Studies of boosted topologies in the search for NMSSM inspired di-Higgs events in the $\tau\tau + $bb final state with the CMS experiment

Martin Marz | September 12, 2022

Table of contents

The Standard Model (SM)

- Lepton and boson masses generated via spontaneous electroweak symmetry breaking
- Prediction of a single Higgs boson
- SM has shortcomings (Gravitation, DM etc.)
- One possible extension is supersymmertry

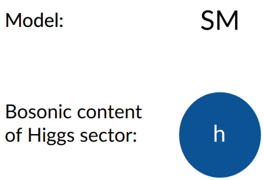


Figure: Courtesy of Felix Heyen

Minimal Supersymmetric Standard Model (MSSM)

- Minimal extension with SU(2) Higgs doublet
- Predicts super partners for all particles with different spin quantum number
- No super partners observed till now

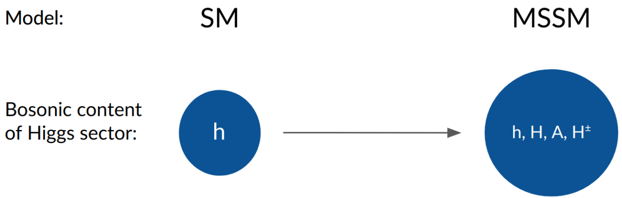


Figure: Courtesy of Felix Heyen

Introduction to the NMSSM

- MSSM symmetry **must** be broken
- Next-to-Minimal Supersymmetric Standard Model **(NMSSM)**: symmetry breaking occurs **naturally**
- Introduce additional singlet scalar field
- Results in **even more** Higgs bosons (7 in total)
- For my thesis h and h_S of importance

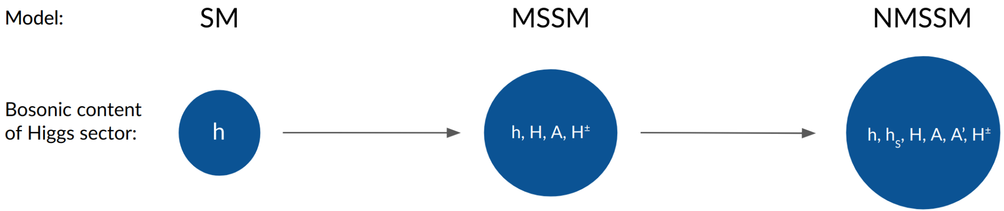


Figure: Courtesy of Felix Heyen

Overview

Karlsruhe Institute of Technology

- Search for di-Higgs events in the $H \rightarrow h(\tau\tau)h_S(bb)$ final state
- $h \,\hat{=}$ Higgs boson with standard model like properties $m_h = 125\,\text{GeV}$
- $h_S \,\hat{=}$ new singlet Higgs boson
- $H \,\hat{=}$ new heavy Higgs boson
- Analysis by Dr. Janek Bechtel at KIT (ETP-KA/2021-04)
- Published analysis: here

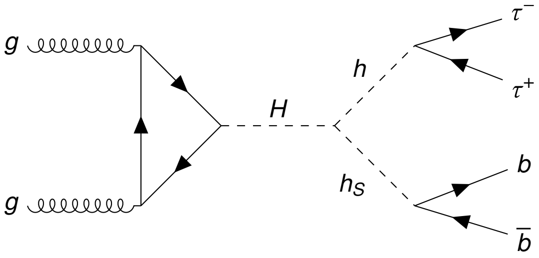


Figure: Feynman diagram for the $H \rightarrow h(\tau\tau)h_S(bb)$ final state.

Current upper limits

- $m_h = 125\,\text{GeV}$
- No theory prediction for m_H, m_{h_S}
  $\implies$ **free parameters**
  $\implies$ large possible mass range
- Only **on-shell** decays
  $\implies$ restriction on m_H, m_{h_S}
- Upper limits on
  $\sigma(gg \to H) B(H \to hh_S \to \tau\tau + bb)$
- Limits in graph scaled for visualization purposes

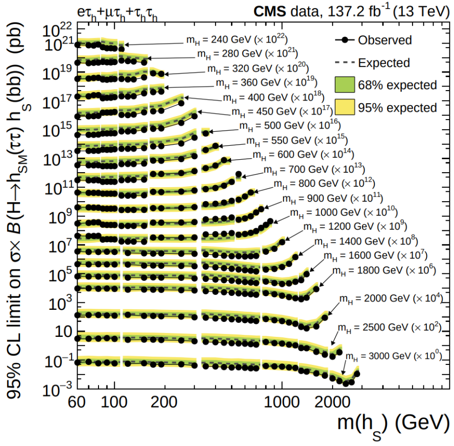


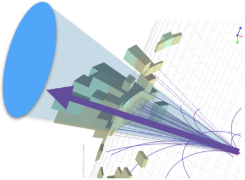
Figure: Taken from published analysis

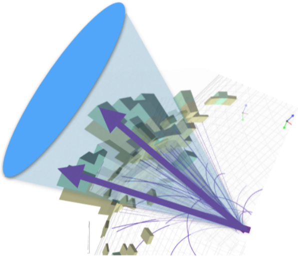
Boosted topologies

- **No dedicated** studies regarding **high p_T** objects which could be useful for large heavy Higgs masses

$$\Delta R(\text{daughters}) \approx \frac{2M(\text{mother})}{p_T(\text{mother})}$$

- $\Delta R \,\widehat{=}\,$ relative distance in $\eta - \phi$ space between two objects
- **Light** objects decay with **high p_T**
  $\implies$ **small ΔR** between decay products
- So for high m_H and/or small m_{h_S} small ΔR are expected
- Objects with **small ΔR** are called boosted





https://indico.cern.ch/event/732102/contributions/3092580/attachments/1759641/2854473/

Swapped final state

- Felix Heyen studied the inclusion of the swapped final state $H \longrightarrow h(bb)h_S(\tau\tau)$ (ETP-KA/2021-18)
- Only the $\tau_h\tau_h$ final state was used

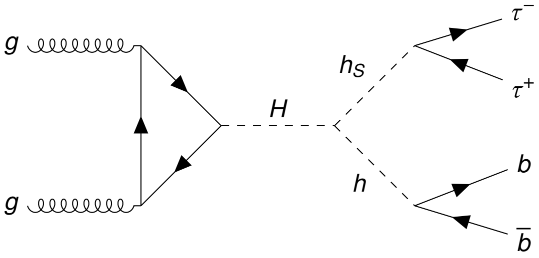


Figure: Feynman diagram for the $H \longrightarrow h(\tau\tau)h_S(bb)$ final state.

Results of Felix’s thesis

- An overall improvement is achieved
- Different exclusion sensitivity for the two final states
- Attributed to efficiency of boosted topologies
- Particularly relevant for boosted $h_S \longrightarrow \tau\tau$ decays with $m_{h_S} < 125\,\mathrm{GeV}$
- $h \longrightarrow \tau\tau$ are boosted for heavy m_{h_S}

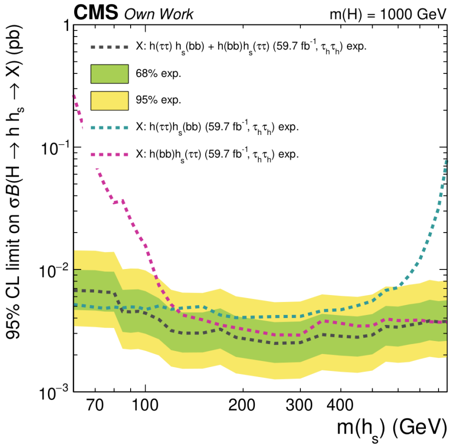


Figure: Taken from Felix Heyen’s master thesis

Goals of my thesis

1. Study the reconstruction efficiency of di-Higgs signal events
2. Scan the mass phase space of m_H and m_{h_S} for boosted topologies
3. Study the impact of a dedicated treatment of events with boosted b-jet topologies

- Multiple differences w.r.t the published analysis due to switch to UL Run-2
- Final results **not** comparable with the published analysis
  $\Longrightarrow$ only relative improvement measured

Reconstruction efficiency

- Signal samples for $H \to h(bb)h_S(\tau\tau)$ and $H \to h(\tau\tau)h_S(bb)$
- $m_H = [600, 700, 800, 900, 1000, 1200]$ GeV and $m_{h_S} = 60$ GeV
- For the reconstruction efficiency calculation a reference set of events (N_{gen}) is needed
- Of interest are boosted topologies
- Reference set limited on fiducial volume (as shown in table below)
- In addition apply matching between generator and reconstruction level objects to determine identification purity

Table: Selection criteria on generator-level particles

Particle	Transverse momentum (GeV)	Pseudorapidity	Mother particle
μ	$p_T \geq 20$	$\lvert\eta\rvert \leq 2.6$	h_S/h
τ_h	$p_T \geq 25$	$\lvert\eta\rvert \leq 2.8$	h_S/h
b-quark	$p_T \geq 15$	$\lvert\eta\rvert \leq 3.0$	h/h_S

Event selection criteria

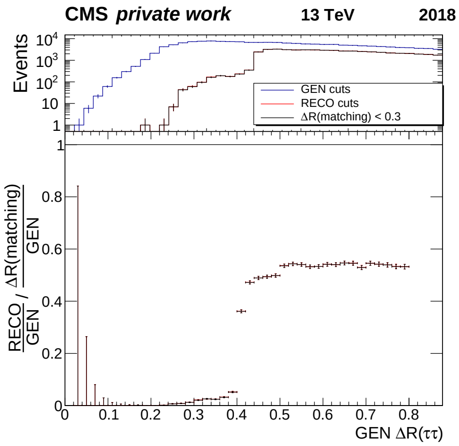
- Each event needs to pass multiple sets of selection criteria
- Most important RECO cuts: based on the published analysis
  1. presence of a $\mu\tau_\mathrm{h}$-pair
  2. $\mathrm{N_{jets}} \geq 2$ with $p_\mathrm{T} \geq 30\,\mathrm{GeV}$
  3. $\mathrm{N_{bjets}} \geq 1$ with $p_\mathrm{T} \geq 20\,\mathrm{GeV}$
- $\Delta R(\mu\tau_\mathrm{h}) \geq 0.5$ cut left out since boosted topologies are of interest
- b-jet pair constructed from jets to reconstruct the respective Higgs boson

GEN $\Delta R(\tau\tau)$ dependency

- Only lepton selection applied to isolate reconstruction efficiency from the b-quark pair
- Only statistical uncertainties shown in plot on the right
- Reconstruction efficiency reveals a steep drop for

$$\text{GEN } \Delta R(\tau\tau) \leq 0.5$$

- Identification purity $\approx 100\,\%$
- Current event selection **not** suited for boosted τ-topologies
- This explains the lower event yield and subsequently the different exclusion sensitivity for boosted $\tau\tau$ events found by Felix Heyen

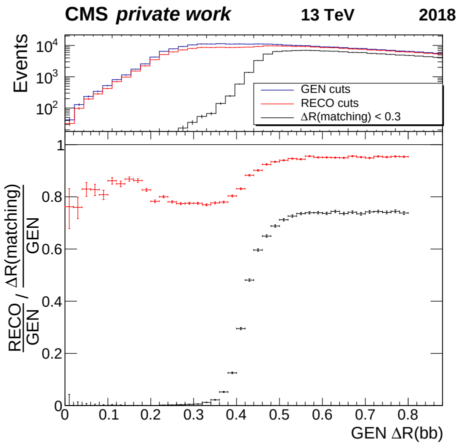


GEN $\Delta R(\mathrm{bb})$ dependency

- Only b-quark and jet selection applied to isolate reconstruction efficiency from the $\mu\tau_\mathrm{h}$-pair
- Only statistical uncertainties shown in plot on the right
- Reconstruction efficiency reveals a small drop for

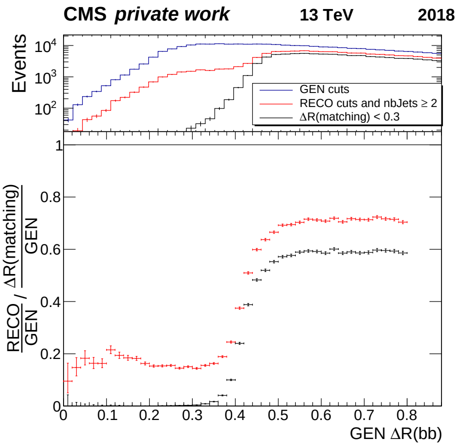
$$\mathrm{GEN}\ \Delta R(\mathrm{bb}) \leq 0.5$$

- Identification purity $\approx 0\,\%$ for this region
- Signal events with boosted b-jet pairs pass current event selection but are not reconstructed correctly



Interpretation

- $N_{bjets} \geq 1$ criteria allows most of the events with boosted b-jet topology
- A dedicated treatment of events with misidentified particles (boosted events) should improve the analysis in concerned kinematic regions



