

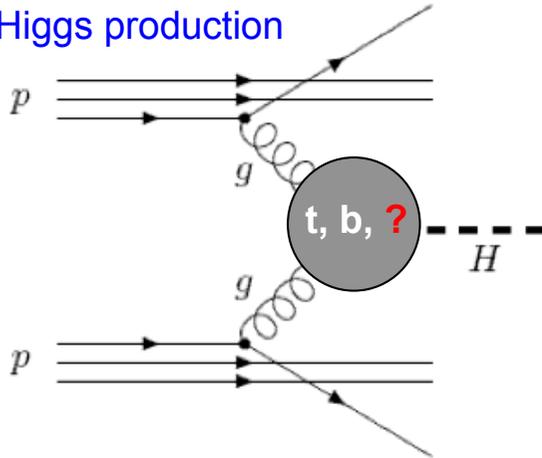
“After the Discovery: Hunting for a Non-Standard Higgs Sector”
Benasque, April 11, 2014

Status and Prospects for ttH Searches at the LHC

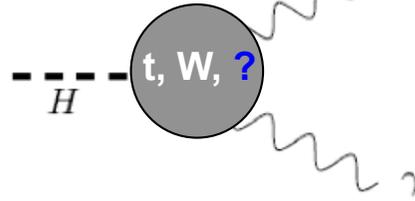
Aurelio Juste
ICREA/IFAE, Barcelona

Motivation

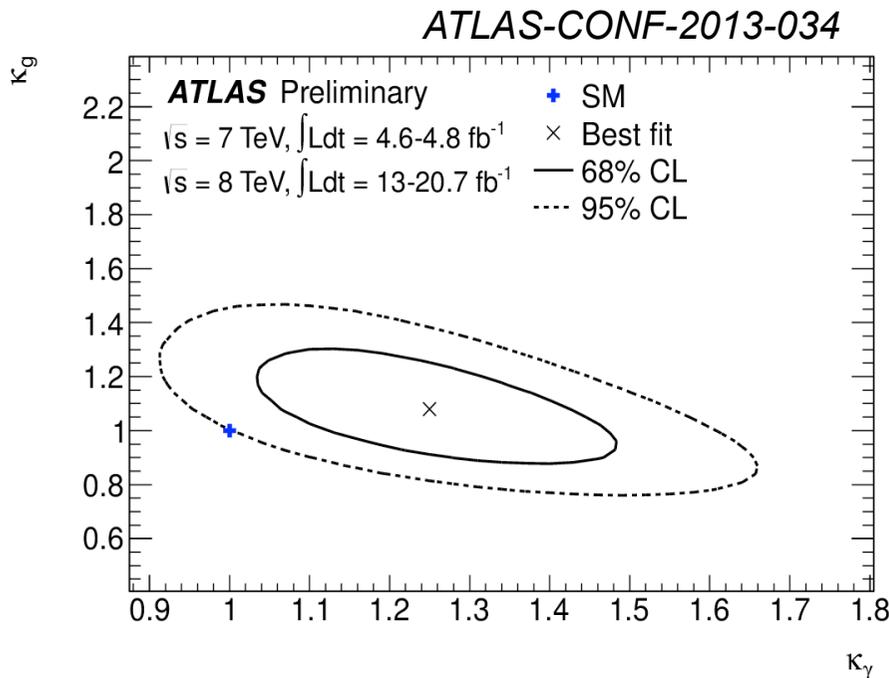
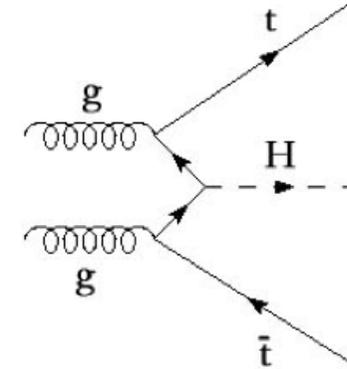
Higgs production



Higgs decay to photons



Higgstrahlung from top quark



- Indirect constraints on the top-Higgs Yukawa coupling can be extracted from channels involving the ggH and $\gamma\gamma H$ vertices
 → assumes no new particles.
- Top-Higgs only Yukawa coupling that can be measured directly:

$$\sigma(t\bar{t}H) \propto g_{t\bar{t}H}^2$$

→ allows probing for NP contributions in the ggH and $\gamma\gamma H$ vertices.

Motivation

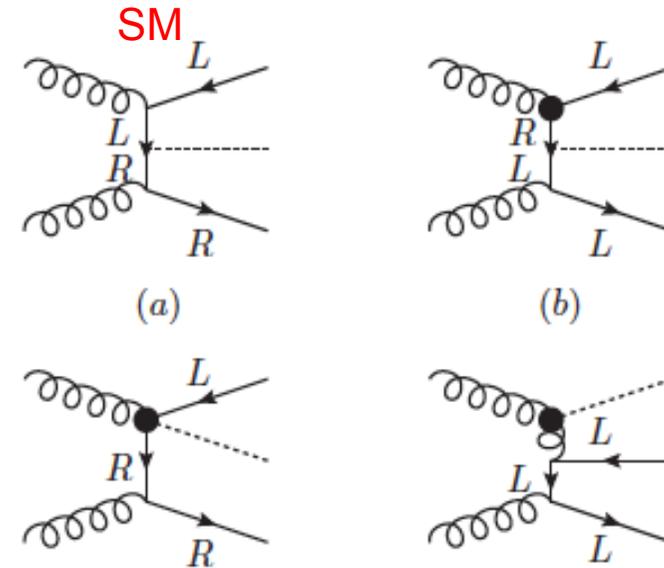
- Higher-dimension operators that involve the top and Higgs fields:
 - are little tested so far, and
 - are particularly sensitive to New Physics associated with EWSB.

- Effective top-Higgs Yukawa coupling can deviate from SM prediction due to contributions from dimension-6 operators.

Example: $\sigma(ttH)$ at $\sqrt{s}=14$ TeV:

$$\begin{aligned}
 \frac{\sigma(pp \rightarrow t\bar{t}h)}{\text{fb}} &= \text{SM} \left[611_{-110}^{+92} + [457_{-91}^{+127} \Re c_{hg} - 49_{-10}^{+15} c_G \right. \\
 &+ 147_{-32}^{+55} c_{HG} - 67_{-16}^{+23} c_y] \left(\frac{\text{TeV}}{\Lambda} \right)^2 \\
 &+ [543_{-123}^{+143} (\Re c_{hg})^2 + 1132_{-232}^{+323} c_G^2 \\
 &+ 85.5_{-21}^{+73} c_{HG}^2 + 2_{-0.5}^{+0.7} c_y^2 \\
 &+ 233_{-144}^{+81} \Re c_{hg} c_{HG} - 50_{-14}^{+16} \Re c_{hg} c_y \\
 &\left. - 3.2_{-8}^{+8} \Re c_{Hy} c_{HG} - 1.2_{-8}^{+8} c_H c_{HG} \right] \left(\frac{\text{TeV}}{\Lambda} \right)^4
 \end{aligned}$$

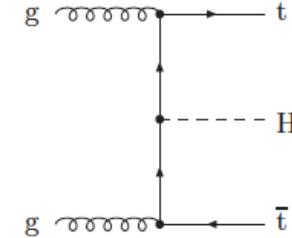
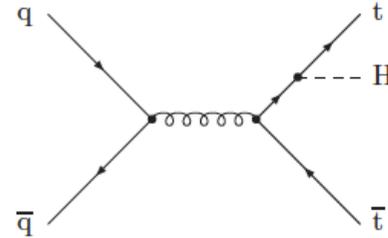
arXiv:1205.1065



Complementary to $gg \rightarrow H$ and $t\bar{t}$ cross section measurements, which are sensitive to a different combination of operators.

ttH Production in pp Collisions

- $t\bar{t}H$ production has the lowest cross section for a SM-like Higgs boson at LHC.
- Interestingly, the phase-space suppression effect is overcome at $\sqrt{s} > 30\text{-}40$ TeV, where $t\bar{t}H$ becomes the 3rd most important production mechanism.



- $\sigma(t\bar{t}H)$ known at NLO in QCD.

For $M_H = 125$ GeV:

$\sqrt{s} = 7$ TeV: $\sigma(t\bar{t}H) = 86$ fb

$\sqrt{s} = 8$ TeV: $\sigma(t\bar{t}H) = 130$ fb

$\sqrt{s} = 14$ TeV: $\sigma(t\bar{t}H) = 611$ fb

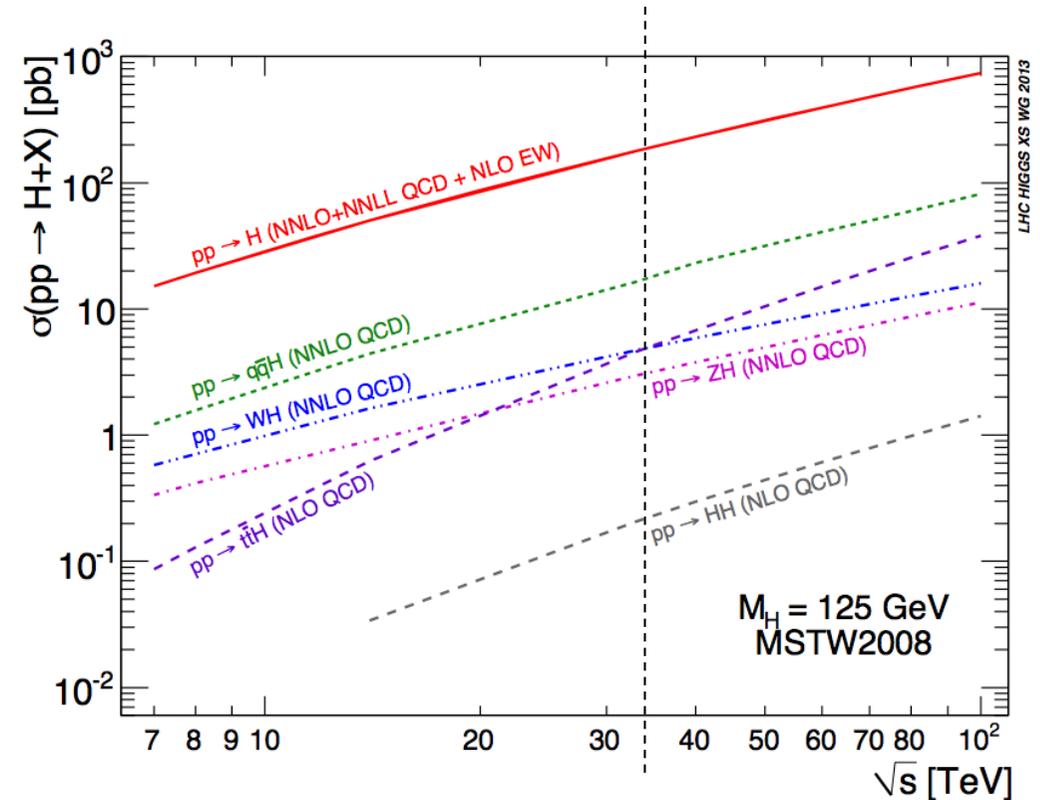
($\sim \times 5$ wrt $\sqrt{s} = 8$ TeV)

Uncertainties:

+5.9%/-9.3% (scale), $\pm 8.9\%$ (PDF)

Adds a $\sim 8\%$ uncertainty to the top-Higgs

Yukawa. **Will need to be improved!**

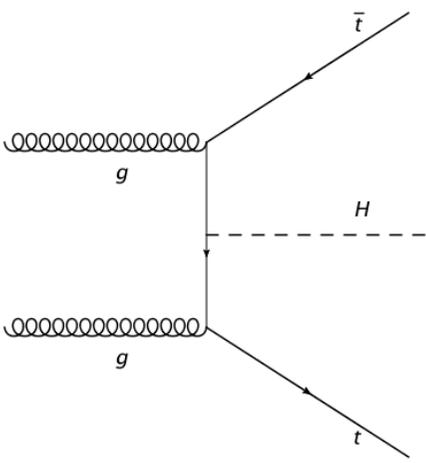
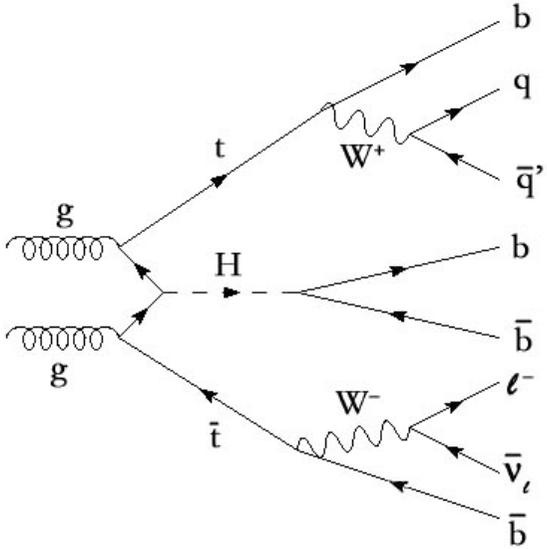


Direct Searches for ttH Production

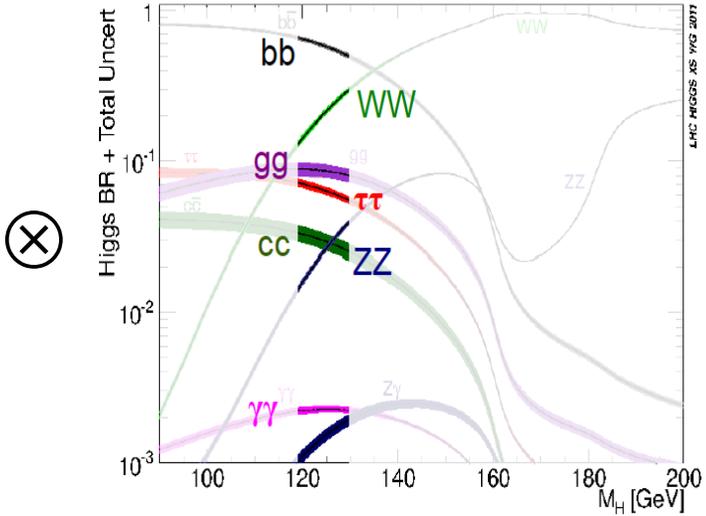
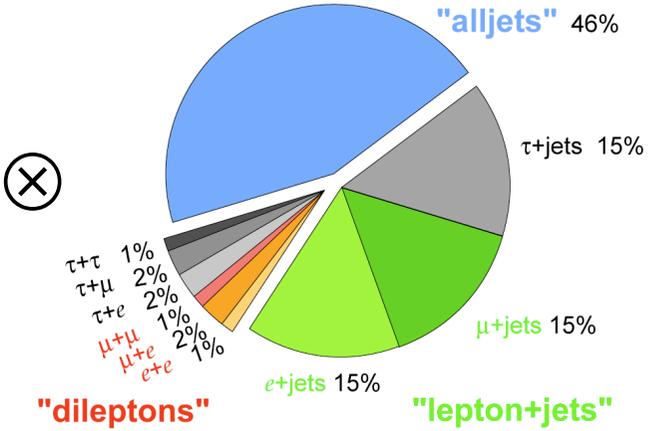
Virtues:

- Distinctive final states with high jet/b-tag multiplicity and multiple heavy resonances
 - A priori many handles against backgrounds!
- For $M_H=125$ GeV, $H \rightarrow b\bar{b}$ dominates, although other decay modes can also be exploited: $H \rightarrow \tau^+\tau^-$, W^+W^- , ZZ , and even $\gamma\gamma$!
- Many possible final states to consider!
Need to find the best combinations of top and Higgs decays to isolate the signal.

Lepton+jets channel ($H \rightarrow b\bar{b}$)



Top Pair Branching Fractions

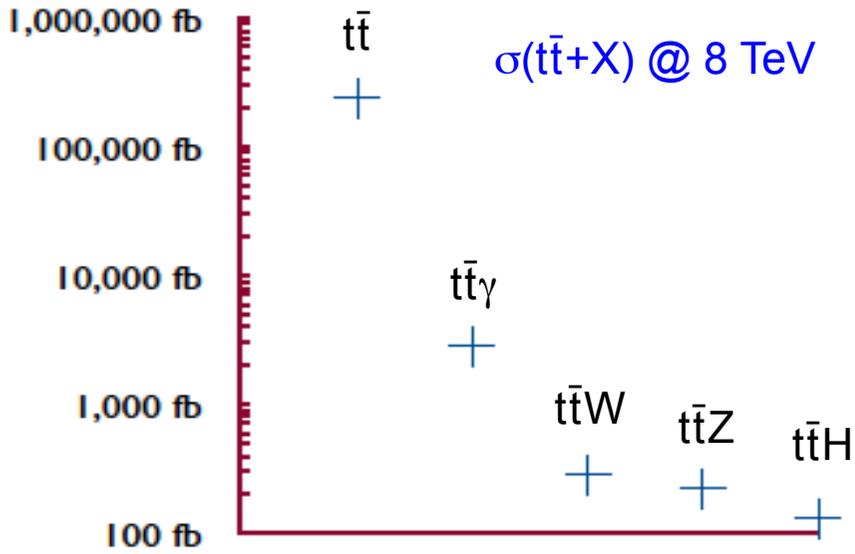
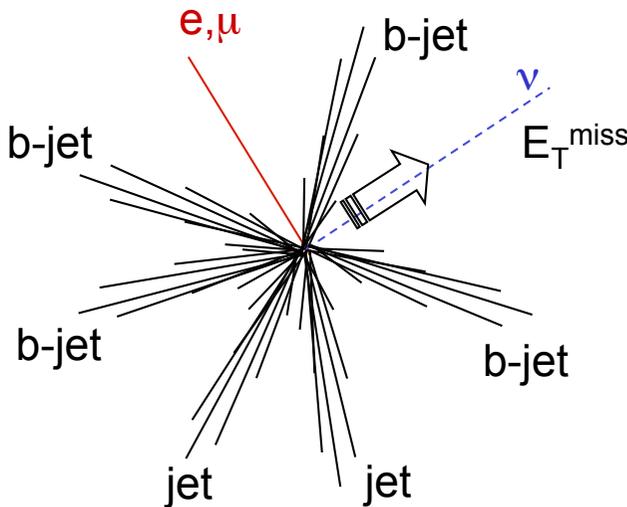


Direct Searches for ttH Production

Challenges:

- Low production cross section.
- $H \rightarrow b\bar{b}$, $\tau^+\tau^-$: large combinatorial and physics backgrounds (mainly $t\bar{t}$ +jets), affected by large systematic uncertainties.
- $H \rightarrow W^+W^-$, ZZ : typically focus on multilepton final states, which have smaller backgrounds but also small signal rate.
- $H \rightarrow \gamma\gamma$: small signal rate.

