Search for the Higgs Boson in the ttH production mode using the ATLAS detector

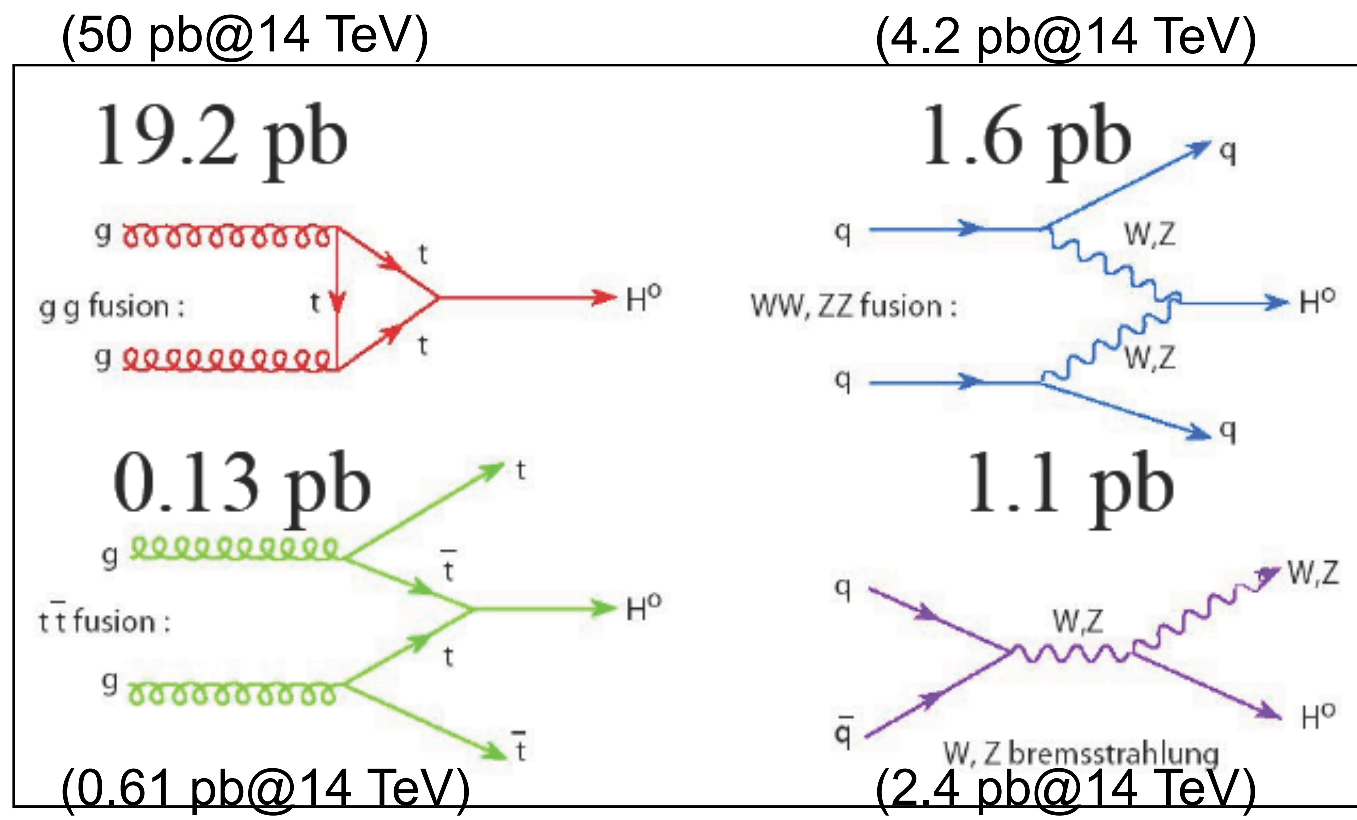
G.Salamanna (Universita' and INFN - Roma Tre)
for the ATLAS Collaboration

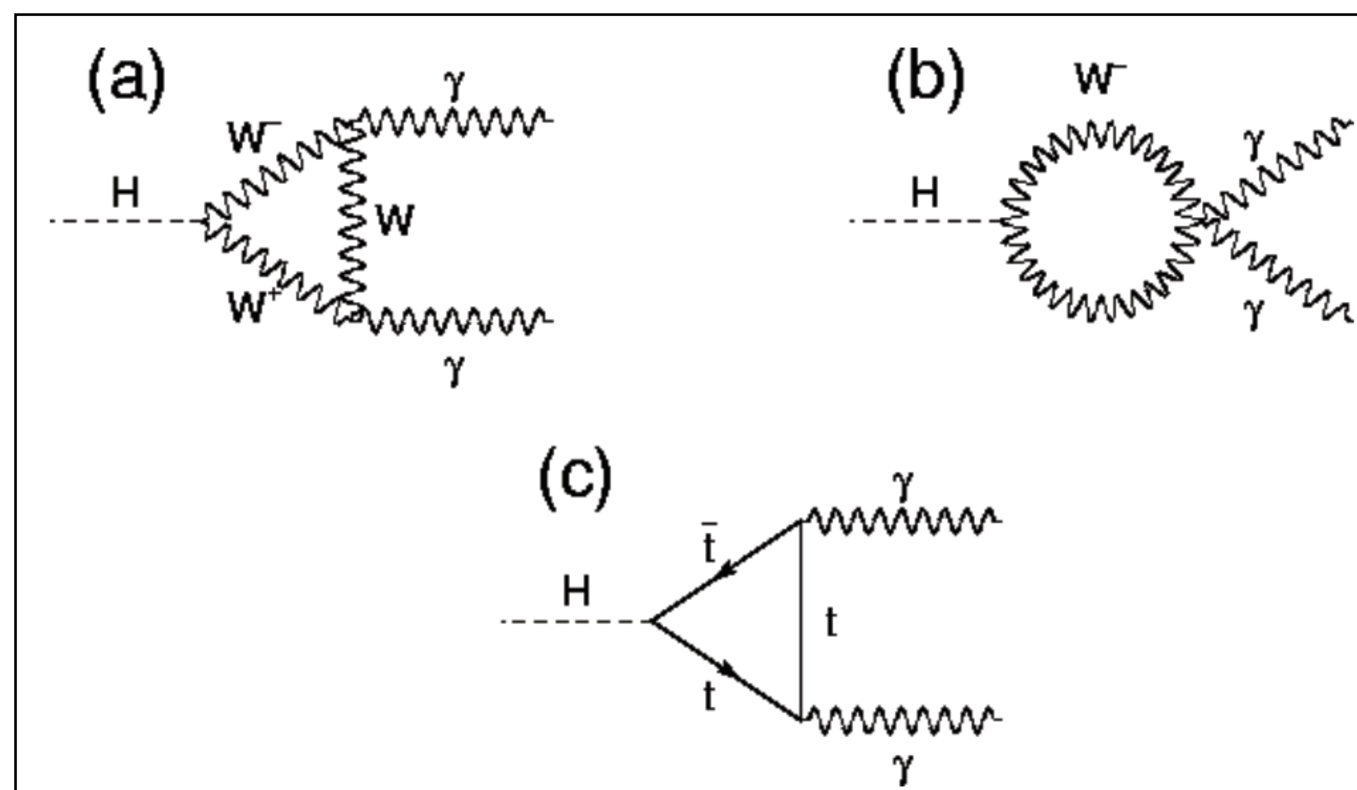
Outline

- General framework, motivations
- di-photon channel
  - interlude on t+H and WtH
- bb channel
- Prospects
- Conclusions

Higgs Boson coupling with top quarks

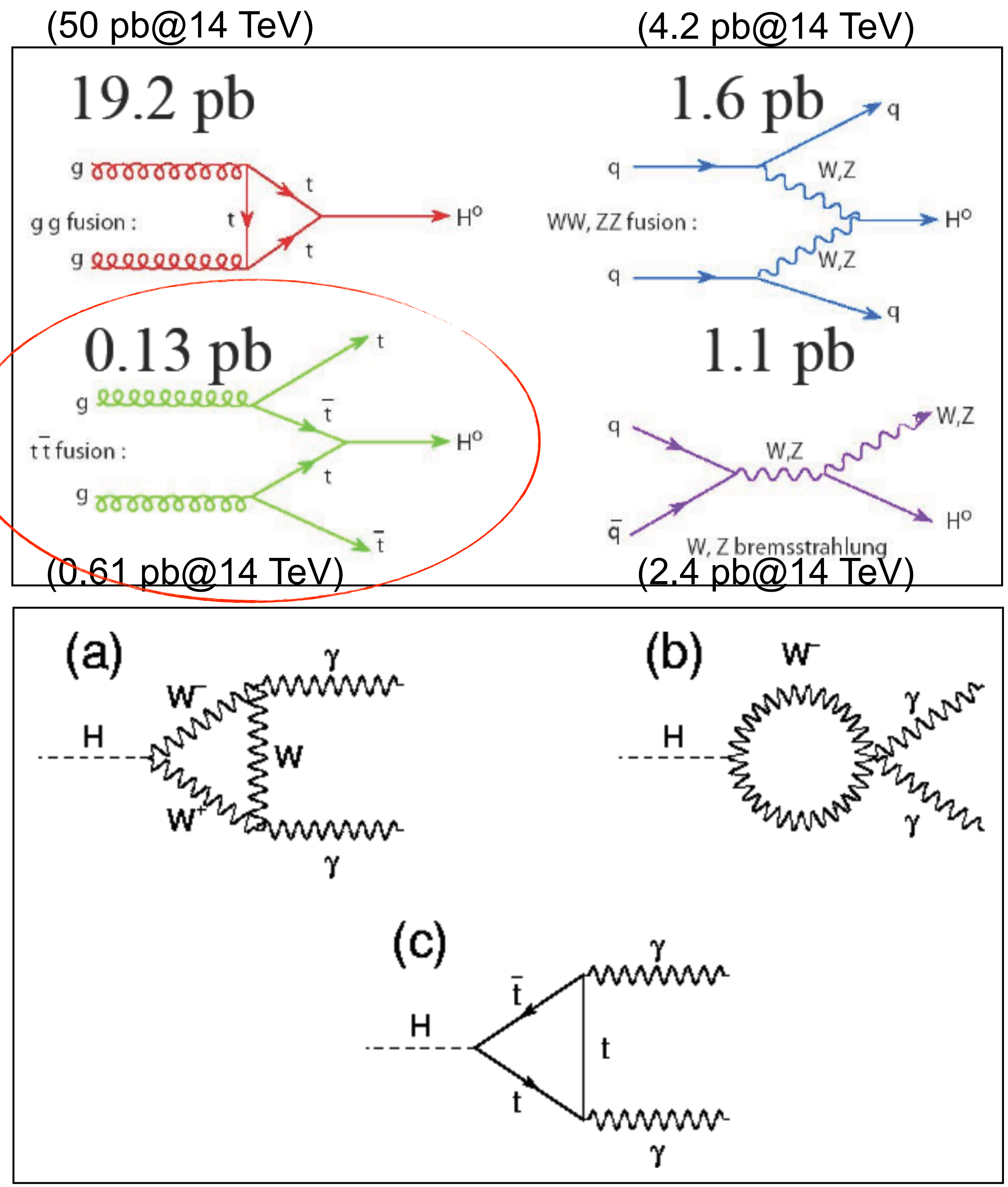
- t-H dominates gg fusion mechanism
  - but through loop process where...
  - ...new physics could also hide
- (t-H contributes to H→γγ, but also in this case it's not the only component)