Parameter spaces with boosted topology

Karlsruhe Institute of Technology

Goals

1. Study the generator event selection of signal events (selection as shown before)
2. Study the regions with boosted topologies
3. Categorization of each m_{H}; $m_{\mathrm{h_S}}$ mass hypothesis

Categorization

- Each $m_{\mathrm{H}}; m_{\mathrm{h_S}}$ mass hypothesis falls into one of four categories
  1. A mass hypothesis is classified as **boosted** $\tau\tau$ if
  $$\text{median } \Delta R(\tau\tau) \leq 0.5$$
  2. Equally a hypothesis is classified as **boosted bb** if
  $$\text{median } \Delta R(\mathrm{bb}) \leq 0.5$$
  3. A hypothesis can be **boosted** $\tau\tau$ and **boosted bb** at the same time
  4. A hypothesis can be neither **boosted** $\tau\tau$ nor **boosted bb**
- The threshold 0.5 is chosen based on the previously observed efficiency drops

Results

Karlsruhe Institute of Technology

- Hypotheses marked in:
  - Red have an event acceptance below 20 %
  - Green are the hypotheses categorized as **boosted bb**
  - Blue are the hypotheses categorized as **boosted** $\tau\tau$
  - Almost all hypotheses contain events with a boosted τ-pair and/or b-pair
- Similar study for the swapped final state $\mathrm{H} \rightarrow \mathrm{h(bb)h_S}(\tau\tau)$ (Manuel Freudig)
- Regions for **boosted** $\tau\tau$ and **boosted bb** hypotheses are also swapped

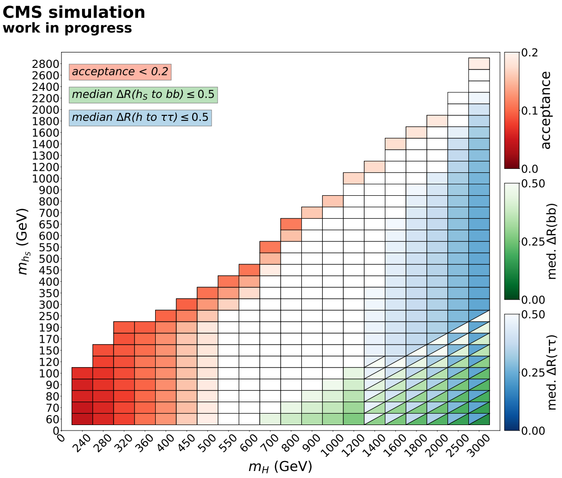


Figure: $\mathrm{H} \rightarrow \mathrm{h}(\tau\tau)\mathrm{h_S(bb)}$

Summary of generator study

Karlsruhe Institute of Technology

- Boosted $\mu\tau_h$ -pairs do **not** pass the event selection criteria
- $N_{bjets} \geq 1$ allows most of the boosted b-jet signatures to pass the event selection
- These events contain to a large fraction mis-identified b-jet pairs
- Boosted signatures are expected for a significant number of m_H-m_{h_S} hypotheses independent of the decay channel
- The kinematic acceptance is bad for small m_H and/or high m_{h_S}

Multiclassification

- Classify signal and background processes
- Event classification with neural network (NN)
- 16 input features
  1. p_T of the μ
  2. visible p_T of the τ_h
  3. ...
- 5 Output nodes
  1. $H \rightarrow h(\tau\tau)h_S(bb)$
  2. $H \rightarrow h(bb)h_S(\tau\tau)$
  3. Zll
  4. W+jets
  5. $t\bar{t}$

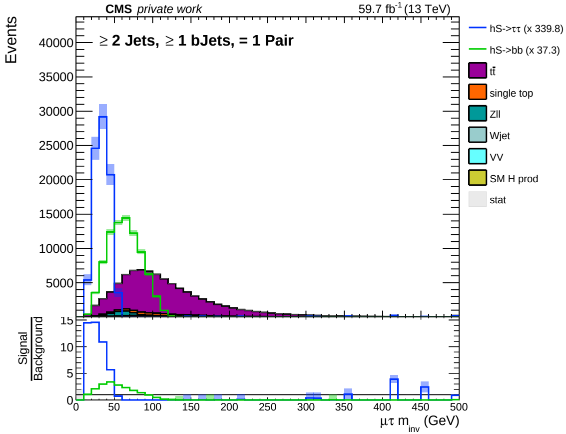


Figure: $m_H = 1200\,\text{GeV}$

Measures to improve multiclassification

- Architecture similar to the original analysis
- Baseline NN trained for relative comparison

Adding input features (Feat.)

- 13 new input features are introduced to the analysis
- Selected due to their high discriminating power between signal and background
- Mostly AK8 (i.e. $R = 0.8$) jet variables are chosen to make use of boosted b-jet topologies

Splitting $\mathbf{H \rightarrow h(\tau\tau)h_S(bb)}$ into two orthogonal groups (Split.)

1. $H \rightarrow h(\tau\tau)h_S(bb)$ signal events with a spacial distance GENΔR(bb) ≤ 0.5 (boosted)
2. $H \rightarrow h(\tau\tau)h_S(bb)$ signal events with a spacial distance GENΔR(bb) > 0.5 (resolved)

One additional selection criterion ($\mathbf{N_{AK8}}$)

- $N_{\text{AK8 jet}} \geq 1$
- Each aiming at optimizing the NN for boosted topologies without losing performance for resolved topologies

Comparison of confusion matrices and Taylor coefficients

Karlsruhe Institute of Technology

- The discrimination between the two final states and signal versus background (ROC score range of 0.7-0.9) for each NN
- All NN's with changes performed better on confusion matrix level compared to the baseline
- The most important features for the baseline NN are:
  1. m of the $\mu\tau_{\mathrm{h}}$-pair
  2. m of $\tau\tau + \mathrm{bb}$
  3. p_{T} of the p_{T}-leading b-jet
- The most important features for the NN with additional set of input features are:
  1. m of the $\mu\tau_{\mathrm{h}}$-pair
  2. **new:** ΔR between the $\mu\tau_{\mathrm{h}}$-pair and b-jet pair
  3. p_{T} of the p_{T}-leading b-jet
  4. **new:** p_{T} of the p_{T}-leading AK8-jet

More realistic comparison method

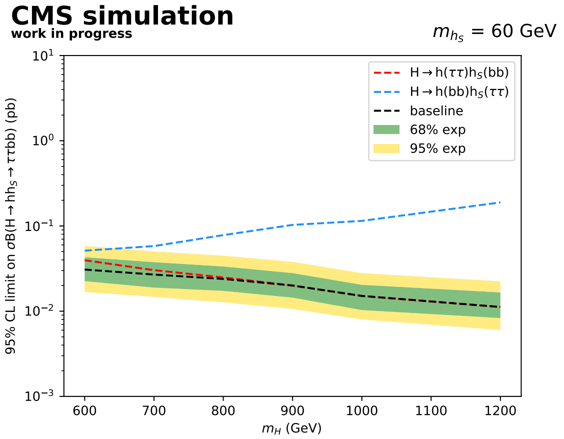
Karlsruhe Institute of Technology

- For each NN signal and background templates are produced
- Statistical model ($\mu \cdot s + b$) is only taking statistical uncertainties into account
- Expected asymptotic 95 % CL upper limit on $\sigma(gg \to \mathrm{H})B(\mathrm{H} \to \mathrm{hh_S} \to \tau\tau + \mathrm{bb})$ calculated for each approach using the $\mathrm{CL_S}$ method

Expected upper limit: baseline

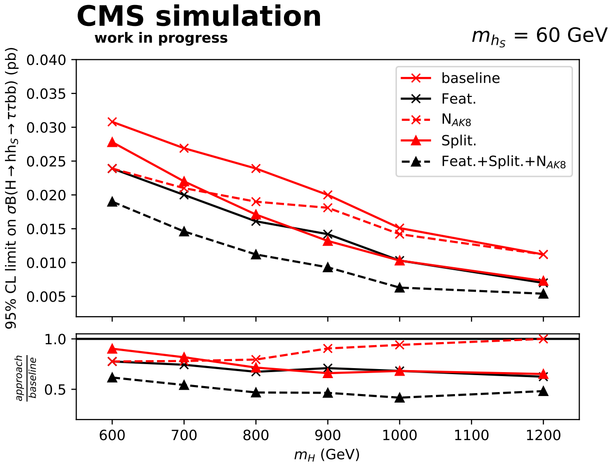
Karlsruhe Institute of Technology

- Limit is decreasing for higher m_H
- Effect of increasing event acceptance
- Stronger effect than event loss due to reconstruction efficiency
- Weaker limit for $H \rightarrow h(bb)h_S(\tau\tau)$ correlates with high fraction of boosted events



Interpretation of upper limits

- Isolated changes:
  1. **Feat.**: Stronger for high m_{H} correlates with stronger discrimination power of variables
  2. **Split.**: Bad discrimination between resolved and boosted signal topology for small m_{H}
  3. $\mathbf{N_{AK8}}$: Loss of events for high m_{H}
- Largest improvement when applying all three changes together
- Average improvement of 50 %



Conclusion

- On its own adding input features resulted in highest improvement between the three treatments
- Highest improvement when applying all three treatments at once
- Effect could be different for other m_H-$m_{\mathrm{h_S}}$ regions
- A dedicated treatment of boosted b-jet topologies is beneficial
- Only statistical uncertainties in this study

Summary and outlook

Karlsruhe Institute of Technology

Summary

- Boosted $\mu\tau_{\mathrm{h}}$-pair topologies do **not** pass the event selection criteria
- Boosted b-jet topologies pass the current event selection criteria
- High number of events with boosted topologies are expected
- $\Delta R(\mathrm{hh_S})$ is a characterizing feature
- Dedicated treatment of boosted topologies is beneficial for the analysis

Outlook

- Extend study to bigger m_{H}-$m_{\mathrm{h_S}}$ parameter phase space
- Improve AK8 jet selection and description
  1. spatial matching with b-jet
  2. make better use of substructure
- Include $\tau_{\mathrm{h}}\tau_{\mathrm{h}}$ and $\mathrm{e}\tau_{\mathrm{h}}$ final states
- Adept event selection criteria for boosted $\mu\tau_{\mathrm{h}}$-pair topologies

$\mu\tau_h$ criteria

Table: Selection criteria for μ in the $\mu\tau_h$ final state

Muon ID	Transverse momentum [GeV]	Pseudorapidity	Distance from PV [cm]	ParticleFlow isolation (ΔR)
medium	$p_T \geq 25$	$\lvert\eta\rvert \leq 2.1$	$d_z \leq 0.2$ $d_{xy} \leq 0.045$	≤ 0.15 (0.4)

Table: Selection criteria for τ_h in the $\mu\tau_h$ final state

Transverse momentum [GeV]	Pseudorapidity	Distance from PV [cm]	DeepTau Identification WP
$p_T \geq 30$	$\lvert\eta\rvert \leq 2.3$	$d_z \leq 0.2$ $d_{xy} \leq 0.045$	Tight vs μ VVLoose vs e Medium vs jets

Pair algorithm

Karlsruhe Institute of Technology

Table: $\mu\tau_\mathrm{h}$ pair algorithm

Importance	Criteria
4.	smallest μ `ParticleFlow` isolation
3.	highest μ p_T
2.	highest `DeepTau vs jets` discriminator
1.	highest $\tau_\mathrm{h} p_\mathrm{T}$

Jet criteria

Table: Selection criteria for different jet types

Jet type	Transverse momentum [GeV]	Pseudorapidity	`DeepFlavour` Identification WP	Seperation to leptons
b-jet	$p_T \geq 30$	$\|\eta\| \leq 2.5$	`medium`	$\Delta R \geq 0.4$
non b-jet	$p_T \geq 20$	$\|\eta\| \leq 2.5$	-	$\Delta R \geq 0.4$
AK8 jet	$p_T \geq 160$	$\|\eta\| \leq 2.5$	-	$\Delta R \geq 0.8$

Veto leptons

Table: Selection criteria for veto leptons

Particle	Lepton ID	Transverse momentum [GeV]	Pseudorapidity	Distance from PV [cm]	`ParticleFlow` isolation (ΔR)
e	`MVA noIso ID V2 WP90`	$p_T \geq 10$	$\lvert\eta\rvert \leq 2.1$	$d_z \leq 0.2$ $d_{xy} \leq 0.045$	≤ 0.3 (0.3)
μ	`medium`	$p_T \geq 10$	$\lvert\eta\rvert \leq 2.1$	$d_z \leq 0.2$ $d_{xy} \leq 0.045$	≤ 0.3 (0.4)

