UPPSALA UNIVERSITET

GEORG-AUGUST-UNIVERSITÄT GÖTTINGEN

Search for pairs of Higgs bosons in the $b\bar{b}\tau^+\tau^-$ decay channel with the ATLAS detector

Petar Bokan
Uppsala University, University of Göttingen
supervisors: Arnaud Ferrari, Stan Lai

Half-time PhD seminar in Uppsala - October 10, 2018

Higgs potential

- Important to measure the shape of the Higgs potential

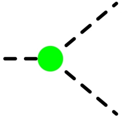
$$V(\phi) = -\frac{1}{2}\mu^2\phi^2 + \frac{1}{4}\lambda\phi^4$$

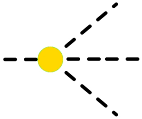
Expanding about minimum: $V(\phi) \to V(v+h)$

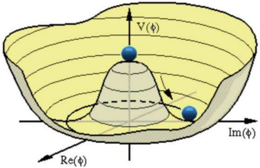
$$V = V_0 + \lambda v^2 h^2 + \lambda v h^3 + \frac{1}{4}\lambda h^4 + \ldots$$

$$= V_0 + \frac{1}{2}m_h^2 h^2 + \frac{m_h^2}{2v^2} v h^3 + \frac{1}{4}\frac{m_h^2}{2v^2} h^4 + \ldots$$

mass term — hh-production — hhh-production







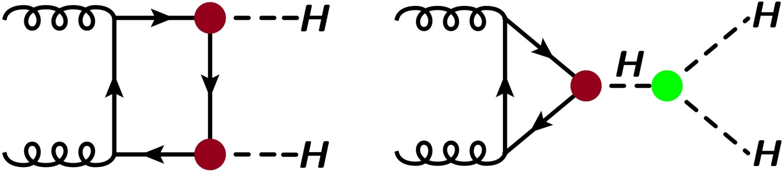
Standard Model (SM):

$$v = \frac{\mu}{\sqrt{\lambda}} = 246\,\text{GeV}$$

$$\lambda = \frac{m_h^2}{2v^2} \approx 0.13$$

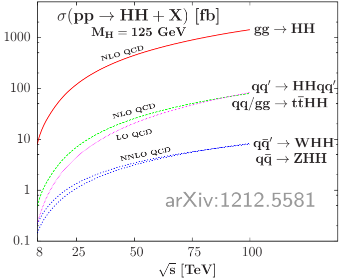
Higgs boson pair production at the LHC

- SM Higgs boson pair production (gluon-gluon fusion - ggF):



Higgs-fermion Yukawa coupling

Higgs boson self-coupling



Small production cross-section:

$\sigma_{\mathrm{SM}}^{\mathrm{ggF}} = 33.41$ fb at $\sqrt{s} = 13$ TeV

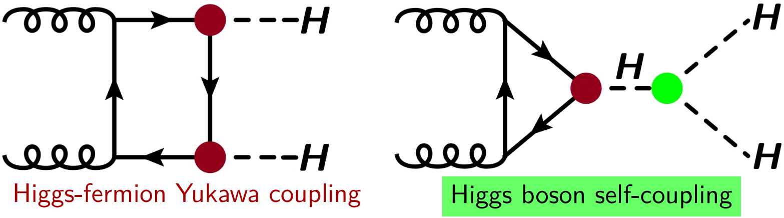
– two massive final state particles

– destructive interference

Phys. Rev. Lett. 117 (2016) 012001
10.23731/CYRM-2017-002 LHCHXSWGHH

Higgs boson pair production at the LHC

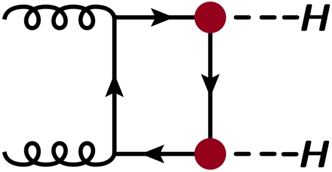
- SM Higgs boson pair production (gluon-gluon fusion - ggF):



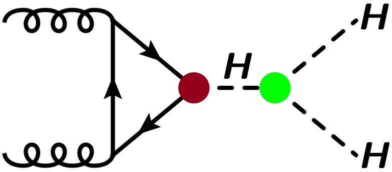
- Potential non-resonant BSM enhancements
  (new couplings, modified Yukawa and/or self-couplings)

Higgs boson pair production at the LHC

- SM Higgs boson pair production (gluon-gluon fusion - ggF):



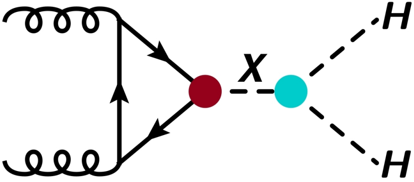
Higgs-fermion Yukawa coupling



Higgs boson self-coupling

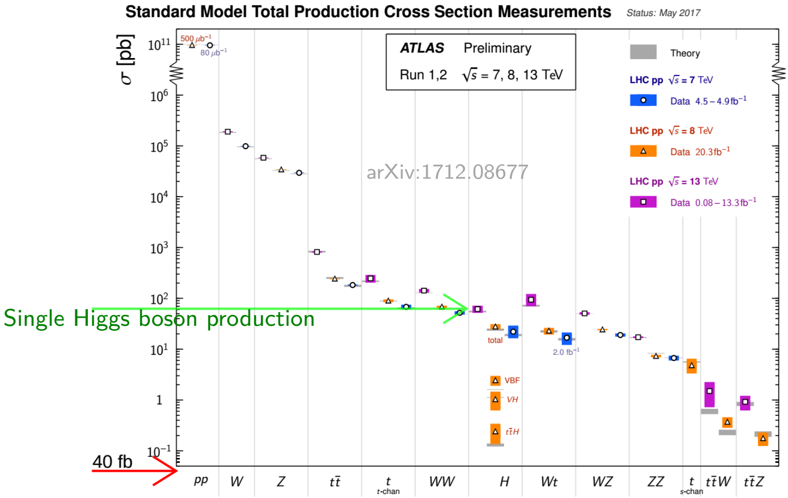
- Potential non-resonant BSM enhancements
  (new couplings, modified Yukawa and/or self-couplings)

- Benchmark BSM resonance hypotheses:
  - Randall-Sundrum graviton $G \to HH$ (spin=2)
  - $S \to HH$ (spin=0)



Resonant production

SM Higgs boson pair production at the LHC



- SM HH-production $\sim 1000\times$ smaller compared to H-production
- Current LHC dataset won't be large enough to reach the sensitivity

Di-Higgs final states

Di-Higgs decay modes and relative branching fractions:

	bb	WW	ττ	ZZ	γγ
bb	33%				
WW	25%	4.6%			
ττ	7.4%	2.5%	0.39%		
ZZ	3.1%	1.2%	0.34%	0.076%	
γγ	0.26%	0.10%	0.029%	0.013%	0.0005%

10.23731/CYRM-2017-002

Some of the most sensitive channels:

$HH \to b\bar{b}b\bar{b}$: the highest BR, large multijet background

$HH \to b\bar{b}\tau^+\tau^-$: relatively large BR, cleaner final state

$HH \to b\bar{b}\gamma\gamma$: small BR, clean signal extraction thanks to a good $\gamma\gamma$ mass resolution

No golden channel! Important to consider a large number of final states!

Di-Higgs final states

Di-H
bran

bb
WW
ττ
ZZ
γγ

e

anks

Some of the relevant Run-2 results:

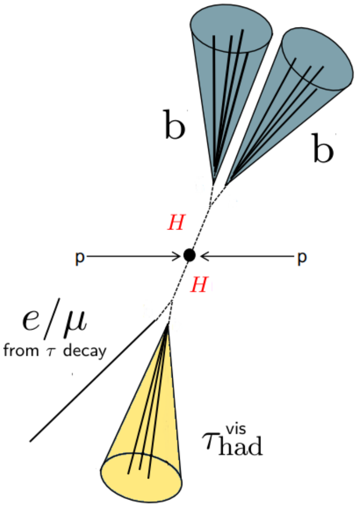
ATLAS $b\bar{b}b\bar{b}$: arXiv:1804.06174
ATLAS $b\bar{b}\tau^+\tau^-$: arXiv:1808.00336
ATLAS $b\bar{b}\gamma\gamma$: arXiv:1807.04873
ATLAS combination: ATLAS-CONF-2018-043
CMS combination: CMS-PAS-HIG-17-030

In this presentation, focusing on:

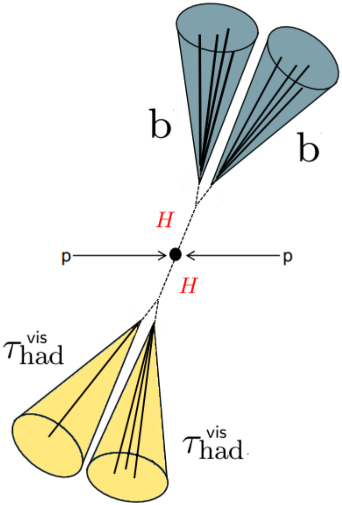
- $b\bar{b}\tau^+\tau^-$ analysis (SM + resonant HH search)
- $b\bar{b}\tau^+\tau^-$ κ_λ scan (included in the ATLAS HH combination)
- High-Luminosity LHC $b\bar{b}\tau^+\tau^-$ prospects (pub note draft ready)
- End of Run-2 prospects, Universal Fake Factor/Rate method

