UC SANTA BARBARA

Preliminary look at HH → $\gamma\gamma\tau\tau$ in nanoAOD

HH Meeting, 30th of March 2021

Matthew Knight, Edward Scott – Imperial College

Claudio Campagnari, Hualin Mei, Francesco Setti - UCSB

Leonardo Giannini, Samuel May, Balaji Sathia-Narayanan, Frank Wuerthwein, Avi Yagil - UCSD

Overview

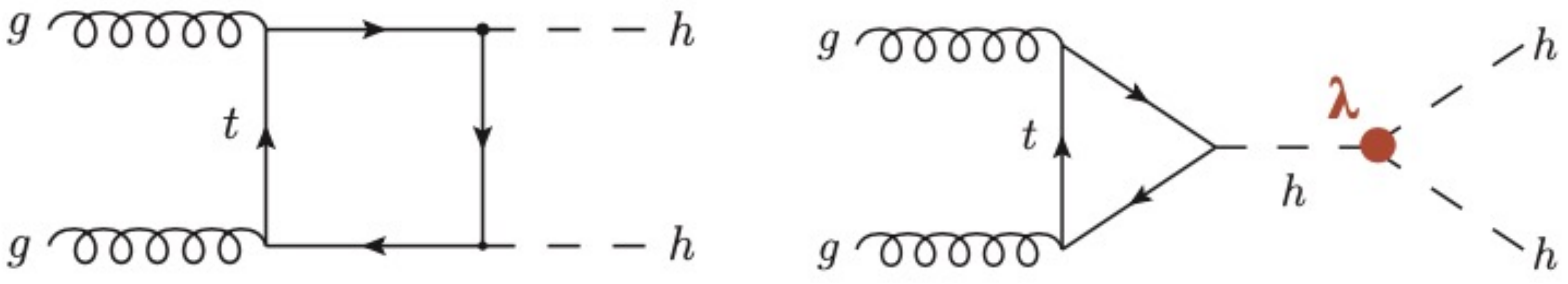
- Di-Higgs production at the LHC
- Inputs & event categorisation
- Multivariate classifier (BDT)
- Signal Region (SR) optimisation
- Extract signal sensitivity

SM HH Production

- Cross section for di-Higgs production from ggF is $\boldsymbol{\sigma(HH)\sim 33}$ **fb** at 13 TeV - ($\sigma(H)\sim 15$ pb)
  - di-Higgs production sensitive to BSM effects

- Higgs pair production probes Higgs trilinear self-coupling

ggF diagrams



Production mode	Cross section (13 TeV)
ggF	$\sim$33 fb
VBF	$\sim$1.7 fb
VHH	$\sim$0.9 fb
ttHH	$\sim$0.8 fb

SM HH Production

- Cross section for di-Higgs production from ggF is $\boldsymbol{\sigma(HH)}$**~33 fb** at 13 TeV - ($\sigma(H)$~15 pb)
  - di-Higgs production sensitive to BSM effects

- Higgs pair production probes Higgs trilinear self-coupling

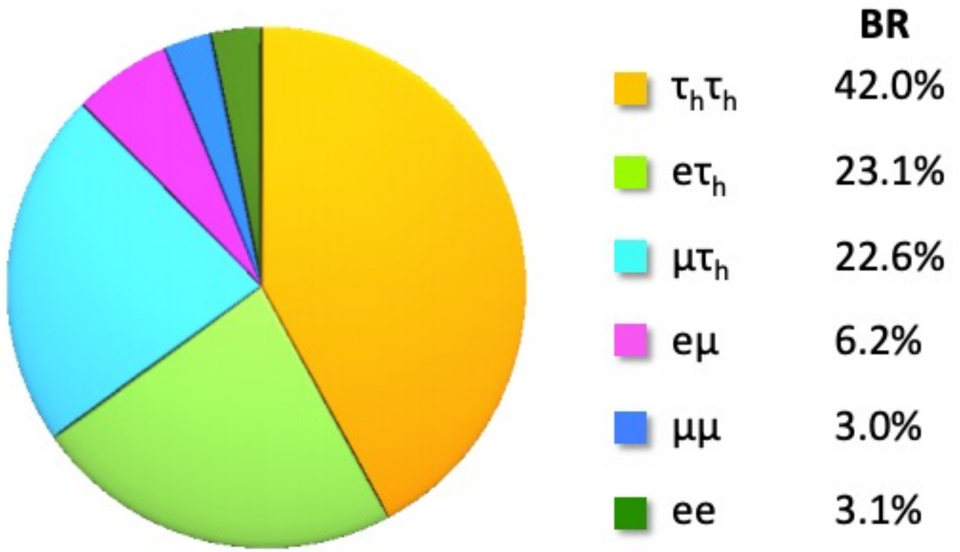
- HH $\rightarrow \gamma\gamma\tau\tau$ final state has not been studied yet

- Higgs branching ratios
  - $\mathcal{B}r(H \rightarrow b\bar{b}) = 58\%$
  - $\mathcal{B}r(H \rightarrow \tau\tau) = 6\%$
  - $\boldsymbol{\mathcal{B}r(HH \rightarrow \gamma\gamma\tau\tau) = 0.027\%}$
  - $\mathcal{B}r(HH \rightarrow \gamma\gamma b\bar{b}) = 0.26\%$
  - $\mathcal{B}r(HH \rightarrow b\bar{b}\tau\tau) = 7\%$

Production mode	Cross section (13 TeV)
ggF	~33 fb
VBF	~1.7 fb
VHH	~0.9 fb
ttHH	~0.8 fb

τ Decays

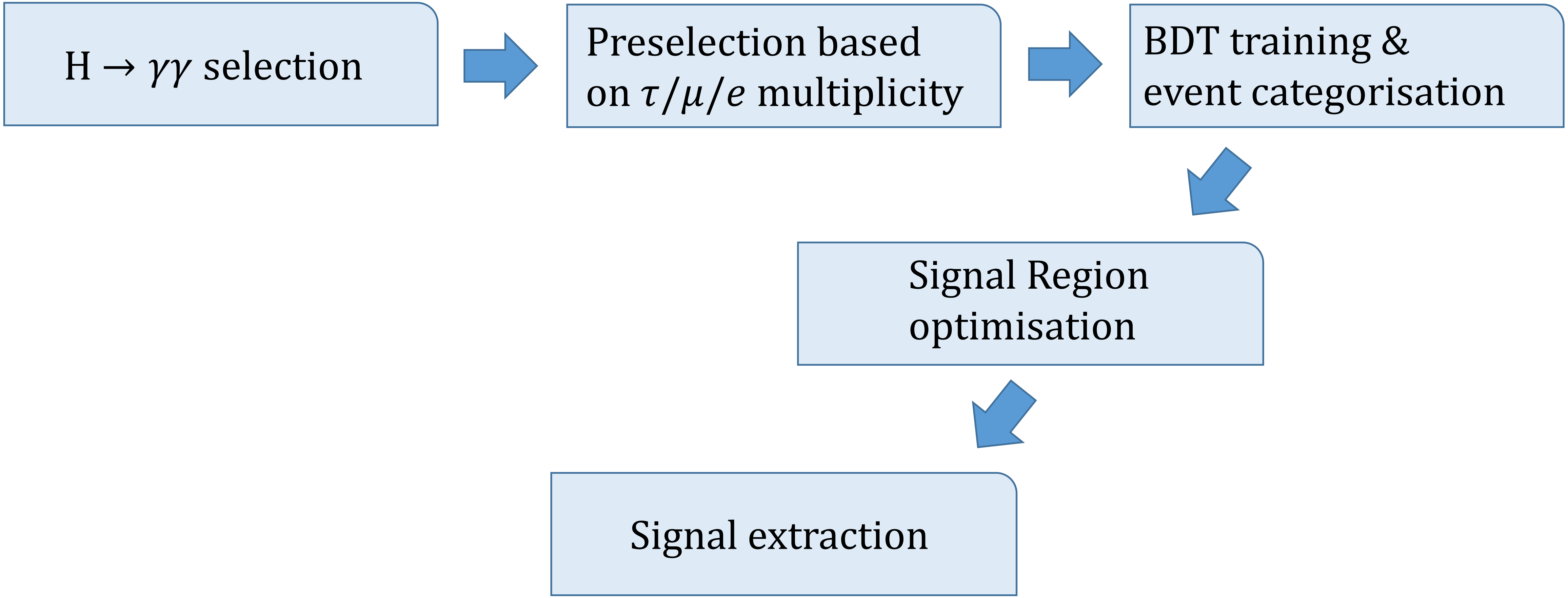
- Dominant decay modes:
  - $\tau \to e\overline{\nu_e}\nu_\tau$ ($\mathcal{BR}$ = 17.8%)
  - $\tau \to \mu\overline{\nu_\mu}\nu_\tau$ ($\mathcal{BR}$ = 17.4%)
  - $\tau \to h^-\nu_\tau + \geq 0\ h^0$ ($\mathcal{BR}$ = 49.5%)
  - $\tau \to h^-h^-h^+\nu_\tau + \geq 0\ h^0$ ($\mathcal{BR}$ = 15.2%)



