ttH Prospect for Run II

Georges Aad, CPPM, Aix-Marseille Université and CNRS/IN2P3, Marseille, France

Top LHC France, 2015-05-19

Disclaimer: due to personal reasons I did not have time to time-travel to 2018 and bring back run II ttH public results; so here is a great presentation about feelings and intuitions

Introduction

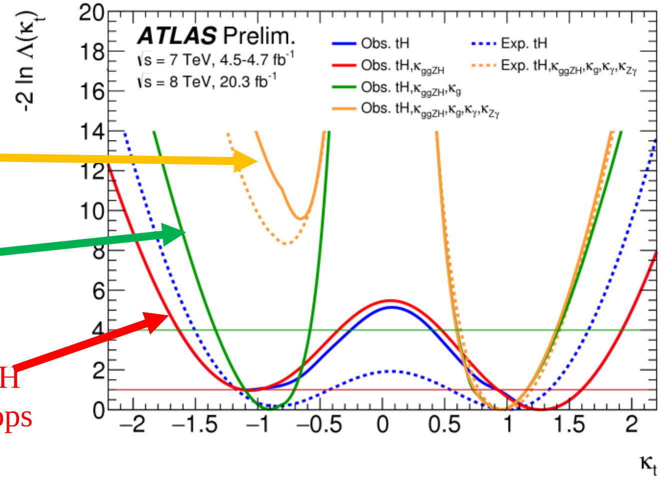
- The top quark is special vis-à-vis of the Higgs boson
  - Largest Yukawa coupling, order of one
  - Just a coincidence or there is something deeper?
- Only coupling that can't be directly measured through Higgs decay
  - Indirect access through loops in Higgs gluon fusion production and $H\rightarrow\gamma\gamma$
  - Well constrained in run I assuming the SM particle content
- Need ttH to constrain top-Higgs coupling with a (more) model-independent assumptions
  - Need tH to resolve the top-Higgs coupling sign (with respect to H-W)
- ttH should be observed for the first time in LHC run II
  - However 5σ discovery very challenging in run II

ATLAS-CONF-2015-007

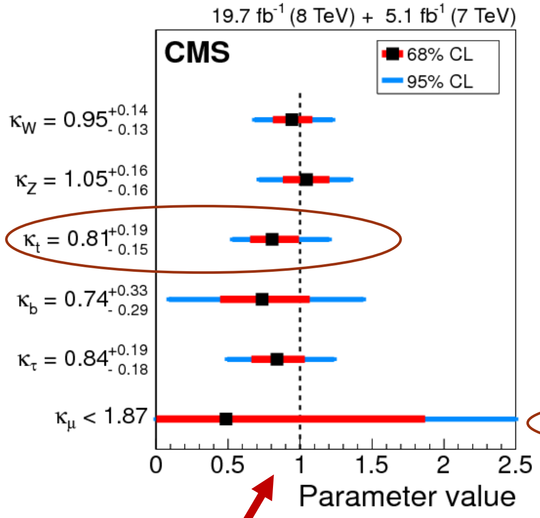
SM content

BSM in gg→H

BSM in gg→H and H→ γγ loops

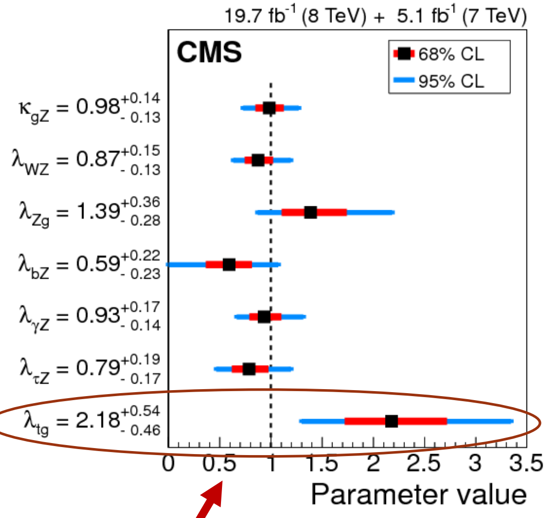


arXiv:1412.8662



SM content

arXiv:1412.8662



BSM in loops and no constraint on total width

ttH Channels Overview

- Challenging channel with small cross section
  - Exploit every accessible experimental signature
- A wide variety of final states
  - Not always corresponding to a specific production/decay mode
    - Reconstructing the event is not always possible
- 3 main groups of channels
  - H→bb
  - Multiplepton (electron, muon)
    - Mainly from H→WW and H→ZZ and $\tau \to e/\mu$
  - $H \to \gamma\gamma$
  - Also exploiting channels with τ_h
- Rich phenomenology of Higgs coupling
  - However hard to exploit with the low statistics in most of these channels
- More exclusive selection/splitting will be available with more statistics in run II
  - Target better S/B

Overview of CMS ttH channels

Category	Signature	Trigger	Signature
H → Hadrons $H \to b\bar{b}$ $H \to \tau_h\tau_h$ $H \to WW$	Lepton + Jets ($t\bar{t}H \to \ell\nu jjbbbb$)	Single Lepton	1 e/μ, $p_T > 30$ GeV $\geq$4 jets + $\geq$2 b-tags, $p_T > 30$ GeV
	Dilepton ($t\bar{t}H \to \ell\nu\ell\nu bbbb$)	Dilepton	1 e/μ, $p_T > 20$ GeV 1 e/μ, $p_T > 10$ GeV $\geq$3 jets + $\geq$2 b-tags, $p_T > 30$ GeV
	Hadronic τ ($t\bar{t}H \to \ell\nu\tau_h[\nu]\tau_h[\nu]jjbb$)	Single Lepton	1 e/μ, $p_T > 30$ GeV 2 τ_h, $p_T > 20$ GeV $\geq$2 jets + 1-2 b-tags, $p_T > 30$ GeV
H → Photons $H \to \gamma\gamma$	Leptonic ($t\bar{t}H \to \ell\nu jjbb\gamma\gamma$, $t\bar{t}H \to \ell\nu\ell\nu bb\gamma\gamma$)	Diphoton	2 γ, $p_T > m_{\gamma\gamma}/2$ (25) GeV for 1st (2nd) $\geq$1 e/μ, $p_T > 20$ GeV $\geq$2 jets + $\geq$1 b-tags, $p_T > 25$ GeV
	Hadronic ($t\bar{t}H \to jjjjbb\gamma\gamma$)	Diphoton	2 γ, $p_T > m_{\gamma\gamma}/2$ (25) GeV for 1st (2nd) 0 e/μ, $p_T > 20$ GeV $\geq$4 jets + $\geq$1 b-tags, $p_T > 25$ GeV
H → Leptons $H \to WW$ $H \to \tau\tau$ $H \to ZZ$	Same-Sign Dilepton ($t\bar{t}H \to \ell^\pm\nu\ell^\pm[\nu]jjj[j]bb$)	Dilepton	2 e/μ, $p_T > 20$ GeV $\geq$4 jets + $\geq$1 b-tags, $p_T > 25$ GeV
	3 Lepton ($t\bar{t}H \to \ell\nu\ell[\nu]\ell[\nu]j[j]bb$)	Dilepton, Trielectron	1 e/μ, $p_T > 20$ GeV 1 e/μ, $p_T > 10$ GeV 1 $e(\mu)$, $p_T > 7(5)$ GeV $\geq$2 jets + $\geq$1 b-tags, $p_T > 25$ GeV
	4 Lepton ($t\bar{t}H \to \ell\nu\ell\nu\ell[\nu]\ell[\nu]bb$)	Dilepton, Trielectron	1 e/μ, $p_T > 20$ GeV 1 e/μ, $p_T > 10$ GeV 2 $e(\mu)$, $p_T > 7(5)$ GeV $\geq$2 jets + $\geq$1 b-tags, $p_T > 25$ GeV