Acceptance

CMS simulation
work in progress

acceptance

m_{h_s} (GeV) \ m_H (GeV)	240	280	320	360	400	450	500	550	600	700	800	900	1000	1200	1400	1600	1800	2000	2500	3000
2800																				0.19
2600																				0.42
2400																				0.52
2200																			0.31	0.59
2000																			0.47	0.63
1800																		0.18	0.55	0.66
1600																	0.18	0.39	0.60	0.68
1400																0.17	0.39	0.50	0.63	0.70
1300																0.29	0.45	0.53	0.65	0.71
1200															0.17	0.38	0.48	0.56	0.65	0.71
1000														0.17	0.37	0.47	0.54	0.59	0.67	0.72
900														0.27	0.43	0.50	0.57	0.60	0.67	0.72
800													0.16	0.36	0.46	0.54	0.58	0.62	0.67	0.73
700												0.16	0.26	0.40	0.49	0.55	0.59	0.63	0.69	0.73
650											0.11	0.21	0.30	0.42	0.50	0.56	0.60	0.63	0.69	0.73
600											0.15	0.25	0.33	0.43	0.51	0.56	0.60	0.63	0.69	0.73
550										0.11	0.20	0.29	0.36	0.46	0.52	0.57	0.60	0.64	0.69	0.73
500										0.15	0.25	0.32	0.38	0.46	0.53	0.57	0.60	0.64	0.69	0.73
450									0.11	0.19	0.27	0.34	0.39	0.47	0.53	0.58	0.61	0.64	0.69	0.73
400								0.11	0.14	0.23	0.30	0.36	0.41	0.48	0.54	0.57	0.61	0.64	0.69	0.73
350							0.11	0.13	0.17	0.26	0.31	0.37	0.42	0.49	0.54	0.58	0.62	0.64	0.68	0.72
300						0.10	0.13	0.17	0.20	0.27	0.34	0.39	0.42	0.49	0.53	0.57	0.61	0.64	0.69	0.73
250					0.10	0.11	0.16	0.20	0.22	0.30	0.34	0.38	0.42	0.49	0.53	0.56	0.61	0.63	0.69	0.73
190			0.09	0.09	0.11	0.15	0.19	0.21	0.25	0.30	0.34	0.39	0.42	0.48	0.53	0.58	0.61	0.64	0.69	0.74
170			0.09	0.10	0.12	0.16	0.18	0.22	0.25	0.30	0.34	0.38	0.42	0.48	0.53	0.57	0.61	0.64	0.69	0.73
150		0.09	0.09	0.10	0.12	0.16	0.18	0.21	0.24	0.30	0.34	0.38	0.42	0.48	0.54	0.58	0.61	0.63	0.69	0.73
120		0.08	0.09	0.11	0.12	0.15	0.18	0.21	0.24	0.29	0.34	0.38	0.42	0.49	0.53	0.58	0.61	0.64	0.70	0.73
100	0.07	0.07	0.08	0.10	0.12	0.15	0.18	0.21	0.24	0.29	0.34	0.39	0.42	0.49	0.54	0.58	0.62	0.64	0.69	0.74
90	0.06	0.07	0.08	0.10	0.12	0.16	0.19	0.21	0.24	0.30	0.34	0.39	0.44	0.49	0.54	0.58	0.62	0.65	0.70	0.74
80	0.06	0.06	0.08	0.10	0.12	0.15	0.19	0.21	0.25	0.30	0.35	0.39	0.43	0.49	0.54	0.59	0.62	0.65	0.70	0.74
70	0.05	0.06	0.08	0.10	0.12	0.15	0.19	0.21	0.25	0.30	0.35	0.40	0.43	0.50	0.54	0.59	0.62	0.65	0.71	0.75
60	0.04	0.06	0.08	0.10	0.13	0.16	0.19	0.22	0.25	0.31	0.35	0.39	0.44	0.50	0.55	0.60	0.63	0.65	0.71	0.75

median ΔR

CMS simulation
work in progress

median $\Delta R(\tau\tau)$

m_{h_S} (GeV) \ m_H (GeV)	240	280	320	360	400	450	500	550	600	700	800	900	1000	1200	1400	1600	1800	2000	2500	3000
2800																				1.62
2600																				0.85
2400																				0.61
2200																			1.10	0.48
2000																			0.71	0.41
1800																		1.63	0.54	0.36
1600																	1.64	0.87	0.46	0.33
1400																1.64	0.87	0.63	0.39	0.30
1300																1.12	0.73	0.56	0.37	0.29
1200															1.64	0.88	0.64	0.51	0.35	0.28
1000														1.65	0.89	0.66	0.52	0.44	0.33	0.27
900														1.13	0.76	0.59	0.48	0.42	0.32	0.26
800													1.66	0.91	0.66	0.54	0.45	0.40	0.30	0.25
700												1.67	1.16	0.78	0.60	0.50	0.43	0.38	0.30	0.25
650											2.33	1.36	1.03	0.72	0.58	0.49	0.42	0.37	0.29	0.24
600											1.68	1.18	0.94	0.68	0.55	0.47	0.41	0.36	0.29	0.24
550										2.32	1.37	1.05	0.86	0.65	0.54	0.46	0.41	0.36	0.29	0.24
500										1.69	1.19	0.95	0.81	0.63	0.52	0.45	0.40	0.35	0.29	0.24
450									2.34	1.39	1.06	0.88	0.76	0.61	0.51	0.45	0.39	0.35	0.28	0.24
400								2.33	1.71	1.21	0.97	0.83	0.72	0.59	0.50	0.44	0.39	0.35	0.28	0.24
350							2.34	1.73	1.43	1.09	0.91	0.79	0.69	0.57	0.49	0.43	0.38	0.35	0.28	0.24
300						2.36	1.74	1.45	1.25	1.01	0.86	0.75	0.67	0.56	0.48	0.42	0.38	0.34	0.28	0.23
250					2.35	1.75	1.48	1.28	1.13	0.95	0.82	0.72	0.65	0.55	0.47	0.42	0.38	0.34	0.28	0.24
190			2.82	2.08	1.73	1.46	1.29	1.15	1.05	0.89	0.78	0.70	0.64	0.54	0.47	0.41	0.37	0.34	0.28	0.23
170			2.40	1.91	1.62	1.41	1.25	1.12	1.03	0.88	0.77	0.69	0.63	0.53	0.47	0.41	0.37	0.34	0.28	0.23
150		2.78	2.10	1.76	1.54	1.36	1.22	1.10	1.02	0.88	0.77	0.69	0.63	0.53	0.47	0.41	0.37	0.34	0.28	0.23
120		2.26	1.88	1.62	1.45	1.30	1.17	1.06	0.99	0.86	0.76	0.68	0.63	0.53	0.46	0.41	0.37	0.34	0.28	0.23
100	2.55	2.08	1.78	1.56	1.42	1.27	1.15	1.05	0.98	0.85	0.76	0.68	0.62	0.53	0.46	0.41	0.37	0.34	0.28	0.23
90	2.39	1.96	1.74	1.55	1.40	1.26	1.14	1.05	0.98	0.84	0.75	0.68	0.61	0.53	0.46	0.41	0.37	0.34	0.28	0.23
80	2.25	1.92	1.69	1.52	1.38	1.25	1.13	1.04	0.97	0.84	0.75	0.68	0.62	0.53	0.46	0.41	0.37	0.34	0.28	0.23
70	2.18	1.92	1.64	1.50	1.36	1.23	1.13	1.04	0.96	0.84	0.75	0.68	0.62	0.52	0.46	0.41	0.37	0.34	0.28	0.23
60	2.14	1.84	1.64	1.49	1.37	1.22	1.13	1.03	0.96	0.84	0.75	0.68	0.62	0.53	0.46	0.41	0.37	0.34	0.28	0.23

CMS simulation
work in progress

median $\Delta R(bb)$

m_{h_S} (GeV) \ m_H (GeV)	240	280	320	360	400	450	500	550	600	700	800	900	1000	1200	1400	1600	1800	2000	2500	3000
2800																				3.15
2600																				3.10
2400																				3.03
2200																			3.12	2.96
2000																			3.05	2.85
1800																		3.12	2.94	2.71
1600																	3.11	3.04	2.79	2.52
1400																3.10	3.01	2.91	2.61	2.31
1300																3.05	2.93	2.82	2.48	2.19
1200															3.09	2.97	2.84	2.70	2.35	2.07
1000														3.07	2.92	2.76	2.58	2.42	2.06	1.79
900														2.97	2.78	2.60	2.42	2.25	1.89	1.62
800													3.03	2.84	2.63	2.40	2.22	2.04	1.70	1.46
700												3.00	2.89	2.63	2.40	2.19	2.00	1.83	1.50	1.30
650											3.05	2.94	2.82	2.55	2.28	2.07	1.89	1.72	1.42	1.20
600											2.97	2.84	2.69	2.42	2.14	1.93	1.76	1.61	1.31	1.11
550										3.02	2.88	2.74	2.58	2.26	2.02	1.79	1.64	1.49	1.21	1.02
500										2.94	2.76	2.59	2.41	2.09	1.84	1.66	1.49	1.35	1.11	0.93
450									2.99	2.80	2.62	2.42	2.24	1.92	1.69	1.51	1.35	1.22	1.00	0.85
400								2.96	2.87	2.64	2.43	2.22	2.05	1.75	1.51	1.33	1.20	1.09	0.88	0.75
350							2.95	2.82	2.69	2.44	2.20	1.99	1.81	1.53	1.34	1.18	1.06	0.96	0.78	0.66
300						2.90	2.76	2.61	2.45	2.16	1.93	1.73	1.58	1.32	1.15	1.02	0.91	0.82	0.66	0.56
250					2.87	2.66	2.52	2.32	2.14	1.86	1.63	1.45	1.33	1.12	0.96	0.85	0.75	0.68	0.55	0.47
190			2.92	2.67	2.47	2.25	2.01	1.83	1.66	1.42	1.25	1.11	1.00	0.84	0.73	0.65	0.58	0.52	0.42	0.36
170			2.70	2.45	2.29	1.99	1.79	1.65	1.49	1.27	1.11	0.99	0.89	0.76	0.65	0.58	0.52	0.47	0.38	0.32
150		2.86	2.50	2.25	2.03	1.78	1.58	1.43	1.30	1.12	0.98	0.87	0.79	0.66	0.58	0.51	0.45	0.41	0.34	0.28
120		2.42	2.12	1.83	1.60	1.43	1.25	1.14	1.05	0.89	0.79	0.70	0.63	0.53	0.46	0.41	0.36	0.33	0.27	0.23
100	2.54	2.07	1.75	1.47	1.33	1.16	1.04	0.95	0.86	0.74	0.65	0.58	0.53	0.44	0.39	0.34	0.30	0.28	0.22	0.19
90	2.30	1.81	1.54	1.33	1.20	1.04	0.92	0.84	0.77	0.66	0.58	0.52	0.47	0.40	0.34	0.30	0.28	0.25	0.20	0.17
80	2.01	1.61	1.34	1.20	1.05	0.92	0.83	0.75	0.70	0.59	0.52	0.46	0.42	0.35	0.31	0.27	0.24	0.22	0.18	0.16
70	1.66	1.39	1.17	1.02	0.93	0.81	0.73	0.66	0.60	0.52	0.45	0.40	0.37	0.31	0.27	0.24	0.21	0.19	0.16	0.13
60	1.42	1.14	1.00	0.87	0.78	0.68	0.62	0.56	0.51	0.44	0.39	0.35	0.31	0.27	0.23	0.21	0.18	0.17	0.14	0.11

