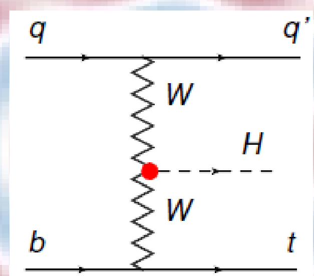
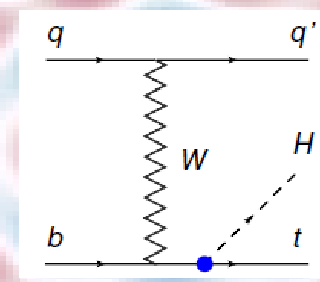
Single top + Higgs: quantum interference and the sign of y_t



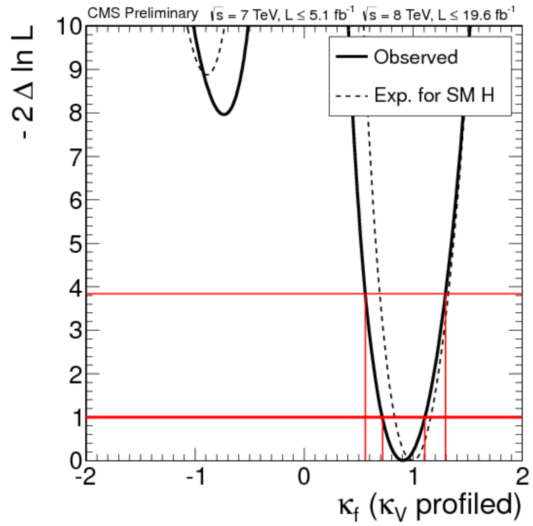


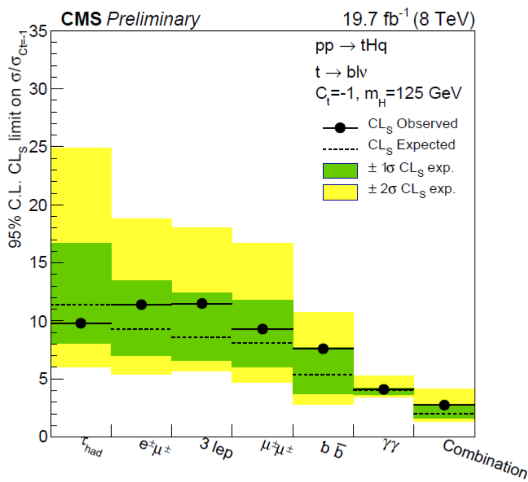
Andrea Giammanco

CP3 - UCL Louvain-la-Neuve

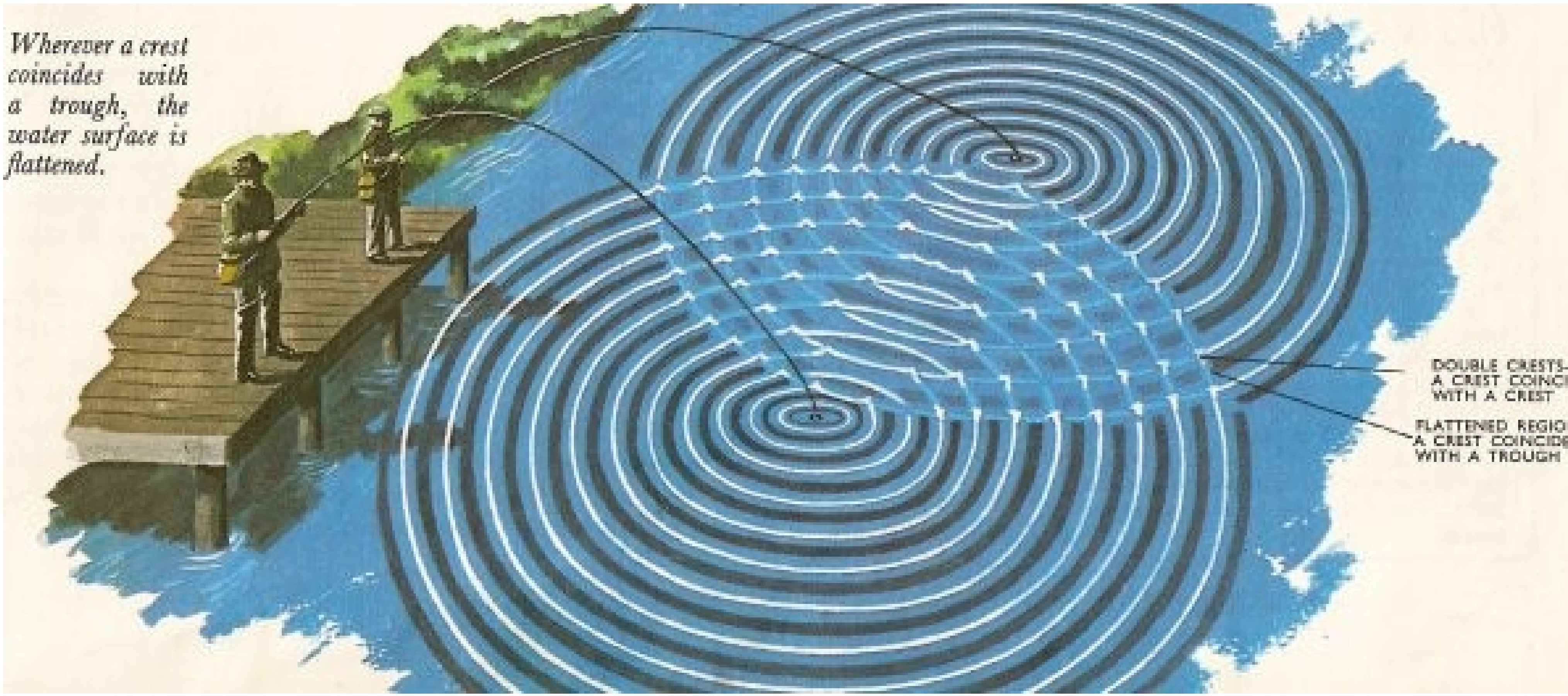
Outline

- Part I: Theoretical motivation
  - Degeneracies
  - A strawman model
  - Digression: other models
- Part II: The hunt in Run I with CMS data
  - Clean and rare: $H \to \gamma\gamma$
  - Plenty and dirty: $H \to b\bar{b}$
  - Somewhere in the middle: $H \to WW, \tau\tau$
  - Combination (preliminary!)
- Part III: Prospects for Run II

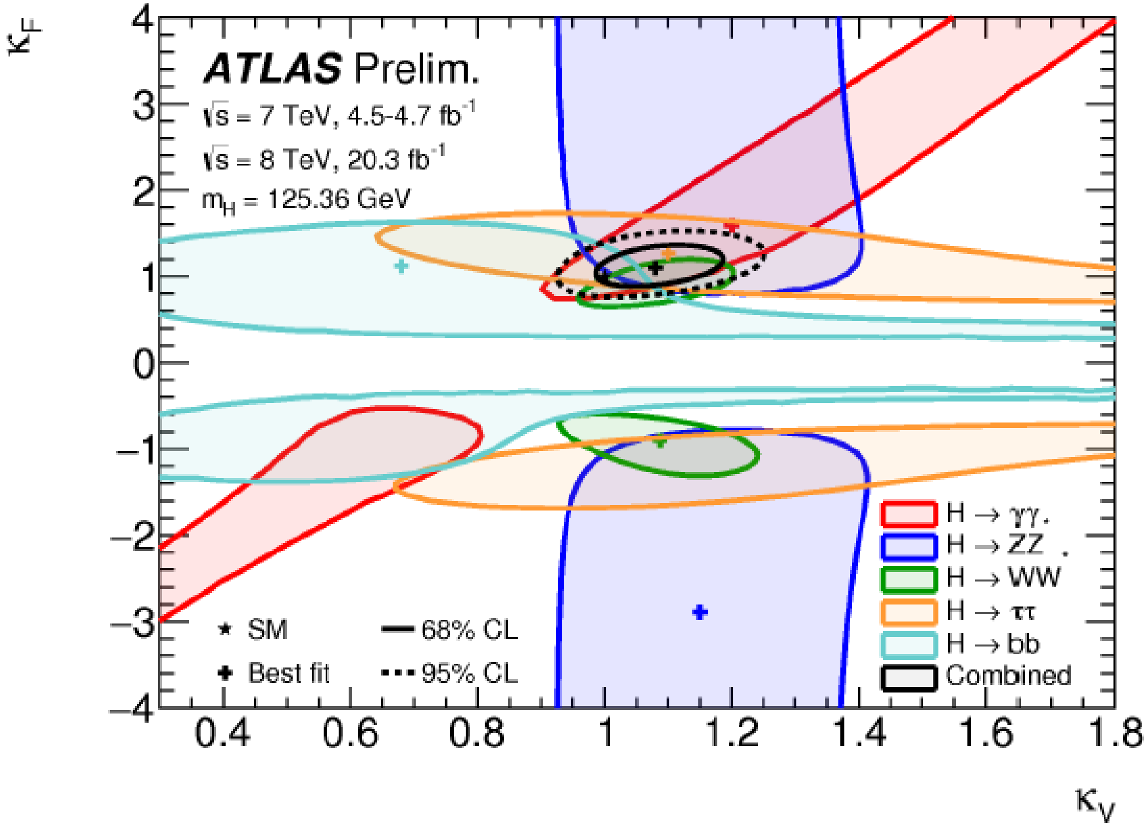


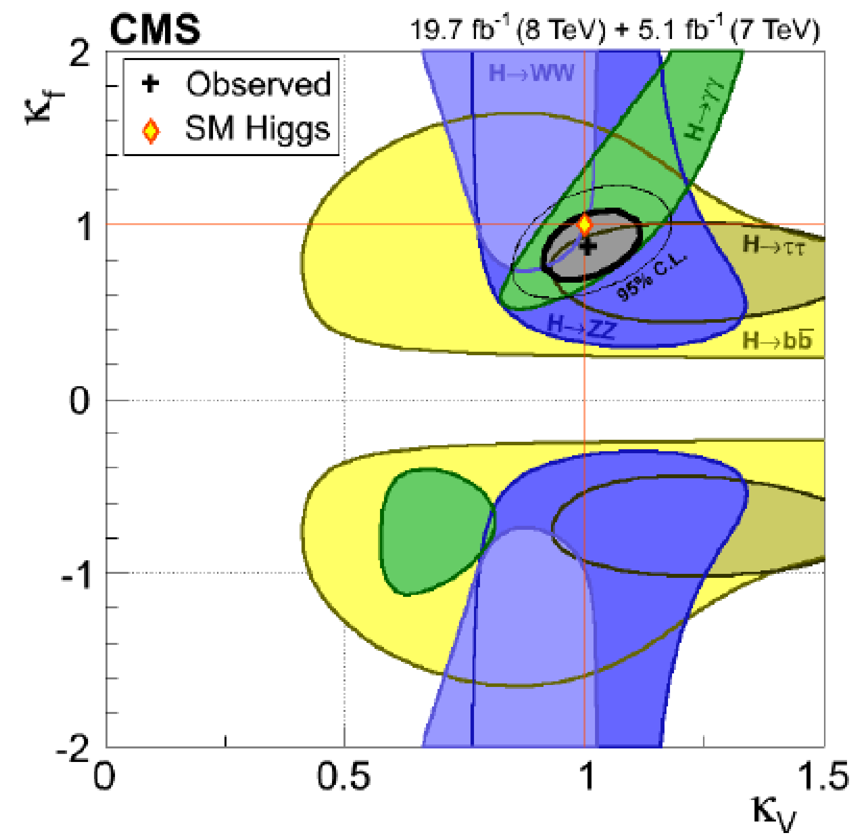


Part I: theory



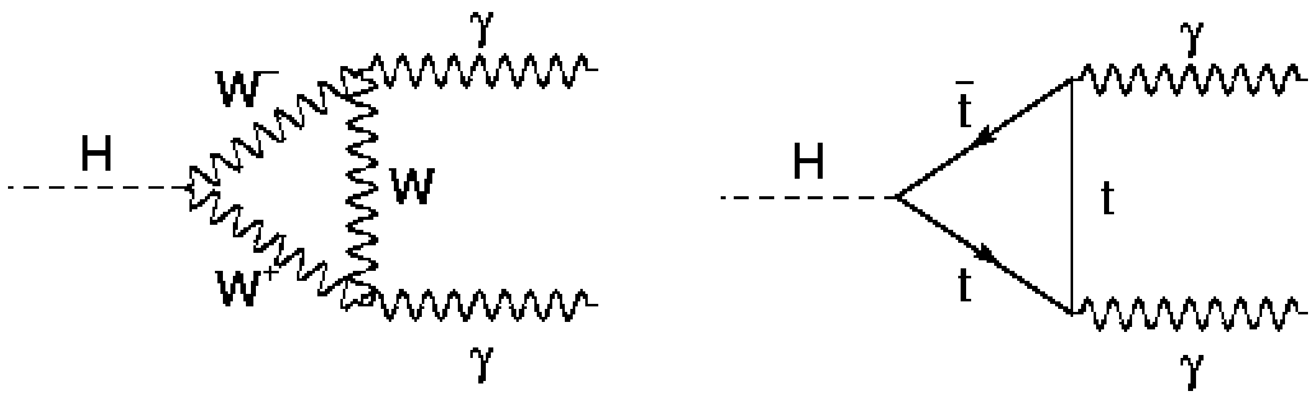
Degeneracies





- Scale factors k_V, k_F multiplying the SM Higgs couplings to bosons and fermions, assuming no new particles
- Most channels show perfect symmetry around k_F=0
- Degeneracy is broken by γγ channel

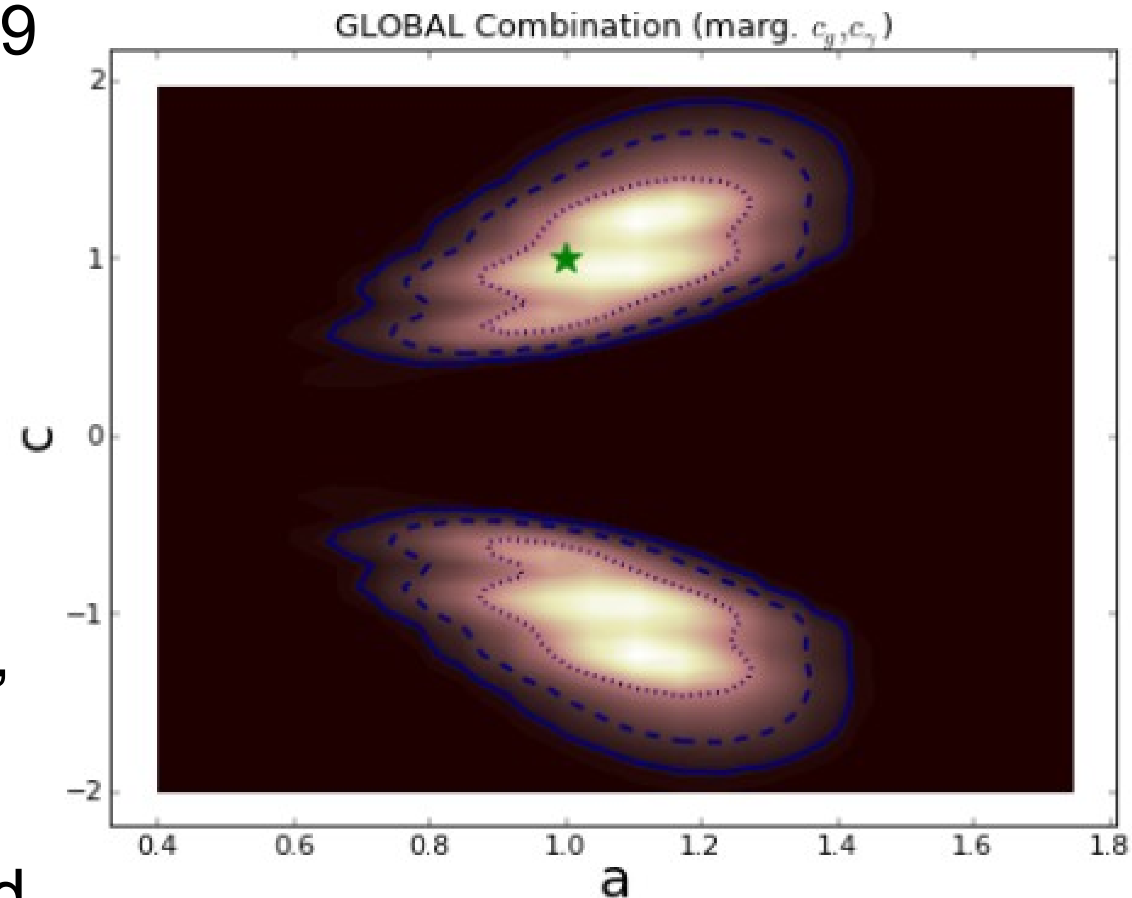
Sensitivity to the sign: from where?



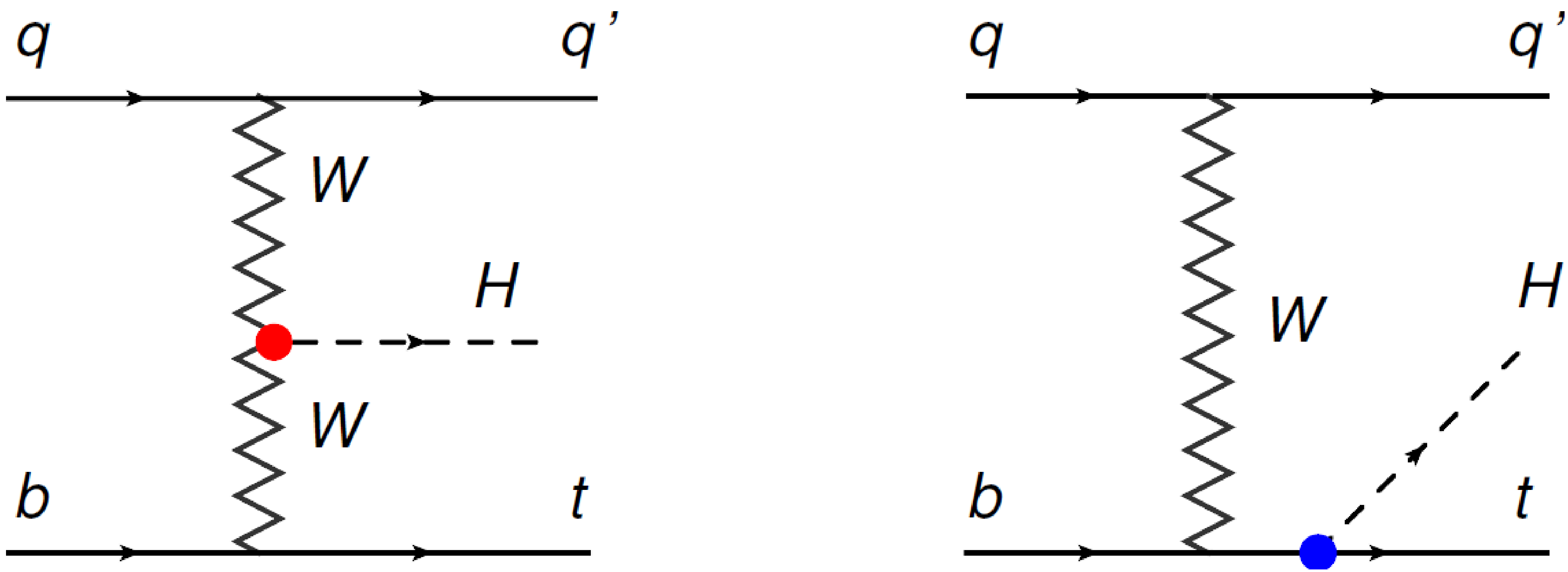
- Opposite sign amplitudes in the SM
  - Hence partial cancellation in the loop
- Flip the relative sign between y_t and g_{HWW} and you get constructive interference
  - It results in a ~2x enhancement in this branching ratio
  - No effect on ggF, VBF, W-associated, $t\bar{t}H$ cross sections
  - Effect on $gg \rightarrow ZH$, but very small wrt $q\bar{q} \rightarrow ZH$

Disfavoured?