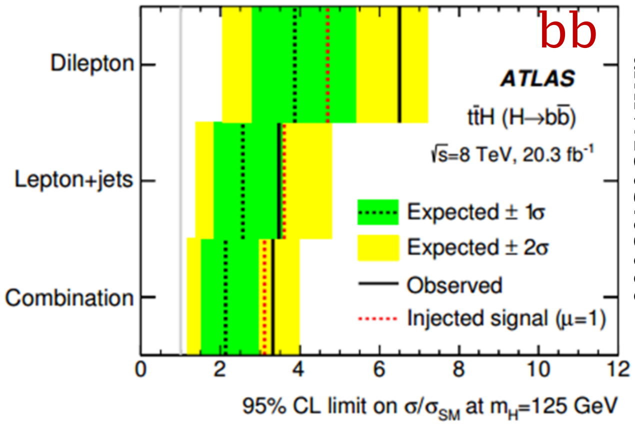
Basically the same for ATLAS in addition to $2l1\tau_h$ final state

Run I Overview

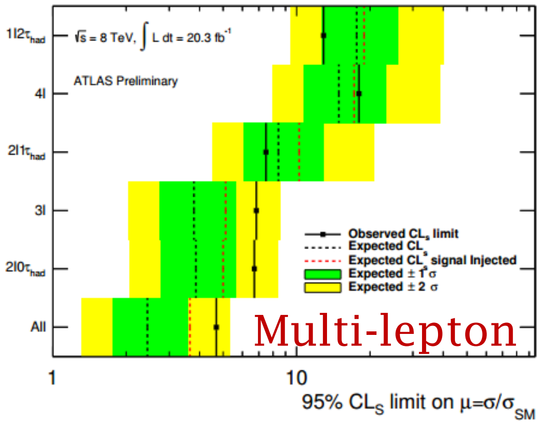
Expected limit summary

	bb	$\gamma\gamma$	2l	3l	4l	$\tau_h\tau_h$	$2l1\tau_h$	Combination
ATLAS	2.2	4.9	3.9	3.8	8.4	18	8.4	-
CMS	3.5 (3.3MEM)	4.7	3.4	4.1	8.8	14.2	-	1.7

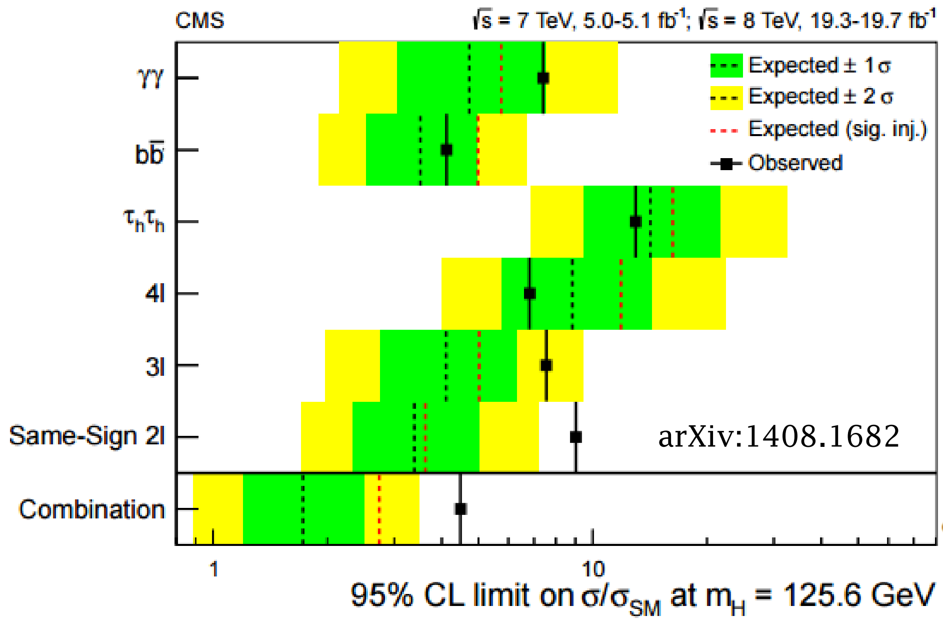
- Legacy results from CMS
- ATLAS still have to publish multilepton and the combination
  - Also expected ttH(bb) all-hadronic
- Higher statistics bb channel provides the best limit (2.2) for ATLAS
  - bb channel in CMS gives similar results to the 2l channel
- Higher purity channels with low stat ($\gamma\gamma$,4l) will become more competitive in run II

