Higgs boson production in association with a top quark pair: recent results from CMS

Joosep Pata
supervised by Prof. G. Dissertori,
G. Kasieczka and L. Bianchini

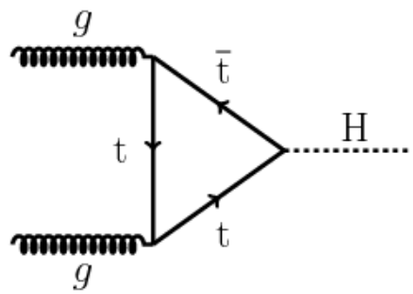
HiggsTools Second Annual Meeting
April 15, 2016

Outline

- What makes ttH(bb) interesting?
- What is the status and present difficulties?
- Closer look at a recently-published ttH(bb) search from the trenches
- Personal view, not for CMS
- Using only public information from HIG-16-004, BTV-15-001 where I have contributed

ttH→bb motivation

- Need direct experimental evidence for Higgs coupling to up-type fermions



- top mass → coupling near unity. Top quark is a Higgs probe!

- goal: verify top pair associated Higgs (ttH) production in Run II

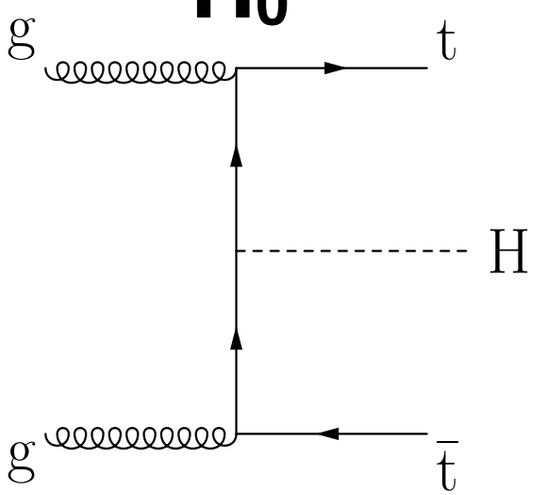
- measure **μ = σ(tt+H) / σ(tt+H in SM)**

- ttH cross-section crucial input for Higgs coupling fits

Challenges

- large background: mainly tt+jets production
  - σ(tt+jets) = **830 pb** vs. σ(tt+Hbb) = **0.3 pb**
- **tt+bb**: irreducible, multiscale process, difficulties in modelling, combinatorial self-background for H(bb)