- J.Ellis, T.You, JHEP 06 (2013) 103, arXiv:1303.3879
  - Based on home-made combination of CMS, ATLAS, and Tevatron (not up to date, but here it doesn't matter)
- Here *a,c* have similar meaning as k_V, k_F
- In the plot reproduced here, BSM contributions are allowed in ggH and Hγγ loops and marginalised, and **the minima are degenerate**



Looking for a better "interferometer"



- In tHq production, accidentally strong cancellation *at tree level* in the SM (only 18 fb @ 8 TeV)
  - As opposed to $H\rightarrow\gamma\gamma$ and $gg\rightarrow HZ$ which are loop-induced
- Hence, strong enhancement (~13x) if the relative sign between HWW and Htt couplings turns out to be negative

What we are looking for

- We call our signal model the tHq production with y_t = -1 or a completely free phase
  - Strawman model where all the rest of the SM is still valid; not a realistic model, because something new must explain a non-SM phase between Htt and HVV couplings
    - But useful to get a well defined prediction
    - A complete model may change numbers a bit, but changing sign (or phase) makes a difference at first order
- Goals
  - First (8 TeV data set): first limits in tHq final state **(done!)**
  - Next (early Run II): once and forever discover or exclude the *"SM with y_t=-1"* model
  - Then: sensitivity to range of phases, up to SM (y_t=+1)

Original inspirations

- I had the luck of chatting with some key theory colleagues at the right moment (E.Gabrielli at NICPB, F.Maltoni at UCL)
- Biswas, Gabrielli, Mele, arXiv:1211.0499, JHEP 01 (2013) 088
  - They proposed to look at tHq with $H \rightarrow \gamma\gamma$ (interference also in decay) and hadronic top decay; topology: 2γ+4j(1b,1fwd)
  - Follow-up paper, arXiv:1304.1822, JHEP 07 (2013) 073, with the inclusion of the channels 2γ+1l+2j(1b,1fwd) and multi-lepton
- Farina, Grojean, Maltoni, Salvioni, Thamm, arXiv:1211.3736, JHEP 05 (2013) 022
  - They proposed $H \rightarrow b\bar{b}$ (best branching ratio) and leptonic top decay; topology: 1l+4/5j(3/4b,1fwd)
  - My group (Louvain-Karlsruhe-Nebraska) chose this strategy, mostly because of our own experimental expertise
- A swarm of pheno papers on the subject followed, see backup

Theoretical framework

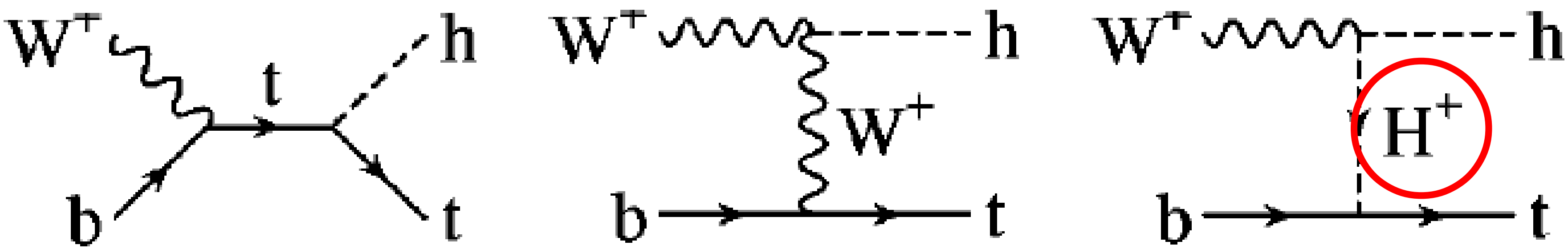
- *Effective approach*: try not to think too much of what specific brand of new physics can give a "wrong sign"
  - Find a balance between maximum sensitivity and minimum assumptions; deliver result in an easy-to-intepret form
- Of course, one always finds models that accomodate it:
  - S.El Hedri, P.J.Fox, J.G.Wacker, arXiv:1311.6488
    - "One possible scenario (...) is a Georgi-Machacek model (*) with one additional Higgs doublet. This model would predict a large number of new charged and neutral Higgses with sizable couplings to the top quark."

      (*) H.Georgi, M.Machacek, Nucl.Phys.B262, 463 (1985)
  - Ellis & You motivate with anti-dilaton model
- More ideas are welcome
  - You can win a citation!

Other models can be tested along the way

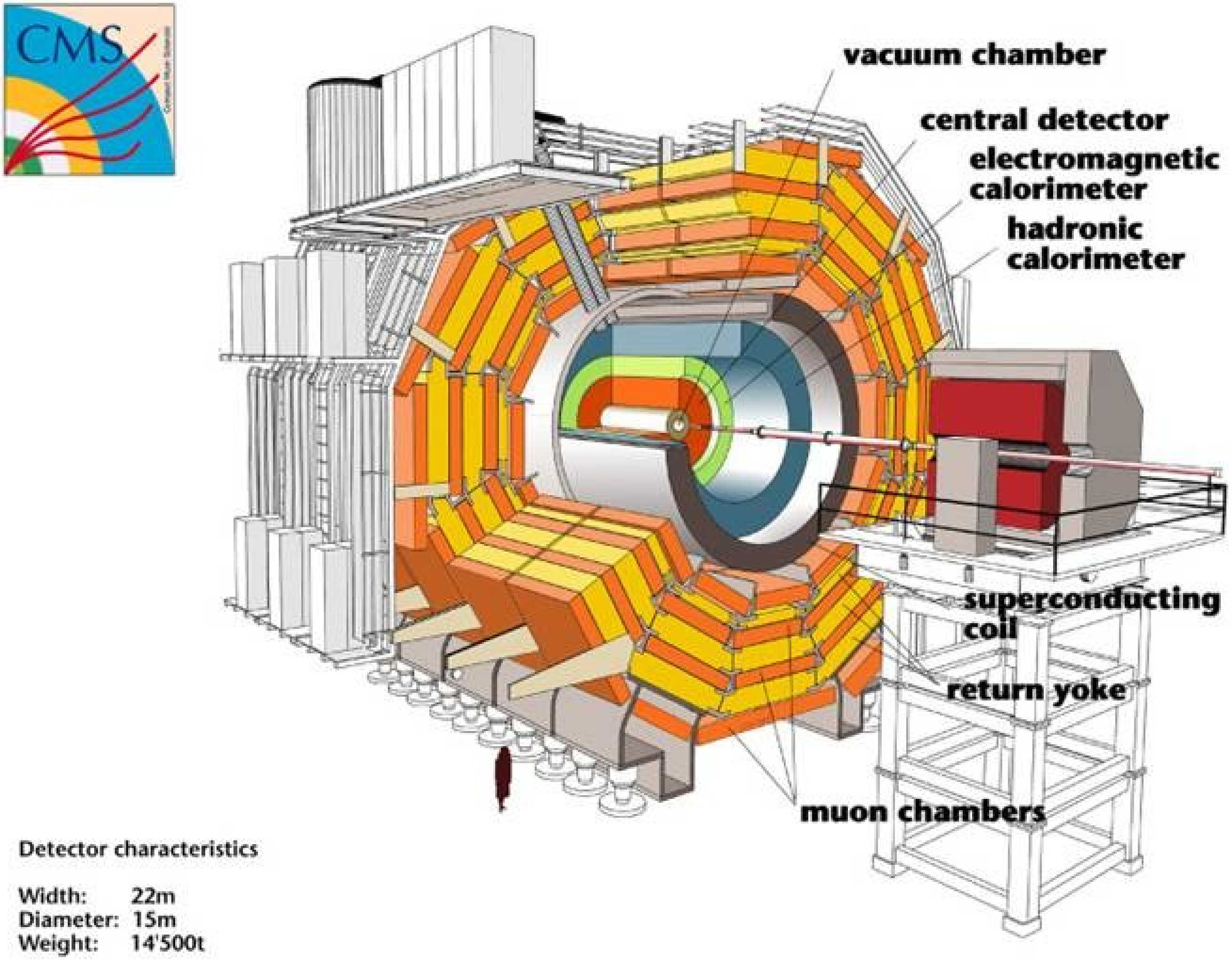
- Other models predict enhancement of single top + Higgs
  - *"Who ordered that?"*, said Rabi of the μ; here *"that"* is a fermion with Yukawa coupling ~1 while all others are ~0
  - Said in another way: mass of the order of the EWSB scale
  - Therefore, quite frequent for model-builders to ponder over the top quark when they try to make EWSB natural
  - ...and to predict larger anomalies in the top sector than for other quarks (quite conveniently, as the top quark entered the precision domain only recently!)
- Example: models with large FCNC in the top sector
  - Process $qg \to tH$(+jets), complementary to search for anomalous decay $t \to Hq$ in $t\bar{t}$ production
- Example: 2HDM

Example: 2 Higgs Doublets Models



- See for example Maltoni et al, arXiv:0106293[hep-ph], Phys.Rev. D64 (2001) 094023
- The charged H is what spoils the cancellation of the SM; in the ansatz of our strawman model, seen as effective phase
- The same diagrams with A instead of h are part of the signal
- MSSM, which is a particular case of 2HDM, gives only mild enhancements (~2x) in the most favorable cases
- Started a pheno project with F.Maltoni on generic 2HDM

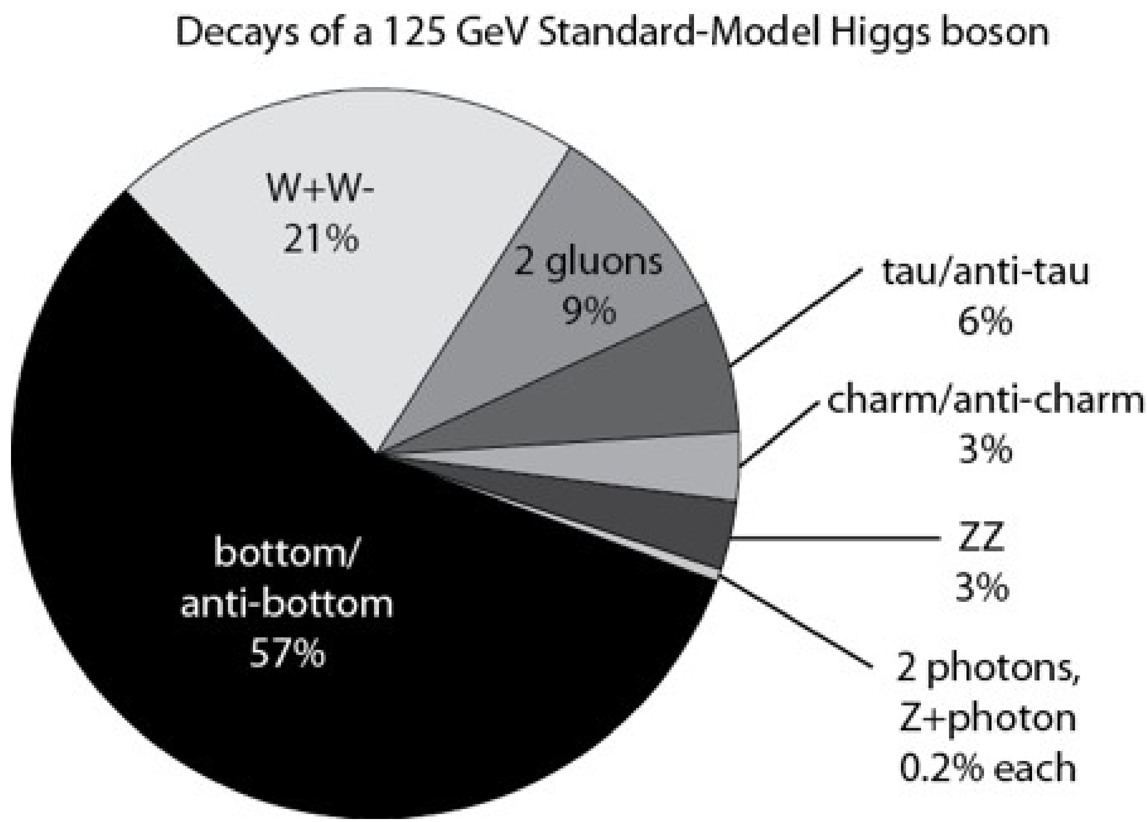
Part II: the experimental results



Analyses in CMS Run I

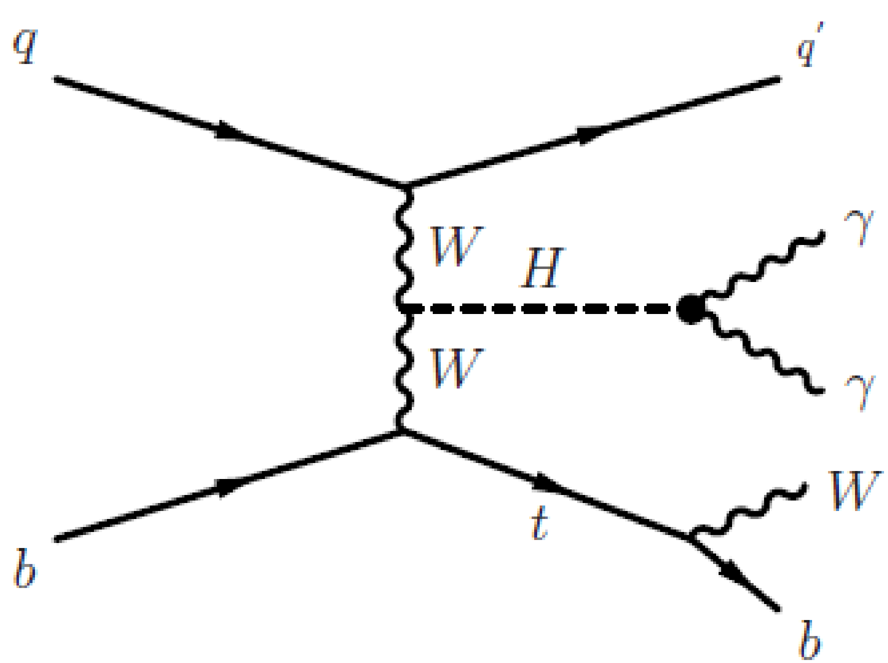
- CMS-HIG-14-001: **$H\rightarrow\gamma\gamma$**
  - BR = 0.2% in SM, which would ~ double if our signal is true
  - Very clean signature
- CMS-HIG-14-015: **$H\rightarrow b\bar{b}$**
  - Largest BR in SM, not much affected in principle
  - Very unclean signature (dominated by $t\bar{t}$ background), and very messy combinatorics
- CMS-HIG-14-026: **$H\rightarrow WW,\tau\tau$**; look at 2l (SS),3l
  - Intermediate features; bkg: mostly $t\bar{t}$+fake leptons
- All these analyses consider semileptonic decay of the top
- Combination paper to be submitted soon
  - Also including a fourth channel ($H\rightarrow\tau\tau\rightarrow e/\mu$+tau-jet)

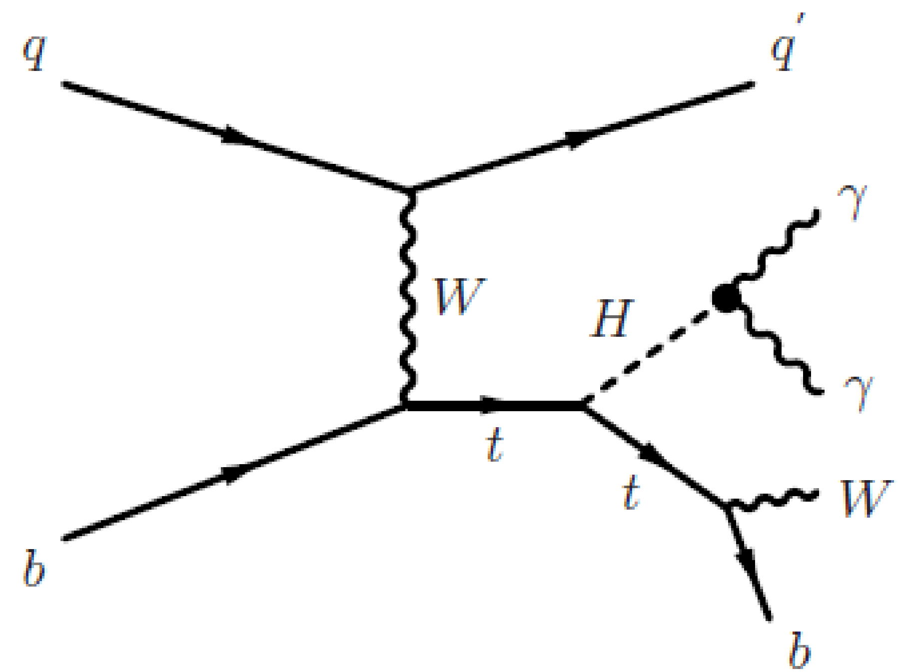
About branching ratios



- We are assuming that these BR's only need to be corrected for the photon channel enhancement
- Strong assumption but supported by direct measurements

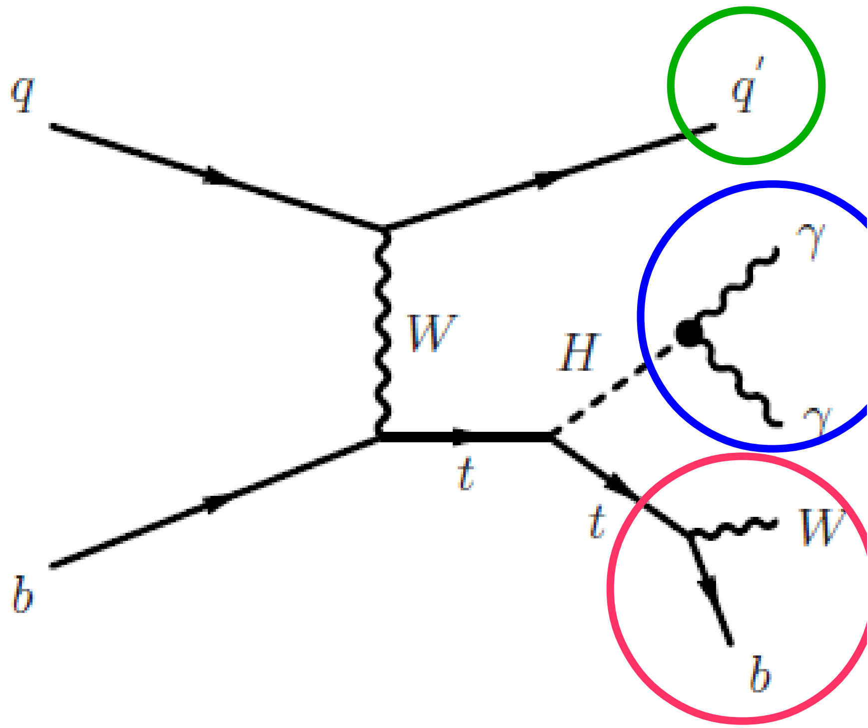
H→γγ analysis





Event pre-selection

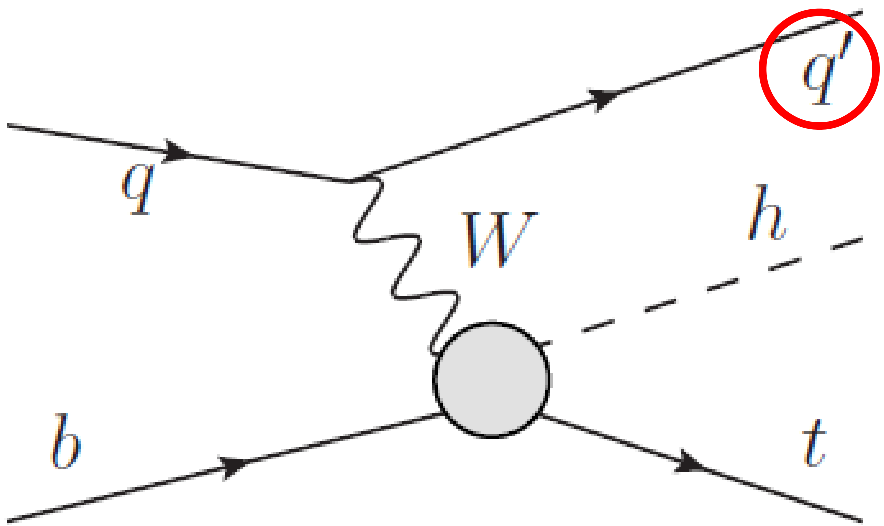
- Di-photon trigger
- Leading photon: $p_T > 50 * m_{\gamma\gamma} / 120$ GeV
- Sub-leading photon: $p_T > 25$ GeV
- Exactly 1 lepton (e/μ), $p_T > 10$ GeV
- At least 2 jets, $p_T > 20$ GeV
- At least one passes a tight b tagging
- The hardest non-b-tagged jet has $|\eta|>1$