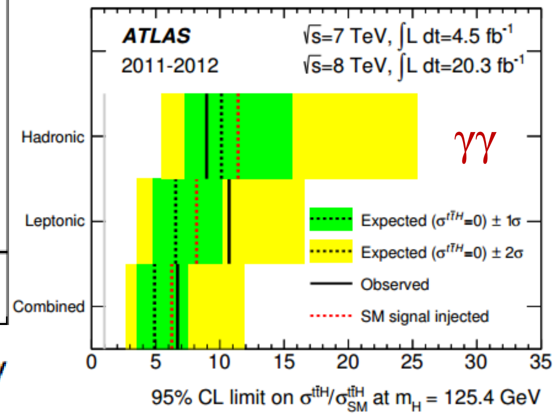


arXiv:1503.05066



ATLAS-CONF-2015-006





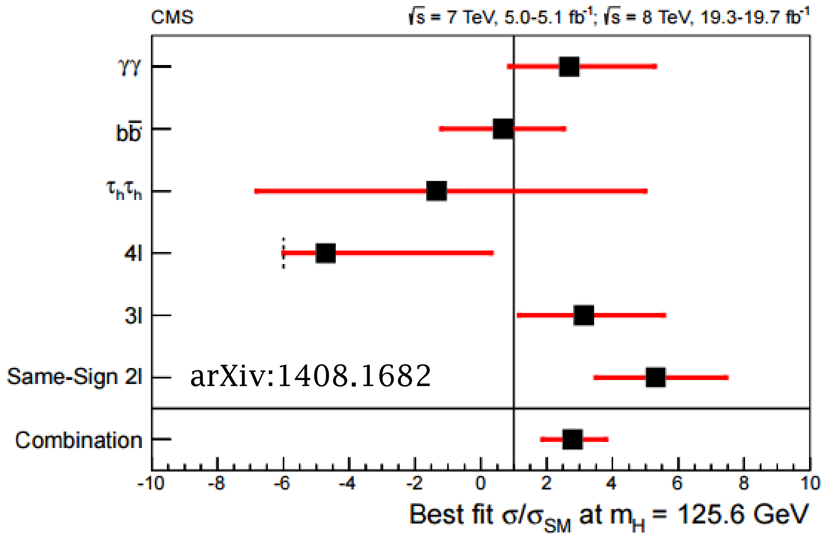
arXiv:1409.3122

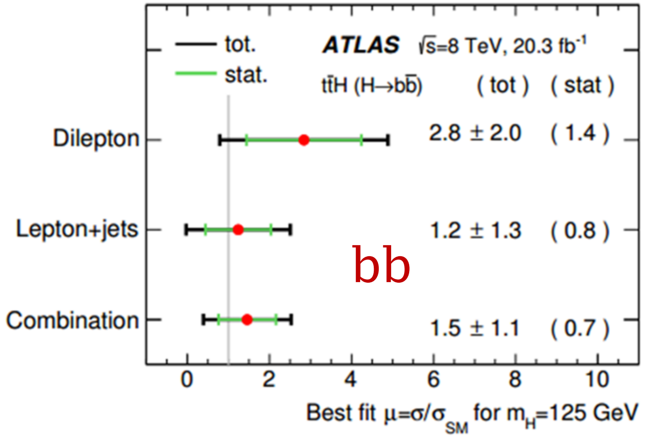
Run I Overview

Expected limit summary

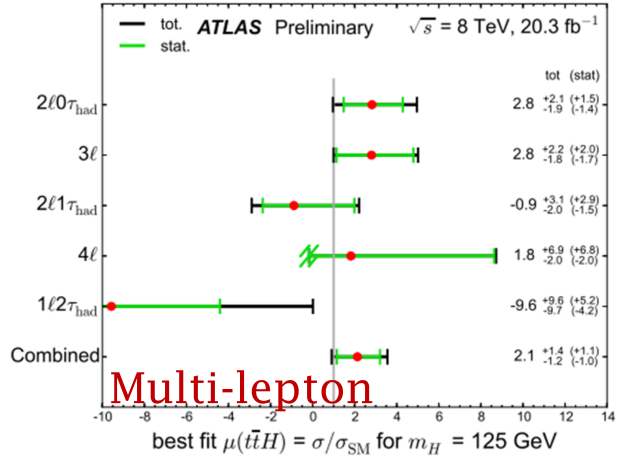
	bb	$\gamma\gamma$	2l	3l	4l	$\tau_h\tau_h$	$2l1\tau_h$	Combination
ATLAS	2.2	4.9	3.9	3.8	8.4	18	8.4	-
CMS	3.5 (3.3MEM)	4.7	3.4	4.1	8.8	14.2	-	1.7

- Legacy results from CMS
- ATLAS still have to publish multilepton and the combination
  - Also expected ttH(bb) all-hadronic
- Higher statistics bb channel provides the best limit (2.2) for ATLAS
  - bb channel in CMS gives similar results to the 2l channel
- Higher purity channels with low stat ($\gamma\gamma$,4l) will become more competitive in run II

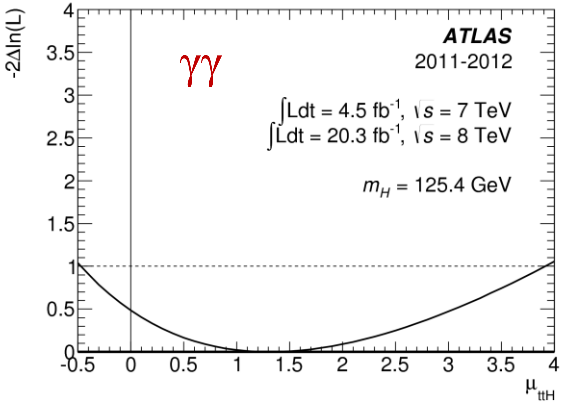




ArXiv:1503.05066



ATLAS-CONF-2015-006



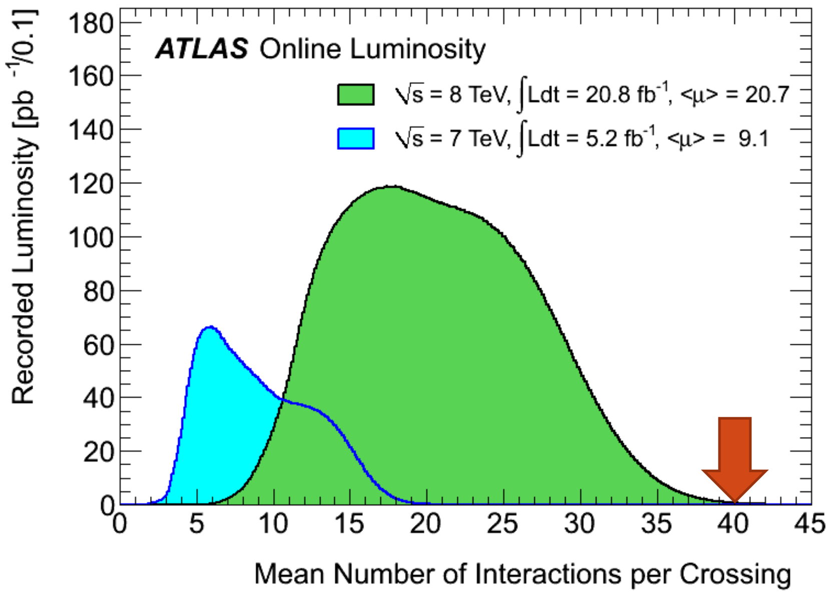
arXiv:1409.3122

Everything compatible with an SM Higgs boson

Cross Sections and Luminosity in Run II

- ttH cross section increases by a factor of ~4 with respect to 8TeV
  - A bit less for ttbar and ttV
    - Will benefit from a slightly higher S/B
  - Reach Run I sensitivity around 5-10 fb^{-1}
- LHC luminosity will increase to ~1.3×10^{34} $cm^{-2}s^{-1}$ but with 25 ns bunch crossing
  - Acceptable pileup (40 per BC)
  - Expected to accumulate 10 fb^{-1} in 2015
  - 100 fb^{-1} in run II
- Produced ttH events in run II multiplied by ~16 with respect to run I
  - Will not show projections done with tons of assumptions especially that most of the ttH channels have large systematics

	ttH (fb)	ttbar (pb)	ttW (fb)	ttZ (fb)
8TeV	130	253	232	206
13TeV	510	831	566	750
ratio	3.9	3.3	2.4	3.6



