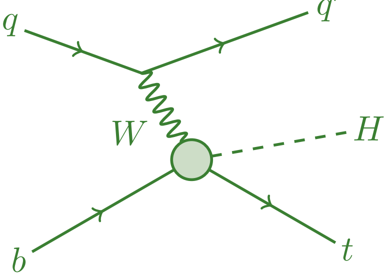
Searches for associated *tH* production with the CMS experiment



Andrey Popov[1,2]
On behalf of the CMS collaboration

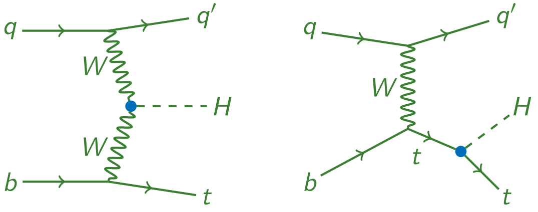
[1]CP3 UCL, Louvain-la-Neuve, BE

[2]also at SINP MSU, Moscow, RU

II CMS single-top workshop,
Naples, Italy,
4–5 Dec 2014

Associated *tH* production

- Two dominant diagrams for *tHq* production in SM:

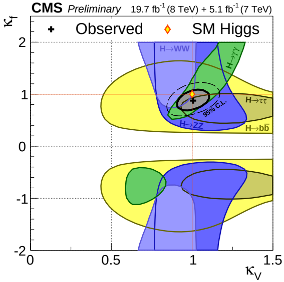


  - Destructive interference in SM ⇒ cross section 18.3 fb†
  - With an inverted sign of Yukawa coupling $y_t = -1$ the interference is constructive, $\sigma = 234\,\text{fb}^\dagger$ ($\times 13$ enhancement)

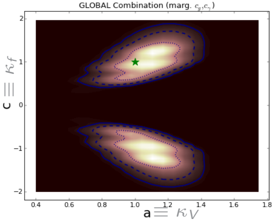
† Farina et al., JHEP **05** (2013) 022

Negative top-quark Yukawa coupling

- Couplings measurements by ATLAS and CMS favour $y_t \approx +1$
  - Sensitivity to the sign comes from $H \to \gamma\gamma$ only
  - Results assume that no new physics affect Hgg and $H\gamma\gamma$ loop-induced couplings
- If BSM contributions to Hgg and $H\gamma\gamma$ are considered, $y_t \approx -1$ is well allowed
  - Bottom fig.: Combined constraints from ATLAS, CMS, and Tevatron marginalised over possible BSM contributions in the loops
- Study of tHq production allows to solve the twofold ambiguity



CMS PAS HIG-14-009



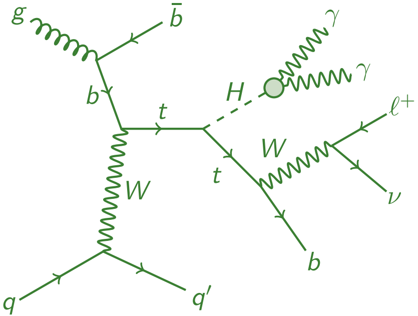
JHEP **06** (2013) 103

Searches for tHq production with negative y_t

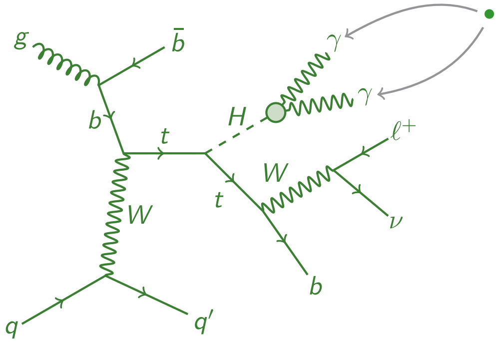
- CMS performed searches for tHq production with negative y_t
  - $H \to \gamma\gamma$ and $H \to b\bar{b}$ decay channels
  - Focused on the $y_t = -1$ case as couplings affect kinematics
  - Whole 8 TeV dataset utilised ($\sim 20\,\text{fb}^{-1}$)
- $H \to \gamma\gamma$ channel (CMS PAS HIG-14-001)
  - Small branching ratio but high purity
  - An additional enhancement of $\mathcal{B}(H \to \gamma\gamma)$ by a factor of 2.4
  - Cut-and-count analysis
- $H \to b\bar{b}$ channel (CMS PAS HIG-14-015)
  - Largest branching ratio but overwhelming $t\bar{t}$ background
  - Complex MVA-based analysis

Search in the $H \rightarrow \gamma\gamma$ channel

Event preselection

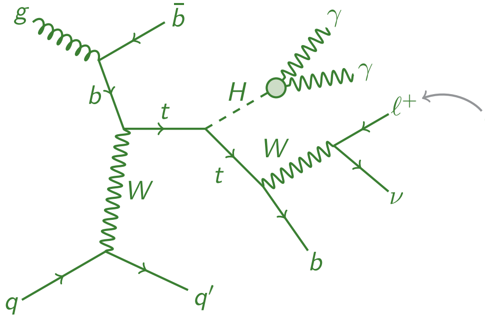


Event preselection



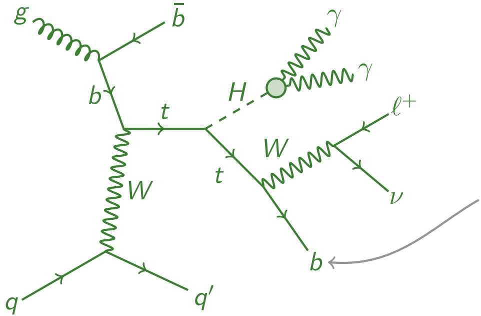
- Two photons
  - $p_T(\gamma_1) > 50 \cdot m_{\gamma\gamma}/120$
  - $p_T(\gamma_2) > 25\,\text{GeV}$
  - Eff. on signal events is 98%