Lepton+jets channel ($H \rightarrow b\bar{b}$)



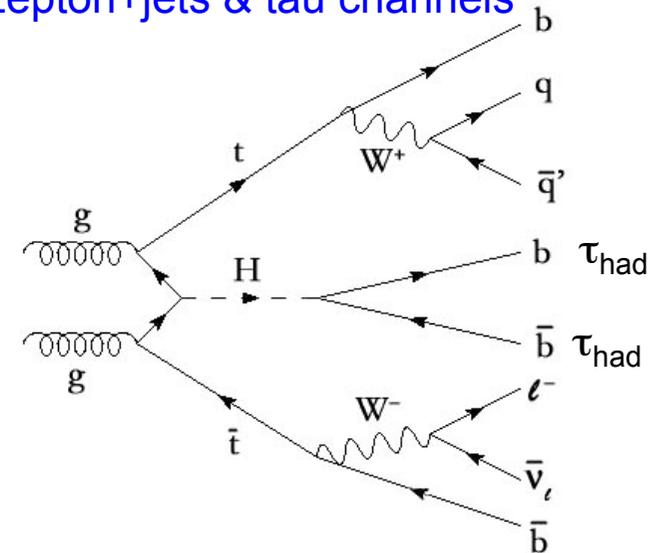
Cross section ratio for $M_H=125 \text{ GeV}$:
LHC: $\sigma(t\bar{t})/\sigma(t\bar{t}H) \sim 2000(1500)$ for $\sqrt{s}=7 \text{ TeV}(14 \text{ TeV})$
LC: $\sigma(t\bar{t})/\sigma(t\bar{t}H) \sim 500(100)$ for $\sqrt{s}=500 \text{ GeV}(1 \text{ TeV})$

$ttH, H \rightarrow bb/\tau\tau$

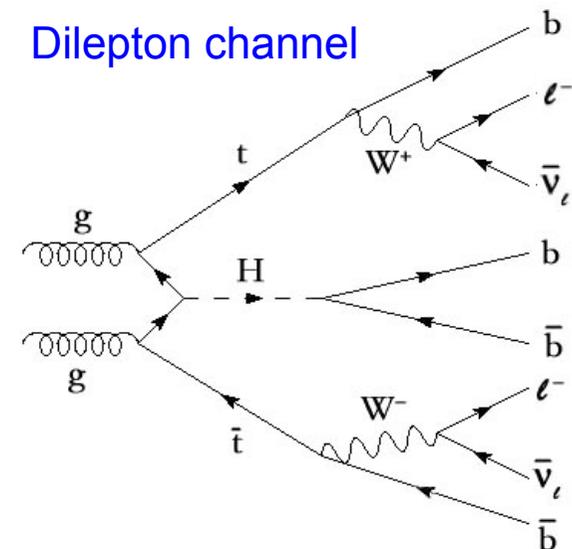
Basic Analysis Strategy

- Select $t\bar{t}$ -enriched samples:
 - Lepton+jets and OS dilepton (ATLAS, CMS)
- Pick signals being targeted:
 - $H \rightarrow b\bar{b}$ (ATLAS, CMS), $H \rightarrow \tau\tau$ (CMS).
- Categorize events by jet and b-tag multiplicities:
 - Improve sensitivity by keeping separate high and low S/\sqrt{B} channels.
 - Signal-depleted channels will be exploited to constrain systematic uncertainties.
- For each analysis channel, choose a discriminant variable:
 - ATLAS: single kinematic variable or multivariate discriminant
 - CMS: multivariate discriminant
- Hypothesis testing including in-situ constraining of systematic uncertainties.

Lepton+jets & tau channels



Dilepton channel



Event Selections

"Signal-rich" categories

Lepton+jets channel



=1 isolated e or μ , $p_T > 25$ GeV
 ≥ 4 jets, $p_T > 25$ GeV (anti- k_T $R=0.4$)
 ≥ 2 b-tags
No E_T^{miss} or $m_{T,W}$ requirements



=1 isolated e or μ , $p_T > 30$ GeV
 ≥ 4 jets, $p_T > 30$ GeV (anti- k_T $R=0.5$)
 ≥ 3 jets, $p_T > 40$ GeV
 ≥ 2 b-tags
No E_T^{miss} or $m_{T,W}$ requirements

Divide into 9 categories:

4 jets (2, 3, ≥ 4 b-tags)
5 jets (2, 3, ≥ 4 b-tags)
 ≥ 6 jets (2, 3, ≥ 4 b-tags)

Divide into 7 categories:

4 jets (3, ≥ 4 b-tags)
5 jets (3, ≥ 4 b-tags)
 ≥ 6 jets (2, 3, ≥ 4 b-tags)

Event Selections

Dileptons channel

“Signal-rich” categories



ee, $\mu\mu$, e μ final states
 lepton $p_T > 25/15$ GeV
 ≥ 2 jets, $p_T > 25$ GeV (anti- k_T R=0.4)
 ≥ 2 b-tags

Divide into 6 categories:

2 jets (2 b-tags)
 3 jets (2, 3 b-tags)
 ≥ 4 jets (2, **3, ≥ 4 b-tags**)



ee, $\mu\mu$, e μ final states
 tight/loose lepton, $p_T > 20/10$ GeV
 ≥ 2 jets, $p_T > 30$ GeV (anti- k_T R=0.5)
 ≥ 2 b-tags

Divide into 3 categories:

3 jets (2 b-tags)
 ≥ 4 jets (2 b-tags)
 ≥ 3 jets (**≥ 3 b-tags**)

Tau channel



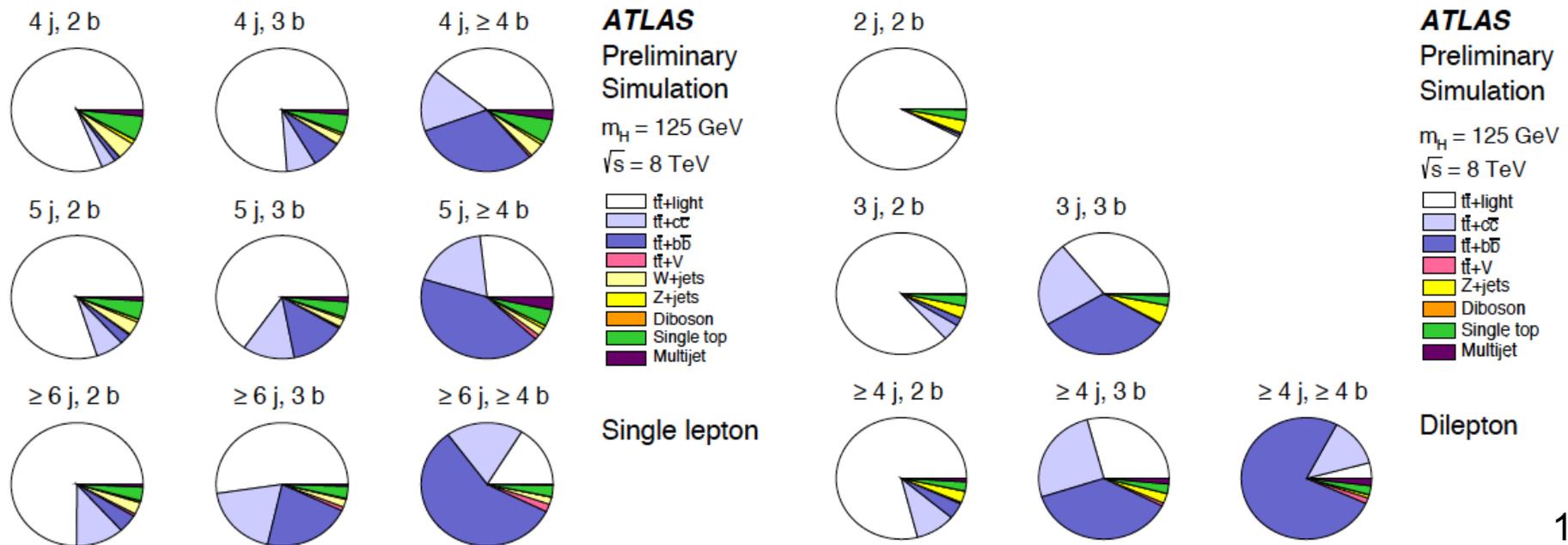
=1 isolated e or μ , $p_T > 30$ GeV and
 ≥ 2 1-prong τ_{had} , $p_T > 20$ GeV
 ≥ 2 jets, $p_T > 30$ GeV (anti- k_T R=0.5)
 1 or 2 b-tags

Divide into 6 categories:

2 jets (1, 2 b-tags)
 3 jets (1, 2 b-tags)
 ≥ 4 jets (**1, 2 b-tags**)

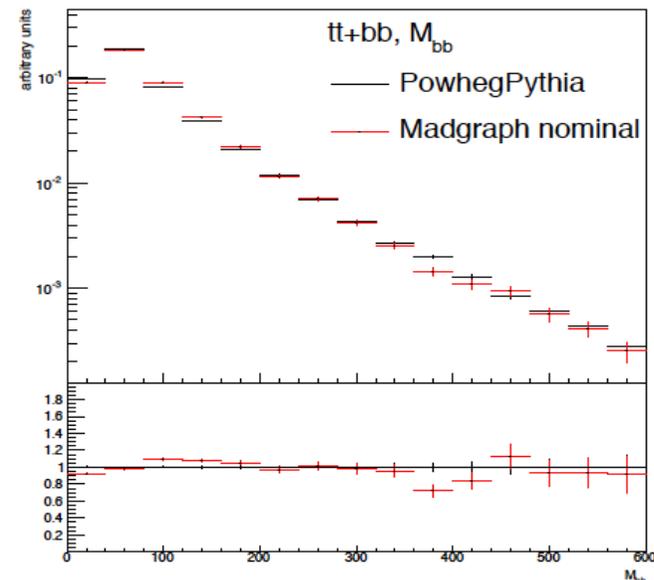
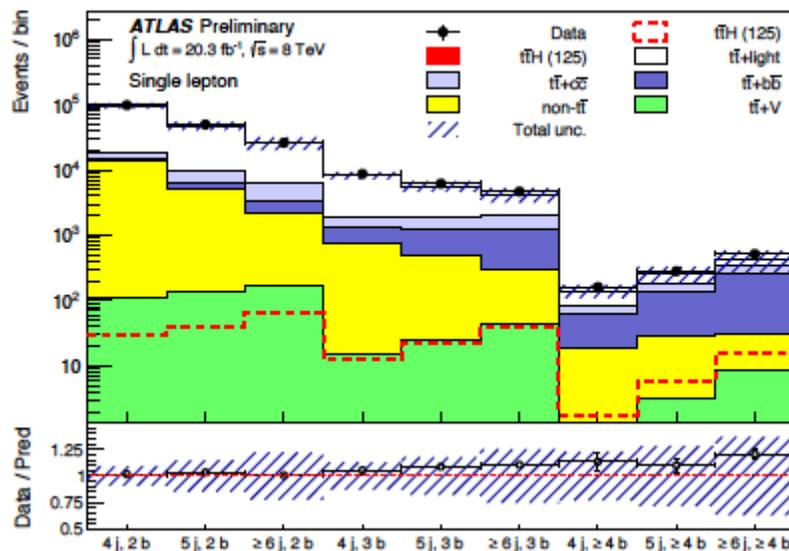
Signal and Background Modeling

- $t\bar{t}H$ Signal: *POWHEL+PYTHIA* (ATLAS), *PYTHIA* (CMS)
- Backgrounds:
 - $t\bar{t}$ +jets: *POWHEG+PYTHIA* (ATLAS), *MADGRAPH+PYTHIA* (CMS)
 - $t\bar{t}W$, $t\bar{t}Z$: *MADGRAPH+PYTHIA*
 - $W/Z/\gamma^*$ +jets: *ALPGEN+PYTHIA* (ATLAS), *MADGRAPH+PYTHIA* (CMS)
 - Single top: *POWHEG+PYTHIA/AcerMC+PYTHIA* (ATLAS), *POWHEG+PYTHIA* (CMS)
 - Dibosons: *ALPGEN+H*(ATLAS), *PYTHIA* (CMS)
 - Multijets: normalization and shape data-driven
- After requiring ≥ 1 b-tag background dominated by $t\bar{t}$.



$t\bar{t}$ +jets Modeling

- Based on matrix element (ME)+parton shower (PS) MCs.
Inclusive $t\bar{t}$ +jets samples normalized to approx NNLO cross section.
- *MADGRAPH+PYTHIA* → used by CMS
 - Separate samples for $t\bar{t}$ +n partons ($n \leq 3$), including heavy quarks (5F scheme).
 - Matched samples. Heavy-flavor overlap removal automatically handled.
- *POWHEG+PYTHIA* → used by ATLAS
 - Good modeling of $t\bar{t}$ +jets production, including jet multiplicity and kinematics.
 - Modeling of $t\bar{t}$ +HF comparable (in normalization and kinematics) to *MADGRAPH*.
- Modeling further improved by applying MC corrections based on differential cross section measurements (e.g. ATLAS: as a function of top p_T and $t\bar{t}$ p_T ; ATL-CONF-2013-099).
- Overall rate of $t\bar{t}+b\bar{b}$ and $t\bar{t}+c\bar{c}$ calibrated to data using background-enriched bins in signal-rich regions.



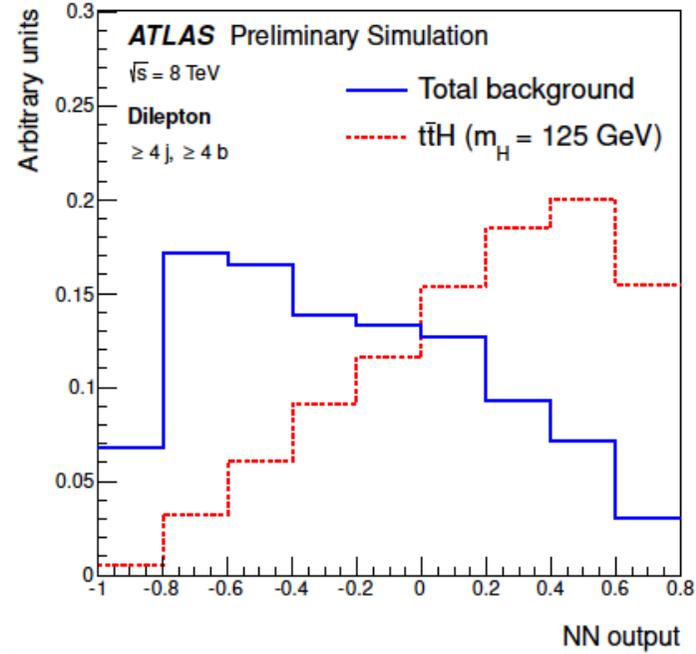
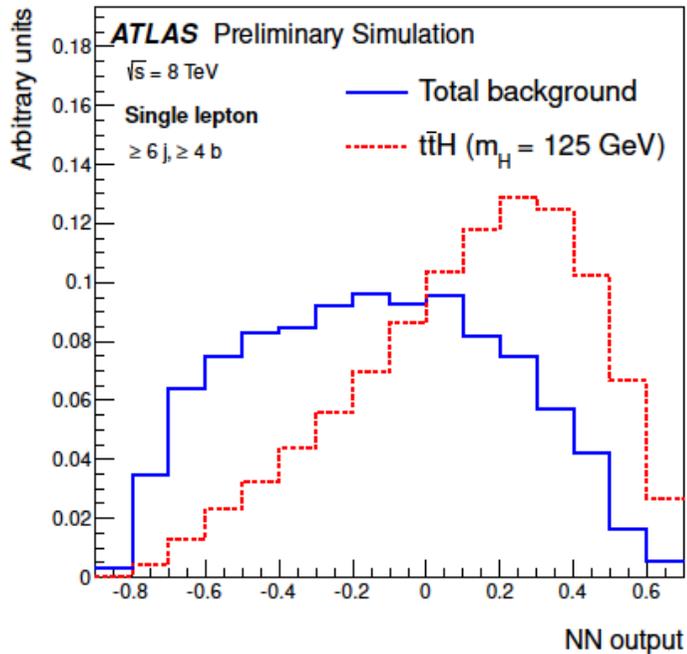
Signal-to-Background Discrimination

- **Signal-depleted regions:** use $H_T^{had} = \sum p_T^{jets}$ for ℓ +jets and $H_T = \sum p_T^{jets} + \sum p_T^\ell$ for dilepton
- **ℓ +jets, 5 jets, 3 b-tags region:** use NN trained to separate $t\bar{t} + b\bar{b}/c\bar{c}$ from $t\bar{t}$ +light jets
- **Signal-rich regions:** use NN trained to separate $t\bar{t}H$ from $t\bar{t}$ +jets in each of the region

Lepton+jets ATLAS-CONF-2014-011

	2 b-tags	3 b-tags	≥ 4 b-tags
4 jets	H_T^{had}	H_T^{had}	H_T^{had}
5 jets	H_T^{had}	NN	NN
≥ 6 jets	H_T^{had}	NN	NN

- Make use of event kinematics
- Object kinematics: p_T^{jet5}, \dots
 - Global event variables: $H_T^{had}, N_{p_T > 40 GeV}^{jet}, \dots$
 - Event shape variables: centrality, Fox-Wolfram moments, ...
 - Object pair properties: $M_{bb}^{min\Delta R}, \Delta R_{bb}^{avg}, \dots$

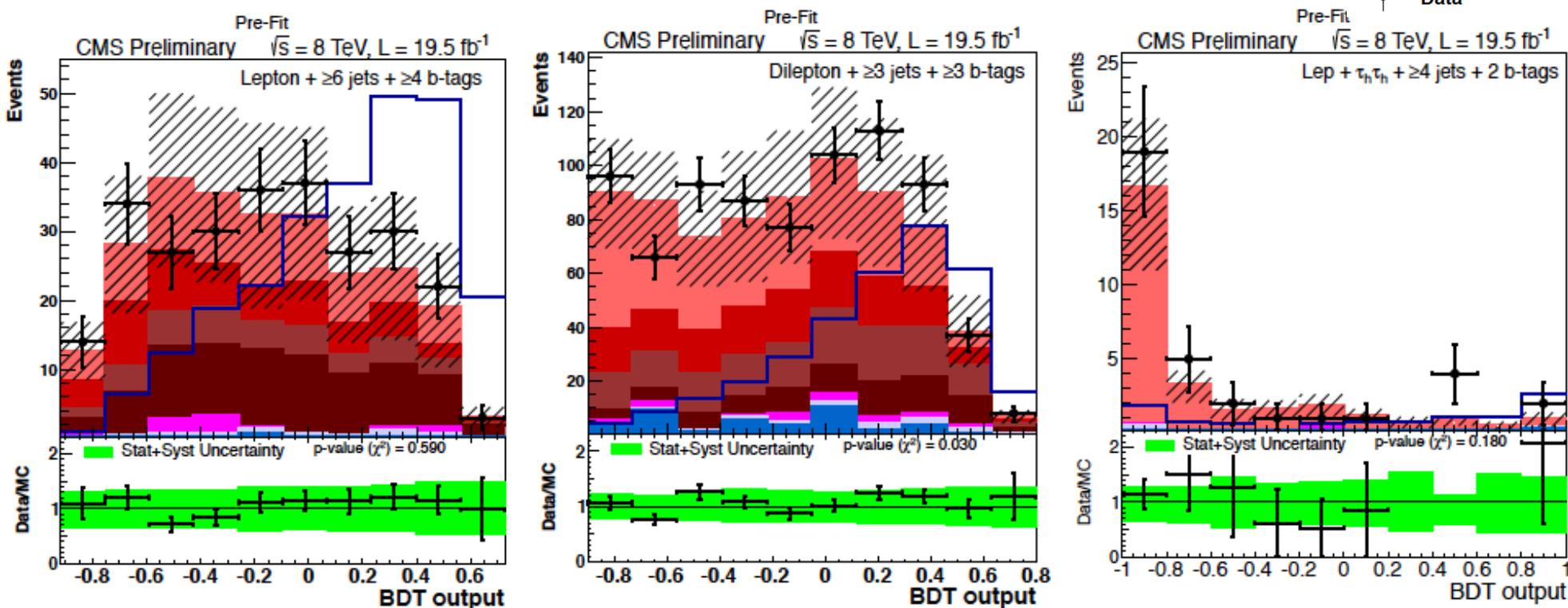


Signal-to-Background Discrimination



- Boosted Decision Trees (BDTs) trained for each category of the analysis.
- A total of 10 / 4-6 / 10 variables are used in the lepton+jets / dilepton / tau channels.
 - Angular correlations: e.g. average $\Delta R(b,b)$
 - Event kinematics: e.g. sphericity
 - B-tagging information: e.g. average b-tagging output variable
 - $t\bar{t}b\bar{b}/t\bar{t}H$ BDT in signal-rich lepton+jets channels
 - Tau isolation and kinematics (tau channel)

CMS PAS HIG-13-019



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
<i>b</i> -tagging efficiency	SN	6
<i>c</i> -tagging efficiency	SN	6
Light jet-tagging efficiency	SN	12
Background Model		
<i>t</i> \bar{t} cross section	N	1
<i>t</i> \bar{t} modelling: p_T reweighting	SN	9
<i>t</i> \bar{t} modelling: parton shower	SN	2
<i>t</i> \bar{t} +heavy-flavour: normalisation	N	2
<i>t</i> \bar{t} +heavy-flavour: HF reweighting	SN	2
<i>t</i> \bar{t} +heavy-flavour: generator	SN	5
<i>W</i> +jets normalisation	N	3
<i>W</i> p_T reweighting	SN	1
<i>Z</i> +jets normalisation	N	2
<i>Z</i> 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
<i>t</i> \bar{t} <i>V</i> cross section	N	1
Signal Model		
<i>t</i> \bar{t} <i>H</i> modelling	SN	2

% change in yield in ≥ 6 jets/ ≥ 4 tags

≥ 6 j, ≥ 4 b

	Pre-fit			
	<i>t</i> \bar{t} <i>H</i> (125)	<i>t</i> \bar{t} + light	<i>t</i> \bar{t} + <i>c</i> \bar{c}	<i>t</i> \bar{t} + <i>b</i> \bar{b}
Luminosity	± 2.8	± 2.8	± 2.8	± 2.8
Lepton efficiencies	± 1.4	± 1.4	± 1.4	± 1.5
Jet energy scale	± 6.5	± 14	± 10	± 8.2
Jet efficiencies	± 1.6	± 5.4	± 2.5	± 2.4
Jet energy resolution	± 0.1	± 8.5	± 4.1	± 4.3
<i>b</i> -tagging efficiency	± 9.0	± 5.8	± 5.1	± 9.2
<i>c</i> -tagging efficiency	± 1.9	± 7.3	± 14	± 2.8
Light jet-tagging efficiency	± 1.0	± 17	± 4.4	± 1.5
<i>t</i> \bar{t} modelling: reweighting	–	± 11	± 13	± 13
<i>t</i> \bar{t} modelling: parton shower	–	± 7.5	± 1.8	± 10
<i>t</i> \bar{t} heavy-flavour: normalisation	–	–	± 50	± 50
<i>t</i> \bar{t} heavy-flavour: reweighting	–	–	± 11	± 12
<i>t</i> \bar{t} heavy-flavour: generator	–	–	± 2.2	± 2.9
Theoretical cross sections	–	± 6.2	± 6.3	± 6.3
<i>t</i> \bar{t} <i>H</i> modelling	± 1.9	–	–	–
Total	± 12	± 30	± 57	± 56

- Many systematic uncertainties, both theoretical and experimental.
- Background systematics much larger than expected signal yield!