Baseline confusion

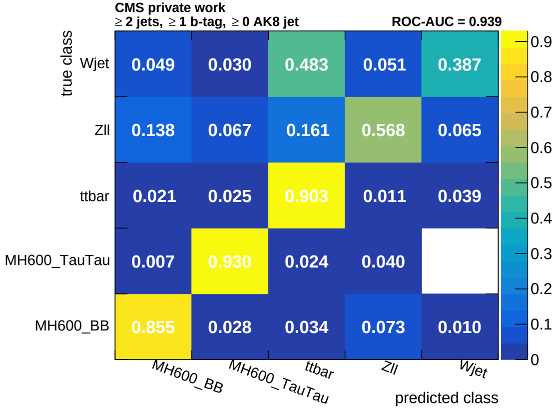


Figure: $m_H = 600\,\text{GeV}$

Baseline Taylor

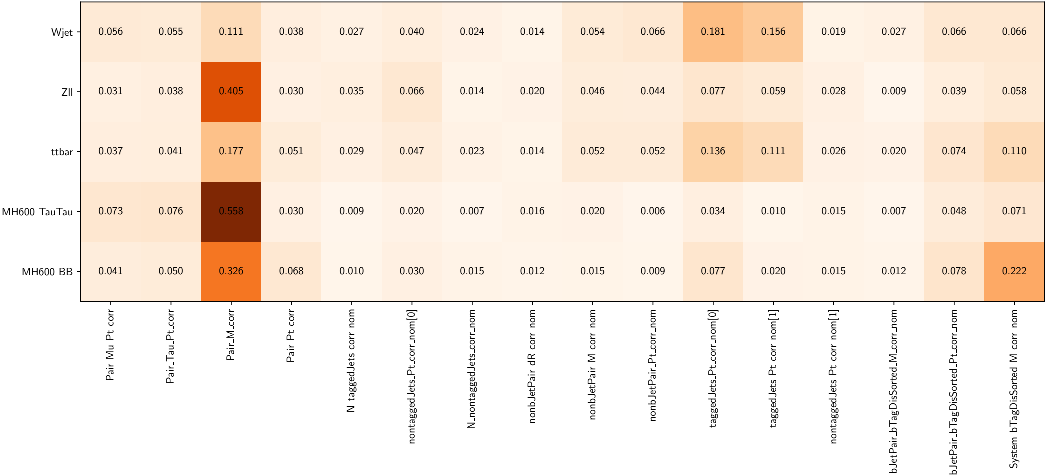


Figure: $m_H = 600\,\text{GeV}$

Add var confusion

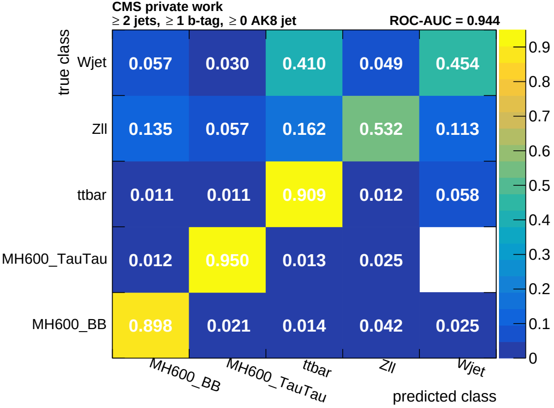


Figure: $m_H = 600\,\text{GeV}$

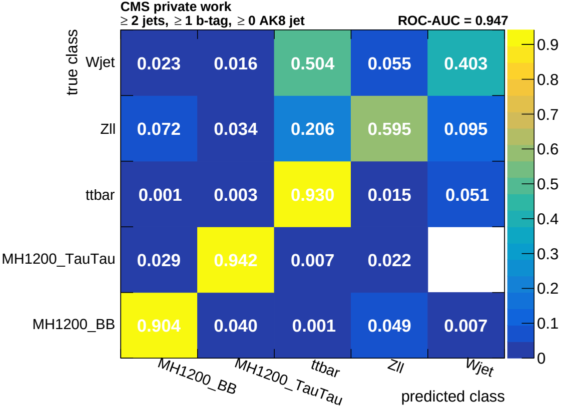


Figure: $m_H = 1200\,\text{GeV}$

Add var Taylor

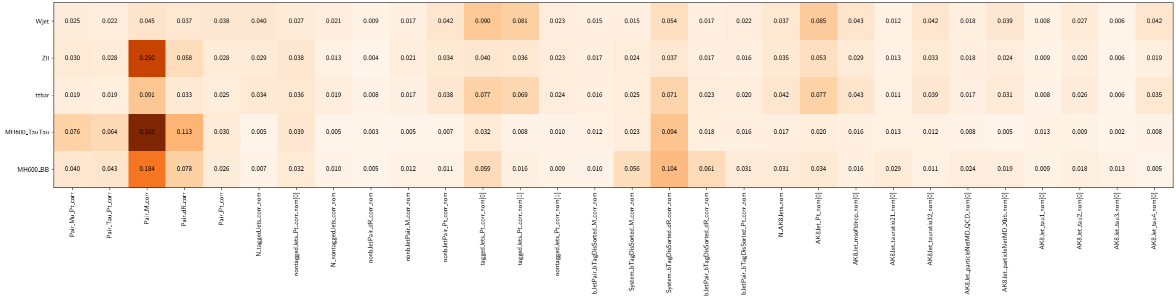


Figure: $m_H = 600\,\text{GeV}$

Split confusion

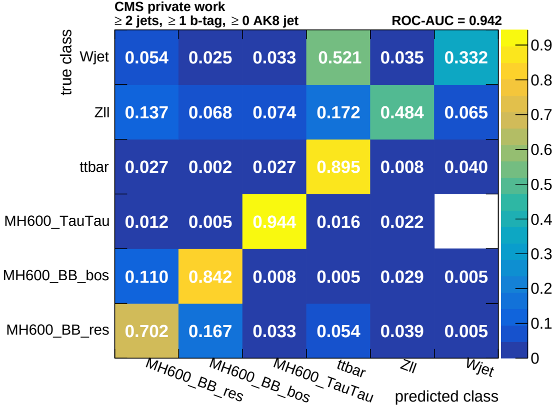


Figure: m_H = 600 GeV

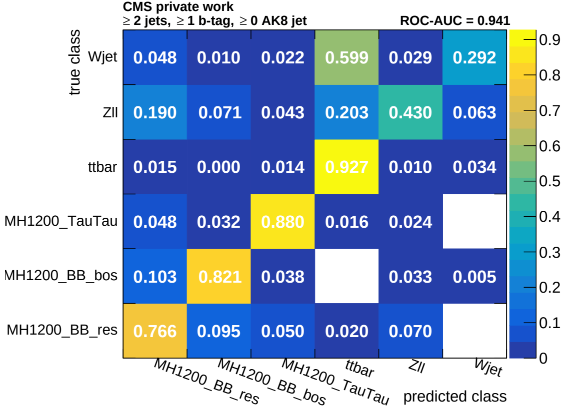


Figure: m_H = 1200 GeV

Cut confusion

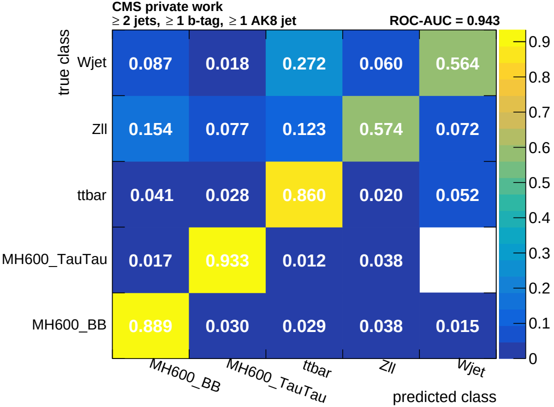


Figure: $m_H = 600\,GeV$

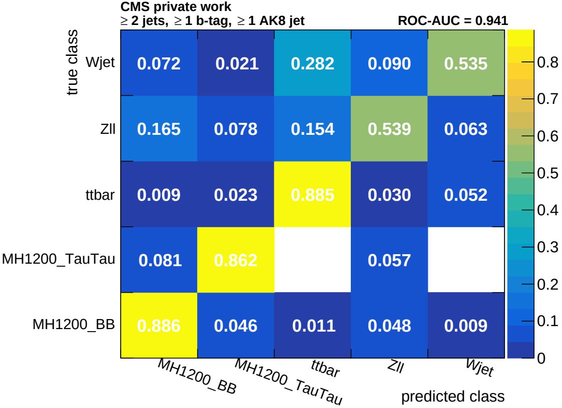


Figure: $m_H = 1200\,GeV$

ΔR ($\mu\tau_{\mathbf{h}}$;bb)