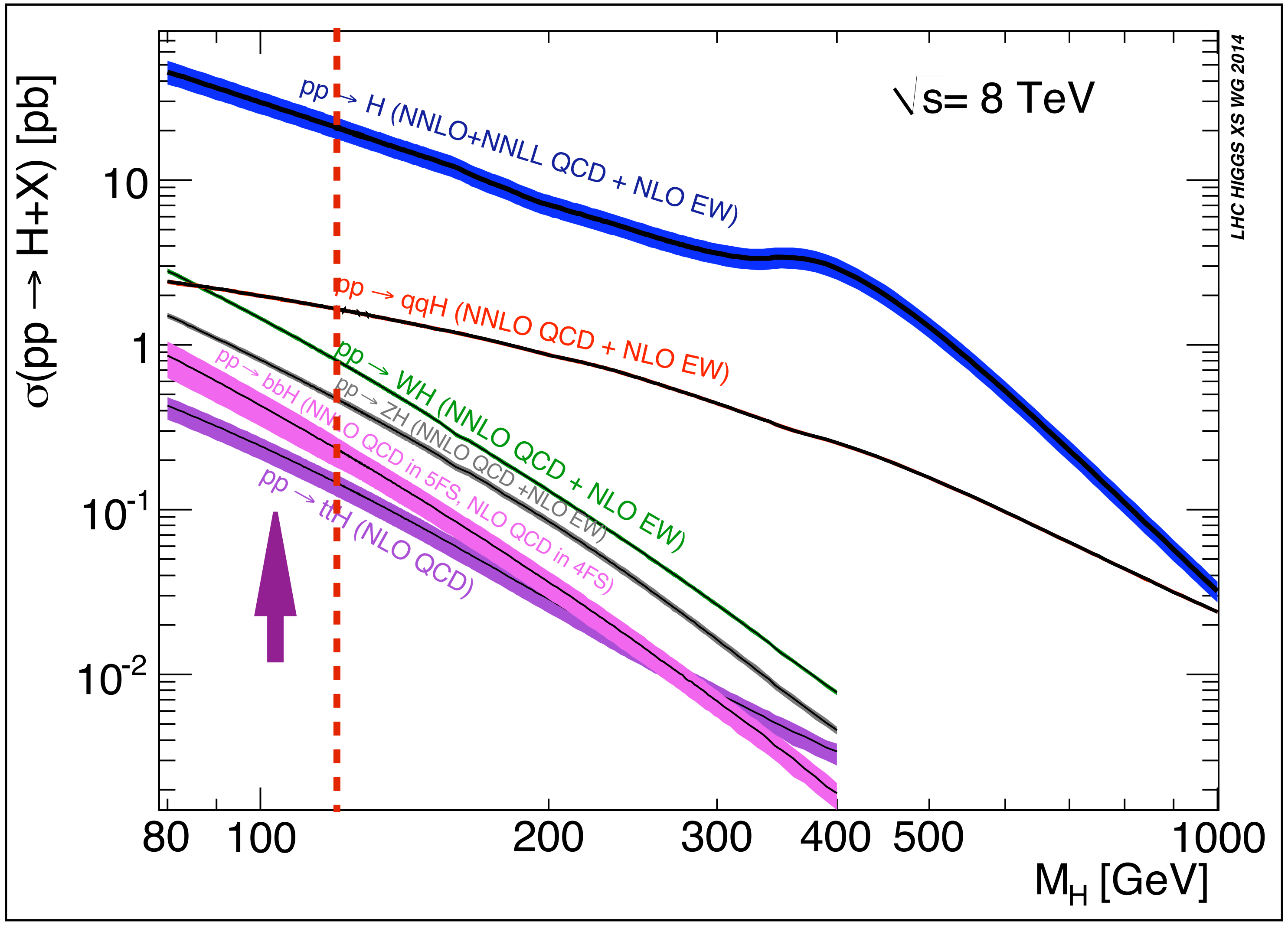


Higgs Boson coupling with top quarks

- t-H dominates gg fusion mechanism
  - but through loop process where...
  - ...new physics could also hide
- (t-H contributes to H→γγ, but also in this case it's not the only component)
  - By measuring ttH we have <u>direct</u> access to the coupling at production instead
    - comparing measured coupling with gg fusion, can unveil new physics



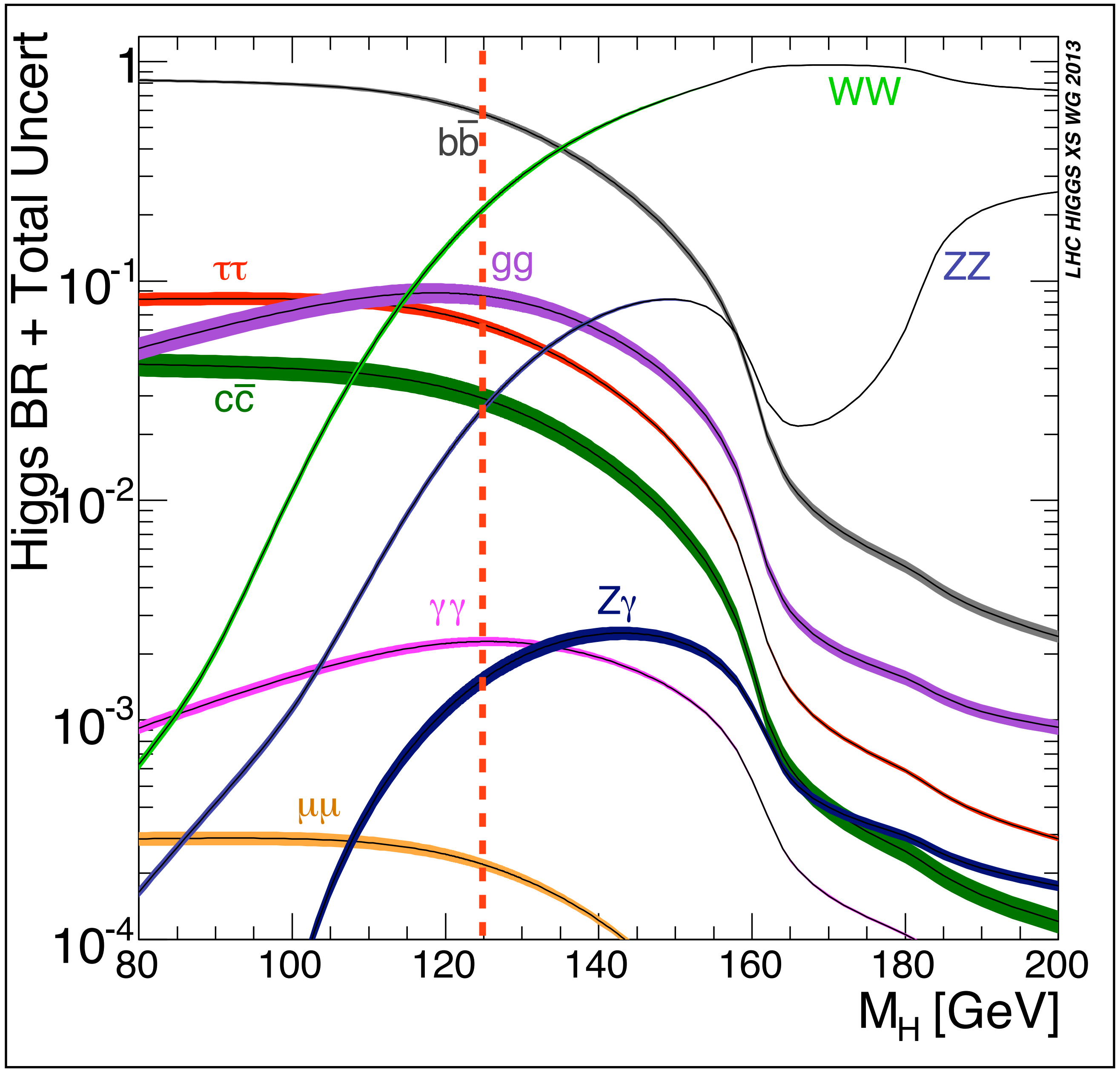
SM cross-section at 8 TeV



- At NLO, σ_{SM}=130 fb for m(H) =125 GeV@ 8 TeV
- @ 14 TeV ~5x increase in σ_{SM}

- Translates into ~ 2600 events *before detector acceptance and selections* with L = 20.3 fb^{-1} collected by ATLAS at 8 TeV
- Main background for analyses in this talk: tt+jets (light or heavy flavour)
  - cross section at 8 TeV: 253 pb (*i.e.* ~2 10^3 times the signal)

Channels considered



- BR(H->γγ) < 0.23%
- BR(H->bb) = 58%
- BR(H->WW+ZZ+ττ) = 30%

- 1 channel
- 2 channels (according to top quark decay mode)
- several channels (according to light lepton and tau multiplicities)

- γγ: clear resonance peak but very scarce
- bb: by far most abundant, but overwhelmed by tt+HF background and less easy bb reconstruction
- multi-lepton channels: good compromise, but sensitive to additional tt+W/Z backgrounds hard to control with data (→ still in progress)

Event generation

- Signal ttH: Helac-One Loop+Powheg interface to parton shower (=”PowHel”)
  - inclusive in Higgs boson decays, cross-section and BR from http://arxiv.org/abs/1101.0593 and updates
  - CT10NLO Parton Distribution Function (PDF)
  - Pythia 8 for parton shower (PS) + CTEQ6L1 PDF
- (W)tH: MadGraph5_AMC@NLO, 5-flavour scheme
  - inclusive in Higgs boson decays, xsec and BR from Yellow Book
  - three different values of k_t = -1, 0, + 1.
  - CT10NLO PDF
  - Herwig++ for parton shower + CTEQ6L1 PDF
- tt+jets: Powheg
  - inclusive in flavour of additional partons
  - CT10NLO PDF
  - Pythia 6 for parton shower + CTEQ6L1 PDF

- Other sources of background in back-up

ttH($\gamma\gamma$)

http://arxiv.org/abs/1409.3122

Overall strategy

- Data collected at 7 (4.5 fb^{-1}) and 8 (20.3 fb^{-1}) TeV
- 2 isolated, high-p_T photons for Higgs boson mass reconstruction
- Categorize events according to top quark decay:
  - Optimized on the expected limit on μ_{ttH}
  - leptonic channel: ≥ 1 leptons (e or μ), ≥ 1 b-tagged jet
  - hadronic channel: ≥ 6 jets, ≥ 2 b-tagged jet
  - Combined signal selection: eff ~ 15%, purity ~80%

Category	N_H	ggF	VBF	WH	ZH	$t\bar{t}H$	$tHqb$	WtH	N_B
7 TeV leptonic selection	0.10	0.6	0.1	14.9	4.0	72.6	5.3	2.5	$0.5^{+0.5}_{-0.3}$
7 TeV hadronic selection	0.07	10.5	1.3	1.3	1.4	80.9	2.6	1.9	$0.5^{+0.5}_{-0.3}$
8 TeV leptonic selection	0.58	1.0	0.2	8.1	2.3	80.3	5.6	2.6	$0.9^{+0.6}_{-0.4}$
8 TeV hadronic selection	0.49	7.3	1.0	0.7	1.3	84.2	3.4	2.1	$2.7^{+0.9}_{-0.7}$