Total background uncertainty: ~37%
 Expected S/B: ~3.8%

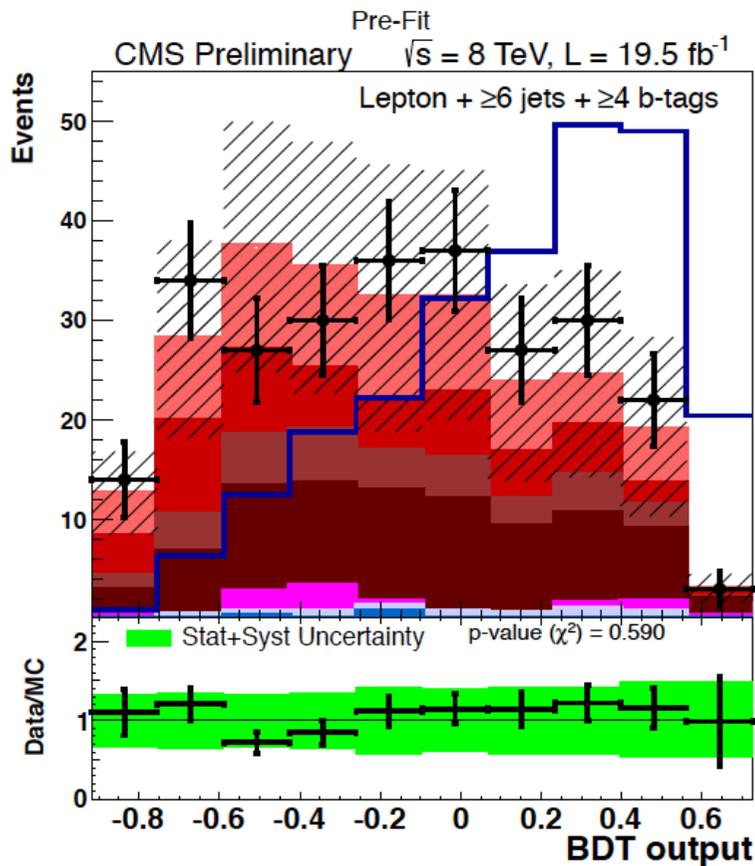
Profiling in Action: Example Plots

- Can exploit high-statistics control samples to constrain the leading syst. uncertainties.
- But need sophisticated enough treatment to not artificially overconstrain them!

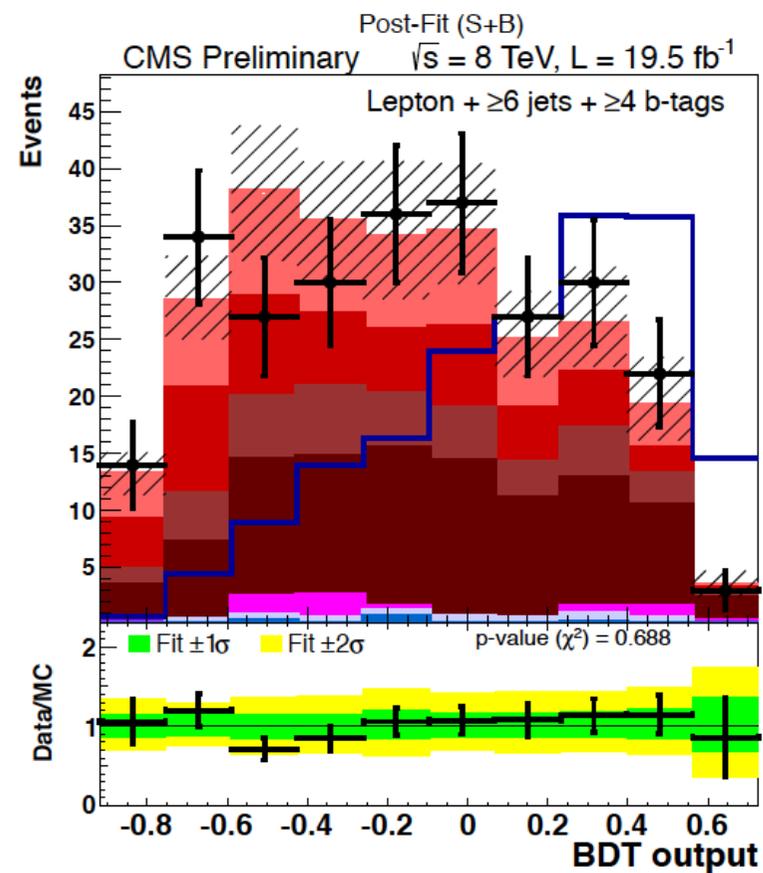
CMS PAS HIG-13-019

Pre-Fit

Post-Fit (S+B)



Total background uncertainty: ~37%



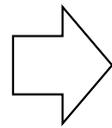
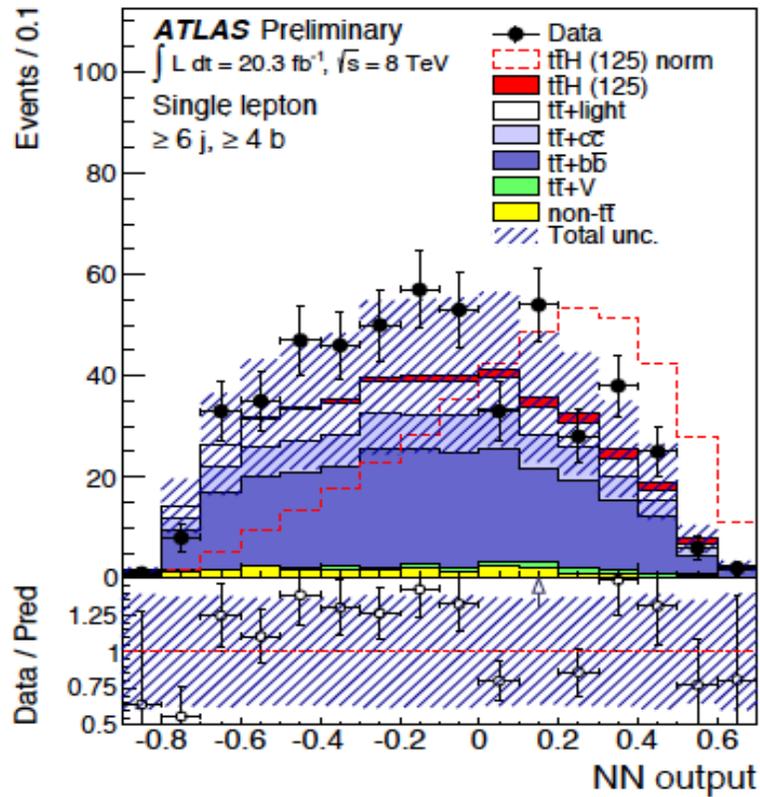
Total background uncertainty: ~7%

Profiling in Action: Example Plots

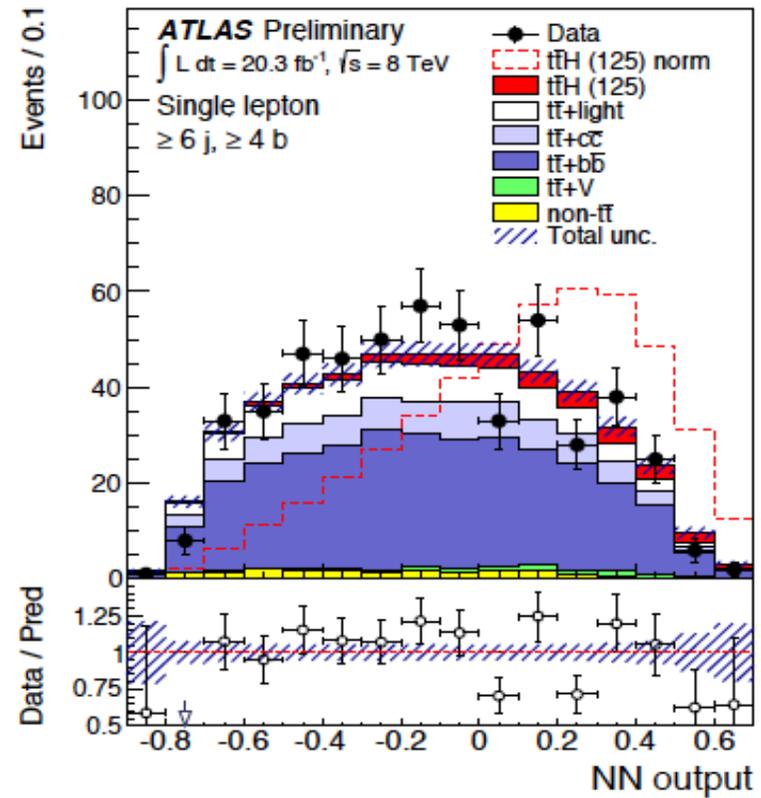
- Can exploit high-statistics control samples to constrain the leading syst. uncertainties.
- But need sophisticated enough treatment to not artificially overconstrain them!

ATLAS-CONF-2014-011

Pre-Fit



Post-Fit (S+B)



Total background uncertainty: ~37%

Total background uncertainty: ~5%

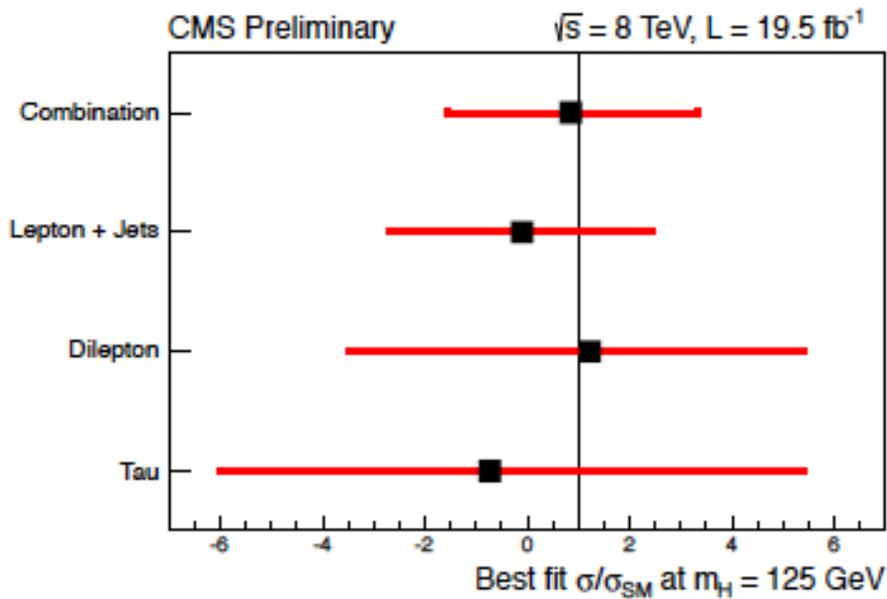
CMS H→bb/ττ Results (8 TeV)



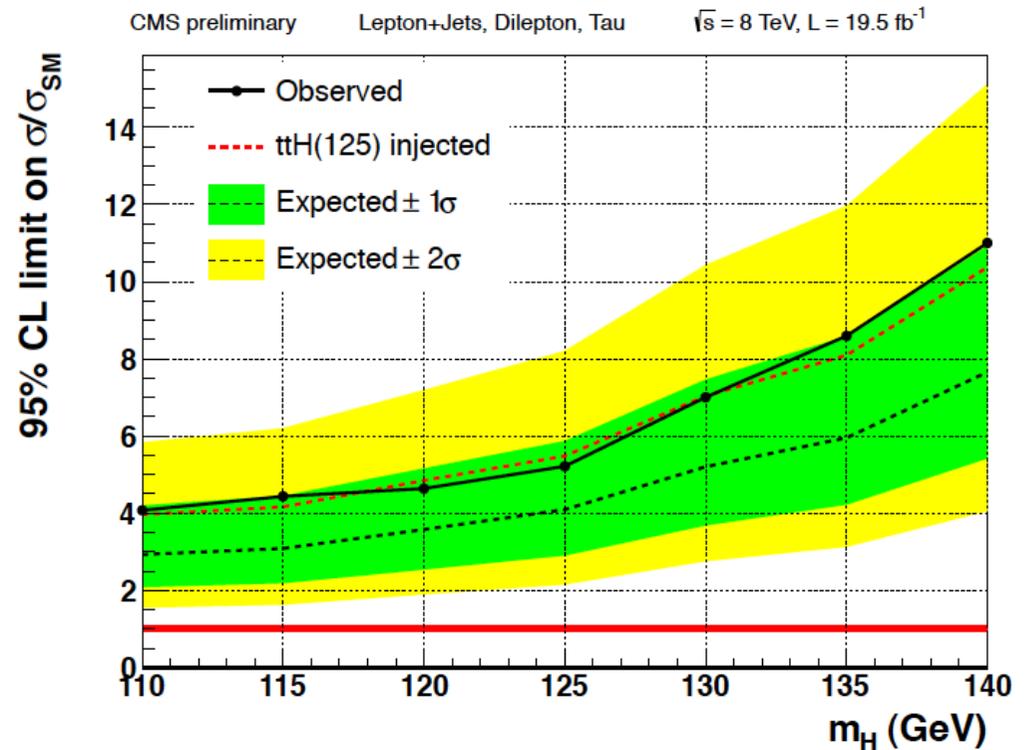
CMS PAS HIG-13-019

- Combination of lepton+jets, dileptons and tau:
Observed (expected) limit @ $M_H=125$ GeV:
5.2xSM (4.1xSM)
- Addition of dilepton and tau channels improves sensitivity by ~15%.

ttH decay mode	Exp	Ob
H→bb, tt→lepton+jets	4.7	4.9
H→bb, tt→dilepton	8.2	9.1
H→tautau, tt→ljets or dilepton	14.2	13.2
Combination	4.1	5.2



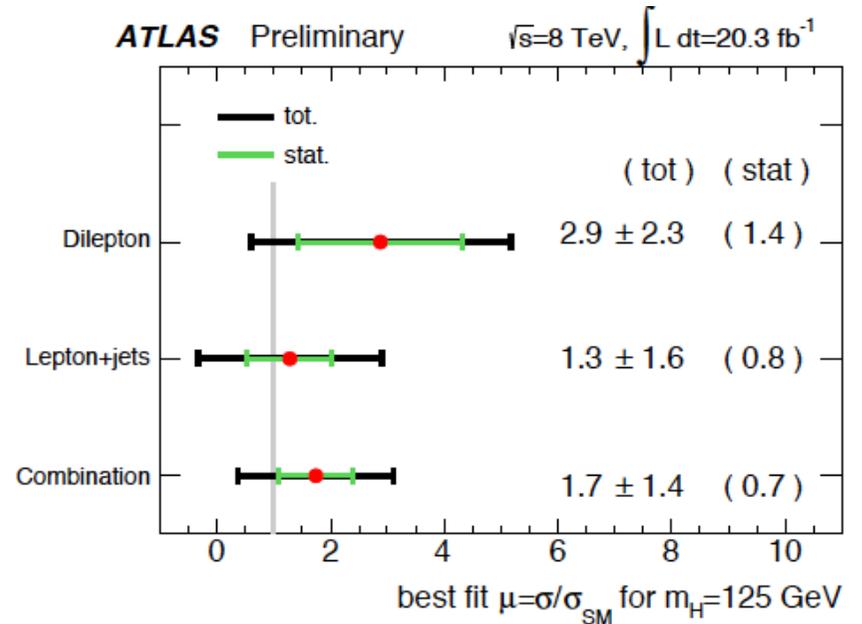
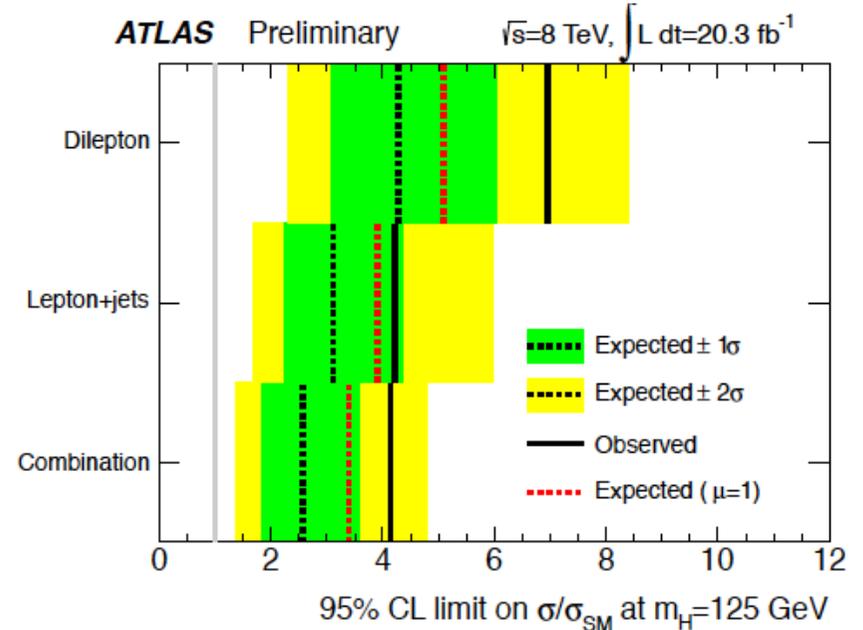
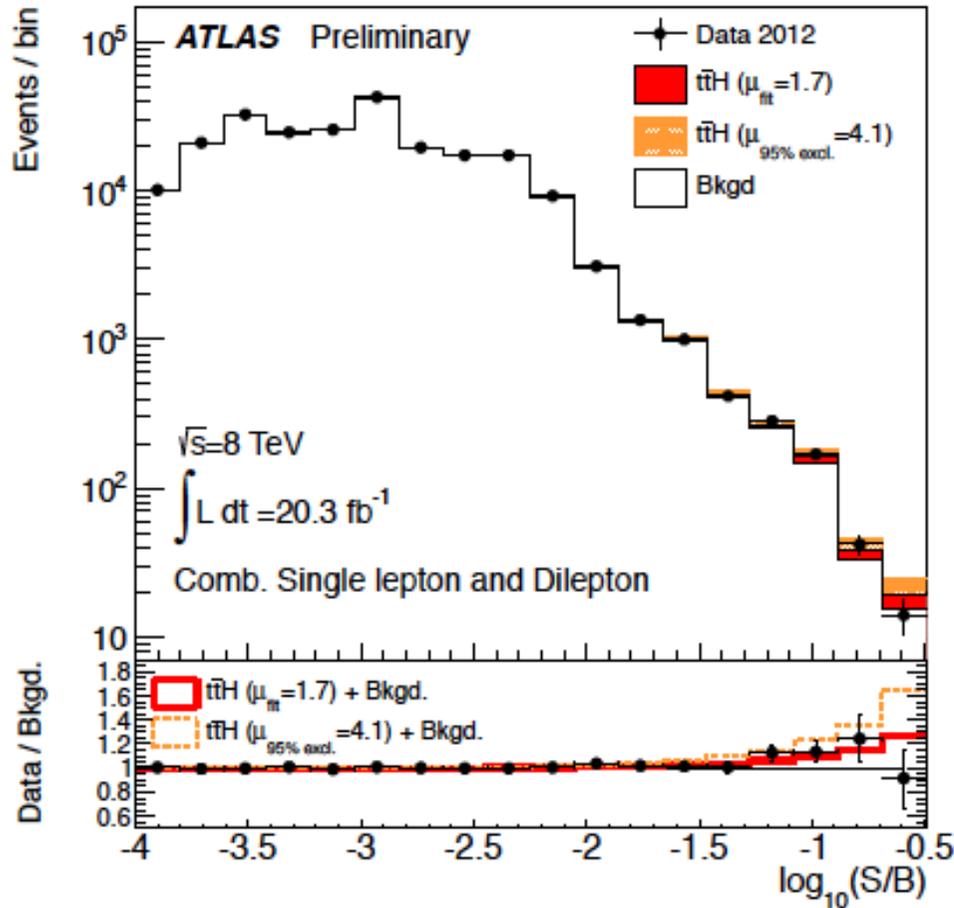
$$\mu_{comb} = 0.85^{+2.47}_{-2.41}$$



ATLAS $H \rightarrow bb$ Result (8 TeV)

- Combination of lepton+jets and dileptons:
Observed (expected) limit @ $M_H=125$ GeV:
4.1xSM (2.6xSM)
- Best-fit signal strength: $\mu_{comb} = 1.7 \pm 1.4$

ATLAS-CONF-2014-011



ttH, Multileptons

Basic Analysis Strategy

- Mainly probe $H \rightarrow WW$, but also non-negligible contributions from $H \rightarrow \tau\tau$ and $H \rightarrow ZZ$.