MC Development

Already detailed in Marco's presentation

- ttH signal
  - Benefit from latest development and NLO+PS generators (PowHel, MD5_aMC@NLO)
    - NLO (QCD) available, NLO (EW) soon
  - Still in a search period, more important is to well model the background
  - But also important to have a good ttH description for MVAs and for μ/couplings determination
- Main background from associated production of top quarks with jets or vector bosons
  - Benefit from ttbar differential measurements done in run I
    - Avoid MC reweighting and tune generators to match data
    - Run II SM measurements will probably follow ttH results
  - Many recent developments with state-of-the-art generators
    - tt+jj (bb) available at NLO matched with parton shower
    - tt+V+1j also available at NLO+PS
- Benefit from latest PDFs and tunes that includes LHC run I data

$t\bar{t}H$, $H \rightarrow b\bar{b}$

Selection

- Low stat even in the $H \rightarrow bb$ channel
  - Will significantly benefit from the larger stat in run II
  - Take advantages of all available channels
    - Work ongoing on all hadronic channel (ATLAS run I)
- Low signal purity
  - Control background is one of the most important tasks
- Increased lepton trigger threshold in run II
  - Reduced signal acceptance
    - However going to higher pT reduces systematic uncertainties
  - Single lepton channel most affected
  - b-jet trigger available in run I (but not used for ttH)
    - Further developed in run II, can we take advantage?
- Signal acceptance will benefit from the jump to 13 TeV
  - Higher pT objects
- Higher pileup (not at the beginning – 25ns vs 50ns)
  - Will mainly affect lepton isolation, JES and jet multiplicities
- Expected improved b-tagging (following slides)

Run I expected signal events (S/B [%])

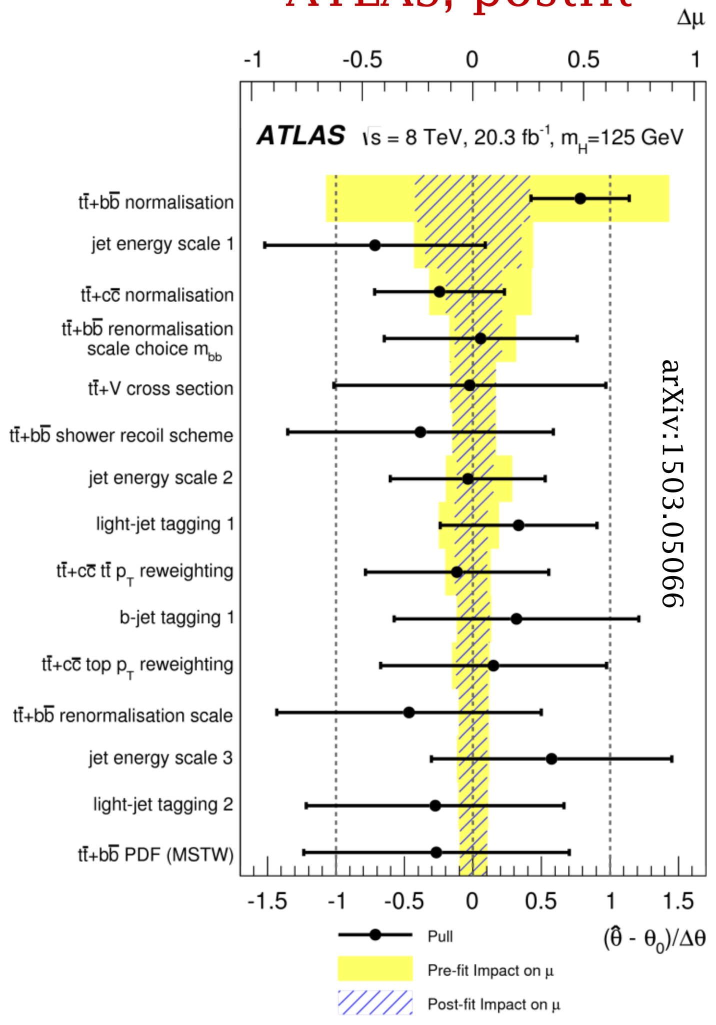
	ttH(bb) lepton+jets			
	4j, 4b	5j, ≥4b	≥6j, 3b	≥6j, ≥4b
ATLAS	2 (1.4%)	6.2 (2.5%)	40 (1%)	16 (4%)
CMS	1.5 (1.2%)	4.4 (2.3%)	18.9 (0.8%)	6.7 (2.6%)

	ttH(bb) di-lepton		
	≥4j, 3b	≥4j, ≥4b	≥3b
ATLAS	8.6 (1.5%)	2.7 (5.9%)	-
CMS	-	-	10 (1.3%)

Should check the interplay between all these factors
Expect better acceptance overall mainly due to b-tagging