absolute numbers

fractions

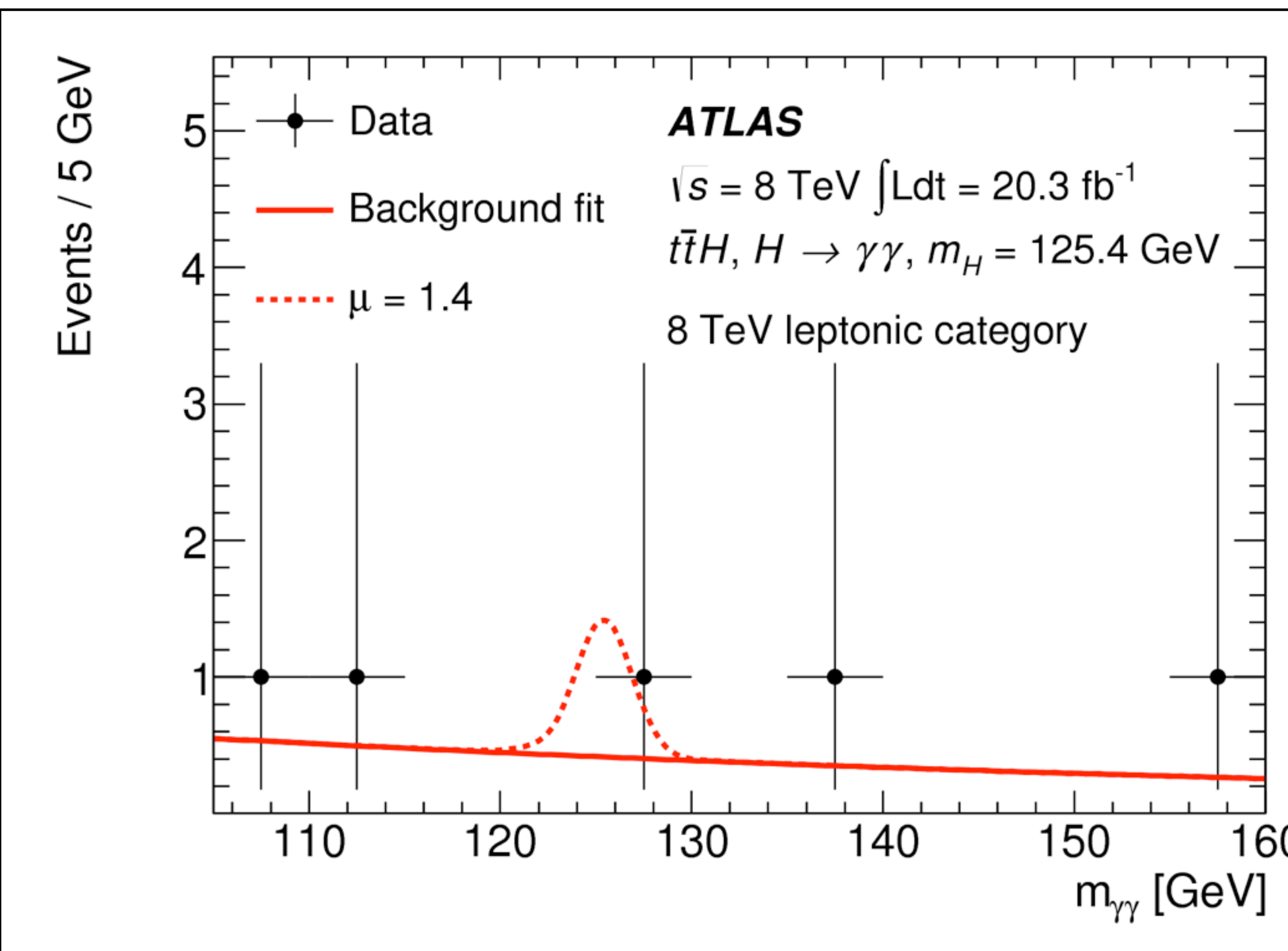
absolute numbers

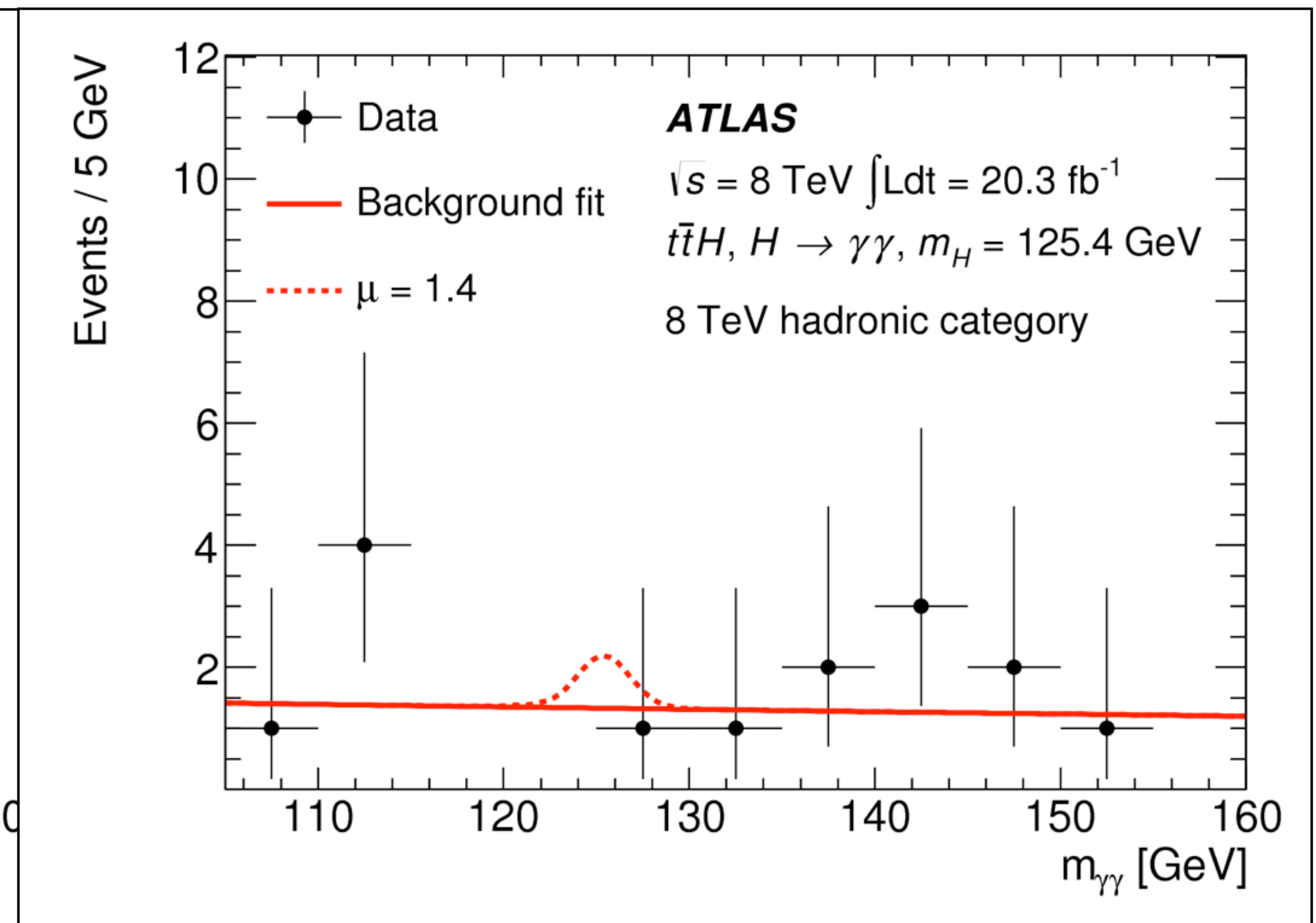
Pre-selections

- Photons passing quality criteria on shower shape and isolated both in tracking and calorimetry
  - 2 photons with a reconstructed vertex required
  - leading (subleading) photon required to have $E_T >$ 0.35 $m_{\gamma\gamma}$ (0.25 $m_{\gamma\gamma}$), and the di-photon mass to be between 105 GeV and 160 GeV ("Signal Region")
- e or μ of good quality (track, track-cluster match) and isolated both in tracking and calorimetry
  - $E_T(e) > 15$ GeV, $p_T(\mu) > 10$ GeV
- Anti-k_T jets with R=0.4, calibrated at hadronic scale
  - $p_T(\text{jet}) > 25$ GeV, central
  - pile-up suppressing selection criteria for jets of $p_T(\text{jet}) < 50$ GeV
- b-quark identification in jets with NN-based algorithm
  - 60, 70, 80% b-tag efficiency working points all used in this analysis

Analysis

- Localized excess looked for around $m_{\gamma\gamma}$=125.4 GeV
- Unbinned LL fit to background and signal in signal region
  - signal: Gaussian core portion and a power-law low-end tail + Gaussian (tails)
  - background: exponential function tested on ad-hoc control region (loosening photon ID) sensitive to jets faking γ



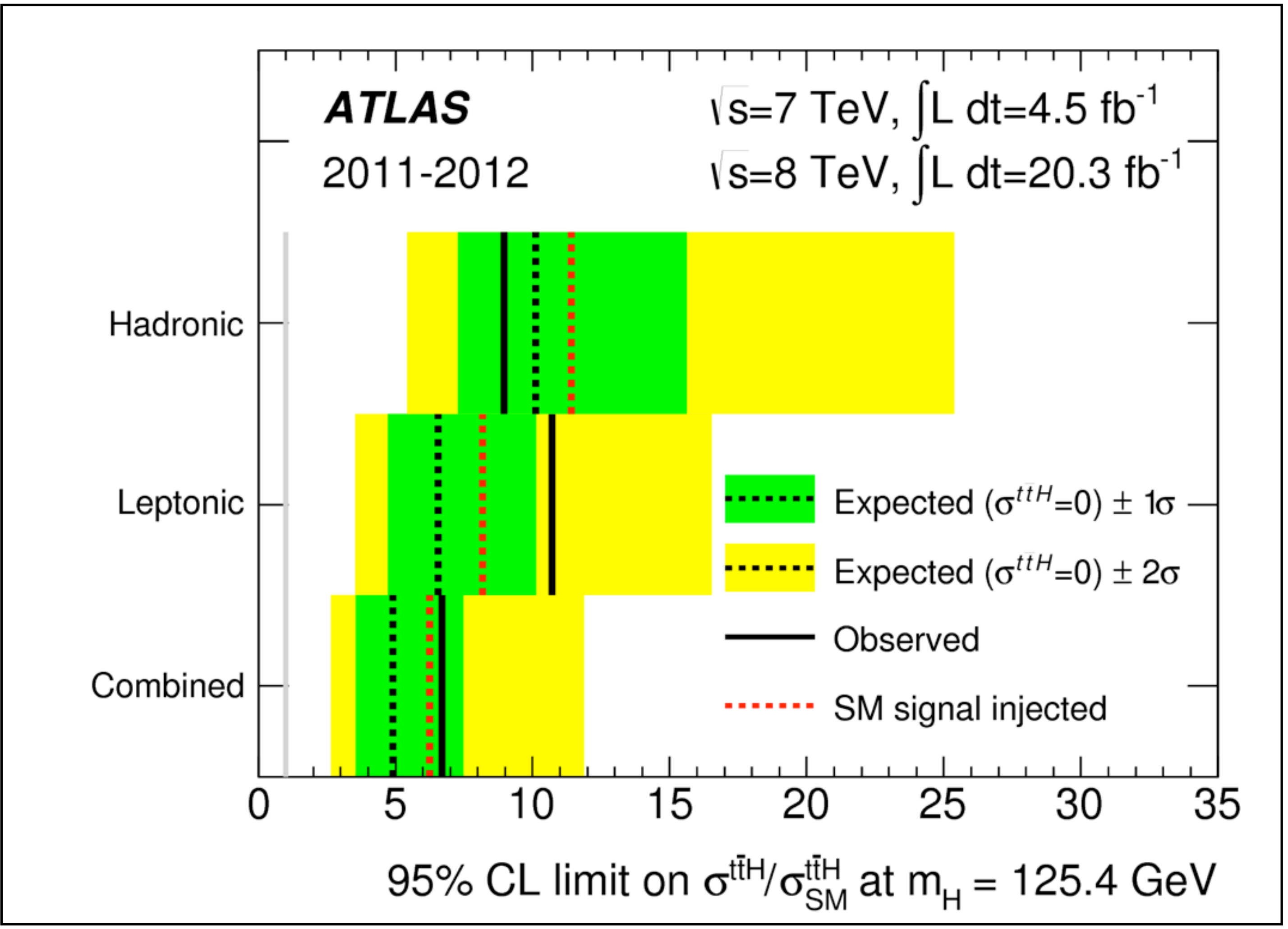


- $m_{\gamma\gamma}$ resolution dominated by resolution on E_γ

Results

http://arxiv.org/abs/1409.3122

$\mu_{ttH} = \sigma/\sigma_{SM}$



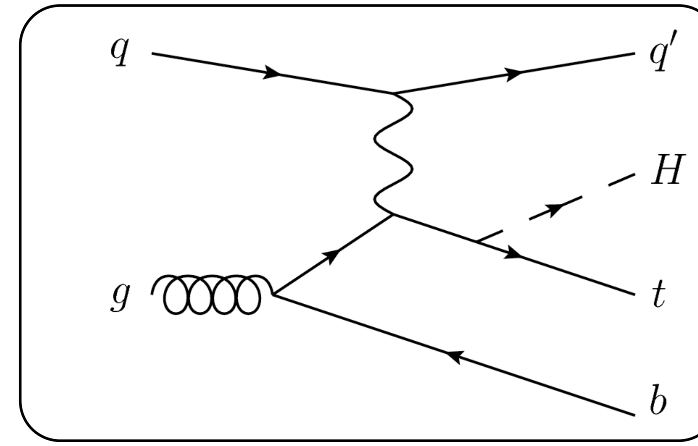
- m_H=125.4 GeV
- Observed limit on μ_{ttH} = 6.5
- Expected: 4.9

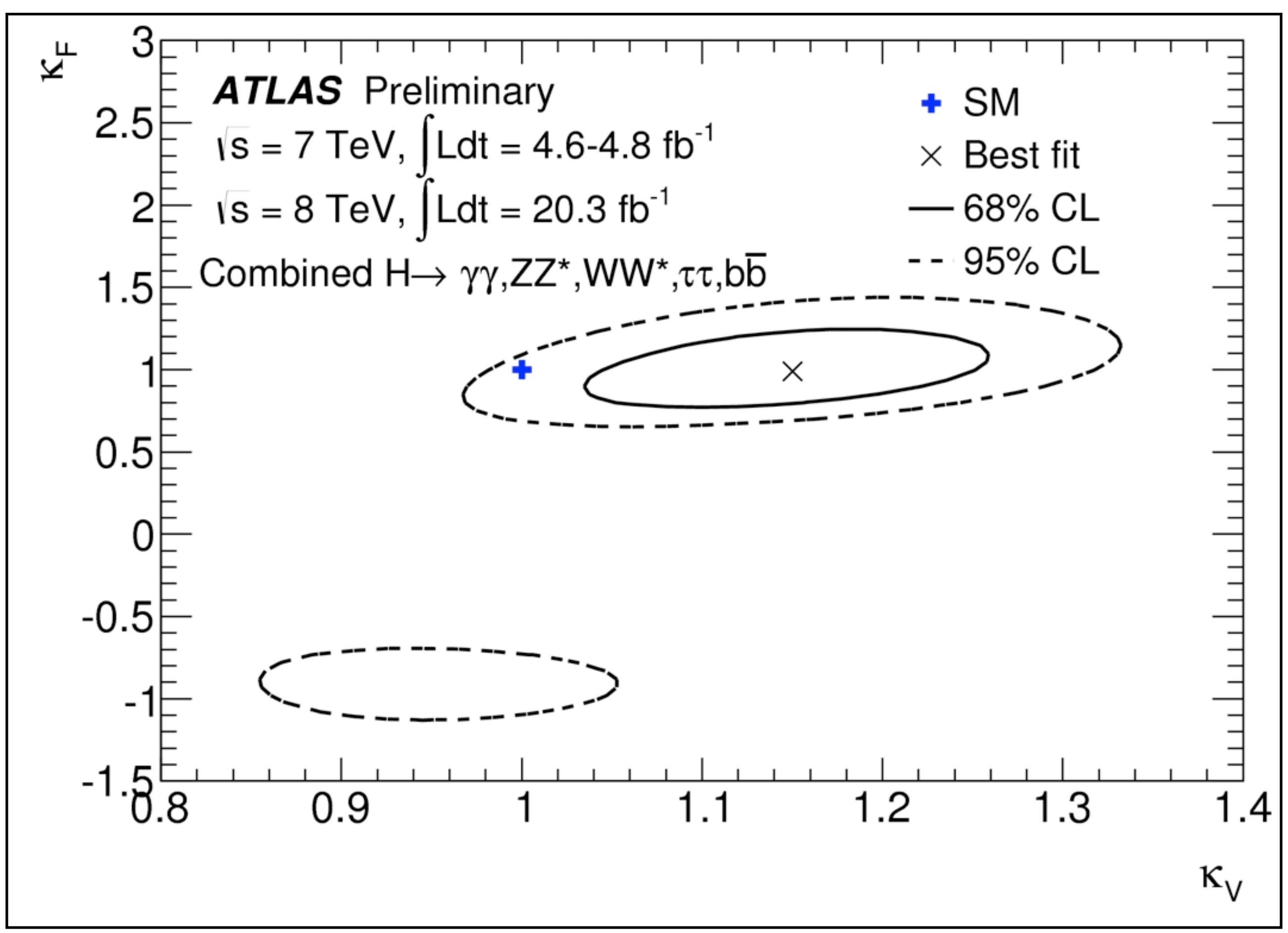
- Comparable impact of theory and experimental systematic uncertainties on final yield of events

