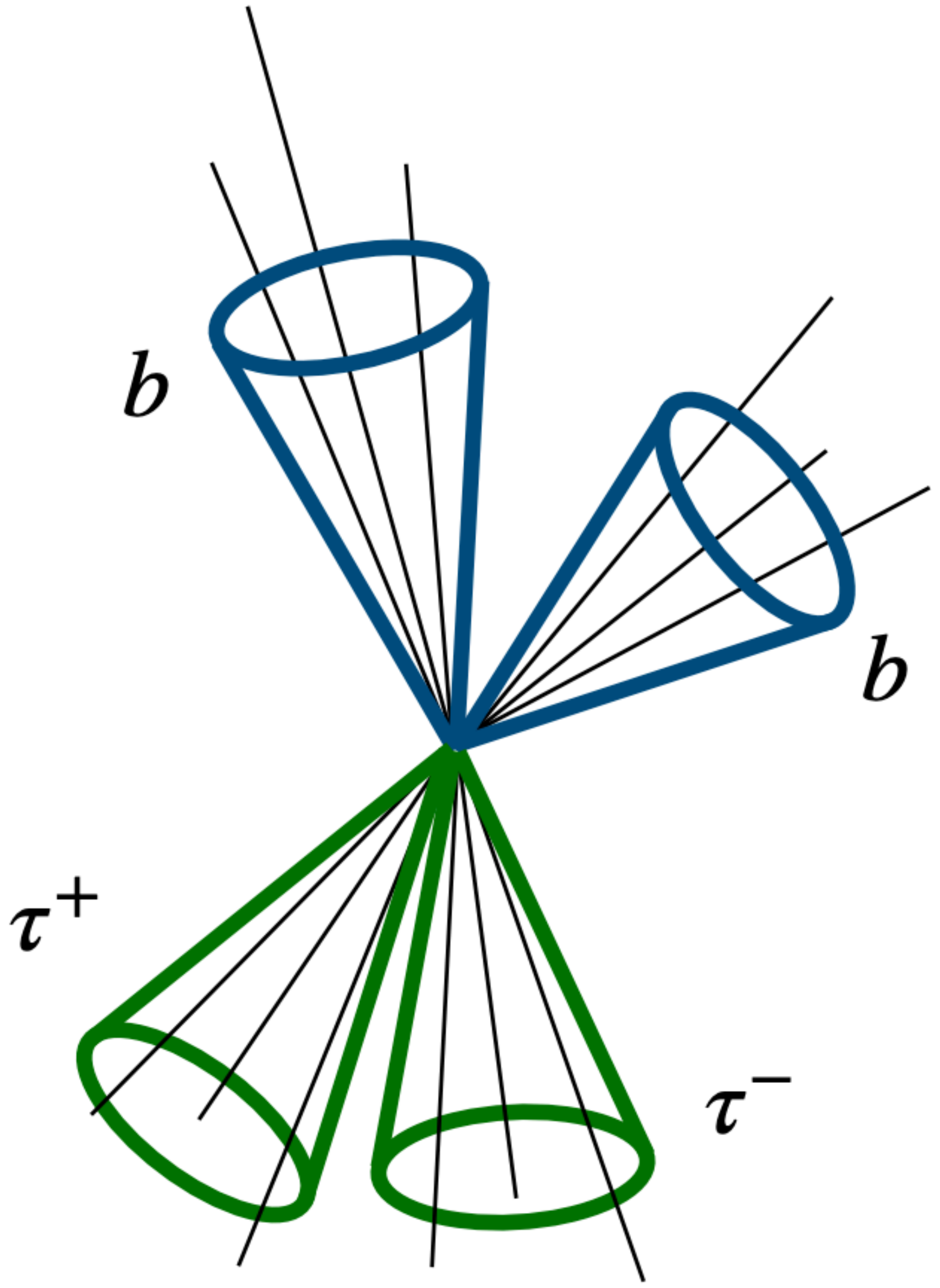
ATLAS non-resonant HH—> bbtautau at Run 2

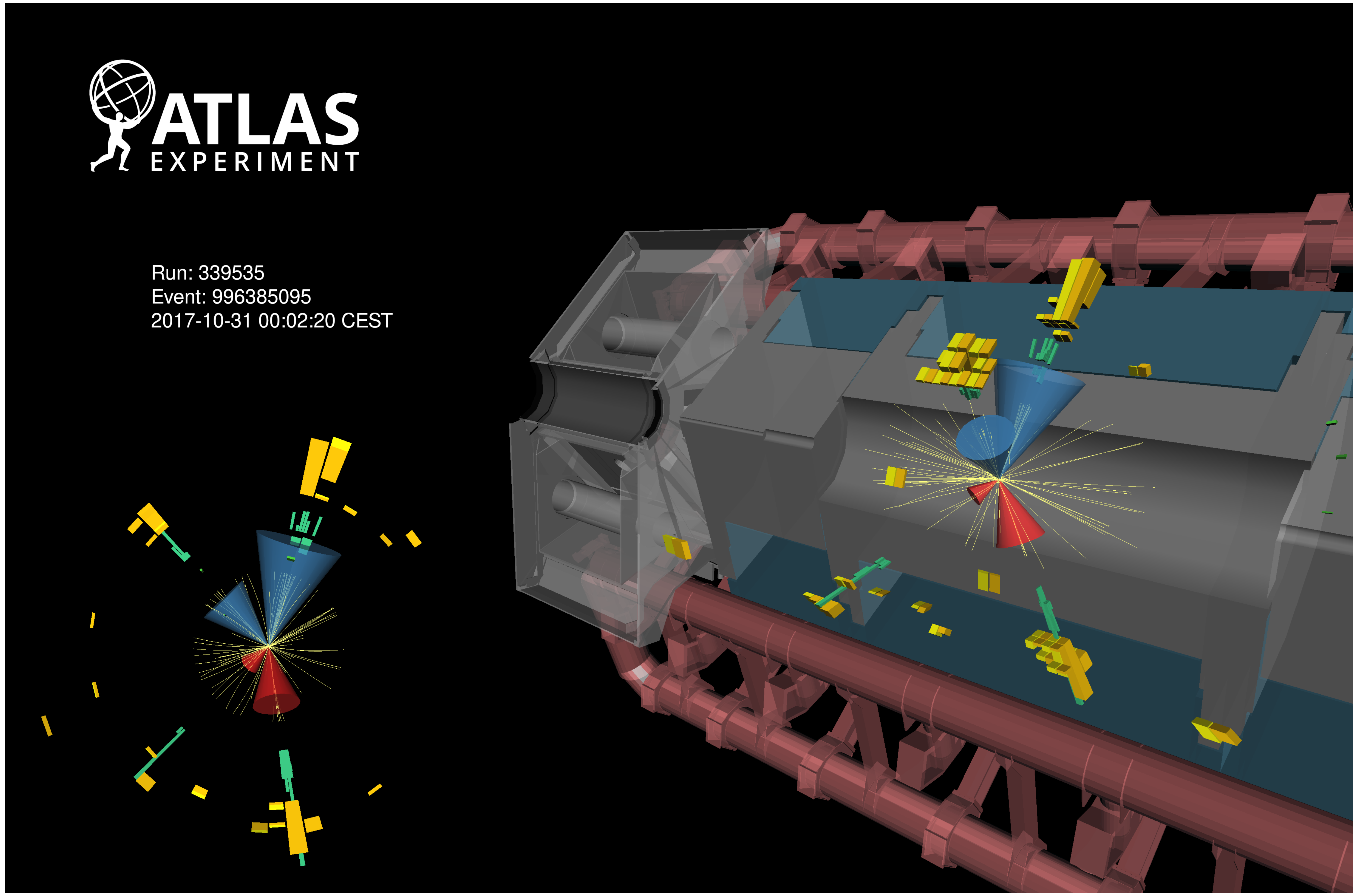
Liangliang Han
Nanjing University



Higgs Potential 2024, Dec 19-23, 2024

Outline

- **Motivation**
- **Analysis strategy**
- **Event selection**
- **Event categorization**
- **Background estimation**
- **Results**
- **Summary**



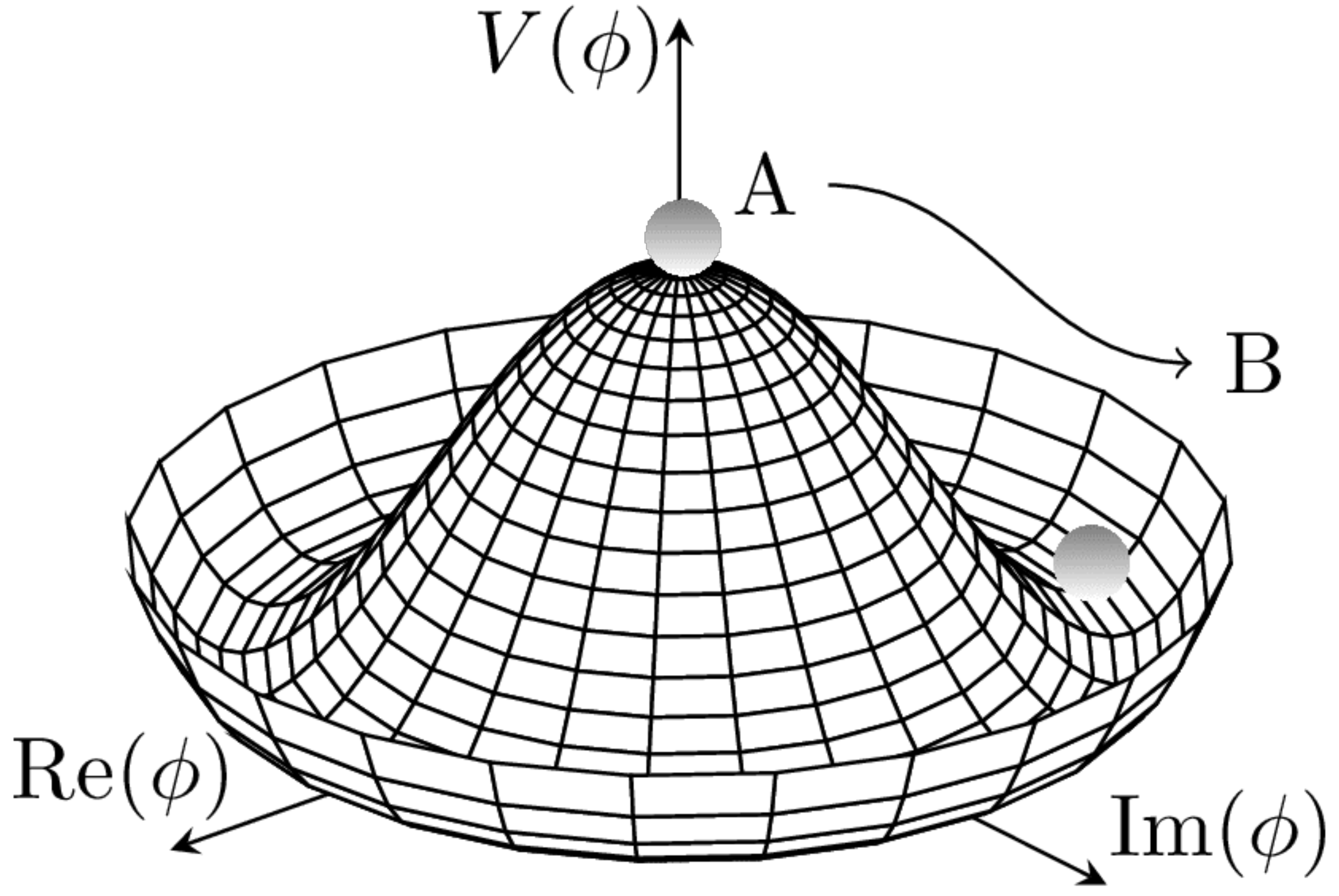
Source

Publication: Phys. Rev. D 110 (2024) 032012 **(this talk)**

Motivation

→ **The structure of Higgs potential is important to**

- Test the electroweak theory of standard model
- Learn more about the thermal evolution in the early universe
- Better understand the stability of the cosmic vacuum



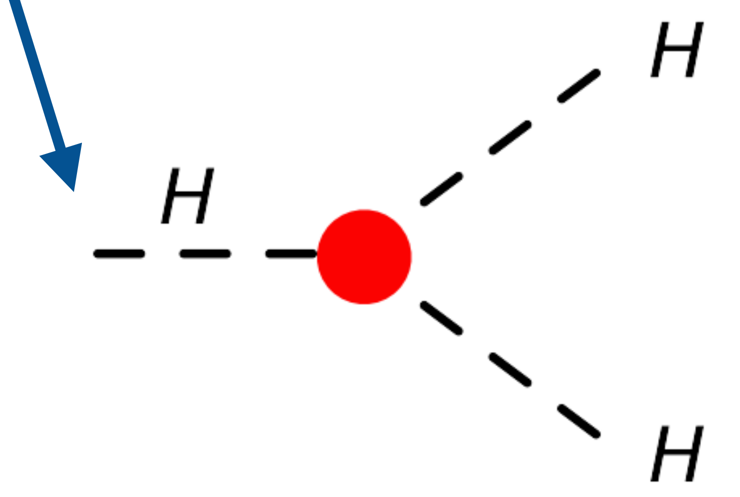
→ **Its shape can be probed by determining the Higgs self-coupling in HH search at LHC**

Trilinear self-coupling term

$$V(\Phi) = -\mu^2(\Phi^\dagger\Phi) + \lambda(\Phi^\dagger\Phi)^2 \xRightarrow{\qquad} = \frac{1}{2}m_H^2 H^2 + \boxed{\lambda v H^3} + \frac{\lambda}{4}H^4$$

Higgs potential

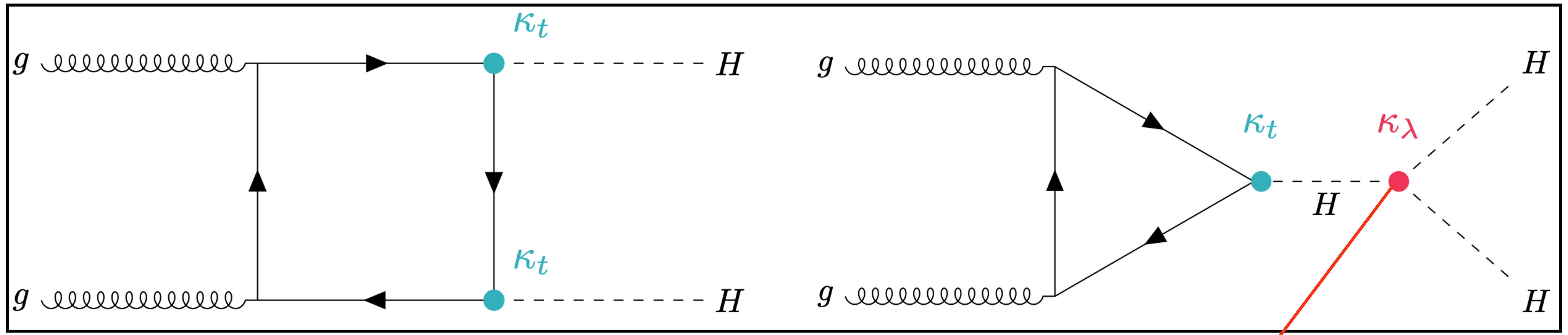
Electroweak symmetry breaking



HH production

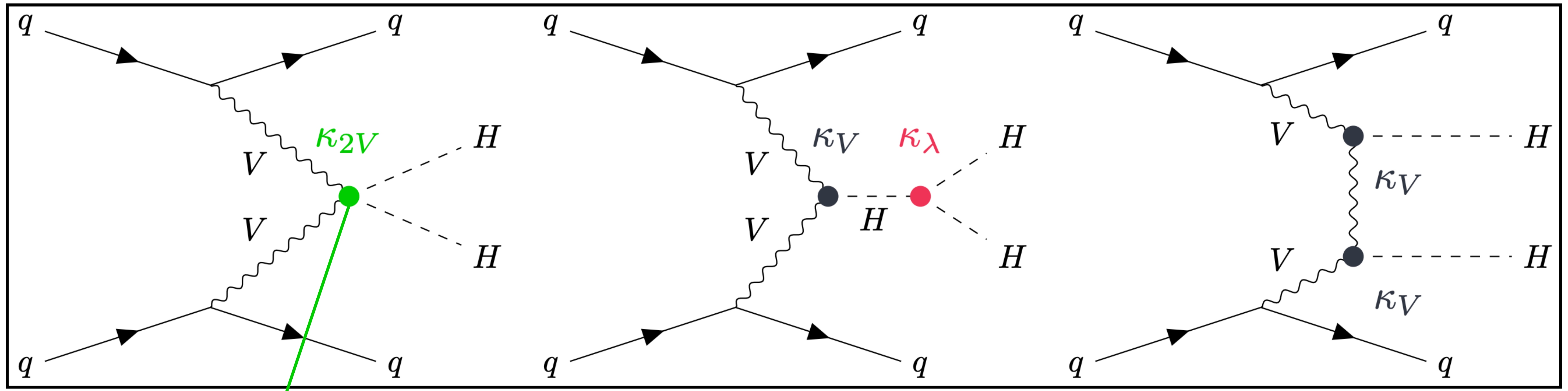
→ SM HH production @ LHC

- Dominant process is gluon-gluon fusion (ggF), $\sigma^{SM}_{ggF} = 31\ fb$



$$\kappa_\lambda = \frac{\lambda_{HHH}}{\lambda^{SM}_{HHH}}$$

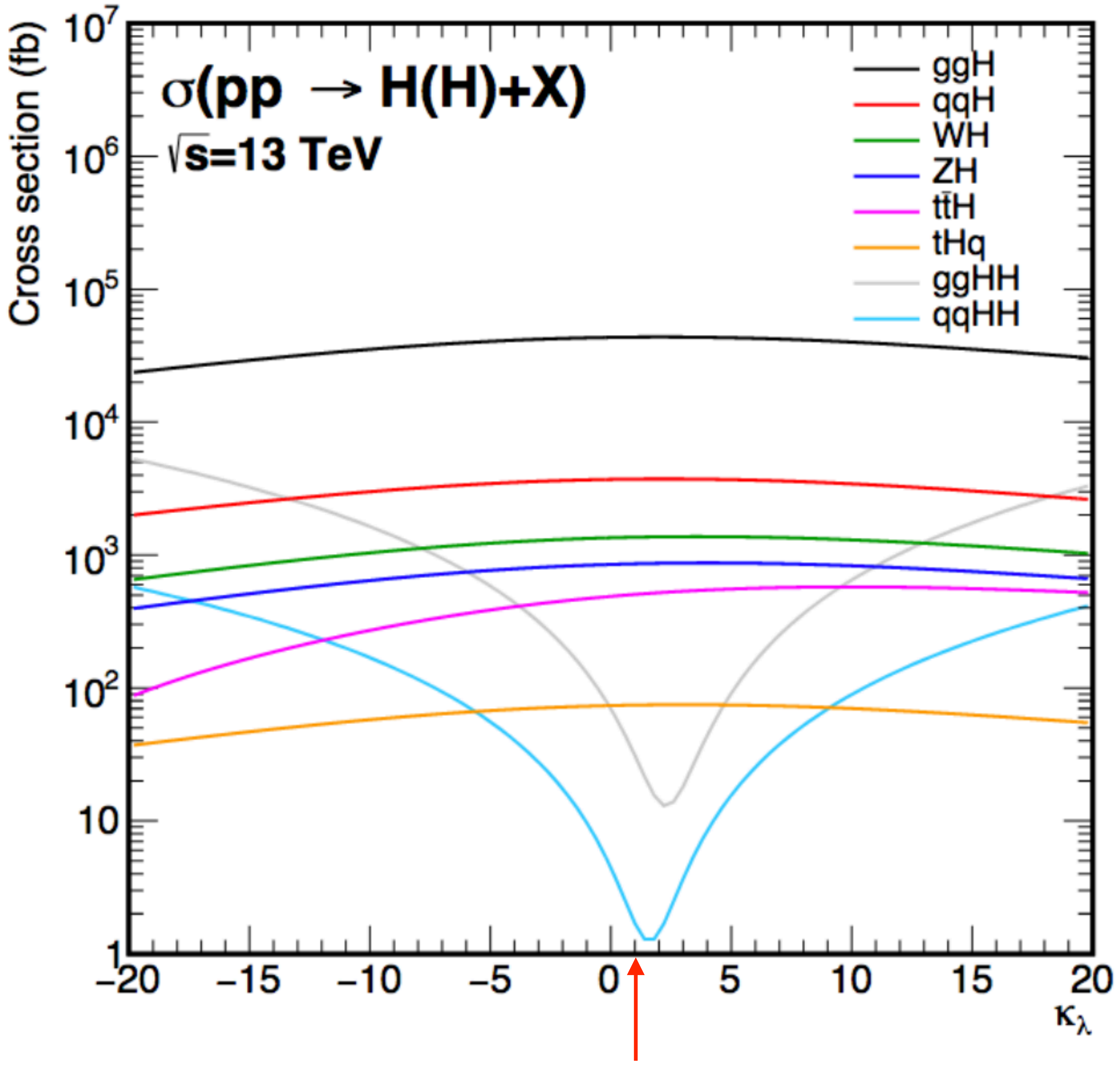
- Sub-dominant process is vector-boson fusion (VBF) , $\sigma^{SM}_{VBF} = 1.7\ fb$



$$\kappa_{2V} = \frac{\lambda_{HHVV}}{\lambda^{SM}_{HHVV}}$$

→ BSM HH production

- Lead to significant enhancement to HH production
- Allowed in many BSM scenarios, which makes it possible to probe new physics



