ttH and tH Searches at CMS

Prof. Chris Neu Department of Physics University of Virginia



Top Quark Physics at the Precision Frontier Fermilab LPC 17 January 2018



Photo: Amy Scroggins, Fermilab Today

Why?

- What place does a discussion on ttH / tH <u>searches</u> have_at a workshop on precision top quark physics?
 - Challenging topology to isolate searches suffer from huge backgrounds and/or difficult to mitigate systematic uncertainties
 - Some systematics are not just difficult to mitigate but difficult to accurately understand their impact
 - CMS and ATLAS have only so far established the ttH process at the 3-4 std. dev. level, and it's been a long road to get to that point
 - *It's very likely there!*
 - Significant uncertainty remains in trying to pin down the ttH content in the data
 - We can't yet precisely say how much though ...
- Looking forward however, the inclusion of top-Higgs analyses in a precision top quark physics program is *essential*





A brief interlude into Higgs physics...







V

Christopher Neu

- Higgs couplings to particles of the SM a primary target of the characterization campaign
- Look for deviations from SM predictions for couplings to vector bosons and fermions

$$g_{HVV} = 2 \frac{m_V^2}{v}$$
 $g_{Hff} = \frac{m_f}{v}$

$$g_{HVV} = \kappa_V \left(2 \frac{m_V^2}{v} \right) \quad g_{Hff} = \kappa_f \left(\frac{m_f}{v} \right)$$

- Here assume same κ_V for each vector boson, and same κ_F for all fermions
- $(\kappa_{V}, \kappa_{F}) = (1.0, 1.0)$ is the SM





- Assume SM only particles participating in loop-mediated processes and BR(BSM)=0
- Examine prominent unique couplings that are accessible



But if one allows for the presence of BSM particles, things look a little less settled.



6

But if one allows for the presence of BSM particles, things look a little less settled.



Loops and Higgs and Top



- Workhorse analyses already probe the top-Higgs coupling, though there are issues...
- Consider gluon fusion:



Loops and Higgs and Top



- Workhorse analysesalready probe the top-Higgs coupling, thoughthere are issues...
- Consider gluon fusion:



But what about this?





Loops and Higgs and Top



The results of the multi-channel fits indicate that the couplings that "are SM-like".

To really know what is going on, **we need a direct probe** of the top-Higgs coupling...

• Similar problems on the decay side:



...but what about...





Motivation: The top-Higgs Coupling

- Within the SM, the Higgs coupling to the top quark, Y_t, is predicted to be by far the largest of all the fermionic couplings
 - m_{top} implies relatively large coupling, $Y_t \sim 1$
 - ~x30 larger than Y_b
 - ~x100 larger than Y_{τ}
 - Strongest coupling among all known SM particles
- Fermionic couplings:
 - LHC analyses have so far only been able to probe the Yukawa couplings Y_b and Y_{τ} directly
- Also Y_t will be the easiest (and perhaps only) up-type fermion coupling to probe
- Best channel: **ttH production**
 - tHX production also accessible sensitivity to sign of Y_t

Loop-induced processes, for example:



ttH production:



<u>Imperative</u>: Absolutely need to measure Y_t directly to know the true nature of the couplings of the new boson.

Summary of CMS ttH Analyses



Summary of CMS ttH Analyses

	H→ bb		$H \rightarrow \tau \tau$ $\tau_{had} \tau_{had}$	$Multilepton \\ \tau_{had} + \tau_{lep}$		$H \rightarrow \gamma \gamma$
7 TeV	CMS- HIG-12-035 (NN)	JHEP 1305 (2013) 145 (NN)				various
8 TeV	EPJC 75 (2015) 251 (ME) CN	IS-HIG-13-0 (BDT)	19	CMS-HIG-13-020 (SS-2lep, 3lep, 4 lep)		various
			JHEP 090	(2014)087		
⁵¹⁰² 13 TeV	²⁰² CMS-HIG-16-004 CMS-HIG-16-038			CMS-HIG-15-008		
2016			CMS- HIG-17-003		CMS- HIG-17-004	CMS- HIG-16-020