- Categories of interest:
  - $\tau\,\tau$ $\rightarrow$ both taus decay hadronically
  - $\tau\,l$ $\rightarrow$ one hadronic τ , one leptonic ($l = e, \mu$)
  - exclude events without τ (e.g. eμ) due to conflicts with other analyses

- $\tau\tau + \tau e + \tau\mu$ cover ~ 88% of all τ decays

Analysis Flow



Inputs

- Use of private production of nanoAOD samples for Run II
  - $HH \rightarrow \gamma\gamma\tau\tau$ sample made from modifying Pythia configuration of official $HH \rightarrow \gamma\gamma b\bar{b}$ sample

- Currently not using ultra-legacy (to be implemented)

Data

- DoubleEG 2016 & 2017
- EGamma 2018

Total Luminosity
$35.9 + 41.5 + 59.8 = 137.2\ fb^{-1}$

Signal

- GluGluToHHTo2TAU2G_* (LO)

Background

Non resonant	Resonant
• Diphoton + Jets • W/Z + Gamma • Gamma + Jets • DY + Jets • QCD • TT + Jets/γ/$\gamma\gamma$	• VH

Physics Objects & Diphoton Selection

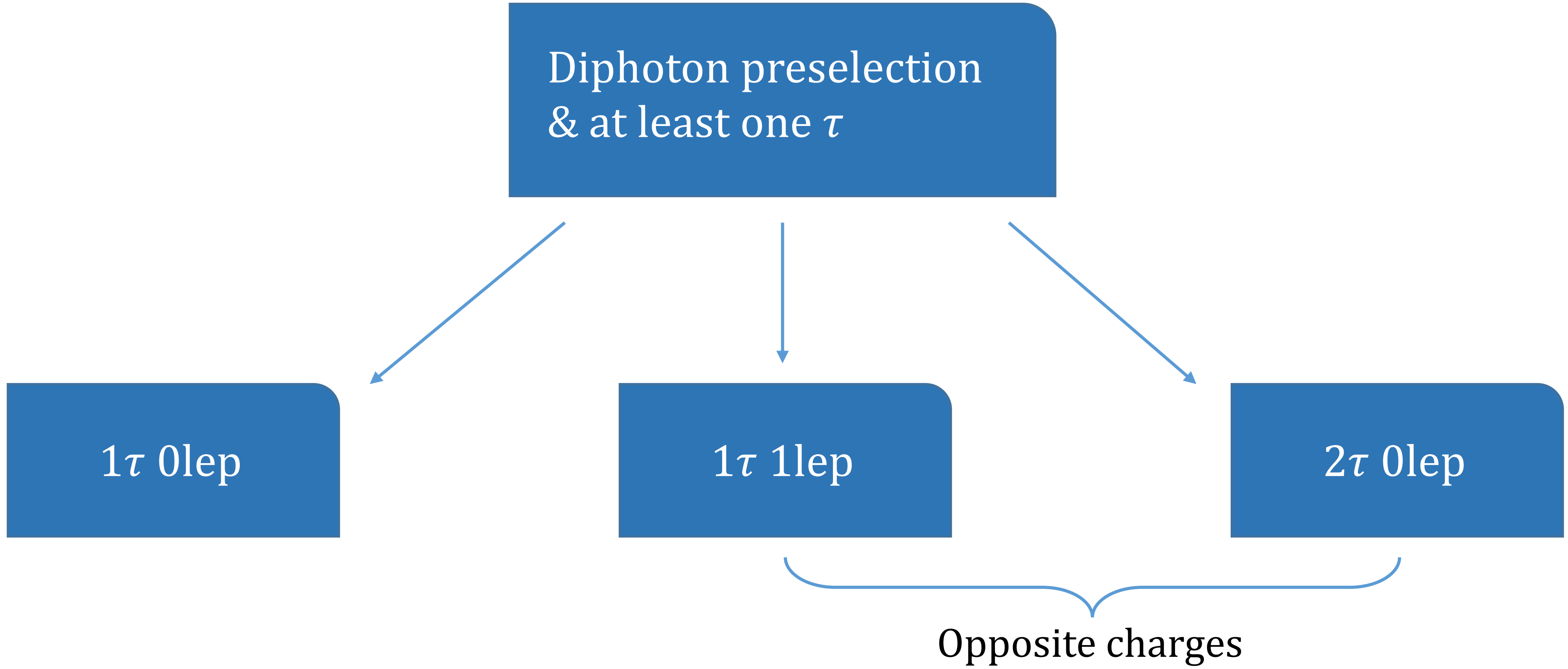
- τ
  - $p_T > 20$ GeV , $|\eta| < 2.3$
  - DecayModeNewDMs
  - $|d_z| < 0.2$ cm
  - deepTauID: vs jet(Loose), vs e(VVLoose), vs μ: VLoose
- μ
  - $p_T > 10$ GeV , $|\eta| < 2.4$
  - $|d_{xy}| < 0.045$ cm , $|d_z| < 0.2$ cm
  - PFRelIso03 < 0.3
- e
  - $p_T > 10$ GeV , $|\eta| < 2.5$
  - $|d_{xy}| < 0.045$ cm , $|d_z| < 0.2$ cm
  - (PFRelIso03 < 0.3 and mvaFall17V2noIso_WP90) or (mvaFall17V2Iso_WP90)

- Standard diphoton preselection plus:
  - $\frac{p_T}{m_{\gamma\gamma}} > 0.33$ (0.25)
  - photon ID $>$ - 0.7

- Overlap removal:
  - $\Delta R > 0.2$ between $\tau/\mu/e$ and selected γs
  - $\Delta R > 0.2$ between pairs of $\tau/\mu/e$

- $\tau/\mu/e$ selection based on loose selection from HH → $bb\tau\tau$ analysis