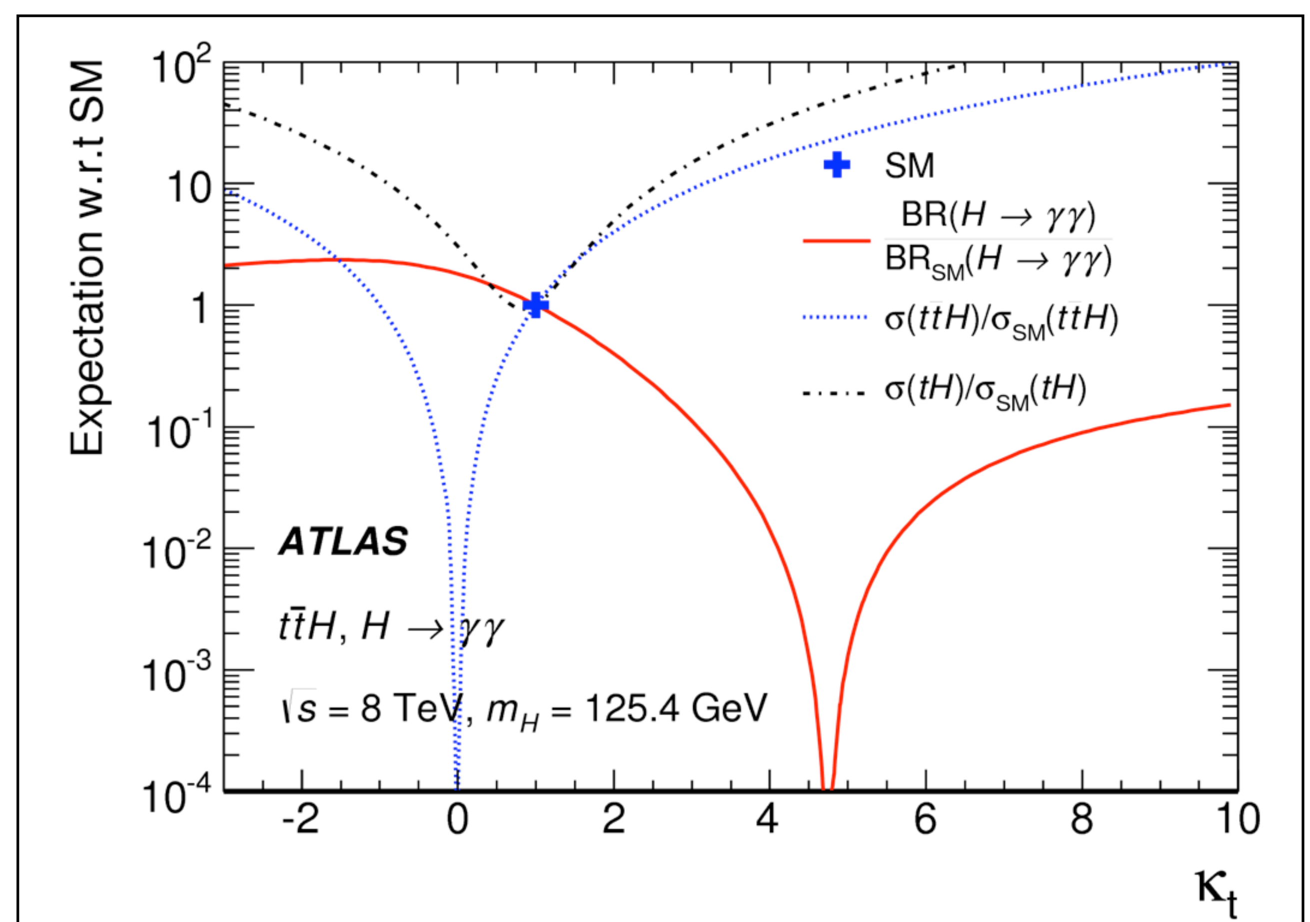
	$t\bar{t}H$ [%] had.	$t\bar{t}H$ [%] lep.	tHqb [%] had.	tHqb [%] lep.	WtH [%] had.	WtH [%] lep.	ggF [%] had.	WH [%] lep.
Luminosity	±1.8							
Photons	±10.0	±10.0	±10.0	±10.0	±10.0	±10.0	±10.0	±10.0
Leptons	< 0.1	±0.7	< 0.1	±0.7	< 0.1	±0.6	< 0.1	±0.7
Jets and E_T^{miss}	±9.1	±1.6	±19	±2.4	±13	±2.9	±30	±10
Bkg. modeling	0.12 evt.	0.01 evt.	applied on the sum of all Higgs boson production processes					
Theory ($\sigma \times$ BR)	+10,−13		+8,−7		+12,−12		+11,−12	+5.5,−5.5
MC Modeling	±11	±3.3	±12	±4.4	±13	±5.2	±130	±100

(W)tH with γγ channel

- Residual sign ambiguity between fermionic and couplings
- Single top + Higgs production probes this sign
  - SM has destructive interference between H emission from top and from W: if relative sign of top coupling flips, large constructive interference
  - Also affects BR(H-> γγ): double-sensitivity of this channel

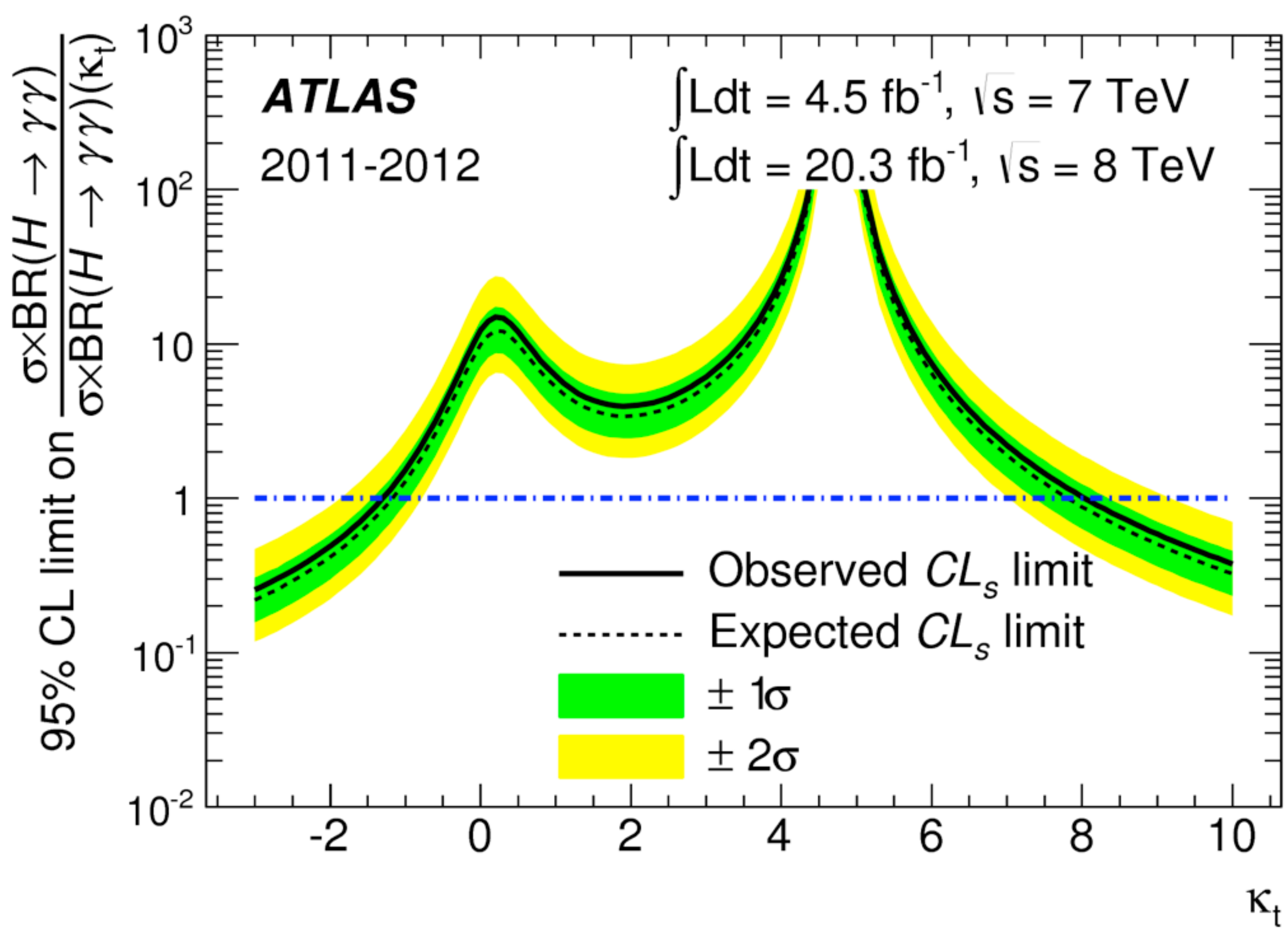






Analysis tHj+WtHj

- Exactly same analysis/samples as ttH
- scanning also limit in top-H Yukawa coupling k_t
  - tH+WtH selection efficiencies extrapolated from 3 benchmark k_t values/MC samples (variations up to 15/20%)



Observed range at 95% CL: [-1.3, 8.0]
expected: [-1.2, 7.8]

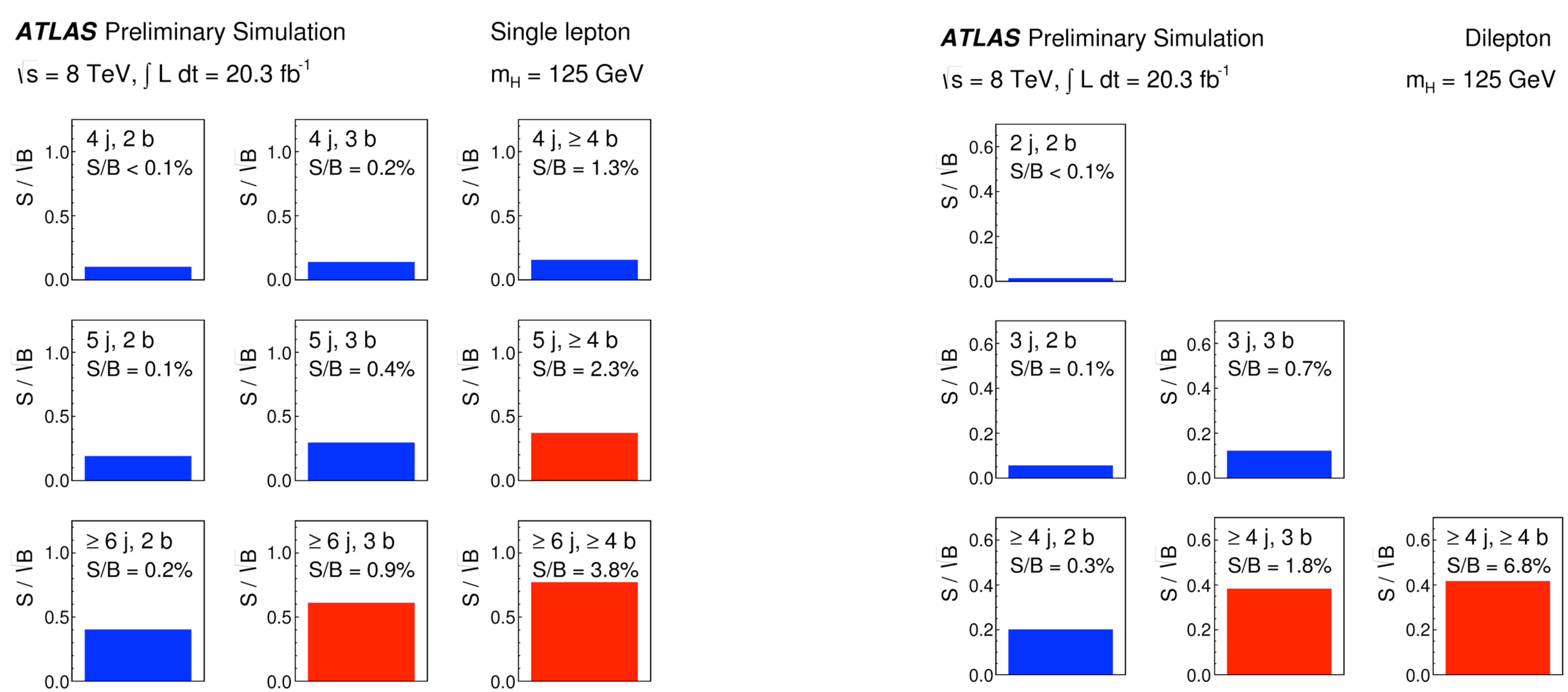
Will repeat in Run II

$t\bar{t}H(b\bar{b})$

ATLAS-CONF-2014-011

Goal and strategy

- Data collected at 8 (20.3 fb^{-1}) TeV
- Multi-variate analysis technique to reduce large bkg from top
  - but this needs attention with modeling of used variables
- Construct matrix of N(jets)-N(b-tags) to characterize B
  - simultaneous fit for N_S (from signal-enriched regions) and N_B (from control-enriched regions)