Christopher Neu

V,

ttH, H \rightarrow bb







Overview: $H \rightarrow bb$

- H→bb is a prime target of ttH analyses:
 - Largest Higgs BR for M_H=125
- CMS considers two topologies:
 - Single-lepton
 channel: one high
 p_T isolated e/µ, ≥4
 jets, ≥2 b tags
 - Dilepton channel: two opposite-sign leptons, ≥2 jets, ≥2 b tags
- Split selected events into categories based on jet, b-tags



A discriminant is devised in each category for signal extraction and a simultaneous fit is performed across all categories.

Low-signal categories serve to help constrain backgrounds.

Details of signal extraction in backup.

Christopher Neu

Big Issue: Understanding the tt+HF Background



- Modeling of tt+jets process:
 - Powheg+Pythia8, normalized to NNLO prediction
 - Separate templates for tt + b, tt + bb, tt +
 2b, tt + cc, tt + LF
 - 50% rate uncertainty per tt + HF process, uncorrelated in final fit
 - Among the leading uncertainties
 - Add. sources include parton shower, hadronisation, PDF, ISR/FSR

- tt+bb production poses irreducible background:
 - Poorly known theoretically
 - Measurements of ttbb CRUCIAL



ttH,H→bb: Results

- Not yet achieving sensitivity to SM-level of ttH,Hbb production
- Upper limit on $\mu = \sigma / \sigma_{SM:}$
 - $\mu < 1.5 (1.7) \text{ obs (exp) at } 95\% \text{ CL}$

			11.4	- 12.9	fb ⁻¹ (13	3 TeV)
	CMS Preliminary					
			μ	tot.	stat.	syst.
Dilepton	⊢+∎		-0.04	+1.50 -1.39	+1.05 -0.96	+1.01 -1.06
Lepton+jets	┞╌┼╼╋╌┥	-4	-0.43	+1.02 -1.02	+0.51 -0.52	+0.88 -0.87
Combined	⊧∔∎-	+-1	-0.19	+0.80 -0.81	+0. 4 5 -0.44	+0.66 -0.68
	-2 0		2	4		6
CMS-HIG-16-038 Best fit $\mu = \sigma/\sigma_{SM}$ at $m_{H} = 125 \text{ GeV}$						



- Best fit: $\mu = -0.19 \, {}^{+0.45}_{-0.44}$ (stat) ${}^{+0.66}_{-0.68}$ (syst)
- Systematics limited...how can we improve?

The Key: Understanding tt+HF

- Canonical ad-hoc 50% rate uncertainty on all tt+HF processes
 - tt+charm even less known than tt+b-jets
 - Huge impact on analyses
- The tt+b-jets process is poorly understood
 - Only recently do we have NLO calculations for the xsec
 - And even more recently NLO ME+PS events for use in analyses
- But NLO ≠ better, necessarily, if the predictions are poor
- Focus currently:
 - Compare various NLO ME+PS events for tt+bb
 - Consistency under well-defined conditions?
 - New scale treatment in MG5_aMC@NLO 2.5.4?
 - How do these state-of-the-art tools compare to CMS data
 - Need control regions independent from ttH signal-extraction campaign



See Section I.6.8 of **Yellow Report 4** of the Higgs Cross Section Working Group

ttH, H \rightarrow multileptons (WW, ZZ, $\tau\tau$)







ttH, H \rightarrow multileptons

- $ttH, H \rightarrow leptons:$
 - − Targeted Higgs decays and BR H→WW* (~20%) , $\tau\tau$ (6%), ZZ (3%)
 - leptons originate from Higgs and top system
- Targeted experimental signatures include multiple leptons
 - 2 same-sign leptons (2lss)
 - 3 leptons
 - 4 leptons