SM : $\kappa_\lambda = 1$

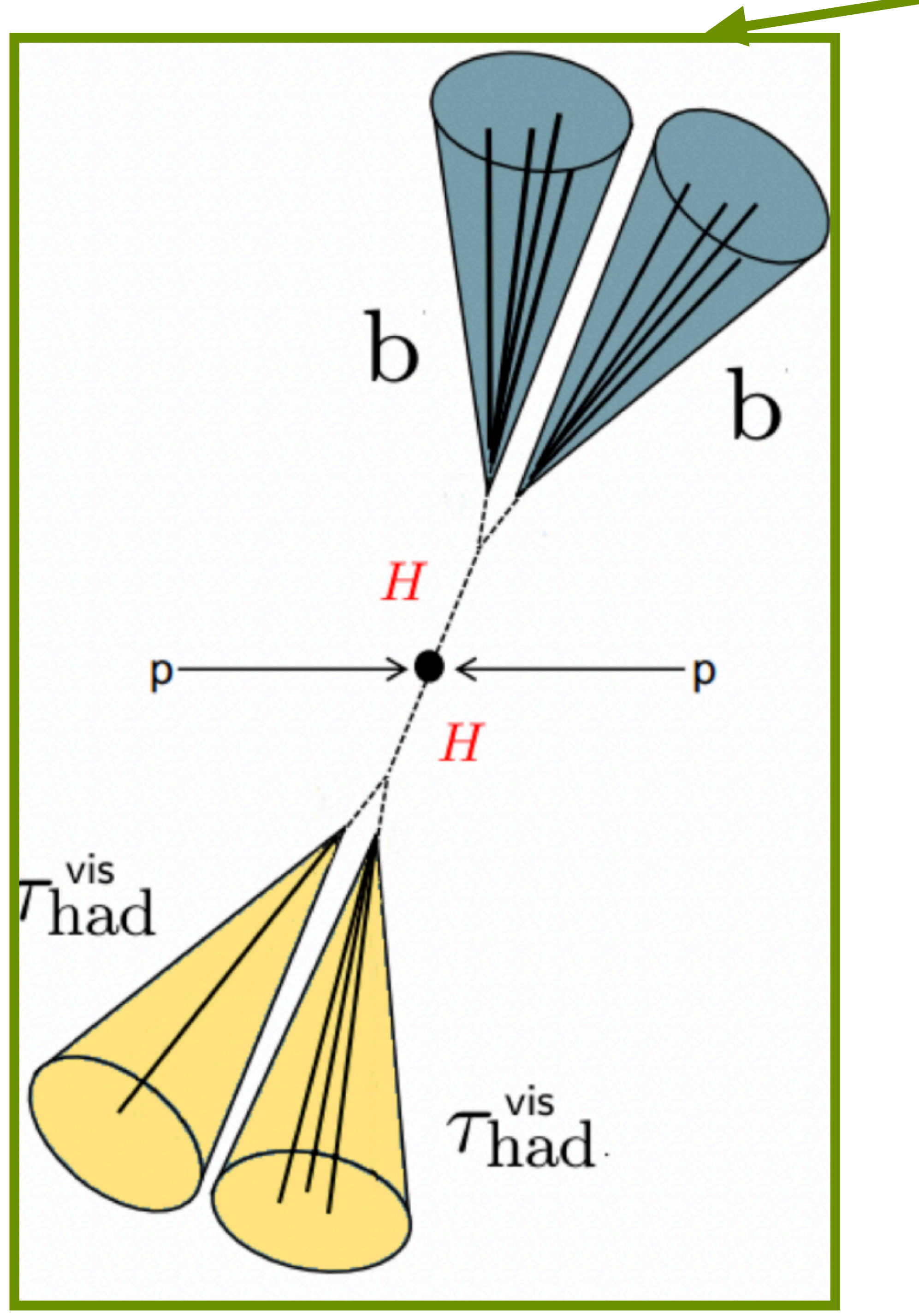
Analysis strategy

- → **Targeting at non-resonant HH—>bbtautau signal (2 b-tagged jets + OS tau-leptons)**
- → **One of the "golden 3" channels, ~7.3% branching ratio**
- → **Two analysis channels, depending on the Di-tau decay mode: hadhad and lephad**

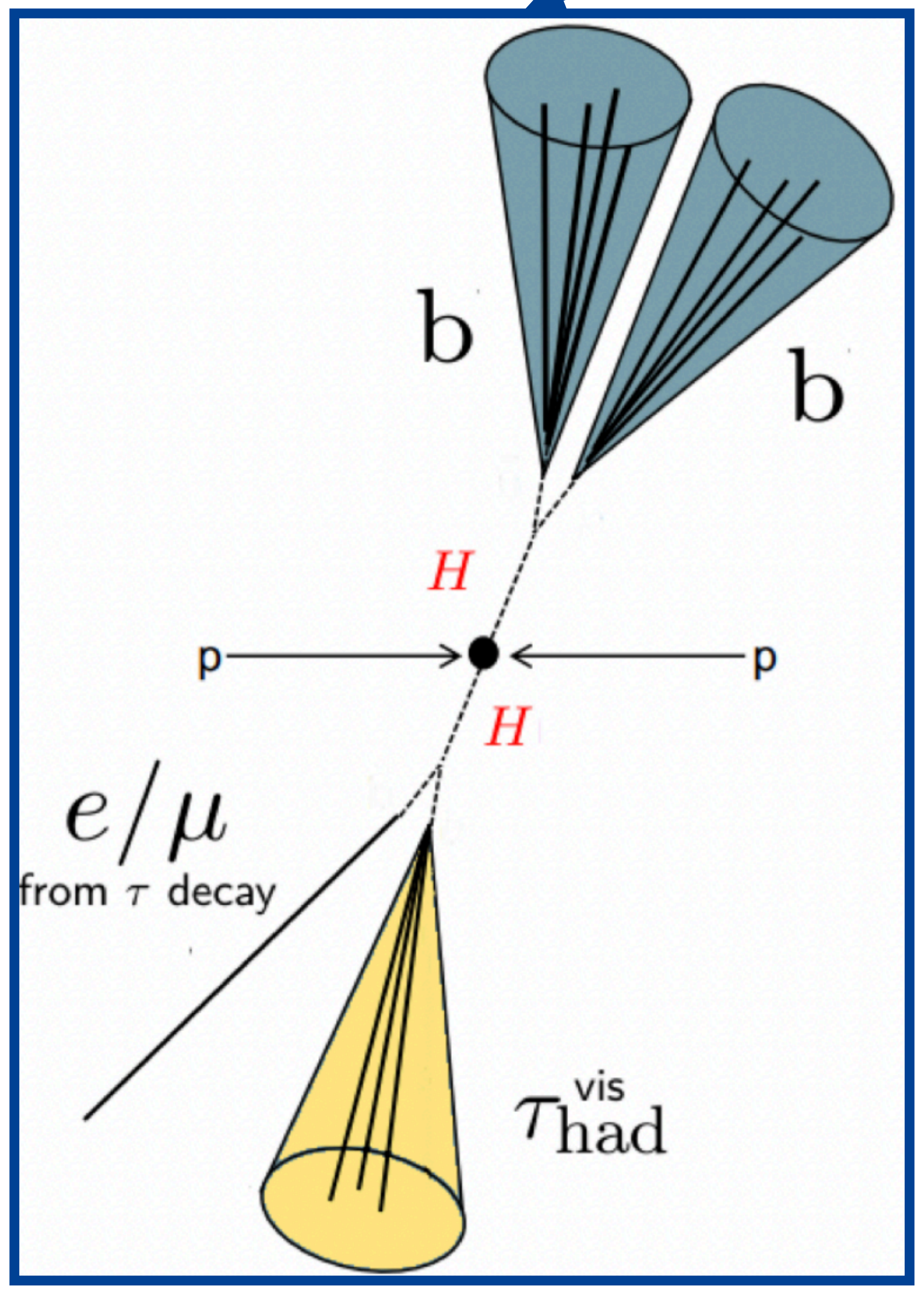
Large branching ratio

Clean final state

	bb	WW	ττ	ZZ	γγ
bb	34 %				
WW	25 %	4.6 %			
ττ	7.3 %	2.7 %	0.39 %		
ZZ	3.1 %	1.1 %	0.33 %	0.069 %	
γγ	0.26 %	0.10 %	0.028 %	0.012 %	0.0005 %



hadhad



lephad SLT (single lepton trigger)

lephad LTT lepton+tau trigger

Analysis strategy

→ **Re-analysis of the full Run 2 dataset based on previous analysis, which:**

- Only optimized for SM HH ggF mode—> inclusive ggF SR
- Signal extraction: BDT in hadhad, Neural Network (NN) in lephad

This analysis:

⇩

→ **Data samples: full Run 2 dataset @ 13TeV**

→ **Signal extraction: Optimized BDT for both lephad and hadhad channel**

→ **Targeting at both ggF and VBF production modes**

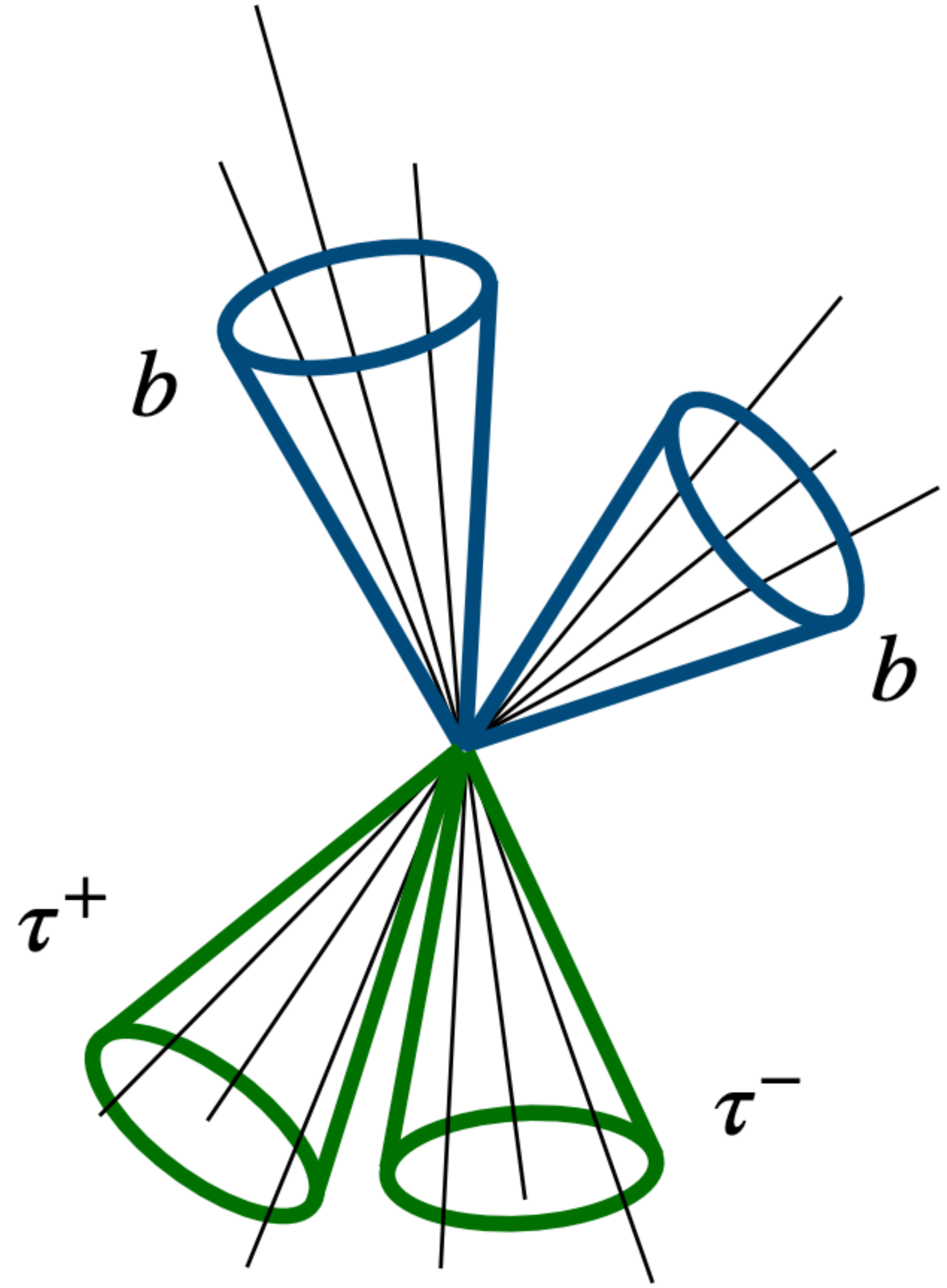
→ **Main backgrounds**

- Top quark, Z boson + jets (heavy flavor), multi-jet, diboson, single Higgs boson

Event selection $HH -> b\bar{b}\tau^+\tau^-$

→ **Search for HH with bbtautau final states**

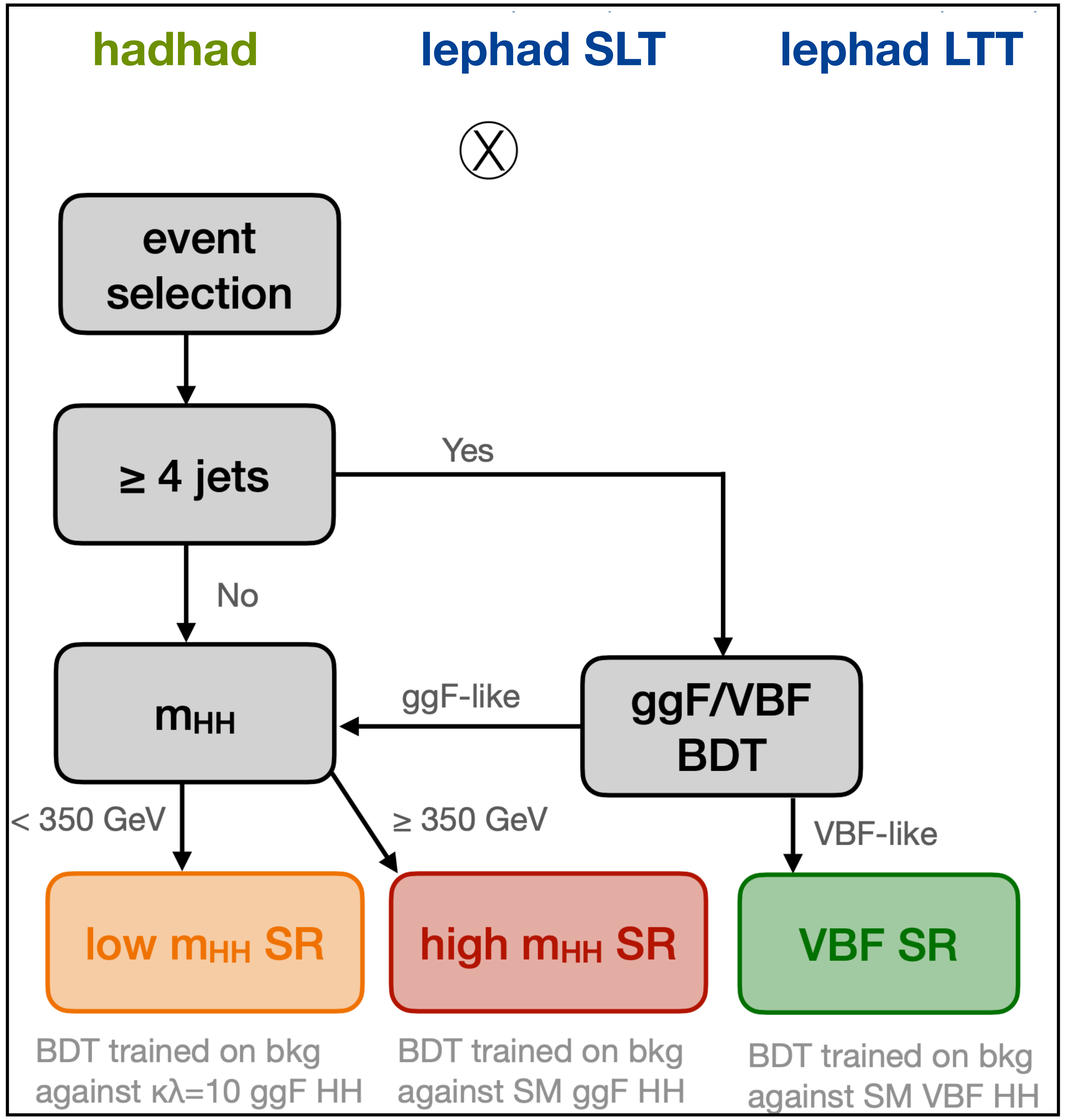
	$\tau_{lep}\tau_{had}$	$\tau_{had}\tau_{had}$
Triggers	Single Lepton (e/μ) triggers (SLT) Or Lepton + τ_{had} triggers (LTT)	Single τ_{had} triggers (STT) Or Di-τ_{had} triggers (DTT)
	Offline Requirements Passed	
Event Selection	$m_{\tau\tau}^{MMC}$ [*] > 60 GeV	
	Opposite-sign of $e/\mu/\tau_{had}$ and τ_{had}	
	Exactly two b-tagged jets	
	One (tight) e or (medium) μ	No loose e/μ
	One (loose) τ_{had}	Two (loose) τ_{had}
	$M_{bb} < 150$ GeV	