Important Issues

ATLAS, postfit



CMS, prefit uncertainties

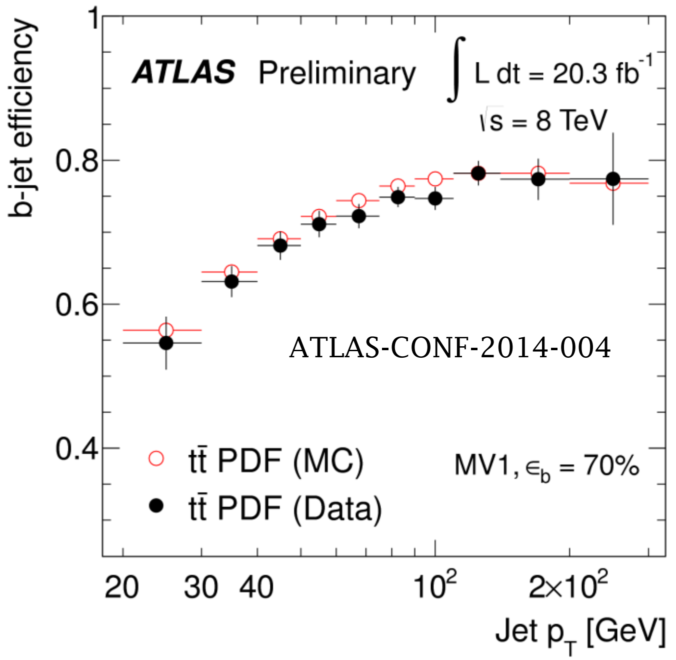
Source	Rate uncertainty	Shape	Process $t\bar{t}H$	Process $t\bar{t}$+jets	Process Others
Experimental uncertainties					
Integrated luminosity	2.6%	No	✓	✓	✓
Trigger and lepton identification	2–4%	No	✓	✓	✓
JES	4–13%	Yes	✓	✓	✓
JER	0.5–2%	Yes	✓	✓	✓
b tagging	2–17%	Yes	✓	✓	✓
Theoretical uncertainties					
Top p_T modelling	3–8%	Yes		✓	
μ_R/μ_F variations	2–25%	Yes		✓	
$t\bar{t}+b\bar{b}$ normalisation	50%	No		✓	
$t\bar{t}$+b normalisation	50%	No		✓	
$t\bar{t}+c\bar{c}$ normalisation	50%	No		✓	
Signal cross section	7%	No	✓		
Background cross sections	2–20%	No		✓	✓
PDF	3–9%	No	✓	✓	✓
Statistical uncertainty (bin-by-bin)	4–30%	Yes	✓	✓	✓

- Main systematic uncertainties:
  - tt+jets (bb) background
    - Profit from significant MC development and run I measurements (following slide)
  - btagging
    - Many developments, will probably improve in run II (following slide)
  - JES
    - Will probably not do that much better for JES in run II
    - However one might gain due to the harder jet pT spectrum at 13 TeV
    - Test in-situ constraint for channels with no neutrinos (maybe also 1ν)

b-tagging

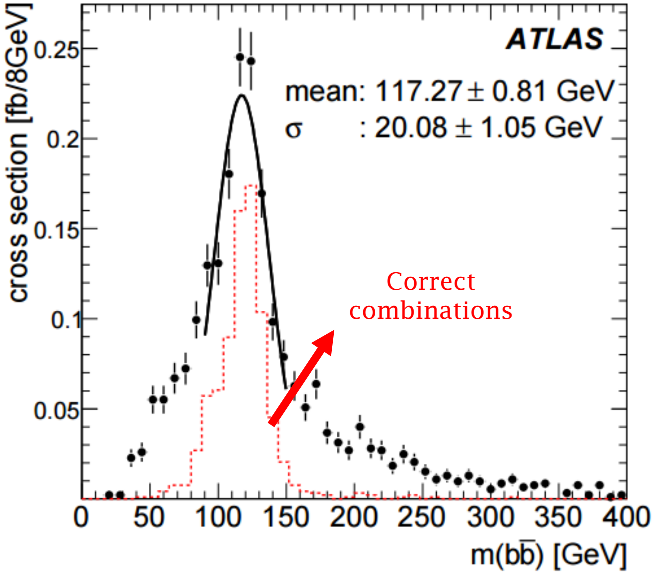
- b-tagging of course crucial for ttH(bb) analysis
- tt+b dominant background when requiring 4b-jets
  - However tt+c and tt+light are large
- Many improvements in the past few years
  - MVAs combining different b-tag techniques
  - Dedicated taggers to identify/reject c-jets
  - B-tagging in dense/boosted environment
  - bb-jets taggers
  - Not everything exploited by ATLAS and CMS for ttH(bb)
- Expect further improvement in b-tagging algorithms in run II
- Large effect expected from increased jet pT when moving to 13 TeV
  - Also improves b-tagging related systematics
- Pileup will slightly degrade b-tagging performance
- New pixel layer for ATLAS
  - Recover losses due to detector aging and to pileup
  - Expect to even improve b-tagging with respect to run I
- All together expect a significant increase of btagging performance in run II
  - Both experiment use the 70% efficiency working point (24% if requiring 4b-jets)
  - Can work at the same operating point and greatly reject tt+light
  - Or keep tt+light efficiency and increase signal
    - ➤ **33% relative signal efficiency increase with only 5% btag efficiency increase**

		4j, 4b	5j, ≥4b	≥6j, 3b	≥6j, ≥4b
ATLAS	tt+b	32%	45%	23%	60%
	tt+c	15%	18%	20%	19%
	tt+l	38%	25%	48%	12%
CMS	tt+b	30%	38%	22%	46%
	tt+c	25%	28%	30%	32%
	tt+l	41%	28%	45	17%

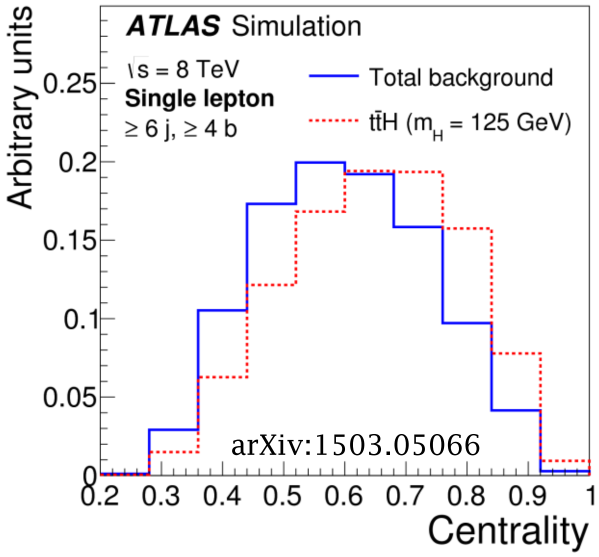


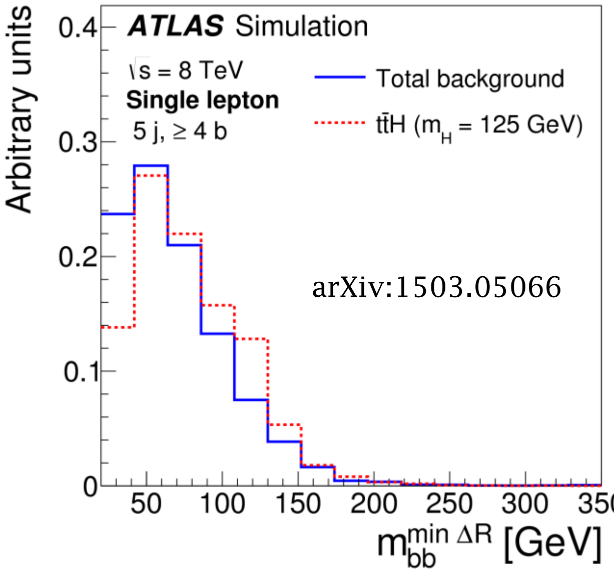
Signal Reconstruction

ATLAS simulation 14TeV (2009)
arXiv:0901.0512



- Both ATLAS and CMS did not explicitly reconstruct the ttH system in run I
  - Very hard due to large number of possible combinations
- Matrix Element Methods (MEM) are used
  - This is somehow close to reconstructing the ttH system
- Only global variables are used to separate the ttH signal from the tt+jets (bb) background
  - H→bb vs continuum bb diluted with jets from ttbar