Event Categorisation



Yields for Run II

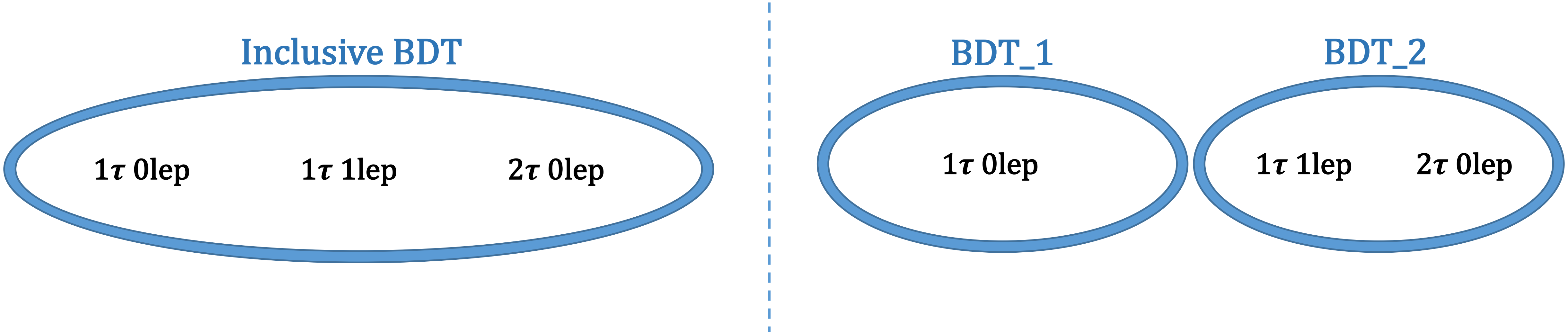
- Preliminary yield table at preselection level

			Inclusive	1tau_0lep	1tau_1lep	2tau_0lep
Bkg.	Non res.	GJet	10937±921	10934±921	3.11±2.53	0.0 ±0.0
		ttbar+ttG/GG	247 ±19	220±18	25.5 ± 5.57	2.1 ±1.2
		ZG	329±10	291±10	27.0±3.0	11.3±1.8
		WG	314±19	309±19	5.42±2.54	0.32±1.26
		Diphoton	6837±35	6801±35	7.77±1.02	27.3±2.22
	Res.	VH	8.96± 0.12	8.34±0.11	0.40±0.03	0.21±0.02
	Inclusive	Bkg	18675±922	18564±922	69.3±7.3	41.2±3.4
		Data	21338±146	21245±145	50.0±7.1	38.0±6.2
		$HH \to \gamma\gamma\tau\tau$	0.298±0.001	0.169±0.000	0.074±0.000	0.054±0.000

- **Work in progress** – improve data/MC agreement by exploring missing backgrounds & apply relevant SFs

Preliminary BDT

- Explored two distinct scenarios:
  - Train a **single BDT** inclusive of **all categories** with at least one τ
  - Train **two BDTs**, one for **1τ 0lep** events and one for **1τ 1lep & 2τ 0lep** events



- BDTs from XGBoost (python) and trained with:
  - Signal: HH $\rightarrow \gamma\gamma\tau\tau$
  - Background: Diphoton - Zγ - VH - Wγ - ttγ - tt$\gamma\gamma$ - tt

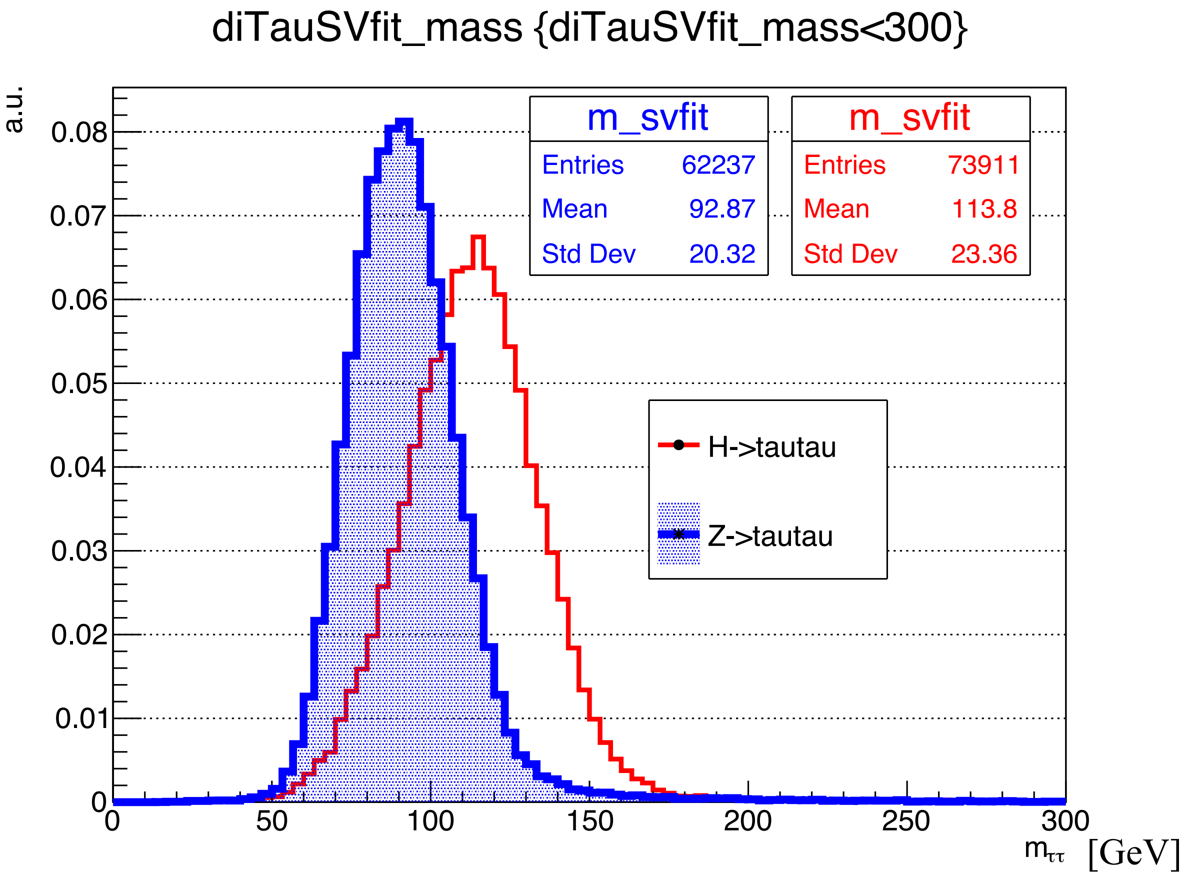
BDT Training Variables

n_taus	sublead. photon $\frac{p_T}{m_{\gamma\gamma}}$	tau_2 pt	jet1_bTagDeepFlavB	$\tau\tau$ pt (SVFit)
n_electrons	sublead. photon η	tau_2 η	jet2 pt	$\tau\tau$ η (SVFit)
n_muons	sublead. photon ID_mva	tau1_id_vs_electron	jet2 η	$\tau\tau$ mass (SVFit)
n_Jets	electron pt	tau1_id_vs_muon	jet2 ID	$\tau\tau$ ϕ (SVFit)
MET pt	electron η	tau1_id_vs_jet	jet2_bTagDeepFlavB	
diphoton $\frac{p_T}{m_{\gamma\gamma}}$	electron tight_ID	tau2_id_vs_electron	$\tau\tau$ ΔR (visible)	
diphoton η	muon pt	tau2_id_vs_muon	$\tau\tau$ ΔR (SVFit)	
diphoton ΔR	muon η	tau2_id_vs_jet	ΔR between $\tau\tau$ (vis) and $\gamma\gamma$	
lead. photon $\frac{p_T}{m_{\gamma\gamma}}$	muon tight_ID	jet1 pt	ΔR between $\tau\tau$ (SVFit) and $\gamma\gamma$	
lead. photon η	tau_1 pt	jet1 η	$\Delta\phi$ between $\tau\tau$ (vis) and $\gamma\gamma$	
lead. photon ID_mva	tau_1 η	jet1 ID	$\Delta\phi$ between $\tau\tau$ (SVFit) and $\gamma\gamma$	