- Categorize channels by number of leptons.
Ideal signatures for $H \rightarrow WW$:

- SS 2-leptons + 6 jets (2 b jets)
- 3-leptons + 4 jets (2 b jets)
- 4-leptons + 2 jets (2 b jets)

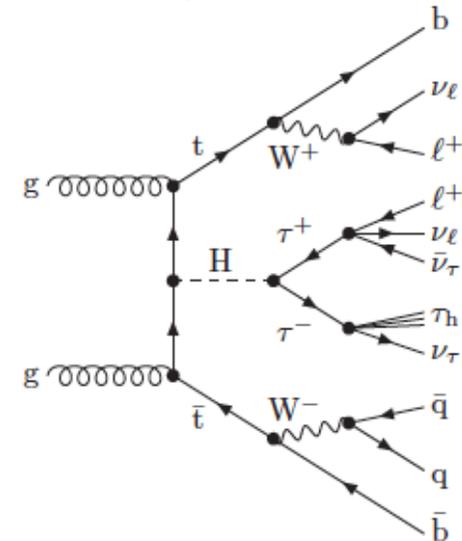
- Low signal rate but also low background, dominated by $t\bar{t}W/Z/\gamma^*$.

Additional contributions from WZ and ZZ .

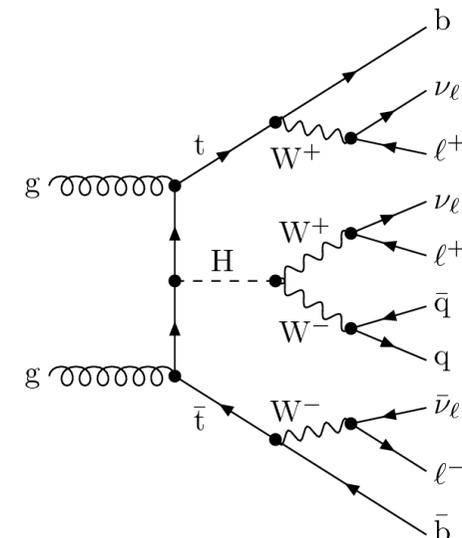
For SS 2-lepton and 3-lepton analyses, sizable contribution from $t\bar{t}$ +jets, with jets misidentified as leptons.

- Use multivariate discriminants to separate signal from backgrounds.

SS 2-leptons channel



3-leptons channel





Event Selections

SS 2-leptons channel

$e^{\pm}e^{\pm}, \mu^{\pm}\mu^{\pm}, e^{\pm}\mu^{\pm}$ final states
2 tight e/μ , $p_T > 20$ GeV; Z veto
 ≥ 4 jets, $p_T > 25$ GeV (anti- k_T $R=0.5$)
 ≥ 2 b-tags
 $E_T^{\text{miss}} \text{ LD} > 0.2, \Sigma p_T^{\text{lep}} + E_T^{\text{miss}} > 100$ GeV

3-leptons channel

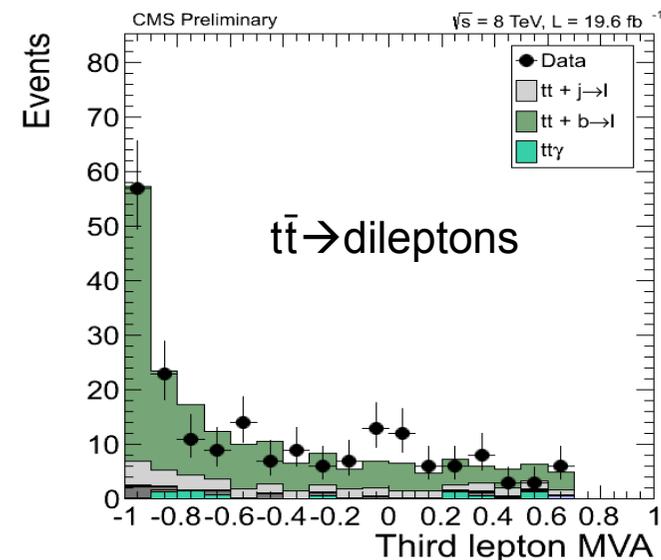
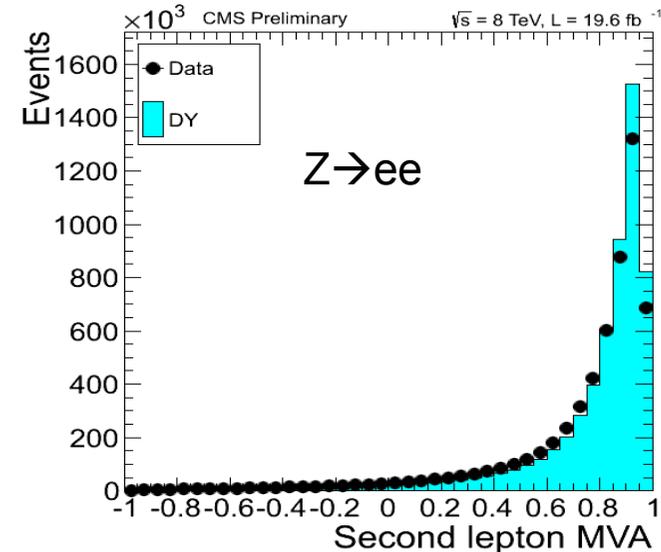
3 tight e/μ , $p_T > 20/10/10$ GeV; Z veto
 ≥ 2 jets, $p_T > 25$ GeV (anti- k_T $R=0.5$)
 ≥ 2 b-tags
 E_T^{miss} LD cut if SF/OS pair and < 4 jets

4-leptons channel

4 loose e/μ , $p_T > 20/10/10/10$ GeV; Z veto
 ≥ 2 jets, $p_T > 25$ GeV (anti- k_T $R=0.5$)
 ≥ 2 b-tags

- MVA-based lepton identification (trained on $t\bar{t}H$ vs $t\bar{t}$ +jets MC):

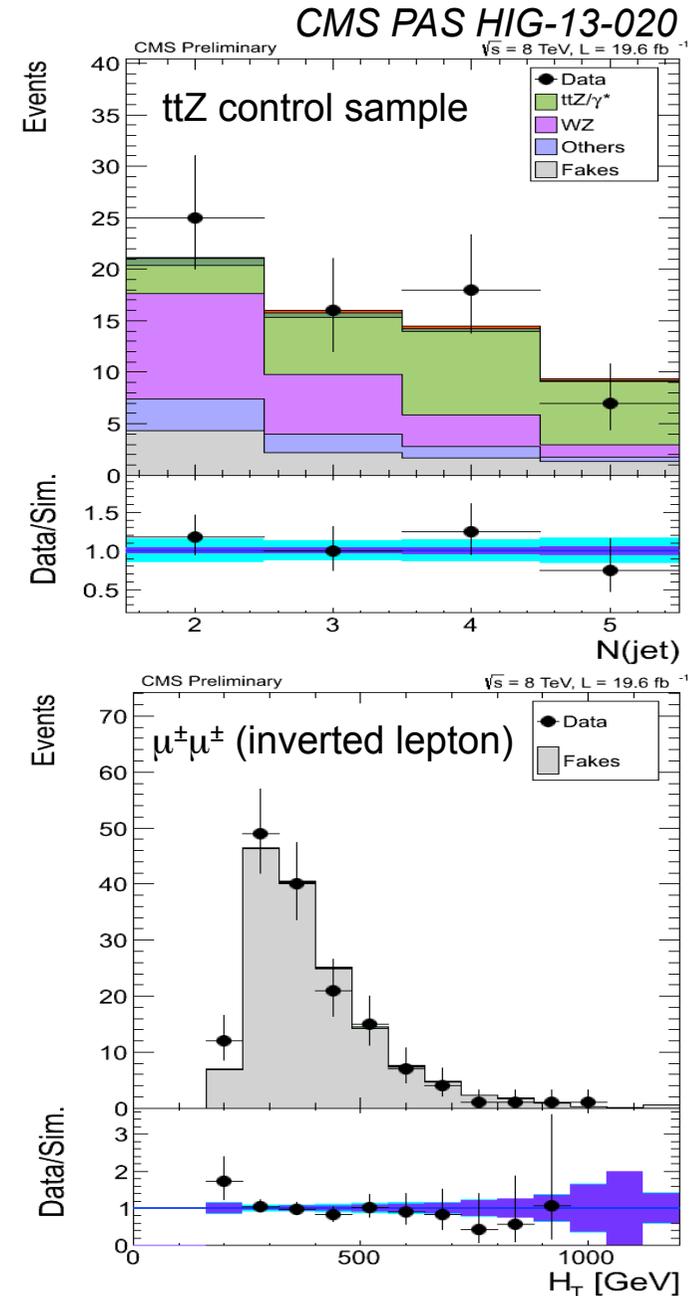
CMS PAS HIG-13-020





Background Estimation

- **ttV+jets (V=W, Z, WW)**
 - Predicted using Madgraph+Pythia MC normalized to NLO cross section ($\sim 13\%$ uncertainty). Additional uncertainties from varying scale choices in the MC.
 - ttZ prediction validated in data control sample ($\sim 35\%$ statistical uncertainty).
- **Dibosons+jets (WZ,ZZ)**
 - Predicted using Madgraph+Pythia MC calibrated in data control samples (WZ+ ≥ 2 +non-b jets and ZZ+ ≥ 1 +non-b jets).
 - Total uncertainty $\sim 20\%$ (includes uncertainty in extrapolation from control to signal region).
- **tt+jets instrumental**
 - Fake leptons: data events with inverted lepton MVA corrected with per-lepton fake rate. Uncert. $\sim 50\%$.
 - Charge misID (SS 2-leptons): OS 2-leptons data events corrected with per-lepton charged misID rate. Uncert. $\sim 30\%$.



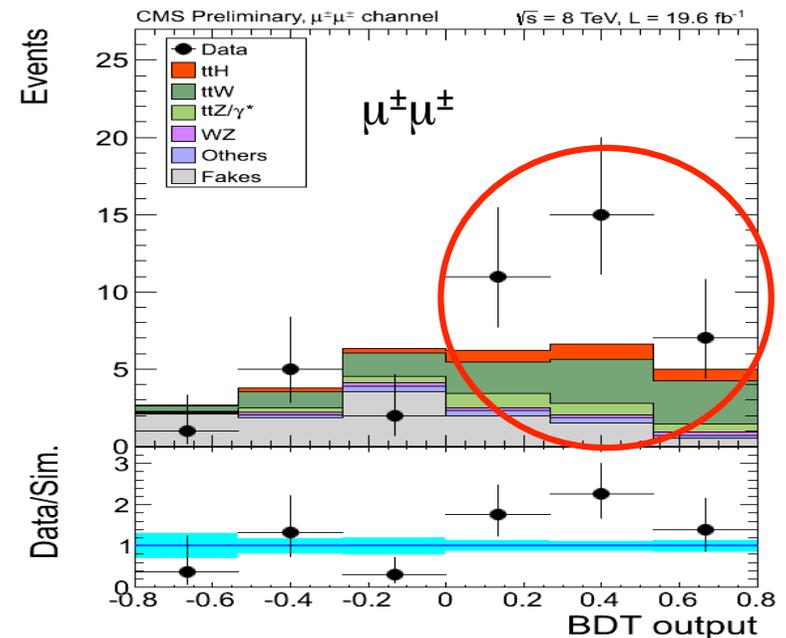
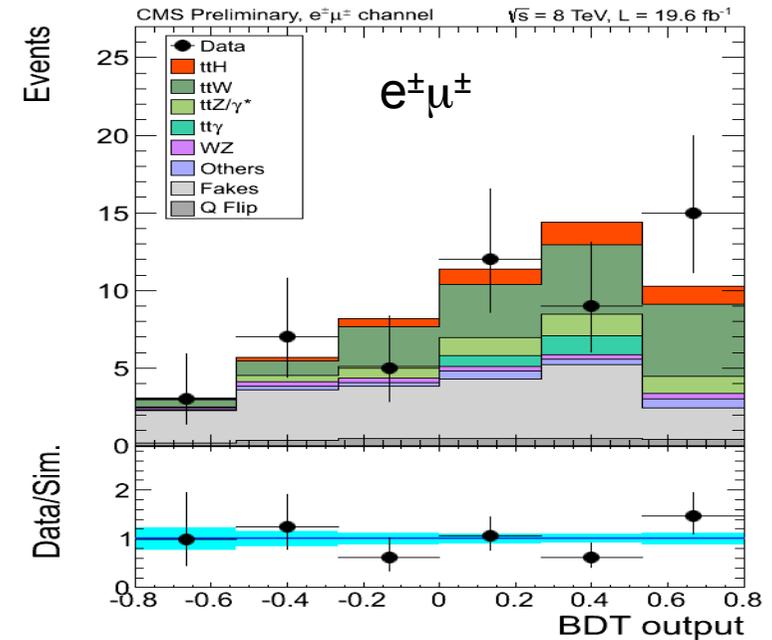


SS 2-Leptons

- Analyze separately e^+e^- , $\mu^+\mu^-$, $e^\pm\mu^\pm$ events.
- Events categorized according to lepton charge sum (exploit charge correlation in ttW).
- BDT trained between ttH signal and tt +jets background MC (6 vars: e.g. H_T).
- $\sim 3\sigma$ excess (wrt SM) in $\mu^\pm\mu^\pm$ channel.
Cross-checks performed show no issues with data quality or background mismodeling.

	$\mu\mu$	ee	$e\mu$
$t\bar{t}H, H \rightarrow WW$	2.0 ± 0.3	0.9 ± 0.1	2.7 ± 0.4
$t\bar{t}H, H \rightarrow ZZ$	0.1 ± 0.0	0.0 ± 0.0	0.1 ± 0.0
$t\bar{t}H, H \rightarrow \tau\tau$	0.6 ± 0.1	0.3 ± 0.0	0.9 ± 0.1
$t\bar{t}W$	8.2 ± 1.5	3.4 ± 0.6	13.0 ± 2.2
$t\bar{t}Z/\gamma^*$	2.5 ± 0.5	1.6 ± 0.3	4.2 ± 0.9
$t\bar{t}WW$	0.2 ± 0.0	0.1 ± 0.0	0.3 ± 0.1
$t\bar{t}\gamma$	-	1.3 ± 0.3	1.9 ± 0.5
WZ	0.8 ± 0.9	0.5 ± 0.5	1.2 ± 1.3
ZZ	0.1 ± 0.1	0.0 ± 0.0	0.1 ± 0.1
rare SM bkg.	1.1 ± 0.0	0.4 ± 0.0	1.5 ± 0.0
non-prompt	10.8 ± 4.8	8.9 ± 4.5	21.2 ± 8.1
charge flip	-	1.9 ± 0.6	2.4 ± 0.8
all signals	2.7 ± 0.4	1.2 ± 0.2	3.7 ± 0.6
all backgrounds	23.7 ± 5.2	18.0 ± 4.7	45.9 ± 8.6
data	41	19	51

CMS PAS HIG-13-020





3-Leptons and 4-Leptons

3-leptons:

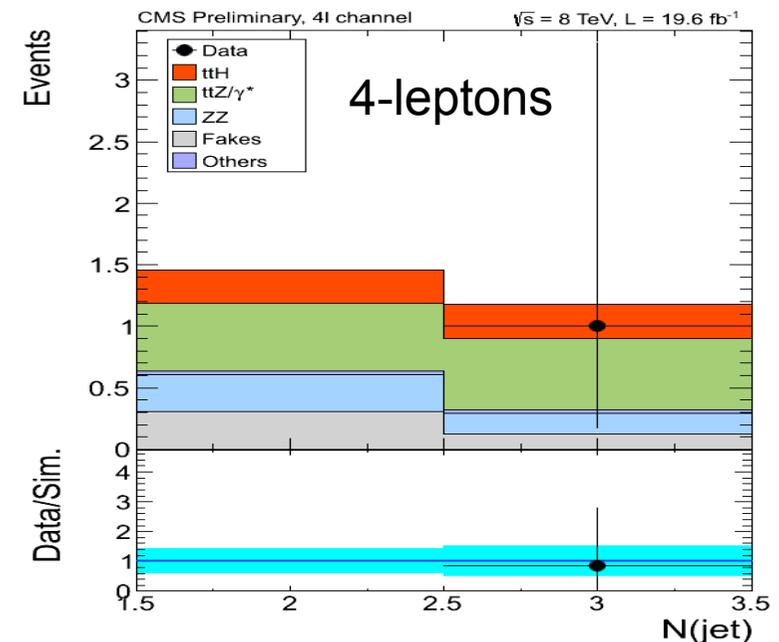
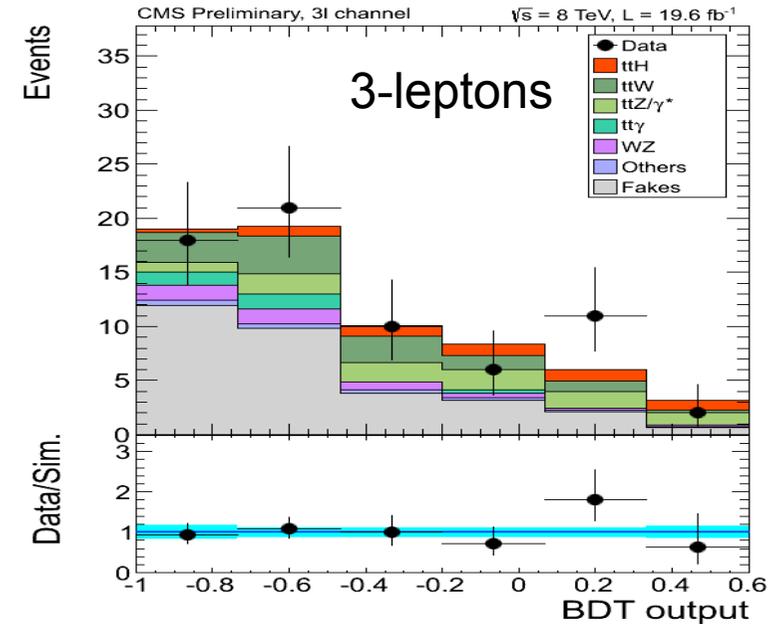
- Events categorized according to lepton charge sum (exploit charge correlation in ttW).
- BDT trained between ttH signal and mixture of tt+jets and ttV+jets background MC (7 vars: e.g. N_{jets}).