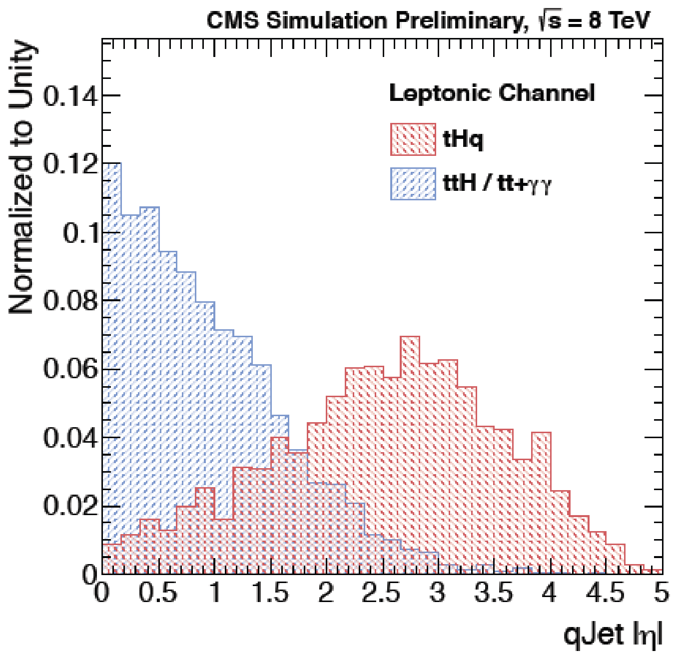


H(125) selection – Top selection – specific of tHq

Very clean selection; the main expected background around the H mass turns out to be $t\bar{t}H$!

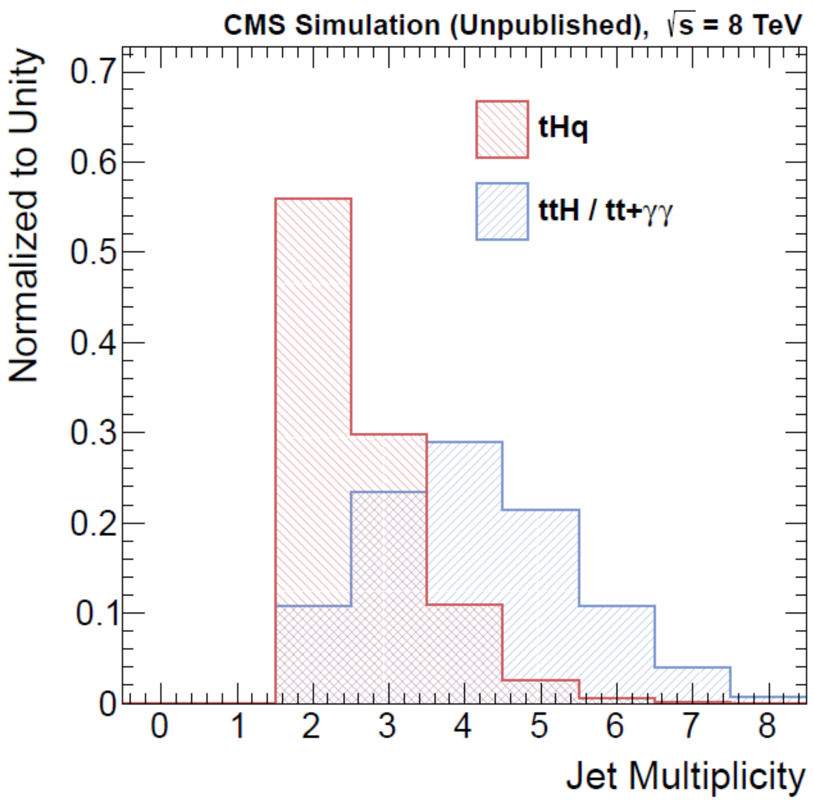
Discriminating variables (1)



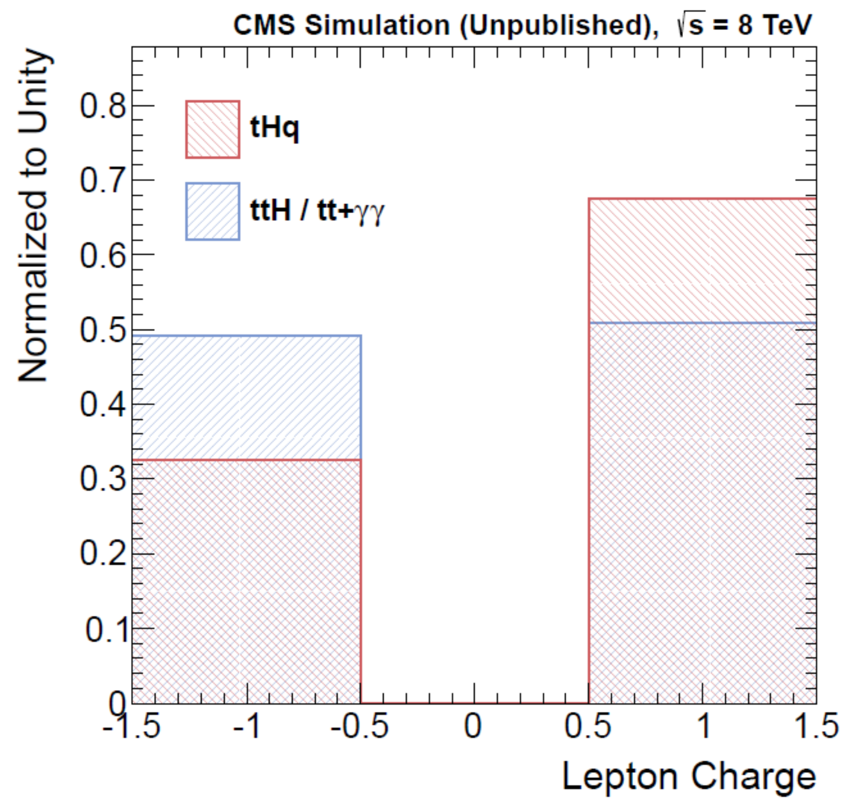


The recoil quark, which gives the only light jet in the event, has a rather characteristic pseudorapidity distribution.
This is mostly dictated by mere kinematics (it depends on the PDF and on the mass of the system it is recoiling against) and it is, therefore, a rather robust prediction of any suitable model.
All the tHq analyses make use of this variable.

Discriminating variables (2)

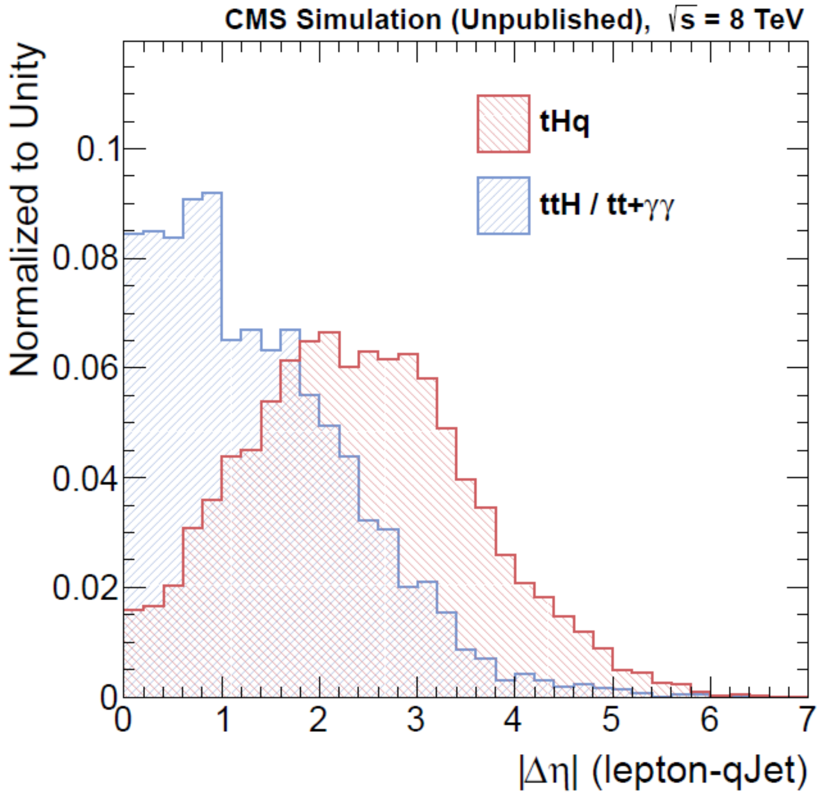


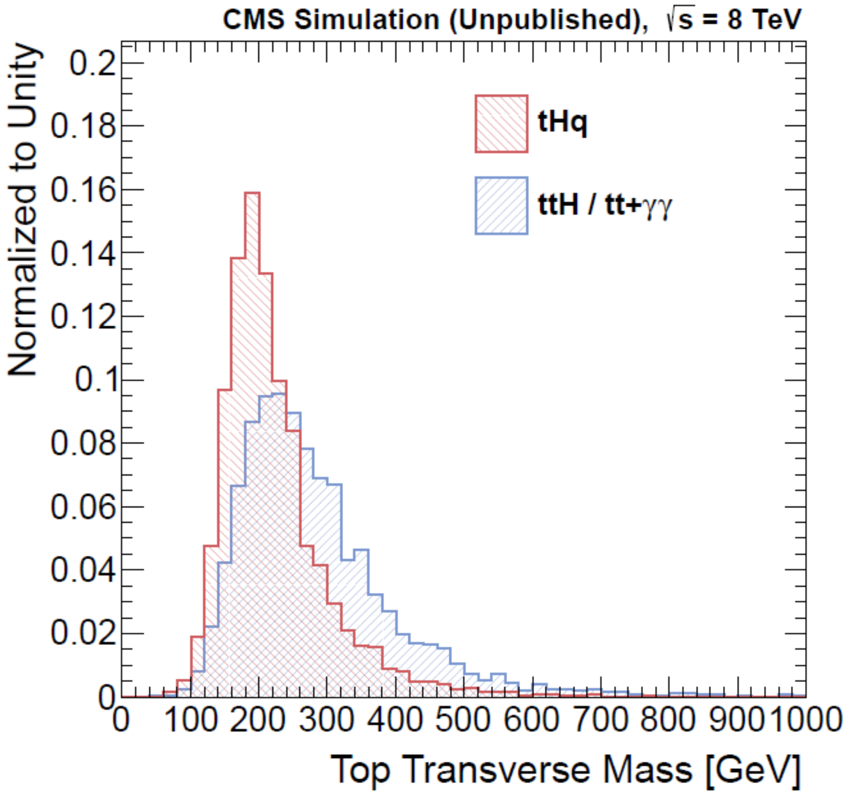
Signal has less jets than $t\bar{t}$+photons and $t\bar{t}H$



More up than down at high x, hence $\sigma(tHd) > \sigma(\bar{t}Hu)$

Discriminating variables (3)

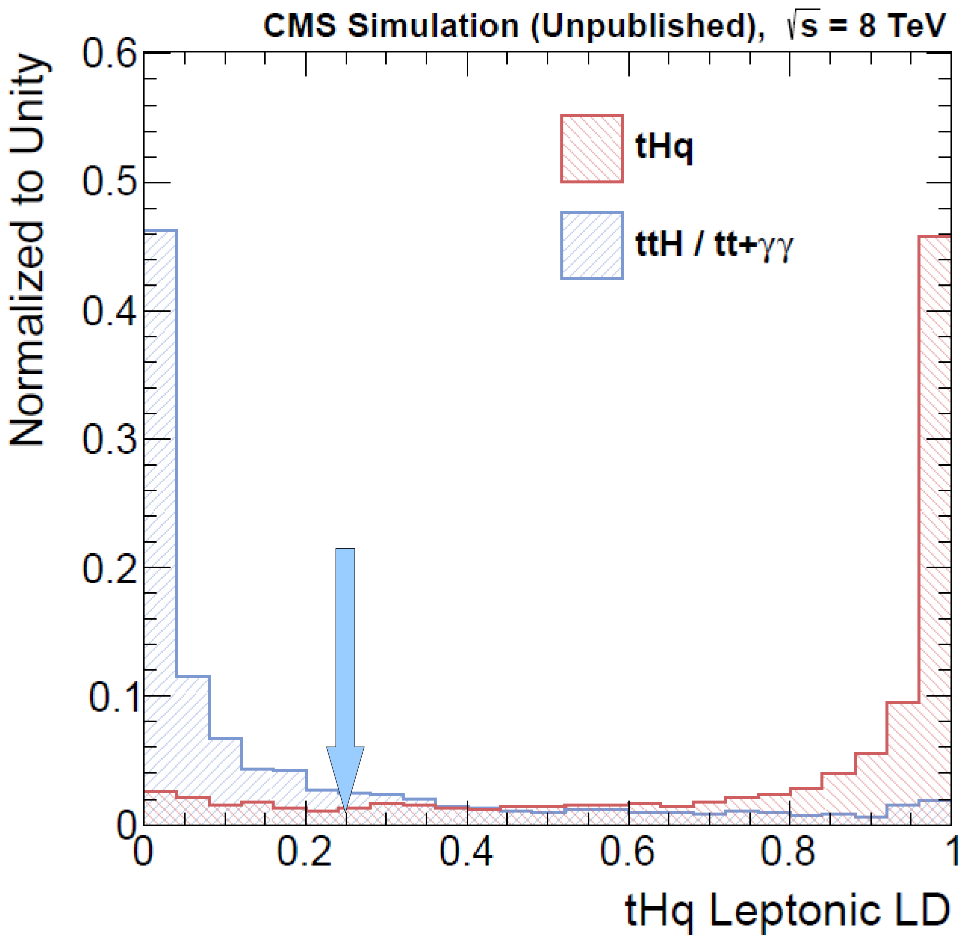




Lepton and light jet tend to have small |Δη| if coming both from top decay, large if the jet is from recoil

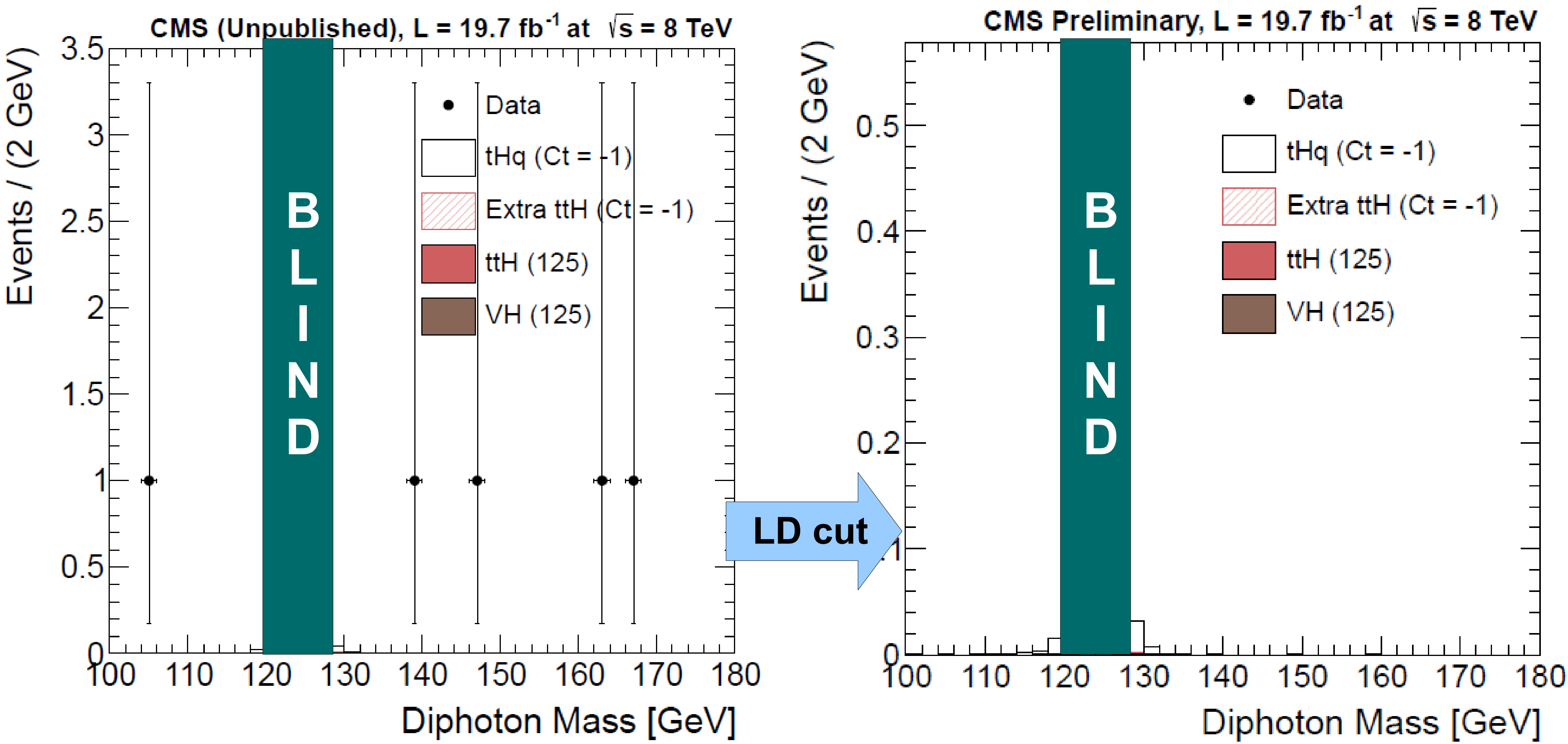
Top transverse mass distribution (from l, b, MET) broader if there are two tops

Likelihood discriminant (LD)



All previous variables are combined in this discriminant.
Cut chosen in MC, before looking at any data, to give
$\#(t\bar{t}H)/\#(tHq) < 10\%$ @ y_t=-1

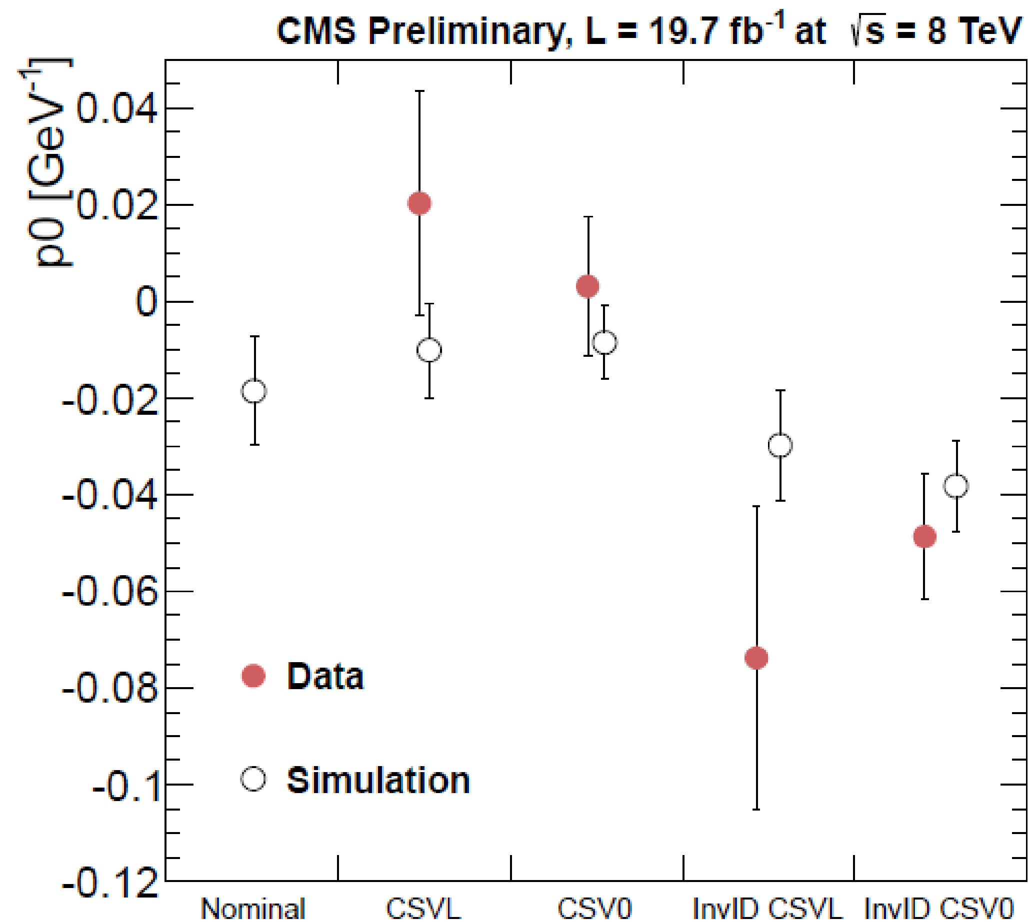
A bit of suspence...