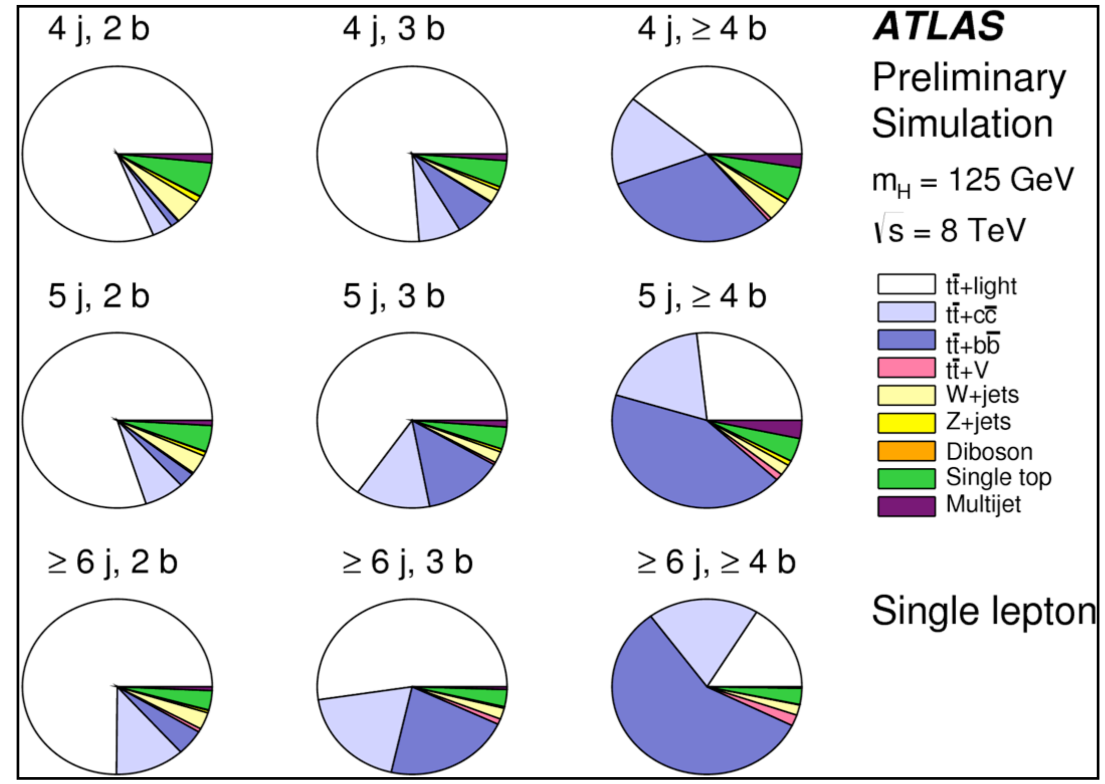
- dilepton channel: ee/μμ/eμ + (2, 3, ≥4 jets) and (2, 3, ≥4 b-tags)
- lepton+jets channel: 1 e or μ + (4, 5, ≥6 jets) and (2, 3, ≥4 b-tags)

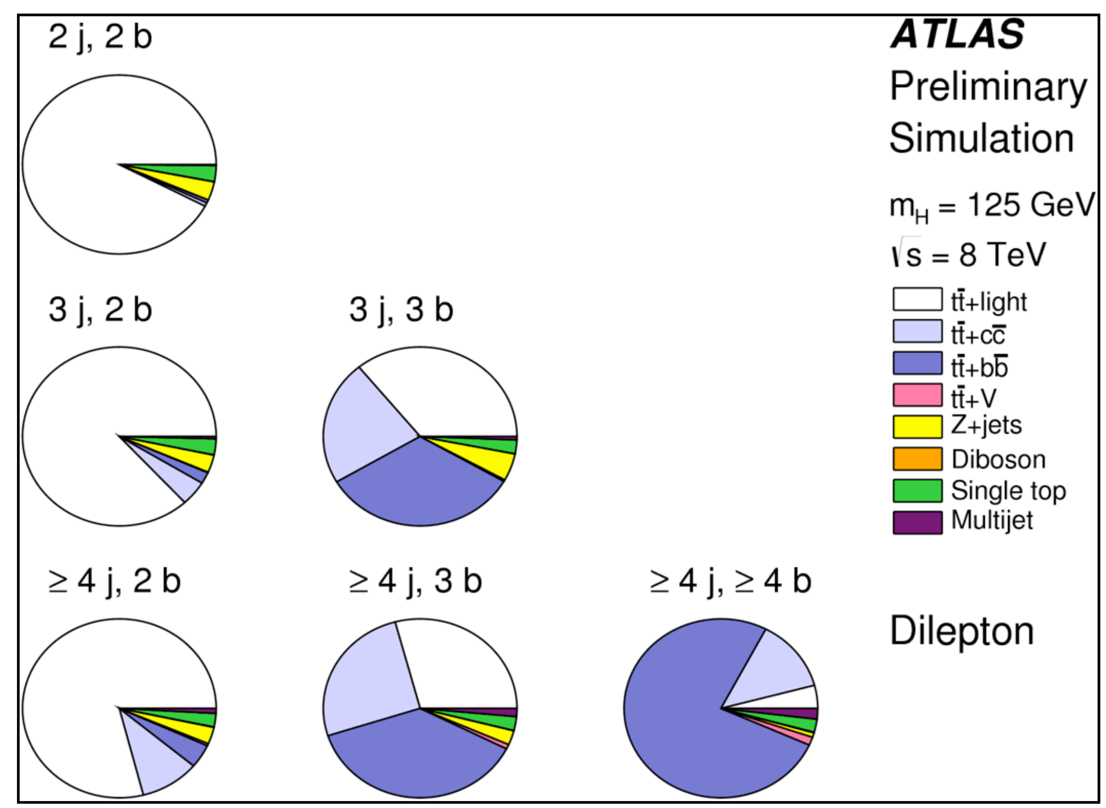
Pre-selections

- Single lepton trigger (OR of low-pt and high-pt to even-out plateau)
- e or μ of good quality (track, track-cluster match) and isolated both in tracking and calo
  - l+jets: $E_T(e)>25$ GeV, $p_T(\mu)>25$ GeV
  - dilepton: leading $E_T(e)/p_T(\mu)$ >25 GeV, subl $E_T(e)/p_T(\mu)$ >15 GeV
- Anti-k_T jets with R=0.4, calibrated at hadronic scale
  - p_T(jet)>25 GeV, central
  - pile-up suppressing cuts for jets of p_T(jet)<50 GeV
- b-quark identification in jets with NN-based algorithm
  - 70% b-tag efficiency working points, ~1% light-jet mistag rate

Background modeling

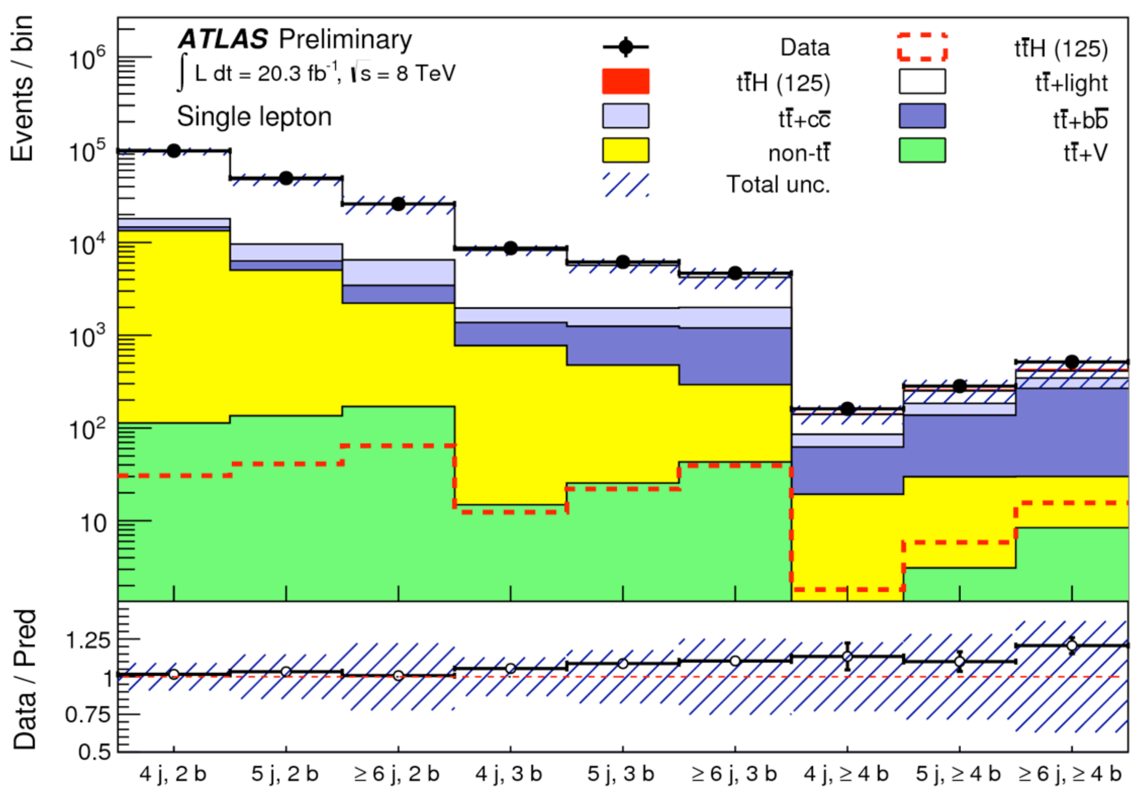
- Main bkg tt+HF in all regions for both channels
  - 50% normalization uncertainty on ttbb/cc
- Powheg + Pythia6 used to model it
  - HF content validated with dedicated studies (ATL-CONF-2013-099)
- Madgraph directly generates tt+bb/cc
  - expected to properly treat ME+PS matching of tt+HF
  - difference between generators taken as systematic uncertainty
- p_T of top quark and tt system reweighted to reproduce spectra obtained in 7 TeV analysis (ATL-CONF-2013-099).



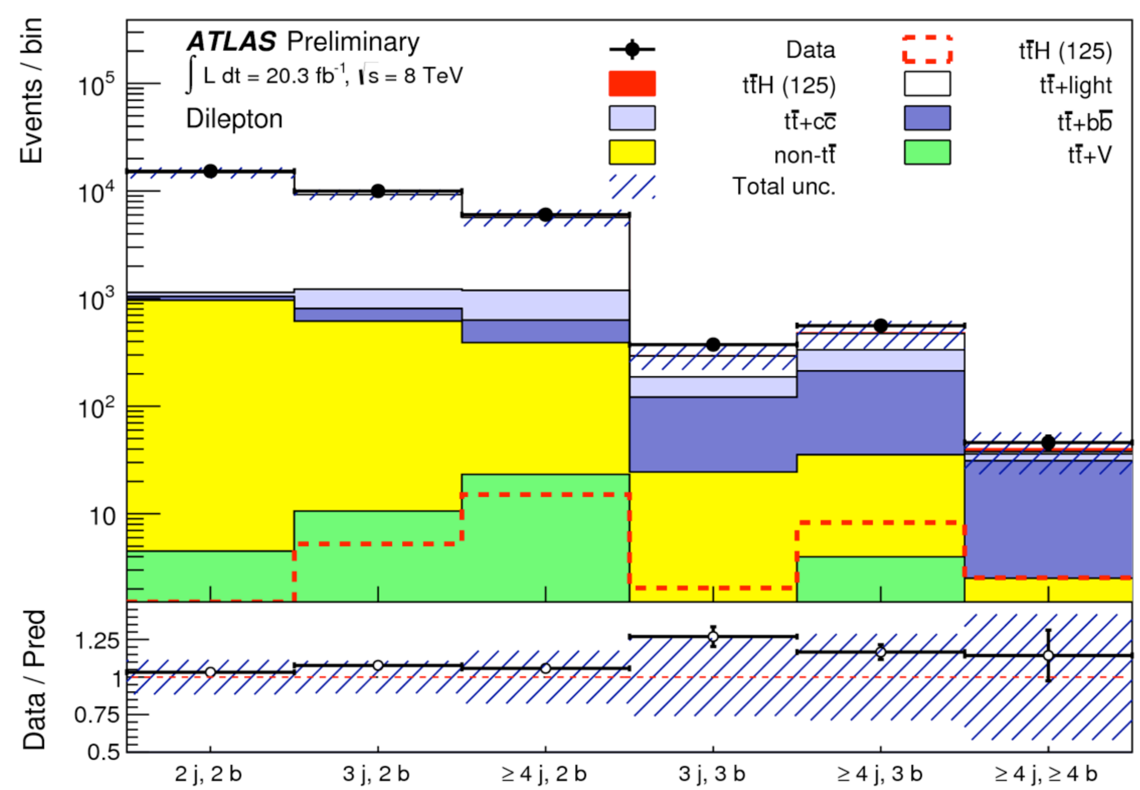


Pre-fit description

- Regions optimized to increase sensitivity with difference S/sqrt(B)
- Best S/B for high N(jets)/N(b-tags)
- Out of the box MC normalizations
- Simulation models data satisfactorily across the board