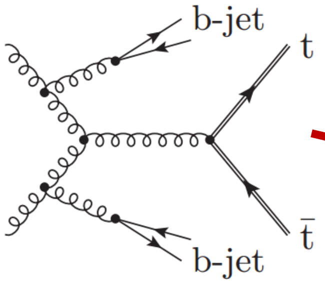


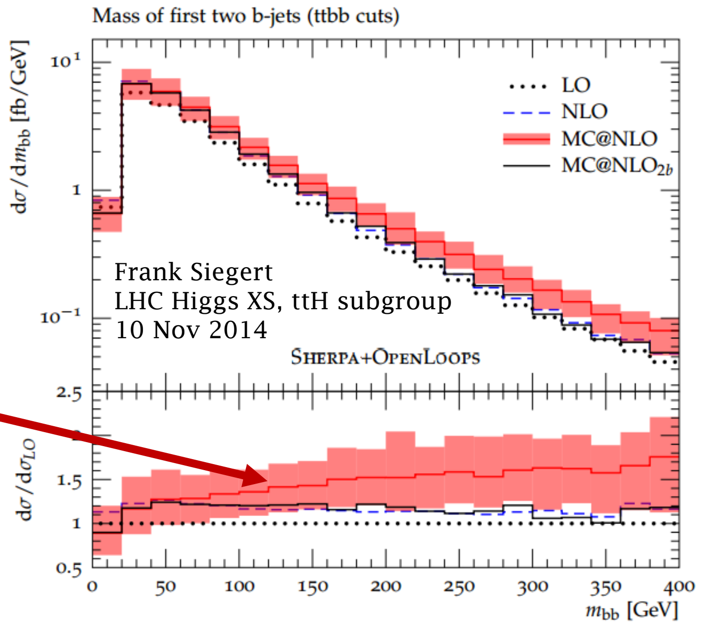
- One should try more powerful techniques to "partially" reconstruct ttH
  - Especially for the lepton+jets and all-hadronic channels
  - More favorable kinematics at 13 TeV
    - missing less jets from ttH decay
  - Will probably never reach high purify but can still be beneficial
  - Can use mass constraint to limit effects from JES systematics

tt+jets (bb) Background

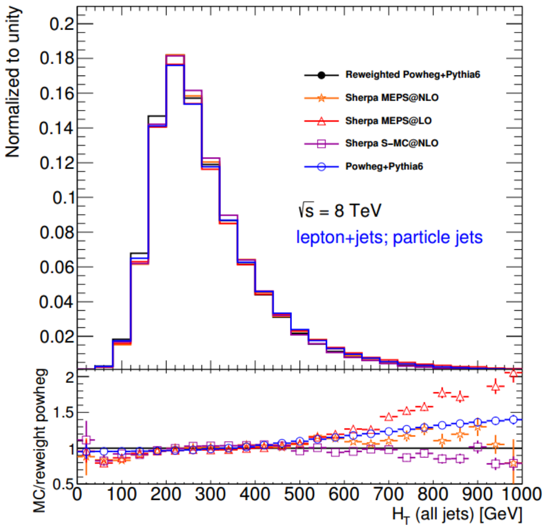
- tt+jets especially tt+HF very hard to model
  - tt+0,1,2 jets @NLO merged with tt+3,4 jets @LO and matched with parton shower
    - MD5_aMC@NLO and Sherpa MEPS@NLO
  - ttbb available at NLO+PS
    - MC@NLO, Sherpa+OpenLoops
  - Unprecedented accuracy for tt+jets, tt+bb
- Already using run I measurements to check new generators
- Still not clear what is best for ttbb
  - Very hard to find a clean control region
  - Inclusive as part of tt+jets
  - Exclusive at NLO (not sure how to merge with above)
  - Massive vs massless b – 4FNS vs 5 FNS
  - …
  - Little constraint from data
    - Limited by statistics in run I
  - Detailed checks are being performed to understand this background

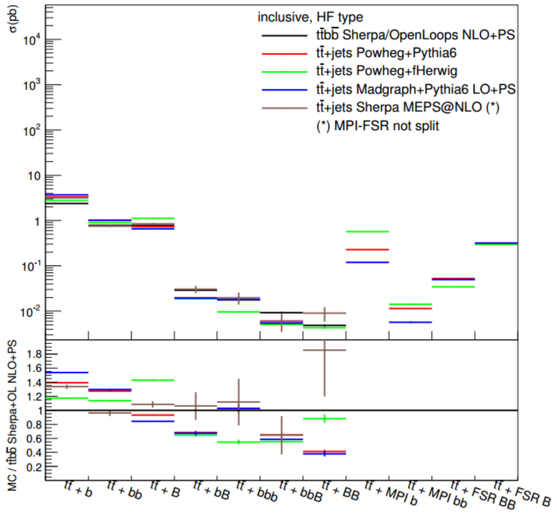
Double collinear g→bb production
Only available at NLO+PS accuracy