- Link to configuration BDT file (full list of training variables and hyperparameters)

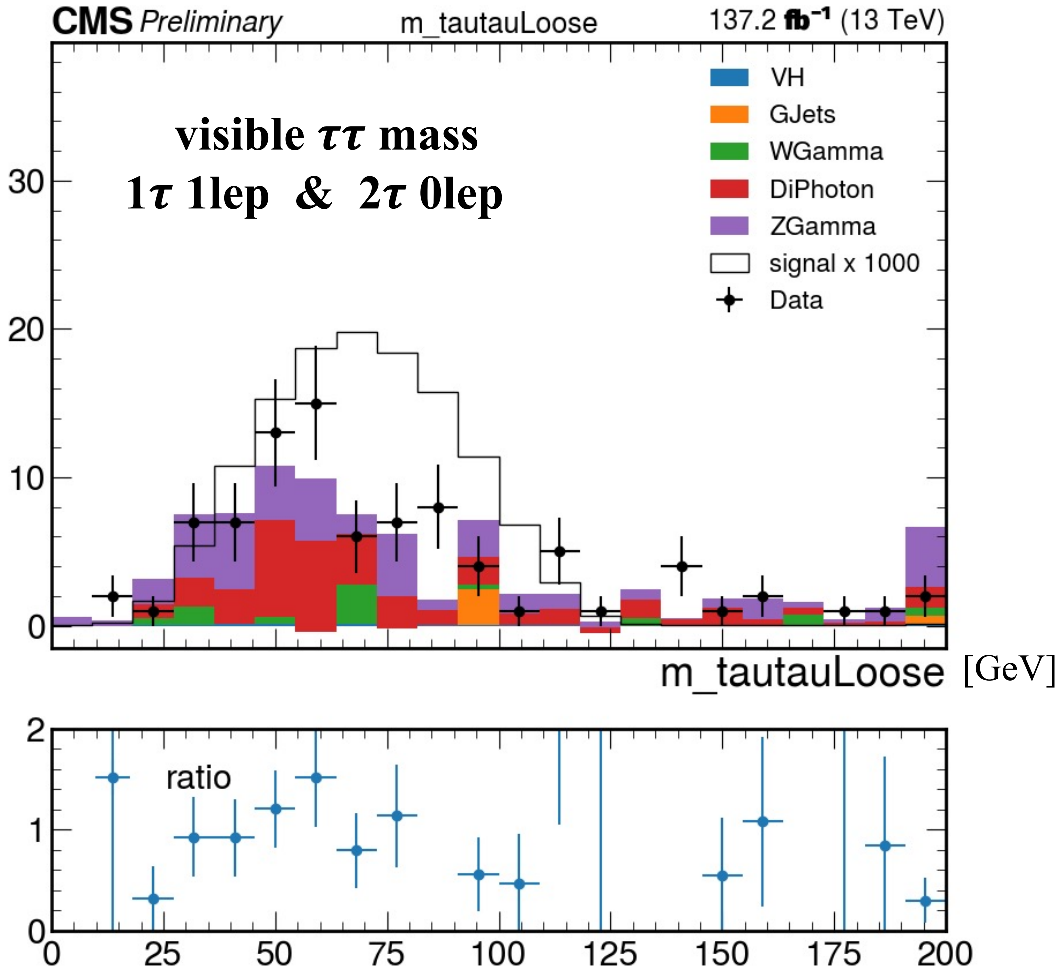
SVFit Tools

- Use of di-tau SVFit invariant mass in BDT
- SVFit tools use Likelihood method to reconstruct di-tau invariant mass from:
  - MET & MET covariant matrix
  - $\tau / e / \mu$ visible Lorentz vectors
- Widely used in CMS analyses, like $HH \rightarrow bb\tau\tau$
- Powerful variable to separate signal from background (requires both $\tau / e / \mu$ reconstructed)

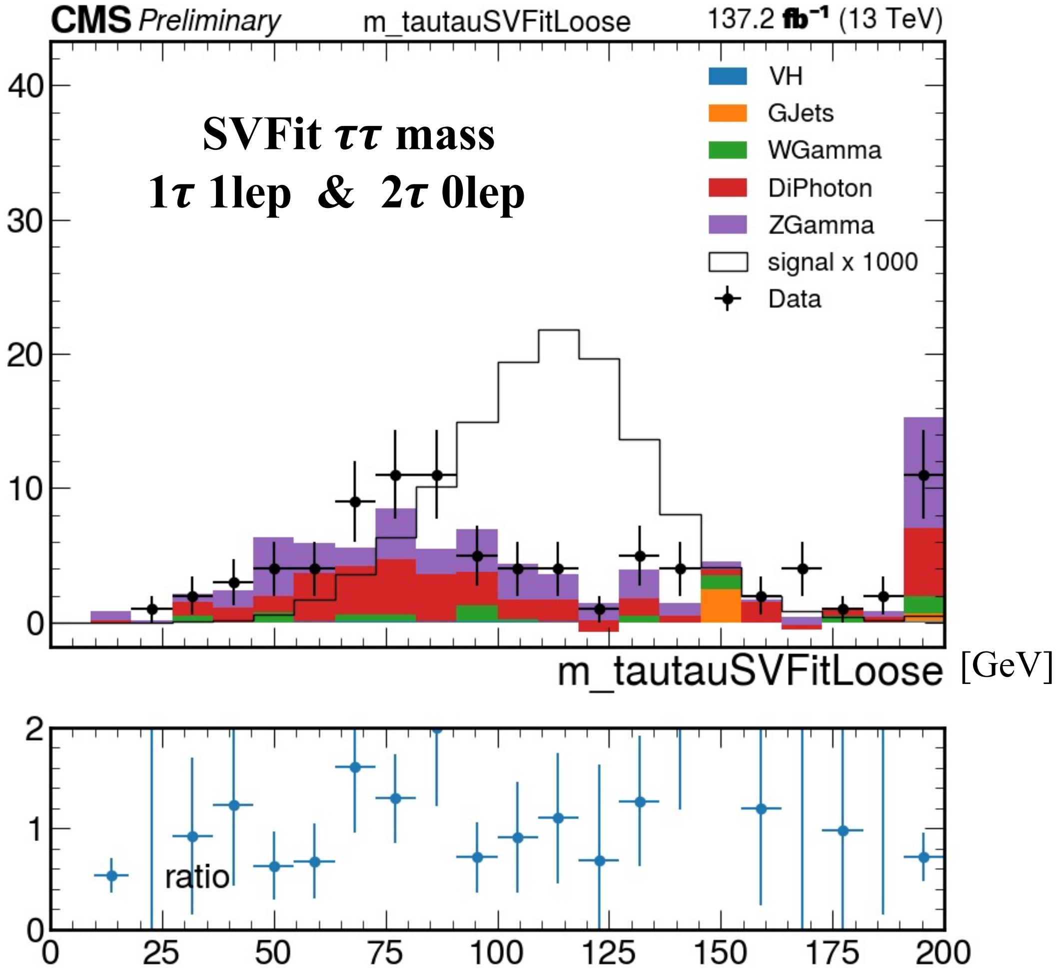


SVFit invariant mass for a system of two hadronic taus.