$HH \to b\bar{b}\tau^+\tau^-$

$\tau_{\text{lep}}\tau_{\text{had}}$ (BR: 45.8%)



$\tau_{\text{had}}\tau_{\text{had}}$ (BR: 41.9%)

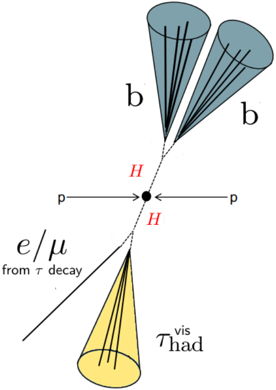


Event pre-selection

$\tau_{\text{lep}}\tau_{\text{had}}$		$\tau_{\text{had}}\tau_{\text{had}}$	
Single lepton trigger SLT	Lepton tau trigger LTT	Single tau trigger STT	Di-tau trigger DTT
1 e/μ and 1 medium τ		**2 medium τs**	
$p_T^{e/\mu} > 25, 27$ GeV (for 24, 26 GeV triggers) $p_T^{\tau} > 20$ GeV	18 GeV $< p_T^{e} <$SLT threshold 15 GeV$< p_T^{\mu} <$ SLT threshold $p_T^{\tau} > 30$ GeV	$p_T^{\text{lead}\tau} > 100, 140, 180$ GeV (for 80, 125, 160 GeV triggers) $p_T^{\text{subl}\tau} > 20$ GeV	$p_T^{\text{lead}\tau} > 40$ GeV $p_T^{\text{subl}\tau} > 30$ GeV
≥ 2 central jets			
$p_T > 45, 20$ GeV	$p_T > 80, 20$ GeV	$p_T > 45, 20$ GeV	$p_T > 80, 20$ GeV $45, 20$ GeV for 2015 data
$m_{\tau\tau}^{\text{MMC}} > 60$ GeV			

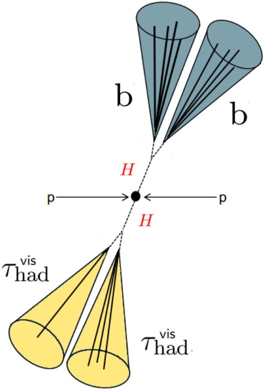
Signal/Control Regions

$\tau_{\text{lep}}\tau_{\text{had}}$		$\tau_{\text{had}}\tau_{\text{had}}$	
Single lepton trigger SLT	Lepton tau trigger LTT	Single tau trigger STT	Di-tau trigger DTT



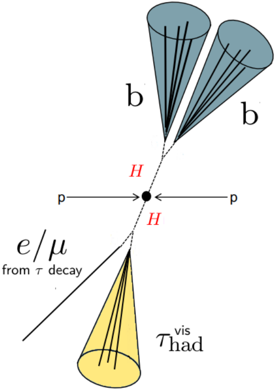
3 Signal Regions:

- Opposite charge of the τ visible decay products
- 2 b-tagged jets



Signal/Control Regions

$\tau_{\mathrm{lep}}\tau_{\mathrm{had}}$		$\tau_{\mathrm{had}}\tau_{\mathrm{had}}$	
Single lepton trigger SLT	Lepton tau trigger LTT	Single tau trigger STT	Di-tau trigger DTT

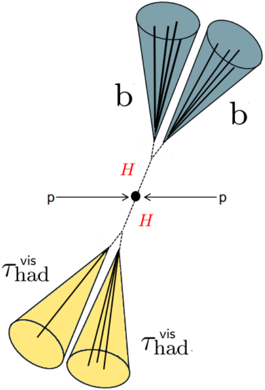


3 Signal Regions:

- Opposite charge of the τ visible decay products
- 2 b-tagged jets

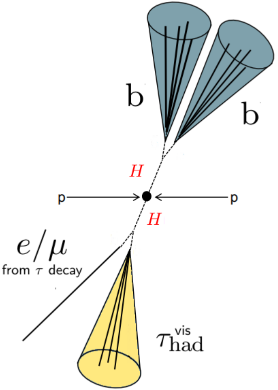
Control Regions:

- 0,1 b-tag
- Same charge
- High m_T^W , $Z + b\bar{b}$, ...



Signal/Control Regions

$\tau_{lep}\tau_{had}$		$\tau_{had}\tau_{had}$	
Single lepton trigger SLT	Lepton tau trigger LTT	Single tau trigger STT	Di-tau trigger DTT

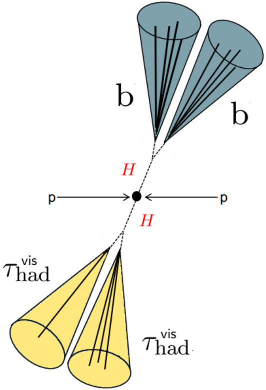


3 Signal Regions:

- Opposite charge of the τ visible decay products
- 2 b-tagged jets

Control Regions:

- 0,1 b-tag
- Same charge
- High m_T^W, $Z + b\bar{b}$, ...

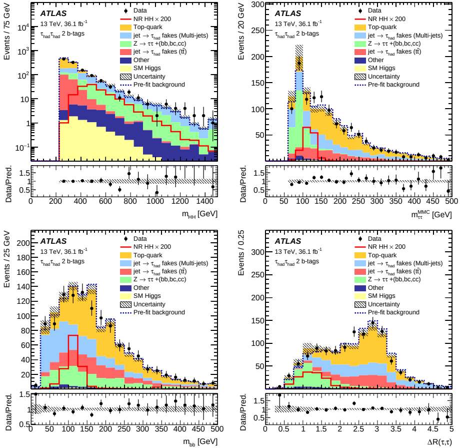


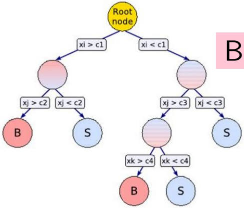
- A **Boosted Decision Tree (BDT)** classification is applied in the SR

Boosted Decision Tree

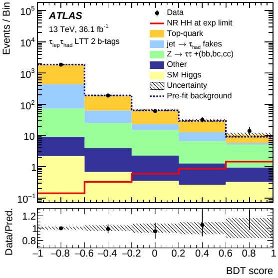
- BDT used to separate signal from background

$\tau_{\rm had}\tau_{\rm had}$ shown here (equivalent for $\tau_{\rm lep}\tau_{\rm had}$)





BDT



BDT Score - final discriminant

Boosted Decision Tree

τ_{ha}

Variable	$\tau_\ell\tau_{\mathrm{had}}$ channel (SLT resonant)	$\tau_\ell\tau_{\mathrm{had}}$ channel (SLT non-resonant & LTT)	$\tau_{\mathrm{had}}\tau_{\mathrm{had}}$ channel
m_{hh}	✓	✓	✓
$m_{\tau\tau}^{\mathrm{MMC}}$	✓	✓	✓
m_{bb}	✓	✓	✓
$\Delta R(\tau,\tau)$	✓	✓	✓
$\Delta R(b,b)$	✓	✓	✓
$E_{\mathrm{T}}^{\mathrm{miss}}$	✓		
$E_{\mathrm{T}}^{\mathrm{miss}}\,\phi$ Centrality	✓		✓
m_{T}^{W}	✓	✓	
$\Delta\phi(\mathrm{h},\mathrm{h})$	✓		
$\Delta p_{\mathrm{T}}(\ell,\tau)$	✓		
Sub-leading b-jet p_{T}	✓		

Table 1: Variables used as inputs to the BDTs for the different channels and signal models.

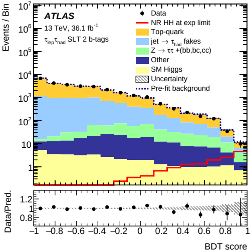
- Separate BDTs trained for each signal (and mass) hypothesis
- In resonant case the BDT is trained on the hypothesis + two neighboring mass points.
- Dedicated BDT used for κ_λ scan.

