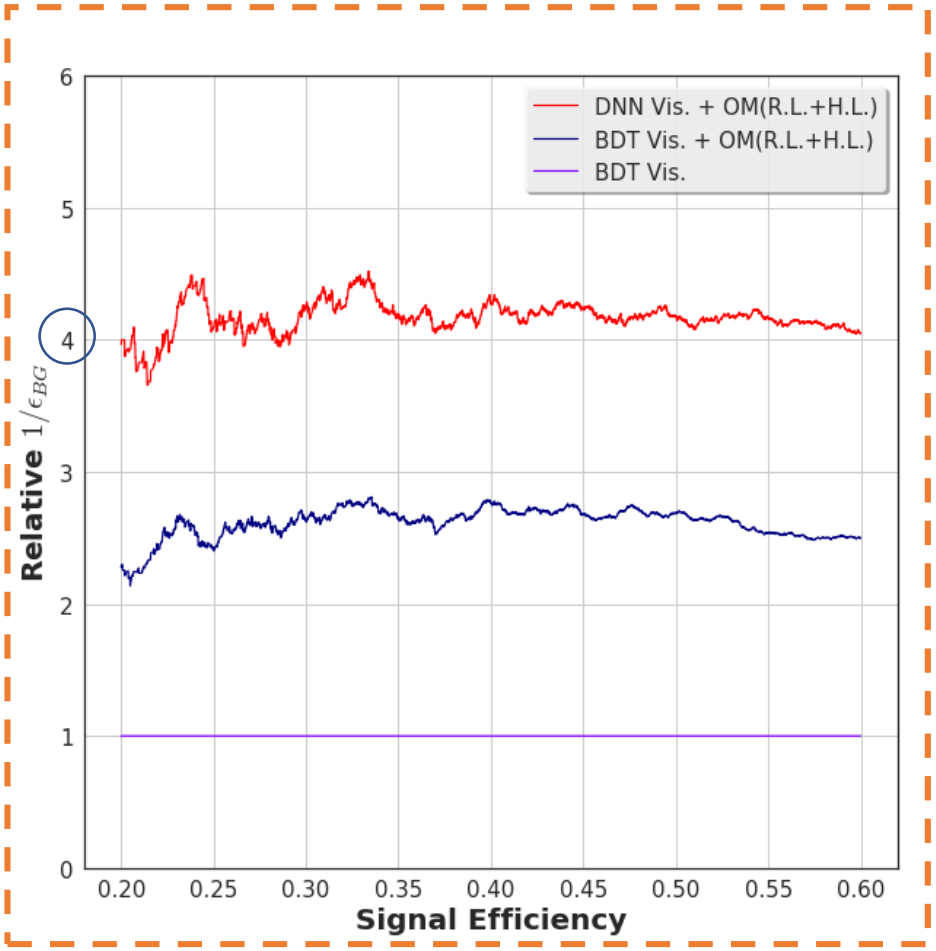
Significance Result for 0 lep (2b+2tau_h)

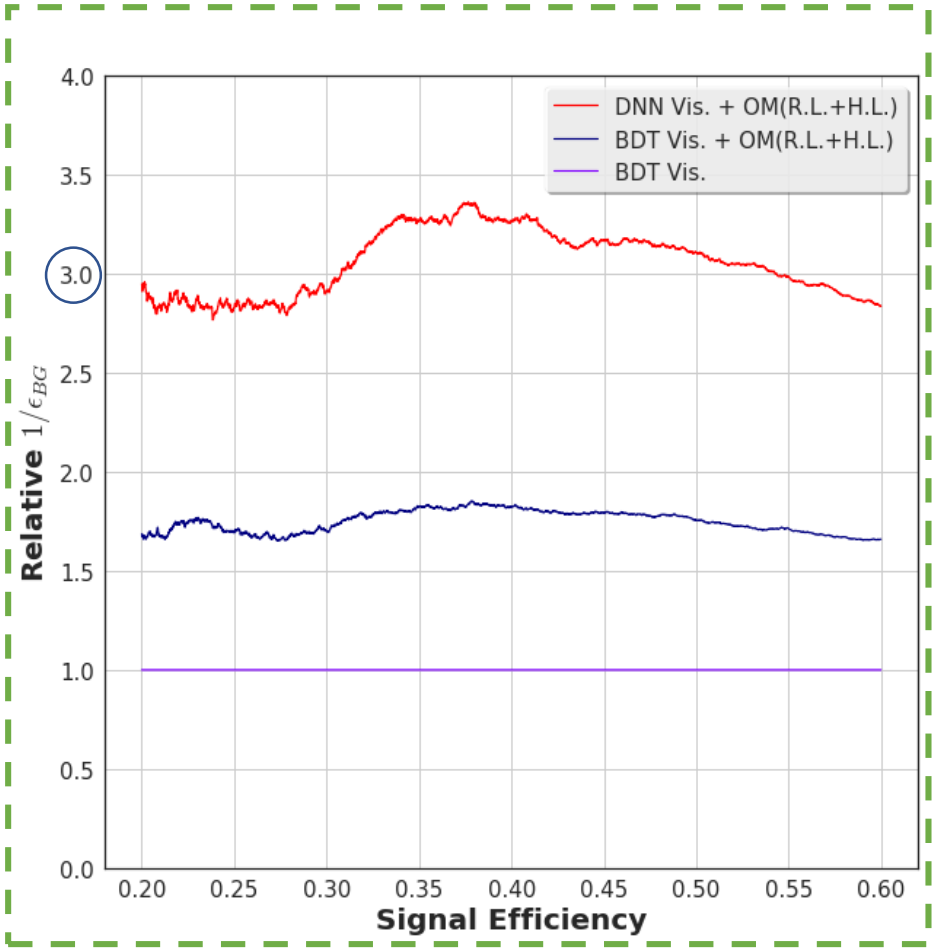


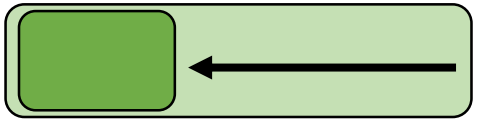
Significance Result for 1 lep (2b+1tau_h)