Five events in the sidebands before LD cut, none left after!
Unusual for hadron collider analyses, more typical of ν and Dark Matter experiments

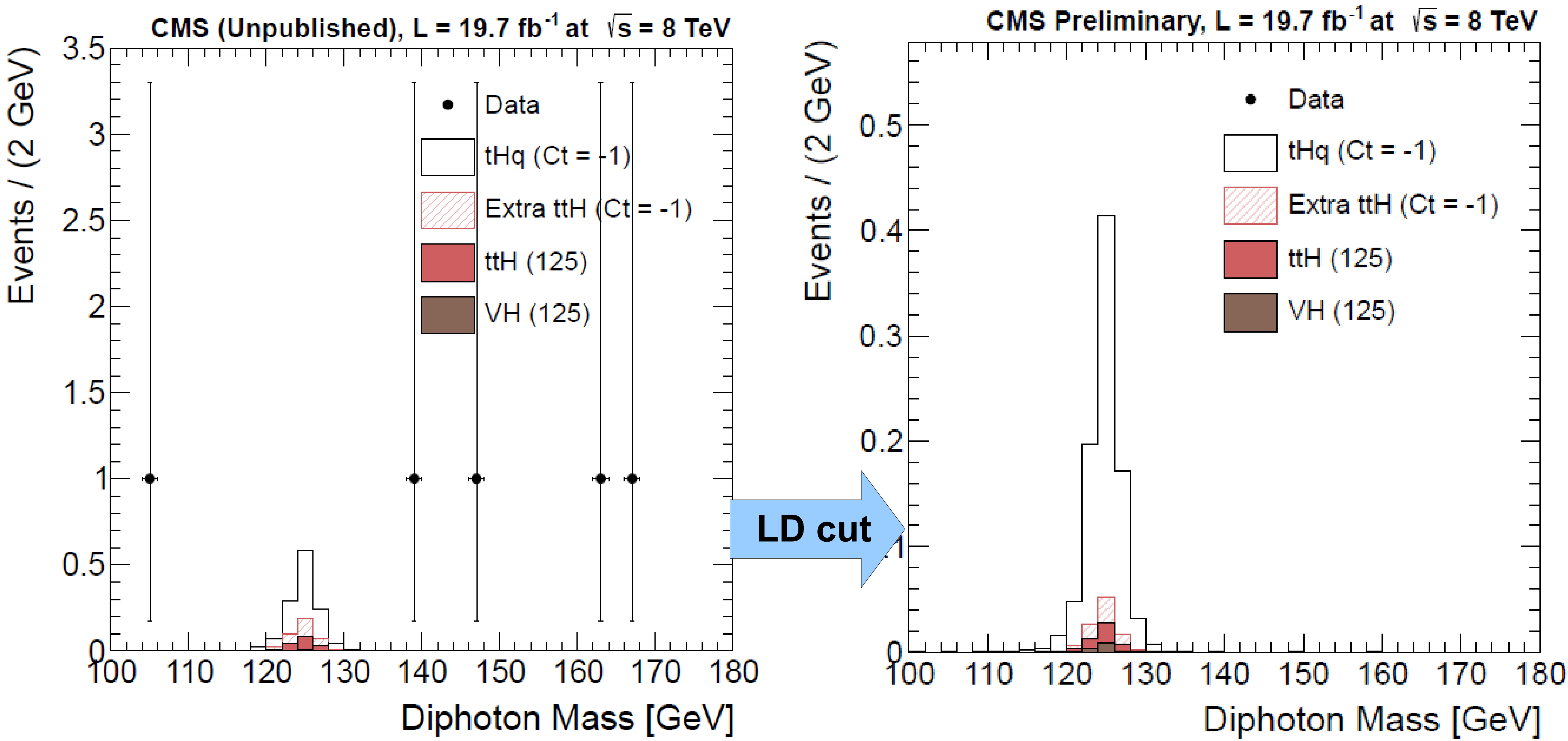
Non-resonant background estimation from data

- Interpolate from sidebands to signal window ([122,128] GeV) via a multiplicative factor α, assuming a $m_{\gamma\gamma}$ shape f_{bg}
- f_{bg} from exponential fit in four control regions
- 2 with looser b-tagging cut but same photon-ID, 2 also with inverted photon-ID
- Fair stability of the slope of the fitted exponential



Maximum $\Delta\alpha$ is 16%, taken as systematic

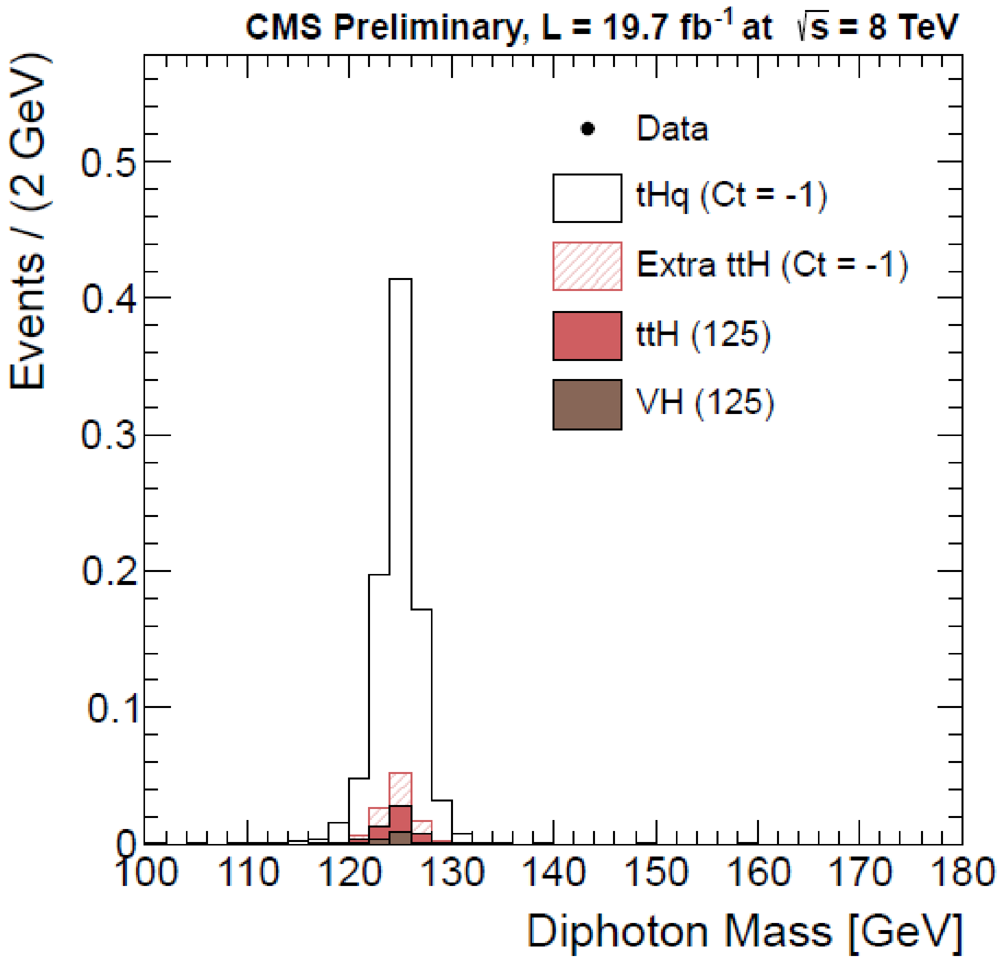
Now let's unblind



No event in signal window

Event yield

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

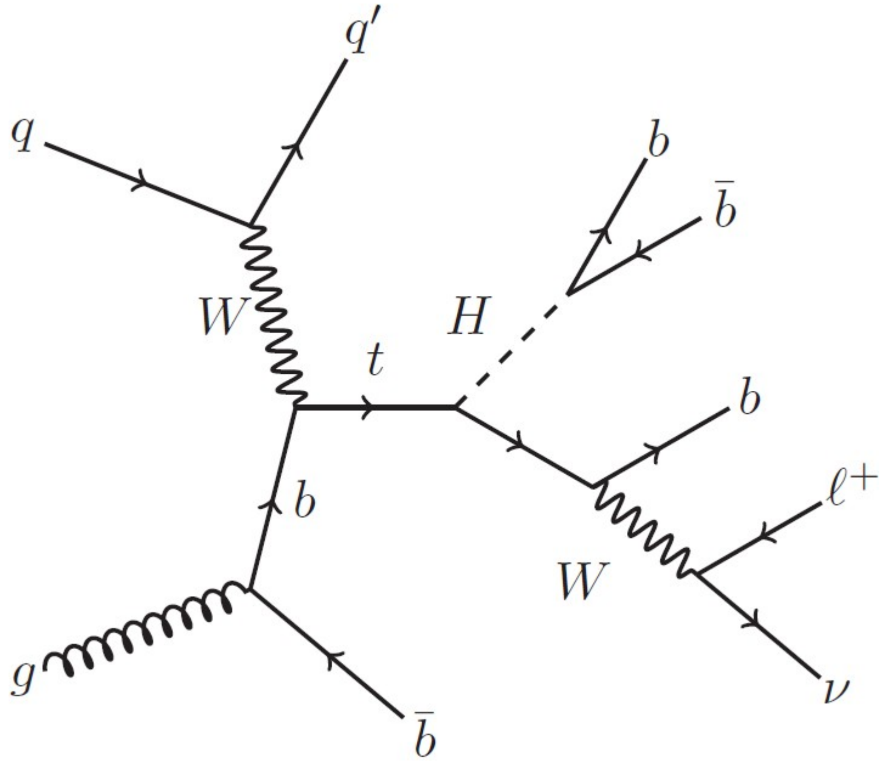


The extra $t\bar{t}H$ and VH (†) are accounted as part of signal.
Yields are counted in [122, 128] GeV.
Less than 1 event expected even in the y_t=-1 hypothesis

Result, and how to improve

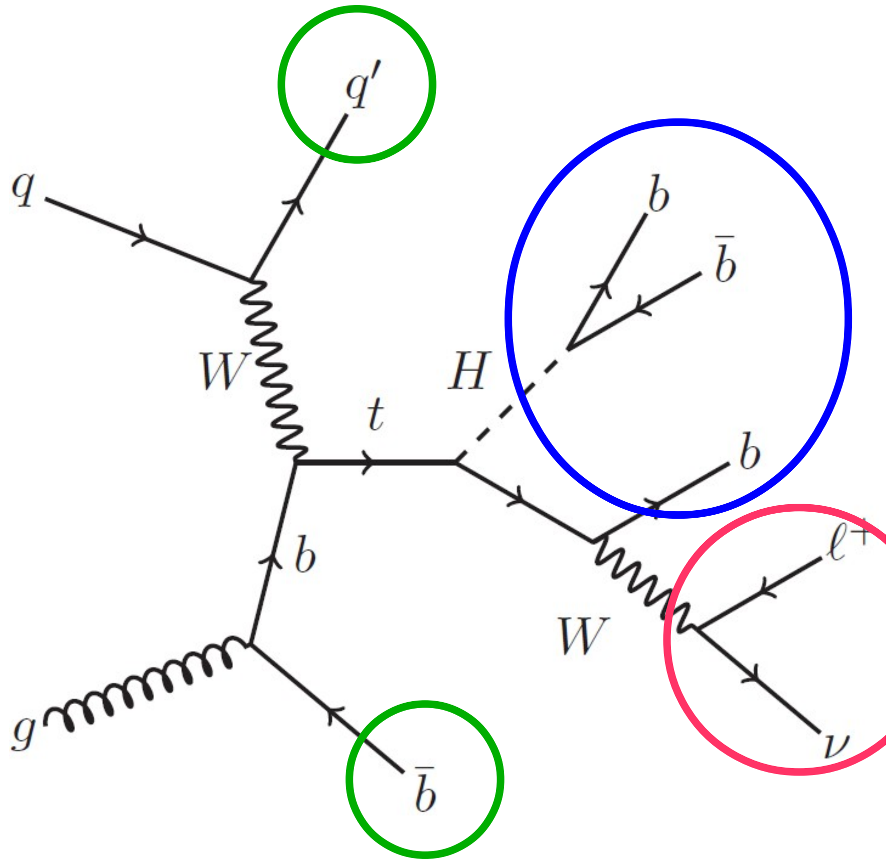
- Upper limit @ 95%CL:
  - σ BR($\gamma\gamma$) < 4.1 $\sigma_{yt=-1}$ $BR_{yt=-1}(\gamma\gamma)$
  - (Expected and observed limit coincide)
- Low-hanging fruits:
  - Add the fully-hadronic top decays
    - Way less pure selection
    - But BR($W \to q\bar{q}$)~6/9, vs BR($W \to l\nu$)~2/9
    - Experience in $t\bar{t}H$ shows that it helps in combination
  - Simultaneous extraction of $t\bar{t}H$ and tHq in two orthogonal signal regions

H→$b\bar{b}$ analysis



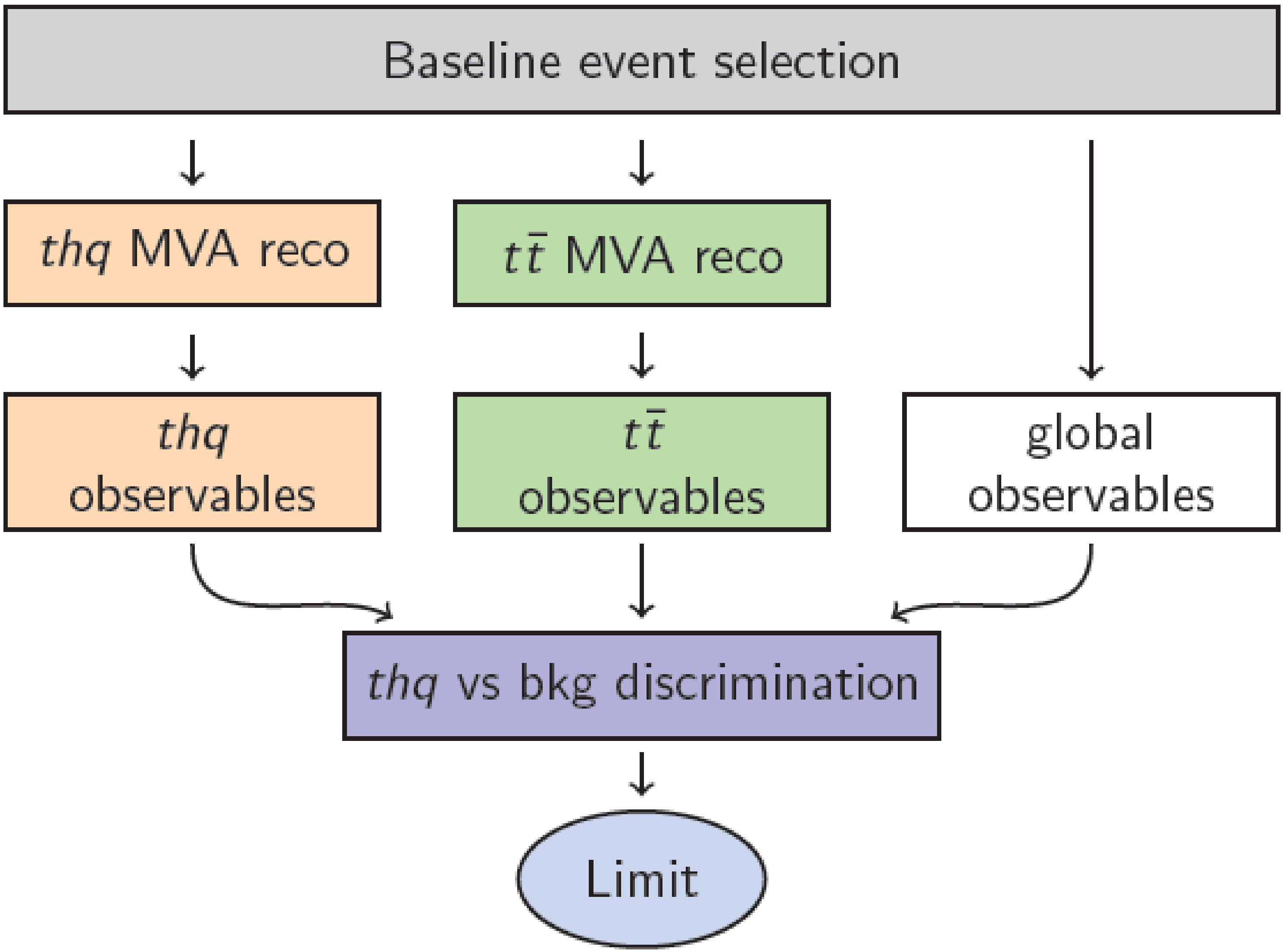
Baseline event selection

- Single-lepton trigger
- Tight muon/electron, $p_T > 26/30$ GeV
  - Veto additional loose leptons
  - MET > 35/45 GeV
- At least 4 jets with $p_T > 30$ GeV
  - + we consider all jets above 20
  - At least 3 must **pass** b tagging
  - At least 1 must **fail** b tagging
  - Jets in $|\eta|>2.5$: $p_T > 40$ GeV
- **From this point on, we classify events by number of tags (3T, 4T)**