Karlsruhe Institute of Technology

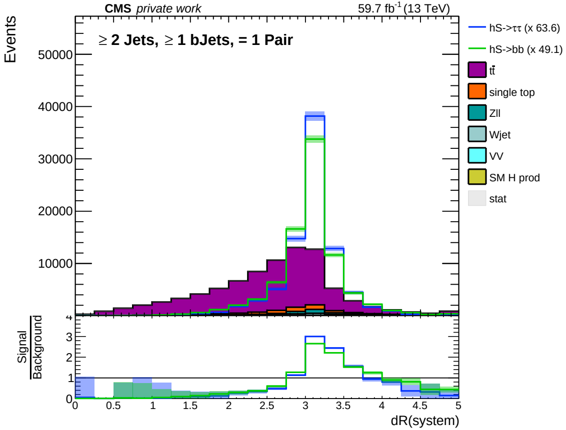


Figure: $m_H = 600\,\text{GeV}$

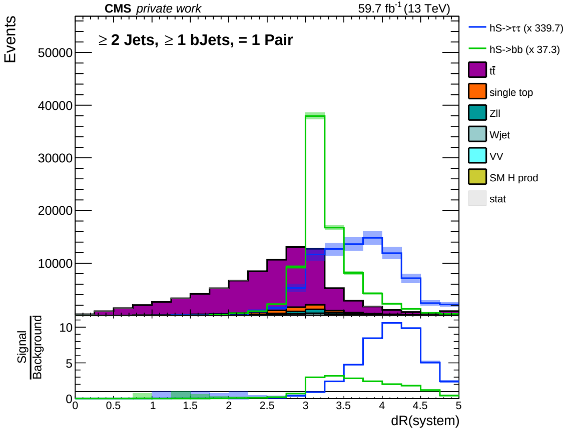


Figure: $m_H = 1200\,\text{GeV}$

Softdrop mass

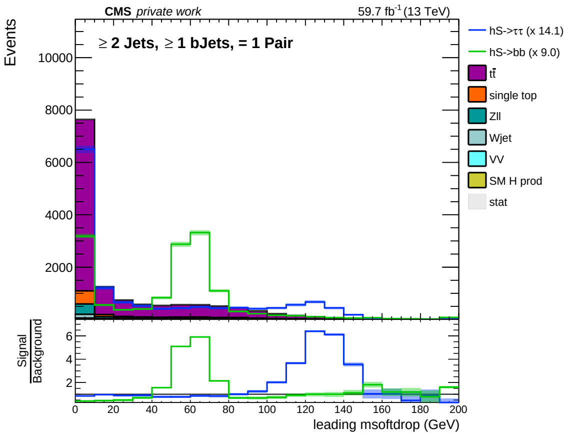


Figure: $m_H = 600\,\text{GeV}$

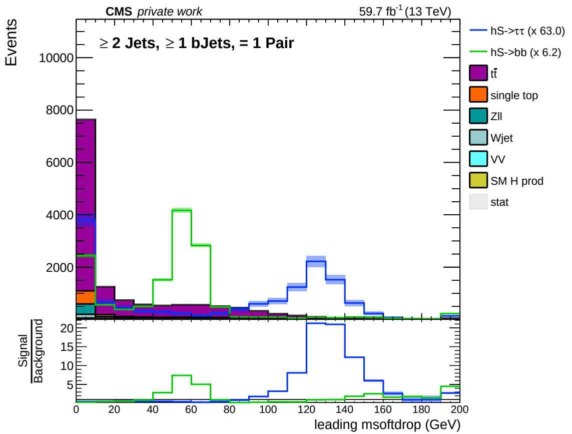


Figure: $m_H = 1200\,\text{GeV}$

QCD disc

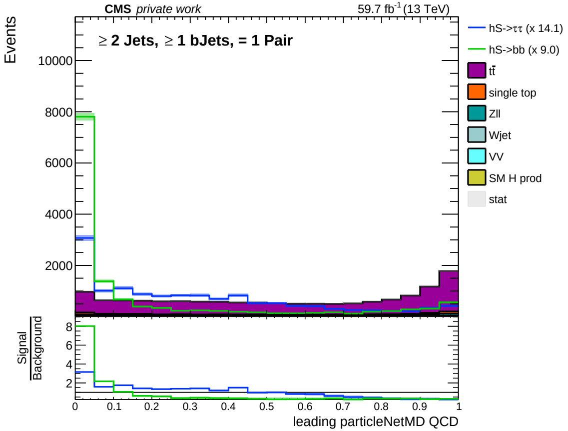


Figure: $m_H = 600\,\mathrm{GeV}$

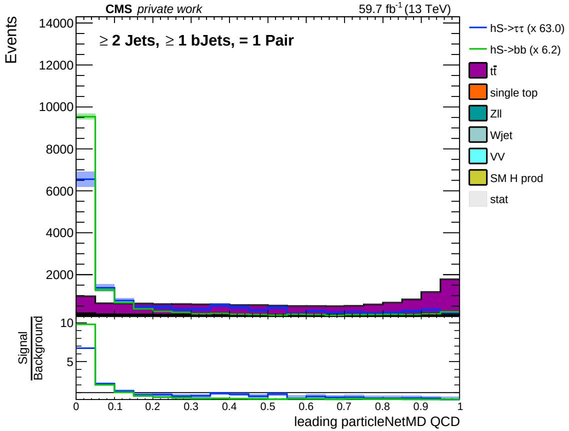


Figure: $m_H = 1200\,\mathrm{GeV}$

X(bb) disc

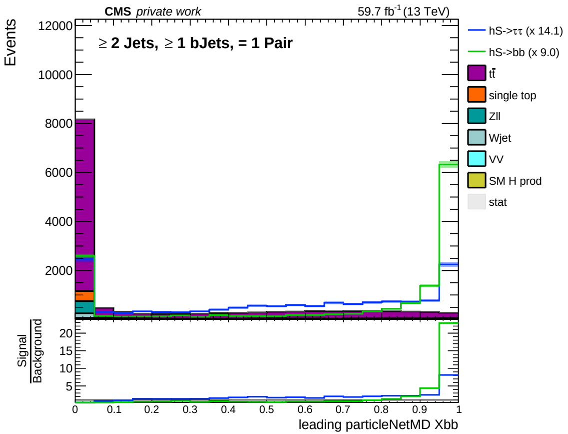


Figure: $m_H = 600\,\mathrm{GeV}$

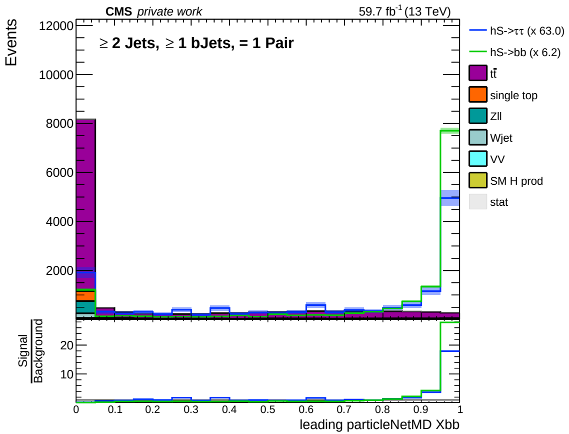


Figure: $m_H = 1200\,\mathrm{GeV}$

AK8 p_T

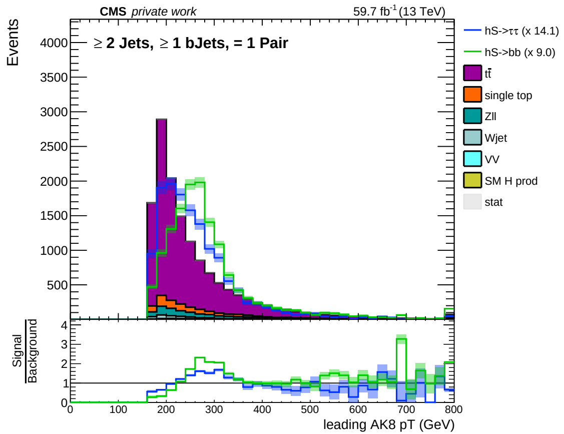


Figure: $m_H = 600$ GeV

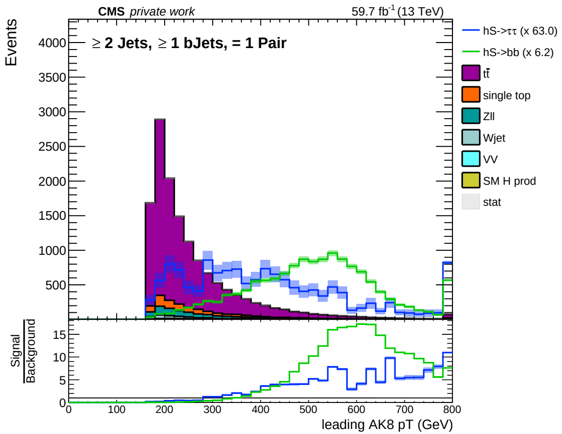


Figure: $m_H = 1200$ GeV

τ_1

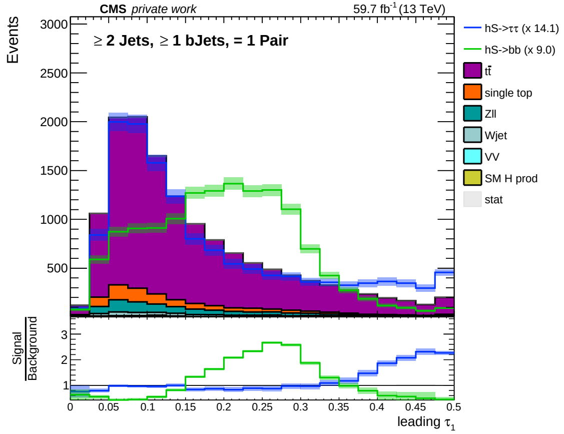


Figure: $m_H = 600\,GeV$

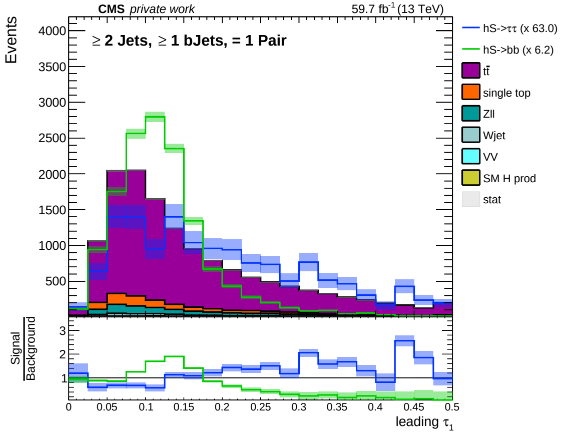


Figure: $m_H = 1200\,GeV$

τ_2

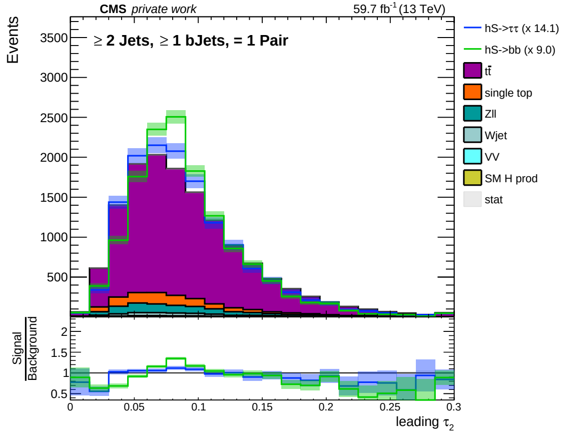


Figure: $m_H = 600\,\mathrm{GeV}$

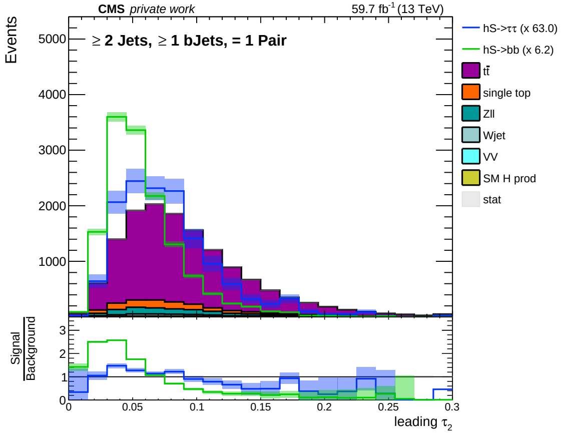


Figure: $m_H = 1200\,\mathrm{GeV}$

τ_2/τ_1

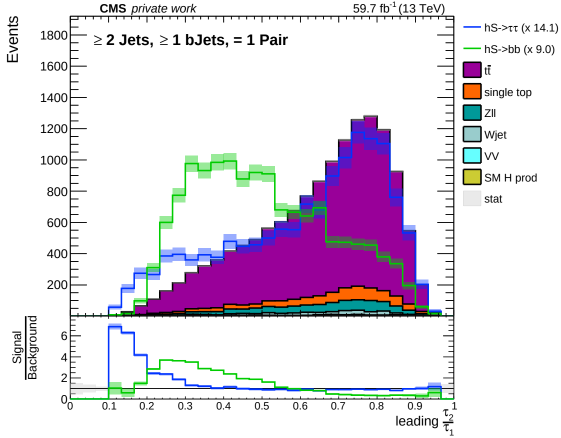


Figure: $m_H = 600\,\mathrm{GeV}$

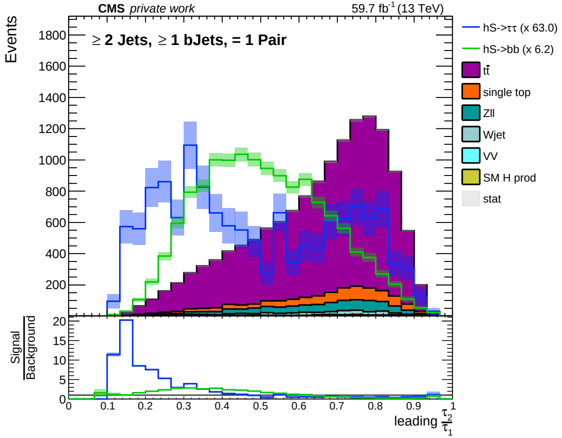


Figure: $m_H = 1200\,\mathrm{GeV}$

Signal nodes baseline

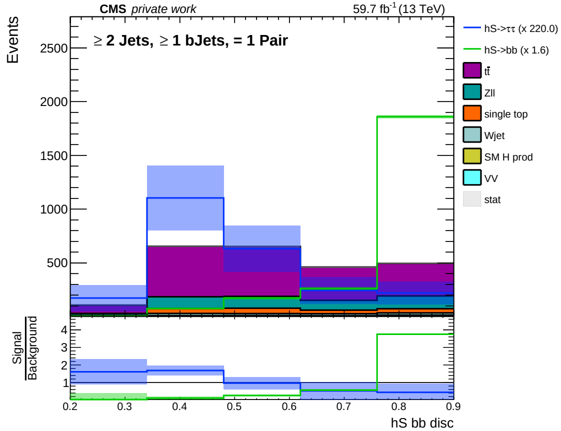


Figure: $m_H = 600\,\text{GeV}$

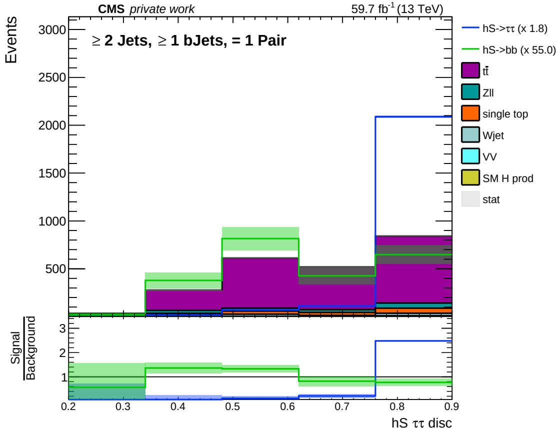


Figure: $m_H = 600\,\text{GeV}$

Background nodes baseline

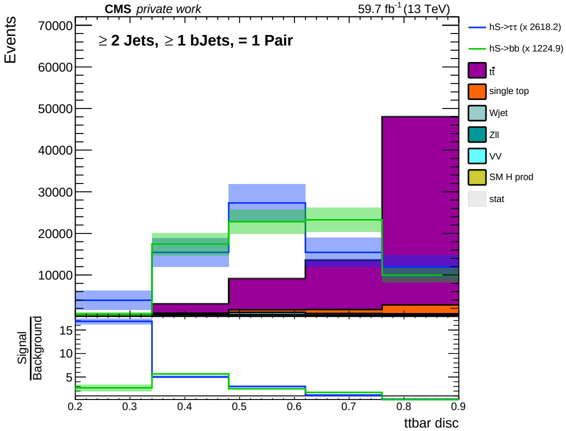


Figure: m_H = 600 GeV

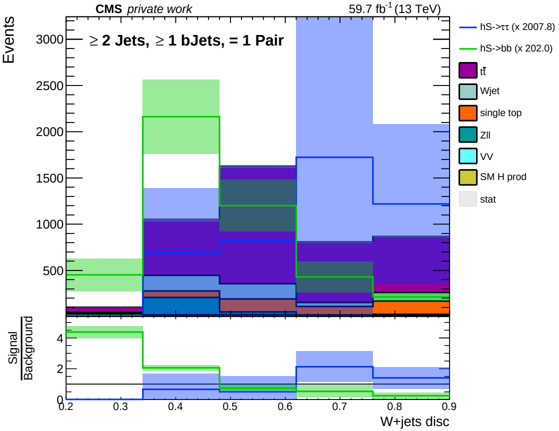


Figure: m_H = 600 GeV

Background nodes baseline

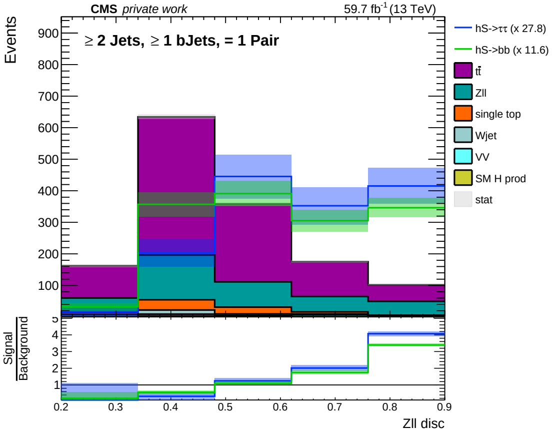


Figure: $m_H = 600\,GeV$

Signal nodes baseline

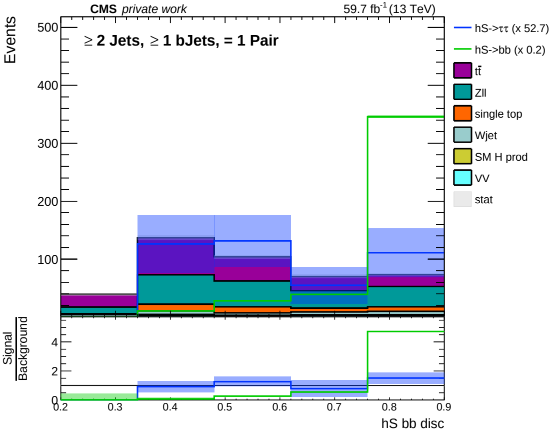


Figure: $m_H = 1200\,\text{GeV}$

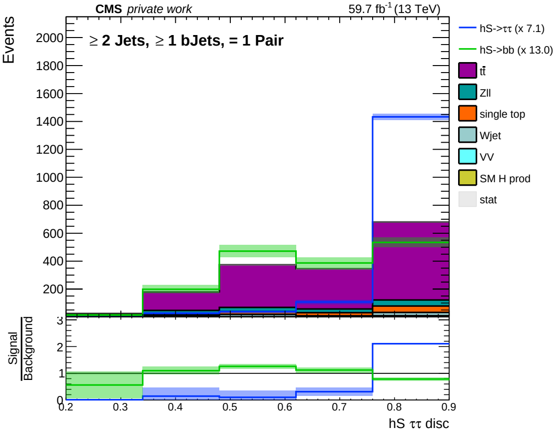