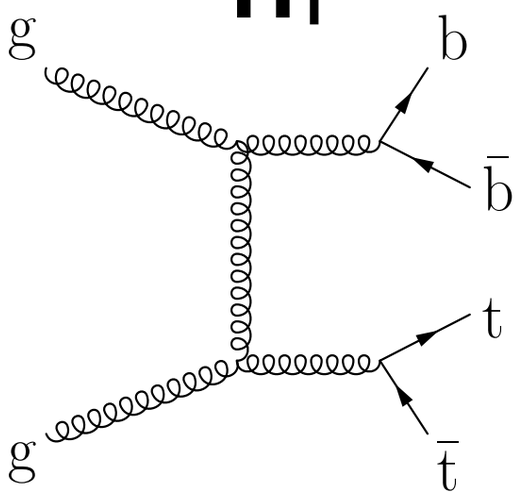
H0



signal

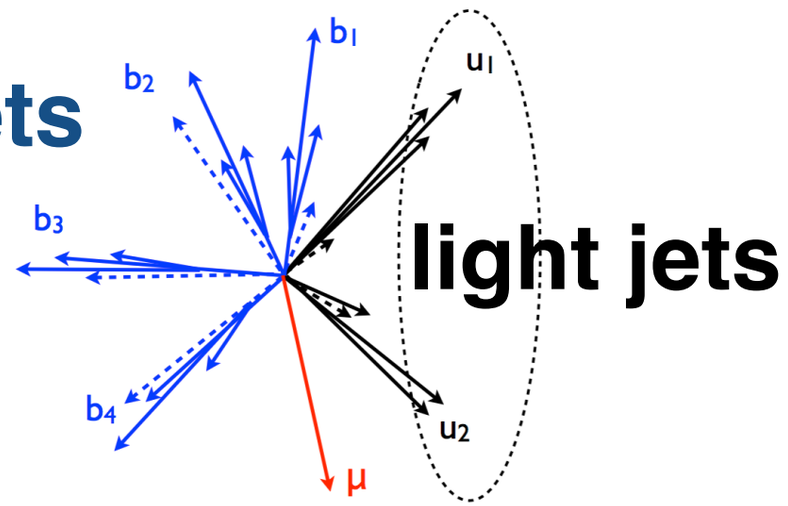
or

H1

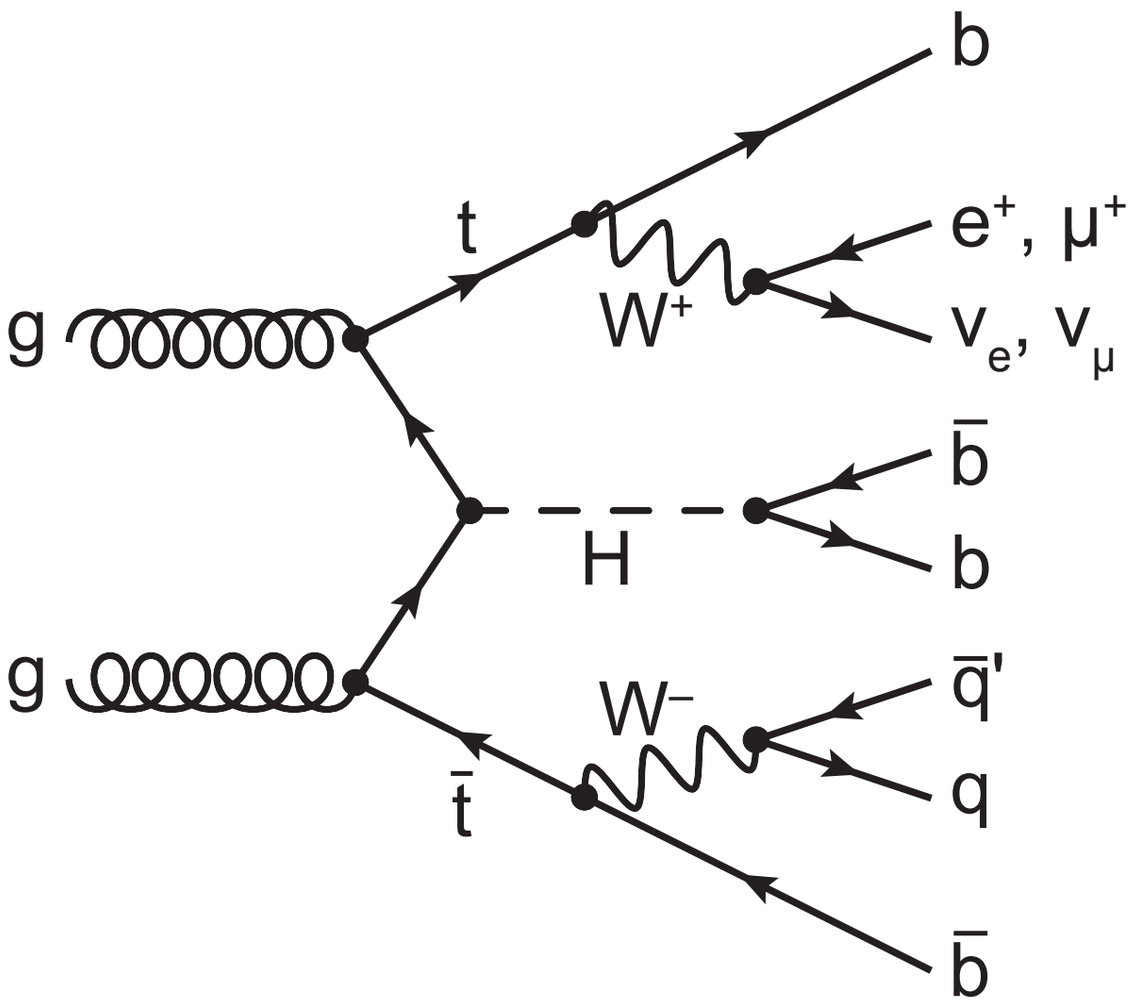


background

both have the same final state

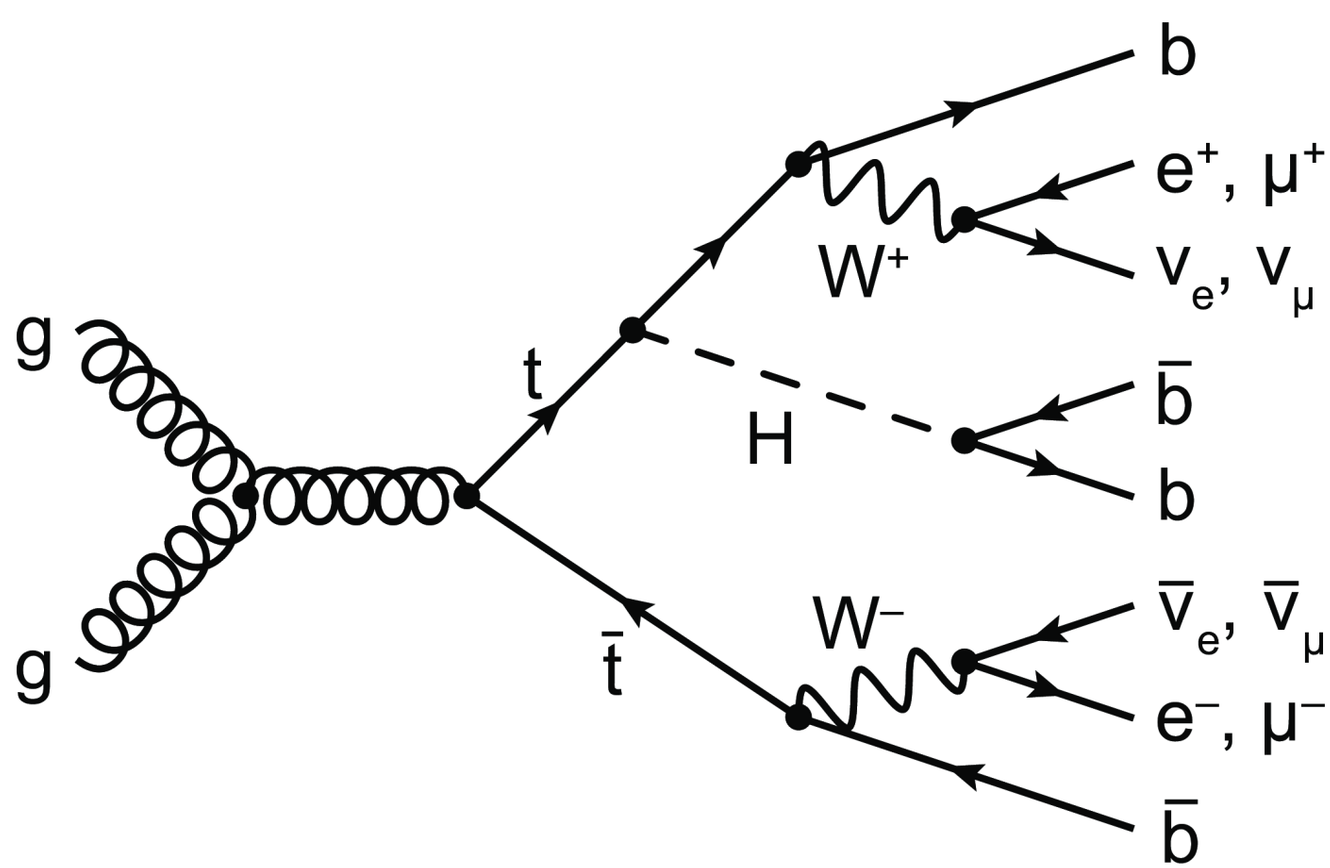


semileptonic



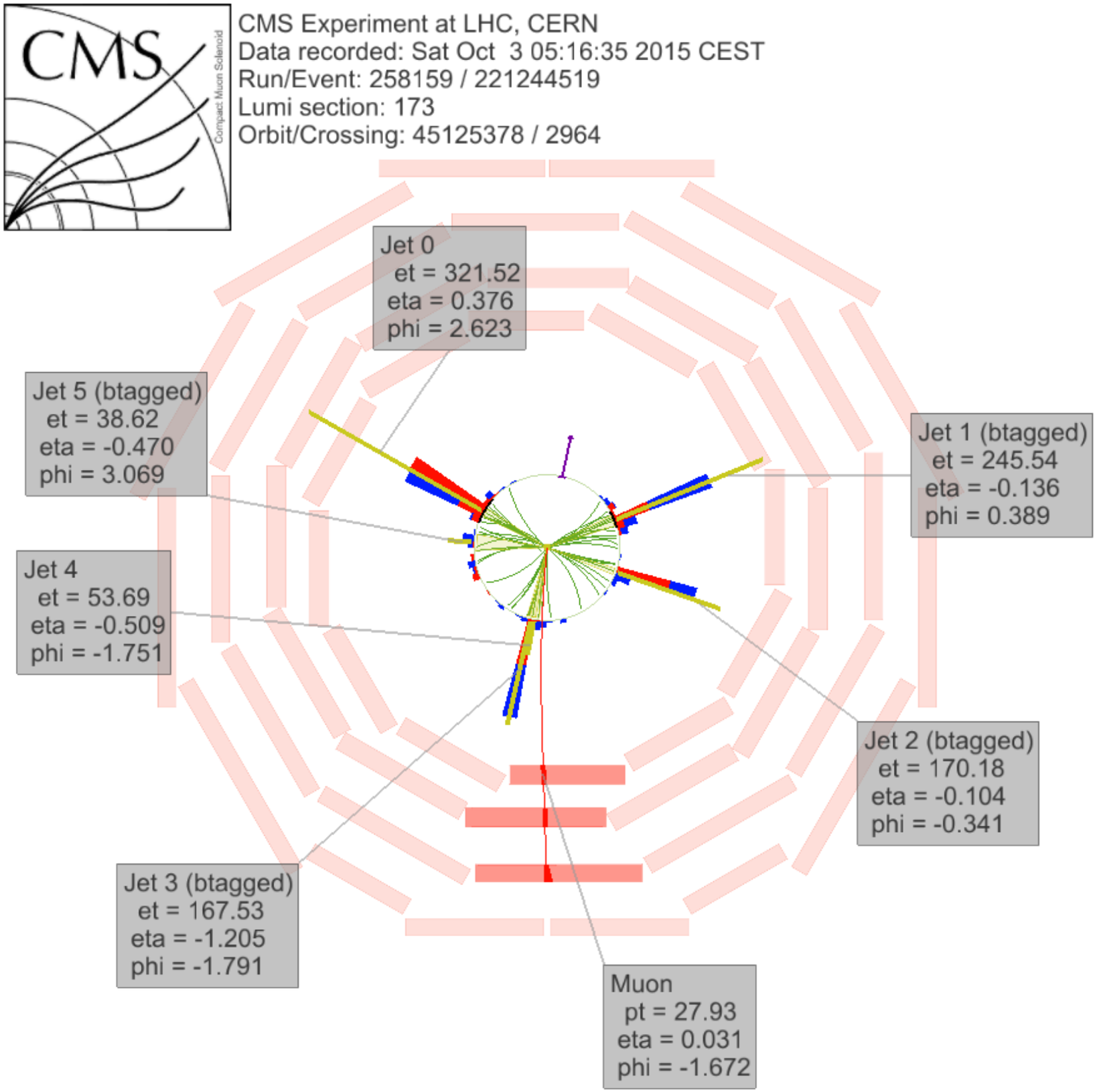
1 lepton, 6 jets, 4 b-tags

dileptonic



2 leptons, 4 jets, 4 b-tags

complicated final state with several b-jets, leptons, missing energy, combinatorial confusion, no clear resonant peak.