Event preselection



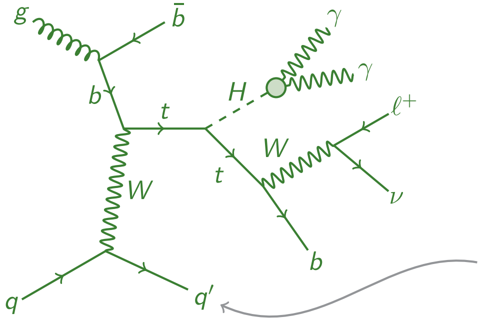
- Two photons
  - $p_T(\gamma_1) > 50 \cdot m_{\gamma\gamma}/120$
  - $p_T(\gamma_2) > 25\,\text{GeV}$
  - Eff. on signal events is 98%
- Exactly one isolated μ or e
  - $p_T > 10\,\text{GeV}$

Event preselection



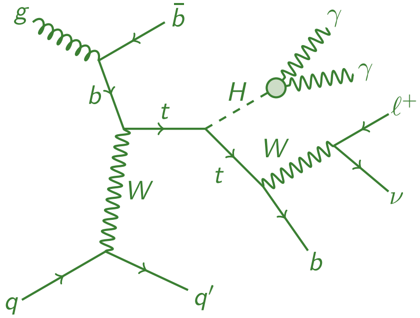
- Two photons
  - $p_T(\gamma_1) > 50 \cdot m_{\gamma\gamma}/120$
  - $p_T(\gamma_2) > 25\,\text{GeV}$
  - Eff. on signal events is 98%
- Exactly one isolated μ or e
  - $p_T > 10\,\text{GeV}$
- A b-quark jet
  - $p_T > 20\,\text{GeV}$

Event preselection



- Two photons
  - $p_T(\gamma_1) > 50 \cdot m_{\gamma\gamma}/120$
  - $p_T(\gamma_2) > 25\,\text{GeV}$
  - Eff. on signal events is 98%
- Exactly one isolated μ or e
  - $p_T > 10\,\text{GeV}$
- A b-quark jet
  - $p_T > 20\,\text{GeV}$
- A forward recoil jet
  - $p_T > 20\,\text{GeV}$, $|\eta| > 1$

Event preselection



- Two photons
  - $p_T(\gamma_1) > 50 \cdot m_{\gamma\gamma}/120$
  - $p_T(\gamma_2) > 25\,\text{GeV}$
  - Eff. on signal events is 98%
- Exactly one isolated μ or e
  - $p_T > 10\,\text{GeV}$
- A b-quark jet
  - $p_T > 20\,\text{GeV}$
- A forward recoil jet
  - $p_T > 20\,\text{GeV}$, $|\eta| > 1$

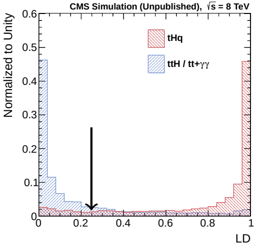
- Consider mass window $122 < m_{\gamma\gamma} < 128\,\text{GeV}$ as the signal region

Resonant backgrounds

- Backgrounds with a Higgs boson contribute to the $m_{\gamma\gamma}$ peak
  - $t\bar{t}H$ (dominant), VH, $H+$jets
- To suppress $t\bar{t}H$, a likelihood product discriminator (LD) is constructed
  - Variables: # jets, $m_T(t)$, $\eta(q')$, $\Delta\eta(\ell, q')$, lepton charge
- A cut on value of LD is added to event selection
- Expected yields in the $m_{\gamma\gamma}$ window after the full selection:

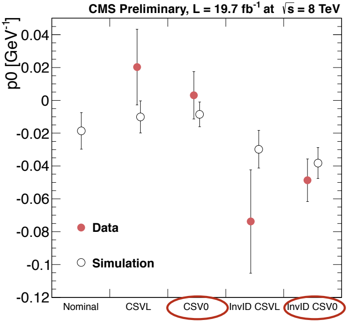
Process	Yield
tHq, $y_t = -1$	0.67
$t\bar{t}H$	$0.03 + 0.05^{\dagger}$
VH	$0.01 + 0.01^{\dagger}$
other H	0

†Increase due to $y_t = -1$, included into signal



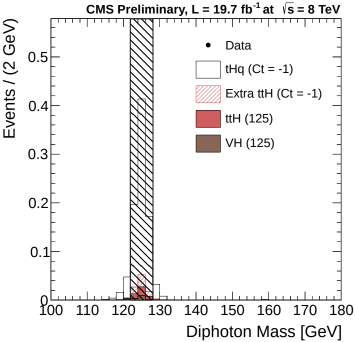
Non-resonant backgrounds

- Remaining backgrounds are smooth in $m_{\gamma\gamma}$
  - $\gamma\gamma + \text{jets}$, $\gamma + \text{jets}$, $t\gamma\gamma$, $t\bar{t}\gamma\gamma$, ...
- Their spectrum in data is fitted with an exponential function
  - Utilise sidebands $m_{\gamma\gamma} \in (100, 122) \cup (128, 180)$ GeV
  - Use four control regions with
    - loosened ("CSVL") or removed ("CSV0") requirement on b-tagging
    - nominal or inverted photon ID
  - Difference between two high-stat control regions used as systematics



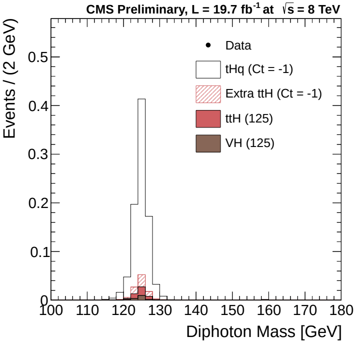
Results

- Zero events observed in $m_{\gamma\gamma}$ sidebands
  - Translates to an estimate of contribution of non-resonant backgrounds under the peak