4-leptons:

- Use N_{jets} as discriminating variable.

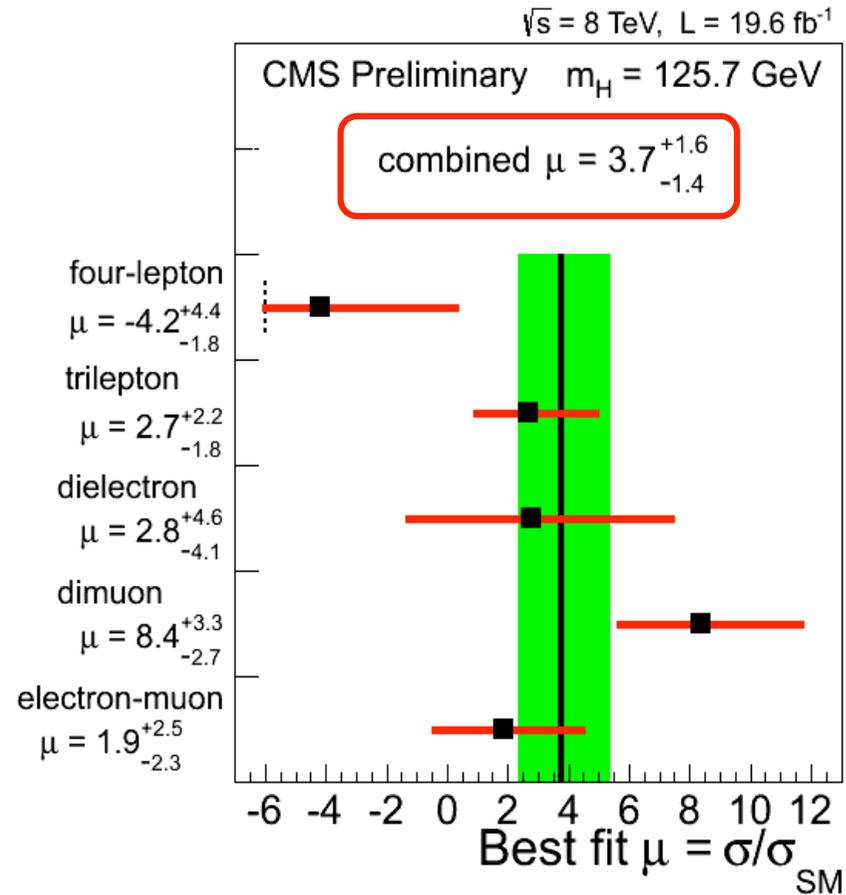
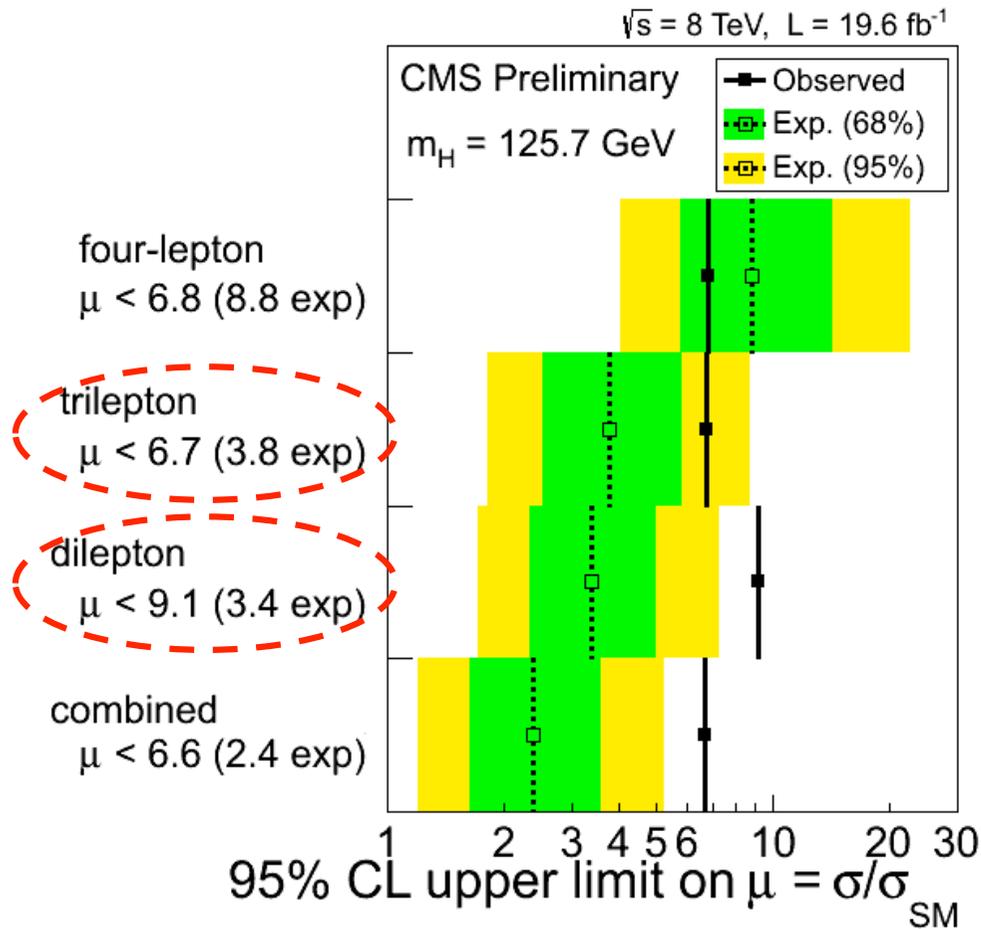
	3ℓ	4ℓ
$t\bar{t}H, H \rightarrow WW$	3.2 ± 0.6	0.28 ± 0.05
$t\bar{t}H, H \rightarrow ZZ$	0.2 ± 0.0	0.09 ± 0.02
$t\bar{t}H, H \rightarrow \tau\tau$	1.0 ± 0.2	0.15 ± 0.02
$t\bar{t}W$	9.2 ± 1.9	-
$t\bar{t}Z/\gamma^*$	7.9 ± 1.7	1.25 ± 0.88
$t\bar{t}WW$	0.4 ± 0.1	0.04 ± 0.02
$t\bar{t}\gamma$	2.9 ± 0.8	-
WZ	4.2 ± 0.9	-
ZZ	0.4 ± 0.1	0.45 ± 0.09
rare SM bkg.	0.8 ± 0.0	0.01 ± 0.00
non-prompt charge flip	33.2 ± 12.3	0.53 ± 0.32
	-	-
all signals	4.4 ± 0.8	0.52 ± 0.09
all backgrounds	58.9 ± 12.7	2.28 ± 0.94
data	68	1

CMS PAS HIG-13-020





CMS Multilepton Result



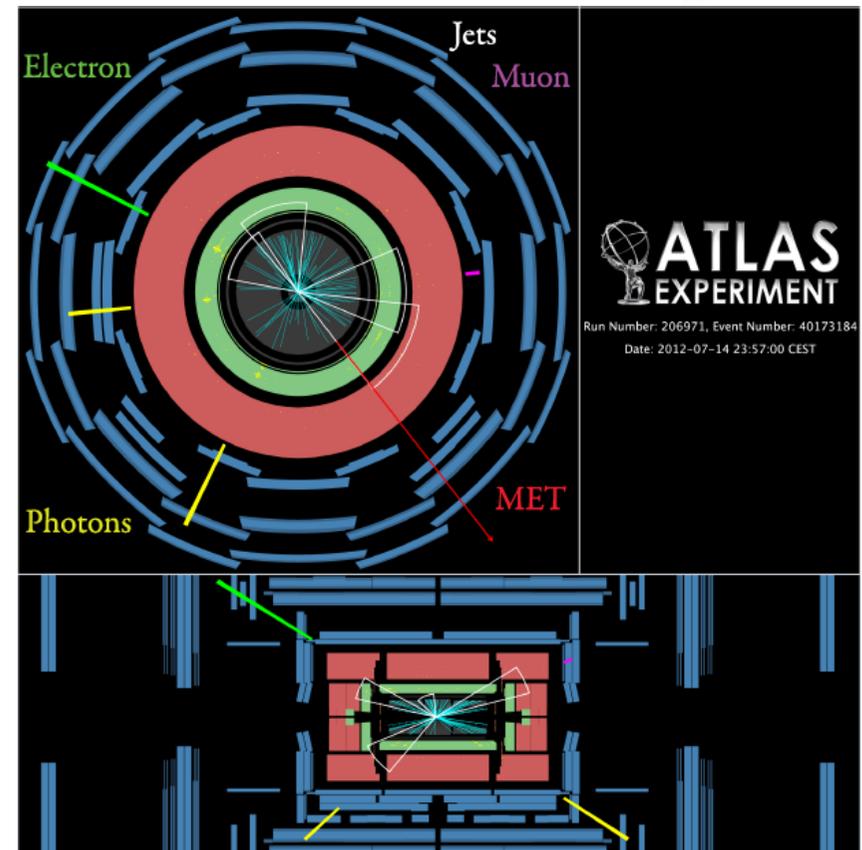
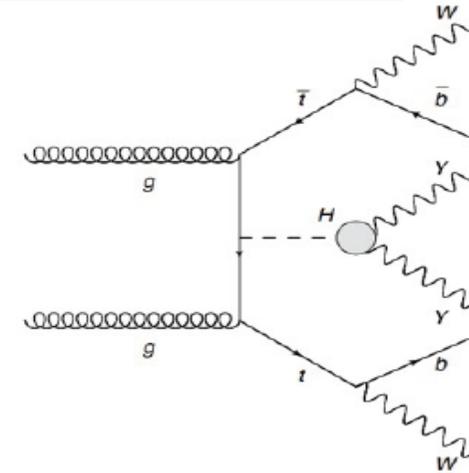
CMS PAS HIG-13-020

- Excellent sensitivity of SS 2-leptons and 3-lepton channels!
- Combination of multilepton analyses:
 Observed (expected) limit @ $M_H=125 \text{ GeV}$: $6.6xSM$ ($2.4xSM$)

$ttH, H \rightarrow \gamma\gamma$

Basic Analysis Strategy

- Small $BR(H \rightarrow \gamma\gamma)$ is compensated by the much smaller backgrounds and the good diphoton mass resolution.
- Capitalize on well-understood $H \rightarrow \gamma\gamma$ analyses.
- Categorize signal events according to the $t\bar{t}$ decay mode (leptonic or hadronic).
Exploit high jet multiplicity and b-jet content of signal to optimize sensitivity.
- Discriminant variable: $m_{\gamma\gamma}$
No need (for now) to estimate complicated background composition. Can perform sideband analysis as in standard $H \rightarrow \gamma\gamma$ analyses.



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

2 photons, $p_{T1}=61$ GeV, $p_{T2}=39$ GeV
 $m_{\gamma\gamma}=126.6$ GeV
 1 electron, $p_T=90$ GeV
 $E_T^{miss} = 43$ GeV
 4 jets, $p_T=75, 71, 50, 39$ GeV, 1 muon-tagged jet

Event Selections

Leptonic channel



2 photons, $p_T > 40 / 30$ GeV
 ≥ 1 e/μ , $p_T > 15/10$ GeV
 $E_T^{\text{miss}} > 20$ GeV
 ≥ 1 jets, $p_T > 25$ GeV
 ≥ 1 b-tags (80% efficiency)



2 photons, $p_T > 0.5m_{\gamma\gamma}/25$ GeV
 ≥ 1 e or μ , $p_T > 20$ GeV
 No E_T^{miss} cut
 ≥ 1 jets, $p_T > 25$ GeV
 ≥ 1 b-tags (70% efficiency)

Hadronic channel



2 photons, $p_T > 40/30$ GeV
 0 leptons
 ≥ 6 jets, $p_T > 25$ GeV
 ≥ 2 b-tags (80% efficiency)



2 photons, $p_T > 0.5m_{\gamma\gamma}/25$ GeV
 0 leptons
 ≥ 4 jets, $p_T > 25$ GeV
 ≥ 1 b-tags (80% efficiency)



High $t\bar{t}H$ purity selections



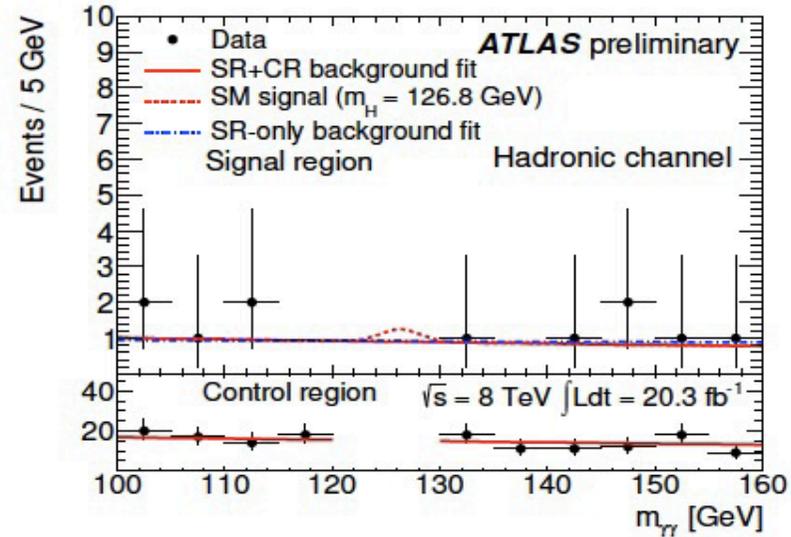
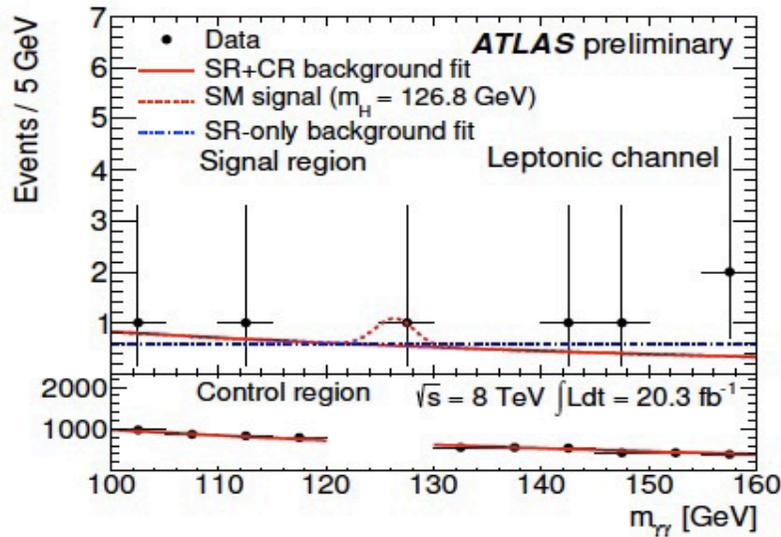
Channel	N_S	$ggF(\%)$	$VBF(\%)$	$WH(\%)$	$ZH(\%)$	$tH(\%)$	$t\bar{t}H(\%)$
Leptonic	0.55	0.6	0.3	7.7	2.4	6.1	82.8
Hadronic	0.36	5.3	1.1	1.1	1.3	—	91.2

Process	Hadronic Channel	Leptonic Channel
$t\bar{t}H$	0.567 (87%)	0.429 (97%)
$gg \rightarrow H$	0.059 (9%)	0 (0%)
VBF H	0.006 (1%)	0 (0%)
WH/ZH	0.019 (3%)	0.013 (3%)
Total signal	0.65	0.44

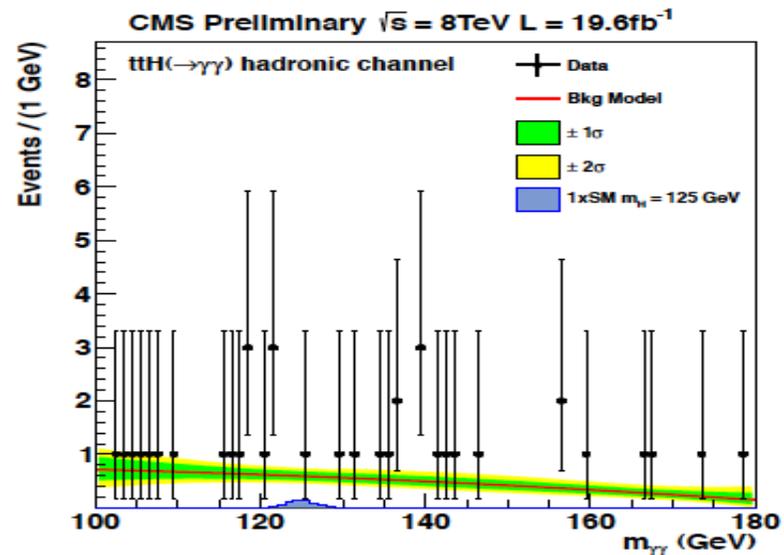
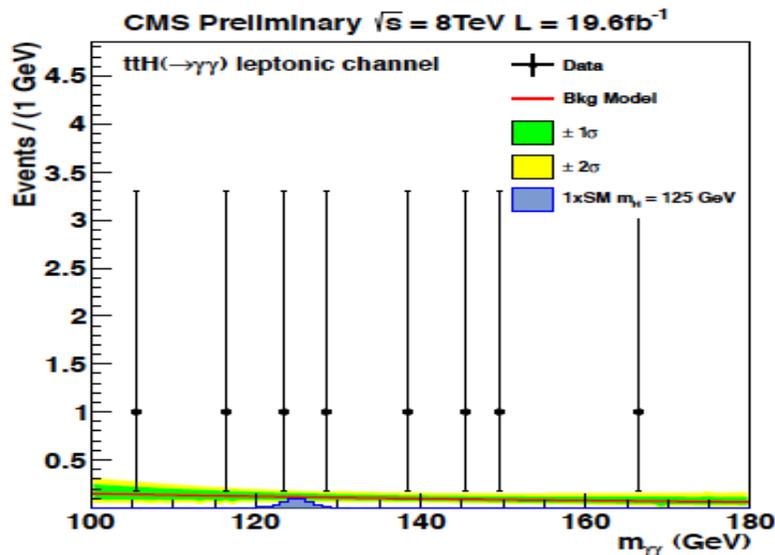
Background Estimation

- ATLAS: exponential fit to both signal region and control regions with relaxed cuts.
- CMS: exponential (leptonic channel) or 2nd order polynomial (hadronic channel) fit in signal region.