CMS Experiment at LHC, CERN
Data recorded: Sat Oct 3 05:16:35 2015 CEST
Run/Event: 258159 / 221244519
Lumi section: 173
Orbit/Crossing: 45125378 / 2964
Jet 0
et = 321.52
eta = 0.376
phi = 2.623
Jet 5 (btagged)
et = 38.62
eta = -0.470
phi = 3.069
Jet 1 (btagged)
et = 245.54
eta = -0.136
phi = 0.389
Jet 4
et = 53.69
eta = -0.509
phi = -1.751
Jet 2 (btagged)
et = 170.18
eta = -0.104
phi = -0.341
Jet 3 (btagged)
et = 167.53
eta = -1.205
phi = -1.791
Muon
pt = 27.93
eta = 0.031
phi = -1.672

ETH Institute for Particle Physics

LHC analyses

higgstools

Run I analyses including the H → bb decay

Experiment	Analysis	Limit	reference
CMS, 7+8 TeV	H → bb+$\tau\tau$+$\gamma\gamma$, BDT	3.4 (3.7)	HIG-13-019
CMS, 8 TeV	H → bb, MEM	4.2 (3.3)	EPJC 75 (2015) 251
ATLAS, 8 TeV	H → bb, MEM+BDT	3.4 (2.2)	EPJC 75 (2015) 349
CMS, 7+8 TeV	ttH combined	4.5 (1.7)	JHEP 09 (2014) 087
ATLAS, 7+8 TeV	ttH fully hadronic	3.1 (1.4)	Moriond16

Analysis strategy well established by both experiments, Run II and 2016 has discovery potential.

ETH Institute for Particle Physics

higgstools

Run II ttH(bb)

Search for $t\bar{t}H$ production in the $H \to b\bar{b}$ decay channel with $\sqrt{s} = 13$ TeV pp collisions at the CMS experiment

The CMS Collaboration

New in Run II

- combined discriminator approach
- boosted topologies included
- modelling with NLO generators (POWHEG)
- New tt+heavy-flavour treatment

- Very collaborative analysis, 40+ people, 5+ different institutes
- Presented by CMS at Moriond 2016

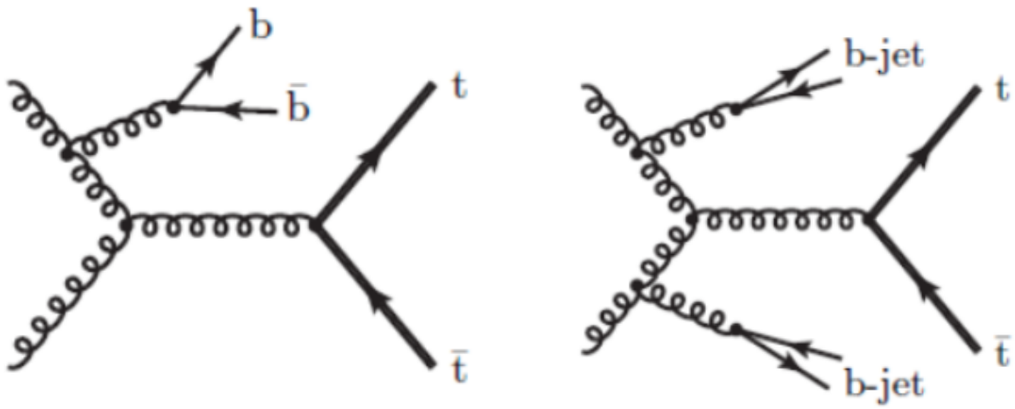
Analysis strategy

- **trigger**: rely on leptonic decay of top for triggering
- **selection**: events compatible with tt+H lepton+jets or dilepton channel and H→bb decay
- **categorization**: based on jet and b-tag multiplicity, presence of boosted top/higgs
- **shape analysis**: ttH vs. tt+jets (bb) against with multivariate tools: BDT or MEM
- **limit**: combined fit in all categories on discriminant shape

Data and simulation

- 2.7/fb of good data, (~50% of Run I stats for ttH)
- NLO modelling of backgrounds with POWHEG, tt+jets inclusive
- ttbar statistics a major issue, esp. for classifier training
- Minor backgrounds modelled by MC: single top (t, tW), diboson, W+jets, tt+W/Z, DY-jets

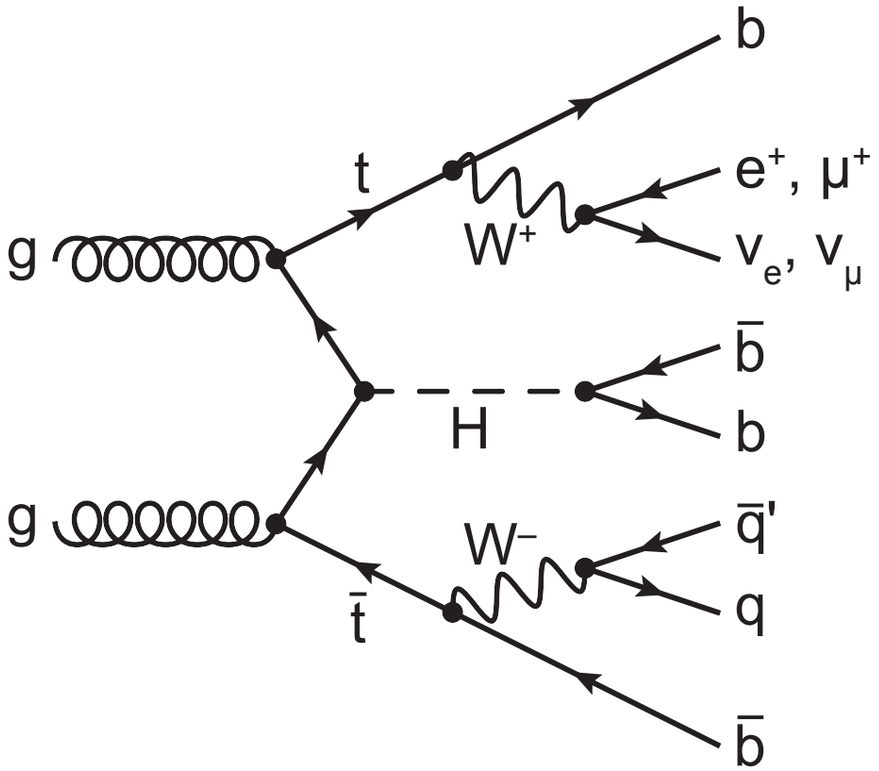
tt+jets subcomponents



tt+jets splitting via hadron matching

details in Tim's talk

Physics objects



leptons

- top pair decaying to lepton +jets: 1 tight muon (electron) with pt > 25 (30) GeV
- dileptonic decay: 1 tight, 1 loose opposite sign lepton