Event categorization

→ Extended categorization (VBF, low ggF and high ggF) for each sub-channel

- To improve the constraint on κ_λ and κ_{2V}



hadhad **lephad SLT** **lephad LTT**

$\otimes$

event selection

≥ 4 jets

Yes

No

m_{HH}

ggF-like

ggF/VBF BDT

< 350 GeV

≥ 350 GeV

VBF-like

low m_{HH} SR

high m_{HH} SR

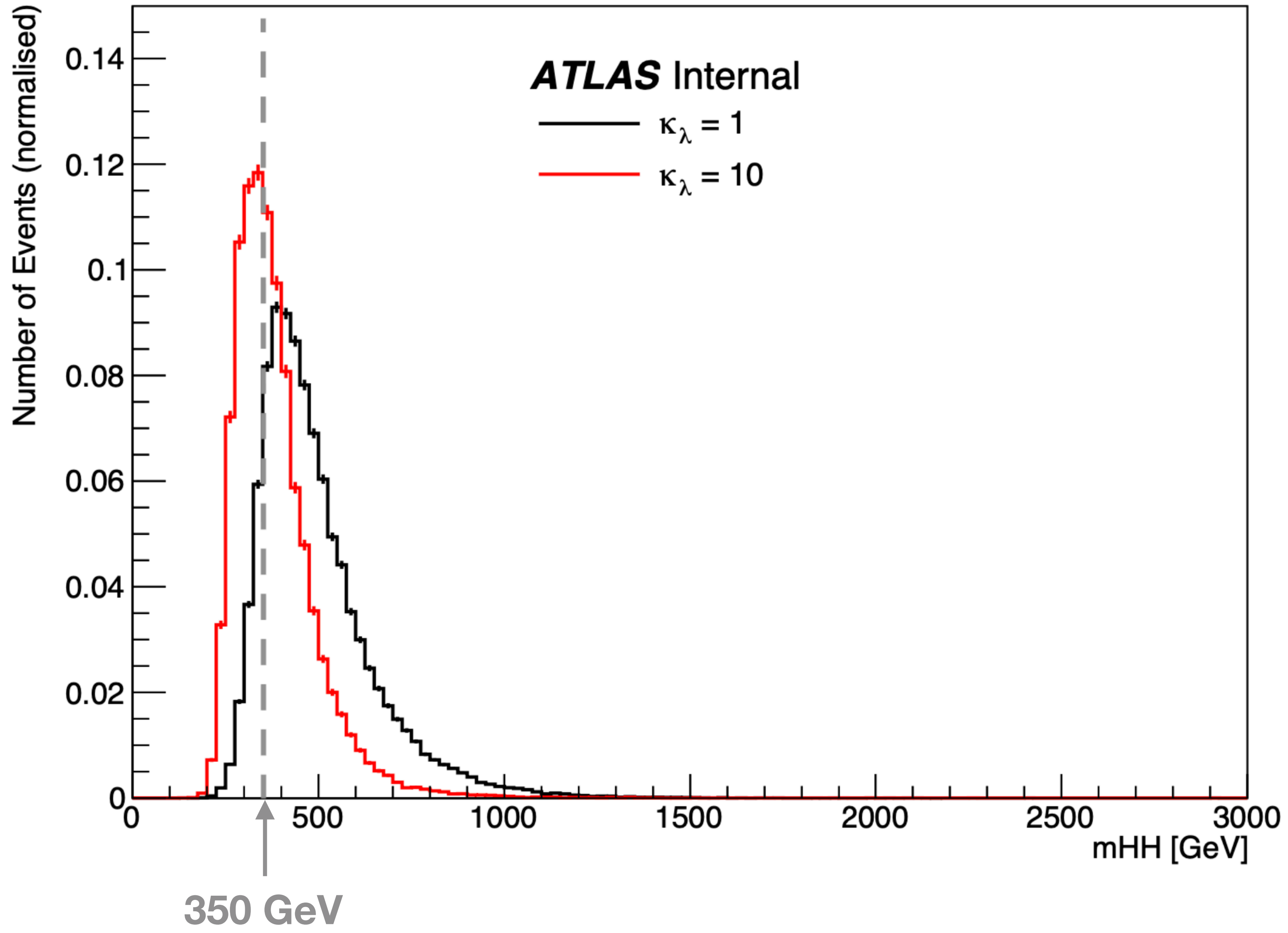
VBF SR

BDT trained on bkg against $\kappa\lambda$=10 ggF HH

BDT trained on bkg against SM ggF HH

BDT trained on bkg against SM VBF HH

mHH categorization to further improve the κ_λ constraint



Train dedicated BDT in the three regions

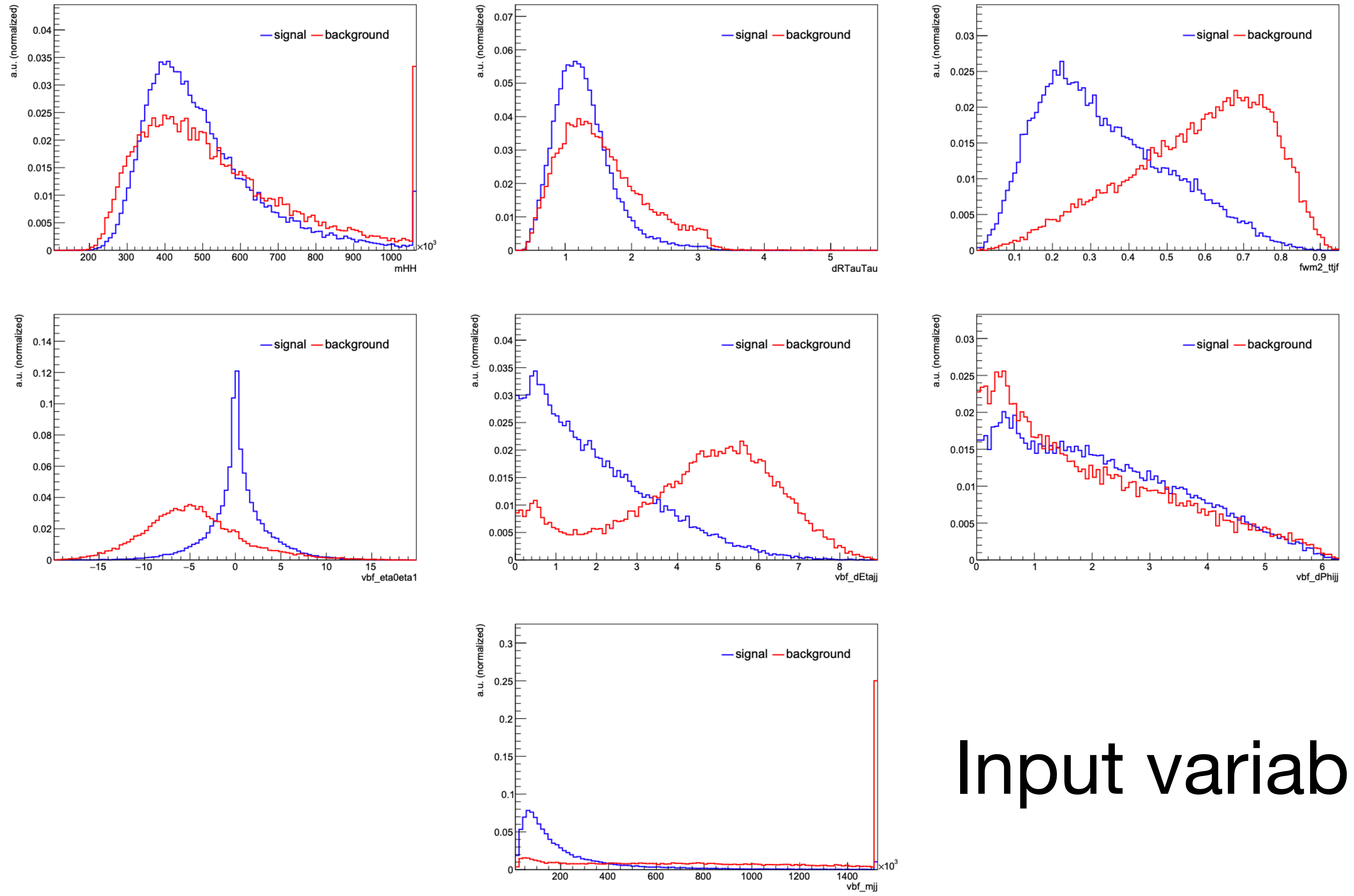
ggF vs VBF categorization BDT

→ A dedicated ggF vs VBF categorization BDT is trained

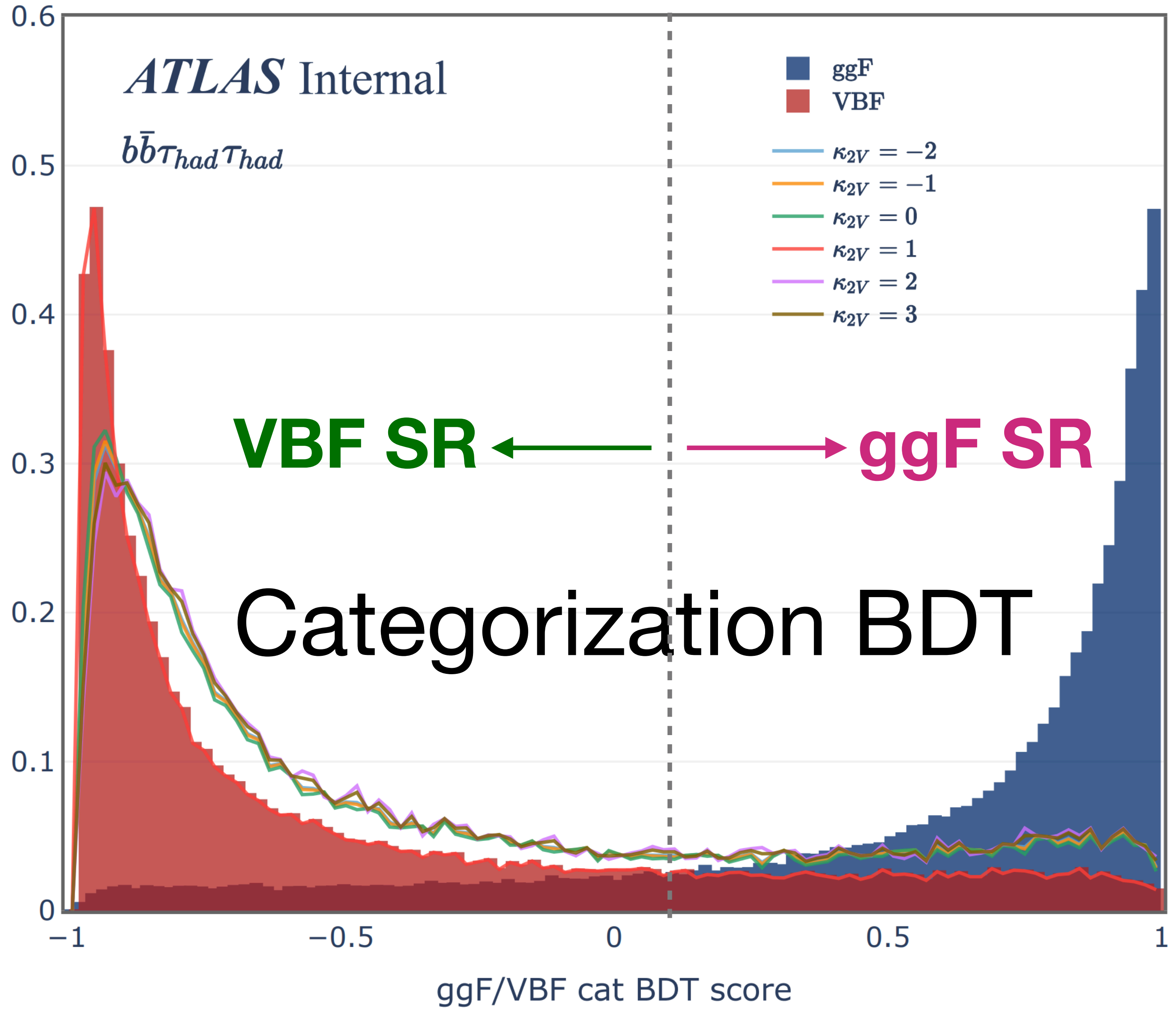
- To separate VBF HH from ggF HH on the events with 4 jets (2 VBF-jet candidates + 2 H—>bb)
- Input variables are typically **VBF-related quantities** and **event shape variables** (Fox Wolfram Moments)

$$m_{jj}^{\mathrm{VBF}} \qquad \Delta\eta_{jj}^{\mathrm{VBF}} \qquad \Delta R_{jj}^{\mathrm{VBF}} \qquad \mathrm{VBF}\ \eta_0 \times \eta_1$$

- Working point: cut value scan on the categorization BDT

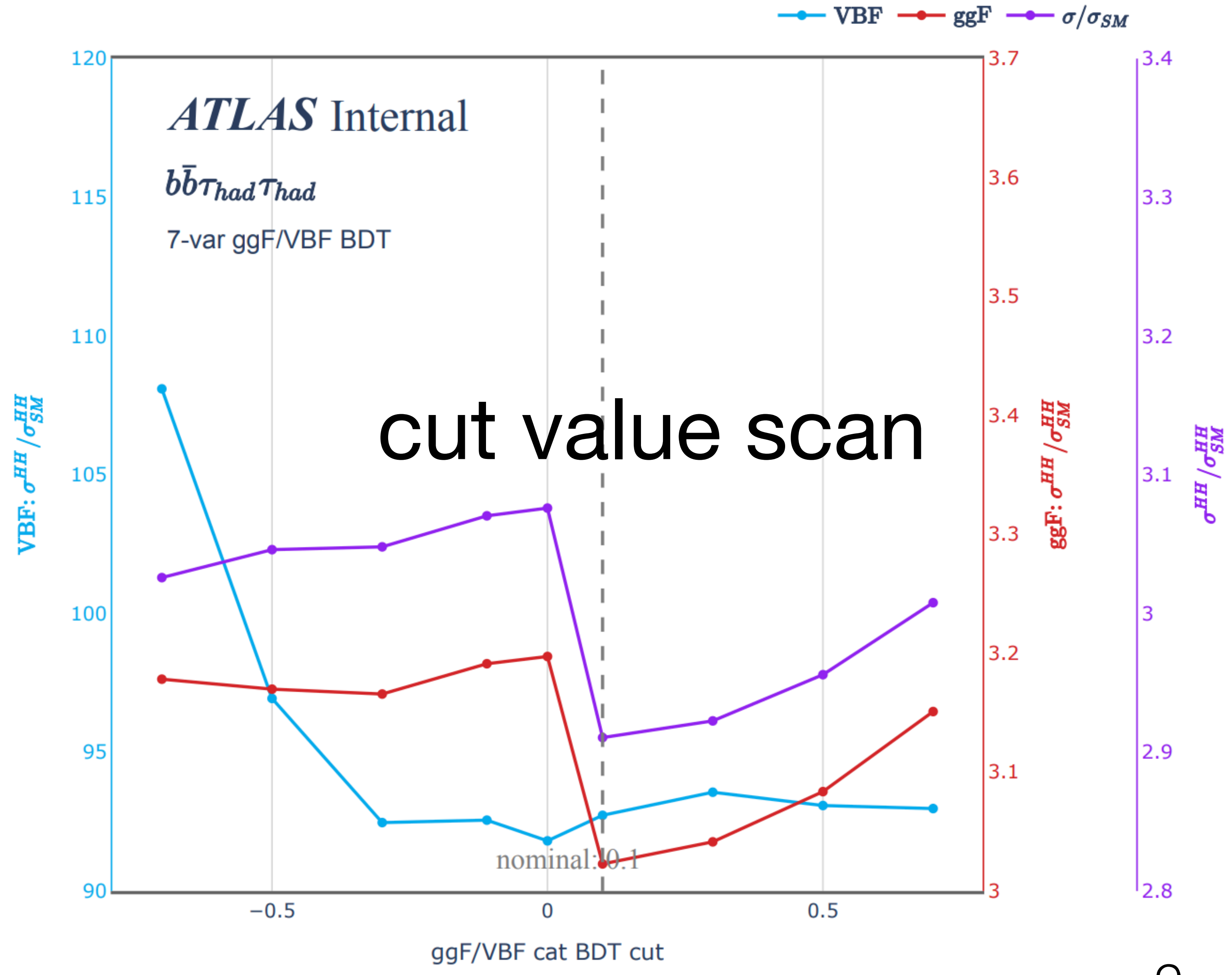


Input variables



VBF SR ← → ggF SR

Categorization BDT



cut value scan

Signal extraction BDT

➔ 3-fold training

All events are used in training, validation and test :)

- Divide events into 3 folds based on the event number
- Train 3 BDTs on each fold, and optimized and applied on other folds

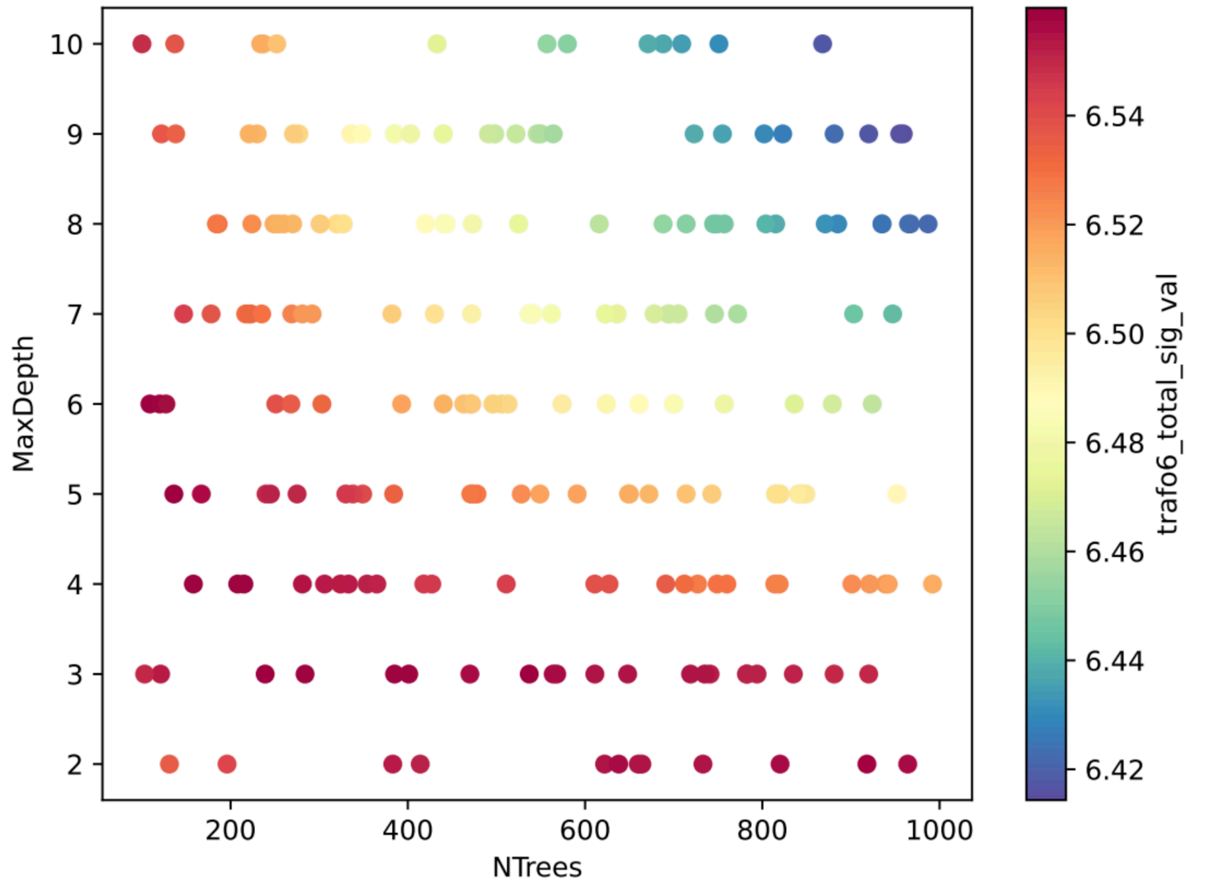
Model	Fold 0 event_number %3 = 0	Fold 1 (event_number %3 = 1)	Fold 2 (event_number %3 = 2)
BDT 0	Training	Validation	Testing
BDT 1	Testing	Training	Validation
BDT 2	Validation	Testing	Training

➔ Input variables selection

- Gradually add one more variable with the most improvement to the sensitivity
- Until reach a plateau where the sensitivity doesn't increase any more

➔ Hyperparameter optimization

- Take the set of hyper parameters that gives best significance



Background estimation

→ Main backgrounds

- Single top
- Single Higgs
- Diboson

⇒ **Estimated from simulation**

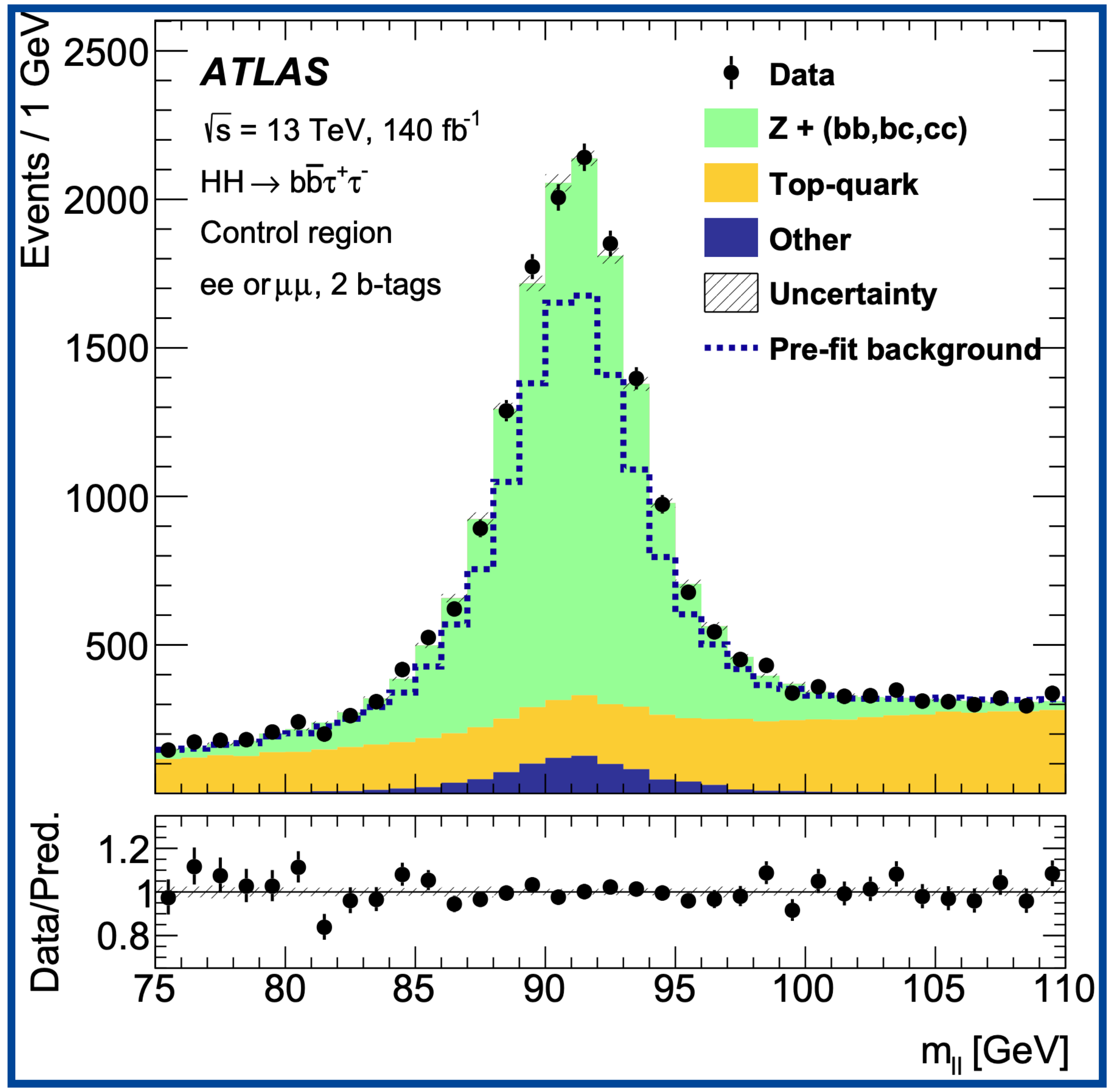
- ttbar
- Z + heavy flavor jets

⇒ **Simulation + dedicated control region**

- ttbar with jet—>tau fakes
- multijet with jet—>tau fakes

⇒ **Data-driven**

- Use fake factor method
  (From anti-ID CR extrapolated into SR)



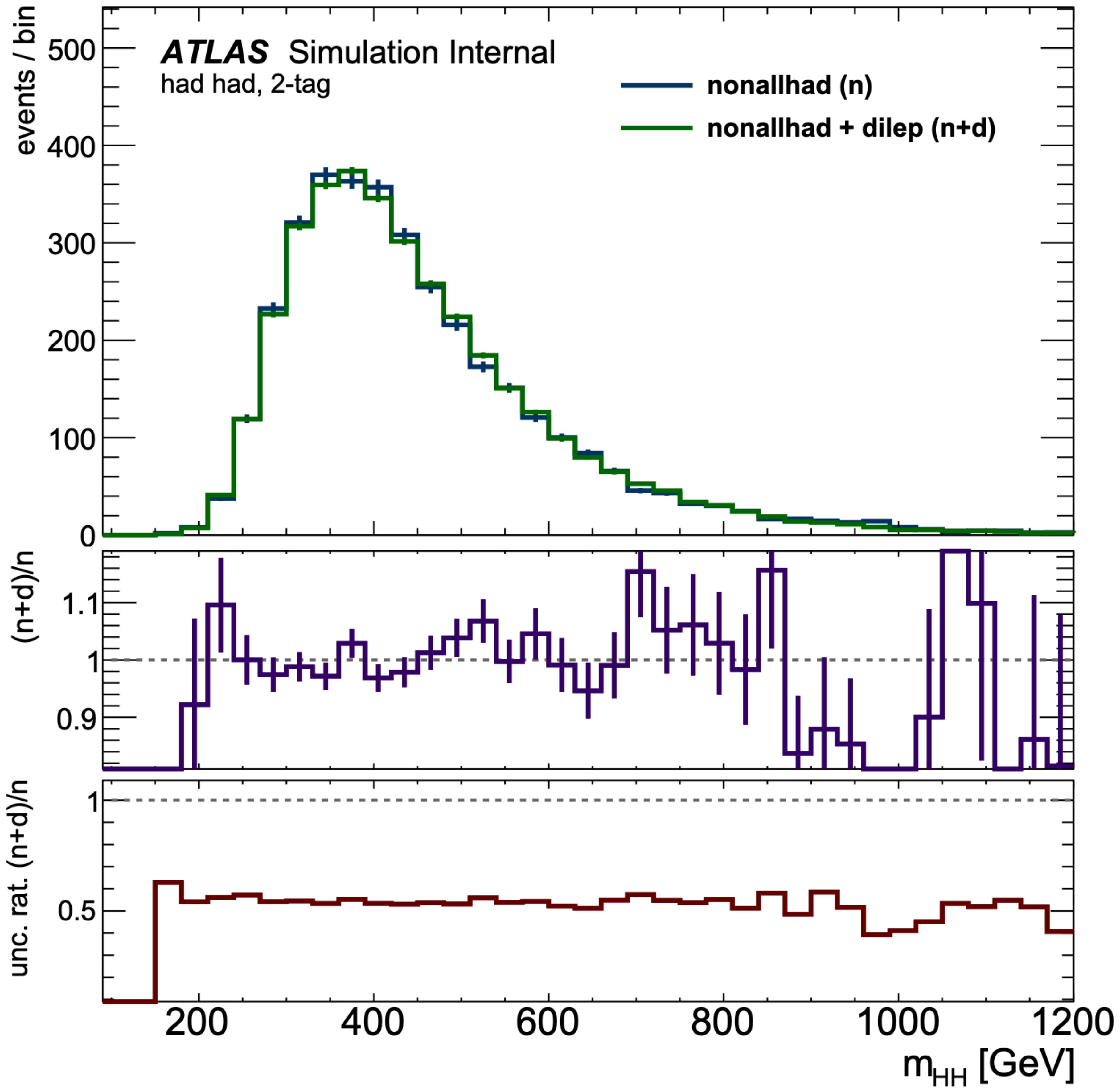
- ZCR: bbll trigger selection
  Exactly 2 OS muons or electrons
  Exactly 2 b-tagged jets
  mll window 75-110 GeV
  mBB < 40 GeV or mBB > 210 GeV
- Typical norm factors

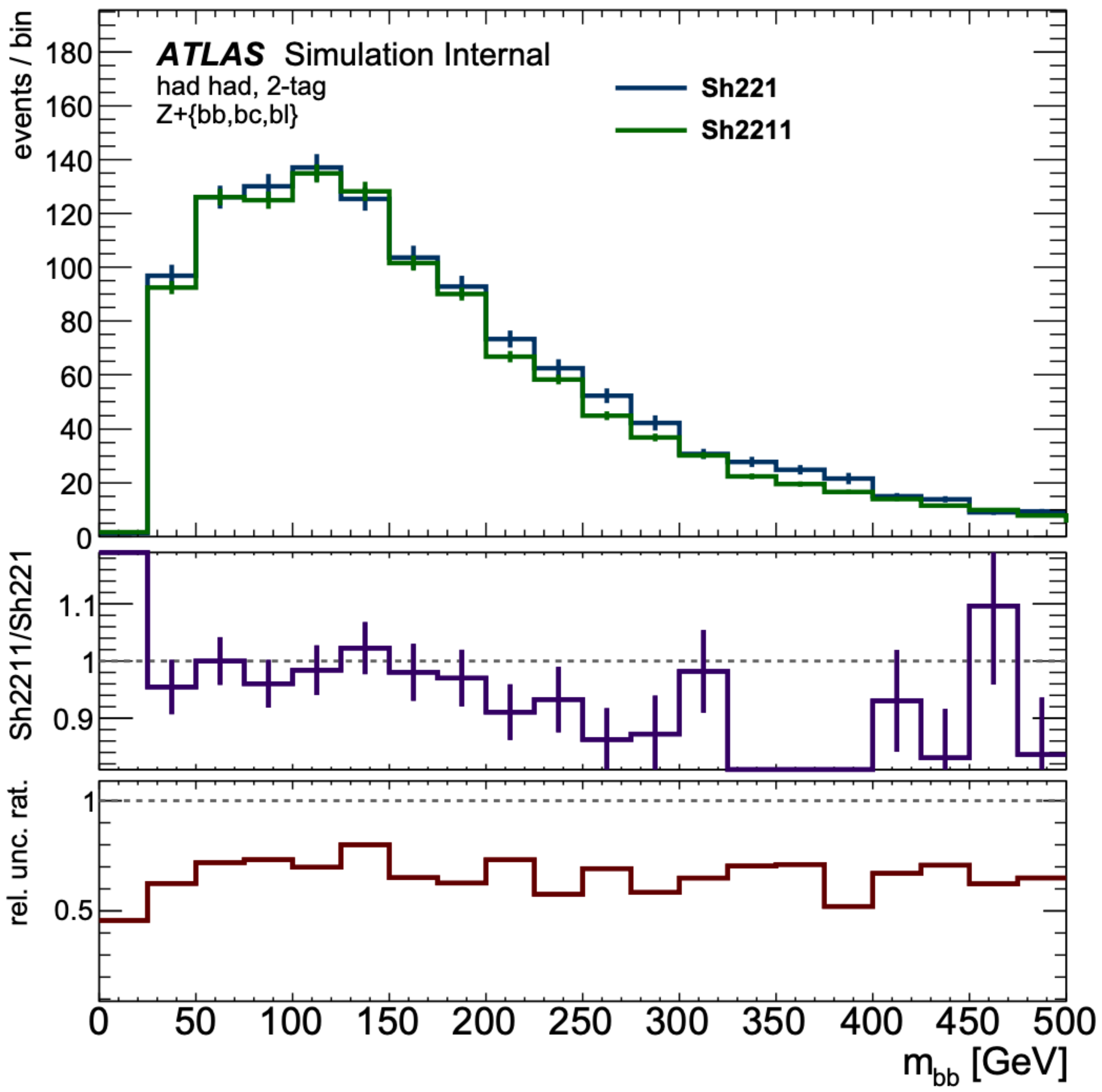
Z+HF	**1.34 ± 0.08**
ttbar	**0.96 ± 0.03**