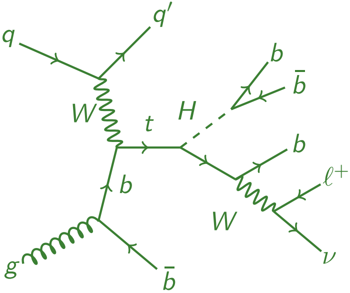
Results

- Zero events observed in $m_{\gamma\gamma}$ sidebands
  - Translates to an estimate of contribution of non-resonant backgrounds under the peak
- Zero events found in the mass window as well
- Observed 95% CL_s upper limit on tHq is 4.1 times the expected cross section
  - Absolute value: $\sigma_{tHq}^{y_t=-1} \times \mathcal{B}_{H\to\gamma\gamma}^{y_t=-1} < 5.2\,\text{fb}$
  - Coincides with the expected limit

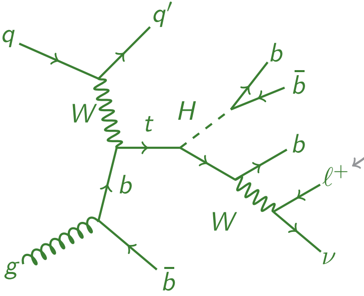


Search in the $H \to b\bar{b}$ channel

Event selection

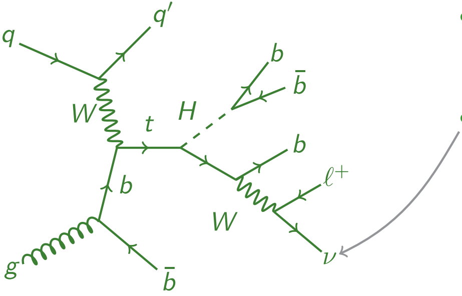


Event selection



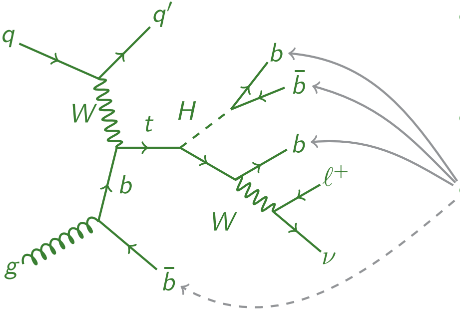
- Exactly one isolated μ or e
  - $p_{\mathrm{T}} > 26\ (30)$ GeV for μ (e)
  - Used to trigger events

Event selection



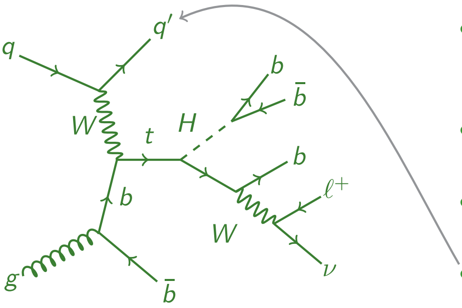
- Exactly one isolated μ or e
  - $p_T > 26$ (30) GeV for μ (e)
  - Used to trigger events
- Moderate $\not{E}_T$
  - $\not{E}_T > 35$ (45) for μ (e)

Event selection



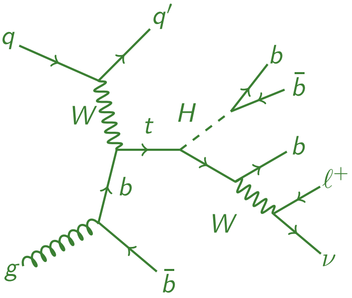
- Exactly one isolated μ or e
  - $p_{\mathrm{T}} > 26$ (30) GeV for μ (e)
  - Used to trigger events
- Moderate $\not{E}_{\mathrm{T}}$
  - $\not{E}_{\mathrm{T}} > 35$ (45) for μ (e)
- Three or four b-quark jets
  - $p_T > 20$ GeV

Event selection



- Exactly one isolated μ or e
  - $p_{\mathrm{T}} > 26$ (30) GeV for μ (e)
  - Used to trigger events
- Moderate $\not{E}_{\mathrm{T}}$
  - $\not{E}_{\mathrm{T}} > 35$ (45) for μ (e)
- Three or four b-quark jets
  - $p_T > 20\,\mathrm{GeV}$
- A non-b-quark recoil jet
  - $p_T > 20\,\mathrm{GeV}$ if $|\eta| < 2.4$
  - $p_T > 40\,\mathrm{GeV}$ if $|\eta| > 2.4$

Event selection

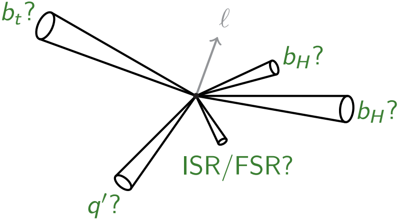


Region	S/B ratio	
3 b-jets	13/1900	0.7%
4 b-jets	1.4/66	2.1%

- Exactly one isolated μ or e
  - $p_{\mathrm{T}} > 26$ (30) GeV for μ (e)
  - Used to trigger events
- Moderate $\not{E}_{\mathrm{T}}$
  - $\not{E}_{\mathrm{T}} > 35$ (45) for μ (e)
- Three or four b-quark jets
  - $p_T > 20\,\mathrm{GeV}$
- A non-b-quark recoil jet
  - $p_T > 20\,\mathrm{GeV}$ if $|\eta| < 2.4$
  - $p_T > 40\,\mathrm{GeV}$ if $|\eta| > 2.4$
- Dominant background is $t\bar{t}$

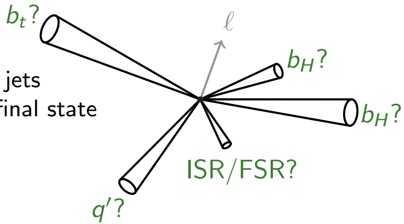
Jet assignment with MVA

- The small S/B ratio makes a multivariate analysis essential
- But construction of input variables from a multijet final state is delicate
  - E. g. which jets stem from H decay?