ATL-PHYS-PUB-2014-022

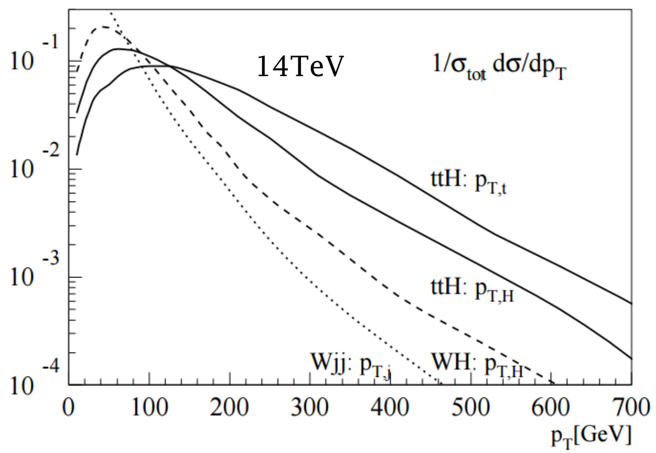


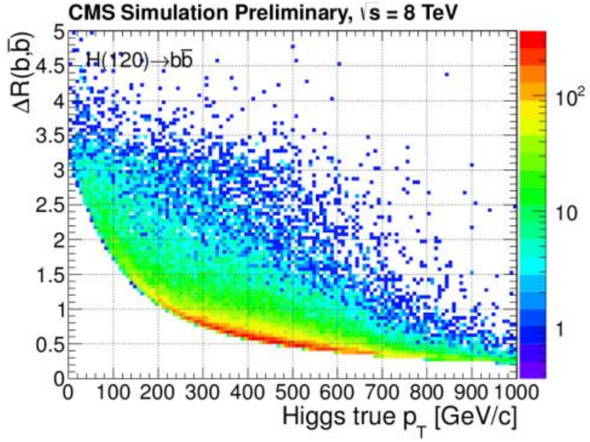


Boosted ttH

- Going to higher top/Higgs pT to increase S/B
  - Merged jets from decay products from boosted objects
- However very low statistics at high pT
  - Not sure if profitable in run II
- Actually using “semi-resolved” regime
  - jets are not really merged but very close

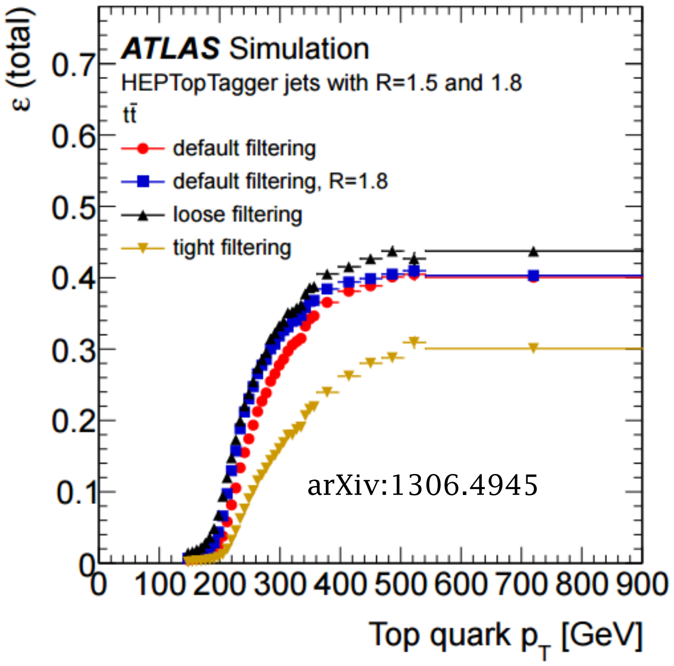
arXiv:0910.5472





	>200 GeV	>400 GeV
Higgs	14%	1.5%
top	31%	5%

Boosted fractions @14 TeV



- Great development for tagging boosted heavy object during run I
- Main idea: building large jets and using properties of sub-jets to tag boosted W, top and Higgs
  - Variety of techniques to “cleanup” the large jet
  - Successfully used in searches for heavy resonance
- Advantages with respect to a ttH(bb) resolved analysis still to be proven in run II

ttH, Multilepton Channels

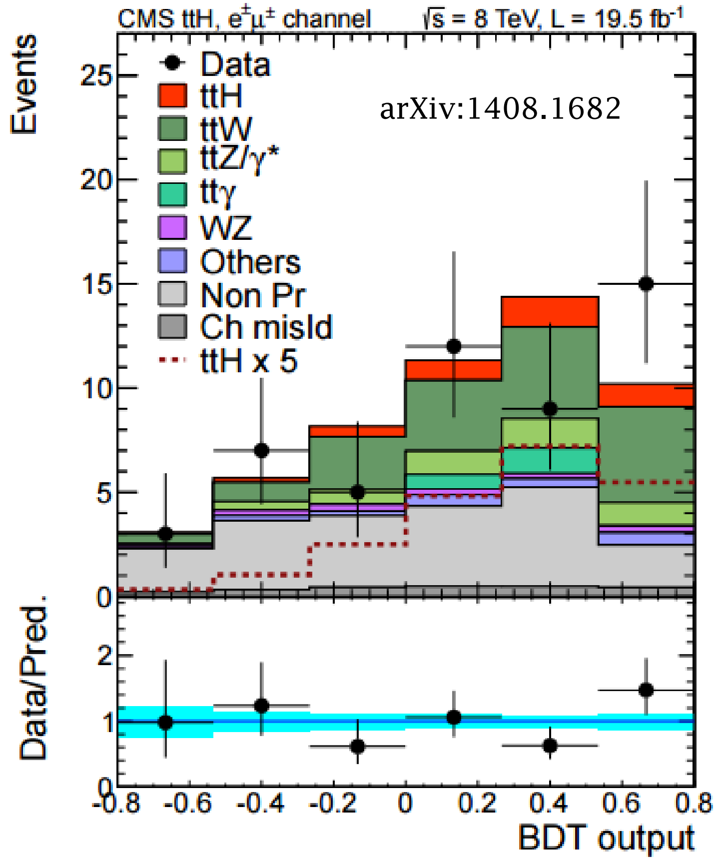
Selection and Signal Extraction

Run I expected signal events (S/B [%])

	2l(ss)	3l	4l	$\tau_h\tau_h$	$2l1\tau_h$
ATLAS	6.6 (8.7%)	2.3 (20%)	0.2 (29%)	0.68 (4.2%)	0.47 (34%)
CMS	8.8 (9.3%)	4.7 (7.9%)	0.54 (26%)	0.8 (0.7%)	-

- Various factors in run II that will improve/worsen ttH(multi-lepton)
  - Increased trigger thresholds
    - Lower signal acceptance
  - Increased pileup (not at the beginning– 25ns/50ns)
    - ➢ Increase fake rate
  - Larger signal lepton pTs at 13 TeV
    - ➢ Efficiency up
    - ➢ Fakes down
    - ➢ Charge mis-id down
- S/B will depend on the interplay of these factors and the design of the analysis cuts

- Shape analysis will greatly benefit from stat increase in run II
  - CMS already exploiting shapes with a BDT analysis
  - ATLAS using cut and count in run I