Background estimation

➔ Improved MC description

- Inclusion of ttbar di-lepton sample
- V+jets changeover from sherpa 2.2.1 to 2.2.11



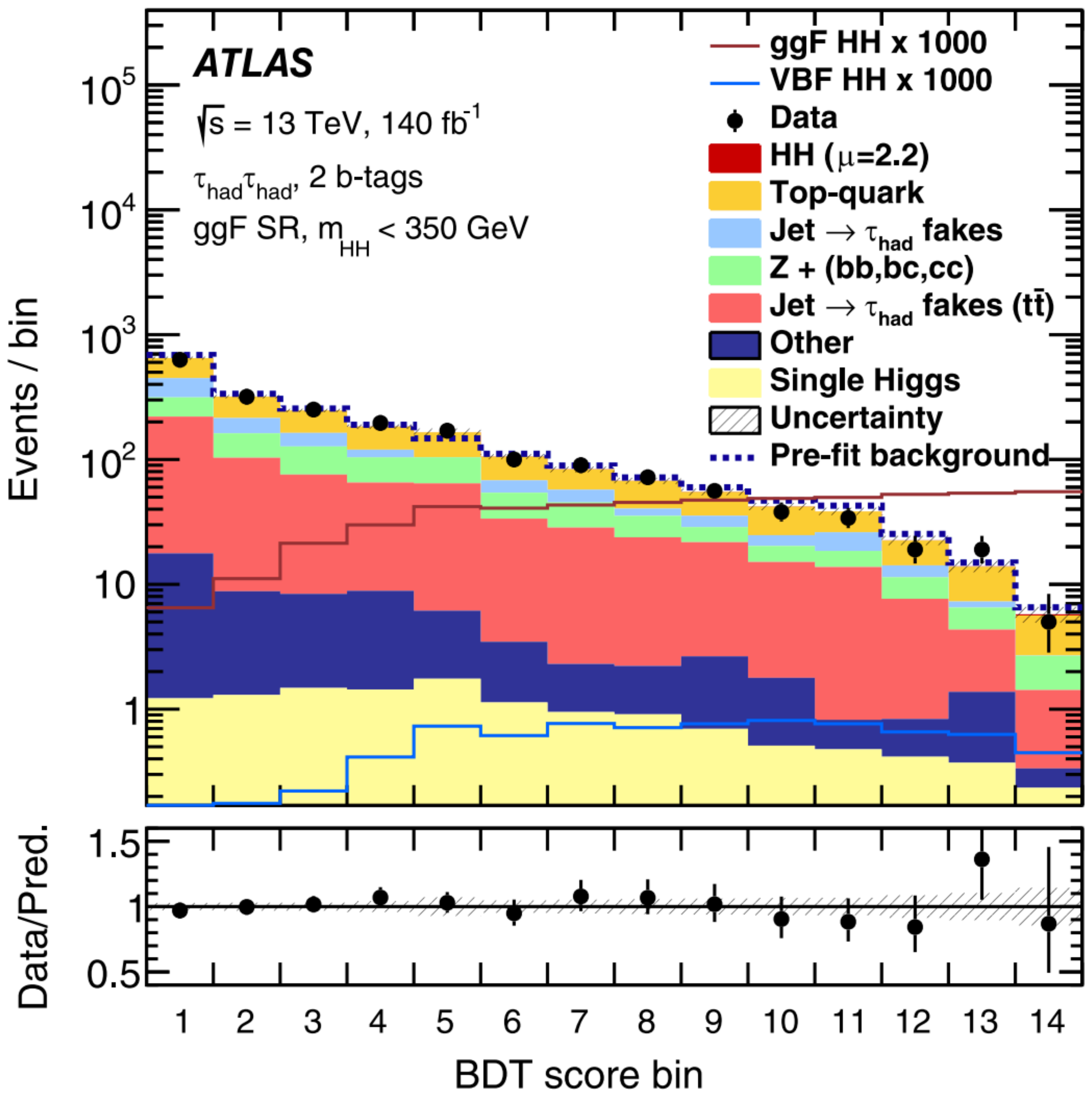


Reduction of the MC statistical uncertainty by a factor of ~2

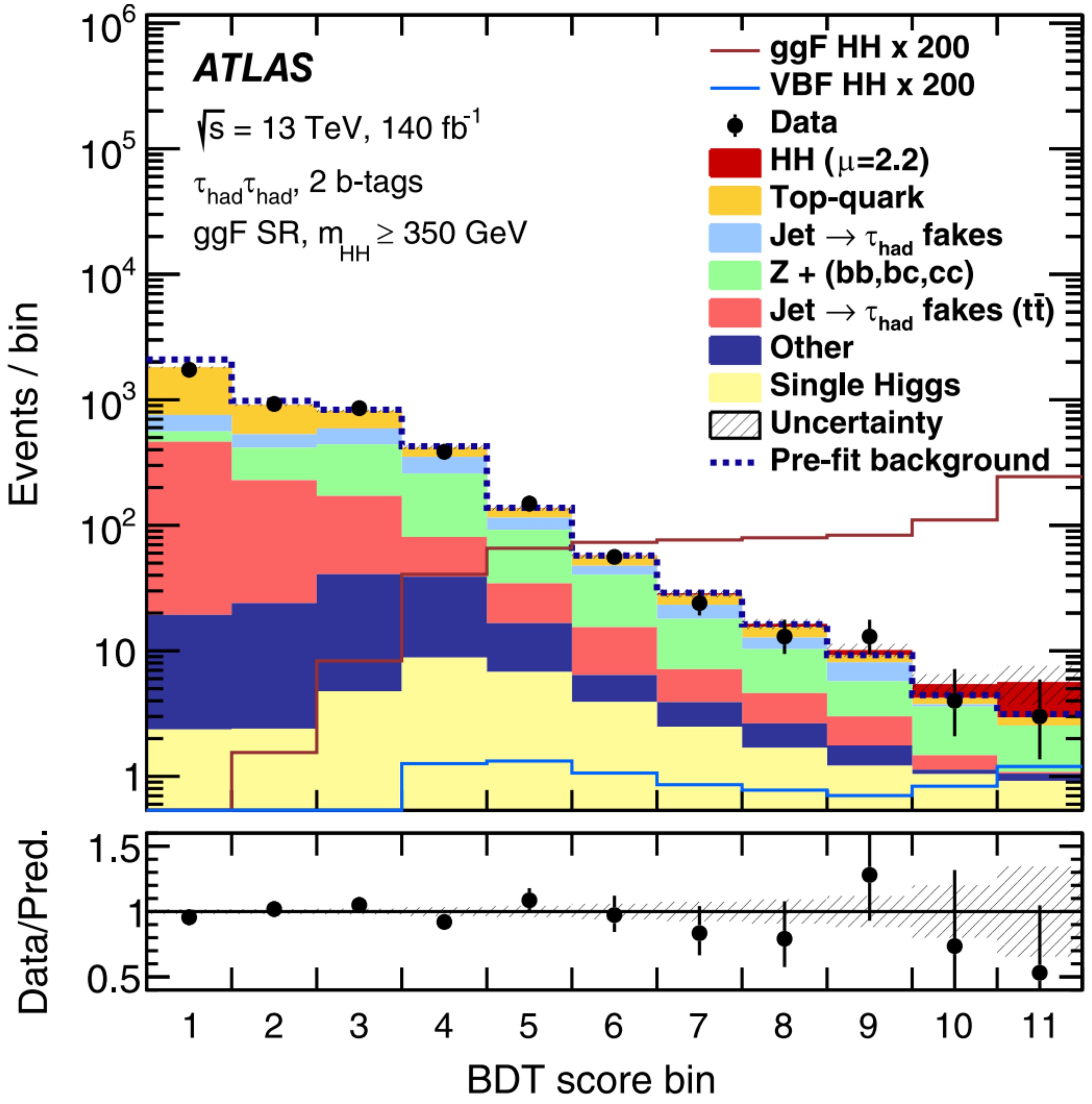
BDT score distributions

→ **hadhad channel**

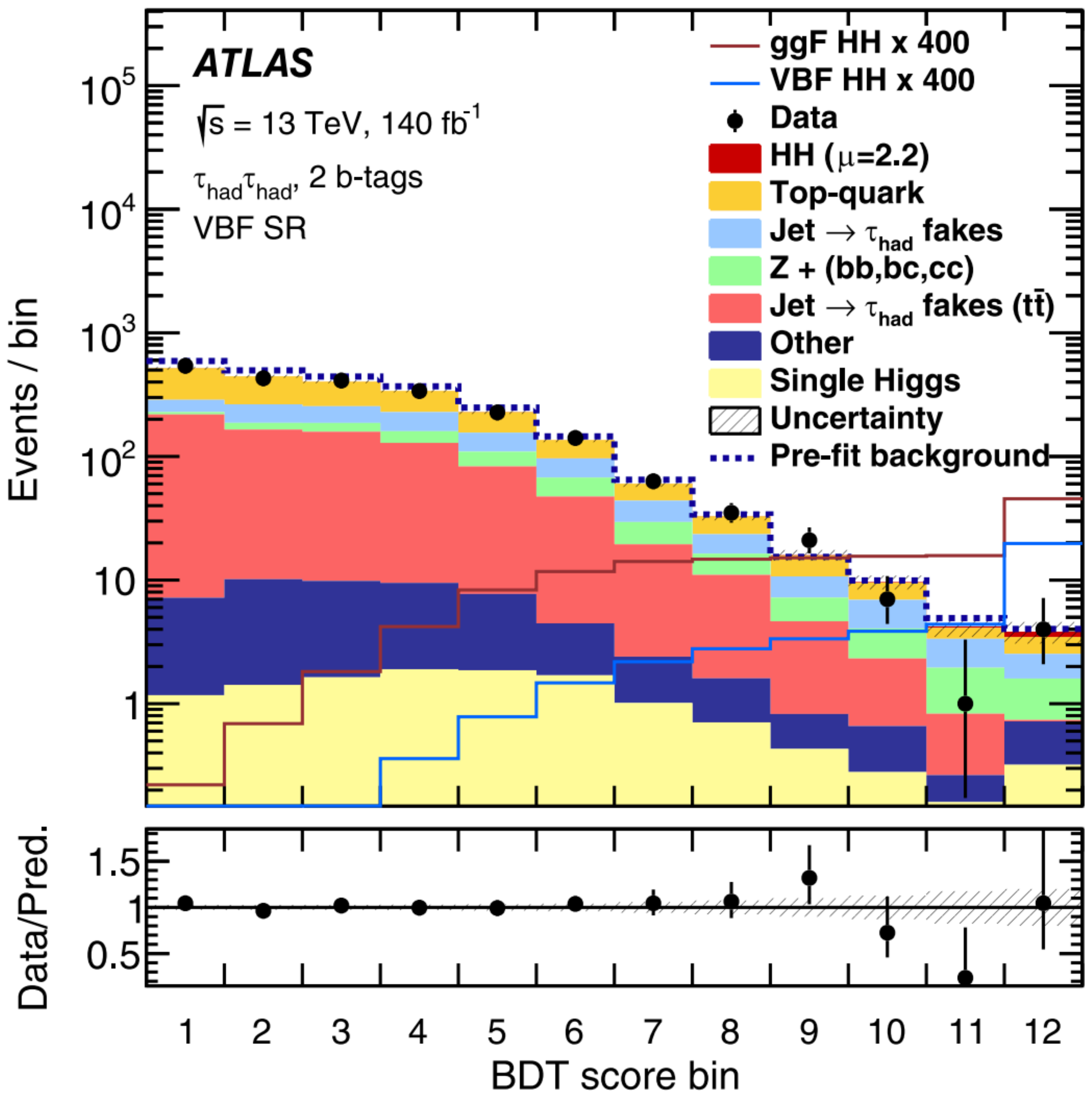
Signal is extracted by a simultaneous fit to all SRs and the CR!



low mHH ggF



high mHH ggF

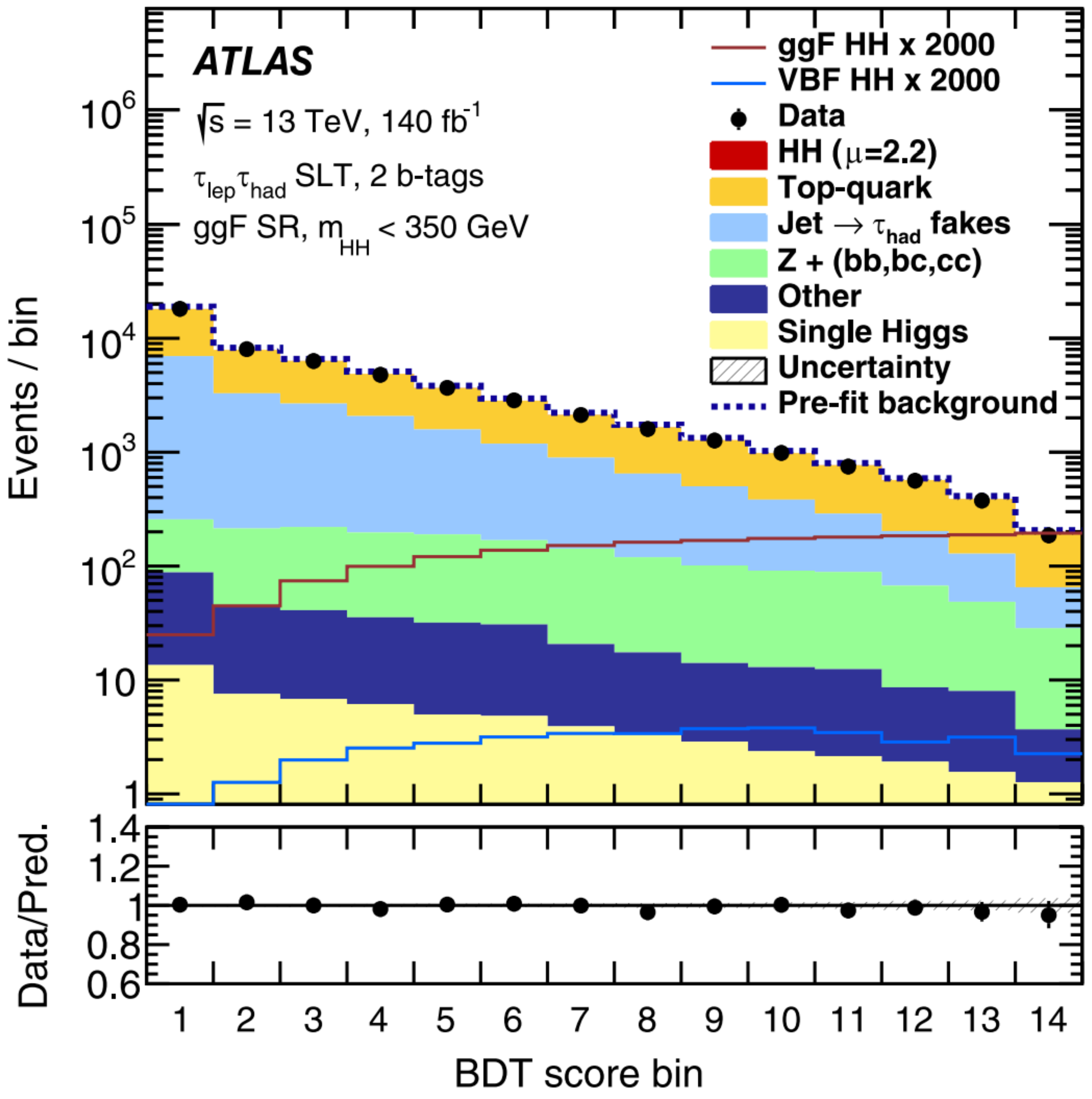


VBF

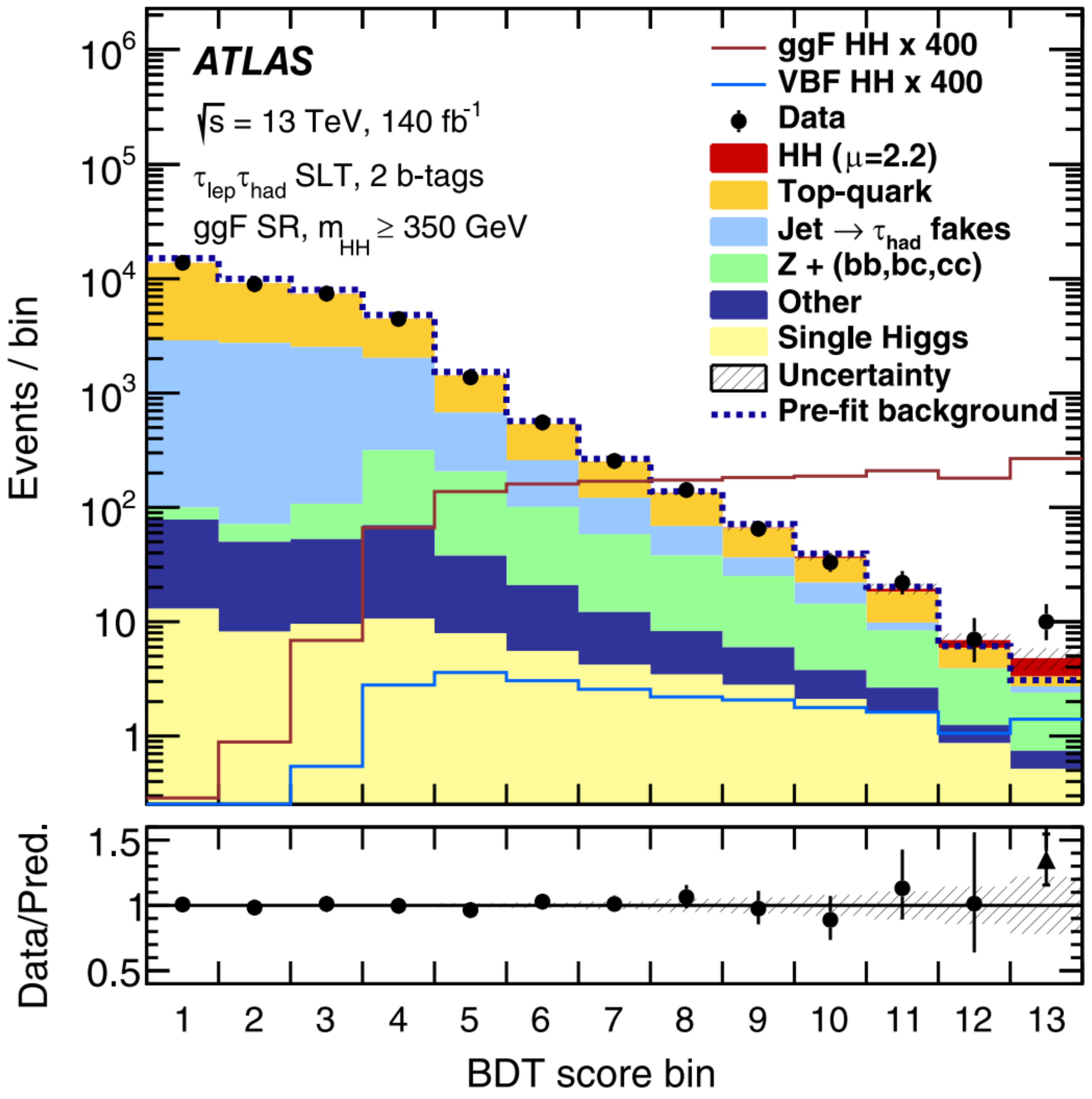
*Binning is determined by algorithms to optimize sensitivity while ensuring valid background stat.