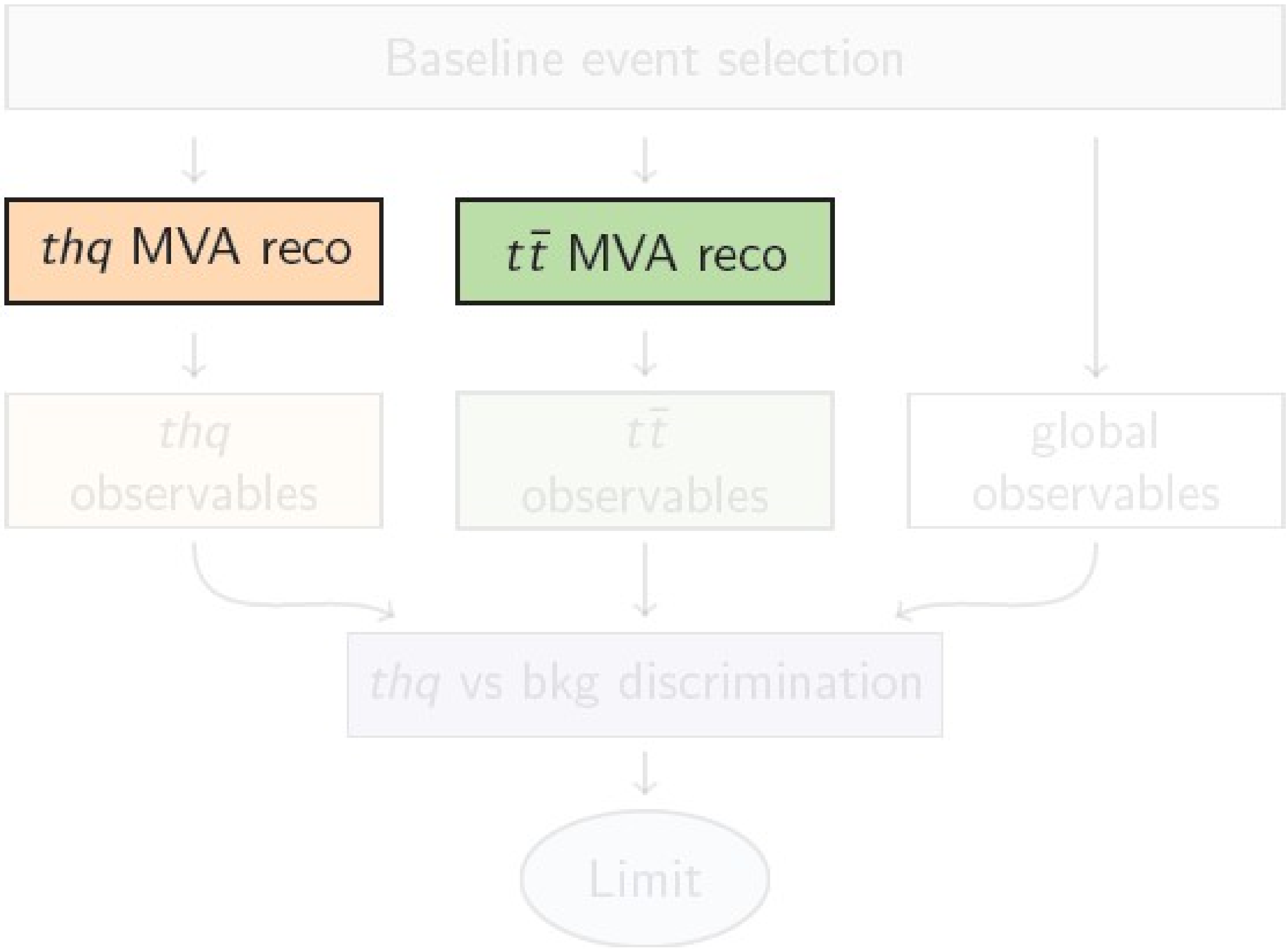
H(125) selection –
Top selection –
specific of tHq

Strategy

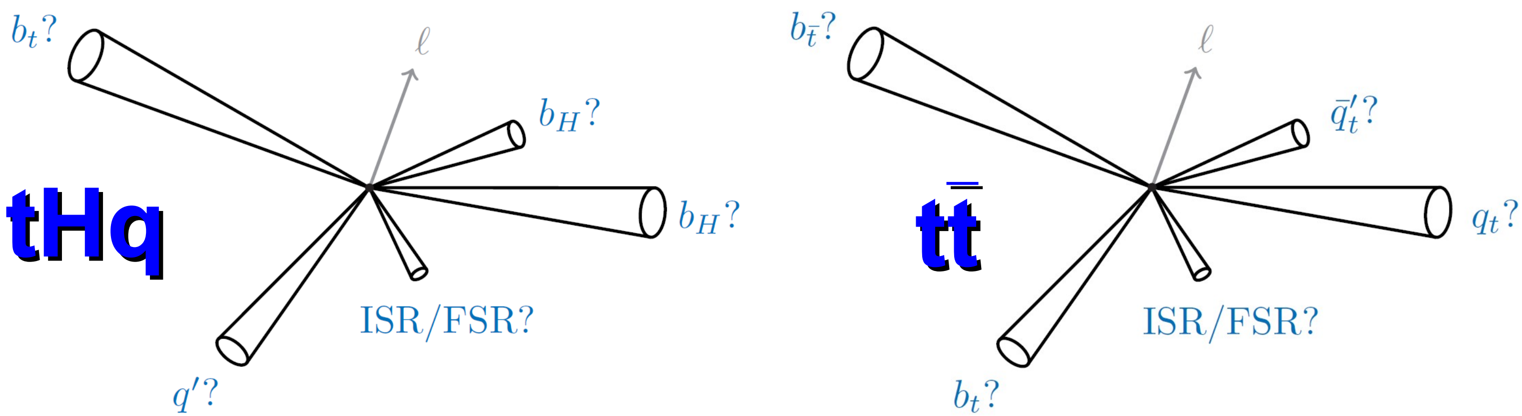


Baseline event selection
thq MVA reco
tt̄ MVA reco
thq
observables
tt̄
observables
global
observables
thq vs bkg discrimination
Limit

Event interpretations



Event interpretations

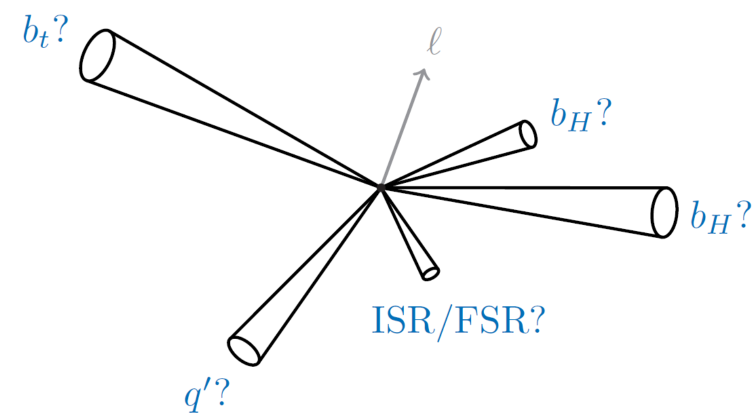


- Jets are associated to their originators (t_{lep}, t_{had}, H, q_{recoil}), in the "tHq hypothesis" and in the "$t\bar{t}$ hypothesis"
  - Goal is to be then able to define observables like "angle between H and t in the tHq hypothesis"
- NN trained to recognize the "best" combination
  - One combination per event treated as "signal" in the training, based on jet-parton matching ($\Delta R<0.3$)
  - At least one jet mismatched → "background"

Table 3: Input variables for the jet-assignment MVA under the tHq hypothesis.

Electric charge of b-quark jet from decay of top quark, multiplied by lepton's charge. The jet charge is defined as in Eq. (1) in Ref. [37], with $\kappa = 1$
ΔR between the two jets from decay of Higgs boson
ΔR between b-quark jet and W boson from decay $t \to bW$
ΔR between reconstructed top quark and Higgs boson
Pseudorapidity of recoil jet
Invariant mass of b-quark jet from decay of top quark and charged lepton
Mass of reconstructed Higgs boson
Pseudorapidity of the most forward jet from decay of H
Tranverse momentum of the softest jet from decay of H
Number of b-tagged jets among the two jets from decay of H
Boolean variable that equals 1 if the b-quark jet from decay of t is b-tagged, 0 otherwise
Relative H_T, $(p_T(t) + p_T(H))/H_T$

tHq



Validation: data/MC comparison of the NN response for random associations in the 2T control region (signal-poor, $t\bar{t}$-rich)

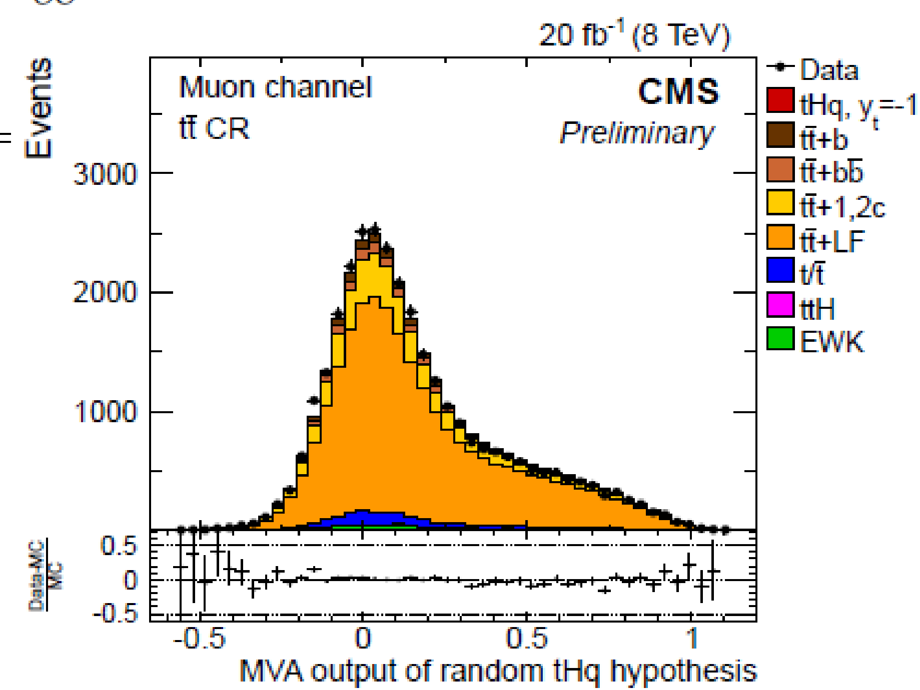
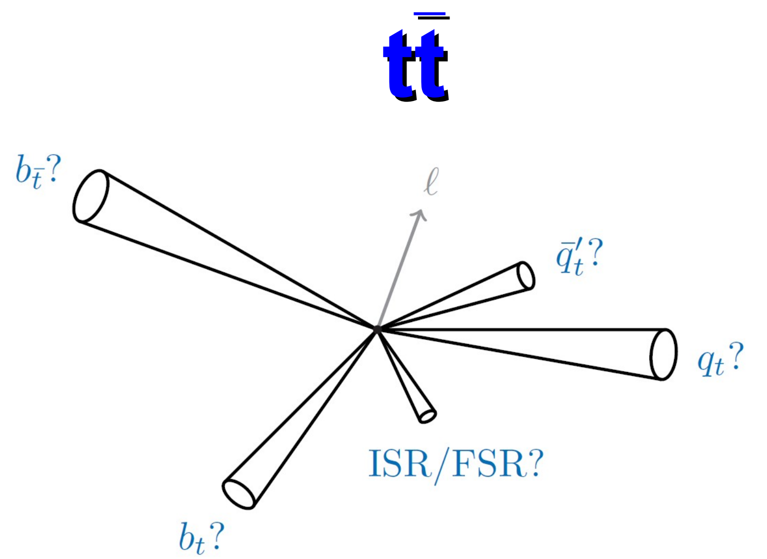
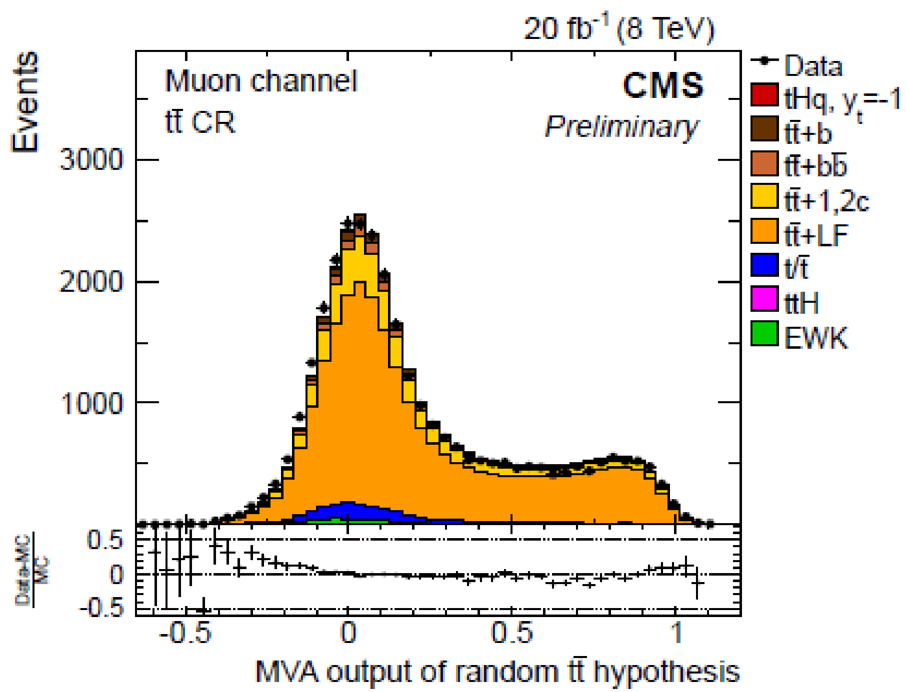


Table 4: Input variables for the jet assignment under the $t\bar{t}$ hypothesis. In the descriptions, t_{had} and t_{lep} stand for hadronically and leptonically decaying top quarks, respectively.

Difference of electric charges of b-quark jets from decays of t_{had} and t_{lep}, multiplied by lepton's charge
ΔR between the two light-flavor jets from decay of t_{had}
ΔR between b-quark jet and W boson from decay $t_{had} \to bW$
ΔR between b-quark jet and W boson from decay $t_{lep} \to bW$
Difference between masses of t_{had} and W from decay of t_{had}
Pseudorapidity of t_{had}
Invariant mass of b-quark jet from decay of t_{lep} and charged lepton
Mass of W from decay of t_{had}
Number of b-tagged jets among the two light-flavor jets from decay of t_{had}
Boolean variable that equals 1 if the b-quark jet from decay of t_{had} is b-tagged, 0 otherwise
Boolean variable that equals 1 if the b-quark jet from decay of t_{lep} is b-tagged, 0 otherwise
Transverse momentum of t_{had}
Transverse momentum of t_{lep}
Relative H_T, $(p_T(t_{had}) + p_T(t_{lep}))/H_T$
Sum of electric charges of the two light-flavor jets from decay of t_{had}, multiplied by lepton's charge



Validation: data/MC comparison of the NN response for random associations in the 2T control region (signal-poor, $t\bar{t}$-rich)



Signal-vs-background MVA

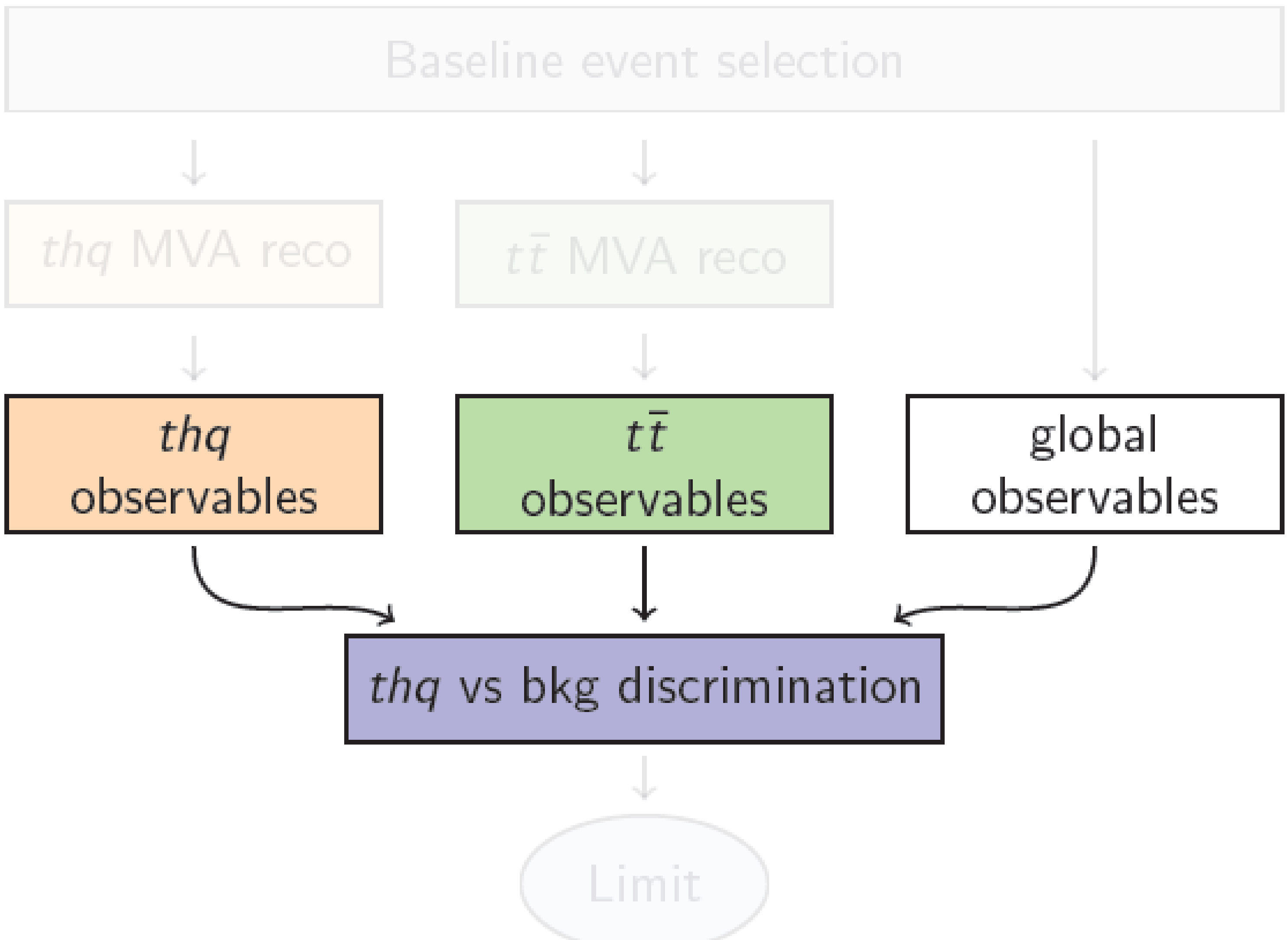
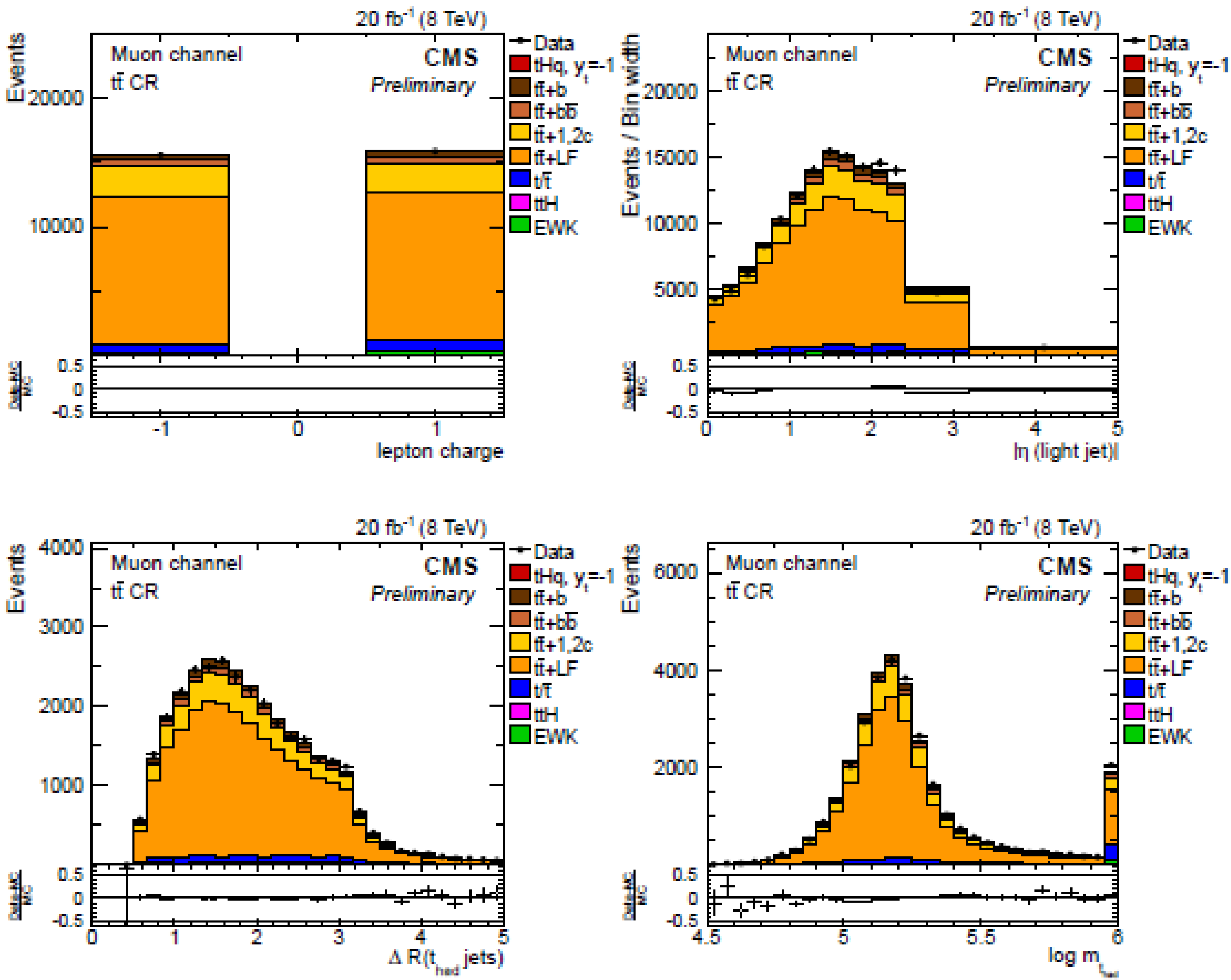