ATLAS-CONF-2013-080



CMS PAS HIG-13-015



ATLAS $H \rightarrow \gamma\gamma$ Result

- Observed (expected) limit @ $M_H=126.8$ GeV:
5.3xSM (6.4xSM)

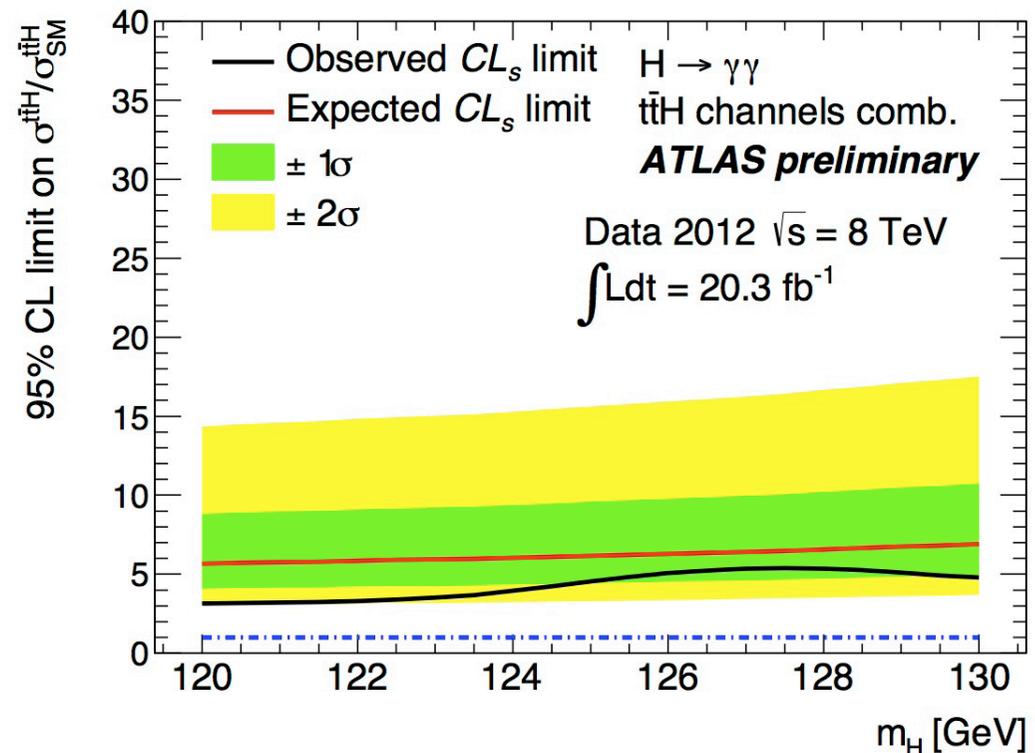
- Search statistically-limited.
Small impact from systematic uncertainties,
in contrast with $t\bar{t}H$, $H \rightarrow b\bar{b}$.

- Good signal-to-background:

Channel	N_S	N_B	N_S/N_B
Leptonic	0.55	$1.2^{+0.6}_{-0.5}$	0.45
Hadronic	0.36	$1.9^{+0.7}_{-0.5}$	0.19

Limits @ $M_H=126.8$ GeV

	Observed limit	Expected limit
Combined (with systematics)	5.3	6.4
Combined (statistics only)	5.0	6.0
Leptonic (with systematics)	9.0	8.4
Leptonic (statistics only)	8.5	8.0
Hadronic (with systematics)	8.4	13.6
Hadronic (statistics only)	7.9	12.6



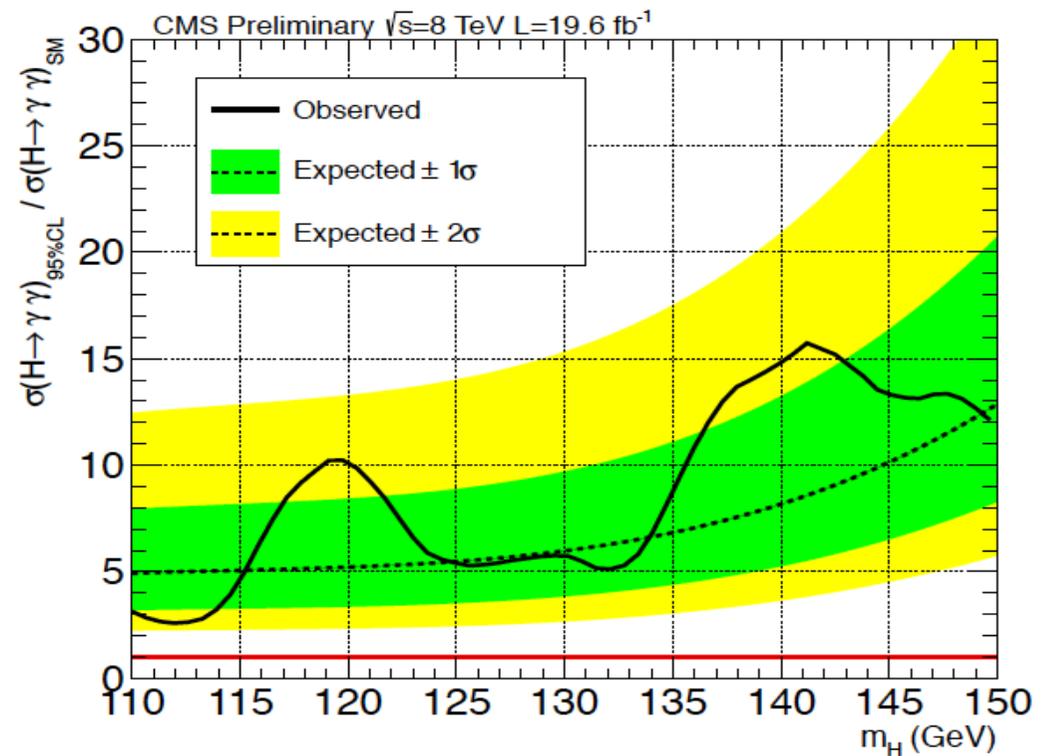
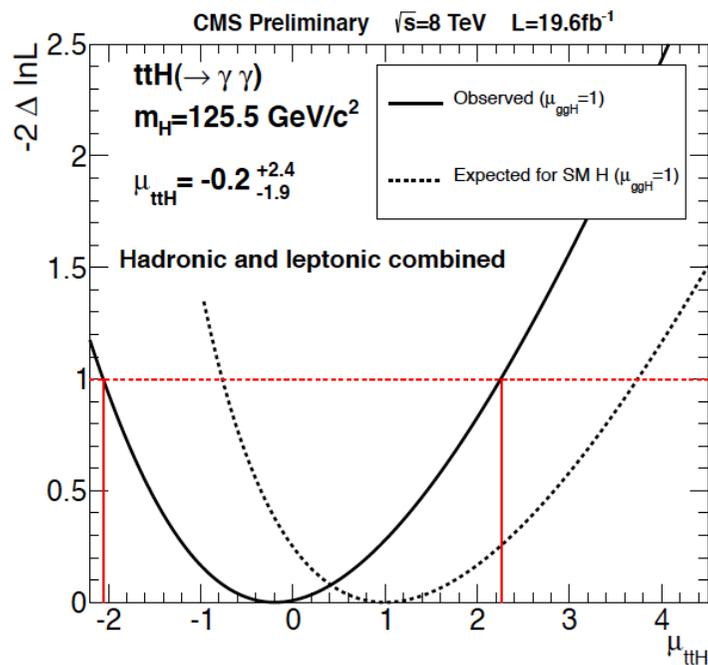
CMS $H \rightarrow \gamma\gamma$ Result

- Observed (expected) limit @ $M_H=125$ GeV:
5.4xSM (5.3xSM)

Limits @ $M_H=125$ GeV

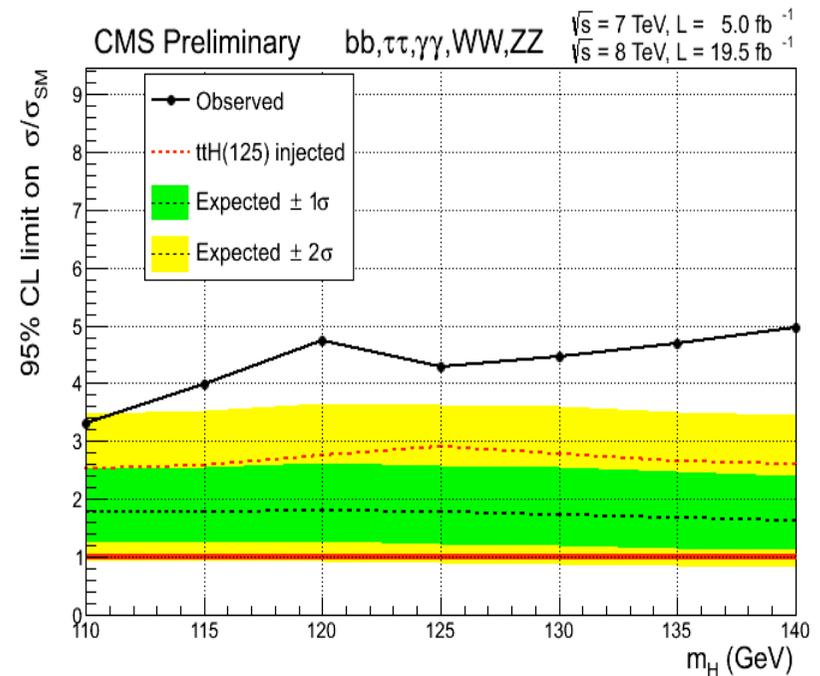
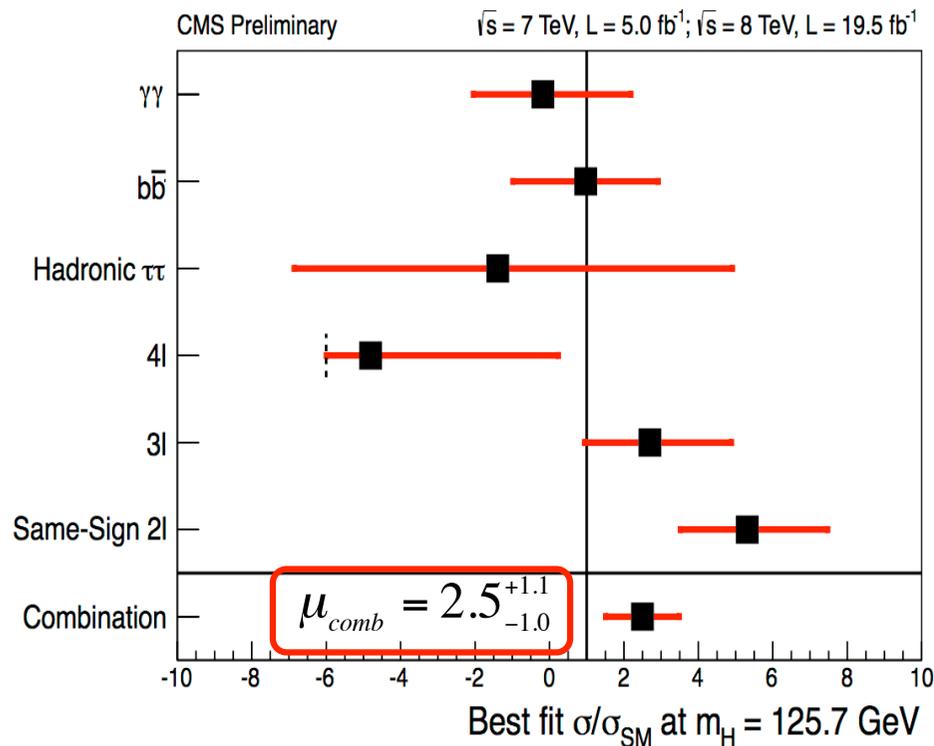
	Observed	Expected	Expected (No Syst.)
Hadronic Channel	6.8	9.2	8.8
Leptonic Channel	10.7	8.0	7.7
Combined	5.4	5.3	5.1

- Search statistically-limited.
Small impact from systematic uncertainties,
in contrast with $t\bar{t}H$, $H \rightarrow b\bar{b}$.



CMS Grand Combination

- **Combination of CMS results:**
 - $H \rightarrow b\bar{b}$ (7 TeV pub + 8 TeV prelim)
 - $H \rightarrow \tau\tau$ (8 TeV prelim)
 - SS 2-lep, 3-lep, 4-lep (8 TeV prelim)
 - $H \rightarrow \gamma\gamma$ (8 TeV prelim)
- **Obs (exp) limit @ $M_H=125$ GeV:**
4.3xSM (1.8xSM).



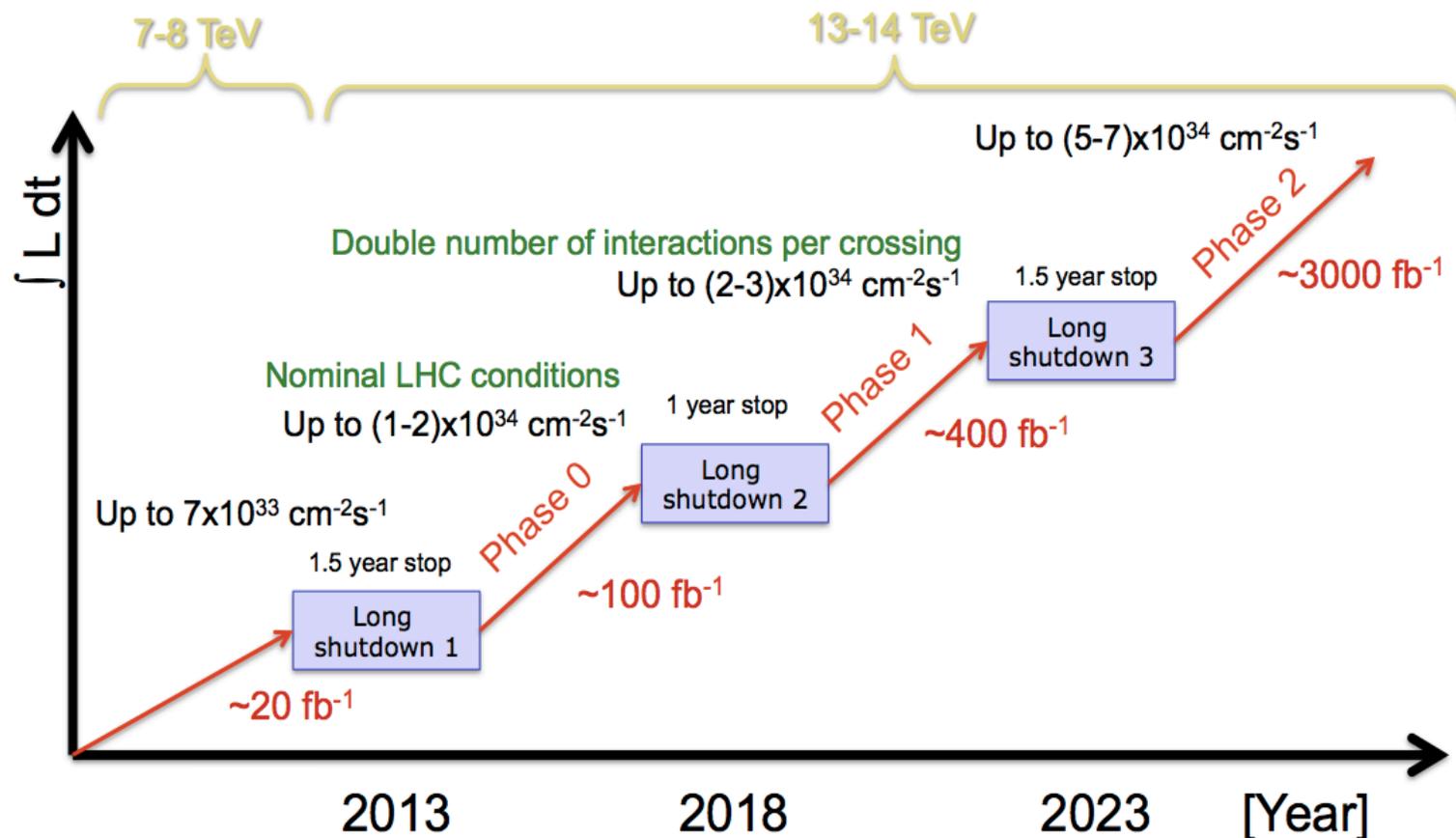
	Observed	Median Signal Injected	Median
$\gamma\gamma$	5.4	6.7	5.5
$b\bar{b}$	4.5	5.2	3.7
$\tau\tau$	12.9	16.2	14.2
4l	6.8	11.9	8.8
3l	6.7	4.7	3.8
Same-sign 2l	9.1	3.6	3.4
Combined	4.3	2.9	1.8

Comparable sensitivity of $H \rightarrow b\bar{b}$ and multilepton channels!

Prospects

LHC Prospects

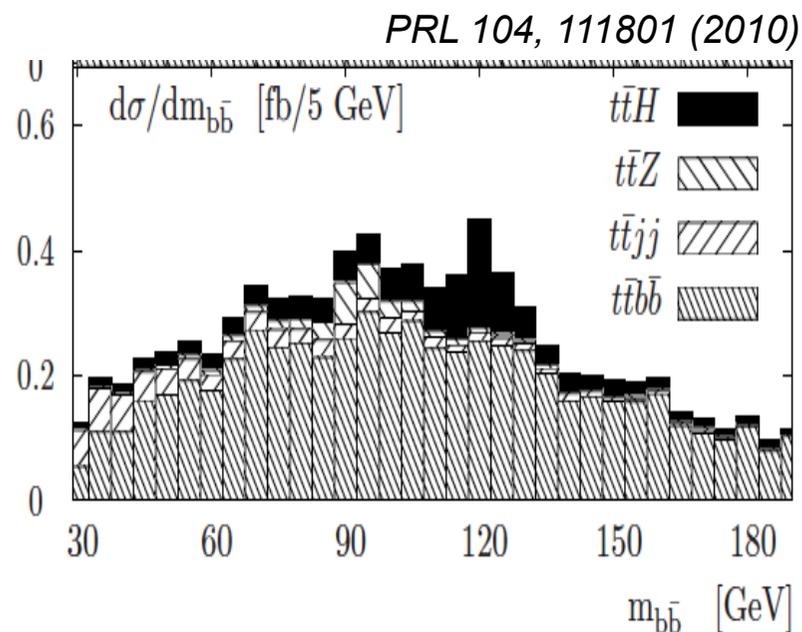
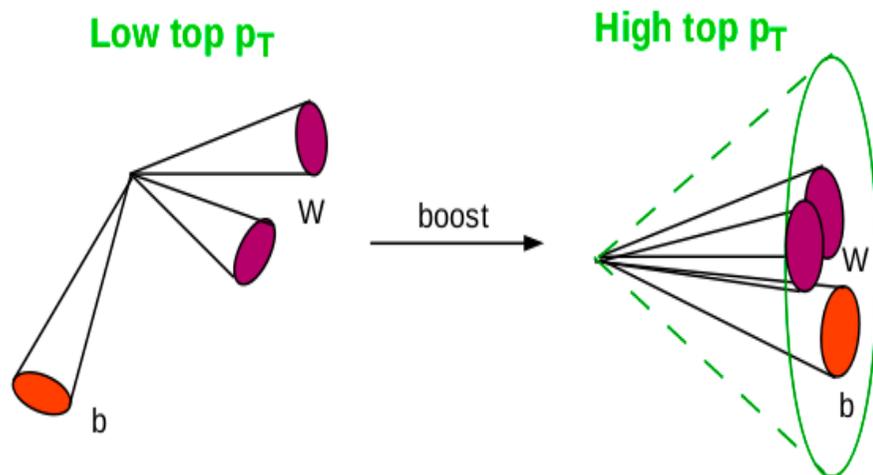
- Short term:
 - ATLAS searches using full Run 1 dataset ongoing in similar channels as for CMS. Expect results to be available during Summer 2014.
 - **Combination of ATLAS+CMS Run 1 results should be close to SM sensitivity!**
- Longer term:
 - We are at the beginning of a 20-year program! Much potential to be exploited.



LHC Prospects

H \rightarrow bb

- It has been pointed out that $\sqrt{s}=14$ TeV “boosted” analyses can potentially achieve $\sim 5\sigma$ significance with 100 fb^{-1} .
 - Requiring high enough p_T for hadronic top quark and/or Higgs boson allows to significantly reduce both physics and combinatorial backgrounds.
 - These techniques have by now become “standard” at the LHC experiments in searches for boosted bosons and top quarks.
 - Experimental searches for boosted ttH just starting. “Resolved” and “boosted” analyses will likely co-exists and be combined at the end.