SVFit Tools

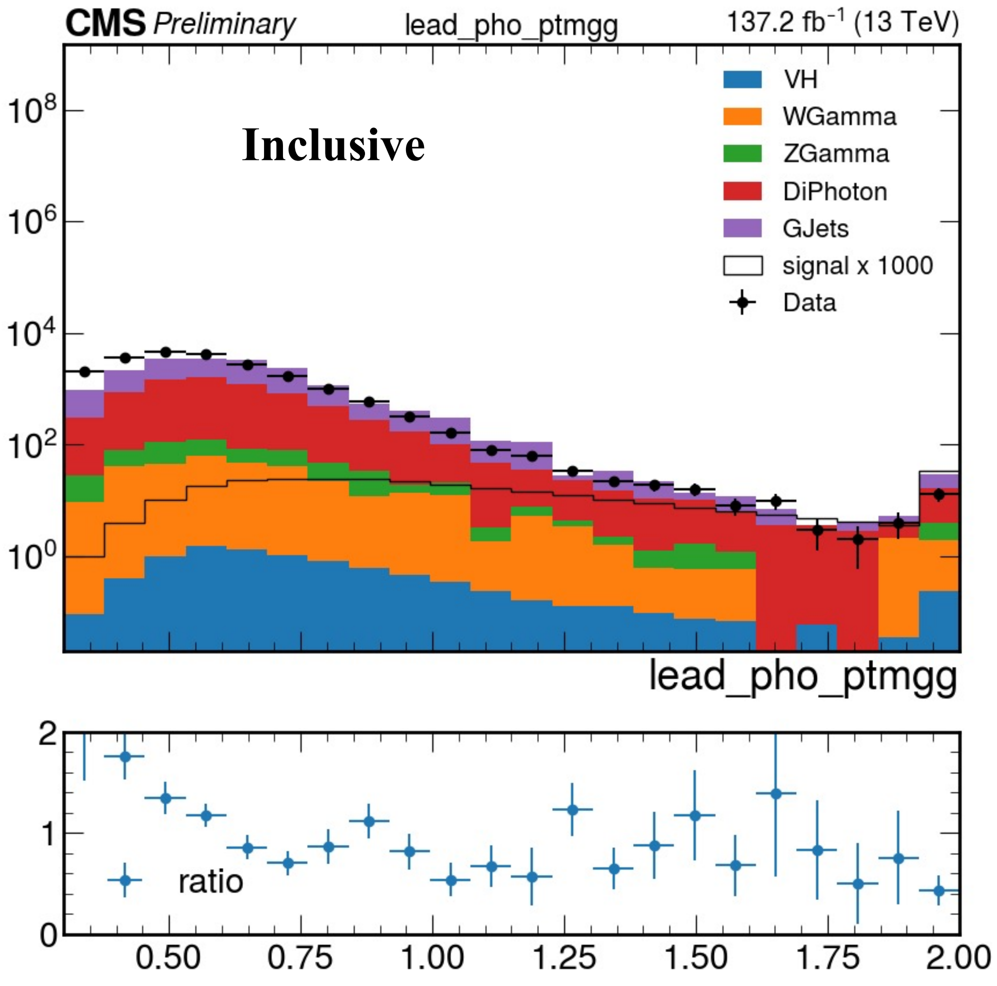


Invariant mass for the $\tau/\mu/e$ candidates from Higgs computed from their **visible** Lorentz vectors. Since we require both Higgs candidates to be reconstructed, only 1τ 1lep & 2τ 0lep events are shown.

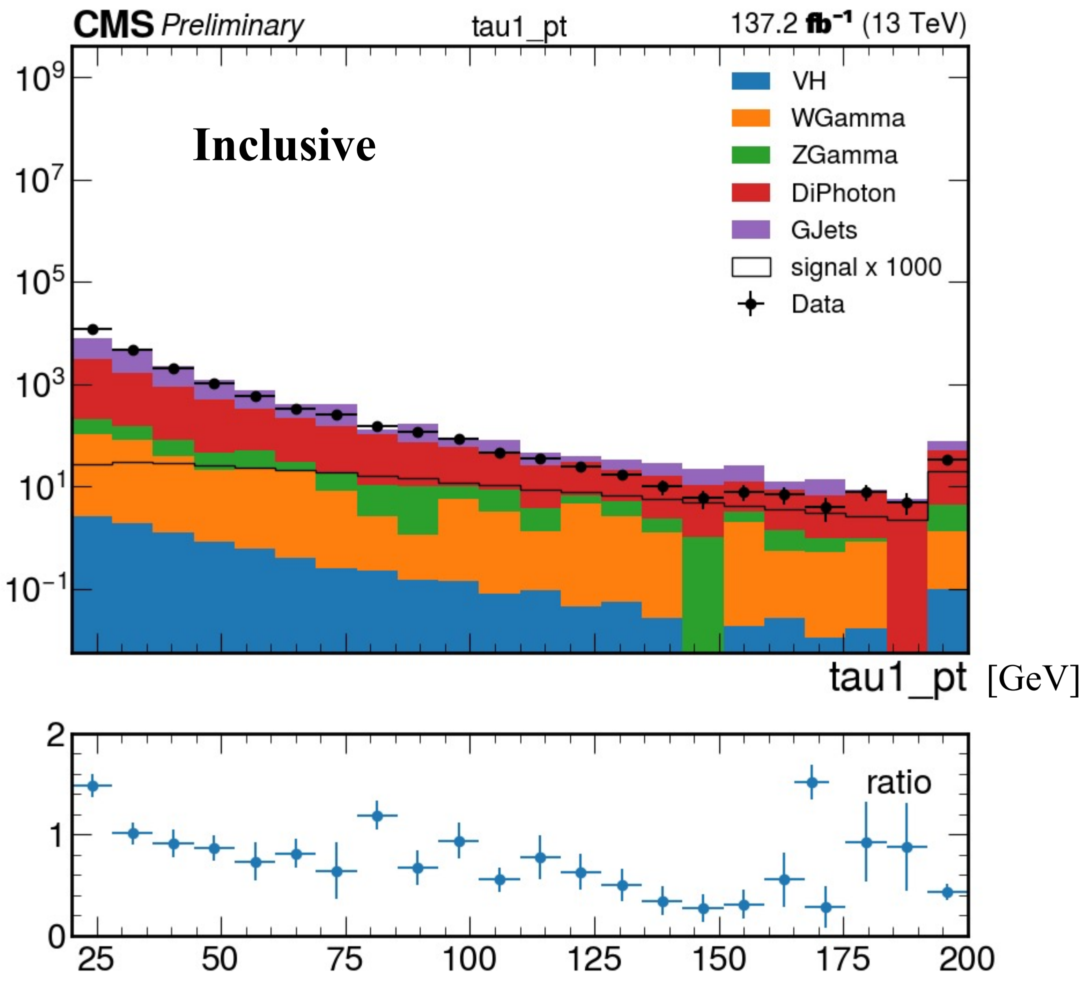


Invariant mass for the $\tau/\mu/e$ candidates from Higgs computed according to the **SVFit** algorithm. Since we require both Higgs candidates to be reconstructed, only 1τ 1lep & 2τ 0lep events are shown.