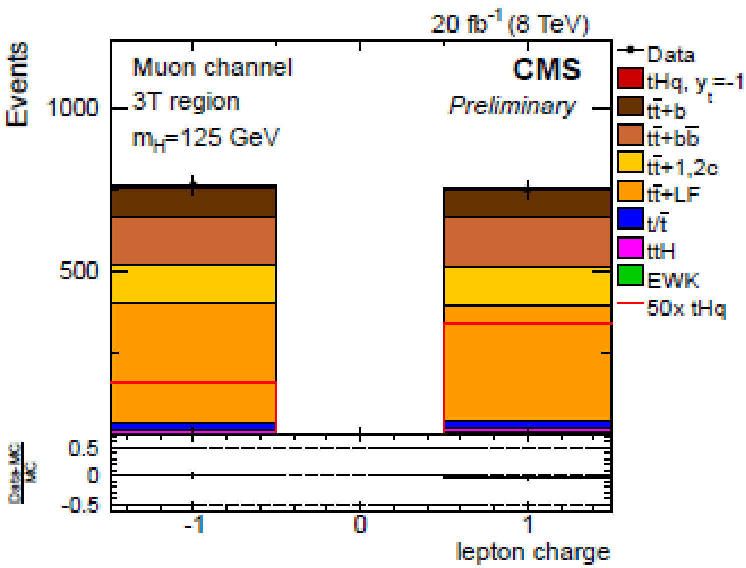
Table 5: Input variables for the classification MVA. The variables are split into three groups: global variables, variables of the jet assignment under the tHq hypotheses, variables of the jet assignment under the $t\bar{t}$ hypothesis. In the descriptions, t_{had} stands for a hadronically decaying top quark.

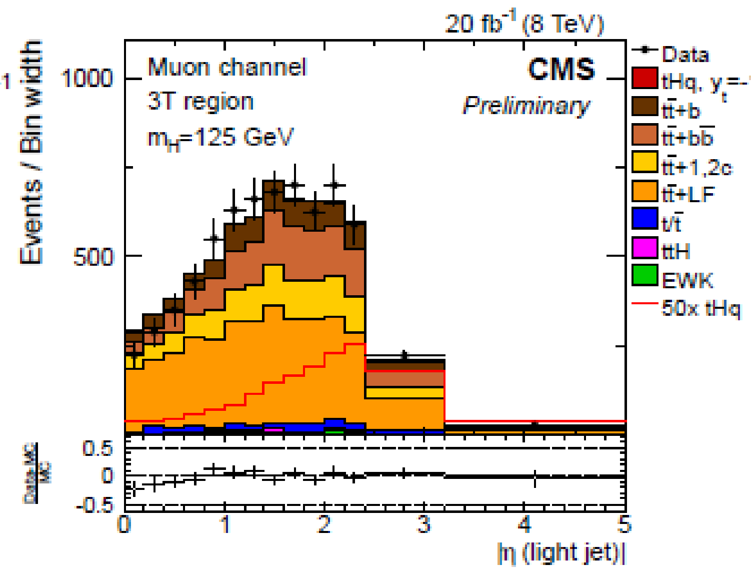
Electric charge of the lepton
Pseudorapidity of the recoil jet
Number of b-tagged jets among the two jets from the Higgs boson decay
Transverse momentum of the Higgs boson
Transverse momentum of the recoil jet
ΔR between the two light-flavor jets from the decay of t_{had}
Mass of t_{had}
Number of b-tagged jets among the two light-flavor jets from the decay of t_{had}

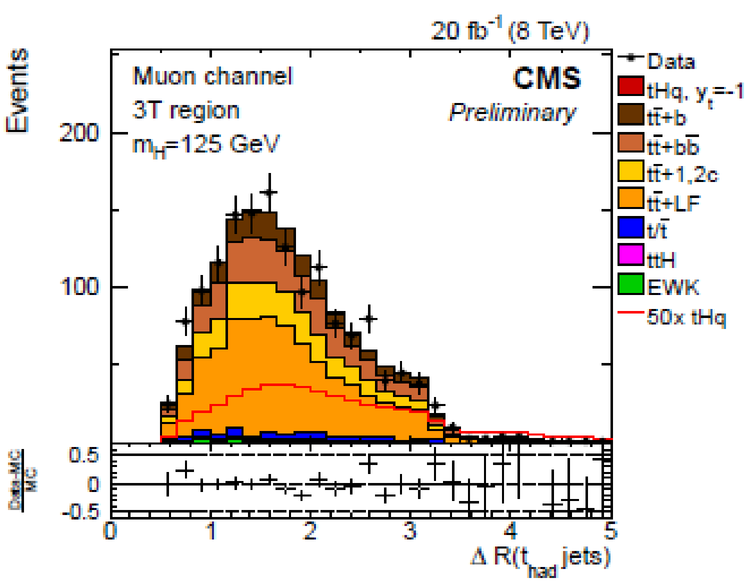
2T region

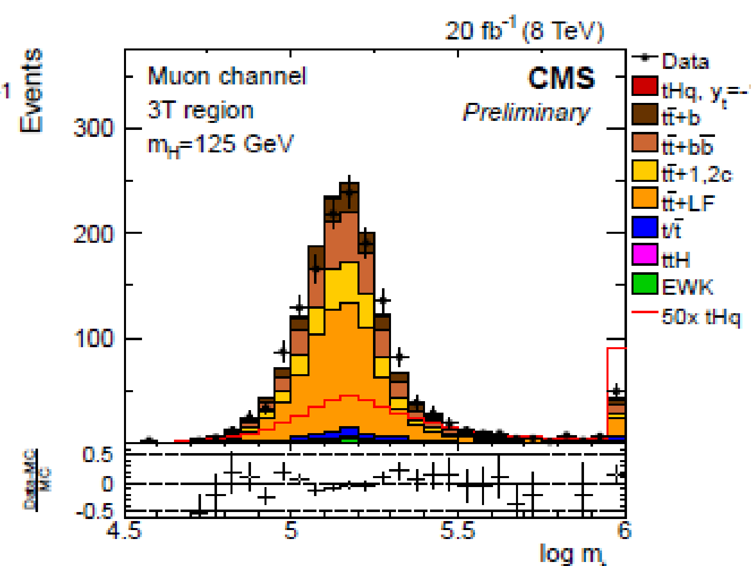

20 fb^{-1} (8 TeV)
Muon channel
$t\bar{t}$ CR
CMS
Preliminary
Data
tHq, y_t=-1
$t\bar{t}$+b
$t\bar{t}$+$b\bar{b}$
$t\bar{t}$+1,2c
$t\bar{t}$+LF
t/$\bar{t}$
ttH
EWK
Events
Events / Bin width
lepton charge
|η (light jet)|
Δ R(t_{had}, jets)
log $m_{t_{had}}$

3T region

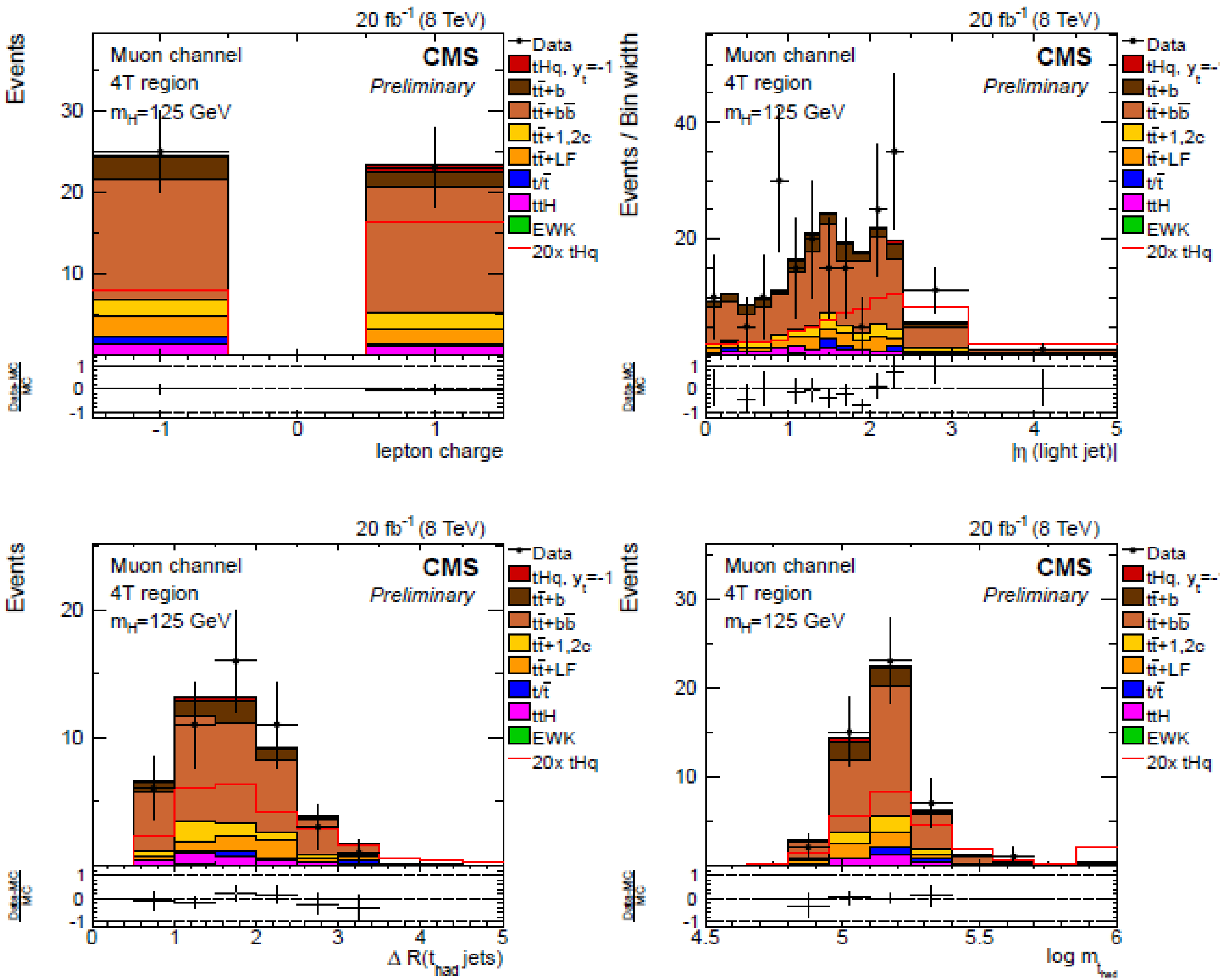

20 fb-1 (8 TeV)
Muon channel
3T region
mH=125 GeV
CMS
Preliminary
Data
tHq, yt=-1
tt̄+b
tt̄+bb̄
tt̄+1,2c
tt̄+LF
t/t̄
ttH
EWK
50x tHq
Events
Data-MC / MC
lepton charge


20 fb-1 (8 TeV)
Muon channel
3T region
mH=125 GeV
CMS
Preliminary
Data
tHq, yt=-1
tt̄+b
tt̄+bb̄
tt̄+1,2c
tt̄+LF
t/t̄
ttH
EWK
50x tHq
Events / Bin width
Data-MC / MC
|η (light jet)|

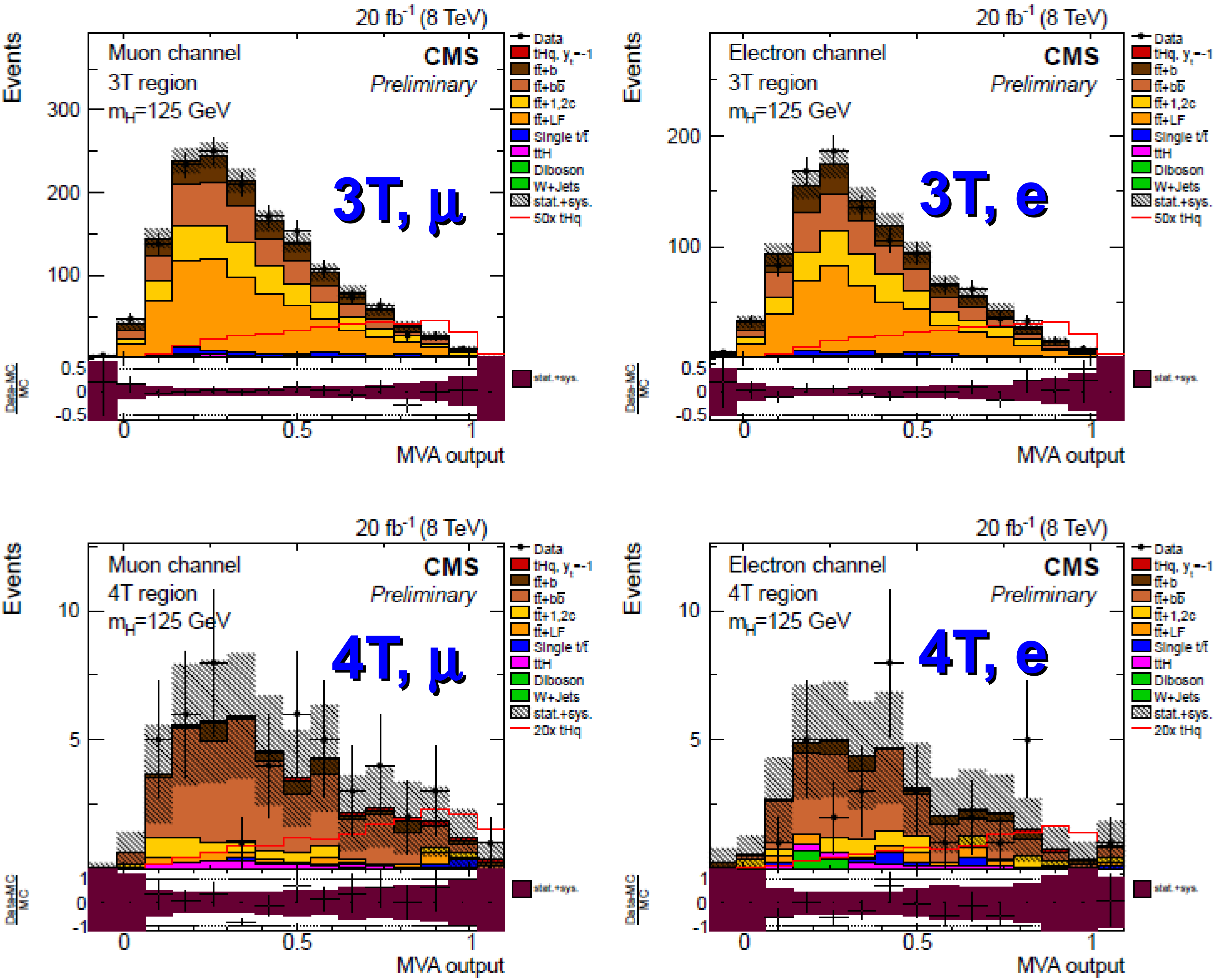

20 fb-1 (8 TeV)
Muon channel
3T region
mH=125 GeV
CMS
Preliminary
Data
tHq, yt=-1
tt̄+b
tt̄+bb̄
tt̄+1,2c
tt̄+LF
t/t̄
ttH
EWK
50x tHq
Events
Data-MC / MC
Δ R(thad jets)


20 fb-1 (8 TeV)
Muon channel
3T region
mH=125 GeV
CMS
Preliminary
Data
tHq, yt=-1
tt̄+b
tt̄+bb̄
tt̄+1,2c
tt̄+LF
t/t̄
ttH
EWK
50x tHq
Events
Data-MC / MC
log mthad