LHC Prospects

arXiv:1307.7280

Multilepton (3-leptons and 4-leptons)

- Recently studied in the context of the European Strategy and Snowmass.
- Analysis still statistically limited with 300 fb⁻¹.
- At very high-luminosity (and pileup), sensitivity may be dominated by 4-leptons due to significant contribution from fake leptons in 3-leptons analysis.
- For 3000 fb⁻¹, experimental uncertainty on top Yukawa ~5%.
- Theoretical uncertainty on $\sigma(ttH)$ adds 8% in quadrature!

Stat. uncert. on $\sigma(ttH)$

Channel	300 fb ⁻¹	3000 fb ⁻¹
3 ℓ only	25%	—%
4 ℓ only	34%	12%
Combined	21%	9%

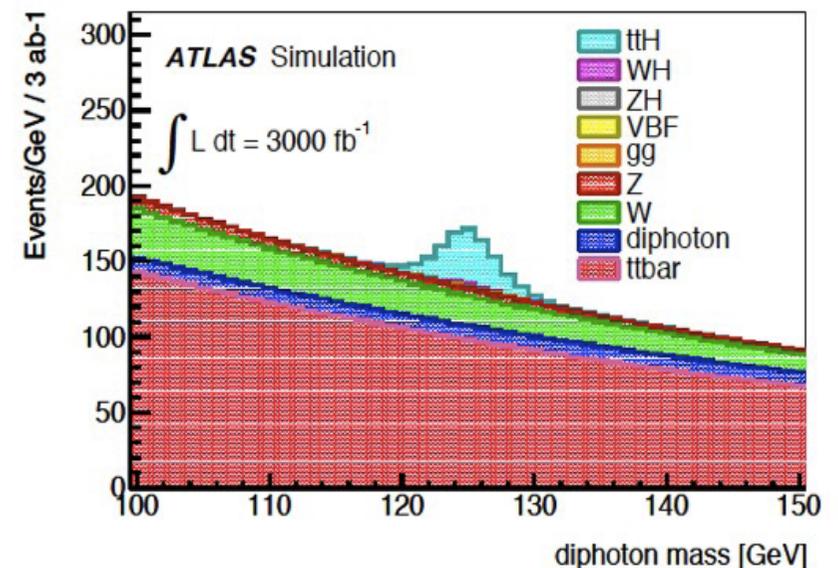
Syst. uncert. on $\sigma(ttH)$

Channel	300 fb ⁻¹	3000 fb ⁻¹
Top fake rate	17%	2%
$\sigma(t\bar{t}H)_{SM}$	16%	16%
Other cross-section systematics	8%	3%
All systematics	27%	17%
Systematics without $\sigma(t\bar{t}H)_{SM}$	18%	4%

H $\rightarrow\gamma\gamma$

- Recently studied in the context of the European Strategy and Snowmass.
 - Analysis statistically limited.
- For 3000 fb⁻¹ at $\sqrt{s}=14$ TeV:
 Expected uncertainty on signal strength ~20%
 → 10% uncertainty on top Yukawa coupling

ATL-PHYS-PUB-2013-007



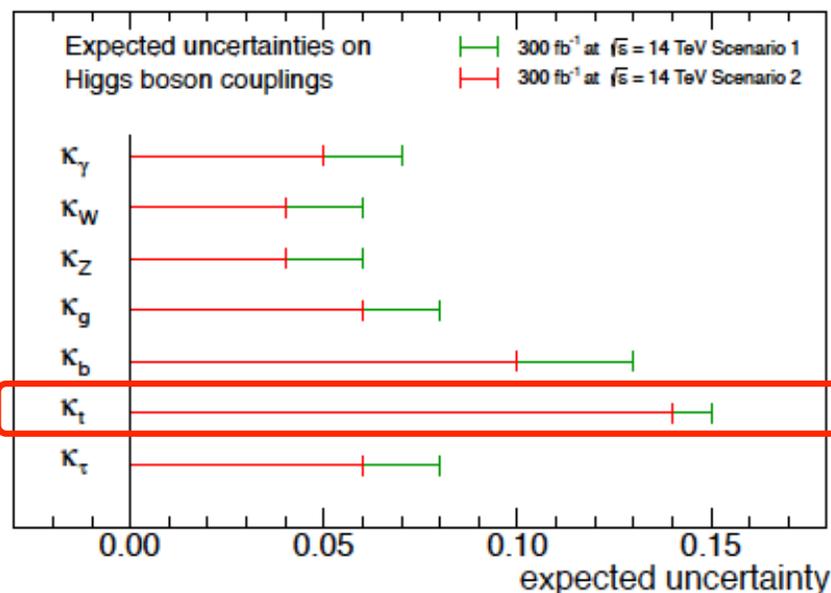
LHC Prospects

arXiv:1307.7135

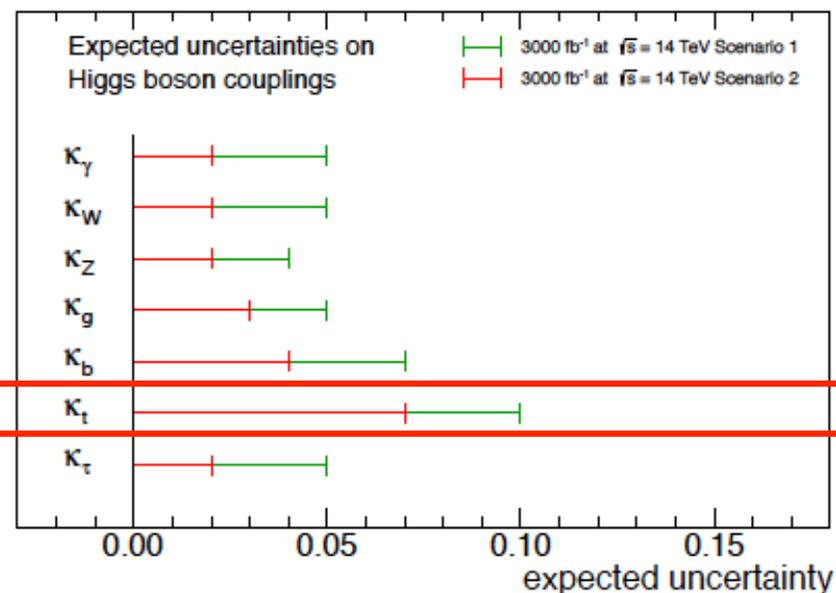
Global Fit Analysis

- Extrapolation global fit to Higgs couplings based on existing CMS Higgs analyses.
 - **ttH: only considering $H \rightarrow \gamma\gamma$ and $H \rightarrow bb$. Will get better after including multileptons.**
- Consider two scenarios:
 - Scenario 1: all systematic uncertainties are left unchanged.
 - Scenario 2: theoretical uncertainties are scaled by a factor of 1/2, while other systematic uncertainties are scaled by the square root of the integrated luminosity.

CMS Projection



CMS Projection



Tentative conclusion: a ~5-10% uncertainty on the top Yukawa coupling at the LHC may be realistic!

Summary

- The precise measurement of the top Yukawa coupling may provide insights on the underlying mechanism for EWSB and whether or not the top quark plays a role in it.
- The program of searches for $t\bar{t}H$ production at the LHC is well underway, with all main decay modes being explored ($H \rightarrow b\bar{b}$, $\tau\tau$, WW and ZZ).
- $H \rightarrow b\bar{b}$ has turned out to be just as challenging as anticipated. Much experience has been gained on how to reduce the impact from systematic uncertainties that led to abandon this channel in the past, by exploiting high-statistics data samples. However, further progress is needed on the theoretical description of the dominant $t\bar{t}$ +jets background. Jet substructure techniques hold great promise in this channel.
- At the same time, searches for $t\bar{t}H$ in diphoton and multilepton final states are showing interesting sensitivity that can be competitive with (or even exceed) that of $H \rightarrow b\bar{b}$.
- A LHC combination could in principle approach SM sensitivity with the Run 1 datasets.

Exciting prospects for the $\sqrt{s}=13/14$ TeV run!

Extra

Motivation

- After the discovery of $h(125)$, the focus is on the precise measurement of its properties, in particular couplings to fermions and gauge bosons.
- The top quark is the most strongly-coupled SM particle to the Higgs boson.

For $m_t = 173$ GeV:

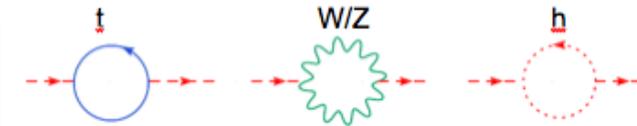
$$\lambda_t = \frac{\sqrt{2}m_t}{v} = 0.996 \pm 0.005$$

- Only quark with a “natural mass”.
- Main responsible for instability of Higgs mass against radiative corrections.

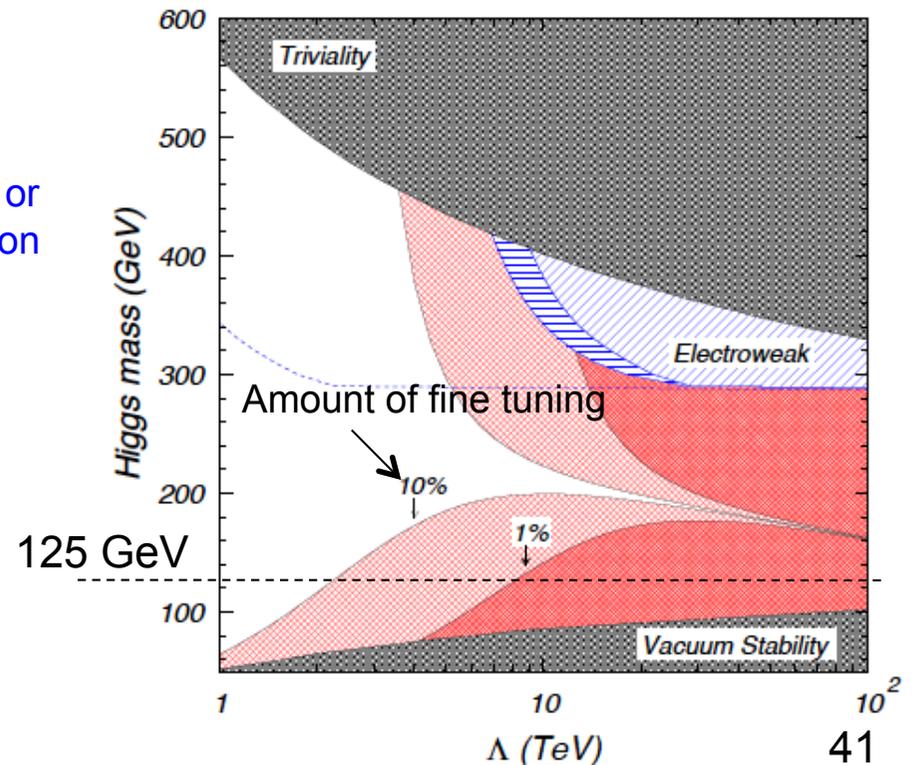
Either New Physics appears at a scale Λ or there has to be a very delicate cancellation

- May either play a key role in EWSB, or serve as a window to New Physics related to EWSB which might be preferentially coupled to it.

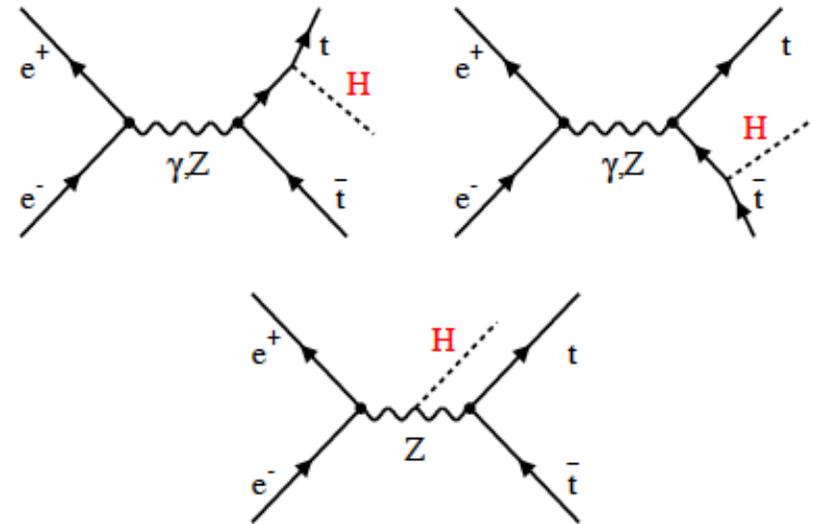
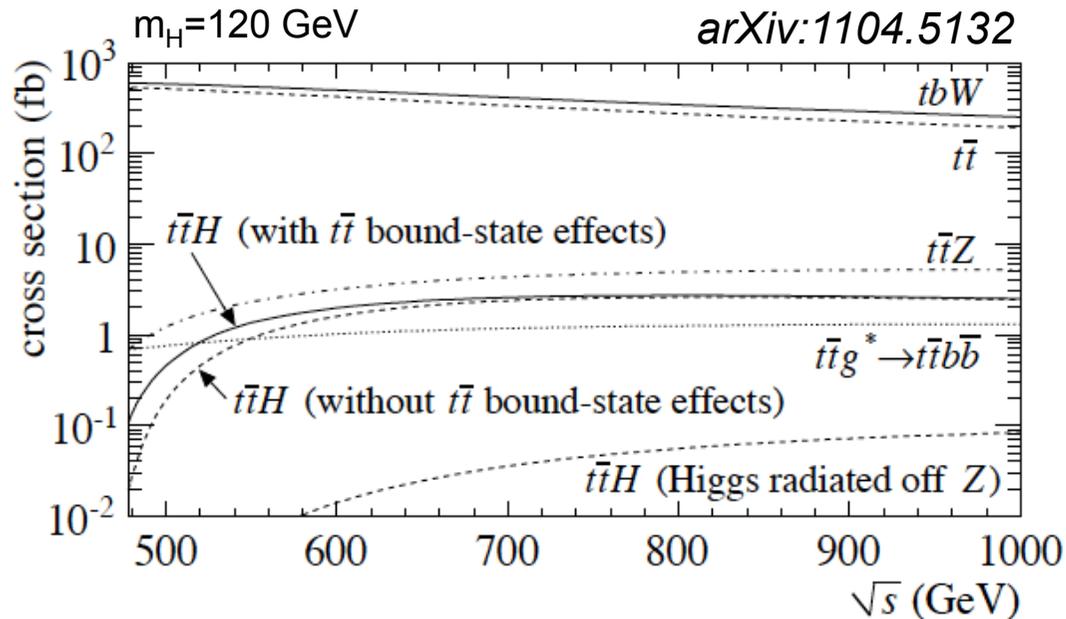
- Big incentive to measure top Yukawa coupling as precisely as possible!



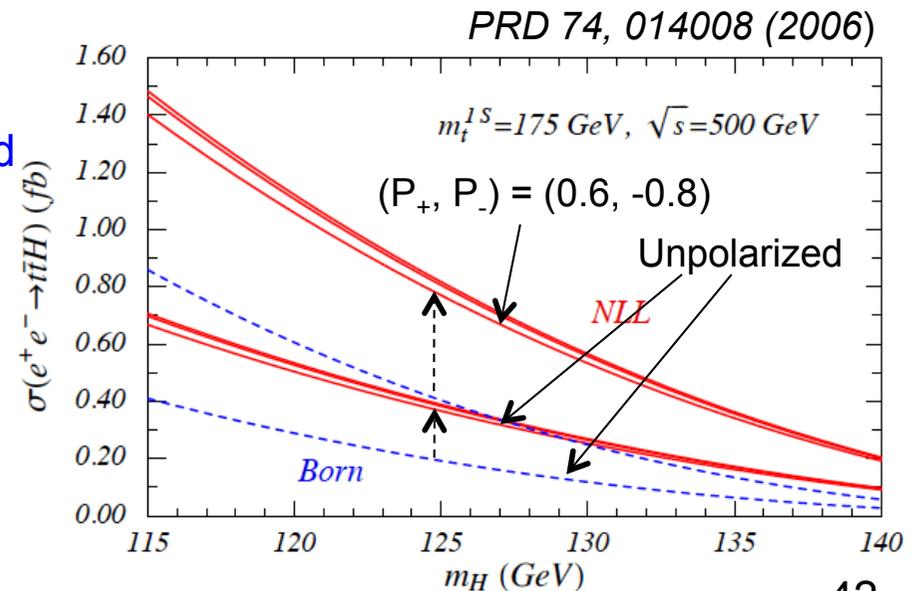
$$(125 \text{ GeV})^2 = m_{H_0}^2 + [-(2 \text{ TeV})^2 + (700 \text{ GeV})^2 + (500 \text{ GeV})^2] \left(\frac{\Lambda}{10 \text{ TeV}} \right)^2$$



ttH Production in e⁺e⁻ Collisions

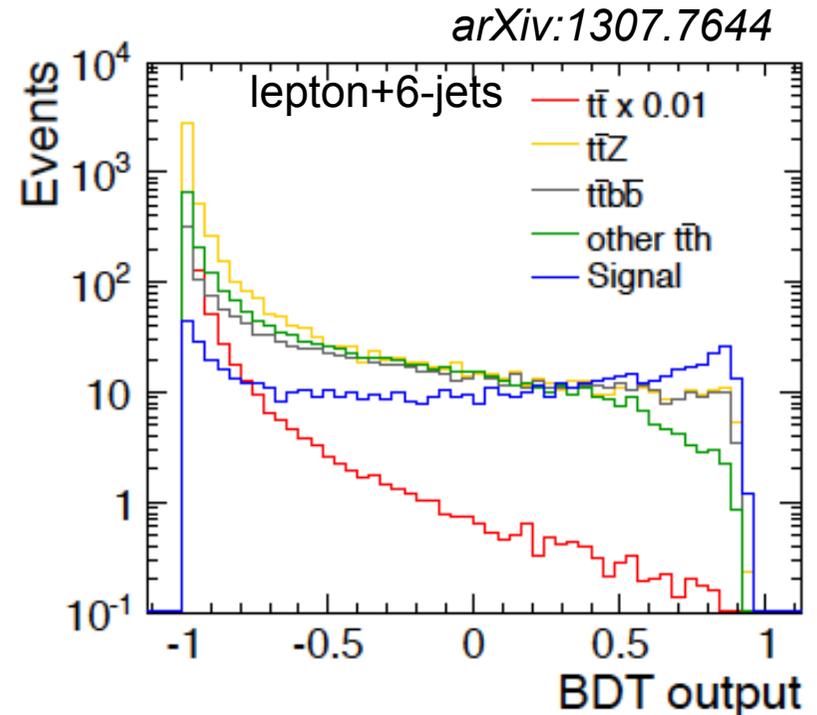


- The optimal \sqrt{s} to extract the top-Higgs Yukawa coupling at an e^+e^- collider is $\sim 800 \text{ GeV}$.
- At $\sqrt{s}=500 \text{ GeV}$, barely enough phase-space and $\sigma(t\bar{t}H)$ significantly reduced by radiative effects in initial state (ISR, beamstrahlung).
- Fortunately, there are a couple of x2 gains possible
 - $t\bar{t}$ bound-state effects near threshold
 - beam polarization
- Still, challenging: $\sigma(t\bar{t}H) \leq 1 \text{ fb}$ for $M_H = 125 \text{ GeV}$.