Data/MC plots of BDT variables

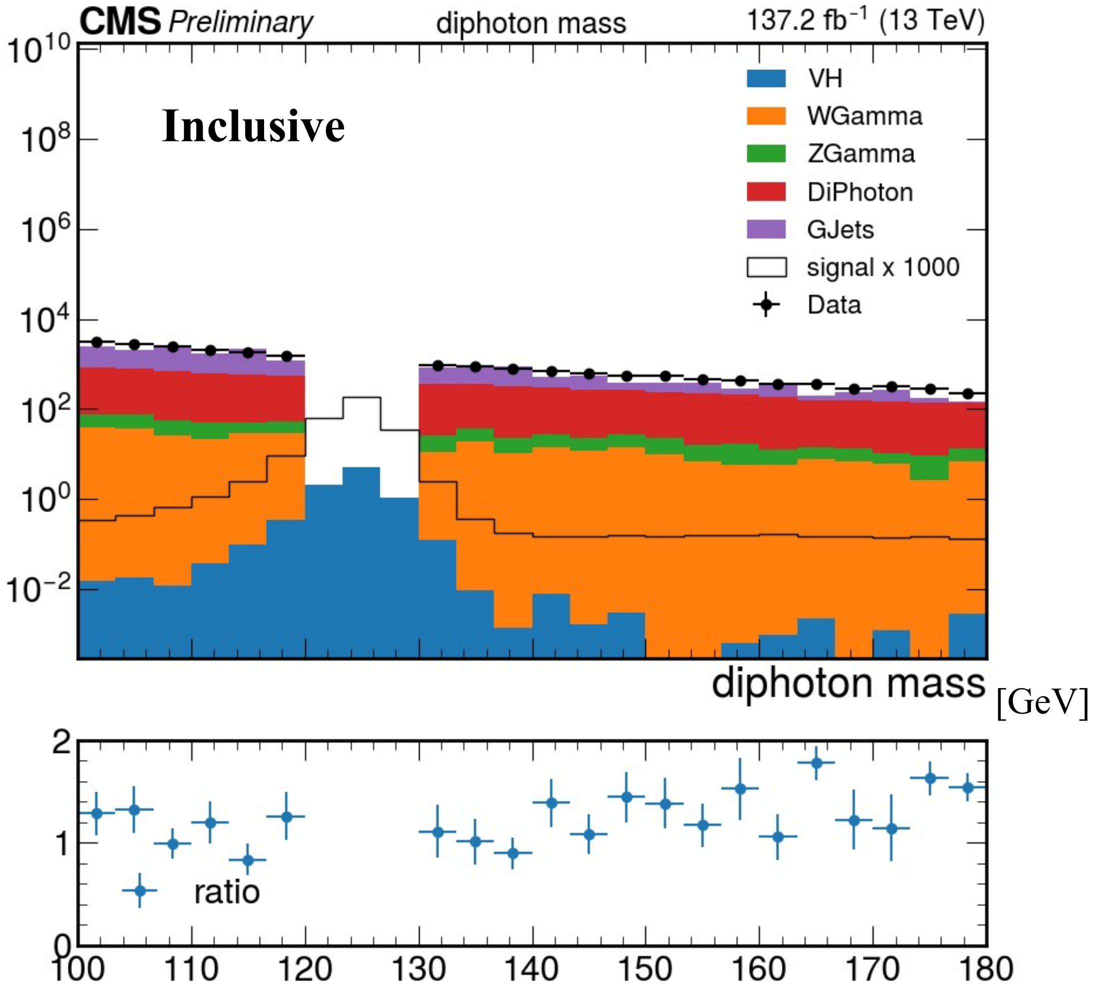


Plot of $\frac{p_T}{m_{\gamma\gamma}}$ for the leading γ (highest pT) from diphoton selection.

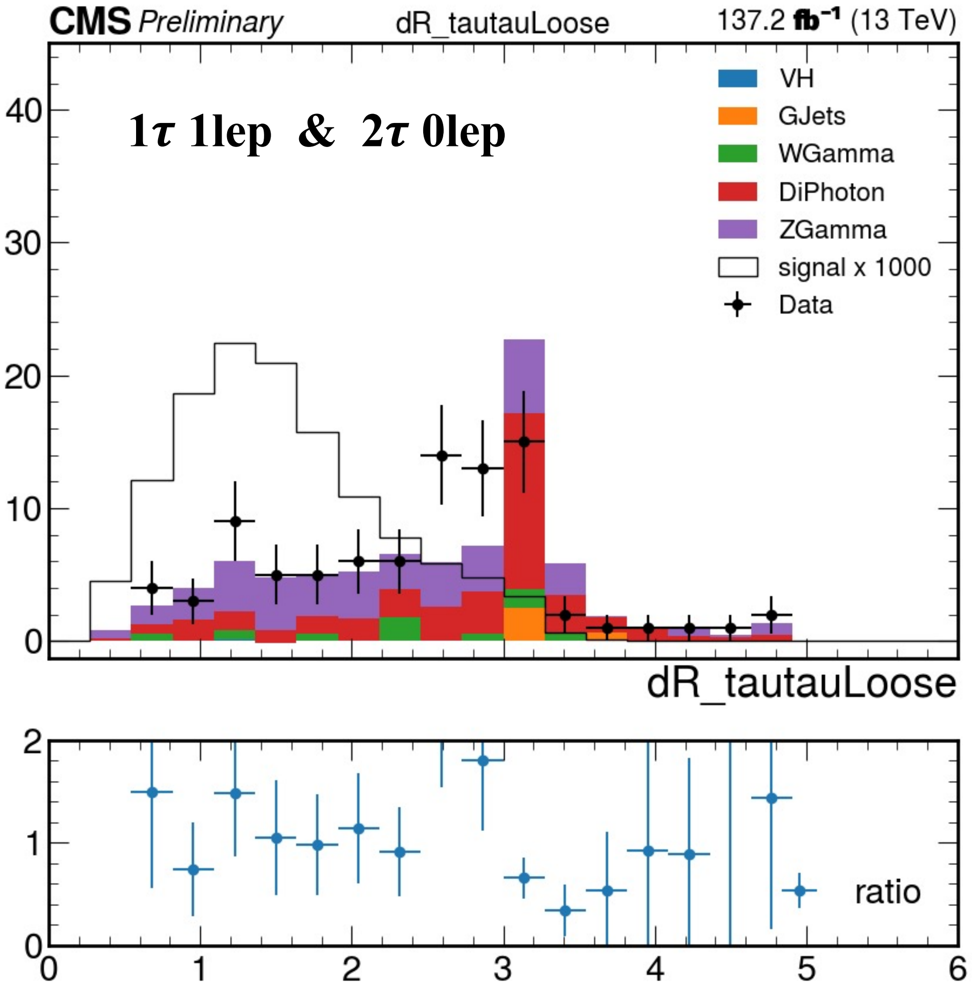


Plot of τ pT for the hadronic tau with largest transverse momentum.