BDT score distributions

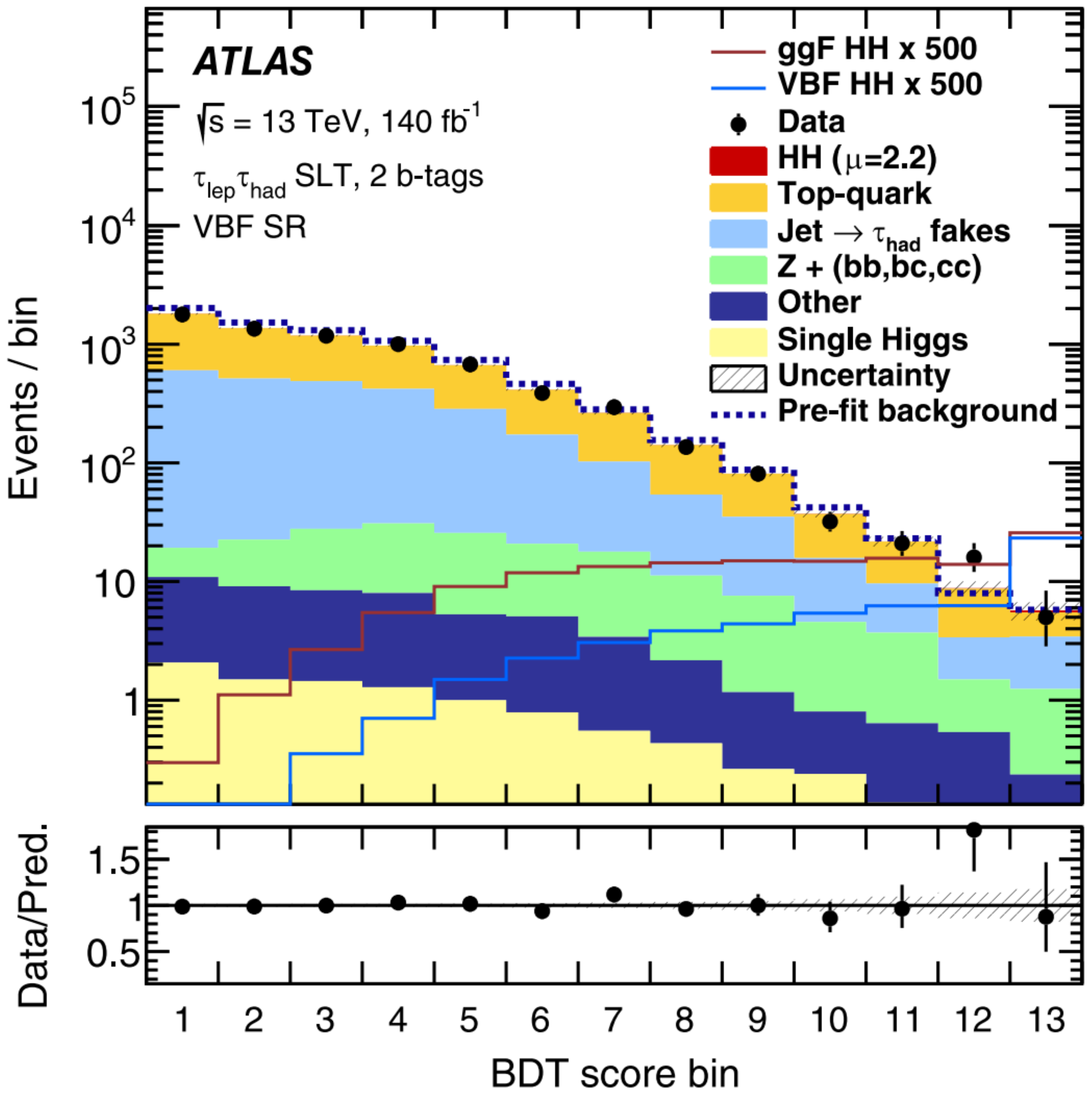
→ lephad SLT channel



low mHH ggF



high mHH ggF

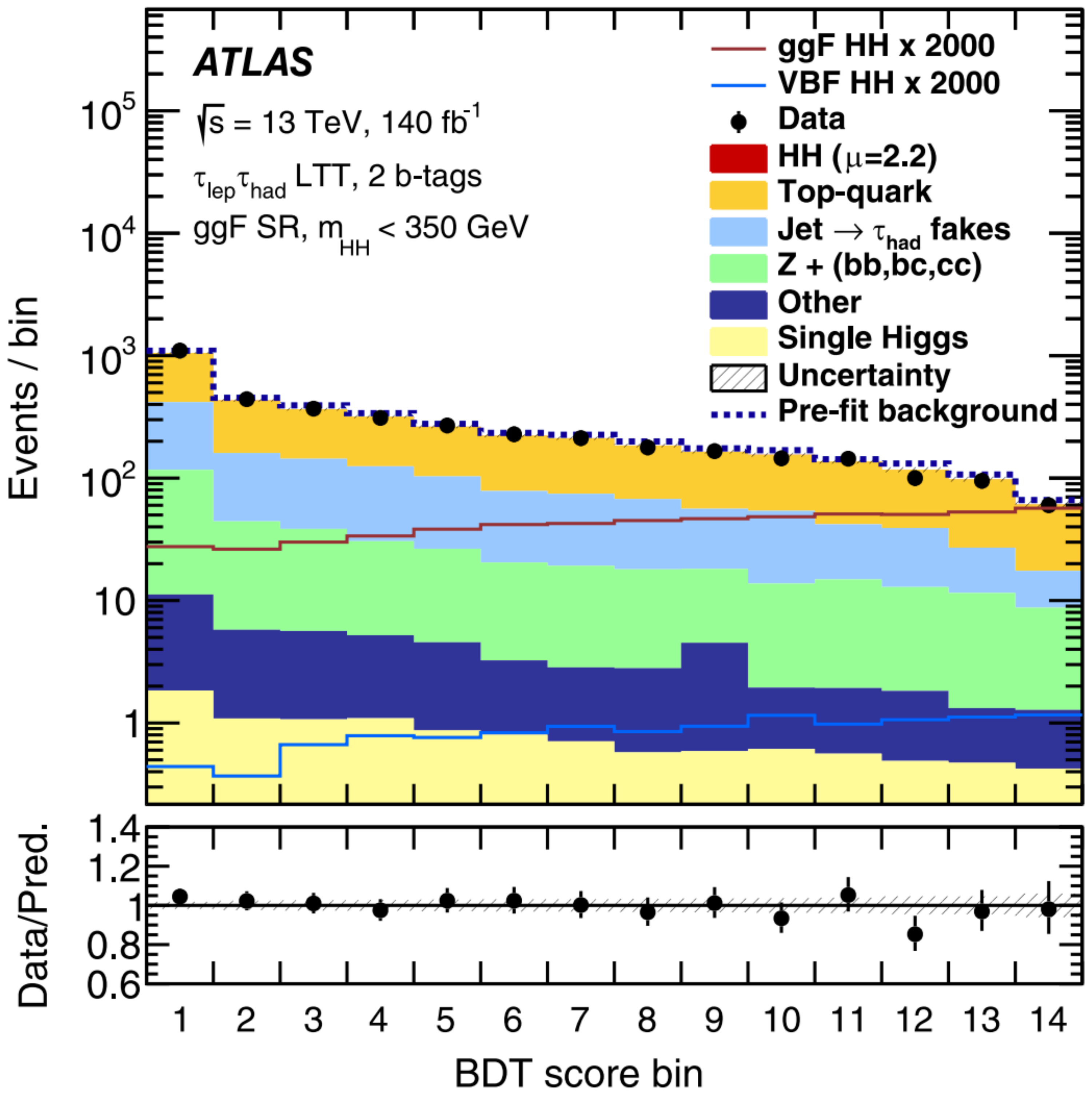


VBF

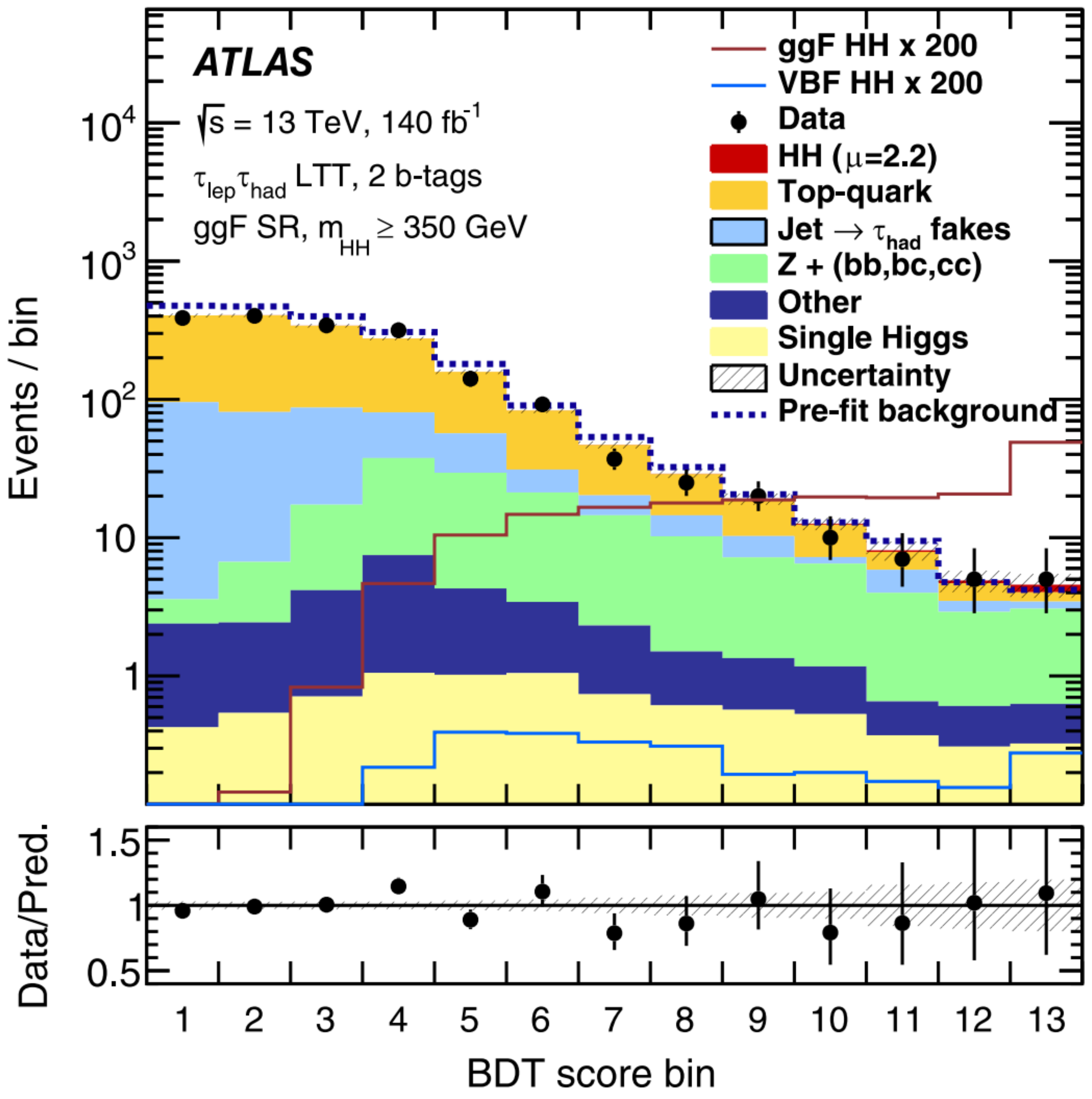
Mild data excess in the last bin of lephad SLT high mHH ggF, statistical fluctuation

BDT score distributions

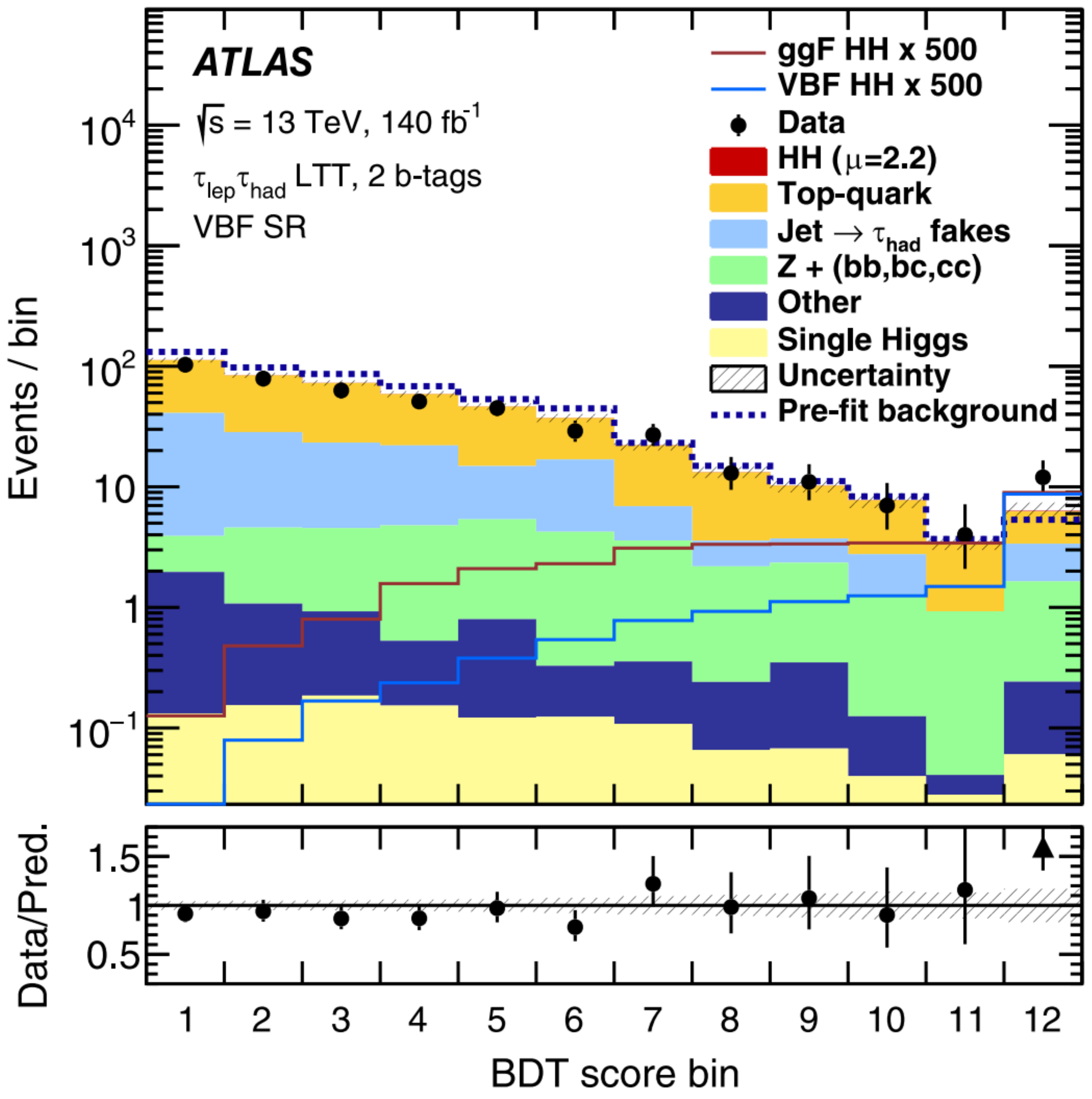
→ lephad LTT channel



low mHH ggF

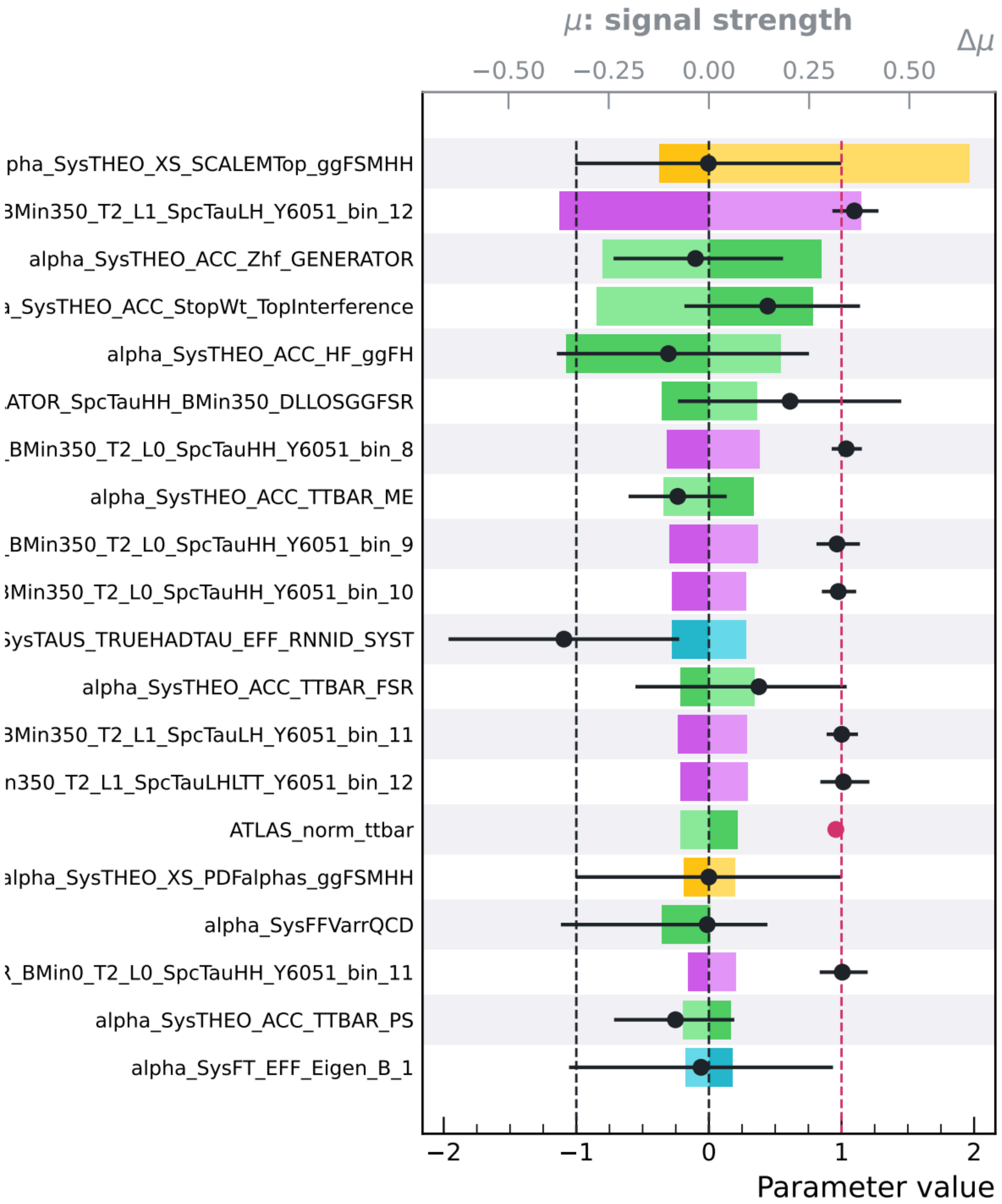


high mHH ggF



VBF

Systematic uncertainty



➔ Leading uncertainty: ggF signal modeling

- Uncertainty in the ggF HH production cross-section arising from variations of the QCD scales and the top-quark mass scheme

➔ Statistical uncertainty of bkg MC samples

➔ Uncertainty related to single-top Wt modeling

Results: HH cross section

→ **No significant excess observed above the expected background**

→ **95% CL upper limit on μ_{HH}**

$\mu_{HH} < 5.9$ observed

$\mu_{HH} < 3.3$ expected, **15% reduction wrt previous analysis**

Observed limit higher than expected due to a statistical fluctuation in the lephad SLT high mHH SR.

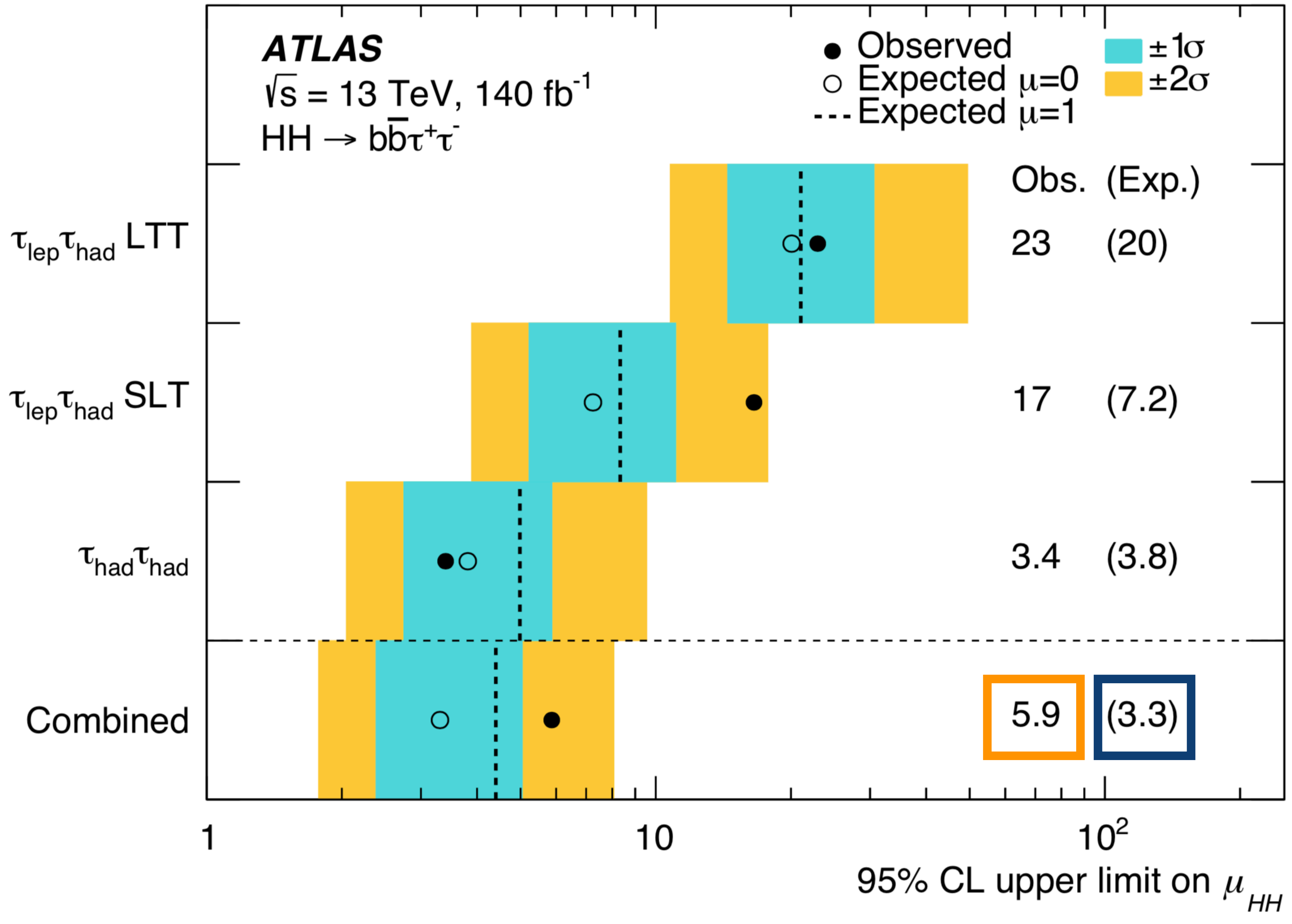
→ **Set 95% CL upper limits simultaneously on ggF and VBF production cross section**