Event selection and signal extraction details in the backup



Low-rate but relatively low-background signatures.

ttH, H → multileptons: Signal Extraction

- Further categorization occurs to exploit different signal purities available via btagging requirements and the charge asymmetry present in ttW
- Signal extraction:
 - 2lss, 3lep: Two BDT classifiers one each targeting tt+jets and ttV (V=W,Z)
 - 4lep: simple counting experiment



V

ttH, H → multileptons: Final Fit



- Final fit simultaneous across all categories, post-fit distributions shown here
- Starting to see sensitivity to SM-level of ttH production no scaling factors on the red histograms here!
- Fit prefers *a slight bit more signal* in the 2lep-ss channels than expected in the SM



ttH, H \rightarrow multileptons: Results



- Best fit: $\mu = 1.5^{+0.3}_{-0.3}(\text{stat})^{+0.4}_{-0.4}(\text{syst})$
- Significance of observation is 3.3 σ , whereas the expectation, assuming SM-level of ttH production was 2.4 σ
- Systematics limited...how can we improve?

Leading systematic uncertainties



- Opportunities:
 - Improve our understanding of authentic leptons but from non-prompt sources
 - Theoretical cross sections on ttW and ttW:
 - NLO currently good to ~±15%, driven by missing higher order terms
 - NNLO tricky computationally





ttH, H $\rightarrow \tau \tau$







ttH, H \rightarrow $\tau\tau$

- Historically, ttH,H \rightarrow $\tau\tau$ analyses were folded in to other endeavors:
 - In 8 TeV, signatures containing two hadronic tau decays were included in a "hadronic decays" analysis, along with H→bb
 - The 2lep-ss + hadronic-tau category was considered in the context of an earlier iteration of the ttH, H→mulitleptons search in 13 TeV.
- We do a fine job at taus here at CMS, no need for these signatures to play a subordinate role in other analyses
- First time in top-Higgs campaign at CMS: Dedicated ttH,H→ττ analysis focusing on the 2016 13 TeV data sample





ttH, H $\rightarrow \tau\tau$: Event Selection

Common selection



Designed to be orthogonal to ttH, multileptons selection to ease future combination.

ttH, H $\rightarrow \tau \tau$: Signal Extraction

• Extraction methods:

- 1lep + $2\tau_h$: BDT, ttH v. tt+jets
- **2lep-ss** + $1\tau_h$: MEM-based discriminant
 - Also split further by whether there is evidence of a hadronically decaying W. aka, "missing jets" v. "no-missing-jets"
- **3lep + 1\tau_h:** Two BDTs, one for tt+jets and one for ttV
 - Significant amount of commonality with ttH,multilepton approach







ttH, H $\rightarrow \tau \tau$: Results



- Best fit: $\mu = 0.72^{+0.62}$ -0.53 (stat \odot syst)
- Significance of observation is 1.4 σ , whereas the expectation, assuming SM-level of ttH production was 1.8 σ





ttH, H $\rightarrow \gamma \gamma$







- Very rare process yet very pure signature
- Important:
 - Completely reconstructible final state
 - No combinatoric background
 - Hence, only ttH search channel in which <u>one can reconstruct a clear mass peak!</u>
- Event selection:
 - 2 photons (requirements on BDT γ ID and EM deposits), $|\eta| < 2.5$
 - (sub)leading $\gamma p_T/m_{\gamma\gamma} > 0.5 (0.25)$
 - $-100 < m_{\gamma\gamma} < 180 \text{ GeV}$
 - Categorize events according to ttbar system decay:
 - Leptonic:
 - $\ge 1 p_T > 20 e \text{ or } \mu \text{ far from } \gamma \text{ and } M_Z \ge 2 p_T > 25 \text{ jets}, \ge 1 \text{ b-tag}$
 - Hadronic:
 - special BDT event classifier
 - − ==0 e or μ , ≥3 pT>25 jets, ≥1 b-tag