Figure: $m_H = 1200\,\text{GeV}$

Background nodes baseline

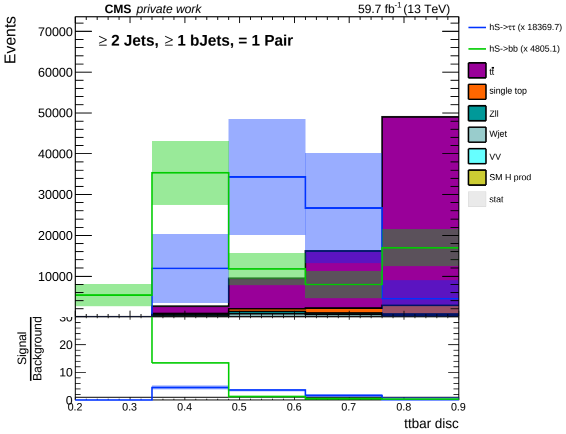


Figure: $m_H = 1200\,\text{GeV}$

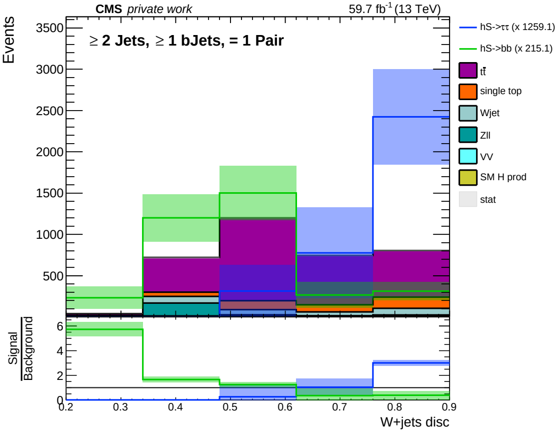


Figure: $m_H = 1200\,\text{GeV}$

Background nodes baseline

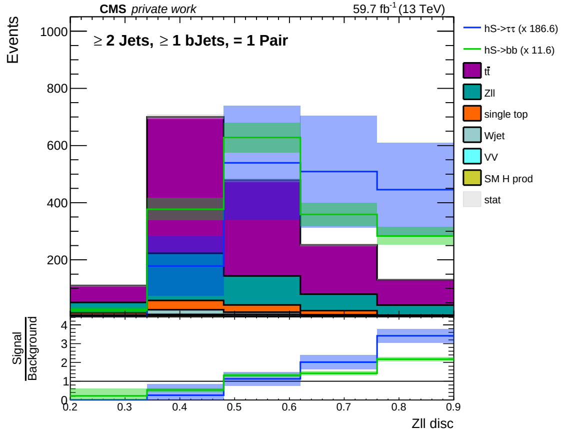


Figure: $m_H = 1200\,\mathrm{GeV}$

Signal nodes add. vars

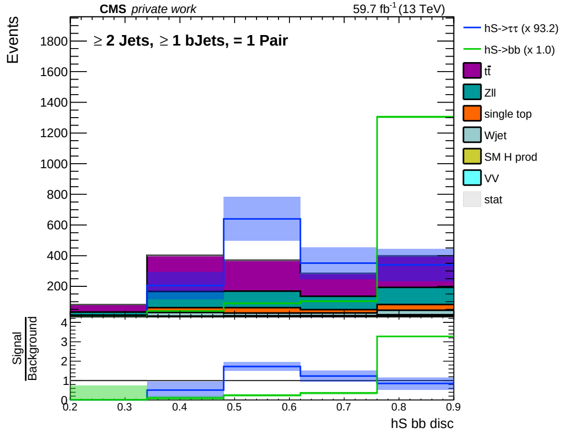


Figure: $m_H = 600\,\text{GeV}$

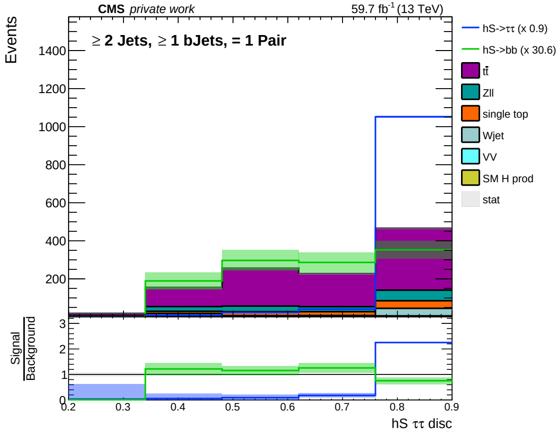


Figure: $m_H = 600\,\text{GeV}$

Background nodes add. vars

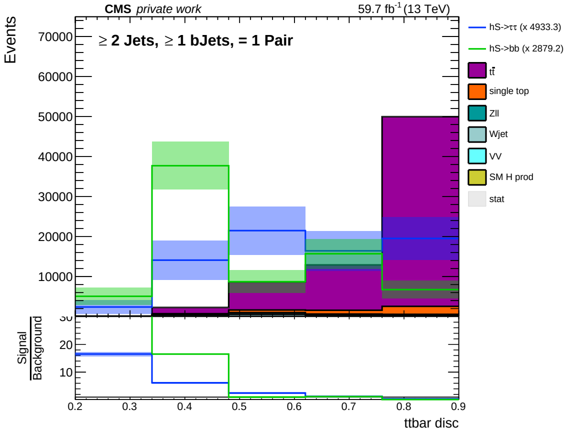


Figure: m_H = 600 GeV

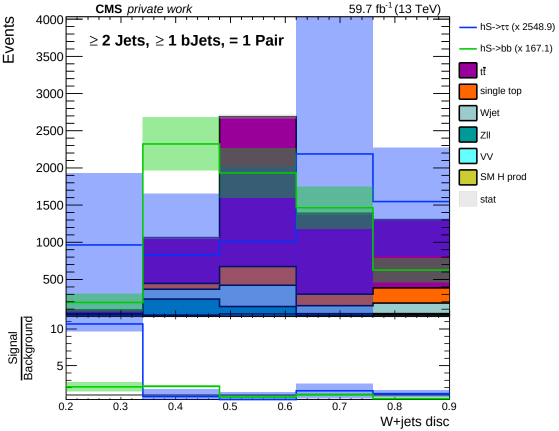


Figure: m_H = 600 GeV

Background nodes add. vars

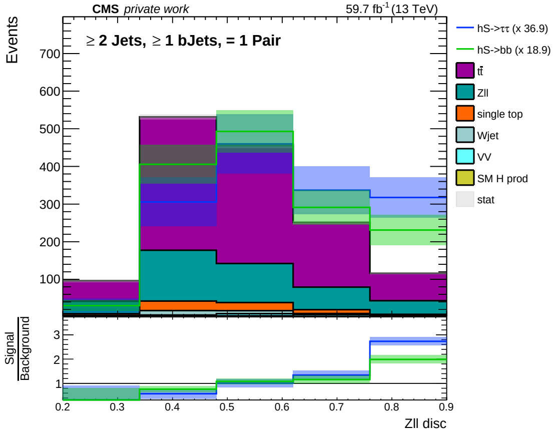


Figure: $m_H = 600\,\text{GeV}$

Signal nodes add. vars

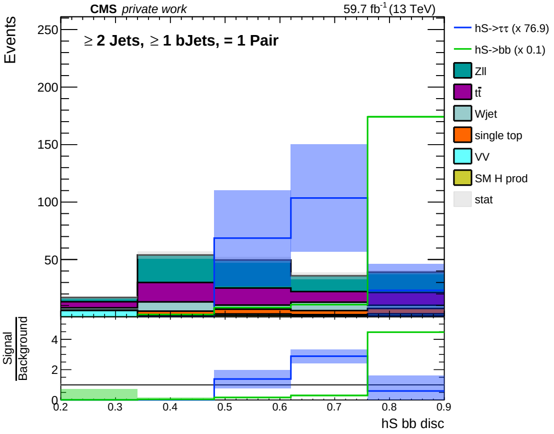


Figure: $m_H = 1200\,\text{GeV}$

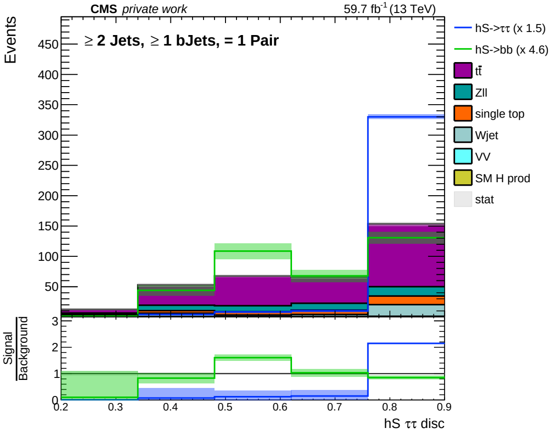


Figure: $m_H = 1200\,\text{GeV}$

Background nodes add. vars

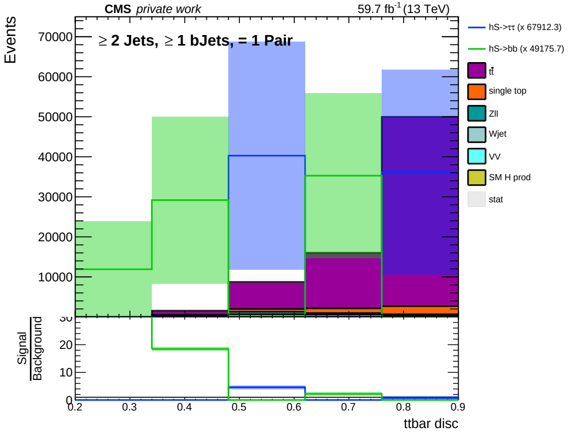


Figure: $m_H = 1200\,\text{GeV}$

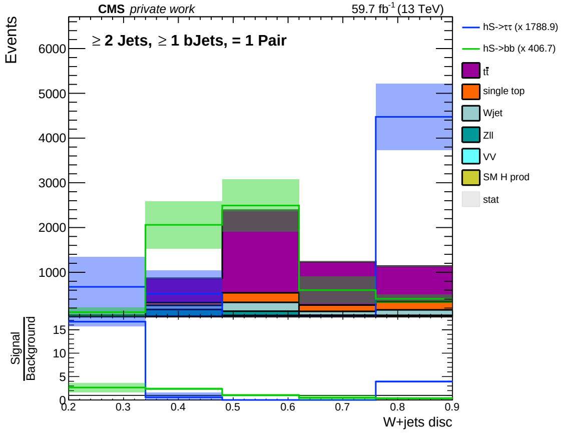


Figure: $m_H = 1200\,\text{GeV}$

Background nodes add. vars

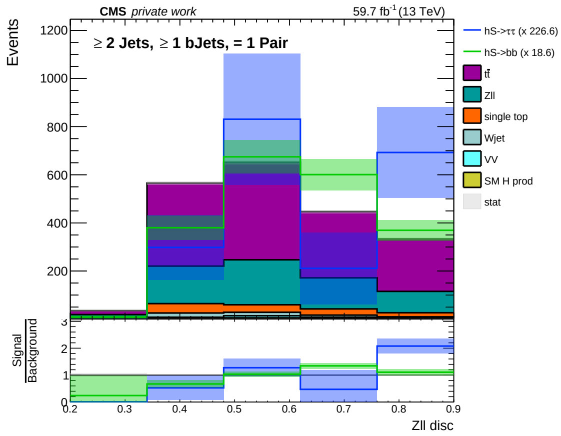


Figure: $m_H = 1200\,\text{GeV}$

Signal nodes split

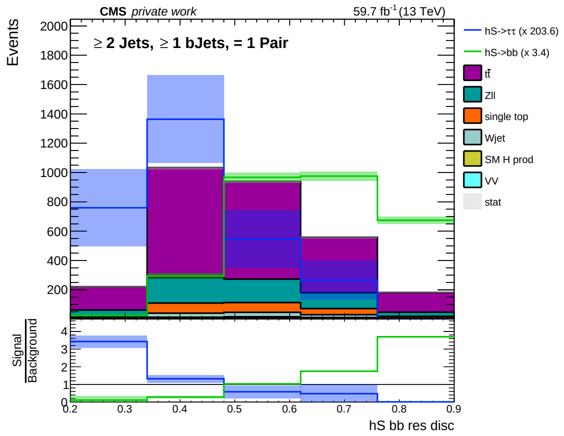


Figure: m_H = 600 GeV

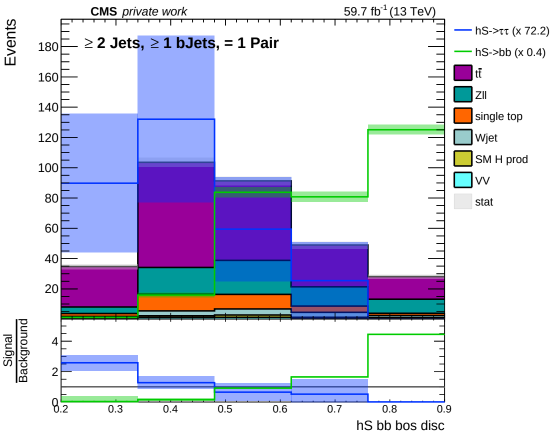


Figure: m_H = 600 GeV

Signal nodes split

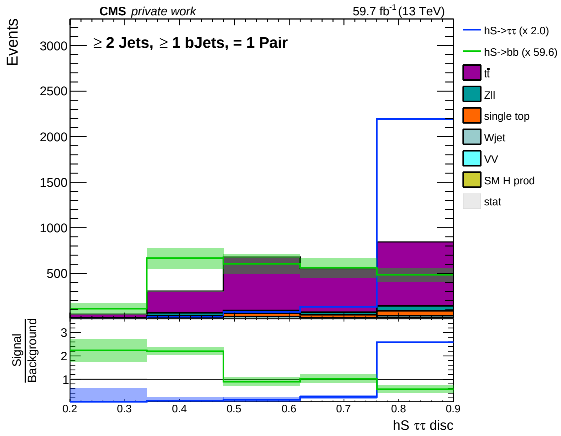


Figure: $m_H = 600\,\mathrm{GeV}$

Background nodes split

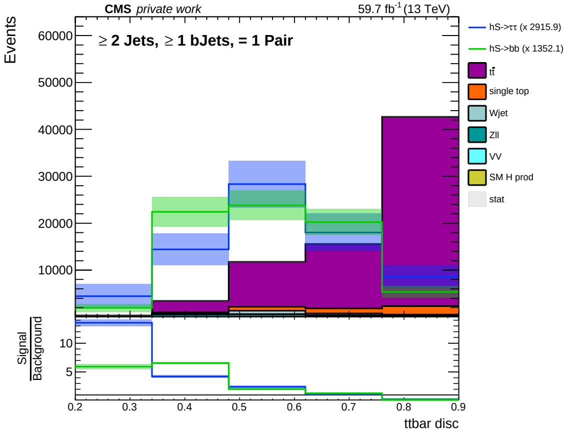


Figure: $m_H = 600\,\text{GeV}$

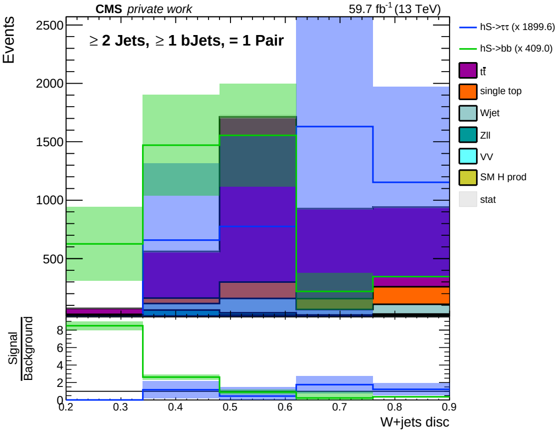


Figure: $m_H = 600\,\text{GeV}$

Background nodes split

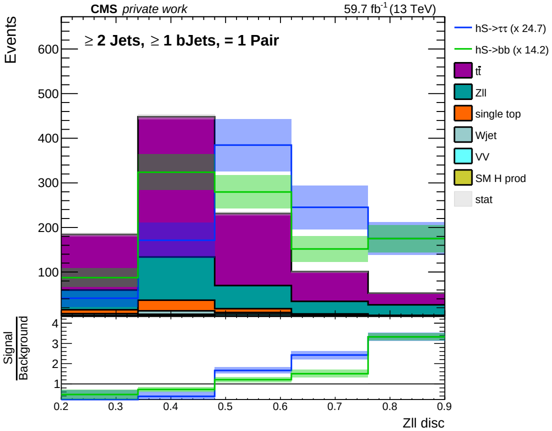


Figure: $m_H = 600\,\mathrm{GeV}$

Signal nodes split

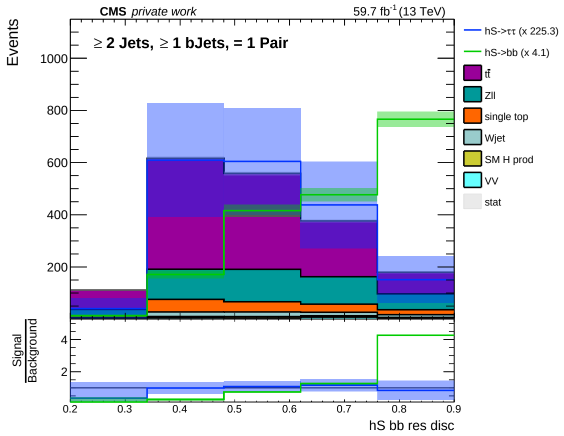


Figure: m_H = 1200 GeV

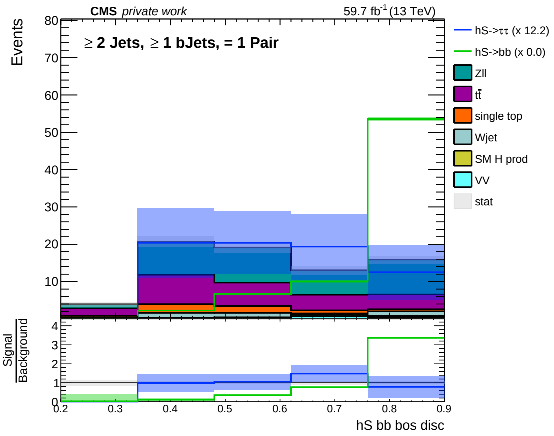


Figure: m_H = 1200 GeV