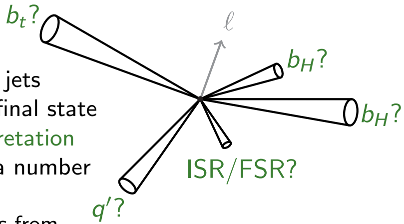
Jet assignment with MVA

- The small S/B ratio makes a multivariate analysis essential
- But construction of input variables from a multijet final state is delicate
  - E. g. which jets stem from H decay?
- Hypothesise origin of RECO jets
  - Consider all possible ways to assign four jets to the four quarks in the $tHq \to \ell\nu 3bq$ final state



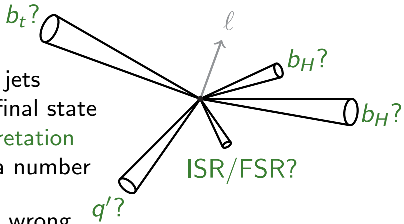
Jet assignment with MVA

- The small S/B ratio makes a multivariate analysis essential
- But construction of input variables from a multijet final state is delicate
  - E. g. which jets stem from H decay?
- Hypothesise origin of RECO jets
  - Consider all possible ways to assign four jets to the four quarks in the $tHq \to \ell\nu 3bq$ final state
  - Each particular way represents an interpretation of the event and can be described with a number of observables
    - Reconstructed m_H, m_t, ΔR between jets from $H \to b\bar{b}$, b-tagging discriminator of the b from $t \to b\ell\nu$, etc.



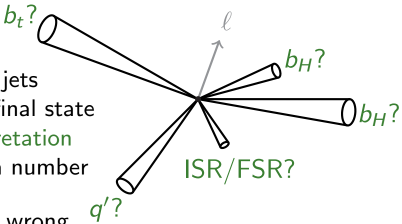
Jet assignment with MVA

- The small S/B ratio makes a multivariate analysis essential
- But construction of input variables from a multijet final state is delicate
  - E. g. which jets stem from H decay?
- Hypothesise origin of RECO jets
  - Consider all possible ways to assign four jets to the four quarks in the $tHq \to \ell\nu 3bq$ final state
  - Each particular way represents an interpretation of the event and can be described with a number of observables
  - Train an MVA to distinguish correct and wrong event interpretations



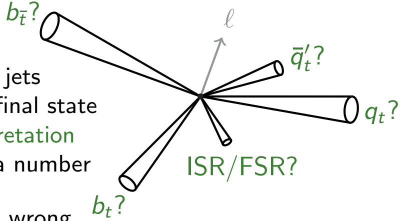
Jet assignment with MVA

- The small S/B ratio makes a multivariate analysis essential
- But construction of input variables from a multijet final state is delicate
  - E. g. which jets stem from H decay?
- Hypothesise origin of RECO jets
  - Consider all possible ways to assign four jets to the four quarks in the $tHq \to \ell\nu 3bq$ final state
  - Each particular way represents an interpretation of the event and can be described with a number of observables
  - Train an MVA to distinguish correct and wrong event interpretations
  - In an unknown event, consider all possible interpretations, evaluate the MVA for each one, and accept the interpretation with the largest MVA response

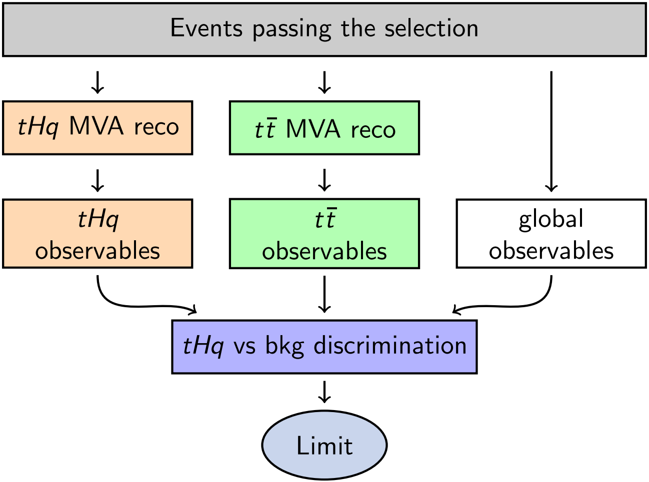


Jet assignment with MVA

- The small S/B ratio makes a multivariate analysis essential
- But construction of input variables from a multijet final state is delicate
  - E. g. which jets stem from H decay?
- Hypothesise origin of RECO jets
  - Consider all possible ways to assign four jets to the four quarks in the $tHq \to \ell\nu 3bq$ final state
  - Each particular way represents an interpretation of the event and can be described with a number of observables
  - Train an MVA to distinguish correct and wrong event interpretations
  - In an unknown event, consider all possible interpretations, evaluate the MVA for each one, and accept the interpretation with the largest MVA response
- A similar procedure is used for semileptonic $t\bar{t}$
  - This is by far the dominant background