T

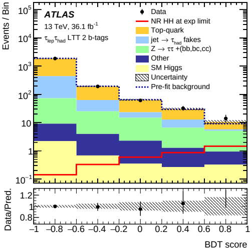
BDT Score - final discriminant

Analysis strategy

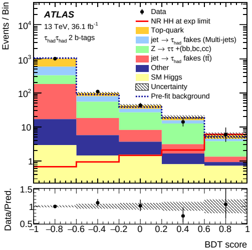
- A BDT score is used as a final discriminant:



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



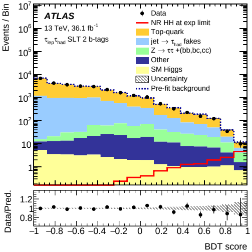
$\tau_{\rm lep}\tau_{\rm had}$ LTT



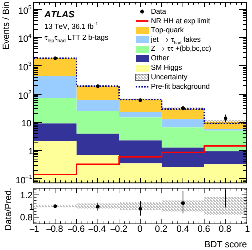
$\tau_{\rm had}\tau_{\rm had}$

Analysis strategy

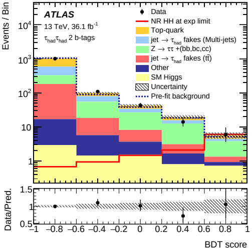
- A BDT score is used as a final discriminant:



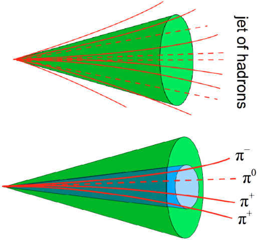
$\tau_{\rm lep}\tau_{\rm had}$ SLT



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



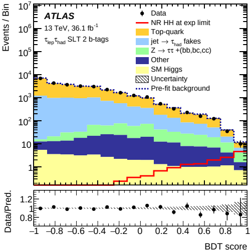
$\tau_{\rm had}\tau_{\rm had}$



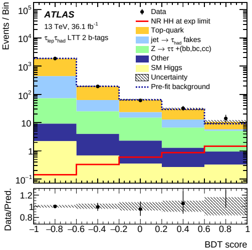
- Jet $\rightarrow$ fake τ background:
  - **Hadronically decaying τs are reconstructed as narrow jets**
  - **Other jets can be misidentified as τs**
  - **Using data-driven methods to estimate jet $\rightarrow$ fake τ from various processes**

Analysis strategy

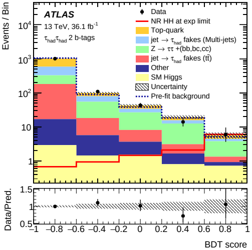
- A BDT score is used as a final discriminant:



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



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



$\tau_{\rm had}\tau_{\rm had}$

Top-quark background with true τs (MC), normalization from data

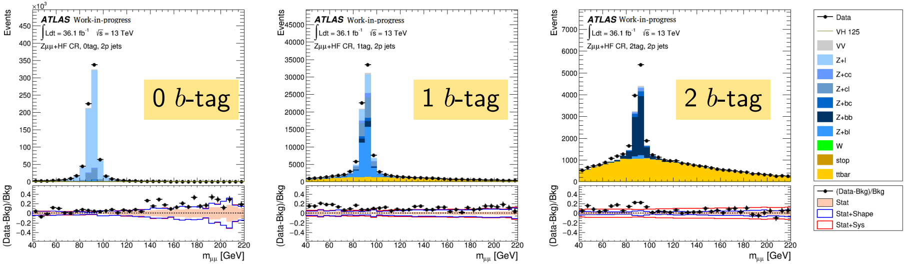
Jet $\rightarrow$ fake $\tau_{\rm had}$ from $t\bar{t}$, multijet and W+jets processes	Jet $\rightarrow$ fake $\tau_{\rm had}$ from multijet processes
	Jet $\rightarrow$ fake $\tau_{\rm had}$ from $t\bar{t}$ processes

$Z \rightarrow \tau\tau + \rm bb/bc/cc$ (MC), normalization from data

Other backgrounds estimated using Monte Carlo

Z+Heavy Flavor normalization

- Cross-section for $Z + HF$ is not well described by Monte Carlo (Sherpa).
- Dedicated $Z \to \mu\mu + \mathrm{bb/bc/cc}$ control region.
- Similar selection to the one in the Signal Region (additionally: $81 < m_{\mu\mu} < 101$ GeV).



- Normalization freely floated in the fit (one bin 2 b-tag region).
- For the SM fit (background only hypothesis): $SF(Z + HF) = 1.34 \pm 0.16$

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

Combined Fake Factor Method

- Jets $\rightarrow$ fake τs background from $t\bar{t}$, multijet and W+jets processes:

$$FF = \frac{N(\tau)}{N(\text{anti-}\tau)}$$

- **parametrized in τ p_T, #prong, trigger**
- **MC events with true τs subtracted**
- **calculated for each process in separate CRs**

(Anti-τ: τ-ID requirement inverted: !Medium and τ-ID BDT score > 0.35)

multijet CR	e/μ fail loose isolation WP, 0/1 b-tag region
$t\bar{t}$ CR	$m_T(l, MET) > 40$ GeV, 2 b-tag region
W+jets CR	$m_T(l, MET) > 40$ GeV, 0 b-tag region

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

Combined Fake Factor Method

- Jets $\rightarrow$ fake τs background from $t\bar{t}$, multijet and W+jets processes:

$$FF = \frac{N(\tau)}{N(\text{anti-}\tau)}$$

- **parametrized in τ p_T, #prong, trigger**
- **MC events with true τs subtracted**
- **calculated for each process in separate CRs**

(Anti-τ: τ-ID requirement inverted: !Medium and τ-ID BDT score > 0.35)

multijet CR	e/μ fail loose isolation WP, 0/1 b-tag region
$t\bar{t}$ CR	$m_T(l, MET) > 40$ GeV, 2 b-tag region
W+jets CR	$m_T(l, MET) > 40$ GeV, 0 b-tag region

- These are used to calculate a combined FF:

$$FF_{\text{COMB}} = FF_{\text{QCD}} \times r_{\text{QCD}} + FF_{t\bar{t}/W+\text{jets}} \times (1 - r_{\text{QCD}})$$

where r_{QCD} is the fraction of multijet events in the anti-τ signal region
QCD FFs calculated in the 1 b-tag region applied to 2 b-tag region due to low stats

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

ABCD and Fake Rate Method

- Data-driven ABCD method used to estimate the multijet background:

A OS τ	B SS τ
C OS anti-τ	D SS anti-τ

(Anti-τ: !Medium and τ-ID BDT score > 0.35)

$$FF = \frac{N(\text{SS}, \tau)}{N(\text{SS}, \text{anti-}\tau)} = \frac{\text{B}}{\text{D}}$$

$$A = FF \times C$$

- 2D FFs (function of τ_1, τ_2 p_T)
- Parametrized in #prong and trigger
- The differential FFs are derived in a 1 b-tag region (overall normalization from the 2 b-tag region)

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

ABCD and Fake Rate Method

- Data-driven ABCD method used to estimate the multijet background:

A OS τ	B SS τ
C OS anti-τ	D SS anti-τ

(Anti-τ: !Medium and τ-ID BDT score > 0.35)

$$FF = \frac{N(\text{SS}, \tau)}{N(\text{SS}, \text{anti-}\tau)} = \frac{\text{B}}{\text{D}}$$

$$A = FF \times C$$

- 2D FFs (function of τ_1, τ_2 p_T)
- Parametrized in #prong and trigger
- The differential FFs are derived in a 1 b-tag region (overall normalization from the 2 b-tag region)

- Fake Rate method used to estimate jet $\rightarrow$ fake τ background from $t\bar{t}$:

$$FR = \frac{N^{\text{passID}}}{N^{\text{total}}}$$

- Use $\tau_{\text{lep}}\tau_{\text{had}}$ $t\bar{t}$ control region to derive FRs
- Binned in τ η and #prong
- Applied to all fake-taus in MC $t\bar{t}$ events.

Contributions

- Z+HF control region, extrapolation systematic uncertainties
- NLO 2HDM sample validation/production
- Signal theory uncertainties
- BDT acceptance studies
- Checks on the statistical analysis

Non-resonant SM HH production - results

($b\bar{b}\tau^+\tau^-$ result, other channels and the combination)

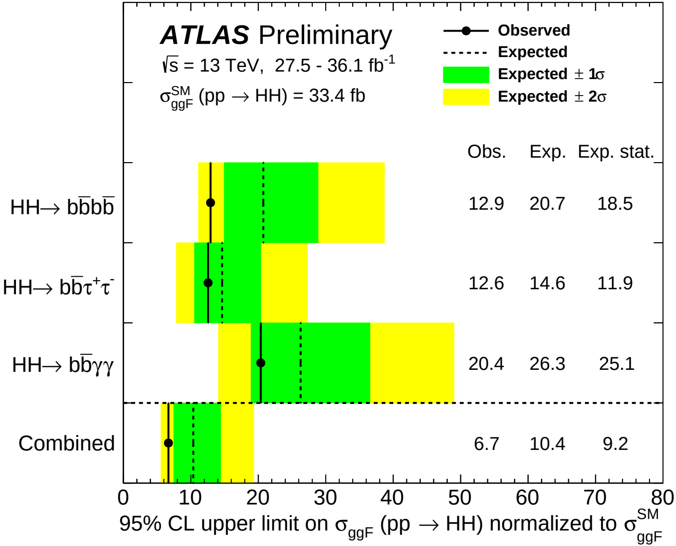
The combination is realized by constructing a combined likelihood function that takes into account data, models and systematic uncertainties

Instrumental and luminosity uncertainties correlated across the channels

The acceptance and the background modeling uncertainties treated as uncorrelated

ATLAS-CONF-2018-043

SM HH production, combined result



obs: 0.22 pb
exp: 0.35 pb

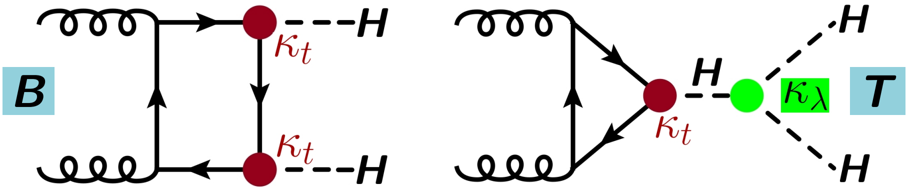
Run-1 ATLAS combination obs (exp): 70 (48) Phys. Rev. D 92, 092004