ILC Prospects

- The precise measurement of the top Yukawa coupling has traditionally been considered as something that only the ILC could do.
- The most recent feasibility studies are finally based on full simulation and employing realistic reconstruction algorithms.
- Example: $t\bar{t}H$, $H \rightarrow b\bar{b}$
 - $\sqrt{s}=1$ TeV, and 1 fb^{-1} equally split between $(P_{e^-}, P_{e^+})=(-0.8, +0.2)$ and $(+0.8, -0.2)$
 - Consider lepton+6-jets and 8-jets channels.
 - Train BDT to separate signal from background.
 - Apply BDT cut to maximize $S/\sqrt{S+B}$.
 - Expected stat. error on top Yukawa: 4.5%
- At $\sqrt{s}=500$ GeV it becomes more challenging:
 - $\sigma(t\bar{t}H)$ down by x10, $\sigma(t\bar{t})$ up by x2.5
 - However, $t\bar{t}H$ is enhanced by x2 due to $t\bar{t}$ bound effects, and $t\bar{t}H$ and $t\bar{t}$ another x2 if using beam polarization.
 - Expected stat. error on top Yukawa: ~10%



After BDT cut arXiv:1307.7644

Final state	BDT trained to select 6 jets
$t\bar{t}H, H \rightarrow b\bar{b}$ (6 jets)	264.9
$t\bar{t}H, H \rightarrow b\bar{b}$ (8 jets)	72.6
$t\bar{t}H, H$ not $b\bar{b}$ (6 jets)	11.7
$t\bar{t}H, H$ not $b\bar{b}$ (8 jets)	4.3
$t\bar{t}H$ (4 jets)	32.8
$t\bar{t}Z$	188.4
$t\bar{t}g^* \rightarrow t\bar{t}b\bar{b}$	185.0
$t\bar{t}$	459.3

S/B~1/3

43

Discussion on ILC Prospects

- Feasibility studies at the LC have shown that a precision on the top Yukawa coupling of 10%(5%) can be achieved at $\sqrt{s}=500$ GeV (1 TeV) with 1 ab^{-1} .

However:

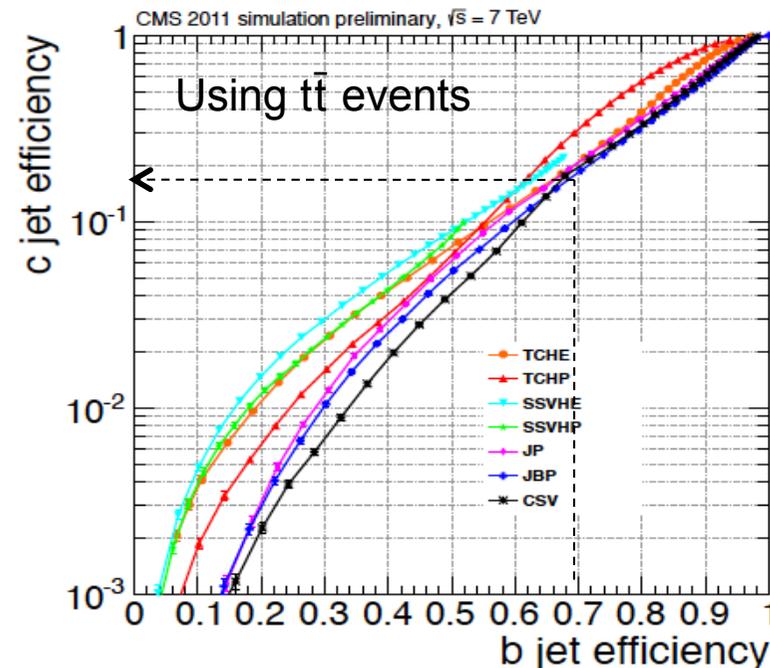
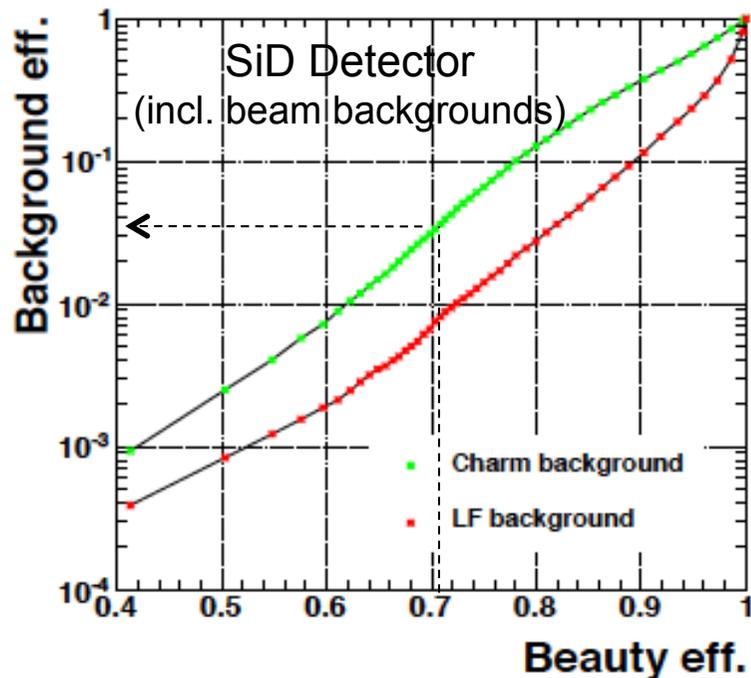
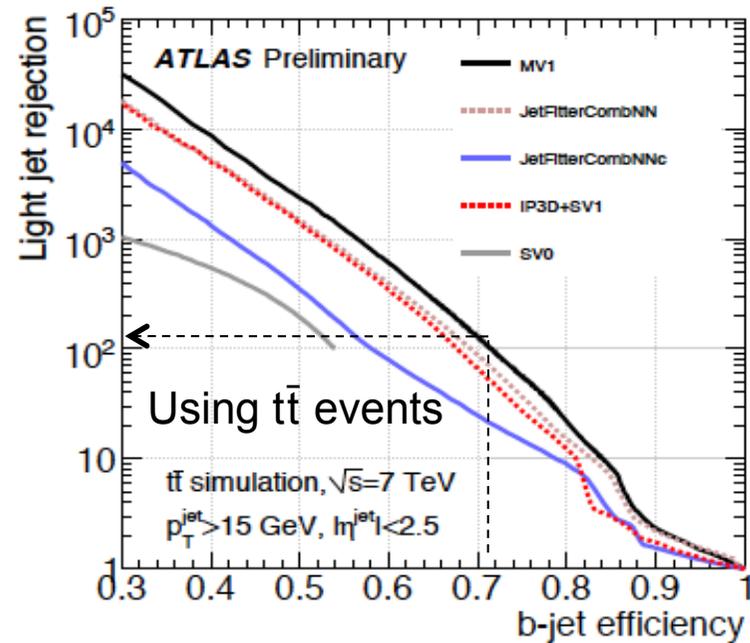
- The signal cross section is only known at NLO with an uncertainty of $\sim 10\%$. Would need to be improved.
- Studies were largely based on fast simulation and not using ME+PS MCs to predict $t\bar{t}$ +jets background
- In the best case scenario ad-hoc uncertainties on the background of 5-10% were assigned. Are those justified?

On the bright side, much of the experience and developments necessary to carry out this measurement at the LHC will be beneficial to the ILC:

- Precise theoretical predictions for signal and backgrounds via e.g. MEPS@NLO.
 - Profiling of systematic uncertainties.
- **Can the LHC measure the top-Higgs Yukawa coupling to $\sim 10\%$ or better?**
 - A 10% measurement means $\sim 5\sigma$ significance.
 - Advanced analysis techniques may resurrect $H \rightarrow b\bar{b}$ as a discovery mode.
 - New channels not considered before ($H \rightarrow \gamma\gamma$, multileptons), have irrupted into the scene with surprisingly good sensitivity and have great potential.

B-Jet Identification

- Using multivariate techniques combining information from:
 - Lifetime: displaced tracks and/or vertices
 - Mass: secondary vertex mass
 - Decay chain reconstruction
 and calibrated in data control samples.
- Performance at LHC: $\epsilon_b \sim 70\%$, $\epsilon_c \sim 20\%$, $\epsilon_{\text{light}} \sim 1\%$.
- Much better b-to-c discrimination at a LC.
 - Important to suppress non- $t\bar{t}b\bar{b}$ background!



$t\bar{t}$ +jets Modeling: A Long Road Ahead

Disclaimer: this is my personal view

- The Problem:**

- No good feeling for how accurate current $t\bar{t}$ +jets/HF ME+PS MCs are.
- Assigned systematic uncertainties are “ad-hoc”. No good understanding for what normalization and shape systematics should be considered and correlations among topologies. Unclear whether we are being too conservative or too aggressive.
- We’ll need a solid quantitative understanding before we can confidently establish a signal in this channel.

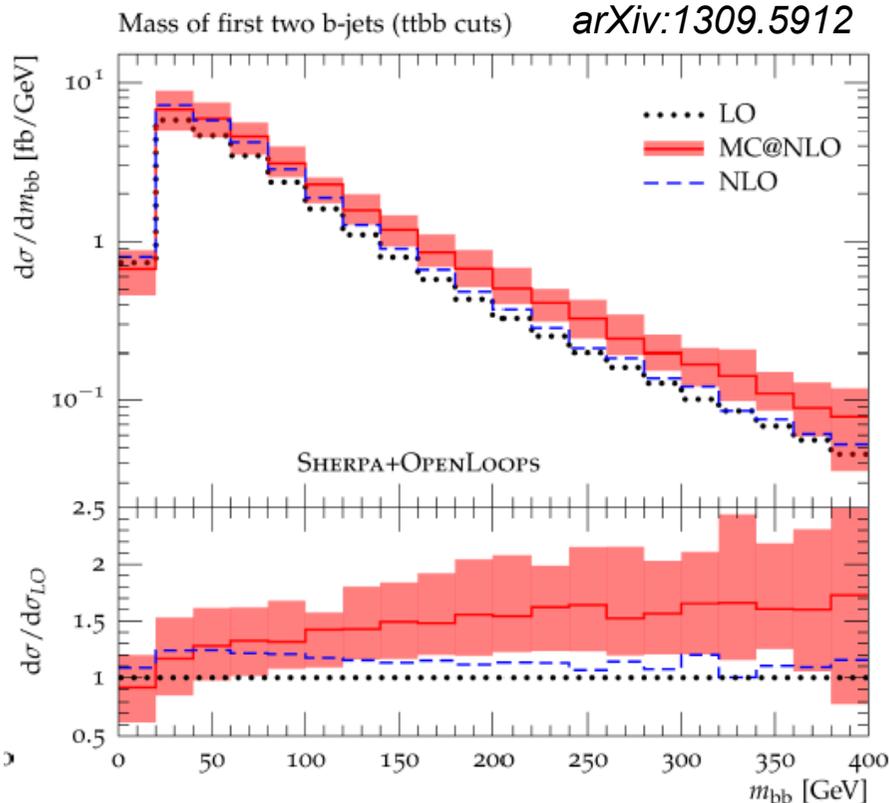
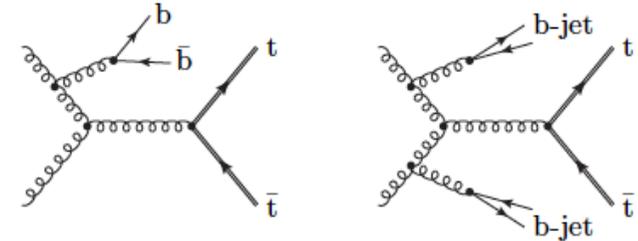
- S. Pozzorini: “What’s needed to quantify $t\bar{t}$ +jets/HF systematics in a meaningful way and reduce it to a decent (say 10-30%) level is MEPS@NLO $t\bar{t}$ +0,1,2 jets (plus extra LO MEs up to 4,5 jets).”

Current
state-of-art

	matrix elements	0j	1j	2j	3j	$\geq 4j$
LO+PS	$t\bar{t} + 2$ jets	-	-	LO	PS	PS
MEPS@LO	$t\bar{t} + 0, 1, 2, 3$ jets	LO	LO	LO	LO	PS
MC@NLO	$t\bar{t} + 2$ jets	-	-	NLO	LO	PS
MEPS@NLO	$t\bar{t} + 0, 1, 2$ jets	NLO	NLO	NLO	LO	PS

$t\bar{t}$ +jets Modeling: A Long Road Ahead

- As a first step towards MEPS@NLO, it'd be quite important to perform detailed comparisons of various LO and NLO accurate simulations (w/ PS).
- Recent progress:**
 - First complete MC@NLO simulation (within Sherpa/OpenLoops) for $t\bar{t}b\bar{b}$ at the LHC, including mass effects.
 - Allows covering the full $t\bar{t}b\bar{b}$ phase space at NLO accuracy including collinear $g \rightarrow b\bar{b}$ splitting.



Parton-level jets:
 $p_T > 25$ GeV, $|\eta| < 2.5$, Anti-kT $R=0.4$
 No hadronization. no underlvinia event

	$t\bar{t}b\bar{b}$	$t\bar{t}b\bar{b}$	$t\bar{t}b\bar{b}(m_{bb} > 100)$
σ_{LO} [fb]	$2547^{+71\%+14\%}_{-37\%-11\%}$	$463.9^{+66\%+15\%}_{-36\%-12\%}$	$123.7^{+62\%+17\%}_{-35\%-13\%}$
σ_{NLO} [fb]	$3192^{+33\%+4.6\%}_{-25\%-4.9\%}$	$557^{+28\%+5.6\%}_{-24\%-4.0\%}$	$141^{+25\%+8.6\%}_{-22\%-3.8\%}$
σ_{NLO}/σ_{LO}	1.25	1.20	1.14
σ_{MC} [fb]	$3223^{+33\%+4.3\%}_{-25\%-2.5\%}$	$607^{+25\%+2.2\%}_{-22\%-2.8\%}$	$186^{+21\%+5.4\%}_{-20\%-4.7\%}$
σ_{MC}/σ_{NLO}	1.01	1.09	1.32
σ_{MC}^{2b} [fb]	3176	539	145
$\sigma_{MC}^{2b}/\sigma_{NLO}$	0.99	0.97	1.03

Significant contribution from double collinear $g \rightarrow b\bar{b}$ splitting at high m_{bb} (one of them from the parton shower)

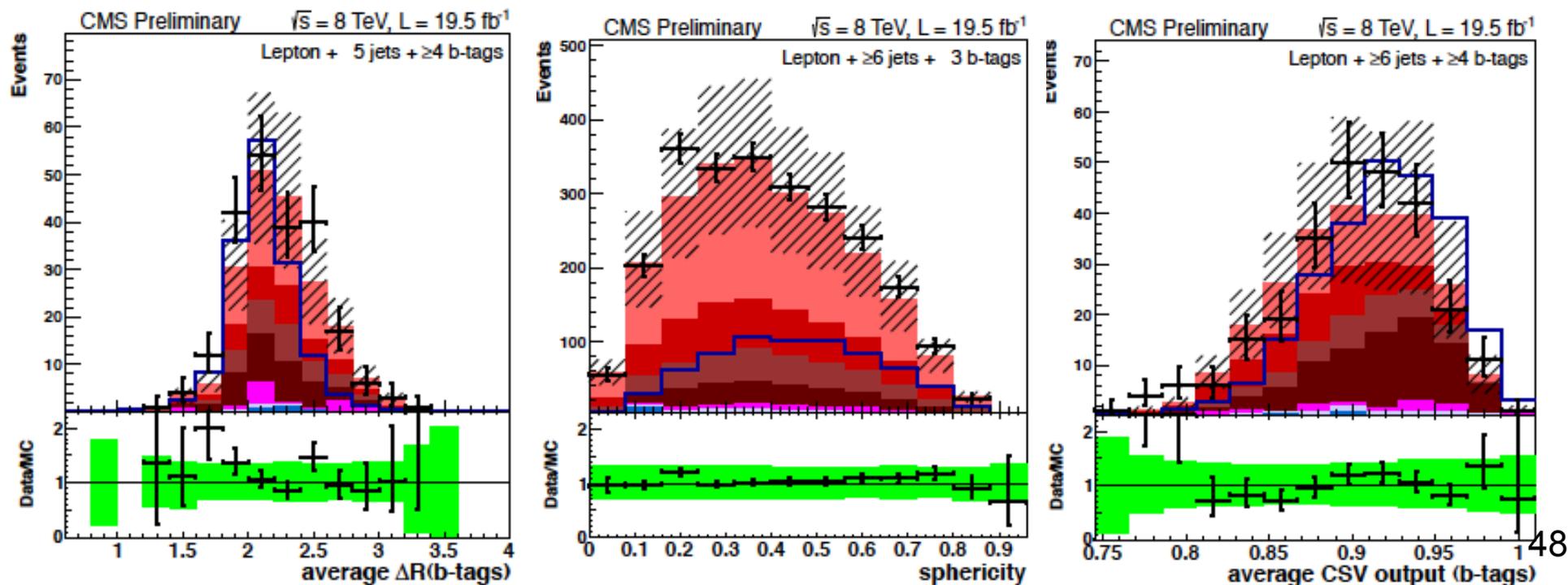
Signal-to-Background Discrimination



- Boosted Decision Trees (BDTs) trained for each category of the analysis.
- A total of 10 / 4-6 / 10 variables are used in the lepton+jets / dilepton / tau channels.
 - Angular correlations: e.g. average $\Delta R(b,b)$
 - Event kinematics: e.g. sphericity
 - B-tagging information: e.g. average b-tagging output variable
 - Tau isolation and kinematics (tau channel)
 - $t\bar{t}b\bar{b}/t\bar{t}H$ BDT in signal-rich lepton+jets channels



CMS PAS HIG-13-019



Systematic Uncertainties



% change in background yield in ≥ 6 jets/ ≥ 4 tags

Uncertainties of the sum of $t\bar{t}+lf$, $t\bar{t}+b$, $t\bar{t} + b\bar{b}$, and $t\bar{t} + c\bar{c}$ events with ≥ 6 jets and ≥ 4 b-tags		
Source	Rate	Shape?
QCD Scale (all $t\bar{t}+hf$)	35%	No
QCD Scale ($t\bar{t} + b\bar{b}$)	17%	No
b-Tag bottom-flavor contamination	17%	Yes
QCD Scale ($t\bar{t} + c\bar{c}$)	11%	No
Jet Energy Scale	11%	Yes
b-Tag light-flavor contamination	9.6%	Yes
b-Tag bottom-flavor statistics (linear)	9.1%	Yes
QCD Scale ($t\bar{t}+b$)	7.1%	No
Madgraph Q^2 Scale ($t\bar{t} + b\bar{b}$)	6.8%	Yes
b-Tag Charm uncertainty (quadratic)	6.7%	Yes
Top p_T Correction	6.7%	Yes
b-Tag bottom-flavor statistics (quadratic)	6.4%	Yes
b-Tag light-flavor statistics (linear)	6.4%	Yes
Madgraph Q^2 Scale ($t\bar{t} + 2$ partons)	4.8%	Yes
b-Tag light-flavor statistics (quadratic)	4.8%	Yes
Luminosity	4.4%	No
Madgraph Q^2 Scale ($t\bar{t} + c\bar{c}$)	4.3%	Yes
Madgraph Q^2 Scale ($t\bar{t}+b$)	2.6%	Yes
QCD Scale ($t\bar{t}$)	3%	No
pdf (gg)	2.6%	No
Jet Energy Resolution	1.5%	No
Lepton ID/Trigger efficiency	1.4%	No
Pileup	1%	No
b-Tag Charm uncertainty (linear)	0.6%	Yes

CMS PAS HIG-13-019

Assume 50% uncertainty on $t\bar{t}b\bar{b}$, $t\bar{t}b$ and $t\bar{t}c\bar{c}$
(uncorrelated among them)
AD-HOC!!

- Many systematic uncertainties, both theoretical and experimental.
- Background systematics much larger than expected signal yield!

Total background uncertainty: ~37%
Expected S/B: ~3.3%

CMS $H \rightarrow bb/\tau\tau$ Results (7+8 TeV)



CMS PAS HIG-13-019

- Combination of lepton+jets, dileptons and tau:
Observed (expected) limit @ $M_H=125$ GeV:
5.2xSM (4.1xSM)
- Addition of dilepton and tau channels improves sensitivity by $\sim 15\%$.
- Analysis not scaling as $1/\sqrt{L}$ (dominated by $t\bar{t}$ +HF syst., revisited since the previous publication result; *arXiv:1303.0763*):
Exp. Limit (7 TeV, 5.0 fb $^{-1}$): 6.9xSM
Exp. Limit (8 TeV, 5.1 fb $^{-1}$): 5.7xSM

ttH decay mode	Exp	Ob
$H \rightarrow bb, tt \rightarrow$ lepton+jets	4.7	4.9
$H \rightarrow bb, tt \rightarrow$ dilepton	8.2	9.1
$H \rightarrow$ tautau, $tt \rightarrow$ ljets or dilepton	14.2	13.2
Combination	4.1	5.2