$\mu_{ggF} < 5.8$ observed

$\mu_{ggF} < 3.4$ expected

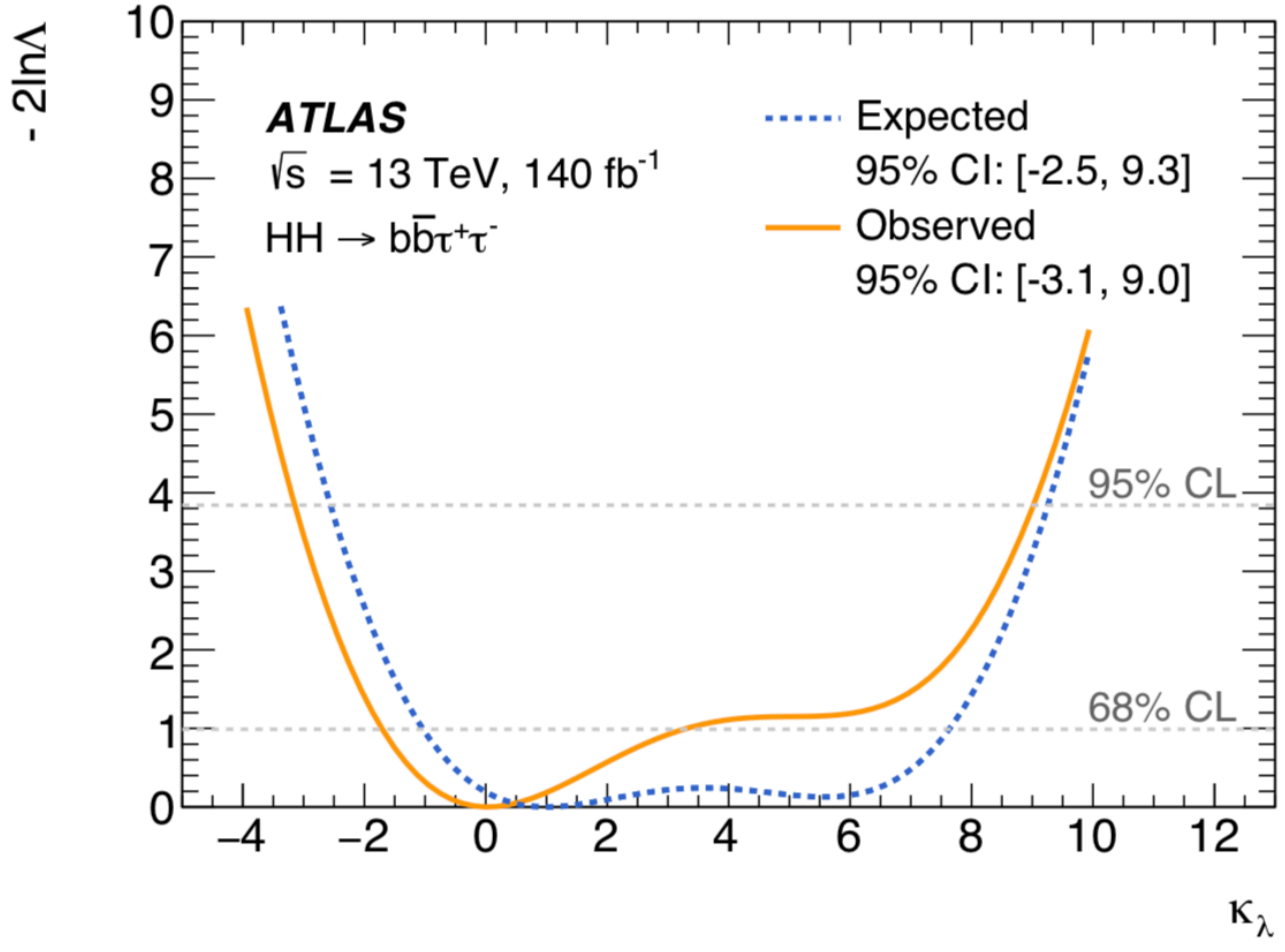
$\mu_{VBF} < 91$ observed

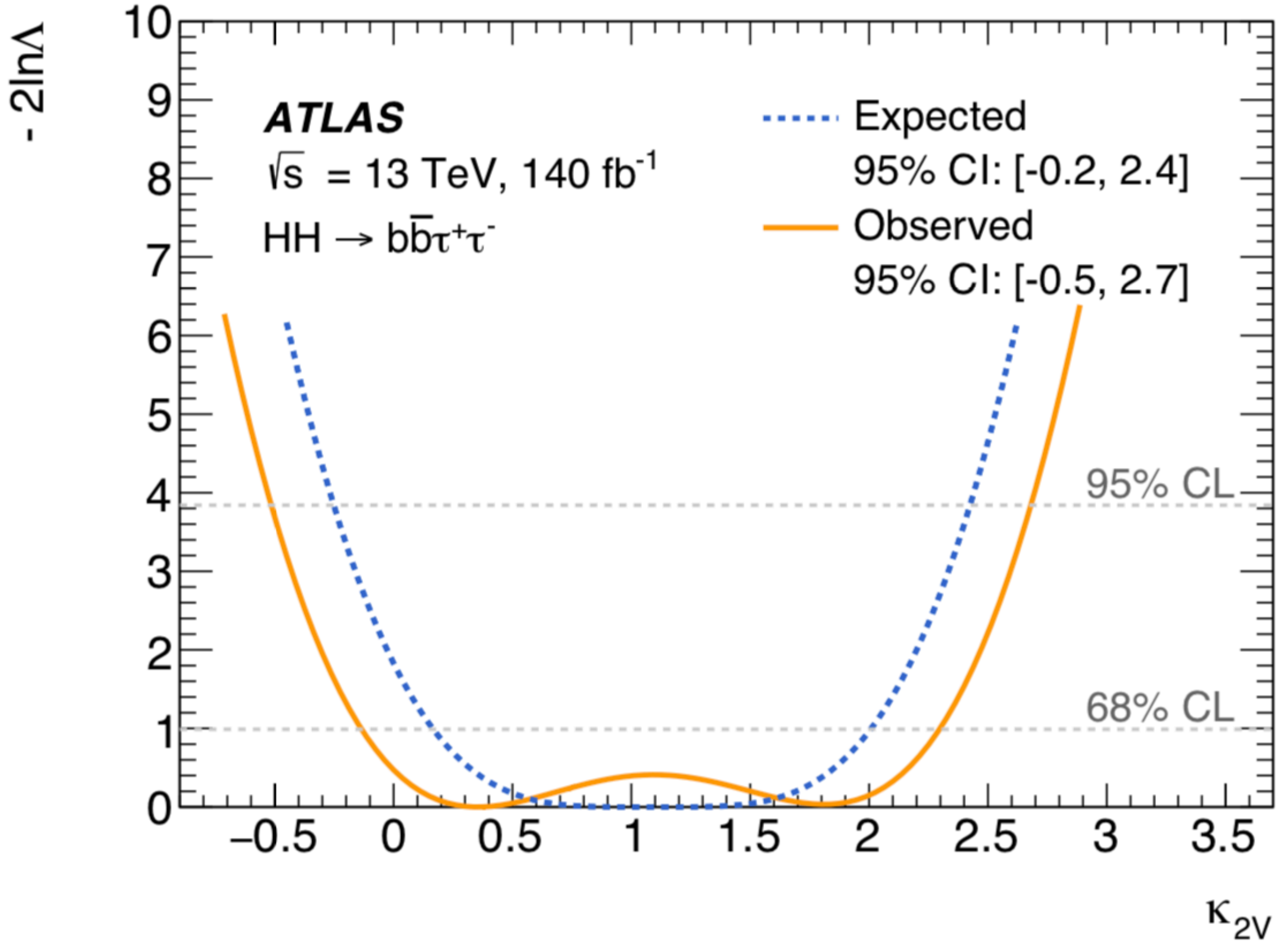
$\mu_{VBF} < 73$ expected



Results: constraint on Higgs self coupling

→ **Constrain the modifier κ_λ and κ_{2V}**





$\kappa_\lambda \in$ [-3.1, 9.0] observed

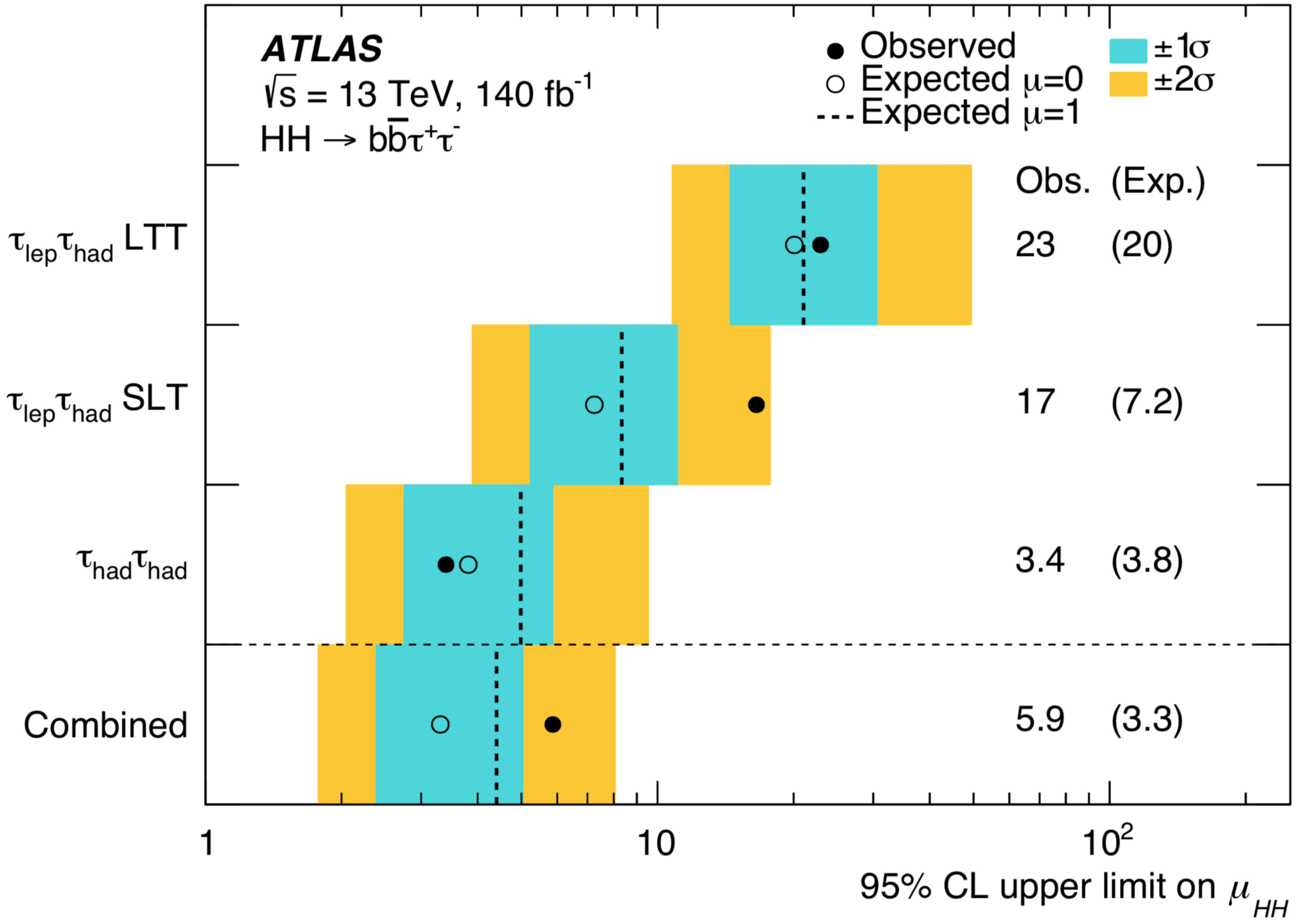
$\kappa_\lambda \in$ [-2.5, 9.3] expected, **11% reduction**

$\kappa_{2V} \in$ [-0.5, 2.7] observed

$\kappa_{2V} \in$ [-0.2, 2.4] expected, **19% reduction**

Summary

- → Overview of the Legacy Run 2 non-resonant HH—>bbtautau analysis
- → No significant excess above the expected background is observed
- → 15% improvement on the expected signal strength
  10% - 20% improvement on the expected κ_λ and κ_{2V} constraint
- → Looking forward to the Run 2 + Run 3 results



Backup

Overview of analysis strategy

→ A sketch depicting the analysis strategy

STT: single $\tau_{had\text{-}vis}$ triggers
DTT: di-$\tau_{had\text{-}vis}$ triggers
SLT: single lepton triggers
LTT: lepton+$\tau_{had\text{-}vis}$ triggers
DLT: di-lepton triggers

hadhad channel **lephad SLT channel** **lephad LTT channel**

Analysis strategy

→ **A sketch depicting the analysis strategy (9 SRs + 1 CR)**

3 channels per di-τ decays

Optimize trigger strategy

$\otimes$

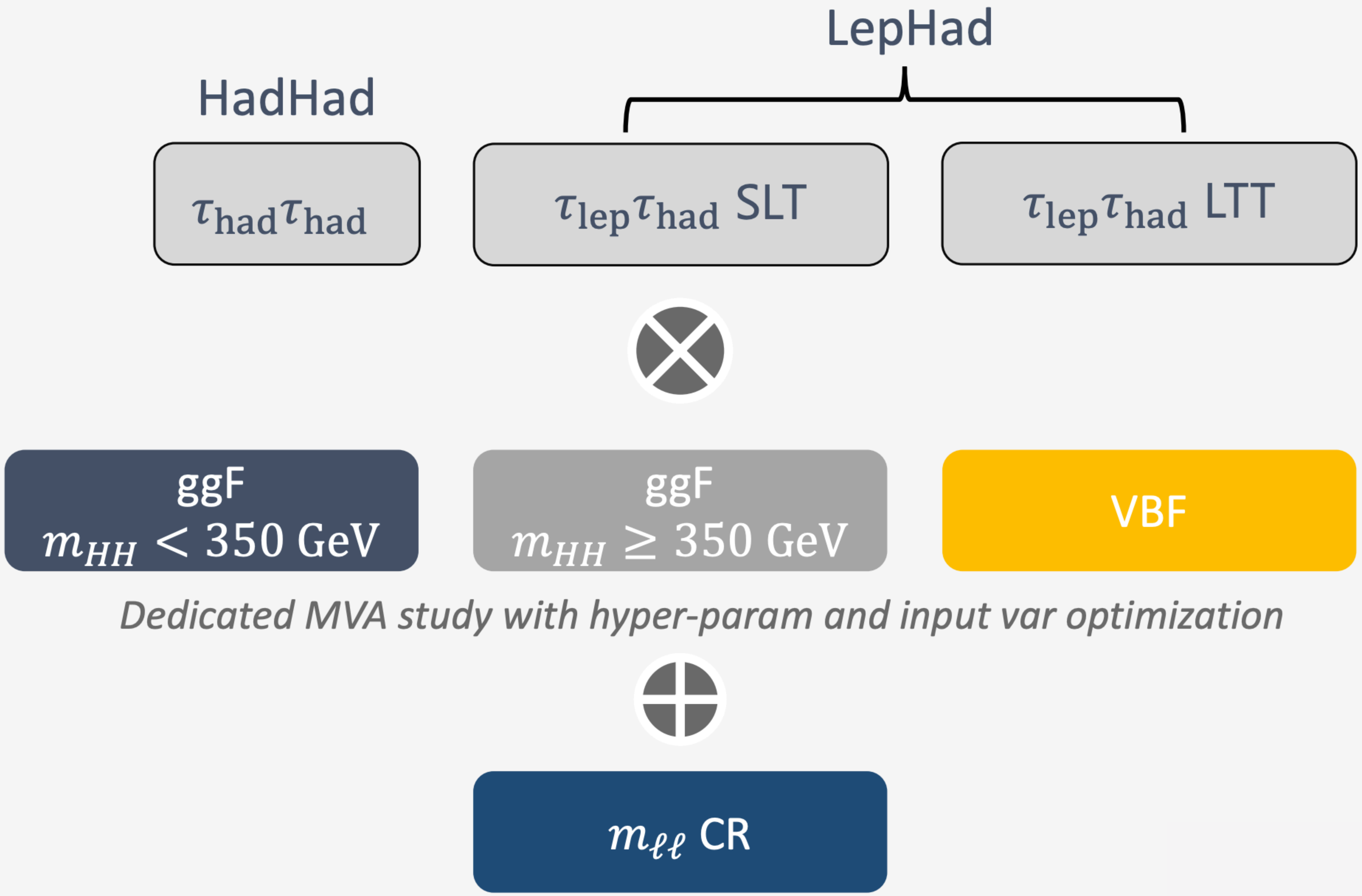
3 signal regions per production mode and m_{HH} split

Improve κ_{2V} constraint

$\oplus$

1 control region

Improve bkg modelling



LepHad

HadHad

$\tau_{\text{had}}\tau_{\text{had}}$

$\tau_{\text{lep}}\tau_{\text{had}}$ SLT

$\tau_{\text{lep}}\tau_{\text{had}}$ LTT

$\otimes$

ggF $m_{HH} < 350$ GeV

ggF $m_{HH} \geq 350$ GeV

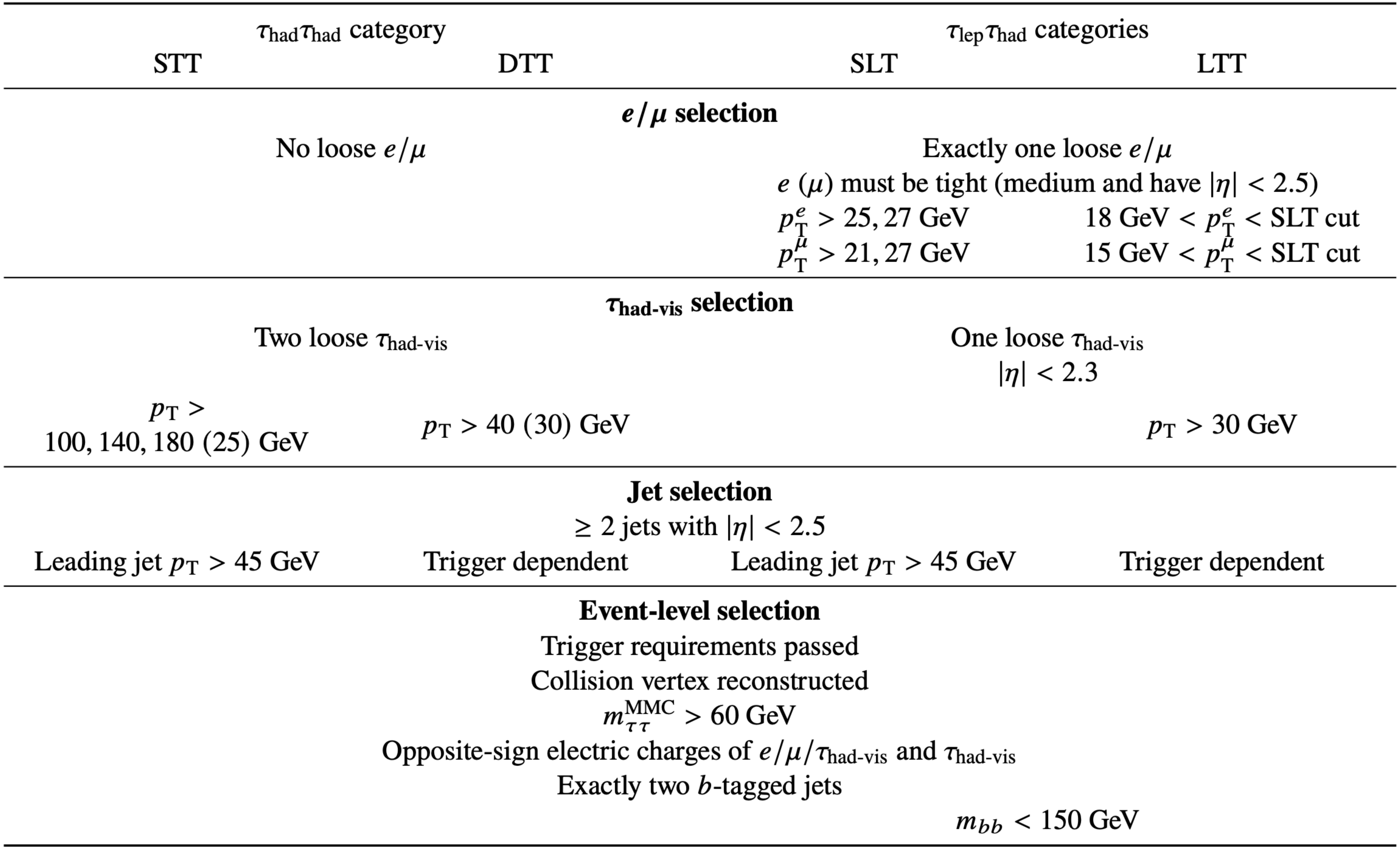
VBF

Dedicated MVA study with hyper-param and input var optimization

$\oplus$

$m_{\ell\ell}$ CR

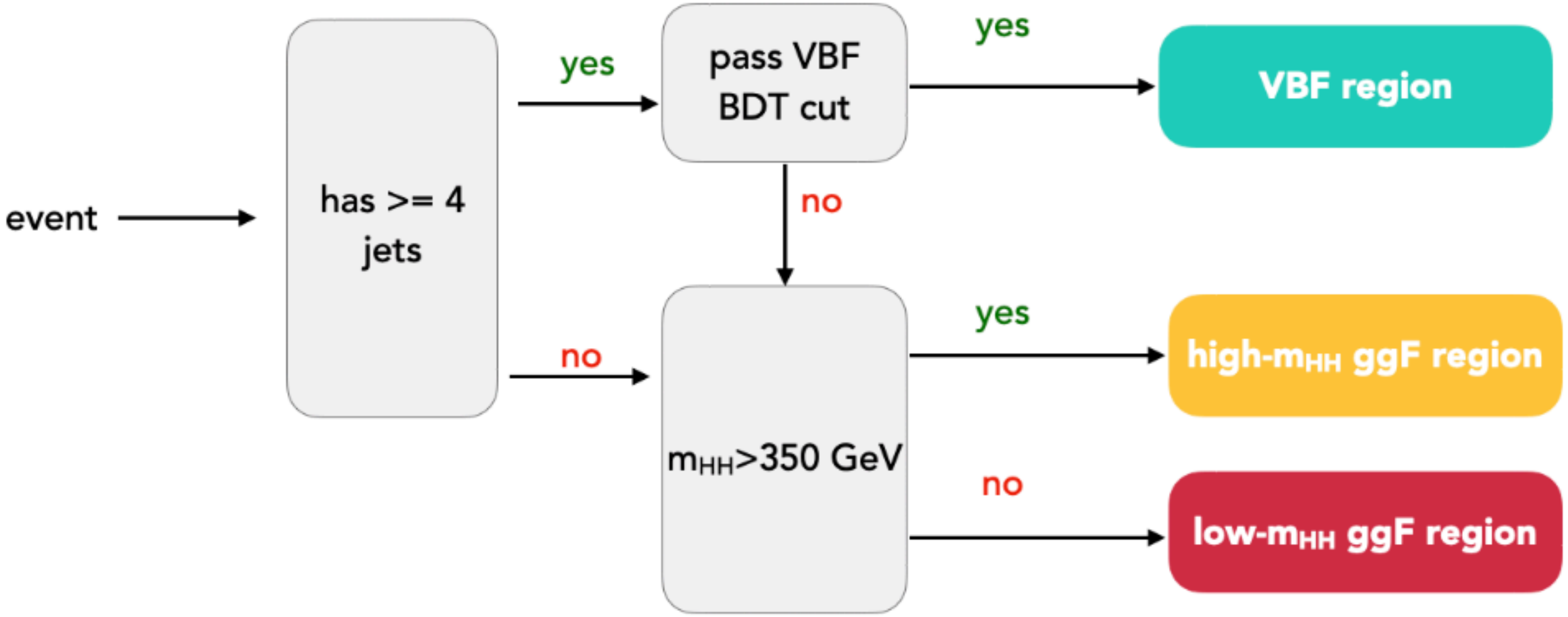
Event selection



$\tau_{\text{had}}\tau_{\text{had}}$ category		$\tau_{\text{lep}}\tau_{\text{had}}$ categories	
STT	DTT	SLT	LTT
e/μ selection			
No loose e/μ		Exactly one loose e/μ	
		e (μ) must be tight (medium and have $\lvert\eta\rvert < 2.5$)	
		$p_{\text{T}}^{e} > 25, 27$ GeV	18 GeV $< p_{\text{T}}^{e} <$ SLT cut
		$p_{\text{T}}^{\mu} > 21, 27$ GeV	15 GeV $< p_{\text{T}}^{\mu} <$ SLT cut
$\tau_{\text{had-vis}}$ selection			
Two loose $\tau_{\text{had-vis}}$		One loose $\tau_{\text{had-vis}}$	
		$\lvert\eta\rvert < 2.3$	
$p_{\text{T}} >$ 100, 140, 180 (25) GeV	$p_{\text{T}} > 40$ (30) GeV		$p_{\text{T}} > 30$ GeV
Jet selection			
≥ 2 jets with $\lvert\eta\rvert < 2.5$			
Leading jet $p_{\text{T}} > 45$ GeV	Trigger dependent	Leading jet $p_{\text{T}} > 45$ GeV	Trigger dependent
Event-level selection			
Trigger requirements passed			
Collision vertex reconstructed			
$m_{\tau\tau}^{\text{MMC}} > 60$ GeV			
Opposite-sign electric charges of $e/\mu/\tau_{\text{had-vis}}$ and $\tau_{\text{had-vis}}$			
Exactly two b-tagged jets			
		$m_{bb} < 150$ GeV	

Signal extraction: Boosted Decision Tree (BDT)

→ **In each sub-channel (hadhad, lephad SLT, lephad LTT), 3 different BDTs trained:**



- Train on SM VBF signal vs bkg
- 3-fold training
- Input variables selection
- Optimized hyperparameters

- Train on SM ggF signal vs bkg
- 3-fold training
- Input variables selection
- Optimized hyperparameters