Trilinear Higgs self-coupling variations

Varied trilinear Higgs self-coupling

HH production modified

(using scale factors: $\kappa_t = g_{t\bar{t}H}/g_{t\bar{t}H}^{SM}$ and $\kappa_\lambda = \lambda_{HHH}/\lambda_{HHH}^{SM}$)

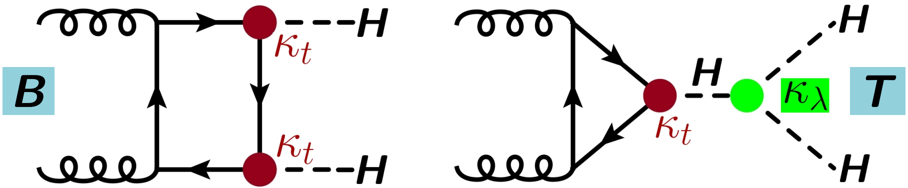


$$A(\kappa_t, \kappa_\lambda) = \kappa_t^2 B + \kappa_t \kappa_\lambda T$$

Varied trilinear Higgs self-coupling

HH production modified

(using scale factors: $\kappa_t = g_{t\bar{t}H}/g^{SM}_{t\bar{t}H}$ and $\kappa_\lambda = \lambda_{HHH}/\lambda^{SM}_{HHH}$)



$$A(\kappa_t, \kappa_\lambda) = \kappa_t^2 B + \kappa_t \kappa_\lambda T$$

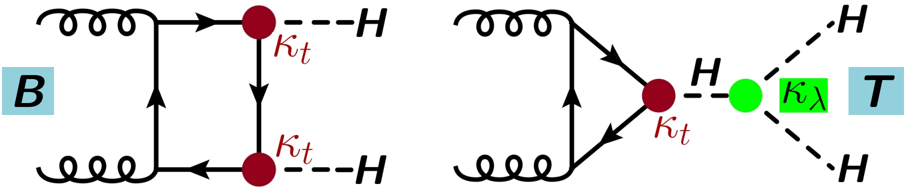
$A(1,0) = B \qquad A(1,1) = B + T \qquad A(1,2) = B + 2T$

Express $|B|^2$, $|T|^2$ and $(BT^* + TB^*)$ in terms of $|A(1,0)|^2$, $|A(1,1)|^2$ and $|A(1,2)|^2$,

Varied trilinear Higgs self-coupling

HH production modified

(using scale factors: $\kappa_t = g_{t\bar{t}H}/g_{t\bar{t}H}^{SM}$ and $\kappa_\lambda = \lambda_{HHH}/\lambda_{HHH}^{SM}$)



$$A(\kappa_t, \kappa_\lambda) = \kappa_t^2 B + \kappa_t \kappa_\lambda T$$

$$A(1,0) = B \qquad A(1,1) = B + T \qquad A(1,2) = B + 2T$$

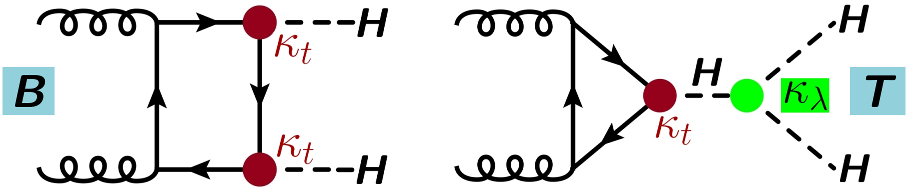
Express $|B|^2$, $|T|^2$ and $(BT^* + TB^*)$ in terms of $|A(1,0)|^2$, $|A(1,1)|^2$ and $|A(1,2)|^2$, which leads to:

$$|A(\kappa_t, \kappa_\lambda)|^2 = a(\kappa_t, \kappa_\lambda)|A(1,0)|^2 + b(\kappa_t, \kappa_\lambda)|A(1,1)|^2 + c(\kappa_t, \kappa_\lambda)|A(1,2)|^2$$

Varied trilinear Higgs self-coupling

HH production modified

(using scale factors: $\kappa_t = g_{t\bar{t}H}/g^{SM}_{t\bar{t}H}$ and $\kappa_\lambda = \lambda_{HHH}/\lambda^{SM}_{HHH}$)



$$A(\kappa_t, \kappa_\lambda) = \kappa_t^2 B + \kappa_t \kappa_\lambda T$$

$$A(1,0) = B \qquad A(1,1) = B + T \qquad A(1,2) = B + 2T$$

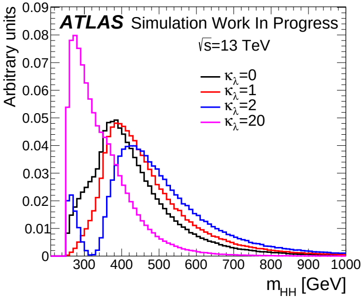
Express $|B|^2$, $|T|^2$ and $(BT^* + TB^*)$ in terms of $|A(1,0)|^2$, $|A(1,1)|^2$ and $|A(1,2)|^2$, which leads to:

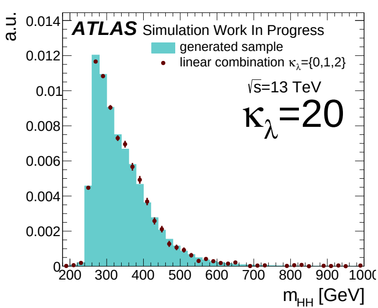
$$|A(\kappa_t, \kappa_\lambda)|^2 = a(\kappa_t, \kappa_\lambda)|A(1,0)|^2 + b(\kappa_t, \kappa_\lambda)|A(1,1)|^2 + c(\kappa_t, \kappa_\lambda)|A(1,2)|^2$$

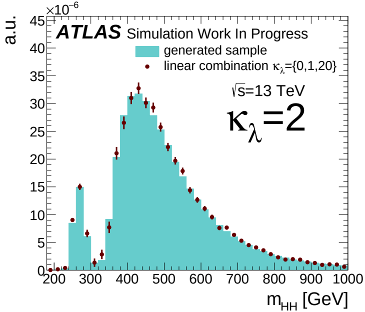
Any $(\kappa_t, \kappa_\lambda)$ combination at LO can be obtained from a **linear combination** of some 3 $(\kappa_t \neq 0, \kappa_\lambda)$ samples!

Linear combination

- Showing generator level m_{HH} for: $\kappa_\lambda = \{0, 1, 2, 20\}$ (other parameters fixed to the SM)
- Different bases tested for linear combination (e.g. $\kappa_\lambda = \{0, 1, 2\}$ vs $\kappa_\lambda = \{0, 1, 20\}$)
- Remaining sample used for validation (very good closure at generator level)



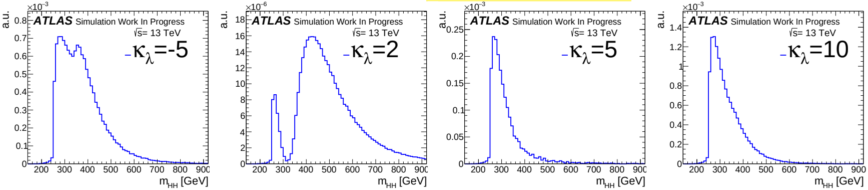




Trilinear Higgs self-coupling scan strategy

1 $m_{HH}^{\kappa_\lambda = x}$, for $x = \{-20, -19, ..., 20\}$, at generator level, at LO

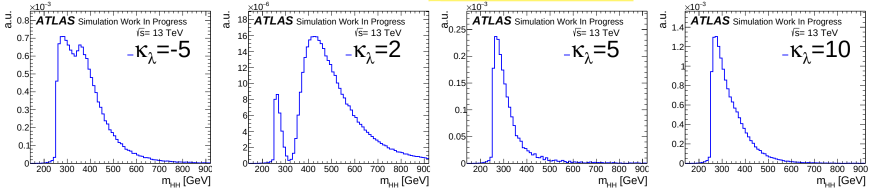
obtained using the linear combination:



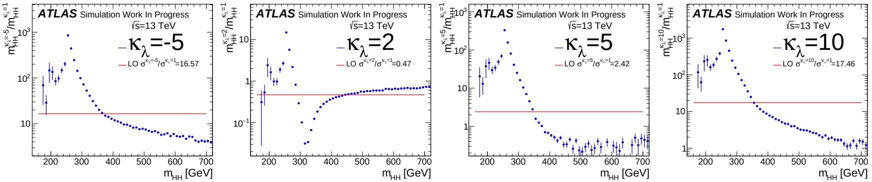
Trilinear Higgs self-coupling scan strategy

1 $m_{HH}^{\kappa_\lambda = x}$, for $x = \{-20, -19, ..., 20\}$, at generator level, at LO

obtained using the linear combination:

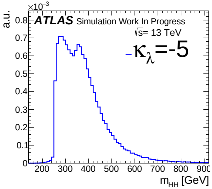


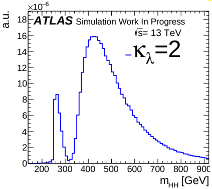
2 Weights, binned in m_{HH}, obtained as: $m_{HH}^{\kappa_\lambda = x}|_{\text{bin}\, i} / m_{HH}^{\kappa_\lambda = 1}|_{\text{bin}\, i}$

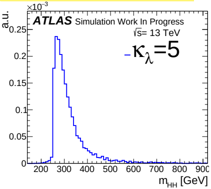


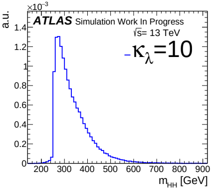
Trilinear Higgs self-coupling scan strategy

1 $m_{HH}^{\kappa_\lambda = x}$, for $x = \{-20, -19, ..., 20\}$, at generator level, at LO obtained using the linear combination:

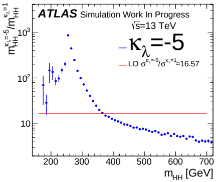


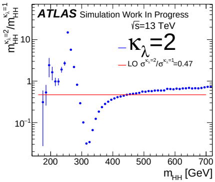


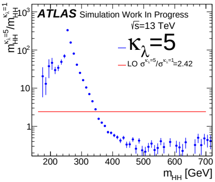


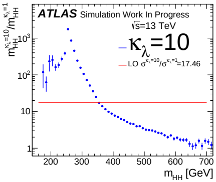


2 Weights, binned in m_{HH}, obtained as: $m_{HH}^{\kappa_\lambda = x}|_{\mathrm{bin}\, i} / m_{HH}^{\kappa_\lambda = 1}|_{\mathrm{bin}\, i}$





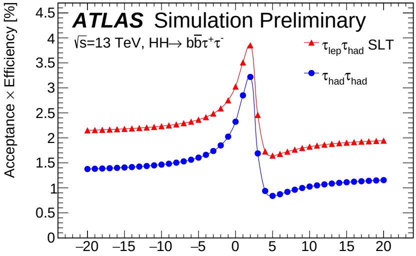




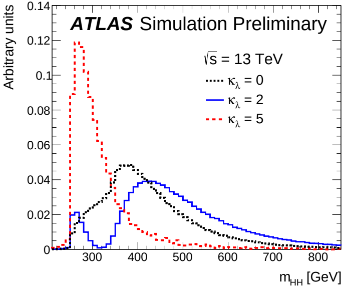
3 These weights are applied to the fully reconstructed NLO SM sample to obtain any κ_λ point, assuming that the LO to NLO factorization does not depend on κ_λ

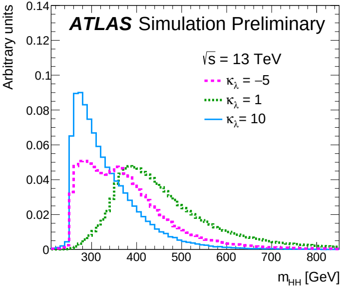
Differences compared to the SM HH search

- Acceptance changes significantly as a function of κ_λ
- A dedicated BDT, trained on $\kappa_\lambda = 20$ signal is used since it performs good for all κ_λ points.



variations of the m_{HH} spectrum with κ_λ:



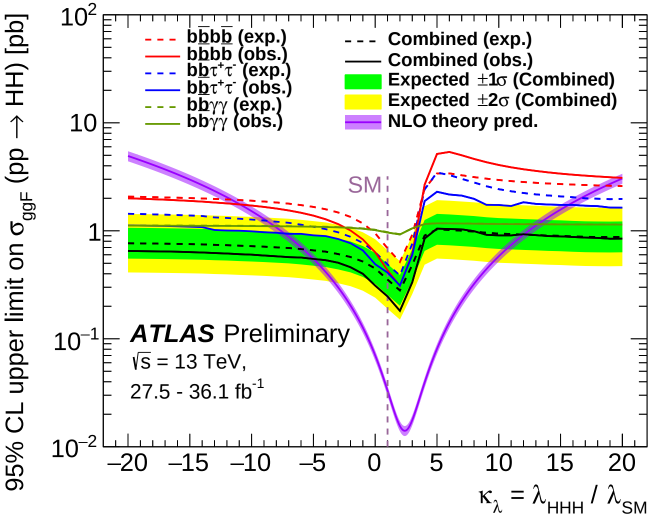


Limits on the cross-section as a function of κ_λ

$4b$
$bb\tau\tau$
$bb\gamma\gamma$
combination

dashed: expected

solid: observed



The scale factor κ_λ is observed (expected) to be constrained in the range:

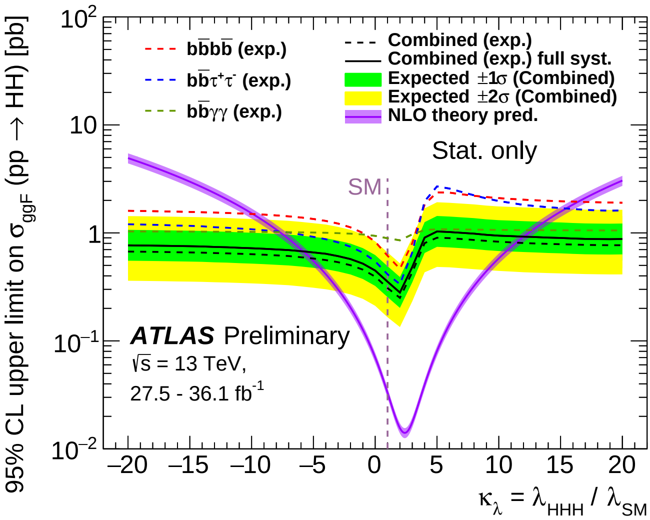
$$-5.0 < \kappa_\lambda < 12.1 \; (-5.8 < \kappa_\lambda < 12.0)$$

Full systematic uncertainty vs data stat-only

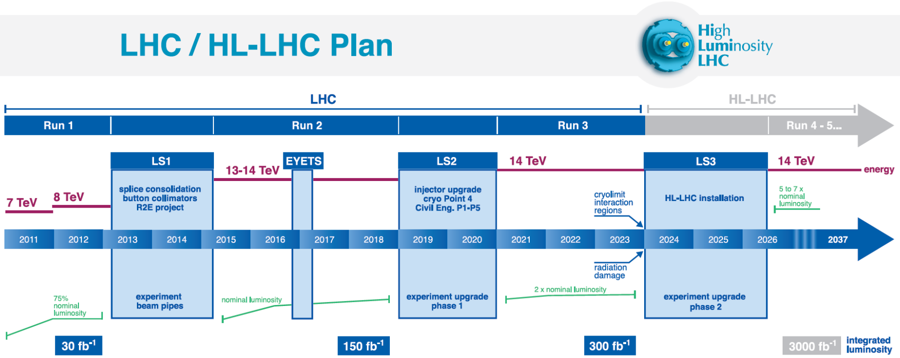
$4b$
$bb\tau\tau$
$bb\gamma\gamma$
combination

solid: expected full syst.

dashed: expected stat-only



Stat. only limits for the individual channels and the combination



$HH \to b\bar{b}\tau^+\tau^-$ HL-LHC prospects

extrapolation of the Run-2 result: $\int Ldt = 36.1 \to \int Ldt = 3000$ fb^{-1}

Signal and background distributions scaled by $f = \int Ldt|_{\text{target}} / \int Ldt|_{\text{current}}$

Signal and background distributions scaled to 14 TeV cross-sections

Normalizations fixed to the best Run-2 fit values

Pixel TDR detector layout $\to$ improved b-tagging performance (8% per b-jet)

Extrapolation strategy

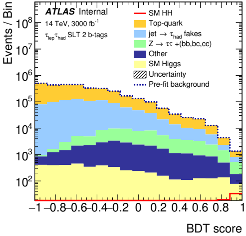
- Normalizations fixed to best Run 2 fit values
  - $Z(\to \tau\tau) + b\bar{b}$ scaled up by 1.34, uncertainty 12%
  - $t\bar{t}$ normalization unchanged, uncertainty 12%
- Signal and backgrounds scaled to 14 TeV cross-sections
- Assuming the same performance, analysis, triggers and $+8\%$ in b-tag efficiency

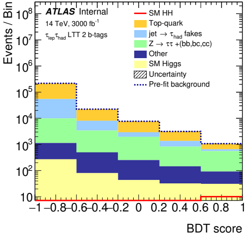
Considering 4 different scenarios:

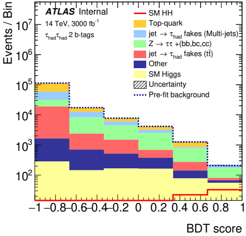
1. current systematic uncertainties
2. current systematic uncertainties, MC statistical uncertainty neglected :
   Fractional impact on $\Delta\mu$ goes from 18% (Run-2) to 84% (HL-LHC)
3. Baseline :
   12% unc on $t\bar{t}$ and $Z + b\bar{b}$ scaled down with lumi, VH scaled to 5%, $t\bar{t}H$ to 10%, all cross-section uncertainties halved, MC statistical uncertainty neglected, stat unc for data-driven bgds scaled to follow Poisson distribution
4. No systematic uncertainties