4T region

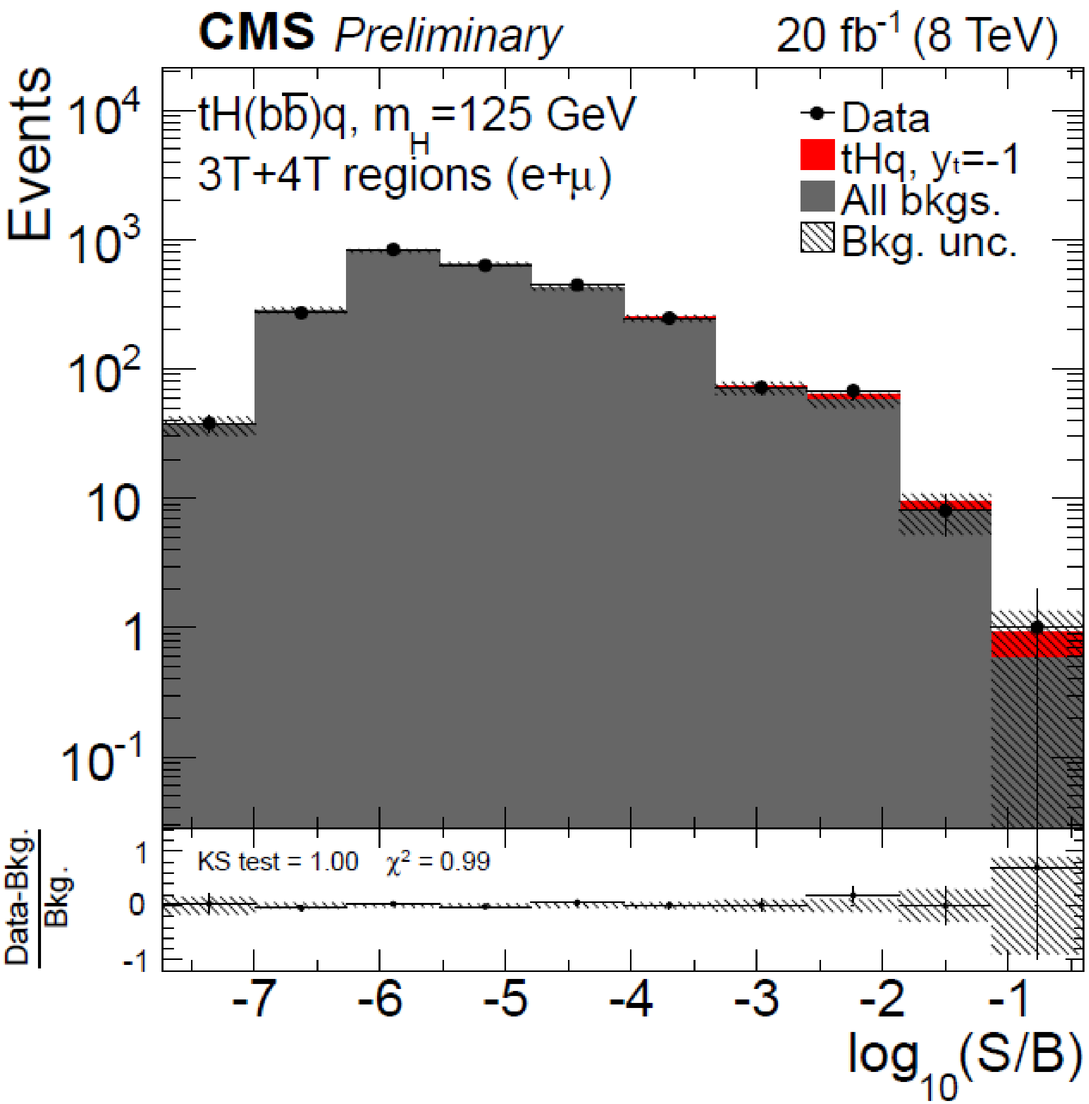


20 fb-1 (8 TeV)
Muon channel
4T region
mH=125 GeV
CMS
Preliminary
Data
tHq, yt=-1
tt+b
tt+bb
tt+1,2c
tt+LF
t/t
ttH
EWK
20x tHq
Events
Events / Bin width
lepton charge
|η (light jet)|
Δ R(thad jets)
log mthad



20 fb-1 (8 TeV)
Muon channel
3T region
mH=125 GeV
CMS
Preliminary
3T, μ
Electron channel
3T region
3T, e
Muon channel
4T region
4T, μ
Electron channel
4T region
4T, e
Data
tHq, yt=-1
tt+b
tt+bb
tt+1,2c
tt+LF
Single t/t
ttH
Diboson
W+Jets
stat.+sys.
50x tHq
20x tHq
Events
Data-MC / MC
MVA output

Result, and cross check

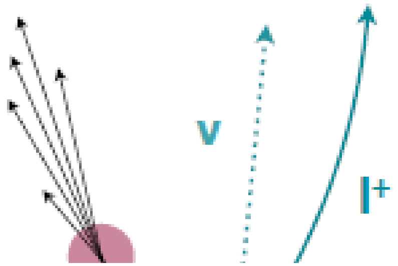


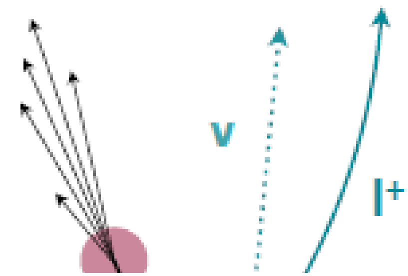
In units of $\sigma_{yt=-1}$:

	Expected	Observed
MC-driven	$\mathbf{5.14^{+2.14}_{-1.44}}$	**7.57**
Data-driven cross-check	$6.24^{+2.26}_{-1.71}$	6.95

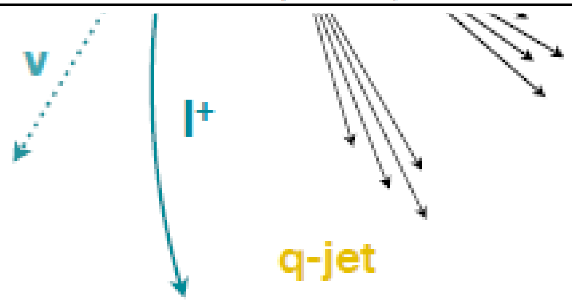
- "Data-driven" x-check makes use of 2T region and b-tag / mistag efficiencies to predict $t\bar{t}$+light jets in 3T and 4T
- Less competitive than default analysis with this dataset
- But scales better with L; to be further considered in Run 2

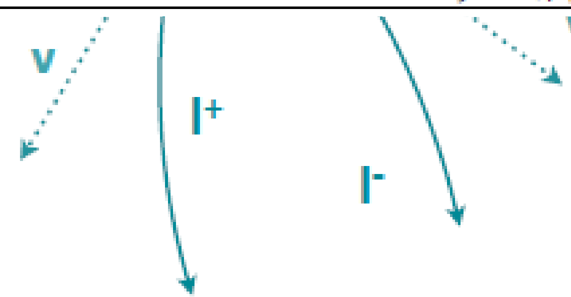
Multileptons analysis

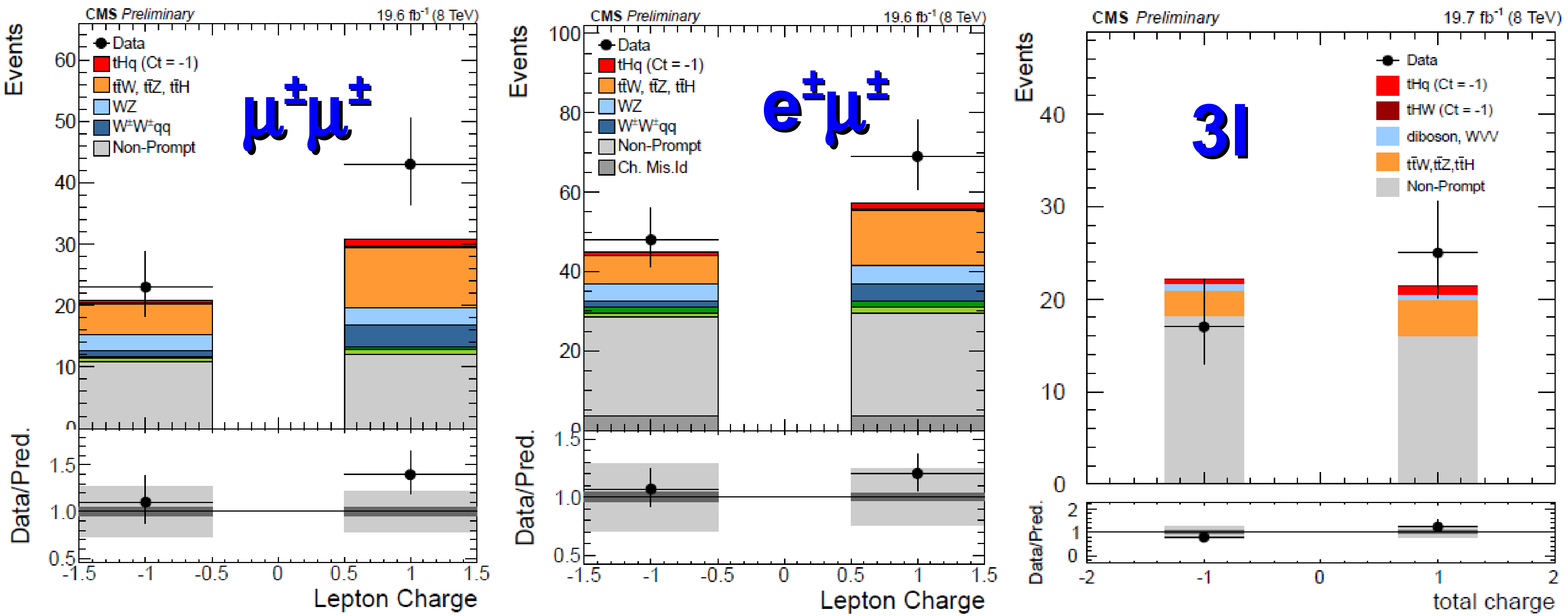




Same-sign $\ell\ell$ channel	**$\ell\ell\ell$ channel**
Two leptons of equal charge, $p_T > 20\,\mathrm{GeV}$	Three leptons with $p_T > 20/10/10$ GeV
No additional leptons with lepton MVA $>$ 0.35	No additional tight leptons
$m_{ll} > 20\,\mathrm{GeV}$	$m_{ll} > 20\,\mathrm{GeV}$
No identified hadronically-decaying τ leptons	Z-veto: $\lvert m_{ll} - m_Z \rvert > 15\,\mathrm{GeV}$
At least one central jet ($\lvert\eta\rvert < 1.0$)	$E_T^{\mathrm{miss}} > 30$ GeV
At least one central jet tagged as CSV-L	Exactly one jet with tagged as CSV-M
At least one forward jet ($\lvert\eta\rvert < 1.0$)	At least one forward jet ($\lvert\eta\rvert > 1.5$)

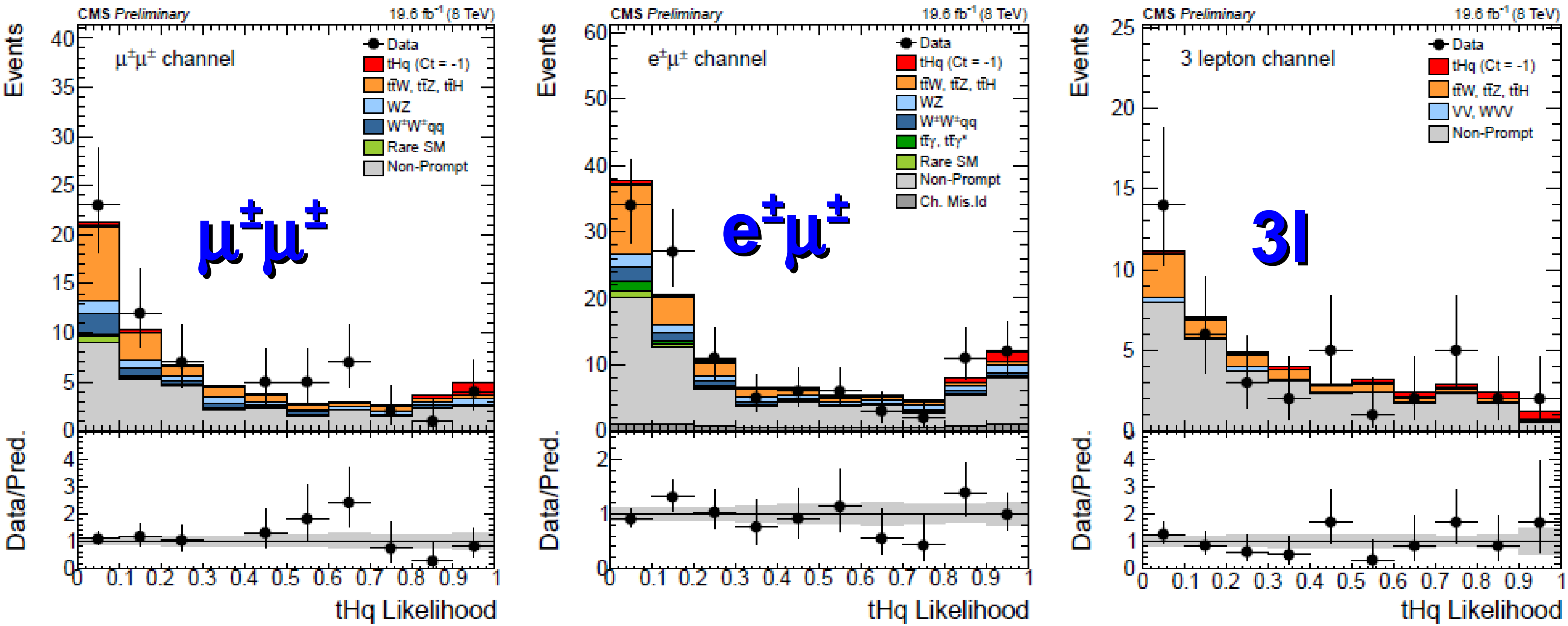






- Non-prompt & fake leptons allow $t\bar{t}$ to contaminate all three signal regions
- Estimated from tight-to-loose ID ratio in control regions
- Charge confusion probability is estimated by $Z \rightarrow ee$
- Finally, event properties are combined in a likelihood

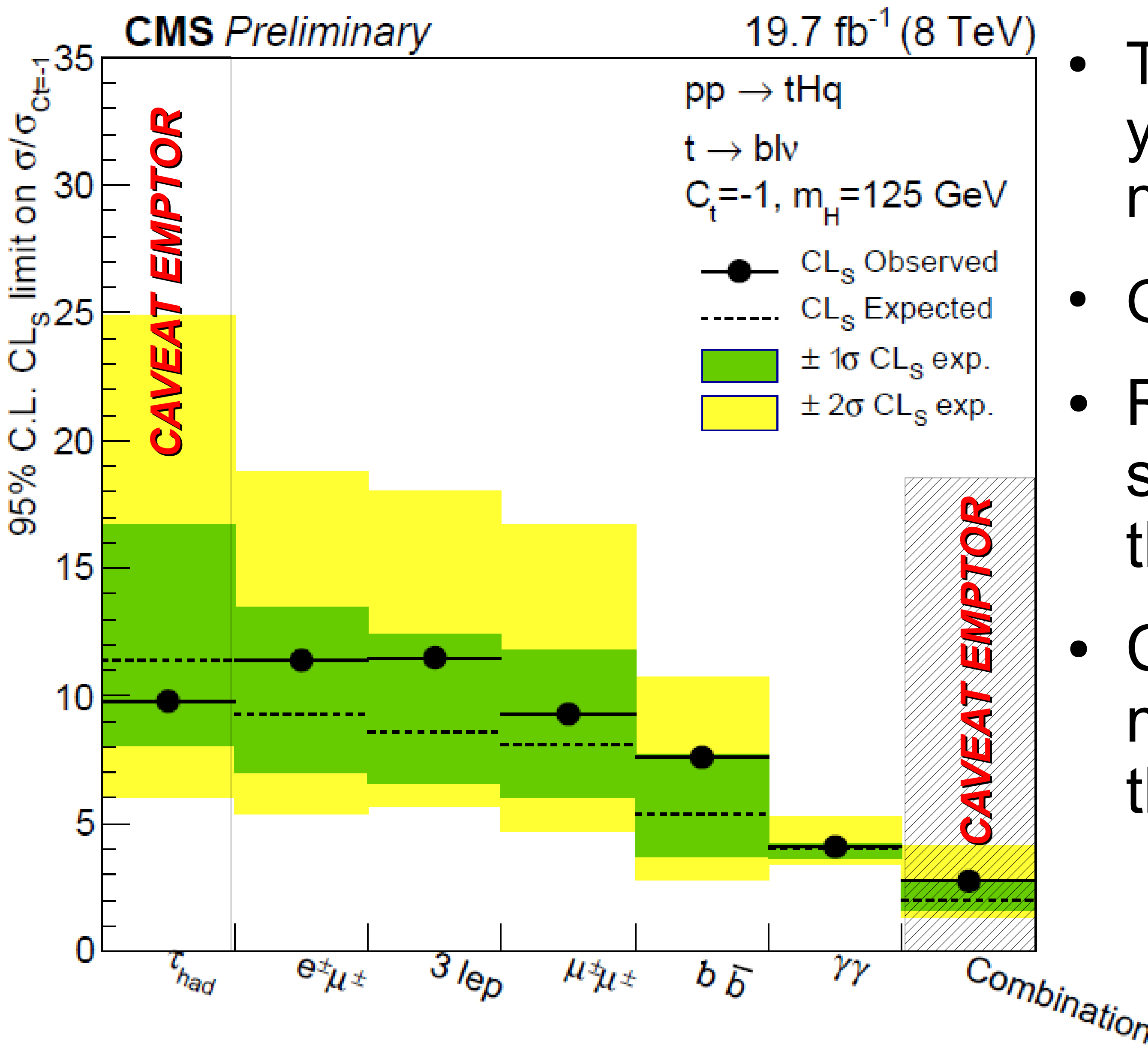
Results



In units of $\sigma_{yt=-1}$:

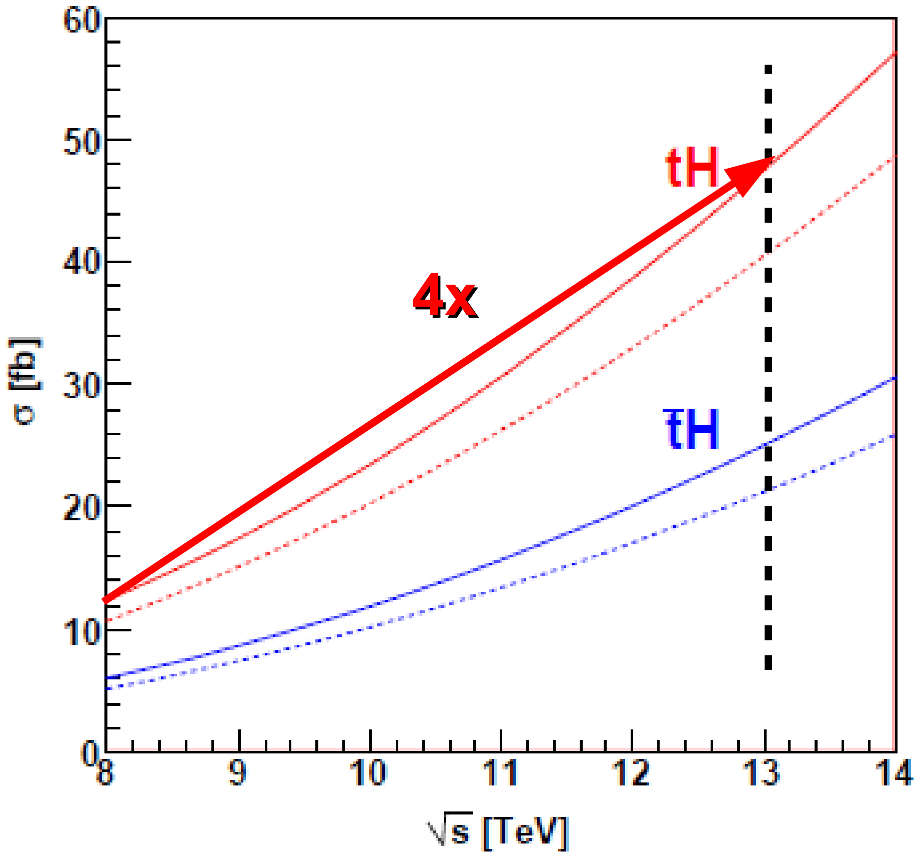
Channel	Observed	Expected	68% prob. band	95% prob. band
SS $\mu\mu$	9.3	8.1	[6.0, 11.8]	[4.7, 16.7]
SS $e\mu$	11.4	9.3	[7.0, 13.5]	[5.4, 18.8]
3ℓ	11.5	8.6	[6.6, 12.4]	[5.7, 18.0]
combined	6.7	5.0	[3.6, 7.1]	[2.9, 10.3]

Combination



- Tau channel not public yet, but conclusions will not change:
- Golden channel is $\gamma\gamma$
- Rather uncorrelated systematics across the channels
- Combination is twice more sensitive than the best channel