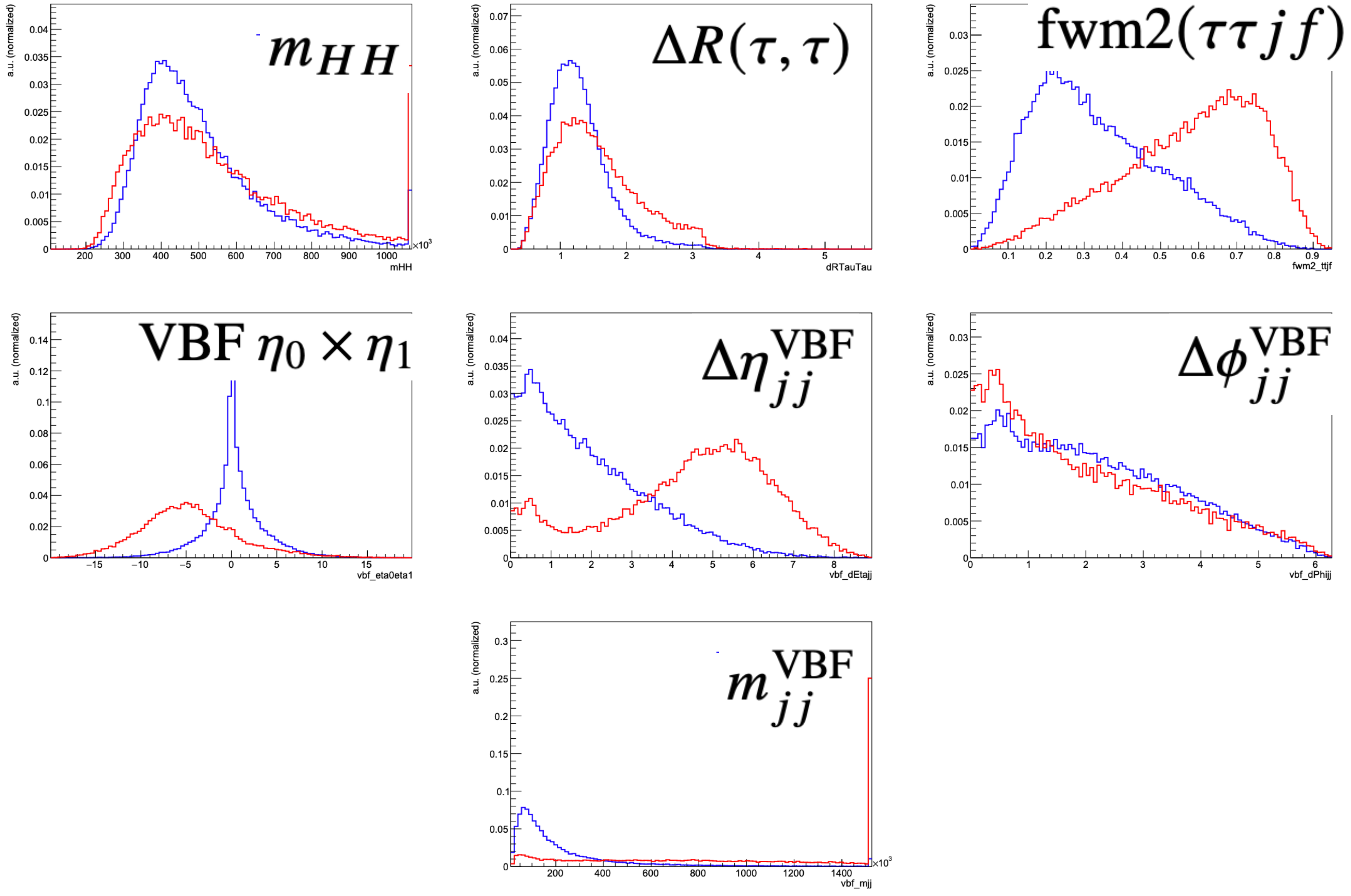
- Train on $\kappa_\lambda = 10$ ggF signal vs bkg
- 3-fold training
- Input variables selection
- Optimized hyperparameters

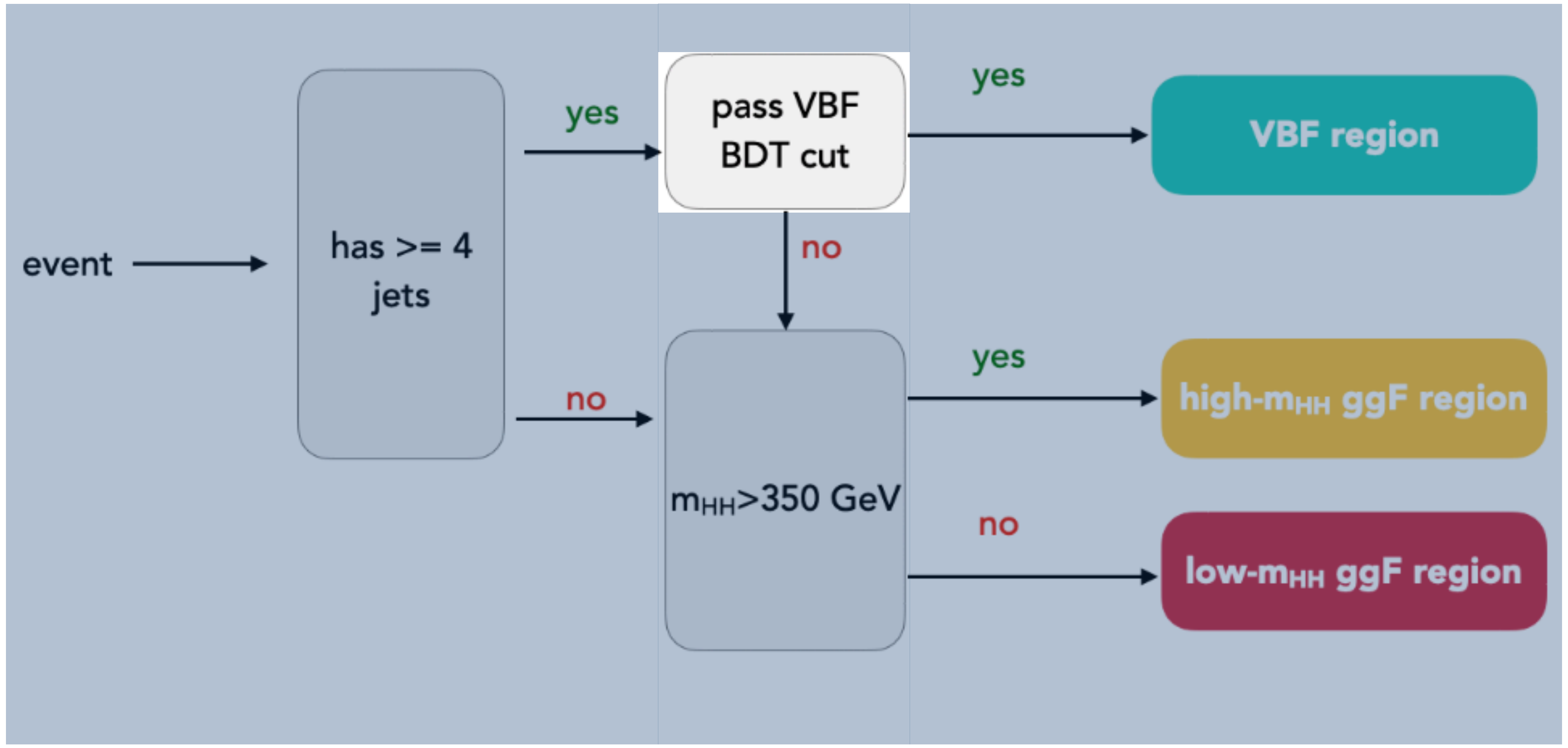
Categorization

→ A dedicated ggF/VBF separation BDT is trained

- Signal vs bkg: SM VBF HH vs SM ggF HH

- Input variables





event

has >= 4 jets

yes

pass VBF BDT cut

yes

VBF region

no

no

m_{HH}>350 GeV

yes

high-m_{HH} ggF region

no

low-m_{HH} ggF region

Variable	Description
m_{HH}	Invariant mass of the HH system, reconstructed from the τ-lepton pair (using the MMC) and b-tagged jet pair
$\Delta R(\tau,\tau)$	The ΔR between the two visible τ decay products
VBF $\eta_0 \times \eta_1$	Product of the pseudorapidities of the leading and sub-leading VBF jets
$\Delta\eta_{jj}^{\mathrm{VBF}}$	The $\Delta\eta$ between the two VBF jets
$\Delta\phi_{jj}^{\mathrm{VBF}}$	The $\Delta\phi$ between the two VBF jets
m_{jj}^{VBF}	Invariant mass of the VBF jet system
$\mathrm{fwm2}(\tau\tau jf)$	2nd order Fox-Wolfram moment, taking into account the τ-lepton pair and central and forward jets

Signal extraction: Boosted Decision Tree (BDT)

→ 3-fold training

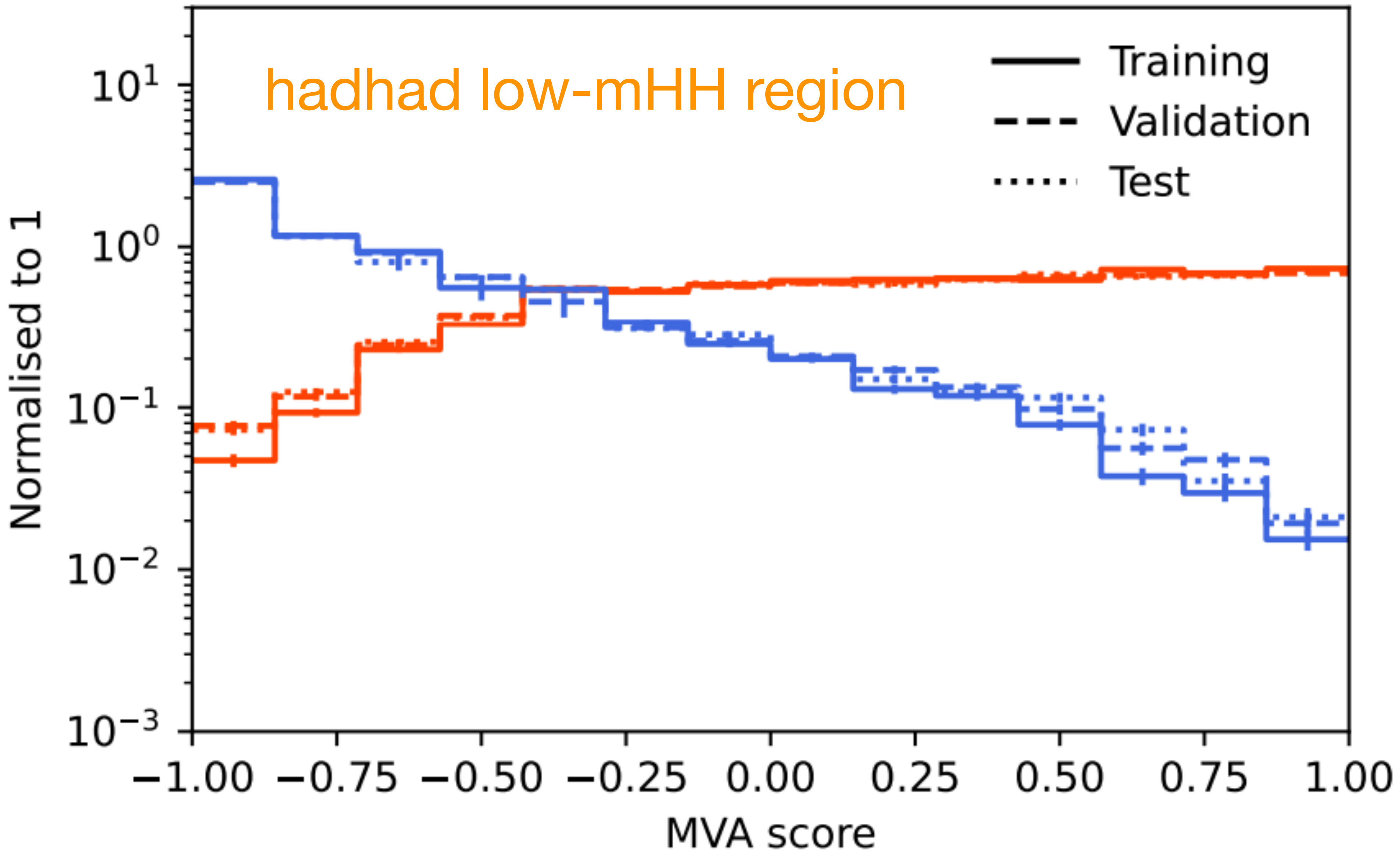
- Divide events into 3 folds based on the event number
- Train 3 BDTs on each fold, and optimized and applied on other folds

Model	Fold 0 `event_number` %3 = 0	Fold 1 (`event_number` %3 = 1)	Fold 2 (`event_number` %3 = 2)
BDT 0	Training	Validation	Testing
BDT 1	Testing	Training	Validation
BDT 2	Validation	Testing	Training

Model	Even-fold	Odd-fold
BDT 0	4/5 for training 1/5 for validation	Testing
BDT 1	Testing	4/5 for training 1/5 for validation

All events are used in training, validation and test :)

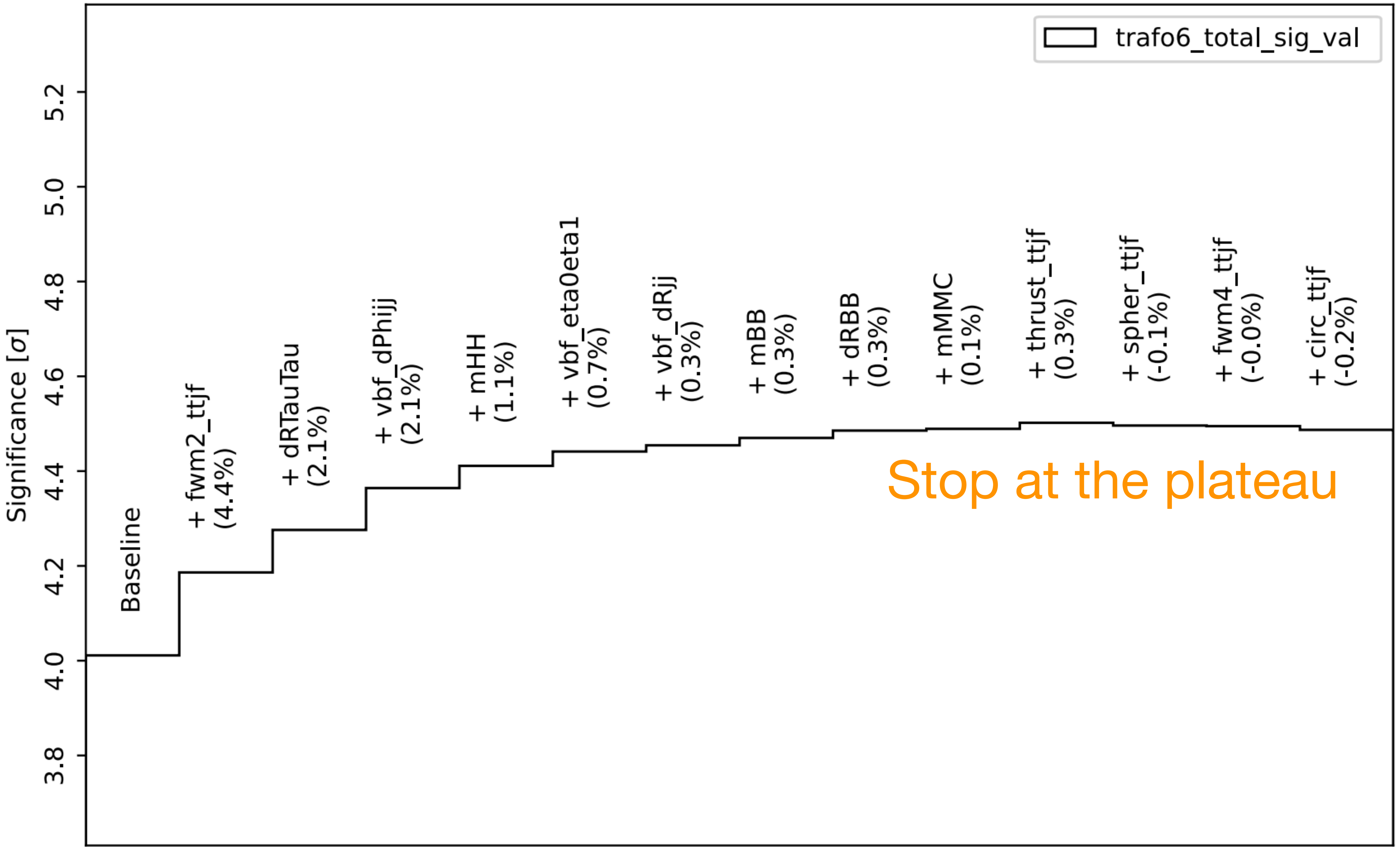
Previous round of the analysis
5-fold cross validation
Only 80% of events used for training



Signal extraction: Boosted Decision Tree (BDT)

→ Input variables selection

- Firstly choose a few variables as the baseline variables
- Gradually add one more variable with the most improvement to the sensitivity
- Until reach a plateau where the sensitivity doesn't increase any more



Stop at the plateau

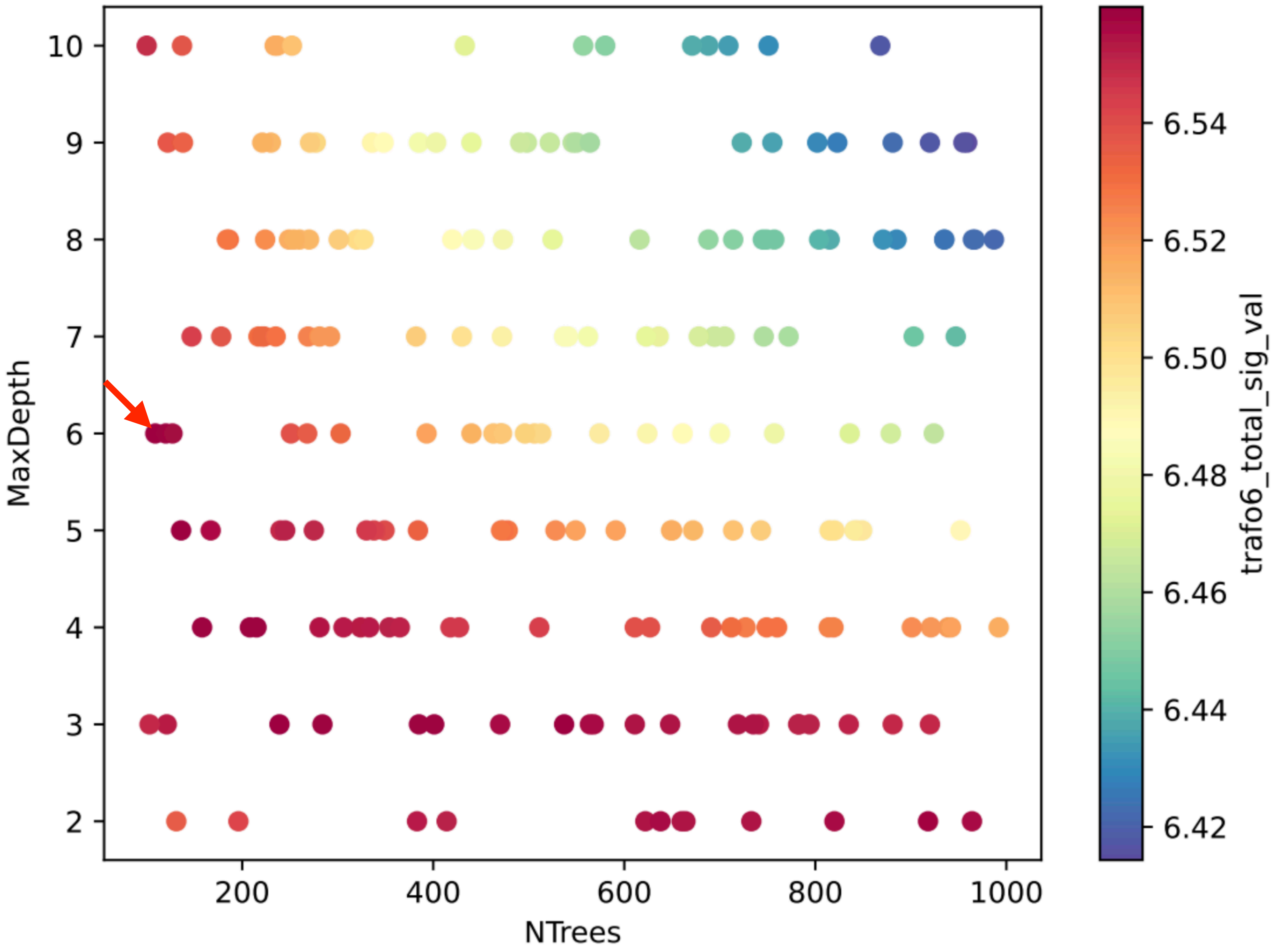
Signal extraction: Boosted Decision Tree (BDT)

→ Hyperparameter optimization (on validation folds)

- Scan the two most important parameters: NTrees and MaxDepth
- The binned signal significance as the figure of merit

$$Z = \sqrt{\sum_{i\in \mathrm{bins}} 2\left((s_i + b_i)\log\left(1 + \frac{s_i}{b_i}\right) - s_i\right)}$$

- Take the set of hyper parameters that gives best significance



Other hyper parameters are set to default values since they are found to have a subheading effect to training

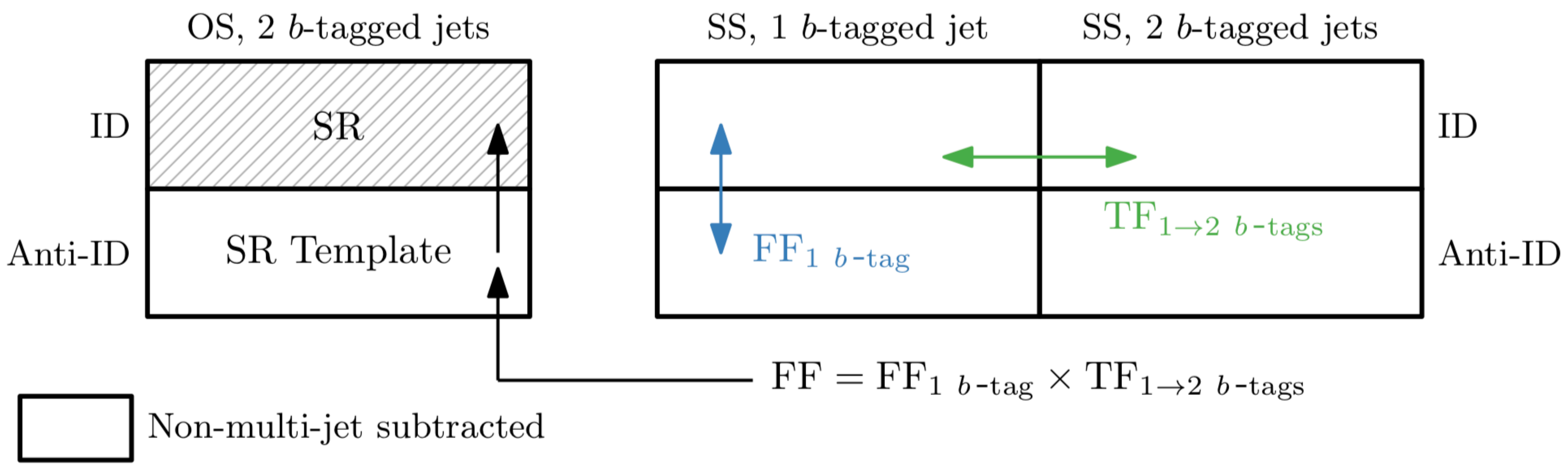
MinNodeSize	1%
BoostType	Grad
Shrinkage	0.2
IgnoreNegWeightsInTraining	True