Results of the extrapolation

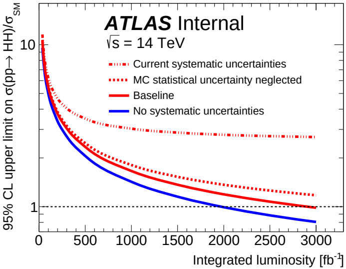
- 3 signal regions: $\tau_{\rm lep}\tau_{\rm had}$ SLT, $\tau_{\rm lep}\tau_{\rm had}$ LTT, $\tau_{\rm had}\tau_{\rm had}$

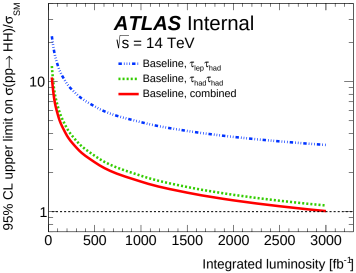






95% CL upper limit on $\sigma(pp \to HH)/\sigma_{SM}$ (background-only hypothesis):

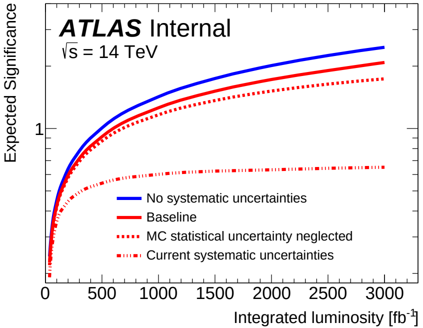




Results of the extrapolation

scenario	-1σ	expected limit	$+1\sigma$	significance
No systematic uncertainties	0.58	0.80	1.12	2.46σ
Baseline	0.71	0.99	1.12	2.08σ
MC statistical uncertainty neglected	0.85	1.18	1.64	1.74σ
Current systematic uncertainties	1.94	2.69	3.74	0.65σ

- Expected discovery significances extrapolated as well
- In the baseline scenario the expected significance is above 2σ

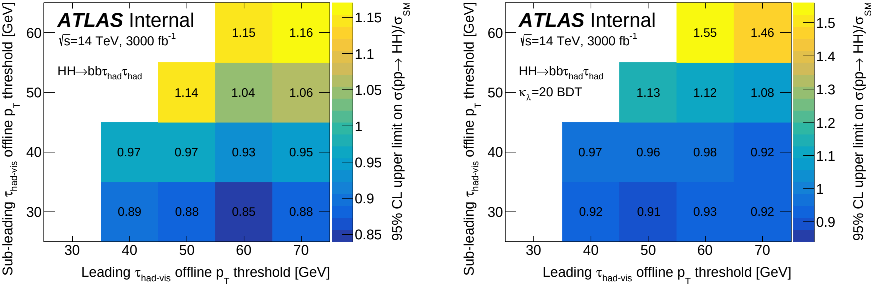


Breakdown of the systematics - baseline

Source	Uncertainty (%)
Total	±52
Data statistics	±42
Full systematic uncertainty	±30
Simulation statistics	±0
Luminosity	±4.2
Pileup reweighting	±5.9
τ_{had}	±14
Fake-τ estimation	±6.3
b- tagging	±7.4
Jets and $E_{\text{T}}^{\text{miss}}$	±3.5
Electron and muon	±5.6
Experimental Uncertainties	±19
Top	±10
Signal	±11
$Z \to \tau\tau$	±14
SM Higgs	±7.5
Other backgrounds	±5.9
Theoretical and Modeling Uncertainties	±19

Table 2: The percentage uncertainties on the simulated SM non-resonant signal strength, i.e. the simulated SM HH yield assuming a cross-section times branching fraction equal to the 95% CL expected limit of 1.0 (limit in baseline scenario) times the SM expectation.

Di-tau trigger studies

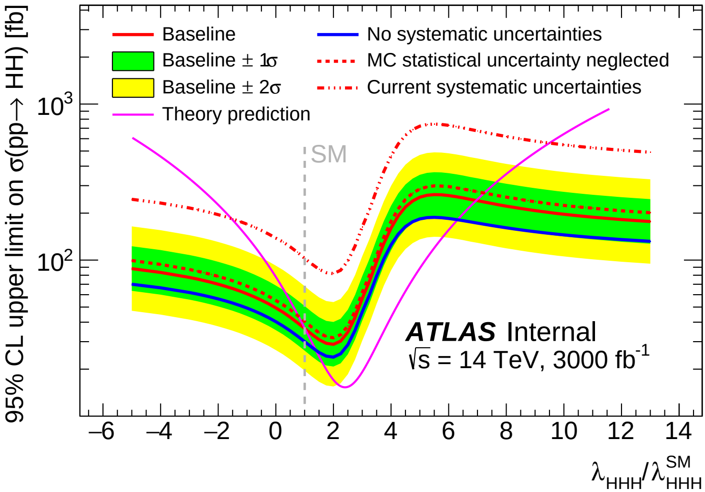


Expected 95% CL upper limit on $\sigma(pp \to HH)/\sigma_{SM}$ (without systematic uncertainties) as a function of the leading and sub-leading $\tau_{had-vis}$ minimum p_{T} thresholds, using the (a) nominal BDT classifier and (b) using the $\kappa_\lambda = 20$ BDT

- The loss in sensitivity is expected to be even more pronounced (the effect masked by $+80$ GeV jet requirement)
- Sensitivity to the Higgs self-coupling is affected more by raising the p_{T} thresholds (softer p_T spectrum), so the study is repeated for $\kappa_\lambda = 20$ BDT

Limits on the cross-section as a function of κ_λ

- Allowed 95% CL κ_λ interval (background-only hypothesis: $\sigma_{HH} = 0$)
  no systematic uncertainties: $1.4 < \kappa_\lambda < 6.3$, baseline: $1.0 < \kappa_\lambda < 7.0$

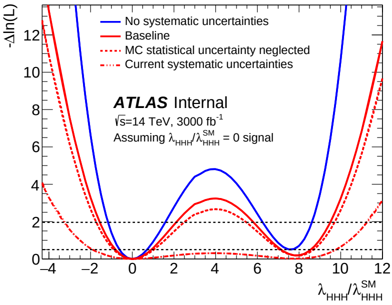


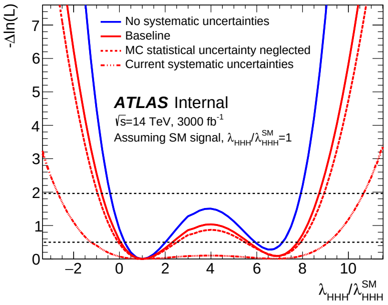
Likelihood ratio test

We can determine allowed κ_λ interval also for assuming:
(left) no Higgs self-coupling (box diagram only, $\kappa_\lambda = 0$) and
(right) assuming SM di-Higgs ($\kappa_\lambda = 1$)

- Allowed 95% CL interval ($\kappa_\lambda = 0$)
  no systematics: $1.2 < \kappa_\lambda < 1.6$ U $6.3 < \kappa_\lambda < 8.7$,
  baseline: $-1.6 < \kappa_\lambda < 2.1$ U $5.8 < \kappa_\lambda < 9.5$

- Allowed 95% CL κ_λ interval (assuming SM signal, $\kappa_\lambda = 1$)
  no systematics: $-0.4 < \kappa_\lambda < 7.9$
  baseline: $-0.8 < \kappa_\lambda < 8.7$





Universal Fake Factor/Rate method

Introduction

- Jet→fake τ estimation in ATLAS done separately for different analyses.
- There are ongoing efforts to measure the Fake Factors (FF) and Fake Rates (FR) centrally and to develop a centralized method that can be used by most of the analyses.

General idea:

- Measure FFs/FRs in regions with different quark/gluon fractions
- Identify well separating variable(s). Estimate q/g fraction in data by fitting MC templates.
- Provide a set of recommendations on which systematic uncertainties to consider
- Consider the impact of b- and c- in respect to light-jets.
- Provide generic tool for:
  (1) measuring the q/g fraction in the analysis SR
  (2) applying the centrally measured fake factors/rates.