Limiting Factors

- Main uncertainties are due to
  - fake ID background
  - ttV backgrounds
  - ttH theoretical uncertainties
  - JES
- Muon and electron fake rate
  - Expected to improve with larger stat in control regions
    - Stat is still a limiting factor
- ttV cross section and acceptance
  - Improve by designing control regions for these backgrounds
    - Possible with more stat?
  - Use state-of-the-art NLO+PS generators
    - ttV+0,1 jets @NLO + ttV+2jets @LO matched with parton shower available
      - Sherpa and MG5_aMC@NLO

ATLAS, postfit source of uncertainties

Source	$\Delta\mu$	
$2\ell 0\tau_{\rm had}$ non-prompt muon transfer factor	+0.38	−0.35
$t\bar{t}W$ acceptance	+0.26	−0.21
$t\bar{t}H$ inclusive cross section	+0.28	−0.15
Jet energy scale	+0.24	−0.18
$2\ell 0\tau_{\rm had}$ non-prompt electron transfer factor	+0.26	−0.16
$t\bar{t}H$ acceptance	+0.22	−0.15
$t\bar{t}Z$ inclusive cross section	+0.19	−0.17
$t\bar{t}W$ inclusive cross section	+0.18	−0.15
Muon isolation efficiency	+0.19	−0.14
Luminosity	+0.18	−0.14

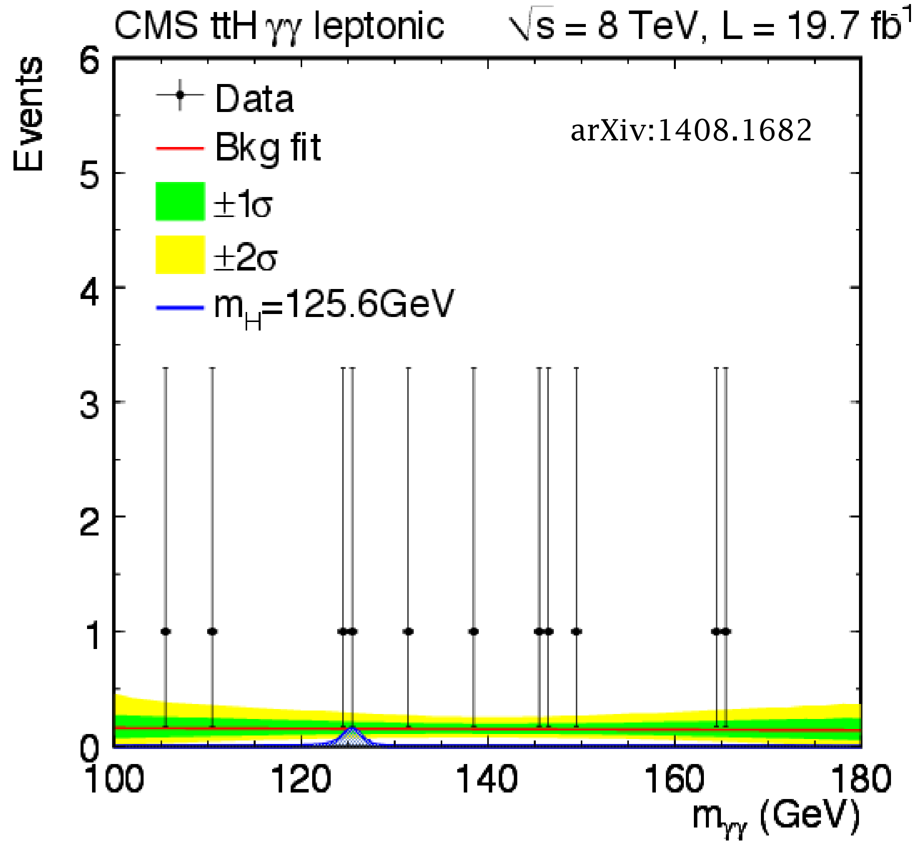
The order/effect of systematics will change once high purity channels (4l, 2l1τ_h) have enough statistics to be competitive with the 2l and 3l channels

$t\bar{t}H$, $H\rightarrow\gamma\gamma$

Main Issues

Expected events run I	ATLAS	CMS
ttH($\gamma\gamma$)	1.9	1.2

- Channels with top coupling at production and decay level
  - Still have some degeneracy between the top Yukawa coupling and an eventual new particle coupling to the Higgs boson
- Analysis based on an invariant mass fit to extract the $\gamma\gamma$ resonance
  - Of course can't do much with ~2 expected events
  - Analysis fully dominated by statistical uncertainties
  - Will greatly benefit from increased statistics in run II
  - Will probably become competitive with other channels at the end of run II
  - Will supersede other channels in HL-LHC



Back-of-envelope projections (since this channel is stat dominated):
With 16 times more signal, can reach a limit near 1 at the end of run II (now it is around 5)

Conclusion

- Top-Higgs coupling is special
  - Is it a gate for new physics?
- ttH needed to search for new particles in gluon/photon coupling to the Higgs boson through loops
  - This channel will be accessible for the first time at LHC run II
- ttH channels are complex channels with low cross sections
  - Challenging analyses
- A lot of progress in run I including many channels
- Very promising in run II
  - Increased cross section at 13 TeV
  - Increased luminosity
  - Hopefully better analysis techniques with a better understanding of the detectors and the tt+X backgrounds

Backup

Projections after run II

Table 1-14. *Expected per-experiment precision of Higgs boson couplings to fermions and vector bosons with 300 fb^{-1} and 3000 fb^{-1} integrated luminosity at the LHC. The 7-parameter fit assumes the SM productions and decays as well as the generation universality of the couplings ($\kappa_u \equiv \kappa_t = \kappa_c$, $\kappa_d \equiv \kappa_b = \kappa_s$ and $\kappa_\ell \equiv \kappa_\tau = \kappa_\mu$). The precision on the total width Γ_H is derived from the precisions on the couplings. The range represents spread from two assumptions of systematic uncertainties, see text.*

Luminosity	300 fb^{-1}	3000 fb^{-1}
Coupling parameter	7-parameter fit	
κ_γ	$5-7\%$	$2-5\%$
κ_g	$6-8\%$	$3-5\%$
κ_W	$4-6\%$	$2-5\%$
κ_Z	$4-6\%$	$2-4\%$
κ_u	$14-15\%$	$7-10\%$
κ_d	$10-13\%$	$4-7\%$
κ_ℓ	$6-8\%$	$2-5\%$
Γ_H	$12-15\%$	$5-8\%$
	additional parameters (see text)	
$\kappa_{Z\gamma}$	$41-41\%$	$10-12\%$
κ_μ	$23-23\%$	$8-8\%$
$\mathrm{BR}_{\mathrm{BSM}}$	$< 14-18\%$	$< 7-11\%$

arXiv:1310.8361
(Snowmass)