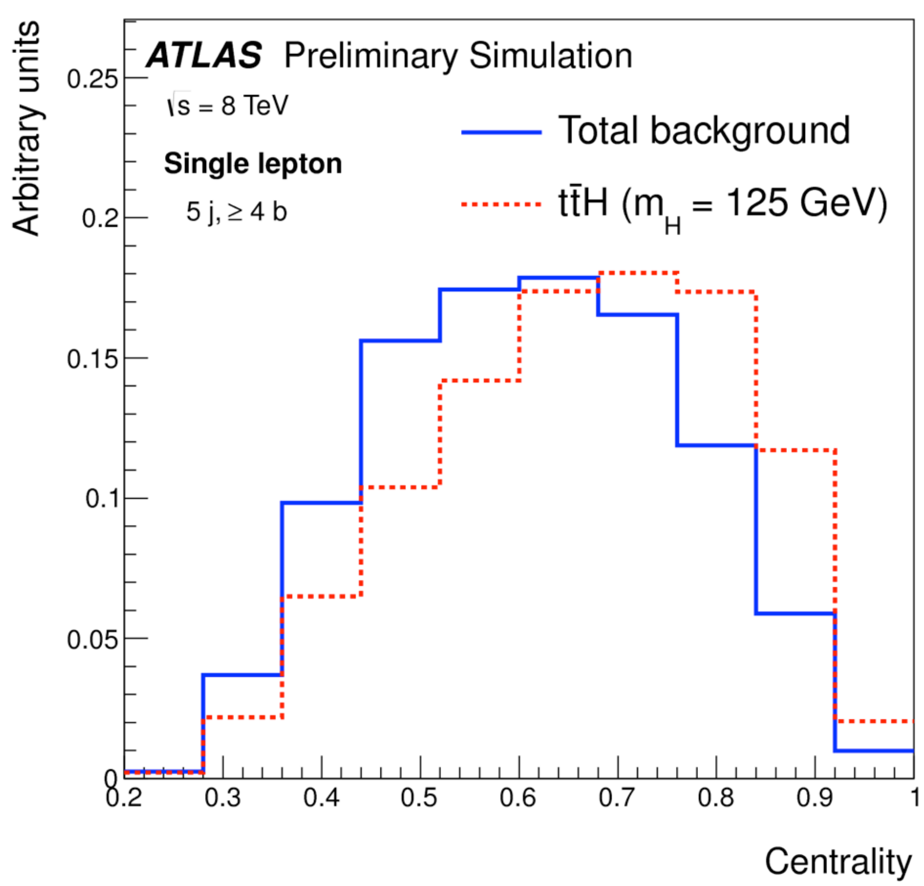
l+jets



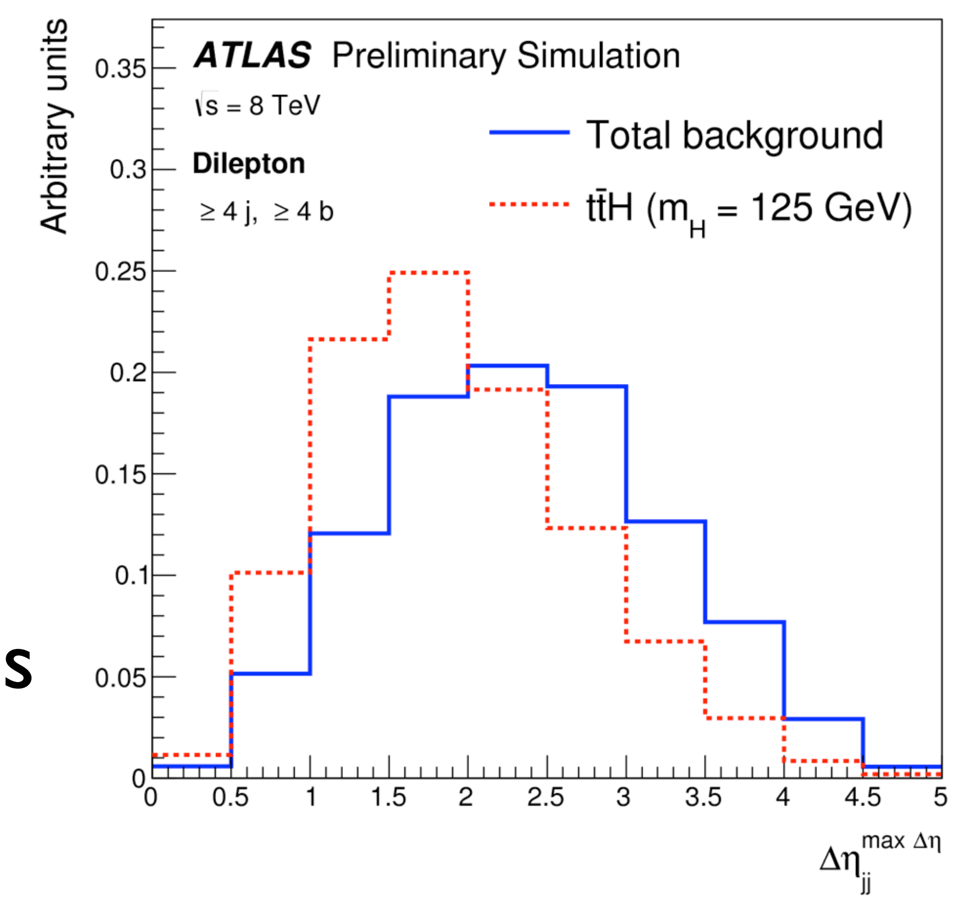
dilepton

MVA: S-B discrimination

- Train a Neural Network (NN) to separate S from B (tt+jets) simulations in each region
  - Uses a suite of variables from event shape and kinematics, single experimental object kinematics
  - proper ranking per channel and per region
- Dedicated background normalization control regions are designed by cutting on **$H_T = \Sigma p_T$(selected jets)**
- Designated signal regions have cut on NN discriminant



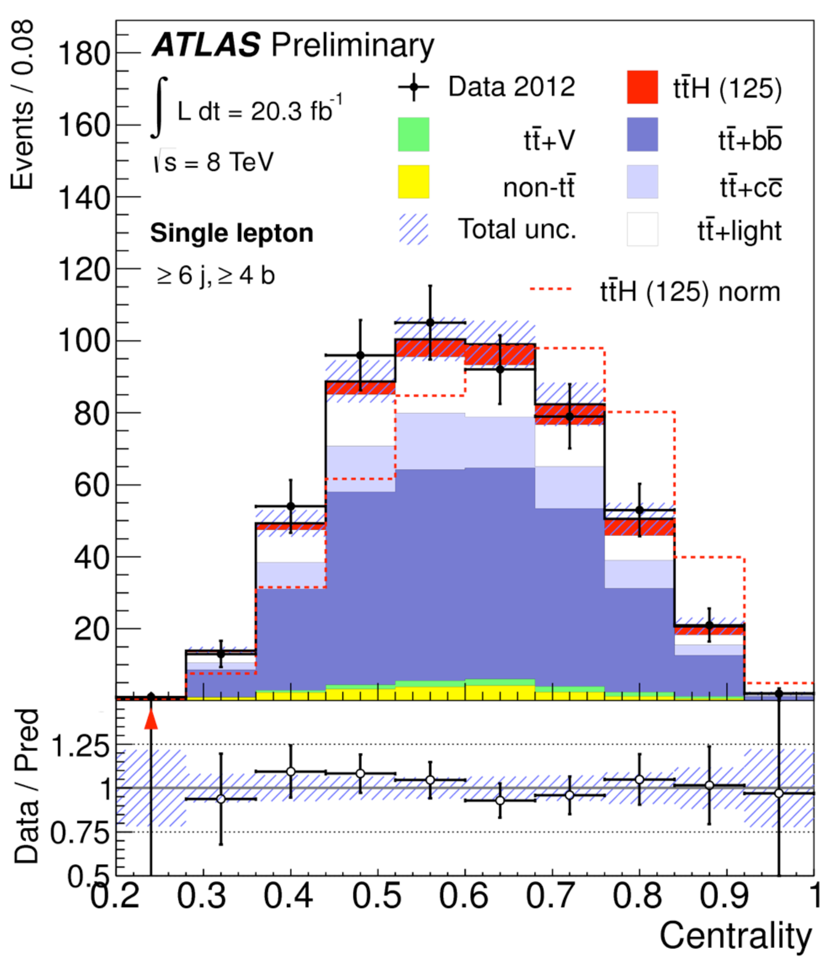
Examples



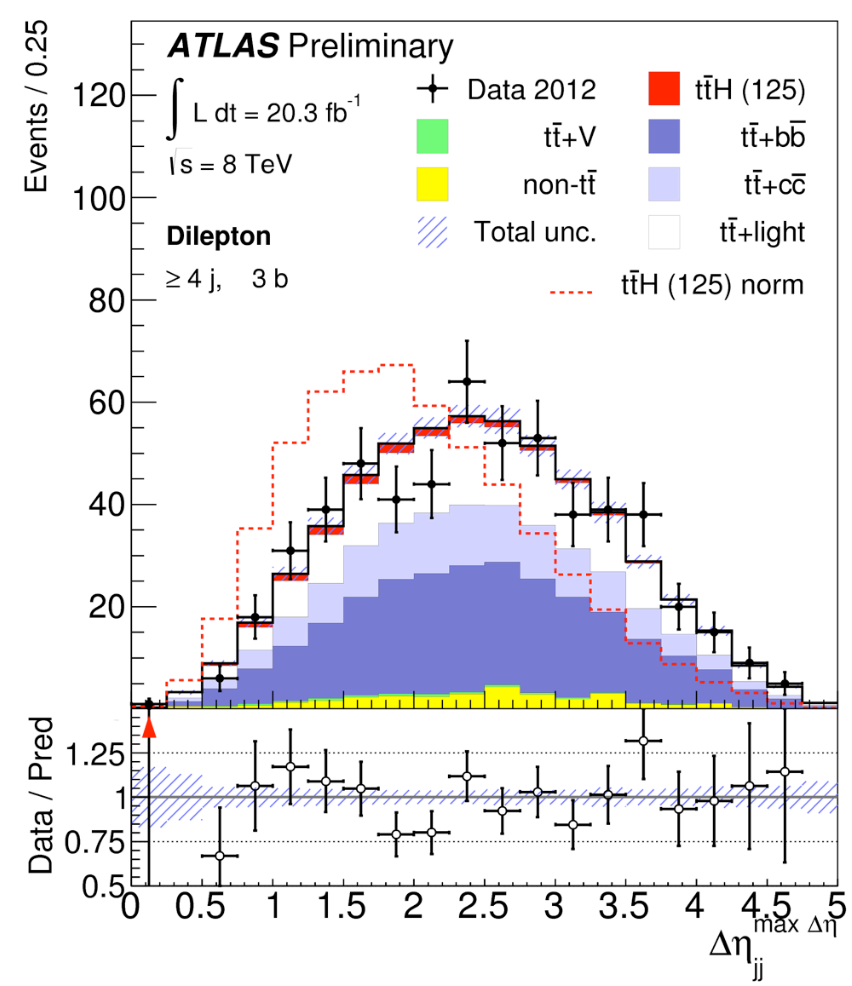
MVA: S-B discrimination

l+jets	2 b-tags	3 b-tags	4 b-tags
4 jets	CR (H_T cut)	CR (H_T cut)	CR (H_T cut)
5 jets	CR (H_T cut)	HF/LF NN	SR (NN cut)
≥ 6 jets	CR (H_T cut)	SR (NN cut)	SR (NN cut)

dilepton	2 b-tags	3 b-tags	4 b-tags
2 jets	CR (H_T cut)		
3 jets	CR (H_T cut)	SR (NN cut)	
≥ 4 jets	CR (H_T cut)	SR (NN cut)	SR (NN cut)