Signal nodes split

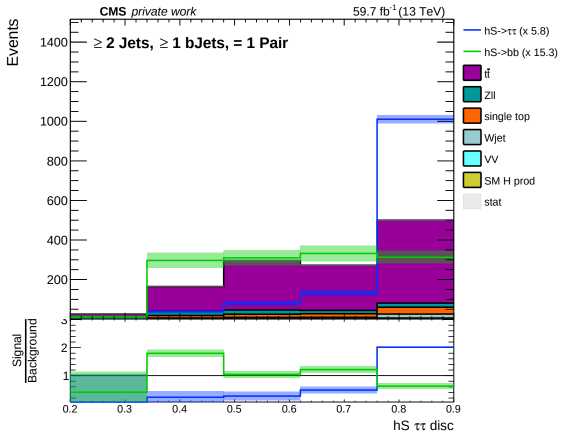


Figure: $m_H = 1200\,\text{GeV}$

Background nodes split

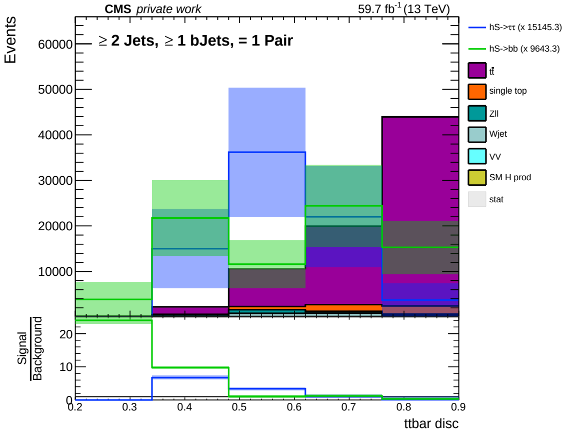


Figure: $m_H = 1200\,\text{GeV}$

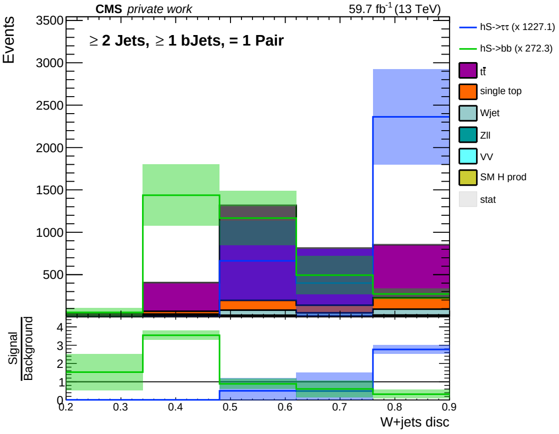


Figure: $m_H = 1200\,\text{GeV}$

Background nodes split

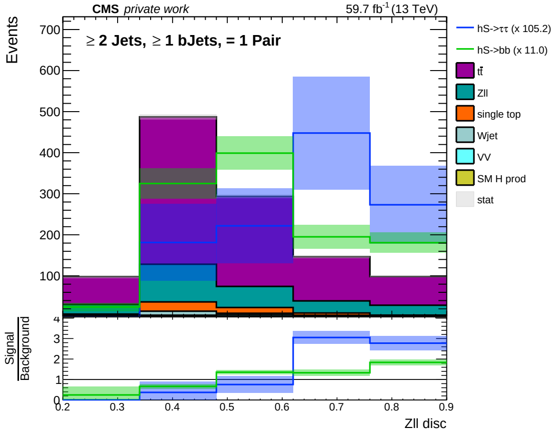


Figure: $m_H = 1200\,\mathrm{GeV}$

Signal nodes cut

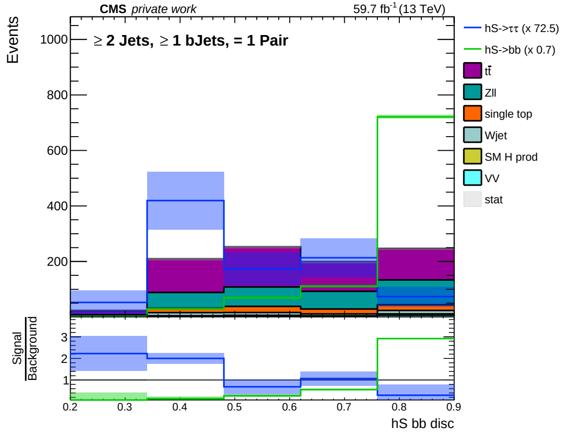


Figure: m_H = 600 GeV

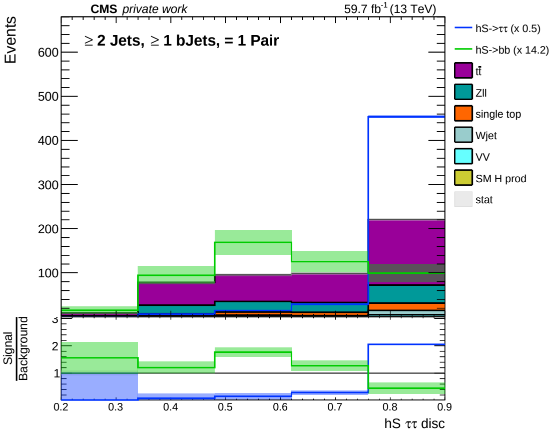


Figure: m_H = 600 GeV

Background nodes cut

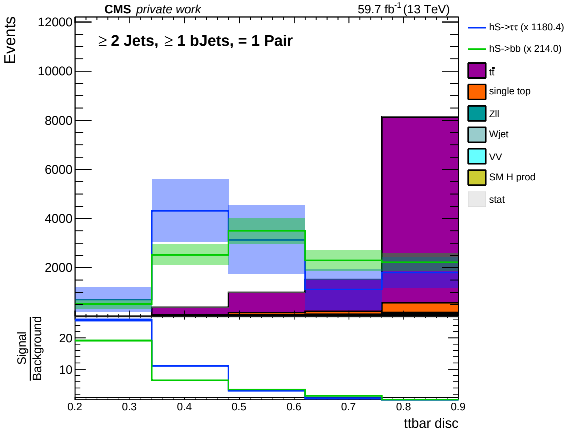


Figure: $m_H = 600\,GeV$

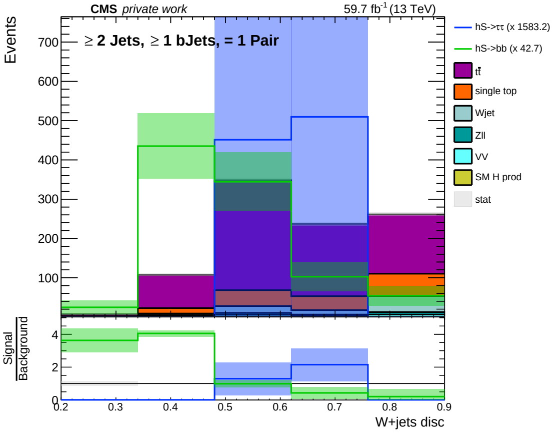


Figure: $m_H = 600\,GeV$

Background nodes cut

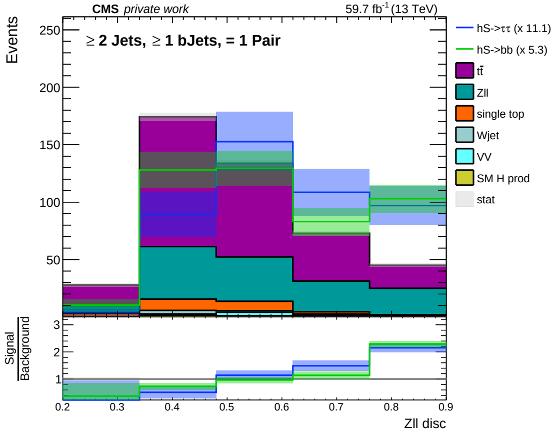


Figure: $m_H = 600\,GeV$

Signal nodes cut

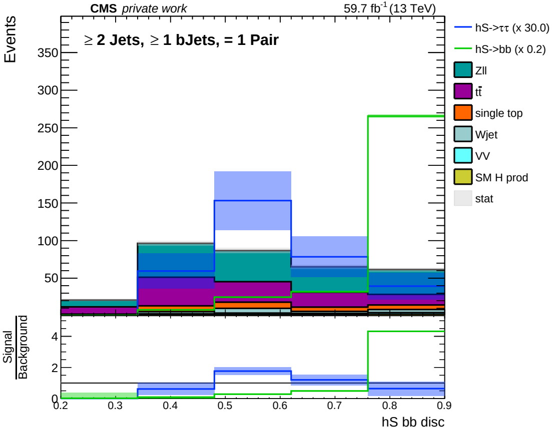


Figure: $m_H = 1200\,\mathrm{GeV}$

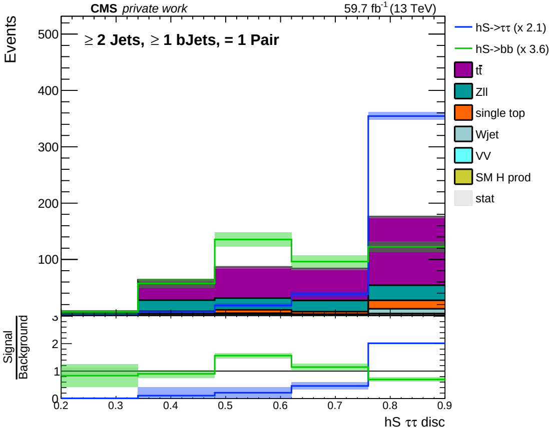


Figure: $m_H = 1200\,\mathrm{GeV}$

Background nodes cut

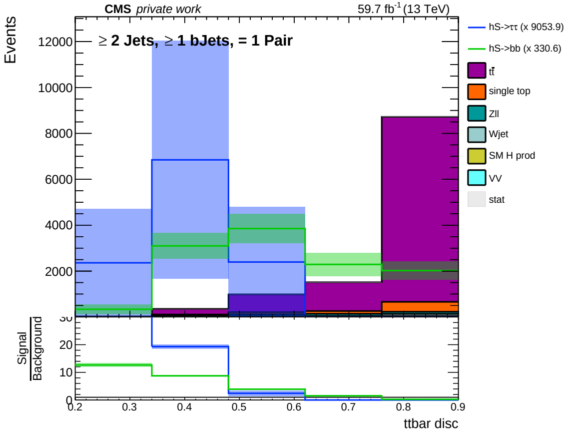


Figure: $m_H = 1200\,\text{GeV}$

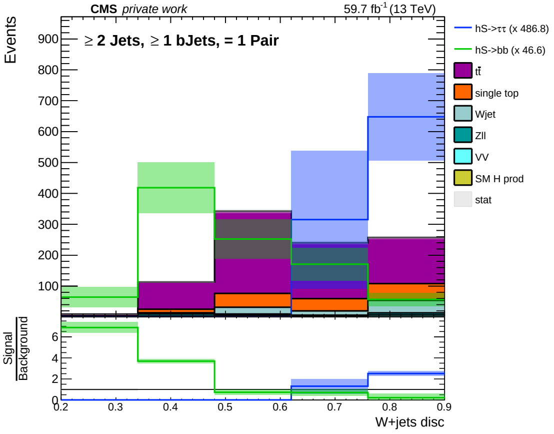


Figure: $m_H = 1200\,\text{GeV}$

Background nodes cut

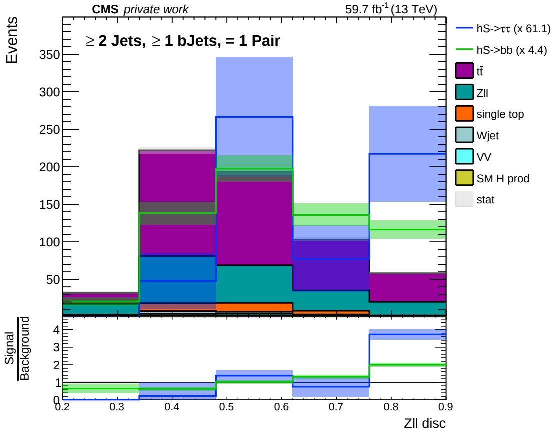


Figure: $m_H = 1200\,\mathrm{GeV}$

Improvements

Karlsruhe Institute of Technology

Table: Improvements on the upper limit on $\sigma B(\mathrm{H} \rightarrow \mathrm{hh_S} \rightarrow \tau\tau + \mathrm{bb})$ for each approach for a dedicated boosted topology description.

Changes	Average [%]	m_H [GeV]	Maximal [%]	m_H [GeV]	Minimal [%]
Feat.	29.83	1200	37.50	600	22.40
Split	26.17	1200	34.82	600	9.74
N_{AK8}	13.38	600	22.40	1200	0.00
Feat.+Split	47.49	1000	58.28	600	31.81
Feat.+N_{AK8}	37.73	700	40.15	1200	34.82
Split+N_{AK8}	34.50	900	39.00	600	28.57
Feat.+Split+N_{AK8}	50.12	1000	58.28	600	38.31

Best fit

- s+b Asimov toy
- example best fit using only specific output nodes