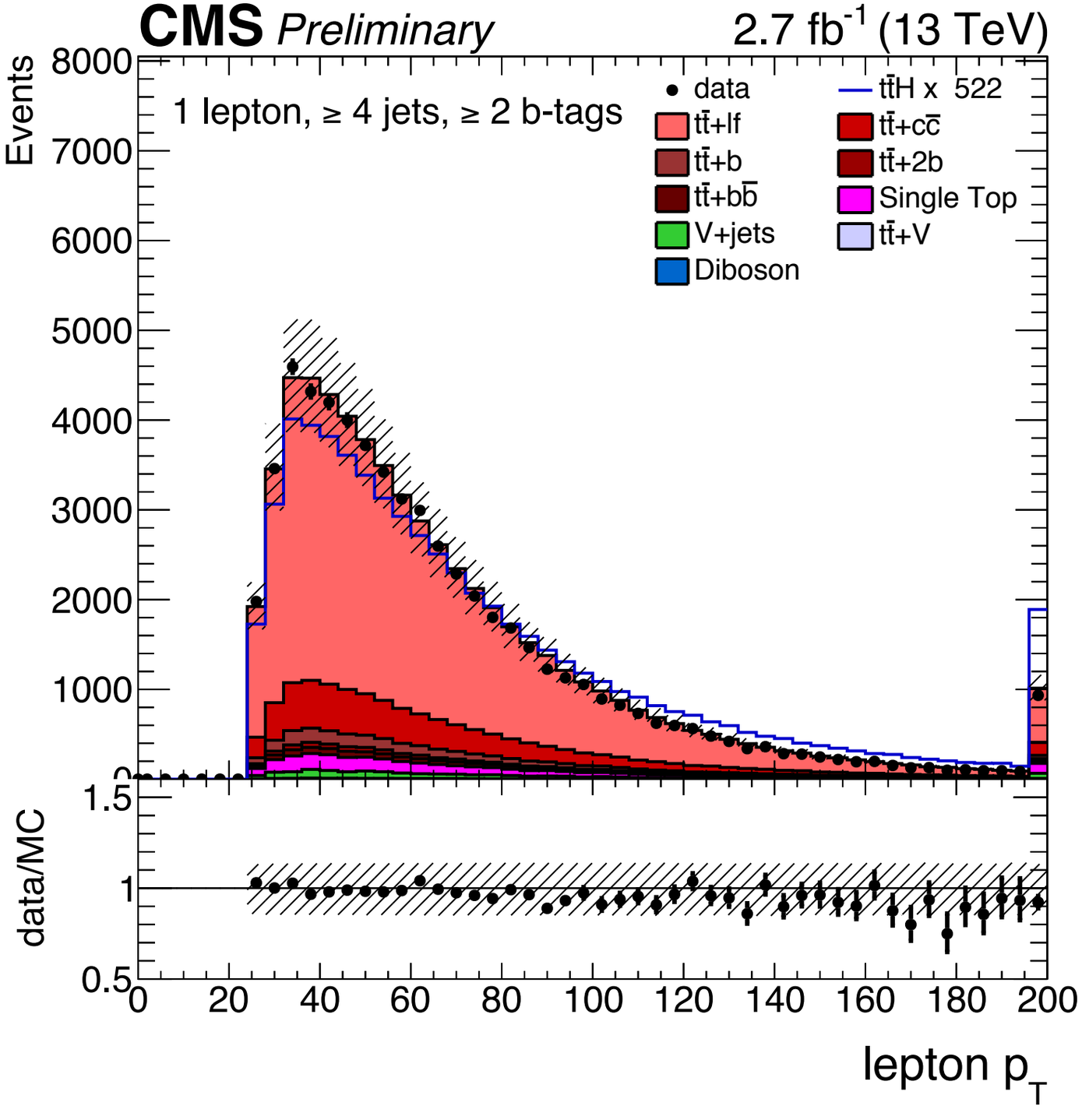
jets

- pt>30 GeV, |η|<2.4
- loosened pt for DL subleading jets

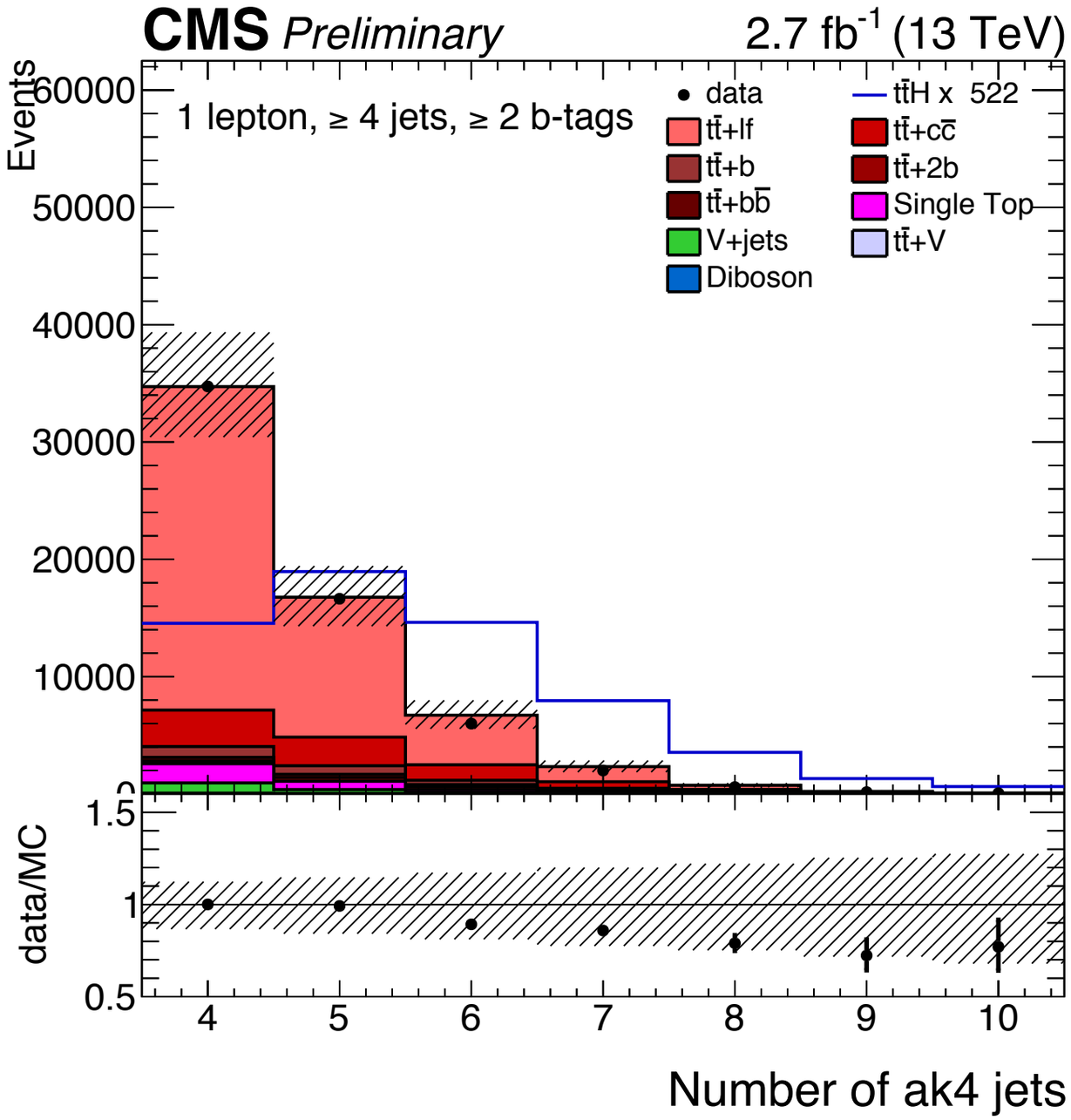
b-tagging

- # of b-tags as categorization
- b-discriminator shape as signal discriminator

Lepton kinematics



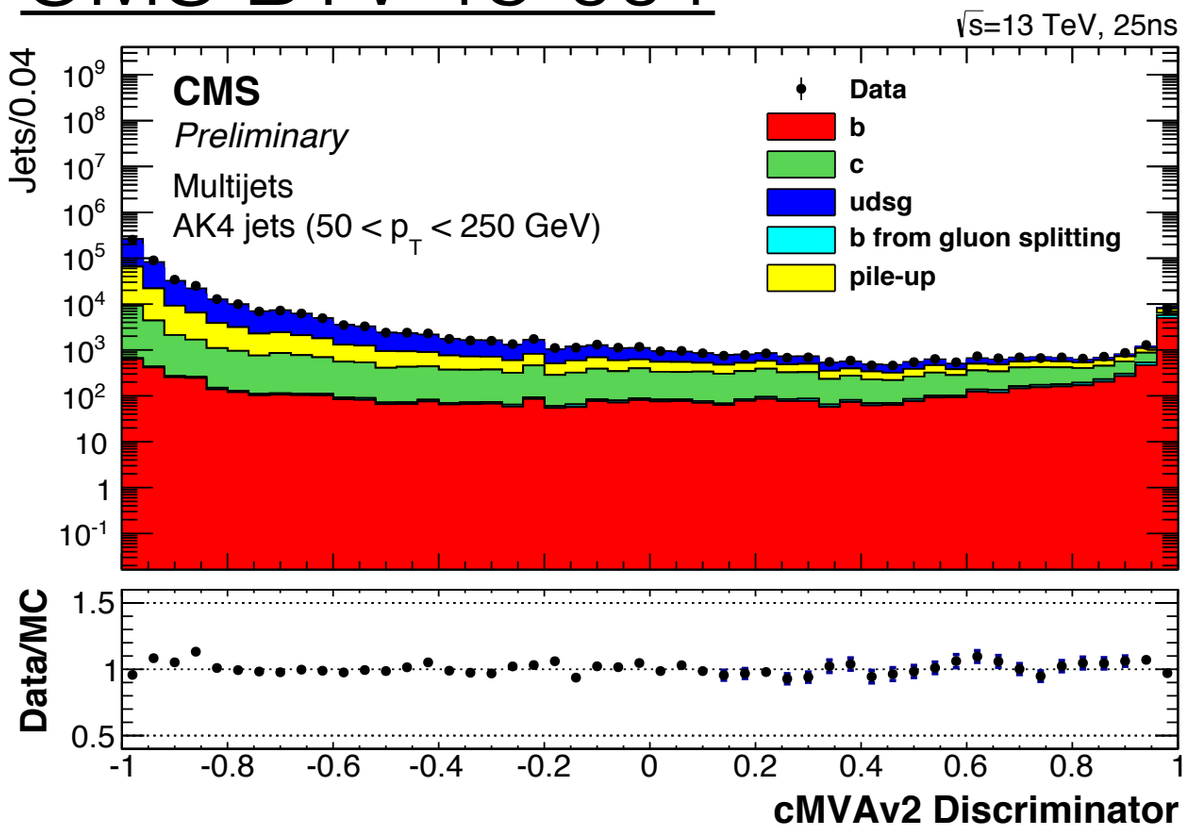
Jet multiplicity



Observe #jets overprediction, well known in CMS, connected to parton shower modelling