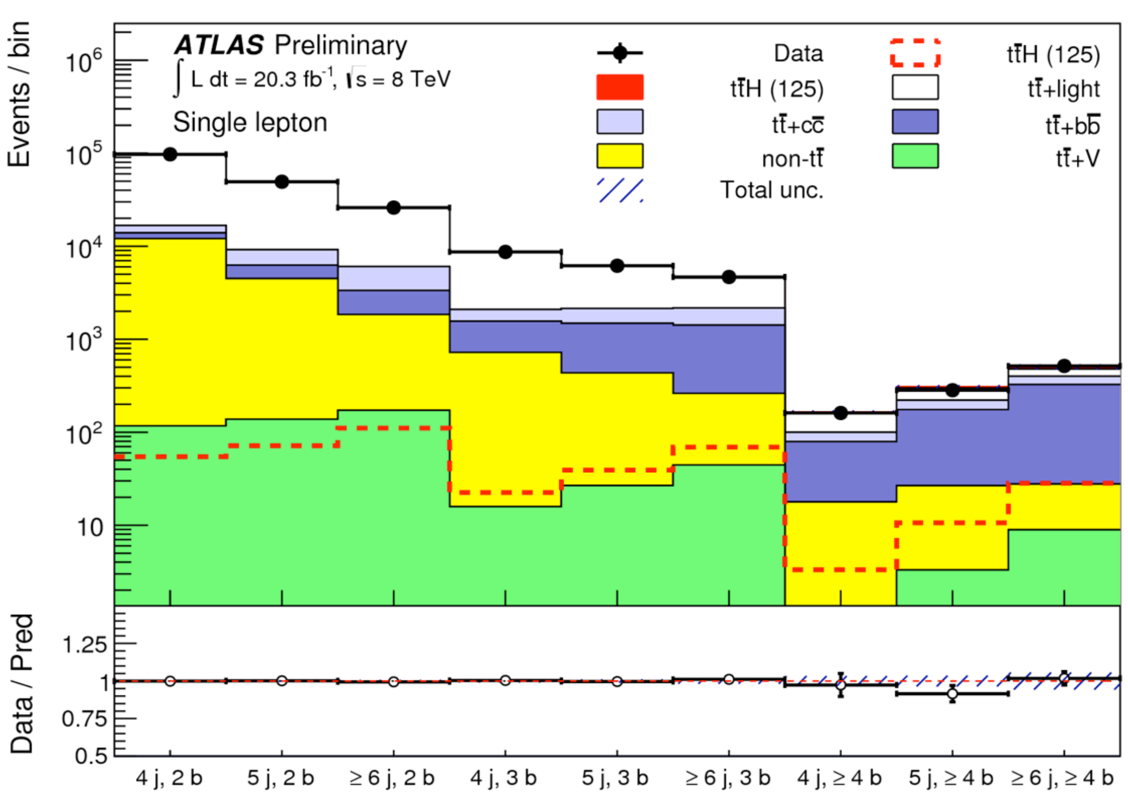


Examples

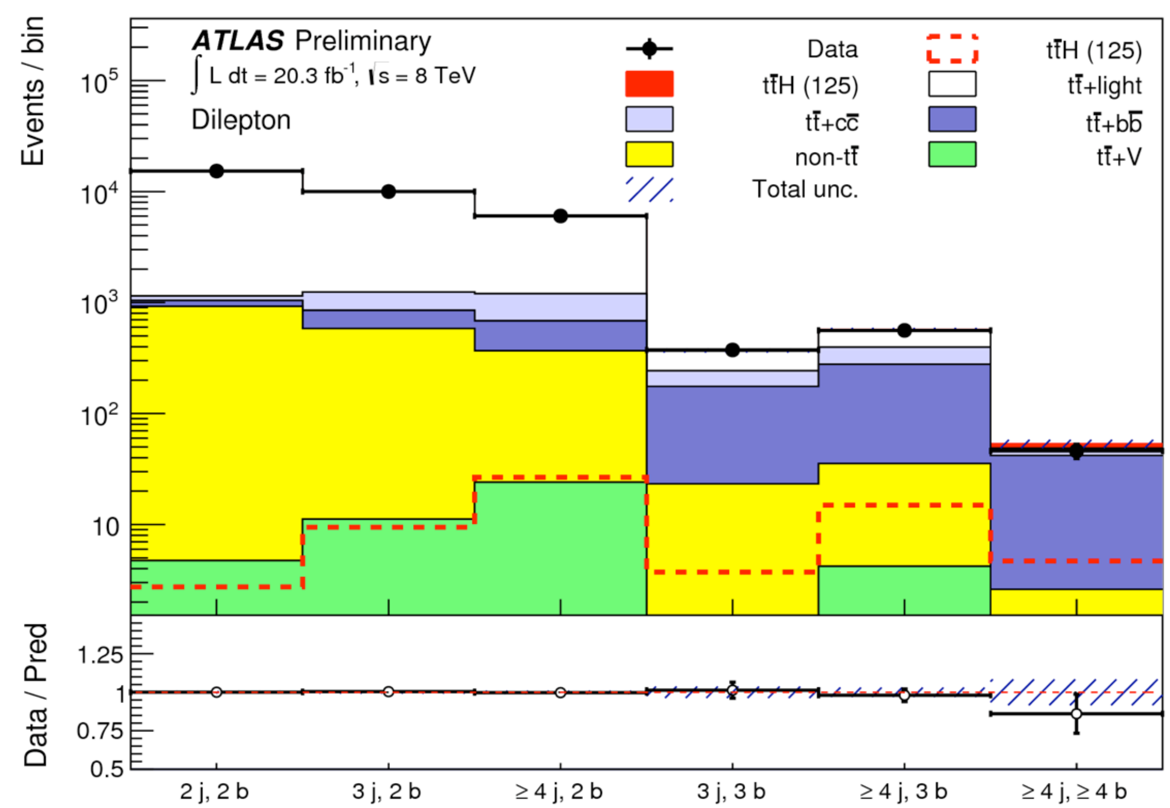


Post-fit description

- A profile likelihood fit is performed considering all regions simultaneously
  - reduces effect of systematic uncertainties thanks to high-stats, bkg enriched control regions



l+jets



dilepton

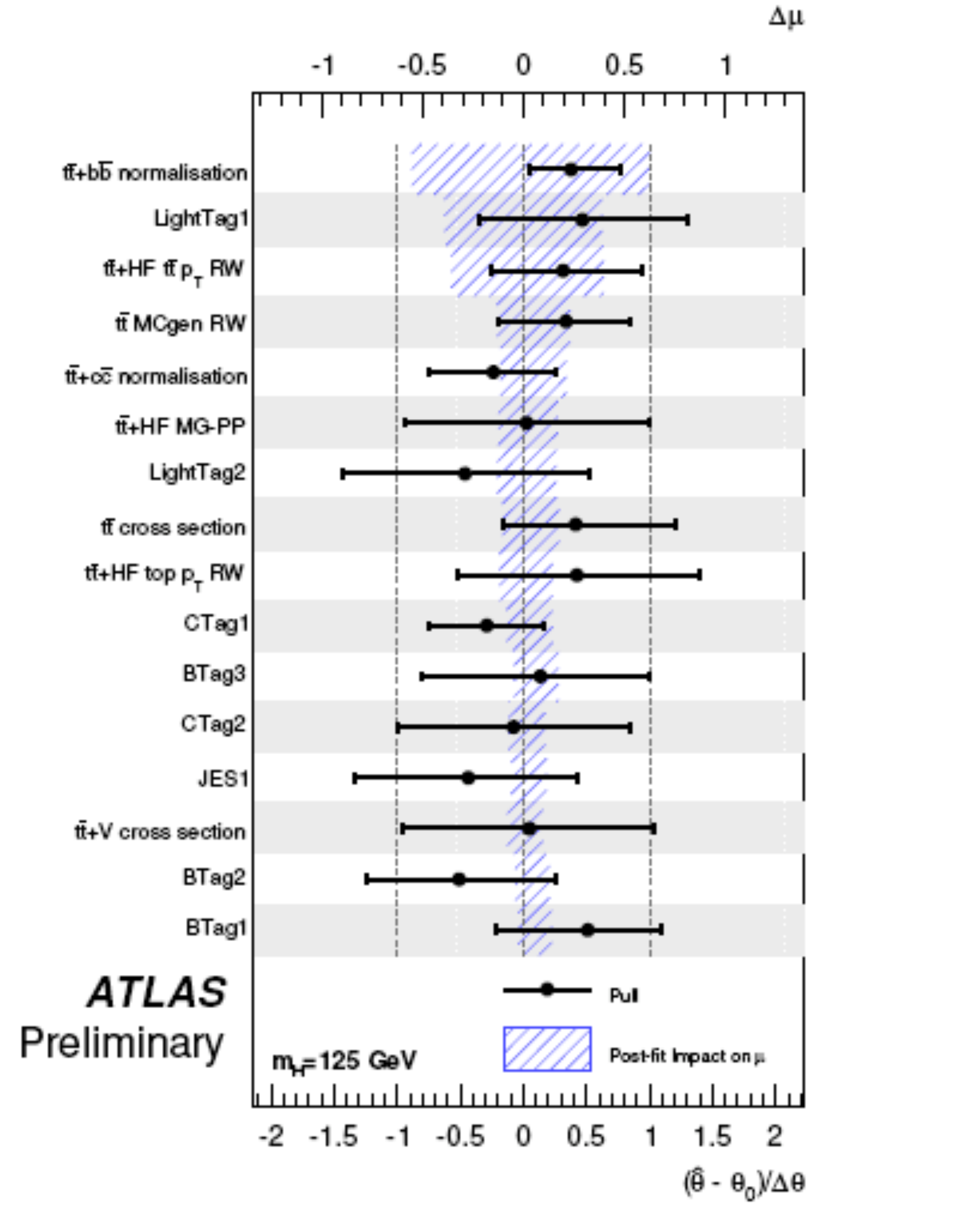
- Fit to data under the signal-plus-background hypothesis.
- Signal normalised to the fitted μ (=1.7)

Systematic uncertainties

Pre-fit

Systematic uncertainty	Type	Components
Luminosity	N	1
Physics Objects		
Electron	SN	5
Muon	SN	6
Jet energy scale	SN	22
Jet vertex fraction	SN	1
Jet energy resolution	SN	1
Jet reconstruction	SN	1
b-tagging efficiency	SN	6
c-tagging efficiency	SN	6
Light jet-tagging efficiency	SN	12
Background Model		
$t\bar{t}$ cross section	N	1
$t\bar{t}$ modelling: p_{T} reweighting	SN	9
$t\bar{t}$ modelling: parton shower	SN	2
$t\bar{t}$+heavy-flavour: normalisation	N	2
$t\bar{t}$+heavy-flavour: HF reweighting	SN	2
$t\bar{t}$+heavy-flavour: generator	SN	5
W+jets normalisation	N	3
W p_{T} reweighting	SN	1
Z+jets normalisation	N	2
Z p_{T} reweighting	SN	1
Multijet normalisation	N	3
Multijet shape dilepton	S	1
Single top cross section	N	1
Dibosons cross section	N	1
$t\bar{t}V$ cross section	N	1
Signal Model		
$t\bar{t}H$ modelling	SN	2