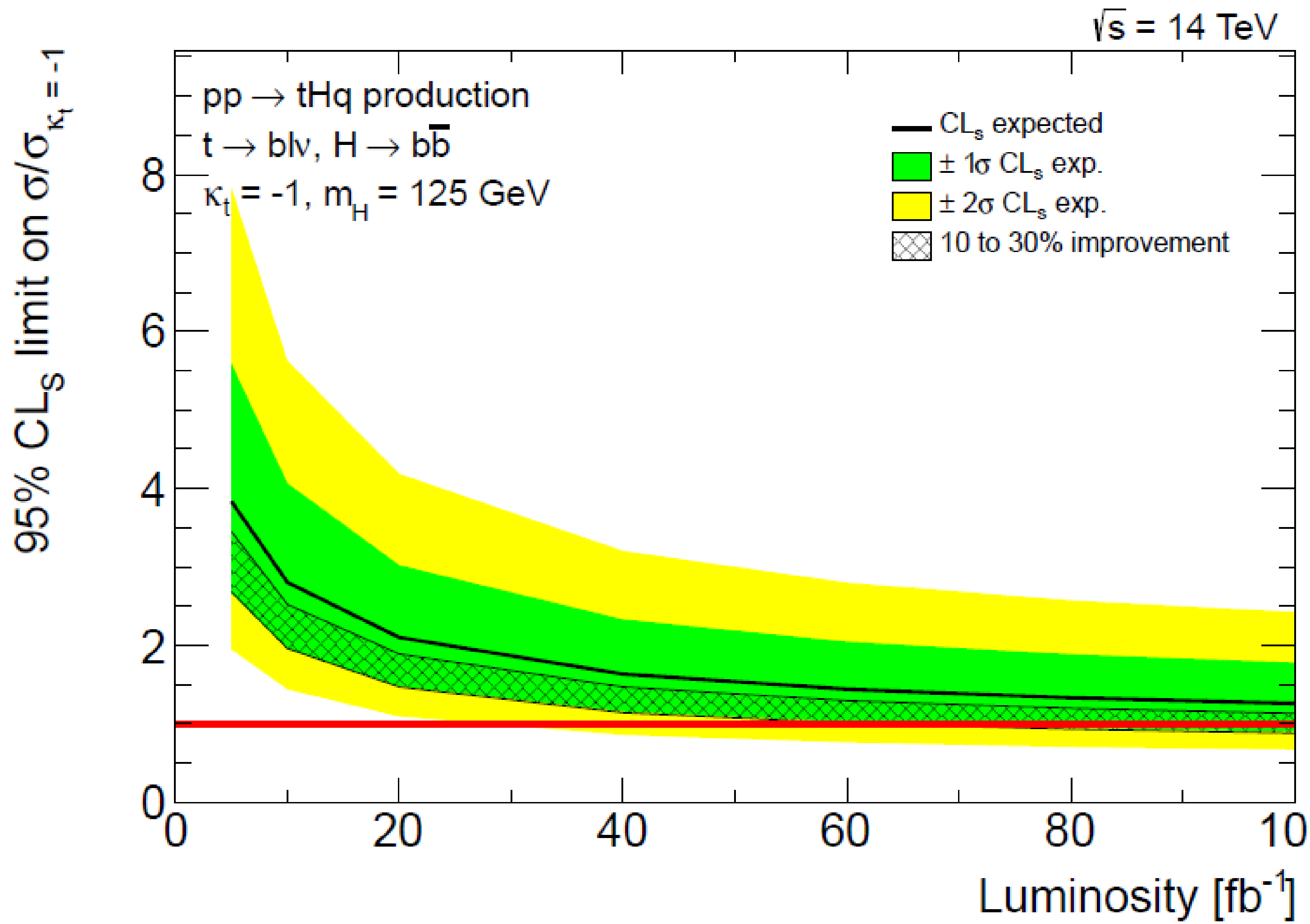
Part III: Prospects



From Campbell, Ellis, Roentsch, arXiv:1302.3856, Phys. Rev. D 87, 114006 (2013)

Slope doesn't depend on y_t; scaling for $t\bar{t}$ is 3.3x

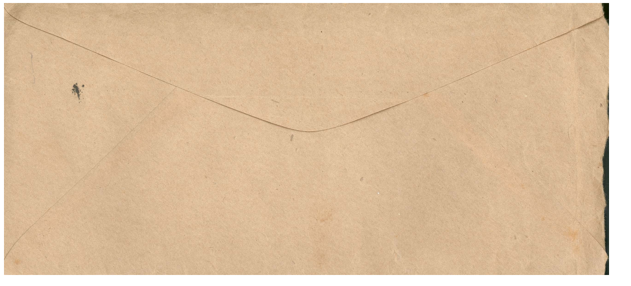
Forecast for $H \rightarrow b\bar{b}$



This analysis, in its current form, hits the *Systematics Wall* soon; significant breakthroughs are needed

Forecast for H→γγ

- We have not run real projections yet, but this channel is in a regime where **back-of-the-envelope** calculations work nicely
  - Analysis limited by statistics
  - Background extremely tiny
- So, **buy an envelope** and assume:

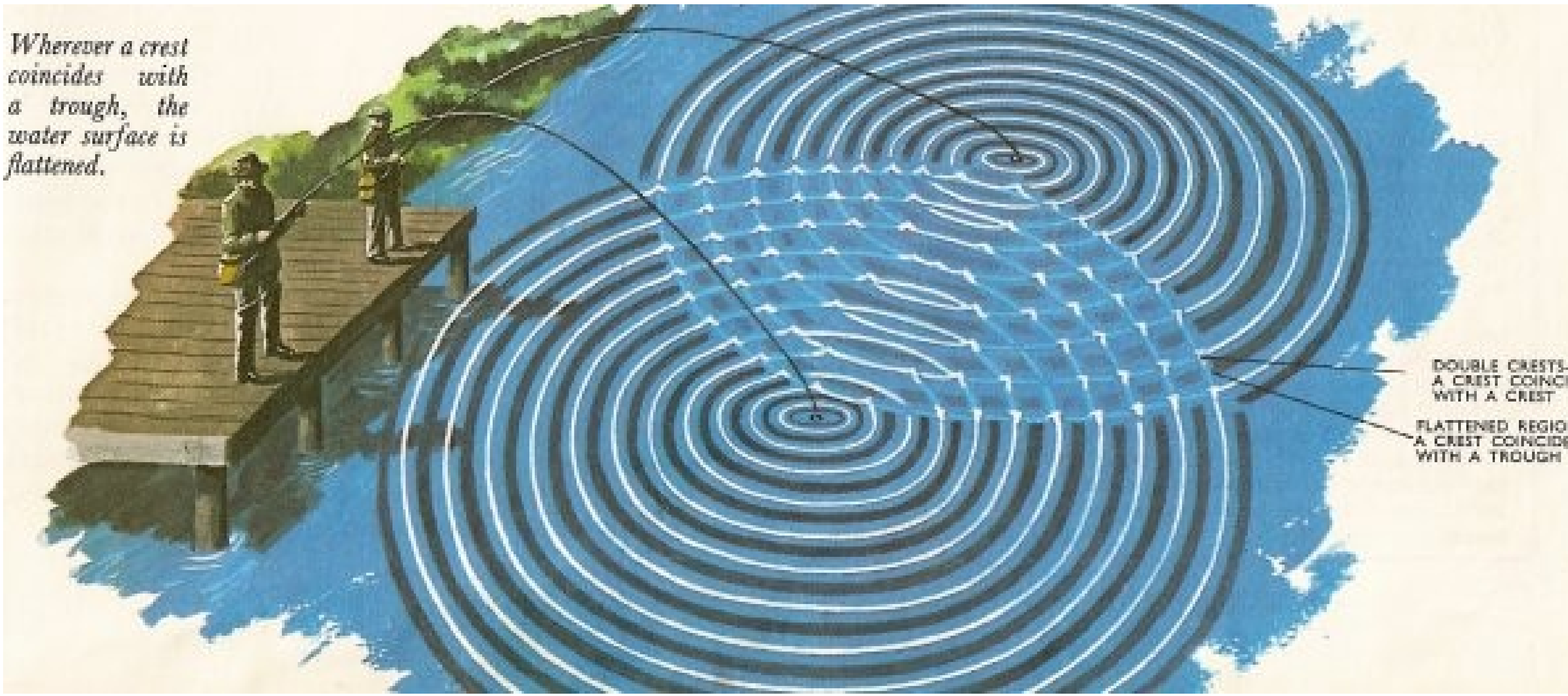


  - No systematics
  - No backgrounds
  - Signal cross section multiplied by 4 with respect to 8 TeV
  - No event observed in data
- Simple Poisson statistics tells:
  - Same upper limit expected with only 5/fb at 13 TeV
  - Sensitive to *"SM with y_t=-1"* with **20/fb**, i.e., early 2016

Conclusions

- Study of tHq production can provide:
  - An unambiguous measurement of the relative sign of Htt and HWW couplings
  - A unique access to their relative phase
  - A test of other models along the way, down to SM
- Started searching for tHq
  - Four channels explored
  - All very challenging, and quite complementary
  - Upper limit still twice our strawman scenario's σ
  - But Run-II is near and we expect to start biting the interesting parameter space very soon

Thanks for your attention

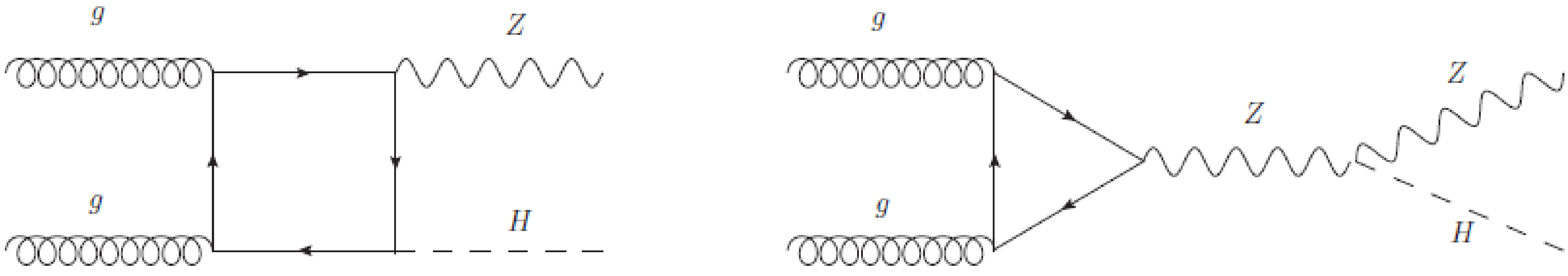


Acknowledgements: Emidio Gabrielli, Fabio Maltoni, Andrea Thamm for patiently explaining their papers to me; my outstanding collaborators from 7 CMS institutes; too many to list in the narrow margin of this slide

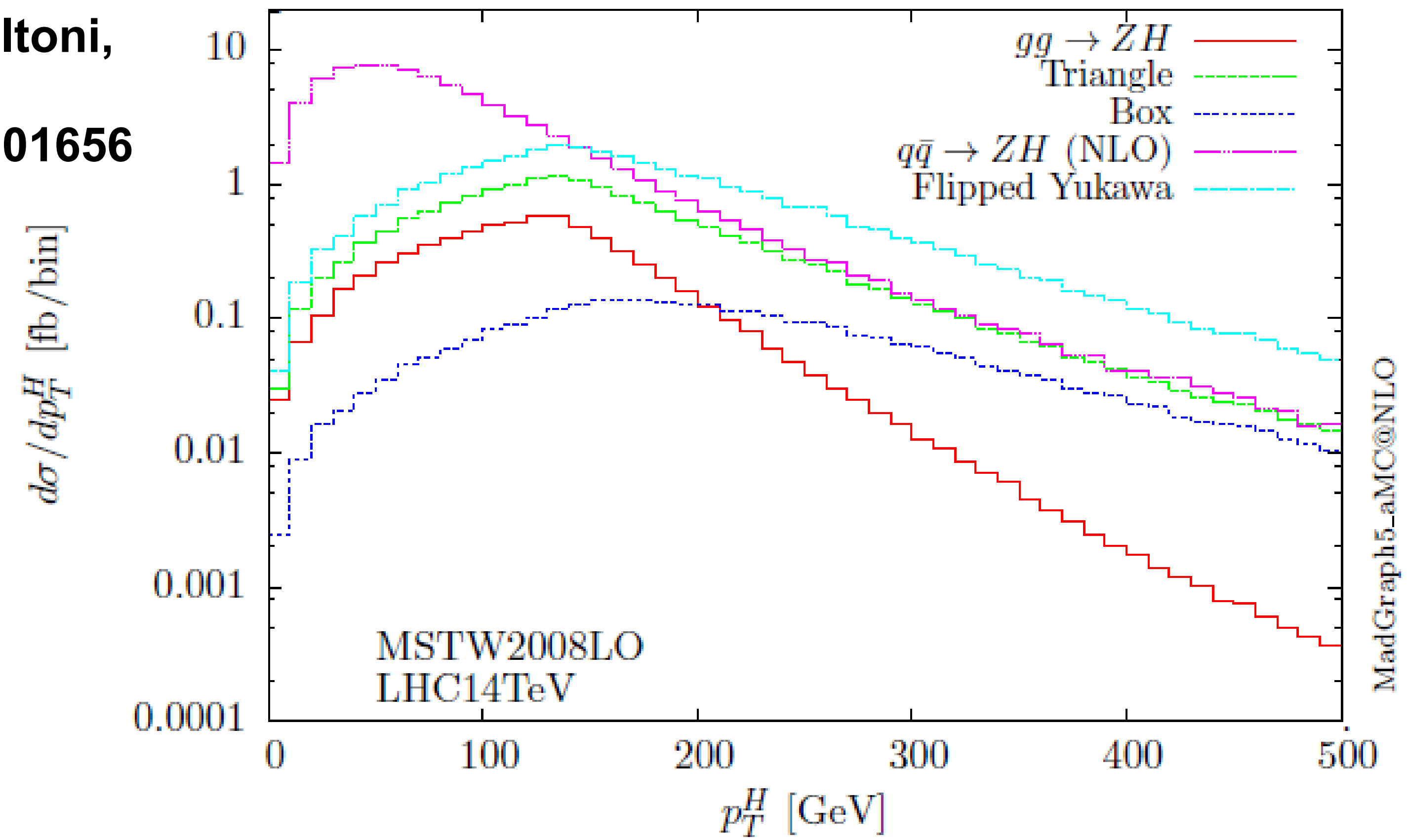
Pheno bibliography (partial!)

- Biswas, Gabrielli, Mele, "Single top and Higgs associated production as a probe of the Htt coupling sign at the LHC", arXiv:1211.0499, JHEP 01 (2013) 088
- Farina, Grojean, Maltoni, Salvioni, Thamm, "Lifting degeneracies in Higgs couplings using single top production in association with a Higgs boson", arXiv:1211.3736
- Biswas, Gabrielli, Margaroli, Mele, "Direct constraints on the top-Higgs coupling from the 8 TeV LHC data", arXiv:1304.1822, JHEP 07 (2013) 073
- El Hedri, Fox, Wacker, "Exploring the dark side of the top Yukawa", arXiv:1311.6488
- Englert, Re, "Bounding the top Yukawa with Higgs-associated single-top production", arXiv:1402.0445, Phys. Rev. D 89, 073020 (2014)
- Ellis, Hwang, Sakurai, Takeuchi, "Disentangling Higgs-Top Couplings in Associated Production", arXiv:1312.5736
- Chang, Cheung, Lee, Lu, "Probing the Top-Yukawa Coupling in Associated Higgs production with a Single Top Quark", arXiv:1403.2053
- Greljo, Kamenik, Kopp, "Disentangling Flavor Violation in the Top-Higgs Sector at the LHC", arXiv:1404.1278
- Kobakhidze, Wu, Yue, "Anomalous Top-Higgs Couplings and Top Polarisation in Single Top and Higgs Associated Production at the LHC", arXiv:1406.1961
- Demartin, Maltoni, Mawatari, Zaro, "Higgs production in association with a single top quark at the LHC", arXiv:1504.00611

gg→ZH



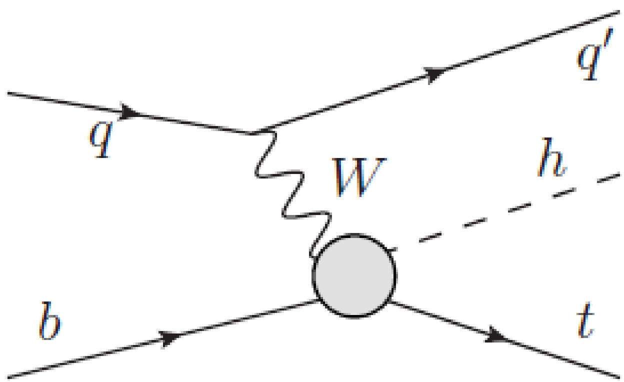
Hespel, Maltoni, Vryonidou, arXiv:1503.01656

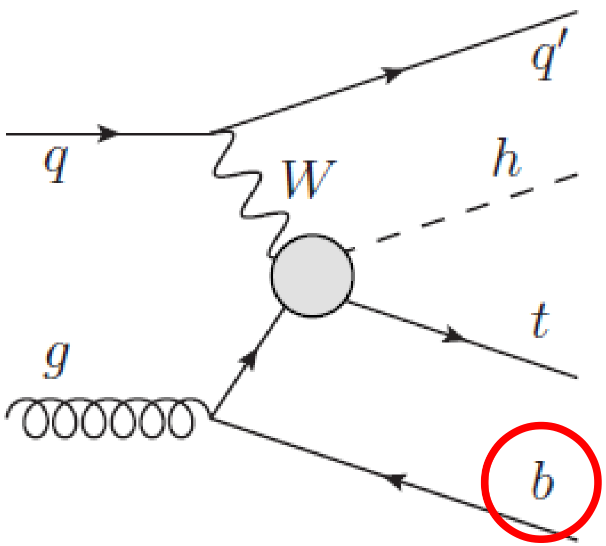


Beyond the "binary analysis": limits as function of a parameter?

- Originally thought about limits as function of the relative phase between k_V and k_F
  - It makes sense if CP is not assumed
- Suggested to provide limits as function of k_F/k_V
  - Technically: by reweighting the events, rather than by the brute force approach of generating many MCs; there is even a routine in aMC@NLO that does that (although at LO)
  - Of course one should validate the reweighting versus the brute force method; e.g. one could just generate the y_t=-1 and +1 samples, reweight the -1 MC to the +1 case and compare to the +1 MC

An important detail





The b must be created with an anti-b. This is usually soft and almost collinear with the beam, but sometimes it enters acceptance:

	$\sigma(pp \to thj)$ [fb]		$\sigma(pp \to thjb)$ [fb]	
	$c_F = 1$	$c_F = -1$	$c_F = 1$	$c_F = -1$
8 TeV	17.4	252.7	5.4	79.2
14 TeV	80.4	1042	26.9	363.5

Data-driven model for $t\bar{t}$, $H \rightarrow b\bar{b}$

- MC modeling of $t\bar{t}$ in signal regions carries large uncertainties (m_F, m_R, JES) that swamp the signal
- Data-driven model has a different set of uncertainties
- We use 2T region and the known b-tagging efficiencies
- Event weights P_3/P_2 and P_4/P_2 calculated from:

$$\mathcal{P}_m = \sum_{\text{comb}} \prod_{i=1}^{m} \epsilon(p_i, f_i) \cdot \prod_{j=m+1}^{n} (1 - \epsilon(p_j, f_j))$$

- This is the probability that an event with *n* jets with momentum p_i and flavour f_i has *m* of them b-tagged
- Here $\varepsilon(p,f)$ is the b-tagging efficiency and the sum is taken over all the possible ways to choose *m* tagged jets

Data-driven model for $t\bar{t}$, $H \rightarrow b\bar{b}$