Improved b-tagging

CMS BTV-15-001

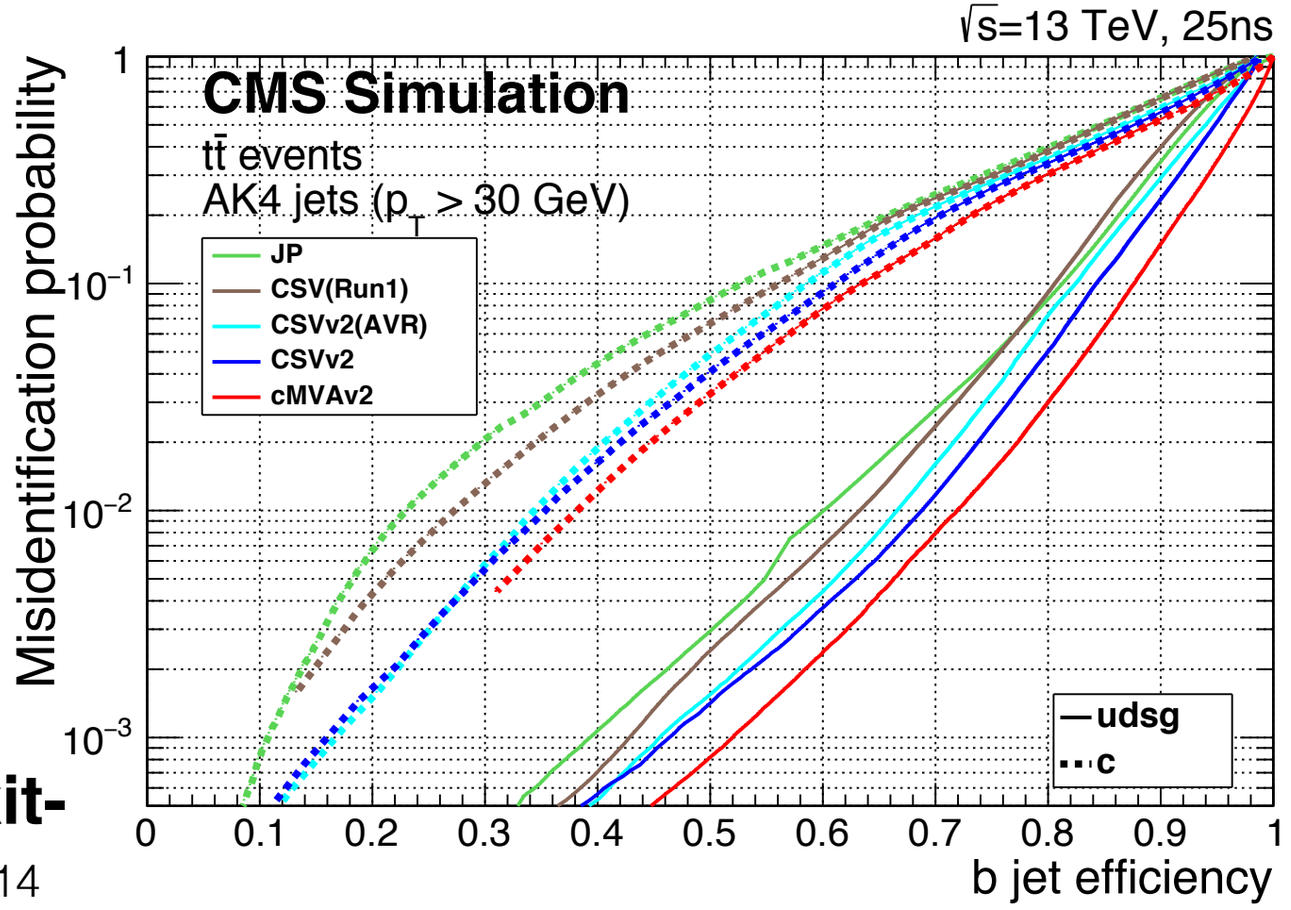


- Integrated into CMS software, deployed in 13 TeV data

- O(30%) better against fake b-quarks

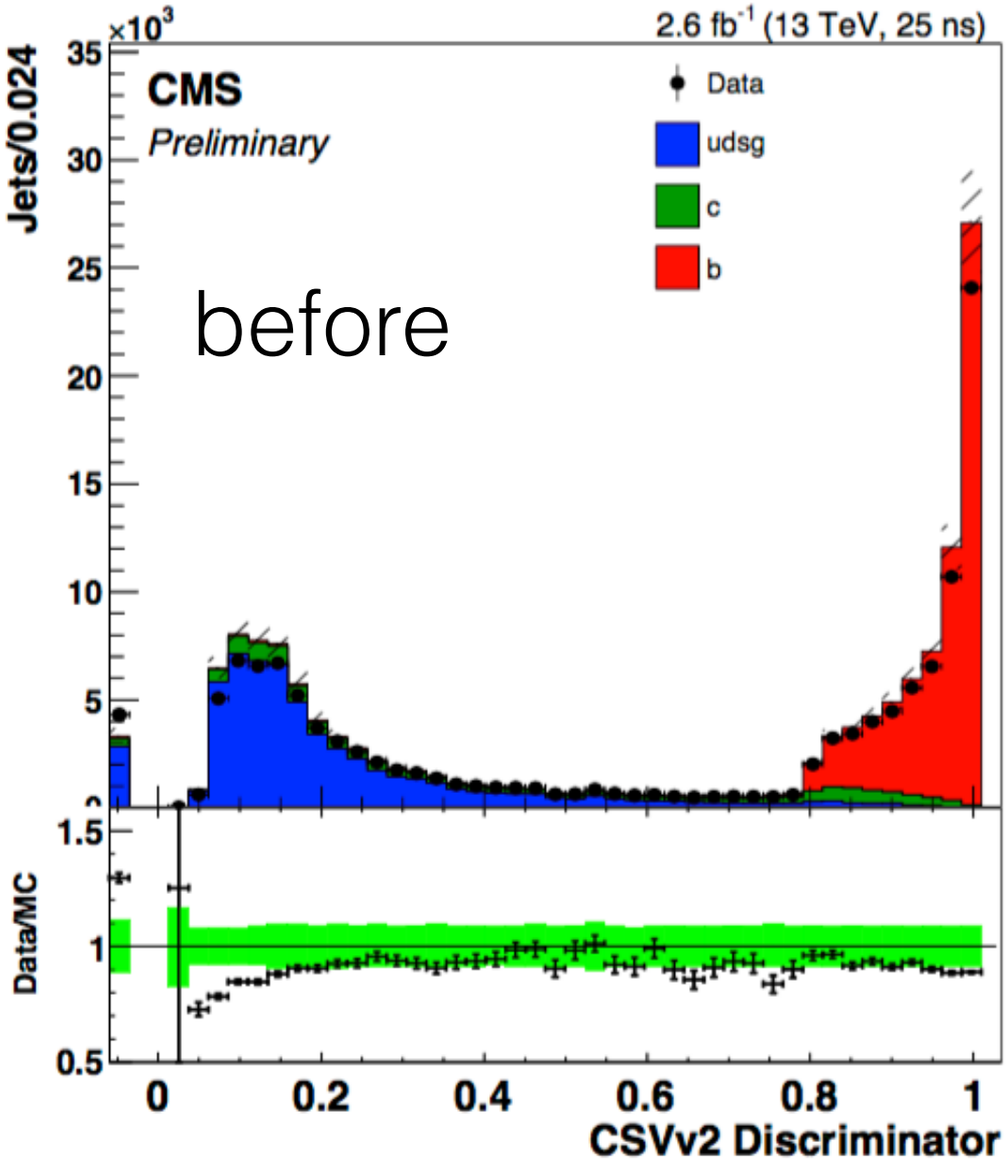
- As service task, developed a new b-discriminator algorithm for CMS