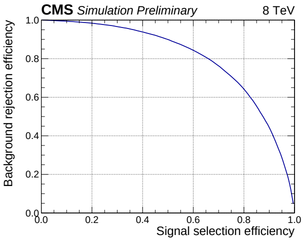


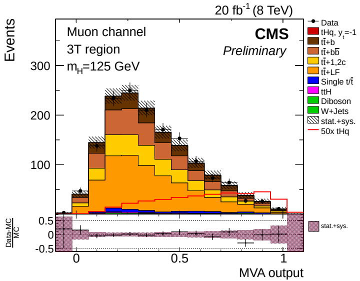
Discrimination between signal and backgrounds



Discrimination between signal and backgrounds

- The jet assignment is only used to construct input variables for the final MVA
  - Trained to distinguish tH events from backgrounds
- Examples of input variables
  - Defined under tH hypothesis: $p_{\mathrm{T}}(H)$, $\eta(q')$
  - Defined under $t\bar{t}$ hypothesis: mass of $t \to$ had, number of b-tagged jets amoung its decay products
  - Independent of jet assignment: $Q(\ell)$
- Response of the final MVA in data is fitted to put an upper limit
- Observed a 95% CL_s upper limit of 7.6 times the expected cross section (exp. limit $5.1^{+2.1}_{-1.7}$)
  - Absolute value: $\sigma_{tHq}^{y_t=-1} \times \mathcal{B}_{H\to b\bar{b}} < 1.0\,\text{pb}$





Summary

- Associated tH production allows to probe for negative y_t
  - A possibility still supported by experimental data if one considers BSM contributions to Hgg and $H\gamma\gamma$ loops
- Searches for tHq with $y_t = -1$ have been performed in the $H \to \gamma\gamma$ and $H \to b\bar{b}$ decay channels
  - Upper limits of 4.1 and $7.6 \times \sigma_{tHq}^{y_t=-1}$ are observed for $H \to \gamma\gamma$ and $H \to b\bar{b}$
- New results from 8 TeV data are coming, including a combination of the searches
  - Stay tuned
- Looking forward LHC Run II
  - Fourfold increase in signal cross section is expected

Additional slides

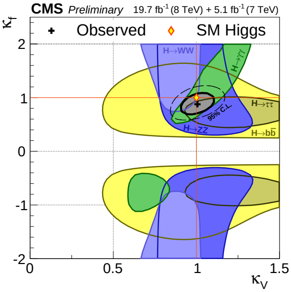
Cross sections

- Cross section is challengingly small
  - The main background is $t\bar{t}$; its cross section is provided for comparison

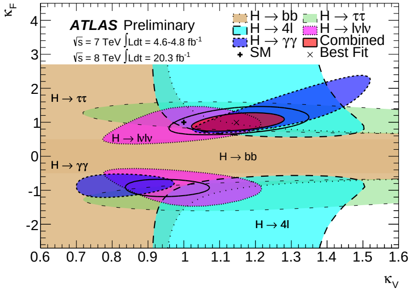
Cross section	8 TeV	14 TeV
tHq, $y_t = +1$ (SM)	18.3 ± 0.4 fb	$88.2^{+1.7}_{-0.0}$ fb
tHq, $y_t = -1$	$233.8^{+4.6}_{-0.0}$ fb	980^{+30}_{-0} fb
$t\bar{t}$	245^{+9}_{-10} pb	950^{+40}_{-30} pb

tHq cross sections are cited according to M. Farina et al., JHEP 1305 (2013) 022 [arXiv:1211.3736]. Cross sections for $t\bar{t}$ are calculated in M. Czakon, P. Fiedler, Phys. Rev. Lett. 110 (2013) 252004 [arXiv:1303.6254]. Uncertainties are combined following R. Barlow, arXiv:physics/0306138

Constraints on Higgs boson couplings from LHC



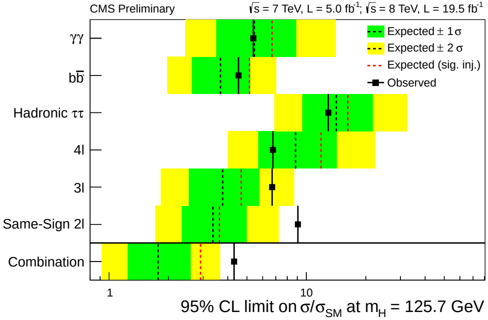
CMS PAS HIG-14-009

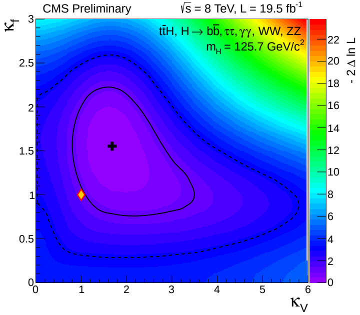


ATLAS CONF-2014-009

Direct measurements of y_t at LHC

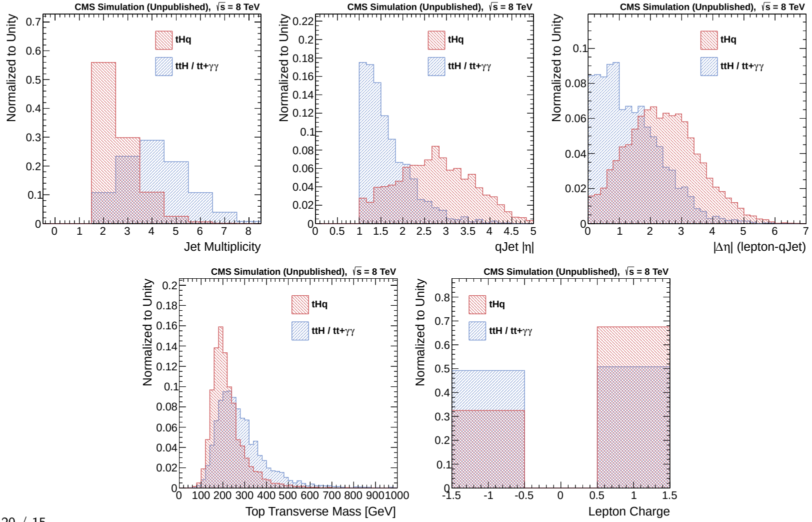
- A number of searches for $t\bar{t}H$ production performed by CMS and ATLAS
- So far, no significant excess over the background prediction has been observed, thus only upper limits on the cross section are set