Data/MC plots of BDT variables



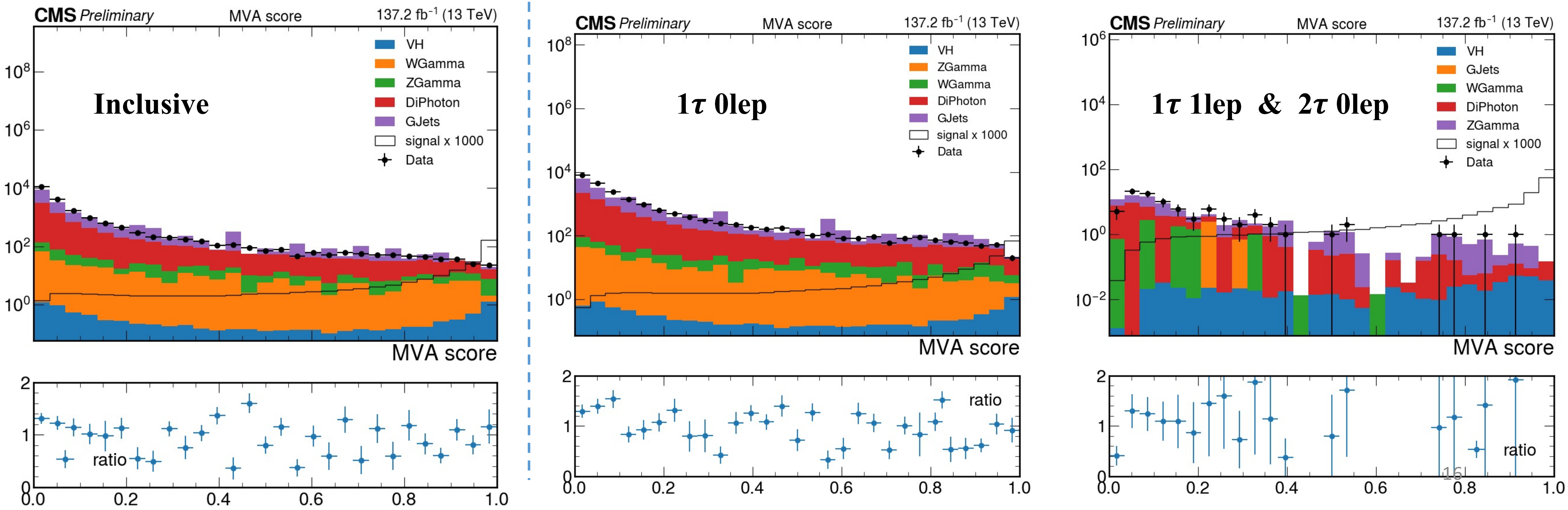
$m_{\gamma\gamma}$ sideband veto: 100< $m_{\gamma\gamma}$<120 &
130 < $m_{\gamma\gamma}$<180 GeV (no sig or res bkg)



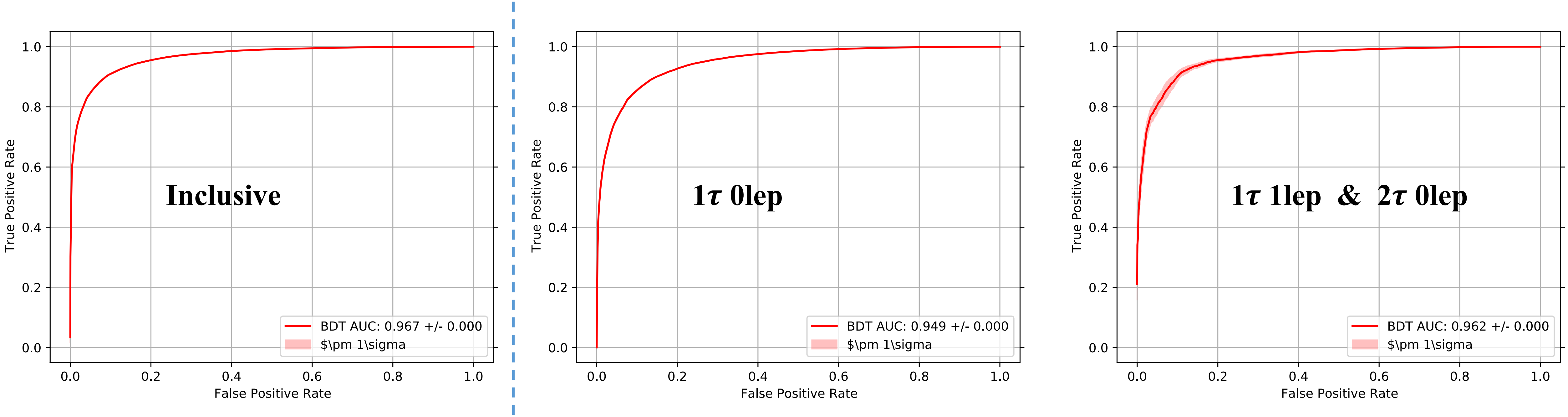
ΔR between $\tau/\mu/e$ candidates from Higgs.

BDT Performance

- BDT scores for the two scenarios explored
  - **Inclusive** BDT trained on 1τ 0lep, 1τ 1lep, 2τ 0lep events
  - One BDT dedicated to **1τ 0lep** events and one dedicated to **1τ 1lep & 2τ 0lep** events



BDT Performance



- Both BDT approaches show similar discrimination power

Signal Region Optimisation

- Adapted SR optimisation tool from ttH analysis:
  - define two bins for each BDT output
  - fit to $m_{\gamma\gamma}$ distribution
  - model signal with Double Crystal-Ball function
  - model background with exponential function
  - figure of merit: Upper Limit on HH cross section
  - during optimisation all signal & background are modelled from MC

Expected Sensitivity

- Based on optimised SR from last slide, we performed 1D fit to $m_{\gamma\gamma}$ to extract the upper limits on HH cross section

- We evaluated two methods to extract limits:
  - Use MC in signal & background fits
  - Use **data** for **non-res. background** and MC for resonant background and signal

Standard Model Upper Limits

non-res. bkg from →	MC	data
Inclusive BDT	31 x SM	29 x SM
Two BDTs	35 x SM	28 x SM

- Preliminary SM limits within a factor of 6 from HH → $b\bar{b}\gamma\gamma$ (5.2 x SM)

Summary

- Plan to deliver the results of the HH $\rightarrow \gamma\gamma\tau\tau$ analysis using an intermediate version of the nanoAOD framework proposed last week at the H$\gamma\gamma$ meeting

- Our current nanoAOD tools take ~20/30 mins to run full analysis from skimmed nanoAOD (diphoton selection)
  - SFs and systematics yet to be applied

- **Upper limit of ~30 x SM** for Higgs pair production

Future Steps

- Synchronise diphoton preselection with respect to Flashgg
- Include processes HH → $\gamma\gamma WW$ & HH → $\gamma\gamma ZZ$ to increase signal yields
- Request official signal MC samples (in progress)
- Include VBF production, potentially also VHH, ttHH
- Dedicated BDT to separate VH (V → ττ) from HH
- Explore BDT/DNN regression to further improve ττ resolution
- Extract signal by fitting 2D distribution ($m_{\gamma\gamma}$ & $m_{\tau\tau}$), if $\tau\tau$ present

Yields for Run II + HH $\rightarrow \gamma\gamma WW$

- Yields after di-photon preselection + 1 hadronic τ **including HH** $\rightarrow \boldsymbol{\gamma\gamma WW}$

	Inclusive	1tau_0lep	1tau_1lep	2tau_0lep
Inclusive bkg	18675 ± 922	18564 ± 922	69.3 ± 7.3	41.2 ± 3.4
Data	21338 ± 146	21245 ± 145	50.0 ± 7.1	38.0 ± 6.2
HH $\rightarrow \gamma\gamma\tau\tau$	0.298	0.169	0.074	0.054
HH $\rightarrow \gamma\gamma WW$ (semi-leptonic)	0.046	0.041	0.005	0.0004
HH $\rightarrow \gamma\gamma WW$ (di-leptonic)	0.021	0.009	0.011	0.0005