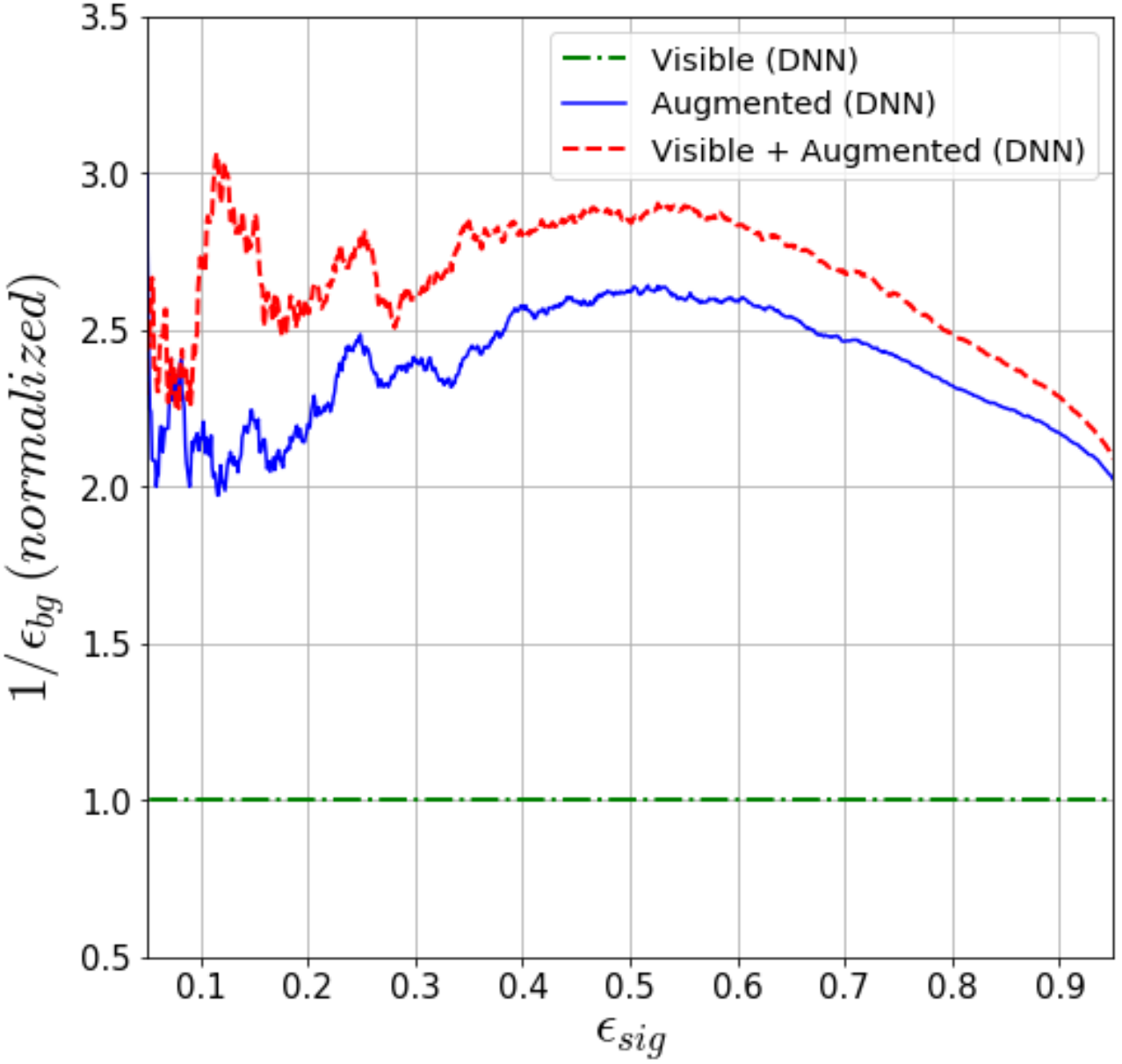
Relative Efficiency Result for 0/1 lep







Relative Efficiency Result for 2 lep



Conclusion

✓ Surveyed **Di-Higgs** searches in **2b+2L(nL+ntau=2)+MET** signatures from **hh->bbWW & bbtautau**

✓ Kinematically categorized signal and bkg processes into **2/5/7 topological classes**, by tau decay kinematics

✓ **Augmented missing data blocks** from invisible d.o.f (in under-constrained system) by all possible physics models involved

✓ And the full augmented data blocks used to **supervised DNN classifier.**

Q&A

BACK UP

❖ **Visible : CMS/ATLAS common features**

- **CMS-HIG 17-002**

- 'dr_tata', 'dr_bb’
- 'dphi_bbtata’, 'dphi_tatamet'', 'dphi_bbmet', 'dphi_ta1met', 'dphi_ta2met'
- 'pt_tata', 'pt_bb’
- 'mt2_bbtata,'mt_ta1', 'mt_ta2'