CMS $t\bar{t}H$ combination based on HIG-12-025, HIG-13-015, HIG-13-019, HIG-13-020

$H \to \gamma\gamma$: Input variables for the LD



$H \to \gamma\gamma$: Systematic uncertainties

	tHq	$t\bar{t}H$	VH	Continuous BG
Luminosity	$\pm 2.6\%$	$\pm 2.6\%$	$\pm 2.6\%$	—
PDF	+3.1/-2.5 %	$\pm 8\%$	$\pm 11\%$	—
QCD scale	+4.8/-4.3 %	+11/-14 %	$\pm 2.3\%$	—
Signal model	$\pm 5.5\%$	—	—	—
Photon energy resolution	+4/-2 %	+4/-2 %	+4/-2 %	—
Photon energy scale	+1/-4 %	+1/-4 %	+1/-4 %	—
Photon ID efficiency	$\pm 2\%$	$\pm 2\%$	$\pm 2\%$	—
Vertex efficiency	$\pm 0.1\%$	$\pm 0.1\%$	$\pm 0.1\%$	—
Trigger	$< 0.1\%$	$< 0.1\%$	$< 0.1\%$	—
JEC	$\pm 1.5\%$	+3/-5 %	$\pm 8\%$	—
JER	$\pm 0.5\%$	$\pm 3\%$	+8/-0 %	—
b-tagging	$\pm 2\%$	$\pm 1.5\%$	$\pm 0.1\%$	—
PU ID	$\pm 2\%$	$\pm 0.5\%$	$\pm 2\%$	—
Lepton reconstruction	$\pm 1\%$	$\pm 1\%$	$\pm 1\%$	—
BG shape	—	—	—	33%

$H \to b\bar{b}$: Expected event yields

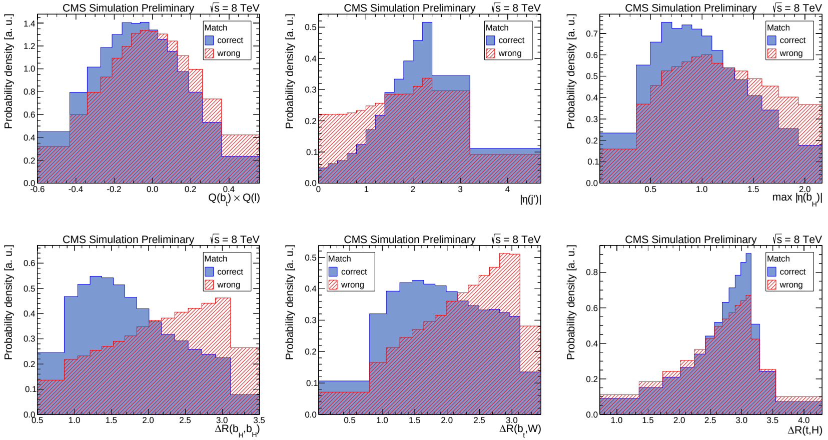
3t region

Process	Muon channel	Electron channel
$t\bar{t}$	1058±5	718±4
Single top	39±3	27±3
Electroweak	17^{+7}_{-5}	11±7
$t\bar{t}H$	12.87±0.17	9.35±0.15
Total background	1128±9	767±10
tHq, $y_t = -1$	7.54±0.03	5.15±0.02
S/B ratio	0.7%	0.7%

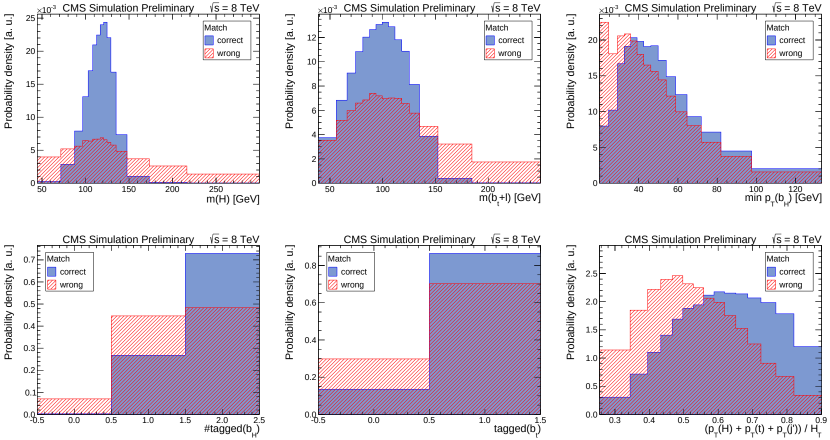
4t region

Process	Muon channel	Electron channel
$t\bar{t}$	29.1±0.8	19.8±0.7
Single top	$1.1^{+0.8}_{-0.6}$	1.2±1.0
Electroweak	4^{+6}_{-4}	5^{+6}_{-4}
$t\bar{t}H$	1.72±0.06	1.43±0.05
Total background	37^{+6}_{-4}	29^{+7}_{-4}
tHq, $y_t = -1$	0.835±0.010	0.580±0.009
S/B ratio	2.3%	2.0%

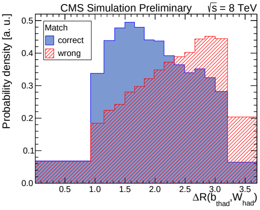
$H \to b\bar{b}$: Input variables for tHq jet assignment