Event Categories	SM 125 GeV Higgs boson expected signal								Bkg			
Event Categories	Total	ggH	VBF	ttH	bbH	tHq	tHW	WH lep	ZH lep	WH had	ZH had	(GeV ⁻¹)
ttH Hadronic	5.85	10.99 %	0.70 %	77.54 %	2.02 %	4.13 %	2.02 %	0.09 %	0.05 %	0.63 %	1.82 %	2.40
ttH Leptonic	3.81	1.90 %	0.05 %	87.48 %	0.08 %	4.73 %	3.04 %	1.53 %	1.15 %	0.02 %	0.02 %	1.50

So considering a window of $M_{\gamma\gamma}$ = 125 ± 1.5 GeV, there will be ~4.5 background events in the ttH Leptonic category.

S/B ~ 0.85

ttH, H $\rightarrow \gamma \gamma$

CMS-HIG-16-040

- Backgrounds so low allows for very simple signal extraction:
 - Determine signal shape in m_{γγ} exploiting superior resolution of CMS crystal ECAL
 - Assume a falling exponential in m_{γγ} for the uncorrelated diphoton background
 - See what amount of signal is favored in the data for a specific M_H hypothesis



CMS-HIG-16-040

ttH, H $\rightarrow \gamma \gamma$



- Results from two ttH categories combined:
 - $-\mu_{ttH} = 2.2 + 0.9_{-0.8}$, assuming M_H = 125.4 GeV
 - Uncertainty driven by statistics
- Largely an afterthought...but will be a workhorse
 - Many recent changes in analysis of full 2016 data sample targeted for improving ttH sensitivity
 - Good things come to those who wait...and build a solid analysis in the meantime

Single top + Higgs Searches







tHq Analyses: Different Approach to Y_t



V

Christopher Neu

q

•

•

•

coupling

 $\kappa_{\rm V} = g_{\rm HVV} / g_{\rm HVV(SM)}$ $Y_t = \kappa_t Y_{t(SM)}$ $\sigma(tHq) \approx a \kappa_t^2 + b \kappa_{\rm V}^2 + c \kappa_t \kappa_{\rm V}$



tHq in Multilepton Signatures at 13 TeV

CMS-HIG-17-005



- Similar event categorization as in ttH,multilepton search
- For $\kappa_t / \kappa_V = -1$, observed (expected) upper limit on $\sigma xBR = 0.64$ (0.32) pb





Summary

- Higgs physics has now moved from the search and discovery phase into a precision measurement era
- Characteristics of this Higgs boson need to be measured with high precision. The measurement campaign has so far revealed no significant deviations from the predictions of the SM
- A few crucial ones remain to be measured the most foremost being the coupling between the top quark and the Higgs boson

2017 Status, 13 TeV analyses	Signal Strength $\mu = \sigma / \sigma_{SM}$
H→bb	-0.19 $^{+0.45}_{-0.44}$ (stat) $^{+0.66}_{-0.68}$ (syst)
multileptons	1.5 ^{+0.3} _{-0.3} (stat) ^{+0.4} _{-0.4} (syst)
Н → тт	0.72 $^{+0.62}_{-0.53}$ (stat \bigotimes syst)
Н→үү	$2.2^{+0.9}_{-0.8}$ (stat \odot syst)

• First direct measurement of the top-Higgs coupling is among the primary goals of the LHC physics program.