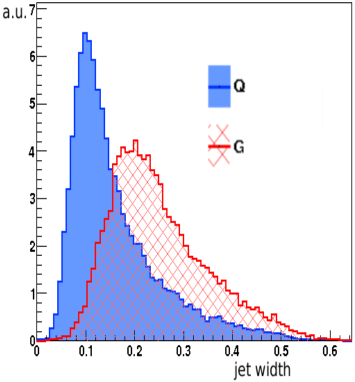
Slide from Lino Gerlach

Jet Width

Jets can be initiated by quarks or gluons

- Quark initiated jets are more τ-like
- Quark/gluon (q/g) fraction has impact on fake rate and fake factor
- Well separating variable: jet width
  - weighted average ΔR of all objects within the jet

$$j = \frac{\sum_i \Delta R^i p_T^i}{\sum_i p_T^i}$$



arXiv:1106.3076v2 [hep-ph] 19 Oct 2011

Approach

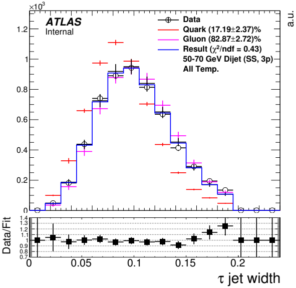
- Conduct Template Fit to estimate q/g fraction in SR
  - Find FF suitable for q/g fraction

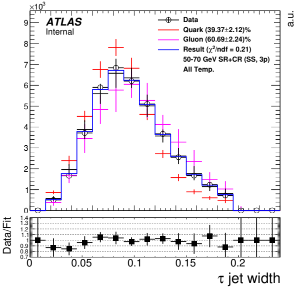
Measuring quark/gluon fraction

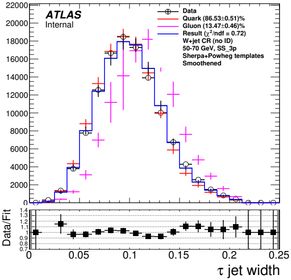
Quark and gluon templates extracted from different MC processes.

Example: 50-70 GeV, SS, 3p

Dijet region	SCR	W+jet region







Slide from Lino Gerlach

Slide from Lino Gerlach

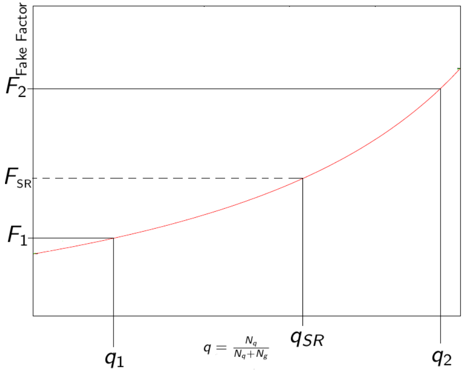
Fake Factor Interpolation

$$F = \frac{\#(\text{quarks passing}) + \#(\text{gluons passing})}{\#(\text{quarks failing}) + \#(\text{gluons failing})} = \frac{N_q p_q + N_g p_g}{N_q(1-p_q) + N_g(1-p_g)}$$

$$= \frac{q(F_q - F_g) + F_g + F_q F_g}{q(F_g - F_q) + 1 + F_q}$$

Fake Factor Interpolation

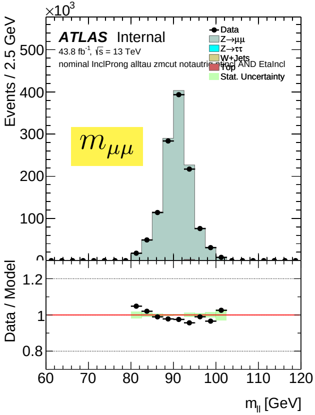
1. Measure q_1 and F_1 in gluon dominated region
2. Measure q_2 and F_2 in quark dominated region
3. Measure q_{SR} in SR and interpolate



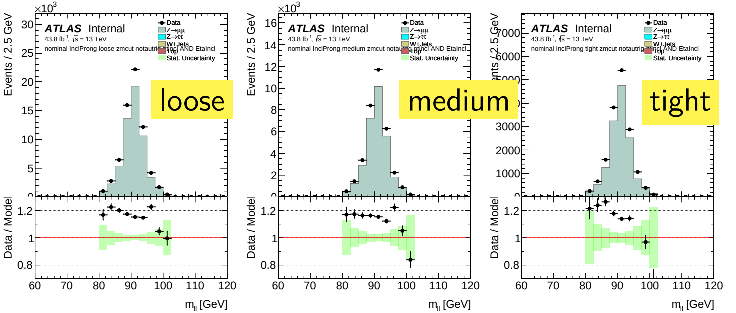
$Z(\rightarrow \mu\mu)$+jets region

- A T&P analysis to measure FF/FR in $Z \rightarrow \mu\mu$+jets channel
- Clean region, easy to validate, small fraction of real τ leptons, dominated by quark-initiated jets
- Important to consider the impact of τ-trigger decision in all regions

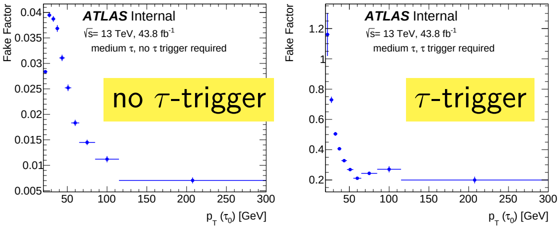
- 1 or 3 prong τ candidate required, $p_T > 18$ GeV
- Single muon trigger
- Leading muon $p_T > 27$ GeV, trigger matched
- Sub-leading muon $p_T > 20$ GeV, opposite electric charge
- N_τ candidates $= 1$, $p_T > 20$ GeV, Electron veto
- $81. < m_{\mu\mu} < 101.$ GeV, $p_T^{\mu\mu} > 15.$ GeV



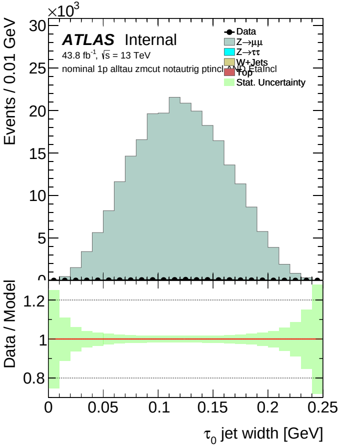
Tau ID and trigger requirements



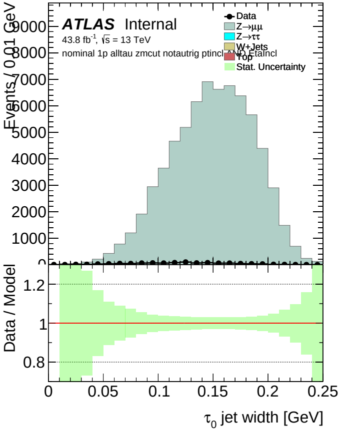
- Different fake rate in MC and data.
- Important to consider the effect of the online τ-ID on the fake rates



Jet Width



quark-initiated



gluon-initiated

- Jet width templates for quark- and gluon-initiated jets

Conclusion & Outlook

- Interesting Run-2 and HE-LHC prospects results from $b\bar{b}\tau\tau$ analysis and HH in general
- Working on the Universal FF/FR methods within the fake-τ task-force
- Ongoing efforts on including the 2017+2018 data within the $b\bar{b}\tau\tau$ analysis
- Many analysis improvements under consideration
- Planned contributions: implementation of the universal FF/FR methods, re-definition of the Signal/Control regions, statistical analysis, $(\kappa_\lambda, \kappa_t)$ scans at NLO, Effective Field Theory re-interpretations (shape benchmarks)
- Work on the boosted $b\bar{b}\tau\tau$ with Christina/Myrto

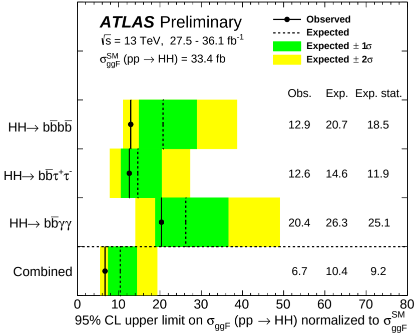
My many supervisors... (+Michel)

Thank you for your attention!

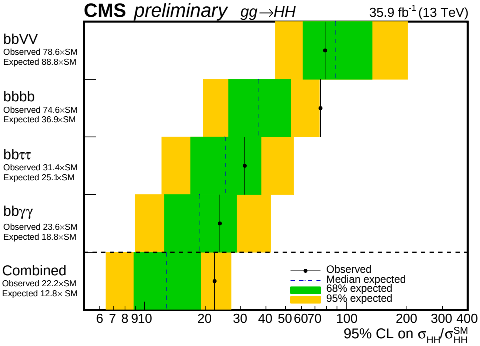
backup slides

SM HH production, combined results

- Most recent ATLAS and CMS combinations of di-Higgs searches
- $bb\tau\tau$ proves to be one of the most sensitive channels



ATLAS-CONF-2018-043



CMS-PAS-HIG-17-030

	$bb\tau\tau$ obs (exp)	combined obs (exp)
ATLAS	12.6 (14.6)	6.7 (10.4)
CMS	31.4 (25.1)	22.2 (12.8)

Allowed intervals for κ_λ

Search channel	Allowed κ_λ interval at 95% CL		
	obs.	exp.	exp. stat.
$HH \to b\bar{b}b\bar{b}$	-10.9 – 20.1	-11.6 – 18.7	-9.9 – 16.4
$HH \to b\bar{b}\tau^+\tau^-$	-7.3 – 15.7	-8.8 – 16.7	-7.8 – 15.4
$HH \to b\bar{b}\gamma\gamma$	-8.1 – 13.2	-8.2 – 13.2	-7.7 – 12.7
Combination	-5.0 – 12.1	-5.8 – 12.0	-5.2 – 11.4

Resonant HH production

(combination in the mass range: 260-1000 GeV)

Differences compared to the SM HH search:

$bb\gamma\gamma$:

looser selection below 500 GeV

final discriminant: $m_{\gamma\gamma jj}$

$b\bar{b}\tau\tau$:

dedicated BDTs

$b\bar{b}b\bar{b}$:

boosted analysis for signal masses > 800 GeV
(combined with the resolved)

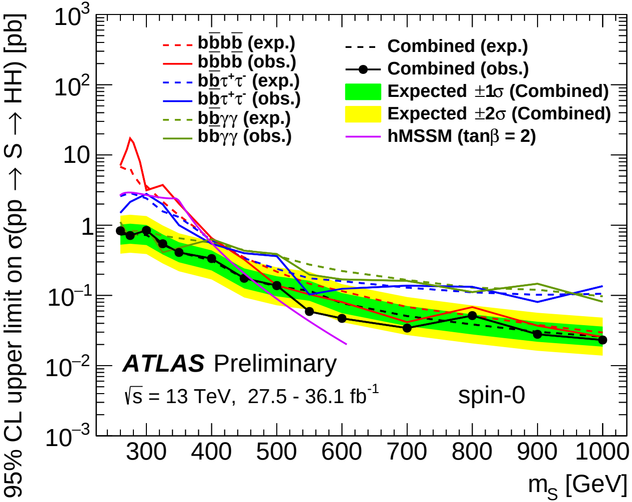
Looking for two Higgs candidates, each composed of a single large-R (1.0) jet with at least one b-tagged track jet associated to it

Scalar resonance

$4b$
$bb\tau\tau$
$bb\gamma\gamma$
combination

dashed: expected

solid: observed



hMSSM, narrow width CP-even Higgs boson (tan$\beta = 2$)*: $m_S < 462$ GeV at 95% CL

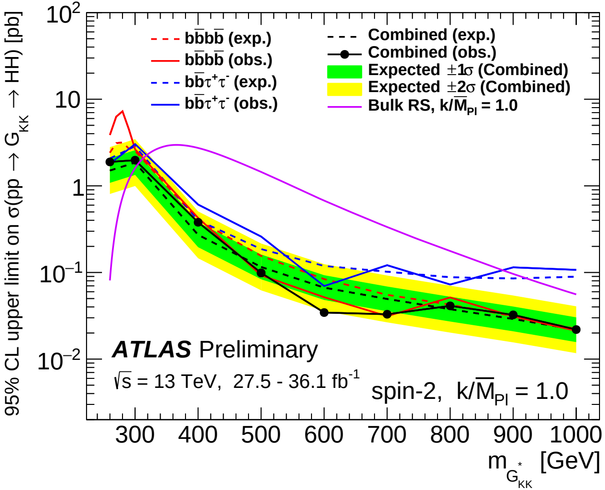
*tan$\beta = 2$: ratio of the vacuum expectation values of the two Higgs doublets

Randall-Sundrum graviton model

$4b$

$bb\tau\tau$

combination



dashed: expected

solid: observed

$(k/\bar{\mathrm{M}}_{\mathrm{Pl}} = 1)^*$ $307 < m_G < 1362$ GeV** at 95% CL

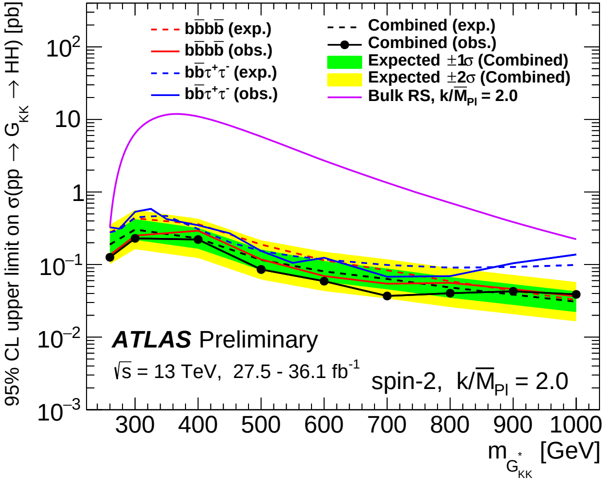
$^*k$: curvature of the warped extra dimension, $\bar{\mathrm{M}}_{\mathrm{Pl}}$: the effective four-dimensional Planck scale
**the upper limit on the mass comes from $4b$ only

Randall-Sundrum graviton model

$4b$
$bb\tau\tau$
combination

dashed: expected

solid: observed



$(k/\bar{\mathrm{M}}_{\mathrm{Pl}} = 2)^* \ m_G < 1744$ GeV** at 95% CL

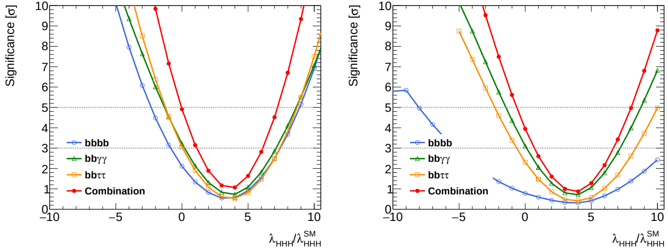
$^*k$: curvature of the warped extra dimension, $\bar{\mathrm{M}}_{\mathrm{Pl}}$: the effective four-dimensional Planck scale
**the upper limit on the mass comes from $4b$ only

HL-LHC HH combination

- Significance (no systematics, baseline):

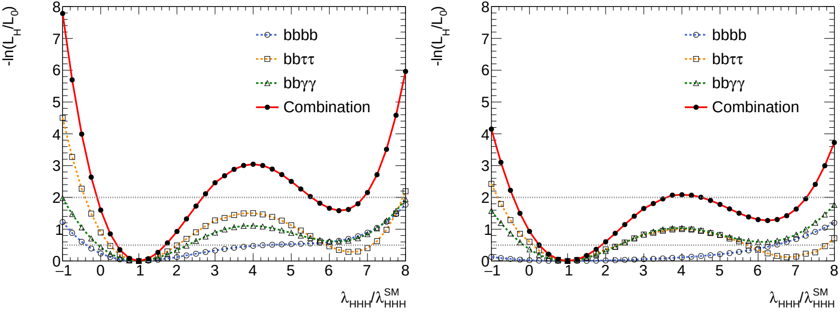
Channel	Statistical-only p_0	Statistical-only Significance	All Systematics p_0	All Systematics Significance
$hh \rightarrow bbbb$	0.0825	1.39	0.271	0.609
$hh \rightarrow bb\tau\tau$	0.00686	2.46	0.0164	2.13
$hh \rightarrow bb\gamma\gamma$	0.0180	2.10	0.0210	2.03
combined	0.000202	3.54	0.00197	3.02

- Significance as a function of κ_λ (no systematics, baseline):



HL-LHC HH combination

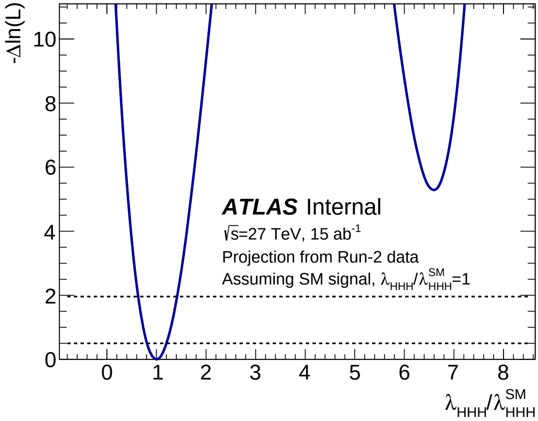
- Limits on the κ_λ, assuming SM signal (no systematics, baseline):



- Confidence intervals on κ_λ from the combination (no systematics):
  - 68%: $0.4 < \kappa_\lambda < 1.7$
  - 95%: $-0.1 < \kappa_\lambda < 2.7$ U $5.5 < \kappa_\lambda < 6.9$
- Confidence intervals on κ_λ from the combination (with systematics):
  - 68%: $0.3 < \kappa_\lambda < 1.9$
  - 95%: $-0.4 < \kappa_\lambda < 3.6$ U $4.5 < \kappa_\lambda < 7.3$

HE-LHC bbtautau prospects

- The discovery significance is expected to be 8.2σ
- The allowed range at 68% (95%) CL for κ_λ with 15 ab^{-1} of $\sqrt{s} = 27$ TeV data is expected to be 0.8<κ_λ<1.2 (0.6<κ_λ<1.4) - assuming SM signal



$$\mathcal{L}(\mathcal{D}, \mathcal{G} | \mu, \alpha) = \prod_{c \in \mathbb{C}} \text{Pois}(n_c | \nu_c(\mu, \alpha)) \prod_{e=1}^{n_c} f_c(x_{ce} | \mu, \alpha) \times \prod_{p \in \mathbb{S}} f_p(a_p | \alpha_p)$$

$$L(\mu, \boldsymbol{\theta}) = \prod_{j=1}^{N} \frac{(\mu s_j + b_j)^{n_j}}{n_j!} e^{-(\mu s_j + b_j)} \quad \prod_{k=1}^{M} \frac{u_k^{m_k}}{m_k!} e^{-u_k}$$