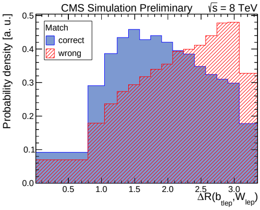


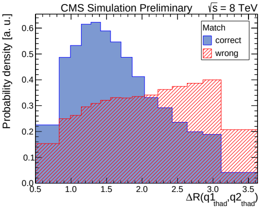
$H \to b\bar{b}$: Input variables for tHq jet assignment

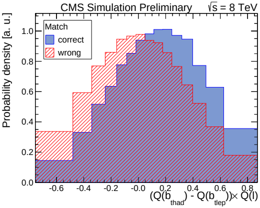


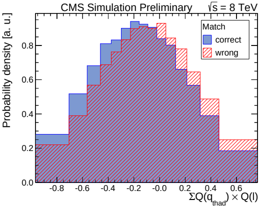
$H \to b\bar{b}$: Input variables for $t\bar{t}$ jet assignment



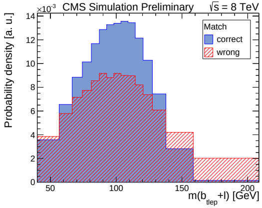


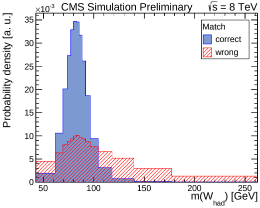


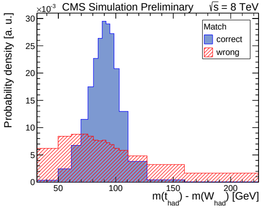


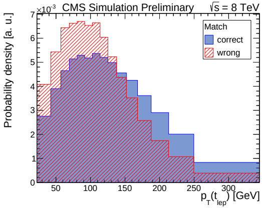


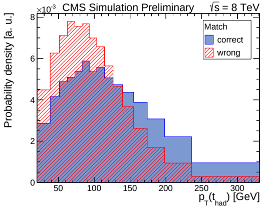
$H \to b\bar{b}$: Input variables for $t\bar{t}$ jet assignment



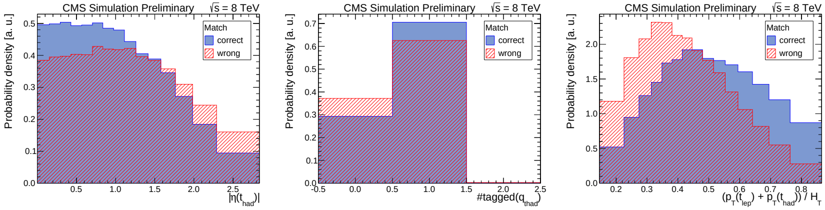








$H \to b\bar{b}$: Input variables for $t\bar{t}$ jet assignment



$H \to b\bar{b}$: Efficiency of jet identification

- Three definitions of identification efficiency are considered:
  - "A": Calculated with all events that pass the selection
  - "B": Use only events in which the parton in question does have an MC-truth match
  - "C": Use events that have a correct interpretation (as defined by the procedure, see the AN)
- A perfect jet assignment would have a 100% efficiency in definition C
  - At the same, time the efficiency in definition A incorporates effects of jet acceptance (especially, p_T cut) as well as jet splitting, merging, and other artefacts and thus might be significantly lower
- Definitions A and C can also be used for groups of more than one jet
  - A group is identified correcly if all jets in it are identified

$H \to b\bar{b}$: Efficiency of jet identification, *tHq* hyp.

- Mean number of jets in a *tHq* event is 5.4
- Mean number of considered event interpretations is 60
- Efficiency of identification of jets and groups of jets:

Object(s)	**Efficiency**		
	A	**B**	**C**
b from $t \to b\ell\nu$	56.7%	61.3%	65.9%
at least one b from $H \to b\bar{b}$	85.6%	—	92.3%
both b from $H \to b\bar{b}$	50.8%	—	64.5%
recoil quark	51.8%	78.3%	78.5%
all four quarks	21.8%	—	44.2%

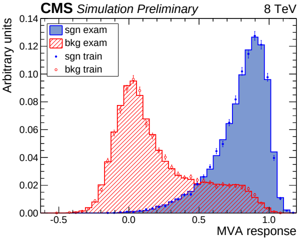
$H \to b\bar{b}$: Efficiency of jet identification, $t\bar{t}$ hypothesis

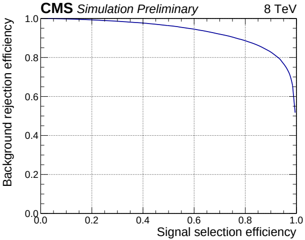
- Mean number of jets in a semileptonic $t\bar{t}$ event is 5.7
- Mean number of considered event interpretations is 35
- Efficiency of identification of jets and groups of jets:

Object(s)	**Efficiency**		
	A	**B**	**C**
b from $t \to b\ell\nu$	63.5%	69.5%	66.0%
b from $t \to$ had	58.3%	63.5%	68.3%
at least one q from $W \to q\bar{q}'$	63.0%	—	90.0%
both q from $W \to q\bar{q}'$	10.9%	—	56.6%
all quarks from $t \to$ had	8.6%	—	47.1%
all four quarks	6.1%	—	37.0%

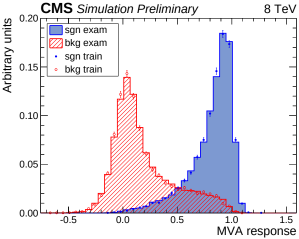
$H \to b\bar{b}$: Responses and ROCs for jet assignment