- $m_H = 600\,\text{GeV}$; $m_{h_S} = 60\,\text{GeV}$
- background nodes have only a small impact
- Zll node also sensitive on signal due to signal migration

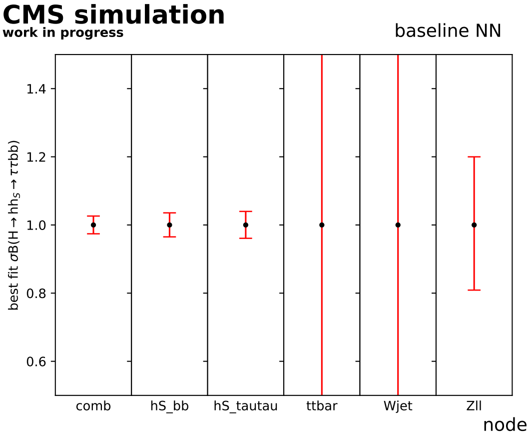
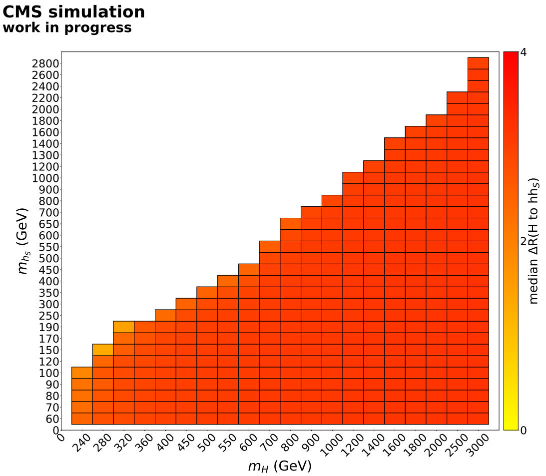


Figure: Statistical uncertainties only

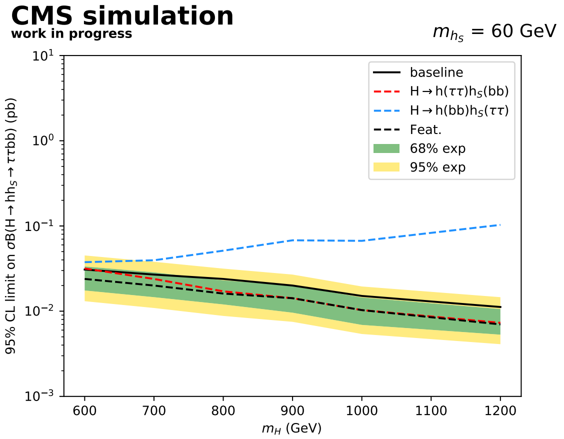
Event signature

- Studied the event signature
- Spacial distance between the two lighter Higgs bosons
- Back-to-back decay regardless of hypothesis
- Characterizing feature for signal events



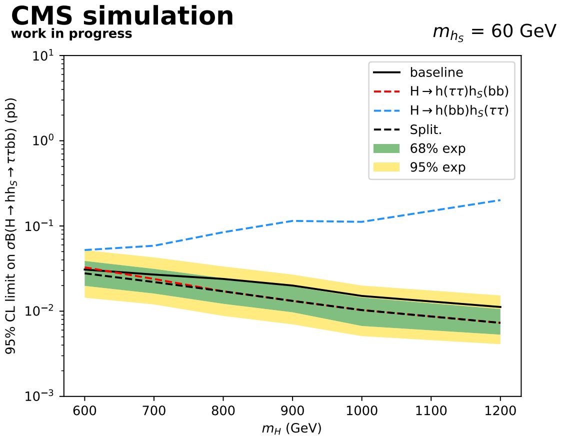
Expected upper limit: Additional input features

- Average improvement of around 30 %
- Stronger improvement for heavier m_H
- Correlates with discriminating power of added variables
- Higher signal over background ratio for heavier m_H



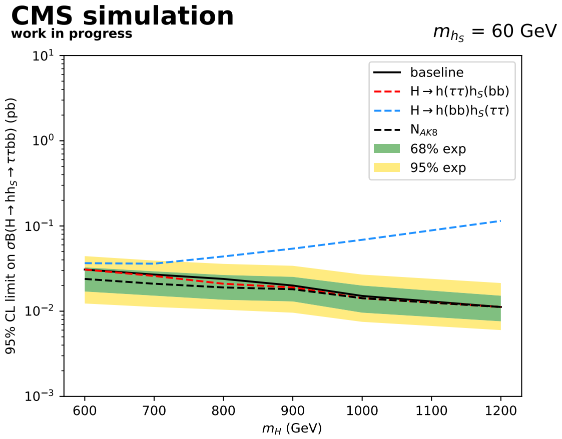
Expected upper limit: Splitting

- Average improvement of 26 %
- Stronger improvement for heavier m_H
- For lighter m_H the signal node discriminator for resolved topologies low
  - Worse discrimination
  - Phase space limited



Expected upper limit: Cut

- Average improvement of 13 %
- Stronger improvement for lighter m_H
- For heavier m_H the imposed cut limits the statistic
- Studied phase space limited



Input features

- p_T of the μ
- visible p_T of the τ_h
- visible m of the $\mu\tau_h$-pair
- visible p_T of the $\mu\tau_h$-pair
- number ob b-tagged jets
- p_T of the p_T-leading b-tagged jet
- p_T of the p_T-sub-leading b-tagged jet
- visible m of the two p_T-leading b-tagged jets
- visible p_T of the two p_T-leading b-tagged jets
- number of non b-tagged jets
- p_T of the p_T-leading non b-tagged jet
- p_T of the p_T-sub-leading non b-tagged jet
- m between the two p_T-leading non b-tagged jets
- p_T between the two p_T-leading non b-tagged jets
- spatial distance ΔR between the two p_T-leading non b-tagged jets
- visible m of the entire $\tau\tau$ + bb system