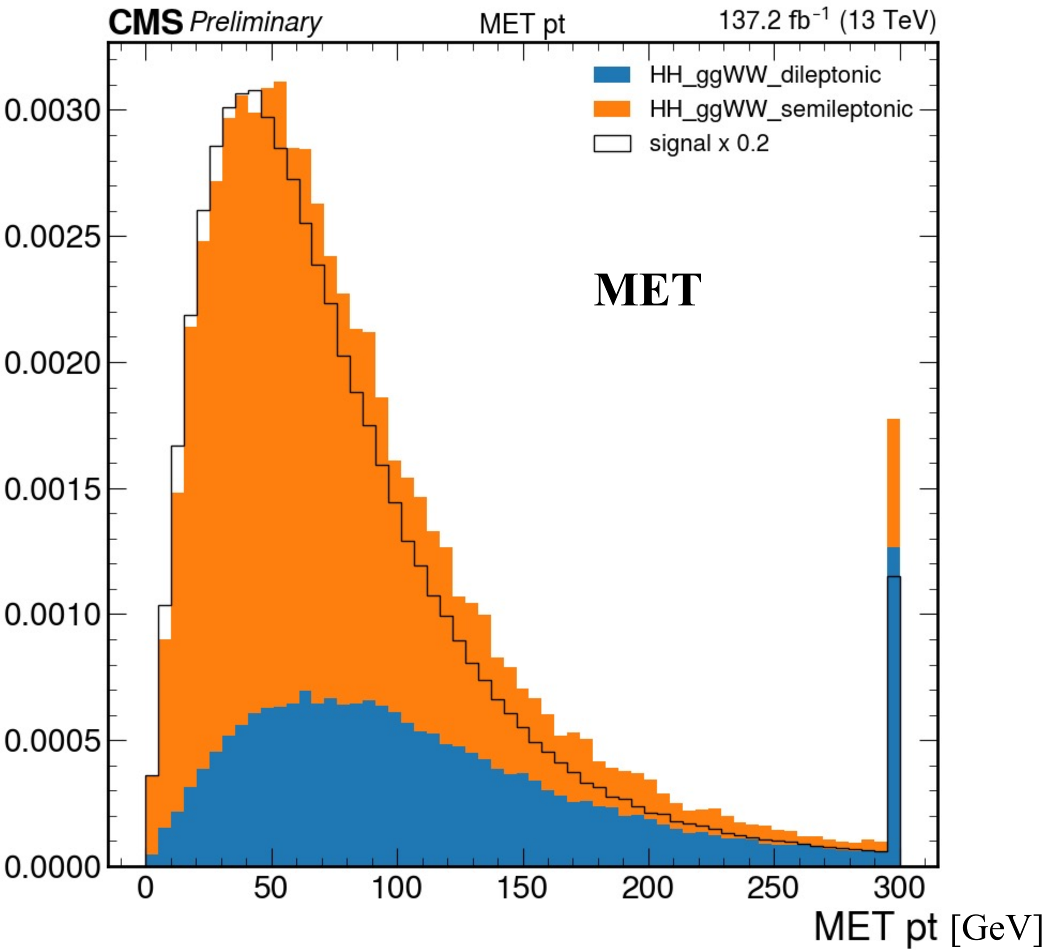
- Signal yields increase by $\sim$20%

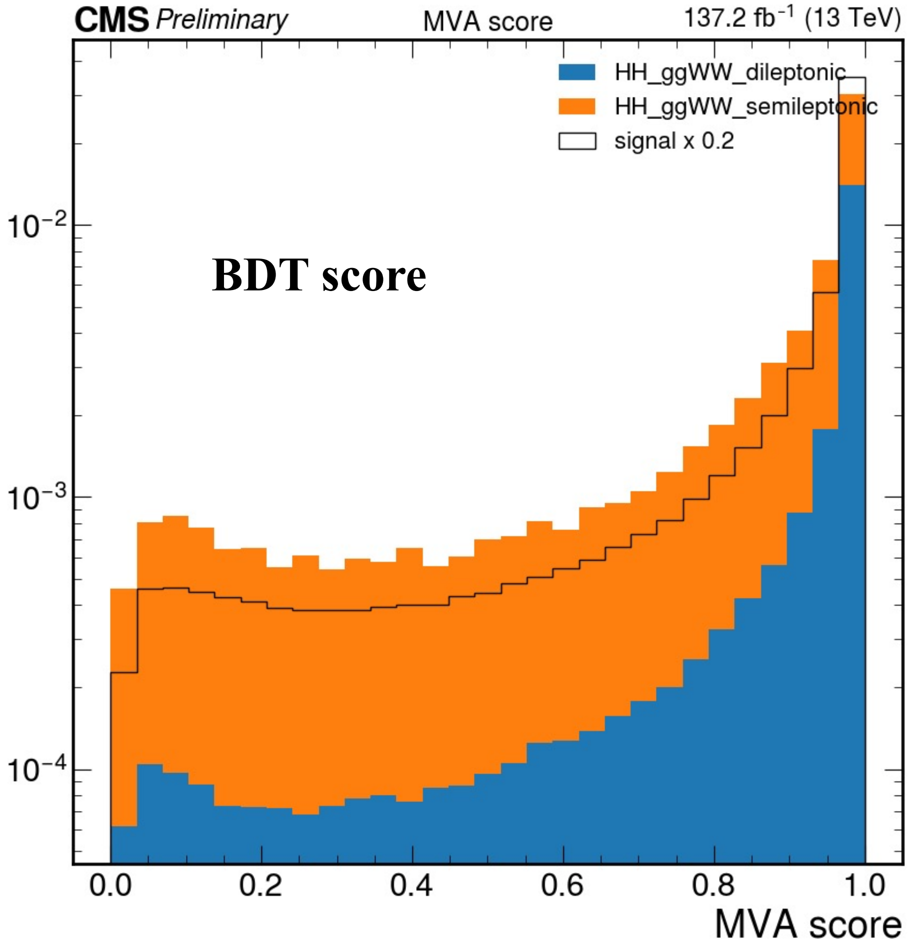
$\gamma\gamma\tau\tau/\gamma\gamma WW$ Comparison

- Compare profiles between HH $\rightarrow \gamma\gamma\tau\tau$ and HH $\rightarrow \gamma\gamma WW$ (semi-leptonic & di-leptonic)
  - HH $\rightarrow \gamma\gamma\tau\tau$ (signal) has been scaled by 0.2 to ease the comparison
  - full set of plots here

- Same inclusive BDT as slides 10-11, now trained on HH $\rightarrow \gamma\gamma\tau\tau$ and HH $\rightarrow \gamma\gamma WW$ (semi-leptonic & di-leptonic) as signal

$\gamma\gamma\tau\tau/\gamma\gamma WW$ Comparison





- No neat separation between HH $\rightarrow\gamma\gamma\tau\tau$ and HH $\rightarrow\gamma\gamma WW$

Expected Sensitivity w/ HH → $\gamma\gamma WW$

- Perform SR optimisation on new BDT including HH → $\gamma\gamma WW$ samples

- Extract limits fitting signal & background to MC

Standard Model Upper Limits

non-res. bkg from →	MC	data
Inclusive BDT	31 x SM	29 x SM
Inclusive BDT + HH → $\gamma\gamma WW$	27 x SM	-

- Adding HH → $\gamma\gamma WW$ lowers the UL from 31 to 27 x SM, within a factor of 5 from HH → $b\bar{b}\gamma\gamma$

Backup