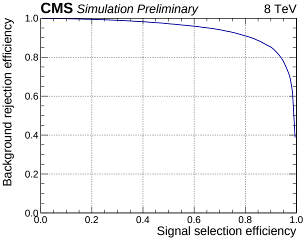
Reco of tHq events





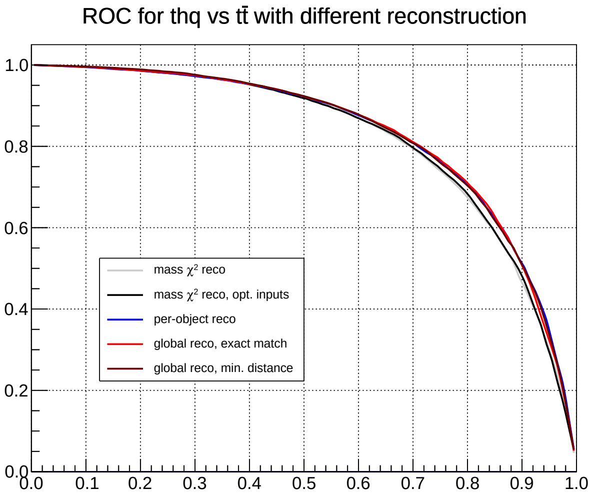
Reco of $t\bar{t}$ events



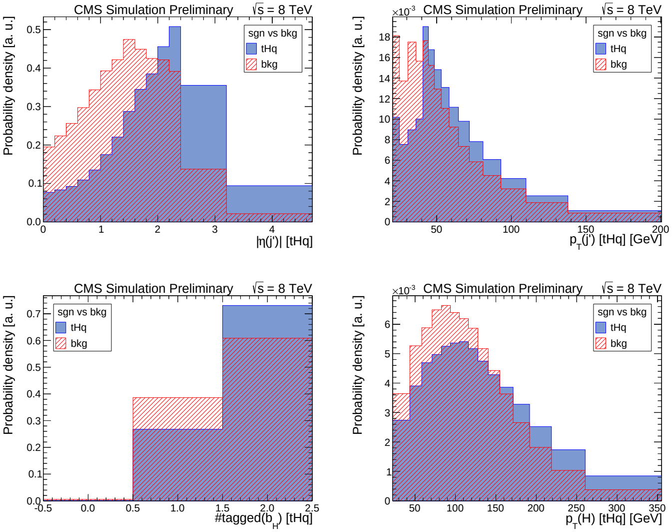


Sum of μ and e channels

$H \to b\bar{b}$: Jet assignment and perf. of class. MVA

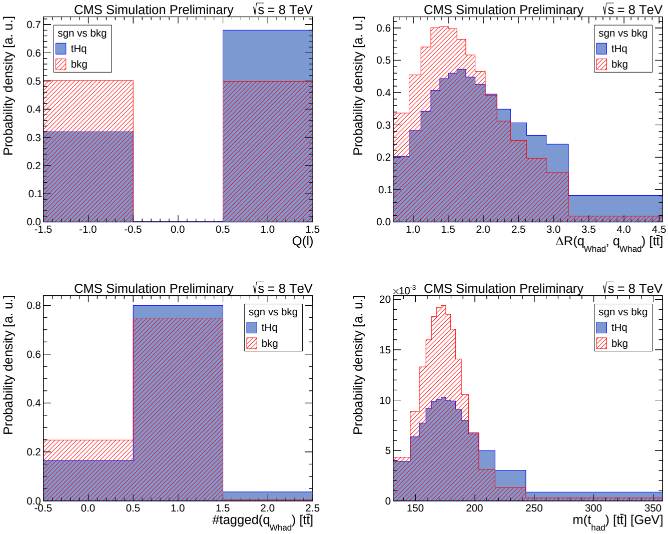


$H \rightarrow b\bar{b}$: Input variables for tHq vs bkgs classification



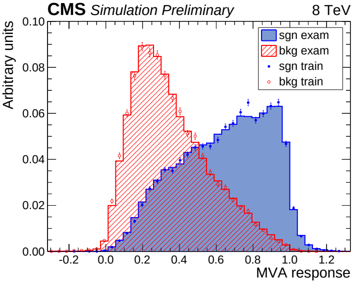
Sum of μ and e channels

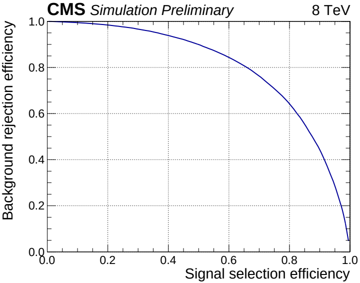
$H \to b\bar{b}$: Input variables for tHq vs bkgs classification



Sum of μ and e channels

$H \to b\bar{b}$: Response and ROC of classification MVA





$H \to b\bar{b}$: Systematics

Source	Type	impact as exclusive source on final limit [%]	improvement of final limit after removal [%]
JES	shape	17	3
JER	shape	< 1	< 1
BTag light flavor	shape	13	< 1
BTag heavy flavor	shape	17	< 1
Pile up	normalization	< 1	< 1
Unclustered energy	shape	3	1
Lepton efficiency	normalization	5	< 1
Luminosity	normalization	10	< 1
Cross section (PDF)	normalization	8	< 1
Cross section (Scale)	normalization	9	< 1
MC Bin-by-Bin unc.	shape	< 1	< 1
Q^2 scale ($tHq + t\bar{t}$)	shape	20	4
Matching	shape	2	2
Top p_T reweighting	shape	19	2
$t\bar{t}$ HF rates (b)	normalization	13	< 1
$t\bar{t}$ HF rates ($b\bar{b}$)	normalization	15	< 1
$t\bar{t}$ HF rates (c / $c\bar{c}$)	normalization	13	1

$H \to b\bar{b}$: Post-fit distributions in MVA response