Background estimation — Fake tau-had in hadhad channel

→ Two sources: Multi-jet process and ttbar process

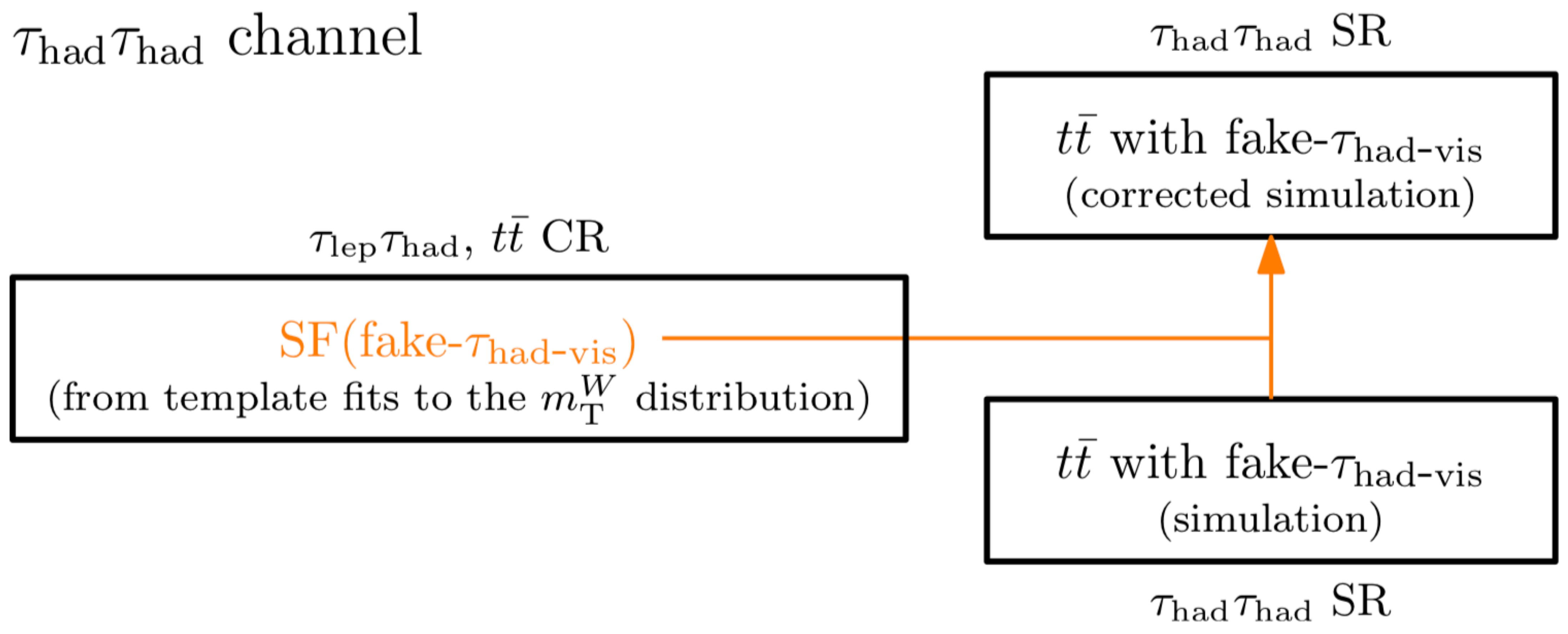
→ Fake tau-had from multi-jet: FF method

- FFs are derived in 1 b-tag SS control region
- Extrapolated to 2 b-tag SS control region by a transfer factors (TFs)

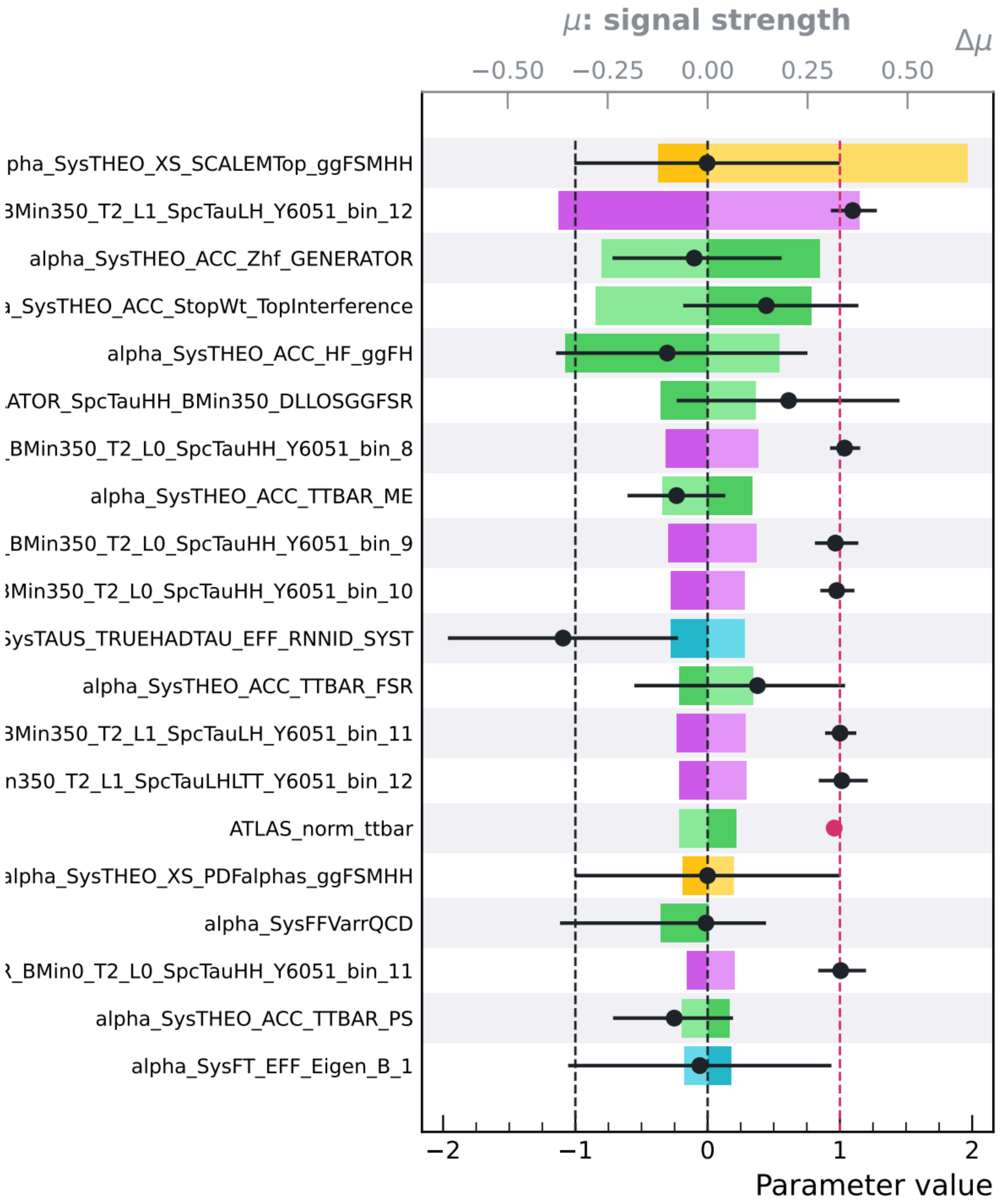


→ Fake tau-had from ttbar: Fake scale factors

- Fake tau-had Scale factors (SF)
- Measured in the lephad ttbar CR by fitting m_T^W to data
- Applied to simulated fake-tau ttbar in SR



Ranking plot



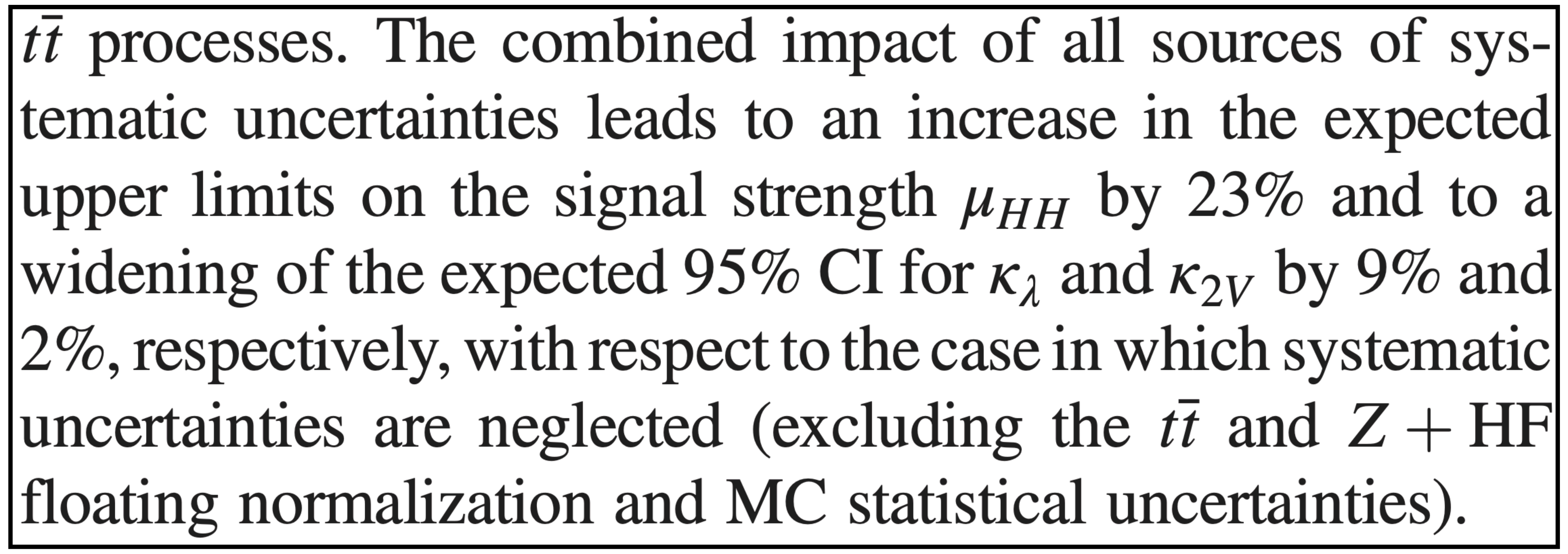
→ Leading uncertainty: ggF signal modeling

- Uncertainty in the ggF HH production cross-section arising from variations of the QCD scales and the top-quark mass scheme

→ Statistical uncertainty of bkg MC samples

→ Uncertainty related to single-top Wt modeling

→ Impact of uncertainties

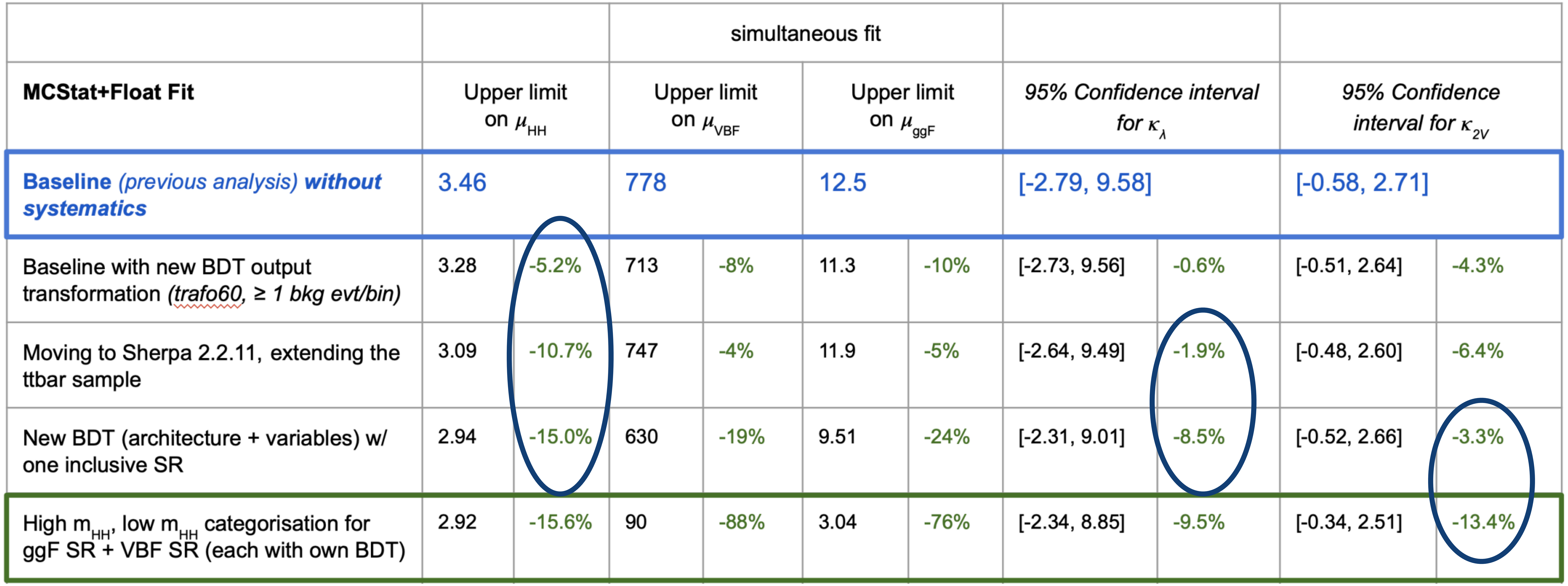


$t\bar{t}$ processes. The combined impact of all sources of systematic uncertainties leads to an increase in the expected upper limits on the signal strength μ_{HH} by 23% and to a widening of the expected 95% CI for κ_λ and κ_{2V} by 9% and 2%, respectively, with respect to the case in which systematic uncertainties are neglected (excluding the $t\bar{t}$ and $Z+\text{HF}$ floating normalization and MC statistical uncertainties).

Breakdown of the improvements

➔ Hadhad channel

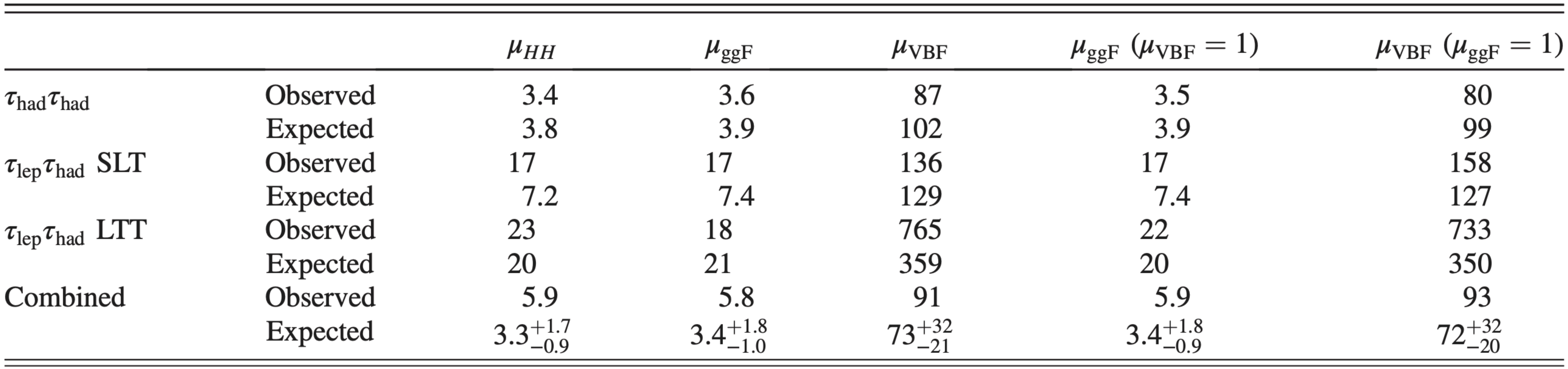
- For upper limit on HH signal strength, the improvement equally comes from new BDT binning, the usage of improved samples, the new optimized BDT
- For κ_λ interval, the new optimized BDT brings largest relative improvement
- For $\kappa 2v$ interval, the introduction of a dedicated VBF SR brings largest relative improvement



			simultaneous fit							
MCStat+Float Fit	Upper limit on μ_{HH}		Upper limit on μ_{VBF}		Upper limit on μ_{ggF}		*95% Confidence interval for κ_λ*		*95% Confidence interval for κ_{2V}*	
Baseline *(previous analysis)* ***without systematics***	3.46		778		12.5		[-2.79, 9.58]		[-0.58, 2.71]	
Baseline with new BDT output transformation *(trafo60, ≥ 1 bkg evt/bin)*	3.28	-5.2%	713	-8%	11.3	-10%	[-2.73, 9.56]	-0.6%	[-0.51, 2.64]	-4.3%
Moving to Sherpa 2.2.11, extending the ttbar sample	3.09	-10.7%	747	-4%	11.9	-5%	[-2.64, 9.49]	-1.9%	[-0.48, 2.60]	-6.4%
New BDT (architecture + variables) w/ one inclusive SR	2.94	-15.0%	630	-19%	9.51	-24%	[-2.31, 9.01]	-8.5%	[-0.52, 2.66]	-3.3%
High m_{HH}, low m_{HH} categorisation for ggF SR + VBF SR (each with own BDT)	2.92	-15.6%	90	-88%	3.04	-76%	[-2.34, 8.85]	-9.5%	[-0.34, 2.51]	-13.4%

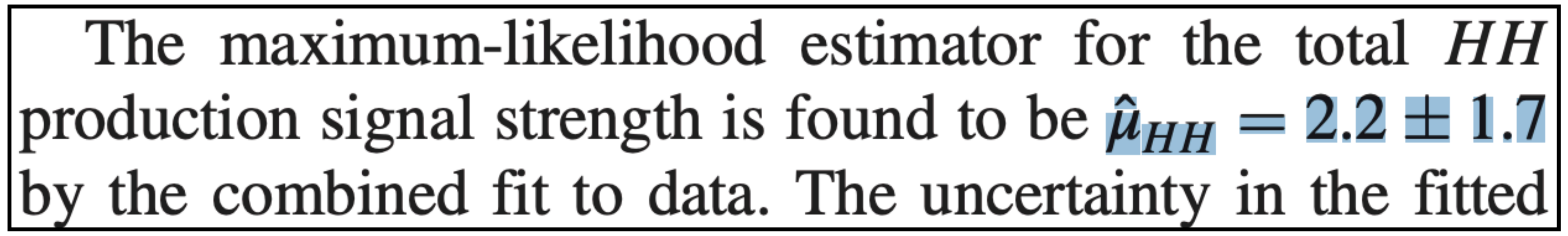
Other results

→ Signal strength upper limits



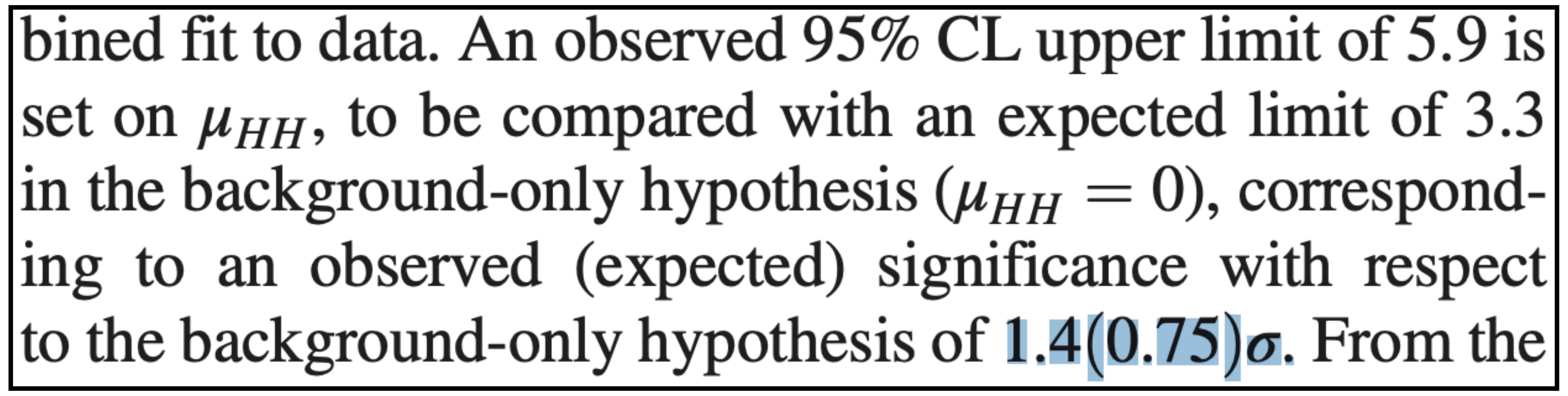
		μ_{HH}	μ_{ggF}	μ_{VBF}	μ_{ggF} ($\mu_{\mathrm{VBF}} = 1$)	μ_{VBF} ($\mu_{\mathrm{ggF}} = 1$)
$\tau_{\mathrm{had}}\tau_{\mathrm{had}}$	Observed	3.4	3.6	87	3.5	80
	Expected	3.8	3.9	102	3.9	99
$\tau_{\mathrm{lep}}\tau_{\mathrm{had}}$ SLT	Observed	17	17	136	17	158
	Expected	7.2	7.4	129	7.4	127
$\tau_{\mathrm{lep}}\tau_{\mathrm{had}}$ LTT	Observed	23	18	765	22	733
	Expected	20	21	359	20	350
Combined	Observed	5.9	5.8	91	5.9	93
	Expected	$3.3^{+1.7}_{-0.9}$	$3.4^{+1.8}_{-1.0}$	73^{+32}_{-21}	$3.4^{+1.8}_{-0.9}$	72^{+32}_{-20}

→ Signal strength



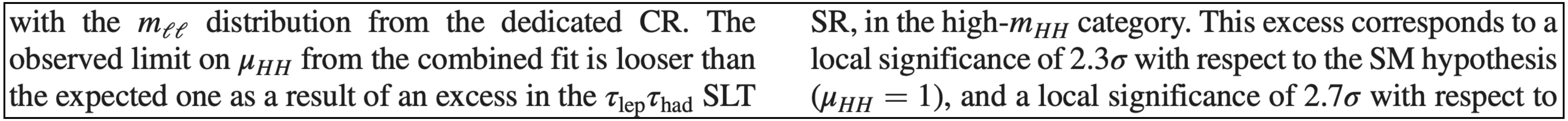
The maximum-likelihood estimator for the total *HH* production signal strength is found to be $\hat{\mu}_{HH} = 2.2 \pm 1.7$ by the combined fit to data. The uncertainty in the fitted

→ Significance



bined fit to data. An observed 95% CL upper limit of 5.9 is set on μ_{HH}, to be compared with an expected limit of 3.3 in the background-only hypothesis ($\mu_{HH} = 0$), corresponding to an observed (expected) significance with respect to the background-only hypothesis of $1.4(0.75)\sigma$. From the

→ Local significance



with the $m_{\ell\ell}$ distribution from the dedicated CR. The observed limit on μ_{HH} from the combined fit is looser than the expected one as a result of an excess in the $\tau_{\mathrm{lep}}\tau_{\mathrm{had}}$ SLT SR, in the high-m_{HH} category. This excess corresponds to a local significance of 2.3σ with respect to the SM hypothesis ($\mu_{HH} = 1$), and a local significance of 2.7σ with respect to

Other results

→ **Two-dimensional contours of κ_λ and κ_{2V}**