What's Next

- Near term:
 - Establish ttH in all accessible decay channels
 - We have some work to do to make this happen:
 - Improve understanding of tt+HF process and uncertainties
 - Improve theoretical understanding of ttV
 - Improve upon already-mature treatment of non-prompt leptons
- Longer term
 - SM-driven backgrounds to ttH, $H \rightarrow \gamma \gamma$, ie tt $\gamma \gamma^*$ at NLO
 - Refine background models
 - Increase purity
 - Differential cross sections

Things like EFTs / top partners / exotic 4th gen / 2HDM / etc look like SM top-Higgs...until you look closely, in the tails. We will enter that regime in the future – best to lay the groundwork now.











ttH: Experimental Challenges

•

Christopher Neu

- Higgs production in association with a top-quark pair
 - Comparatively small production cross section wrt other Higgs production channels
 - Spectacular signature rich final state
- ttH production cross section dwarfed by main background, tt+jets:

root(s) [TeV]	7	8	13
σ (ttH (125)) [fb]	90	130	510
σ (tt+jets) [fb]	177000	253000	830000
Ratio	5.0E-4	5.1E-4	6.1E-4

ttH,H→bb: Signal Extraction

- Single variable not very sensitive
 - Jet energy resolution
 - Combinatorics in jet assignment
- Different choices and combinations of more advanced classifiers
 - Boosted-Decision-Trees (BDTs) and Neural Networks (NNs)
 - Combination of various input variables, trained per category
 - Separation against all tt +X processes
 - Used in signal vs. background separation and in event reconstruction
 - Matrix-Element-Method (MEM) classifiers
 - Likelihood of event kinematics under signal or background hypothesis
 - Particularly powerful against difficult tt + bb background



BDT inputs in SL ≥6j, ≥4t category:

- Best Higgs candidate mass
- $M(t_i, t_j)$ closest to 125
- M_{l,v,jets}
- 4th and 5th highest b-tag disc score
- Sum P_T of all jets, lepton, MET

ttH,H→bb: Signal Extraction

 Great deal of power added to the final analysis by combining these different multivariate approaches.

- For instance, the MEM discriminant can be used as an input to a BDT along with other kinematic variables
- Or as is done here, further categorize according to BDT score and then use the MEM in each subcategory...



V

ttH,H→bb: Signal Extraction



ttH,H→bb: Results



= highest signal/background, hence highest sensitivity categories

ttH,H→bb: Results



- Highest signal purity / sensitivity bins show upward fluctuation
- Compensated by overall deficits elsewhere



Search for ttH Production, H→Multileptons

Same-sign dilepton

2 e/ μ with $p_T > 20$ ≥4 jets with $p_T > 25$ ≥1 b-tagged jet

Three lepton

1 e/ μ with $p_T > 20$ 1 e/ μ with $p_T > 10$ 1 e(μ) with $p_T > 7(5)$ ≥ 2 jets with $p_T > 25$ ≥ 1 b-tagged jet Veto $M_{11} \sim M_Z$

g 700000

g 000000

Four lepton

1 e/ μ with $p_T > 20$ 1 e/ μ with $p_T > 10$ 2 e(μ) with $p_T > 7(5)$ ≥ 2 jets with $p_T > 25$ ≥ 1 b-tagged jet Veto $M_{11} \sim M_Z$





Low-rate but relatively low-background signatures.