- based on subdetector combination

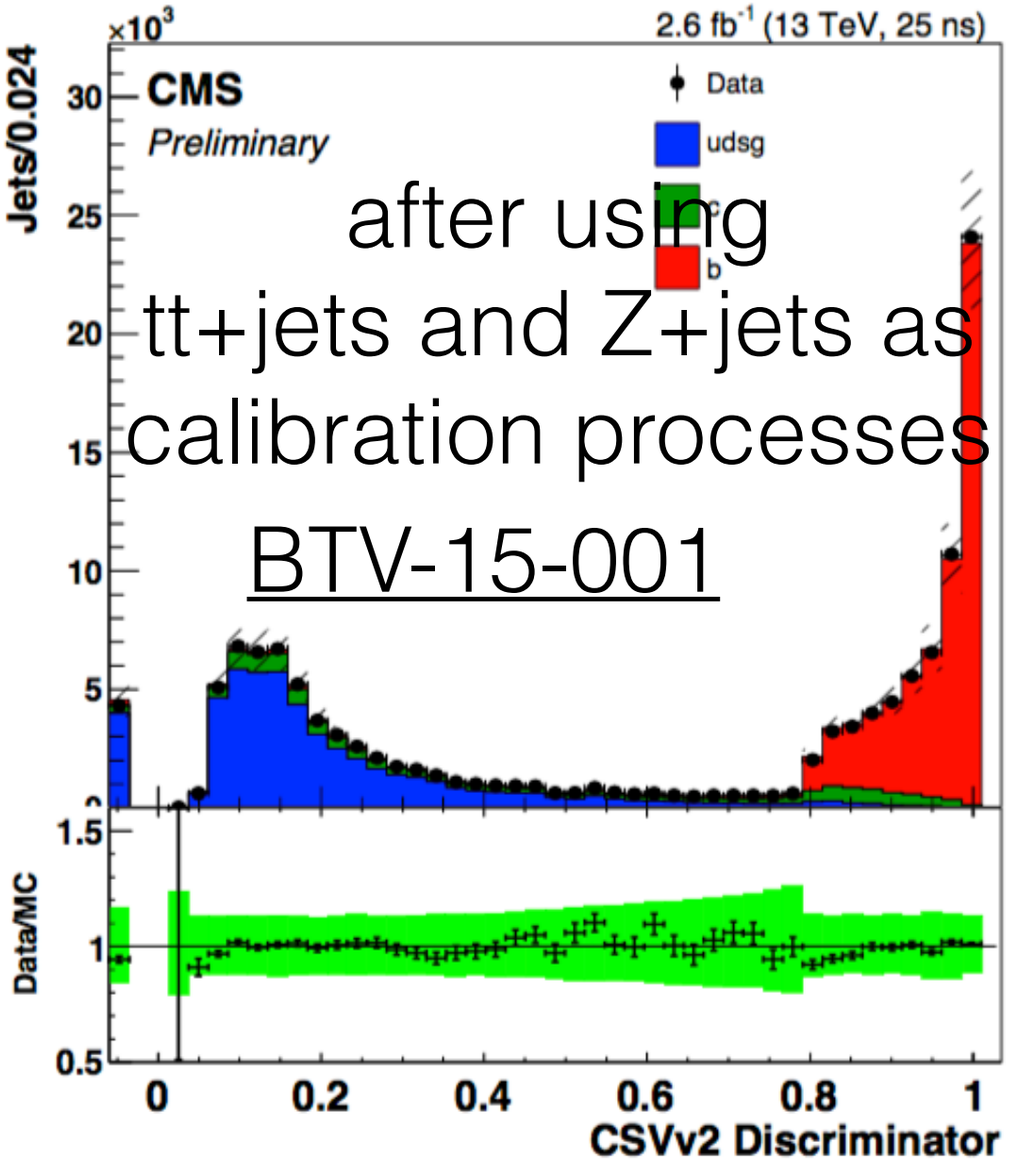


First major use of scikit-learn in LHC!

CSV reshaping



before



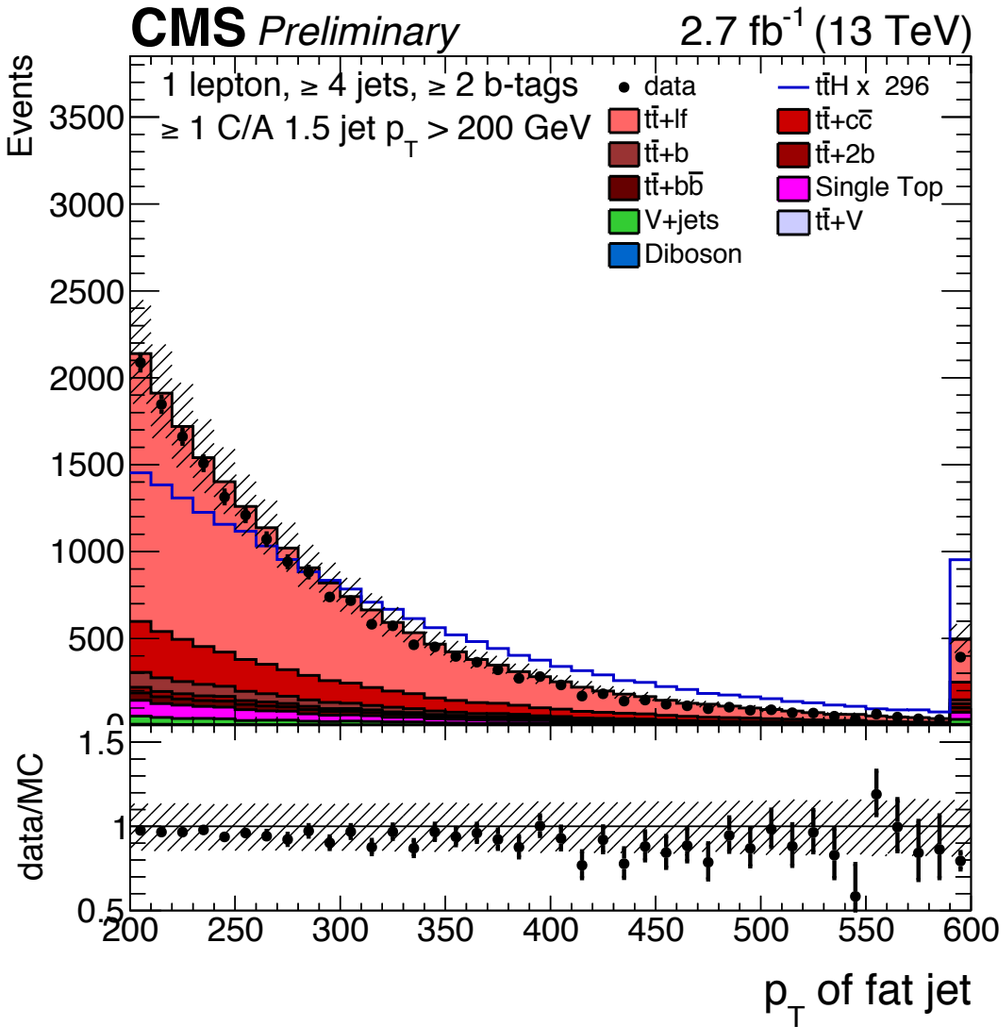
after using tt+jets and Z+jets as calibration processes