S=shape, N=normalization

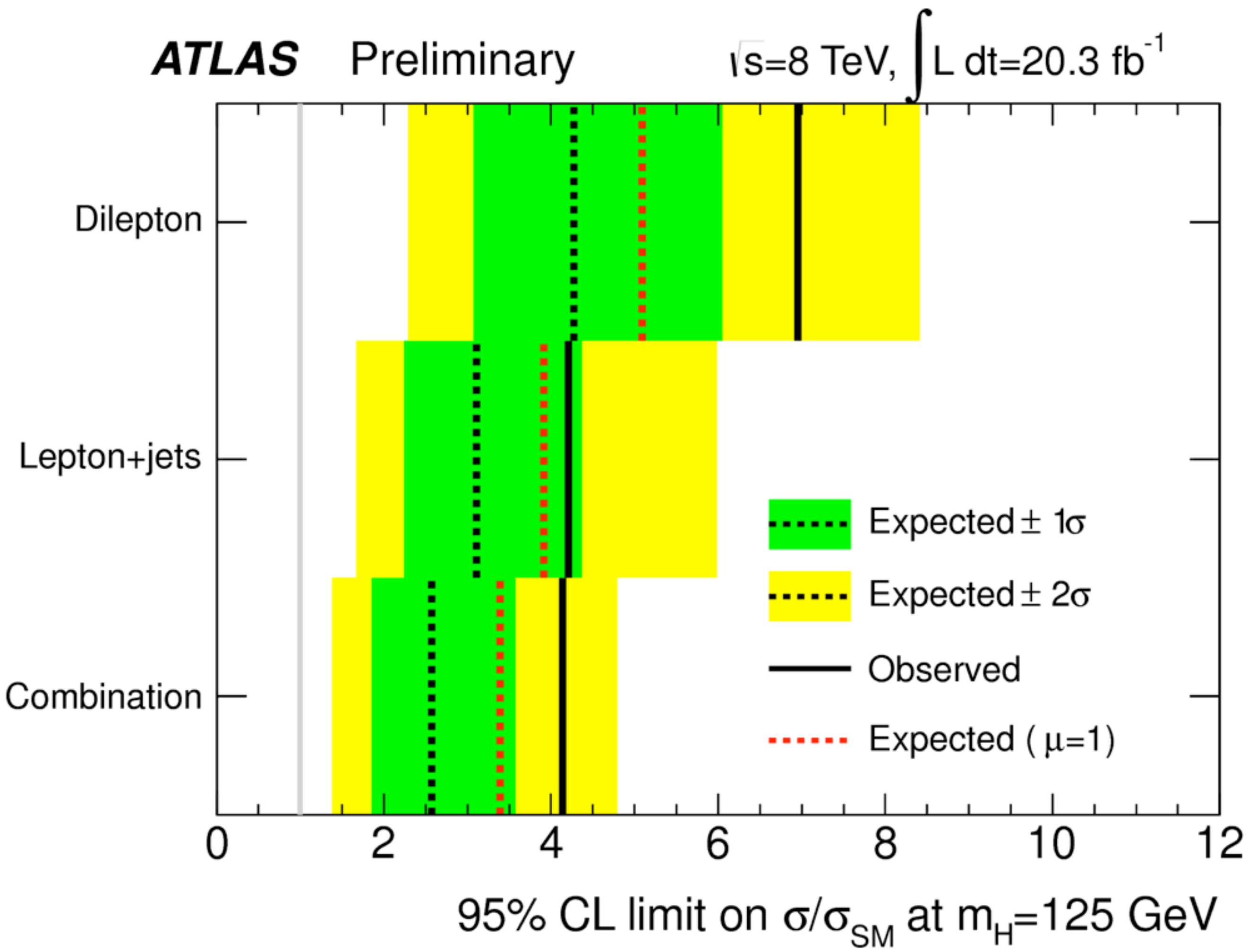


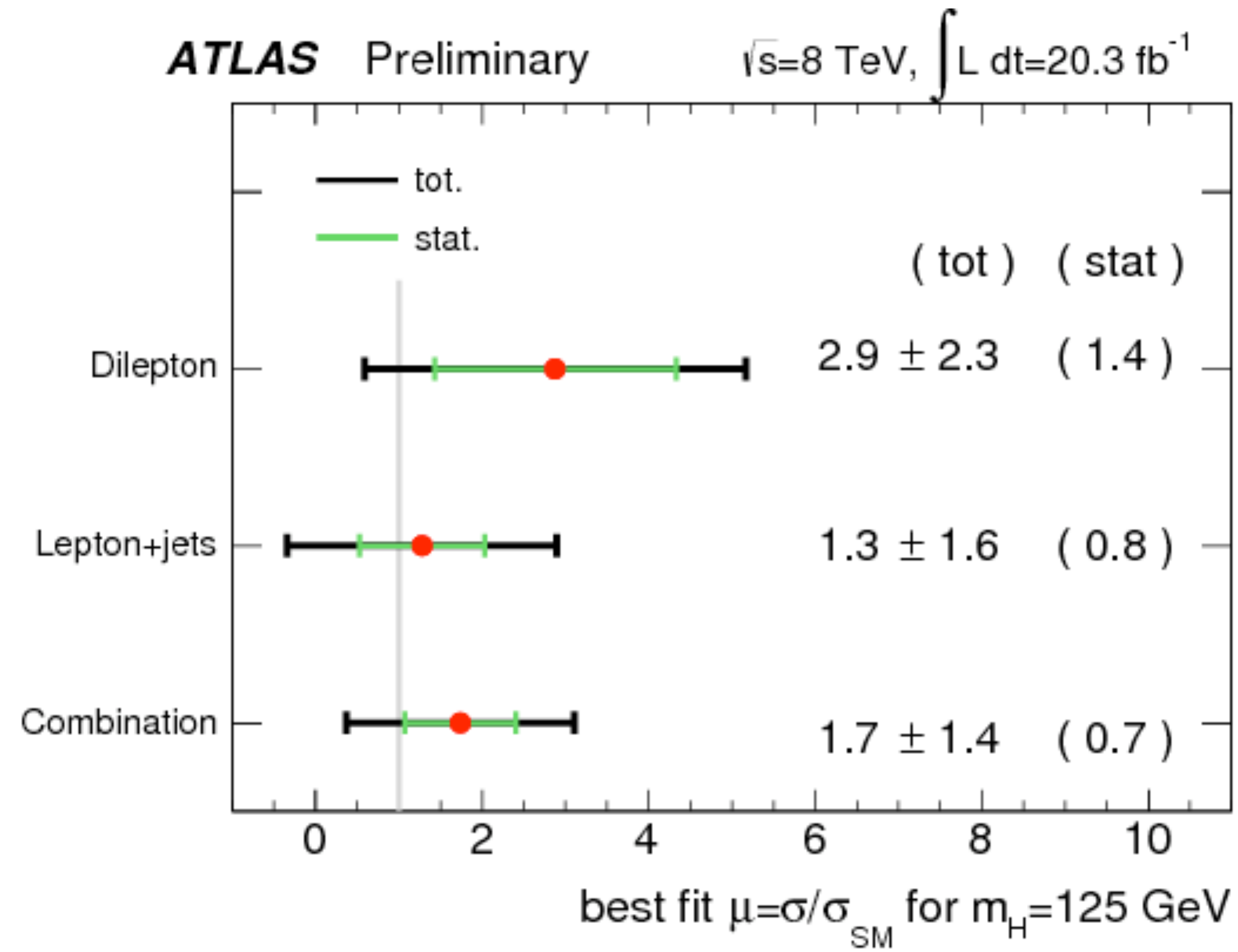
Sources with most impact on μ, as constrained by the data

Results

ATLAS-CONF-2014-011

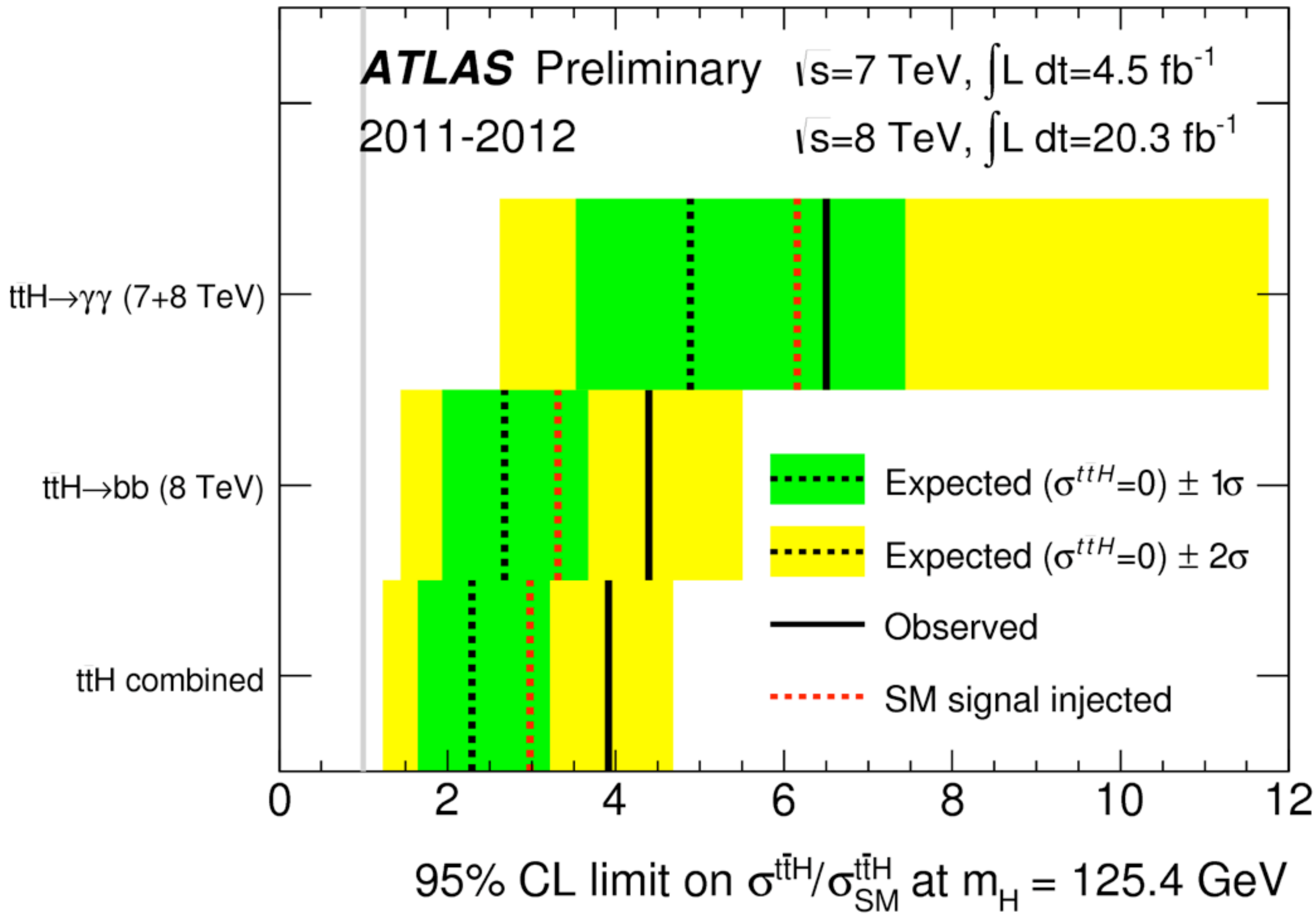
$\mu_{ttH} = \sigma/\sigma_{SM}$





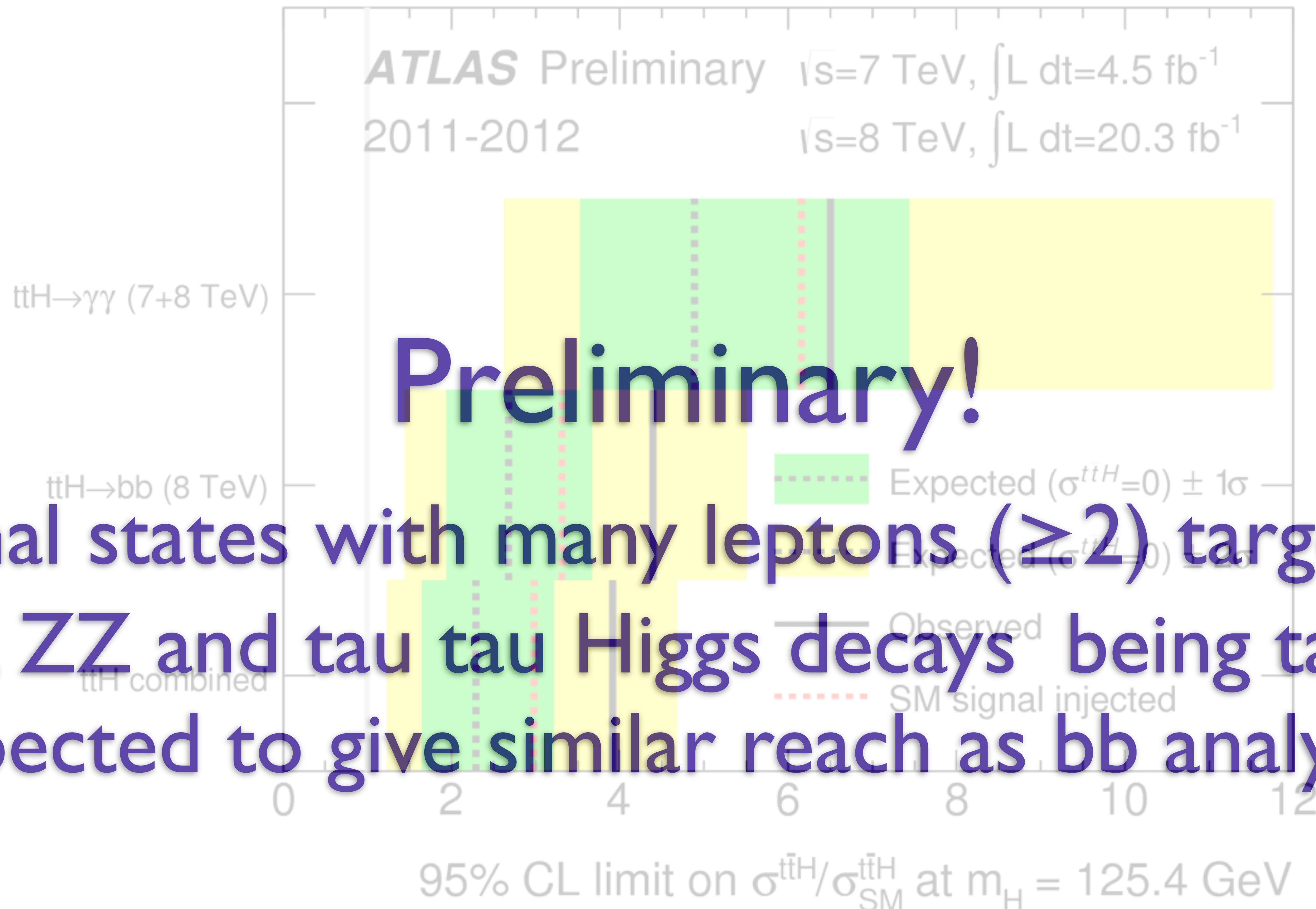
- Observed (expected) 95% CLs limit 4.5 x SM (2.6 x SM) for m(H) = 125 GeV.
- Best fit signal strength $\mu_{ttH} = 1.7 \pm 1.0$
- Significance Observed (expected) : 1.3 σ (0.8 σ)

Current Run I combo



- Observed (expected) 95% CLs limit : 3.9 x SM (2.3 x SM) for m(H)=125.4 GeV
- Best fit signal strength: μ (ttH) = 1.6 ± 0.6 (stat)$^{+1.1}_{-1.0}$ (syst.)

Current Run I combo



Preliminary!

- Final states with many leptons (≥2) targeting WW, ZZ and tau tau Higgs decays being tackled
- Expected to give similar reach as bb analysis

- Observed (expected) 95% CLs limit : 3.9 x SM (2.3 x SM) for m(H)=125.4 GeV
- Best fit signal strength: μ (ttH) = 1.6 ± 0.6 (stat)$^{+1.1}_{-1.0}$ (syst.)

Run II prospects

- Higher E_{CM} beneficial for ttH search:
  - @ 13(14) TeV ~4x (5x) increase in σ_{SM} for signal
  - ~3x (4x) increase in σ_{SM} for tt+jets background
- Additional techniques (e.g. more multi variate analysis, Matrix Element search) in development to maximize signal-bkg discrimination

- Not many prospect studies exist for RunII/future

- expect to measure ttH at 3σ with 300 fb^{-1} from di-photon channel alone (ATL-PHYS-PUB-2014-016)
- much earlier from bb and multi-lepton channels
- *tens fb^{-1} at 13 TeV?*

$\Delta\mu/\mu$	300 fb^{-1}		3000 fb^{-1}	
	All unc.	No theory unc.	All unc.	No theory unc.
$gg \to H$	0.12	0.06	0.11	0.04
VBF	0.18	0.15	0.15	0.09
WH	0.41	0.41	0.18	0.18
$qqZH$	0.80	0.79	0.28	0.27
$ggZH$	3.71	3.62	1.47	1.38
ttH	0.32	0.30	0.16	0.10

Outlook

- ttH production mode being looked at with several, different final states
- Challenging both from detector and backgrounds point of view
- Closing on to SM xsec value
- This will be one of the first hot topics in Run II

Back-up

Other sources of background: simulation

- ttZ, ttW : Madgraph +Pythia
- W/Z +jets : Alpgen + Pythia
- Dibosons : Alpgen + Herwig
- Single top : PowHeg / Acer +Pythia
- Multijets : Estimated by using data driven methods