Η

ttH, H → multileptons: Per-Category Yields

CMS-HIG-17-004							
	μμ	eµ		ee			
tĪW	51.0 ± 0.6 (stat.) ± 6.9 (syst.)	$72.8\pm0.7~{ m (stat.)}\pm$	10.2 (syst.)	20.5 ± 0.4 (stat.) ± 3.1 (syst.)			
$t\bar{t}Z/\gamma^*$	$17.7\pm0.8~{ m (stat.)}\pm2.9~{ m (syst.)}$	47.3 ± 1.6 (stat.) \pm	⊦ 9.0 (syst.)	17.5 ± 1.0 (stat.) ±3.6 (syst.)			
WZ	$4.2\pm0.6~\mathrm{(stat.)}\pm4.1~\mathrm{(syst.)}$	7.0 ± 0.8 (stat.) \pm	⊢ 6.8 (syst.)	1.8 ± 0.4 (stat.) ±1.7 (syst.)			
Rare SM bkg.	4.2 ± 1.5 (stat.) ±3.0 (syst.)	13.3 ± 1.9 (stat.) \pm	⊦ 9.3 (syst.)	4.8 ± 1.1 (stat.) ±3.6 (syst.)			
WWss	3.5 ± 0.6 (stat.) ± 2.5 (syst.)	4.1 ± 0.6 (stat.) \pm	∃ 3.2 (syst.)	1.4 ± 0.3 (stat.) ±1.2 (syst.)			
Conversions		7.8 ± 2.5 (stat.) \pm	⊢ 2.3 (syst.)	3.6 ± 3.5 (stat.) ±1.7 (syst.)			
Charge mis-meas.		16.4 ± 0.2 (stat.) \pm	⊦ 9.1 (syst.)	$10.5\pm0.2~\mathrm{(stat.)}\pm5.9~\mathrm{(syst.)}$			
Non-prompt leptons	$38.7 \pm 1.6 \text{ (stat.)} \pm 20.5 \text{ (syst.)}$	61.8 ± 2.0 (stat.) \pm	13.0 (syst.)	$17.7 \pm 1.1 \text{ (stat.)} \pm 5.4 \text{ (syst.)}$			
All backgrounds	120.3 ± 2.5 (stat.) ± 11.7 (syst.)	231.2 ± 4.3 (stat.) \pm	13.3 (syst.)	$77.9 \pm 4.0 \; ({ m stat.}) \pm 9.0 \; ({ m syst.})$			
tīH signal	20.1 ± 0.5 (stat.) ±2.1 (syst.)	27.9 ± 0.5 (stat.) \pm	∃ 3.0 (syst.)	8.0 ± 0.3 (stat.) ±1.1 (syst.)			
Data	150		268	89			
	3L			4L			
tĪW	32.8 ± 1.0 (stat.)	$)\pm4.9$ (syst.)					
$t\bar{t}Z/\gamma^*$	49.8 ± 3.9 (stat.)	\pm 11.1 (syst.)	2.15 ± 0.2	24 (stat.) \pm 0.44 (syst.)			
WZ	9.1 ± 0.9 (stat.)	$)\pm4.0$ (syst.)					
Rare SM bkg.	8.8 ± 4.3 (stat.)	$)\pm5.9$ (syst.)	0.27 ± 0.1	6 (stat.) \pm 0.19 (syst.)			
WWss							
Conversions	5.3 ± 1.2 (stat.)	$)\pm4.0$ (syst.)					
Charge mis-meas.		-					
Non-prompt lepto	ons 30.8 ± 1.5 (stat.)	\pm 10.9 (syst.)					
All backgrounds	137.3 ± 6.2 (stat.)	\pm 12.4 (syst.)	2.42 ± 0.2	$28 \text{ (stat.)} \pm 0.56 \text{ (syst.)}$			
ttH signal	19.5 ± 1.0 (stat.)	$)\pm3.0$ (syst.)	1.00 ± 0.0	09 (stat.) \pm 0.11 (syst.)			
Data		148		3			



tHq 8TeV Analyses

JHEP 06 (2016) 177

 $\underline{tHq}, H \rightarrow bb$ Leptonic W decay 3-,4-tag categories MVA for tHq v. ttbar

tHq, H \rightarrow WW, $\tau\tau$

Same-sign 2lep and 3lep MVA for tHq v. bkgd

$tHq, H \rightarrow \gamma\gamma$ Enhancement on production and decay side. No events survive in data.



Expected (observed) upper limits:

 $\sigma/\sigma(Y_{t} = -1) < 5.4$ (7.6)





$\sigma/\sigma(Y_t = -1) < 5.0$ (6.7)

 $\sigma/\sigma(Y_t = -1) < 4.1 (4.1)$

Combined upper limit $\sigma/\sigma(Y_{t} = -1) < 2.0$ (2.8)