BTV-15-001

- Need to reshape CSV distribution, major systematic

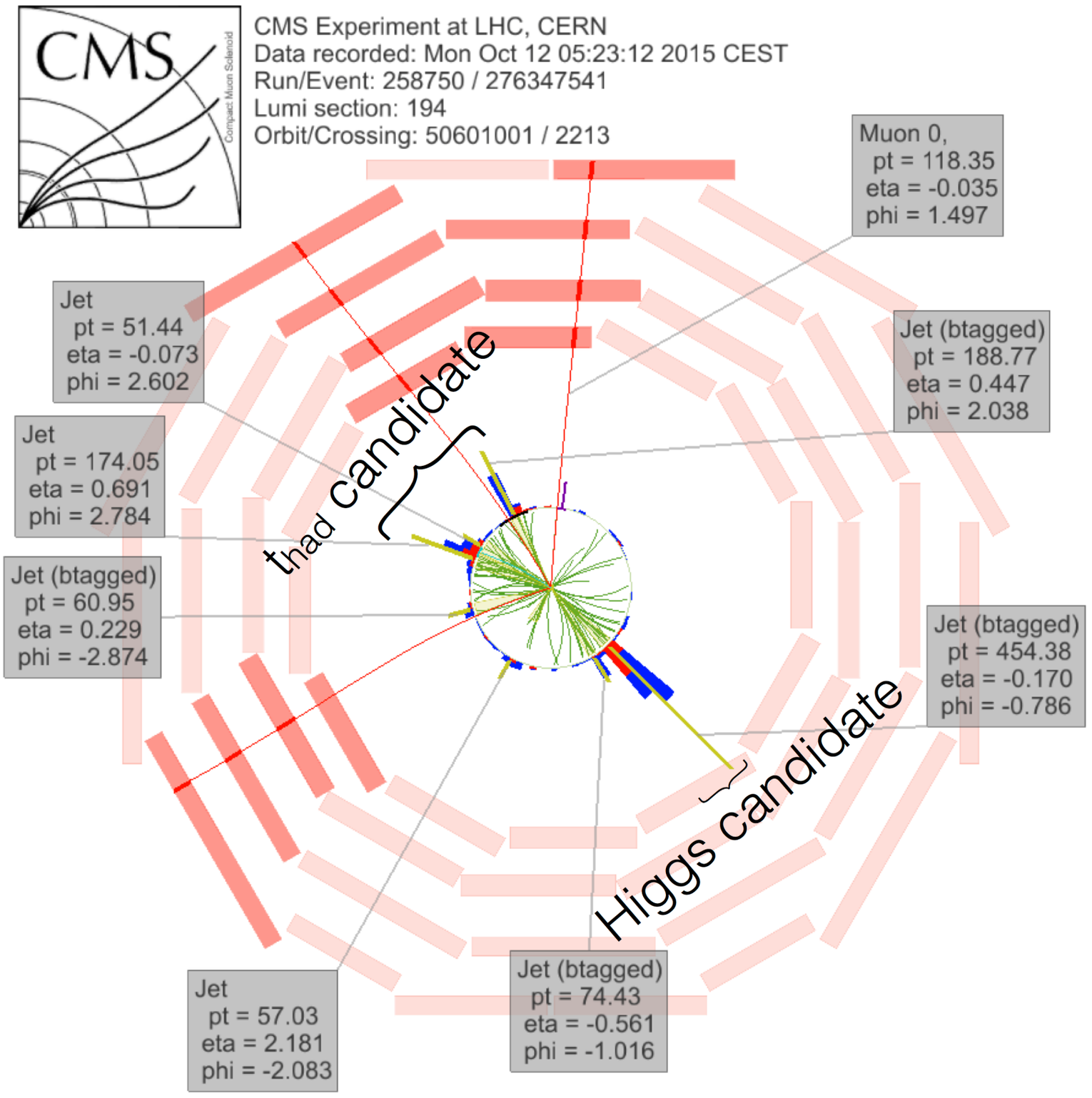
Boosted objects

- CA R=1.5 jets, excluding tight leptons, pt>200 GeV, |η| < 2.0
- use a dedicated MVA based on HepTopTaggerV2 for top-tagging
- Same b-discriminant as for resolved jets on subjets
- Details in Gregor's talk



We have already hundreds of high-momentum boosted jets!

Boosted candidate

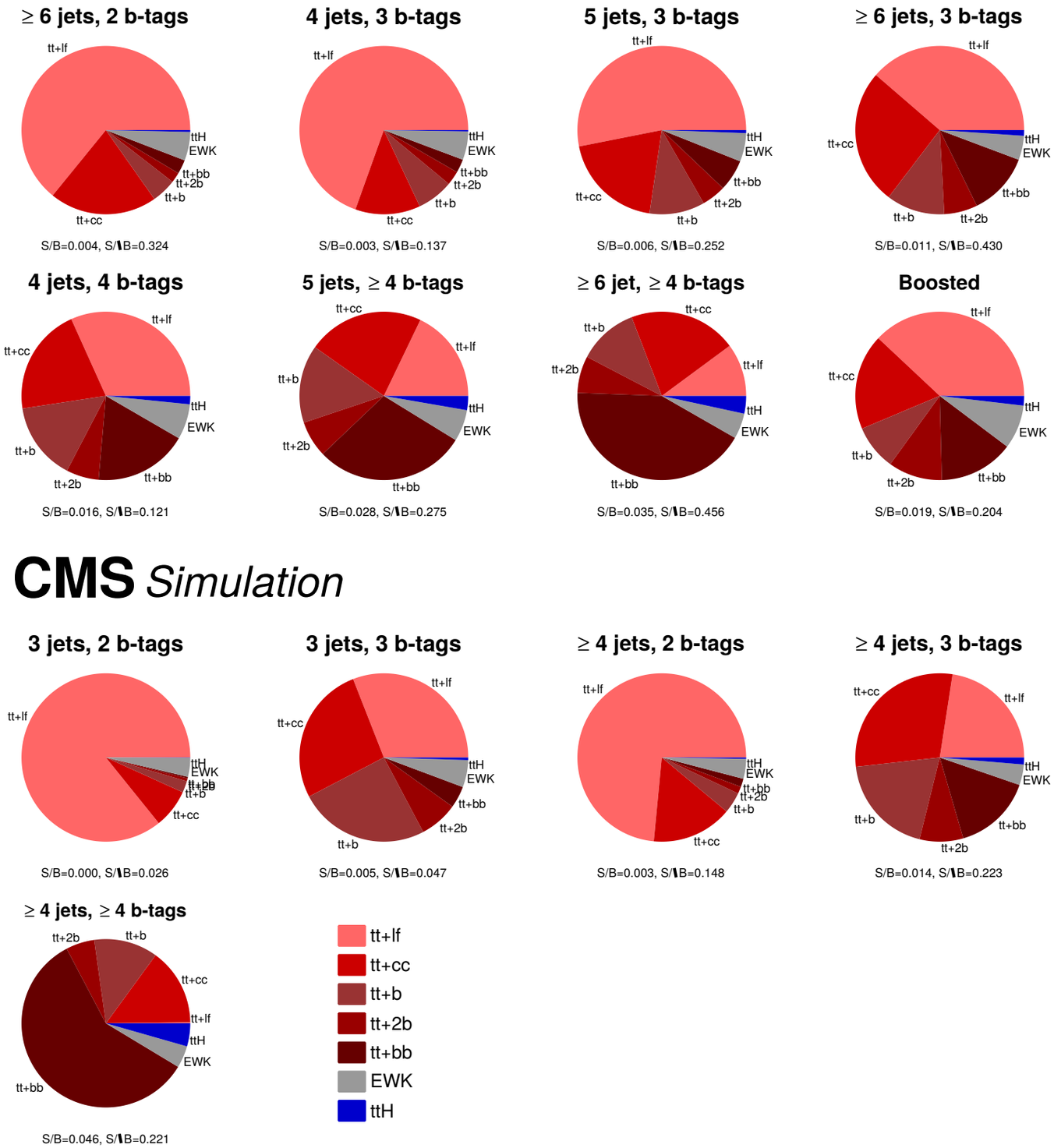


Now that we have some events, how to optimally use them in the analysis?

Currently, create categories with different background components based on jet/b-tag multiplicity

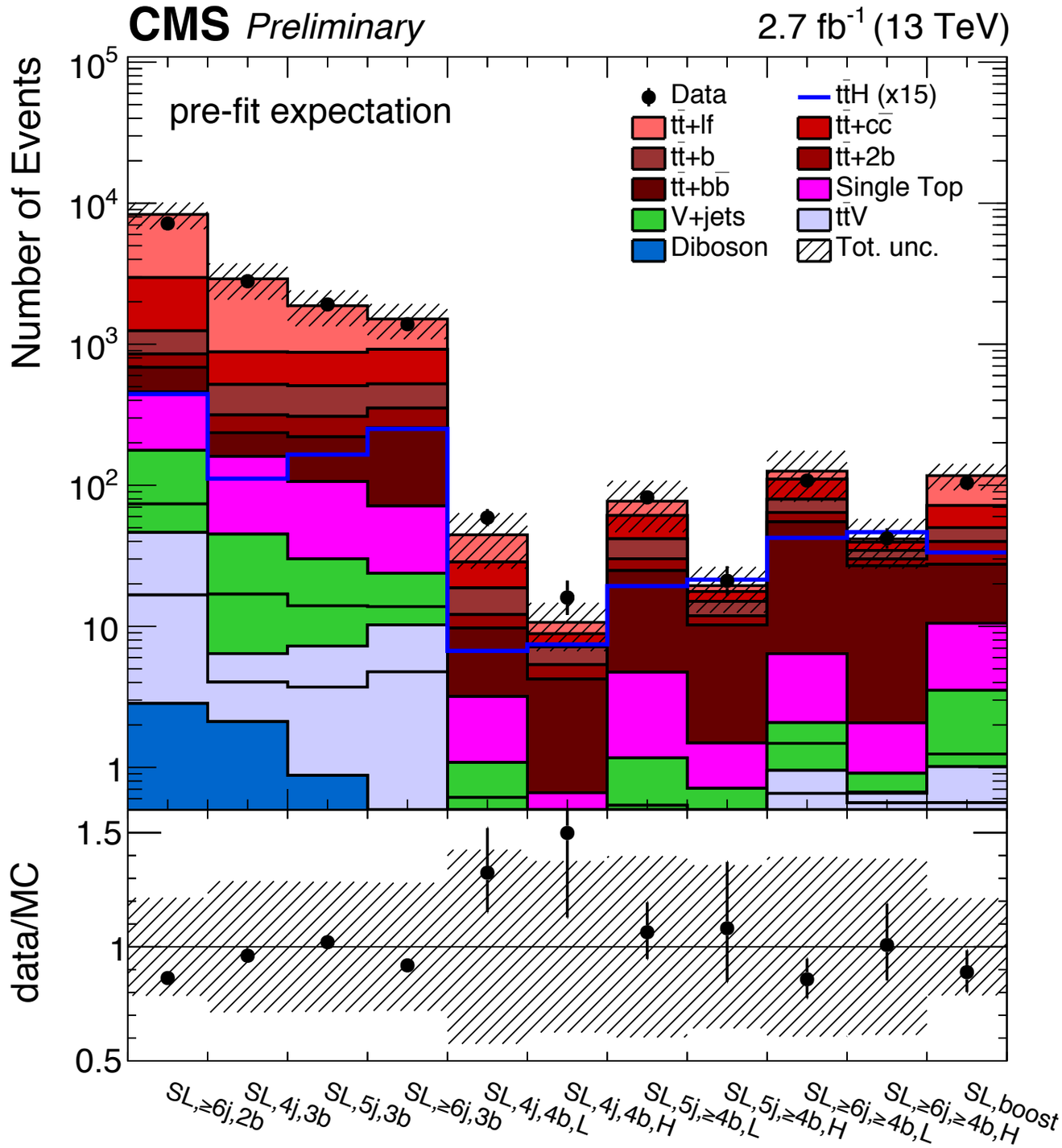
Analysis channels

- 13 final categories
- **Lepton+jets:** exactly 1 tight lepton, veto loose, ≥4j ≥3t or ≥6j ≥2t, 7 baseline categories
- **Boosted**: if l+j and ≥4j ≥2t and both boosted Higgs and top present
- **Dilepton:** exactly 2 opposite sign leptons (1loose + 1tight), 4 baseline categories (3j2t, 3j 3t, ≥4j 2t, ≥4j 3t, ≥4j ≥4t)
- categories defined to be non-overlapping (also SL & DL)



lepton + jets

- Subdivide to 7 jet/tag categories + 1 boosted category
- Evaluate MEM and BDT discriminators, tuned category-by-category
- Study correlations of discriminators
- combined discriminators if motivated



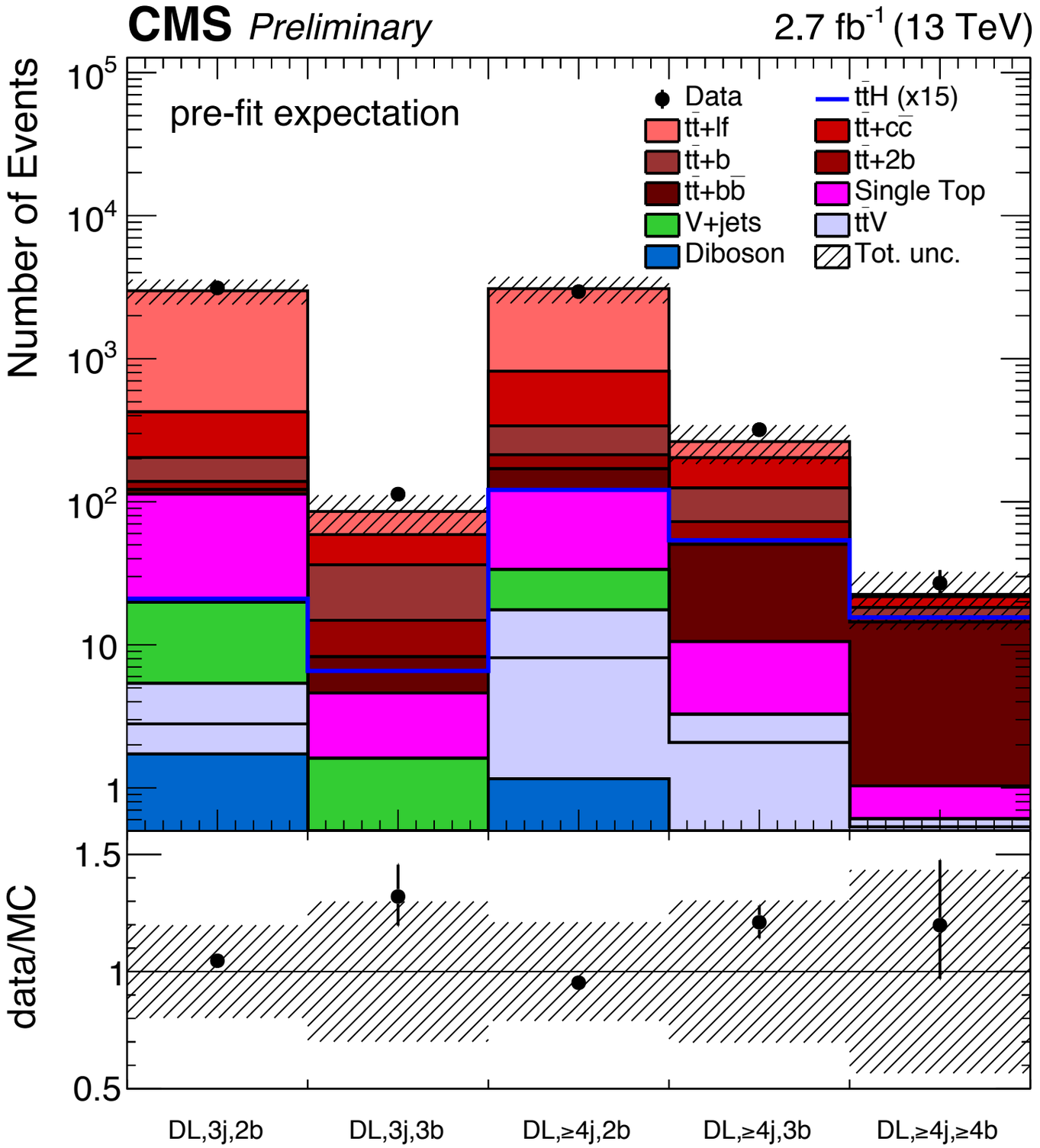
prefit yields

dilepton channel

- **5 categories**: 3j2t, 3j 3t, 3j ≥4t, ≥4j 2t, ≥4j ≥4t

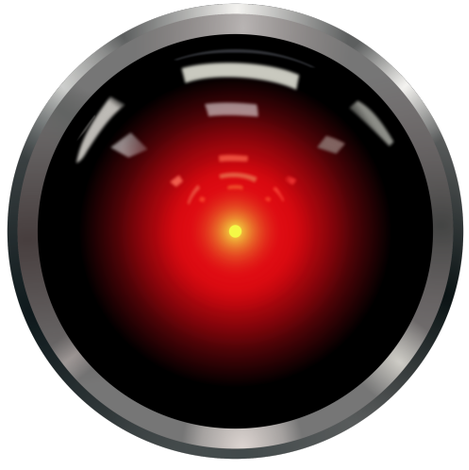
- Low data statistics (also MC), S/√B ~ 0.2 in best category

- Same approach for discriminators as in lepton+jets



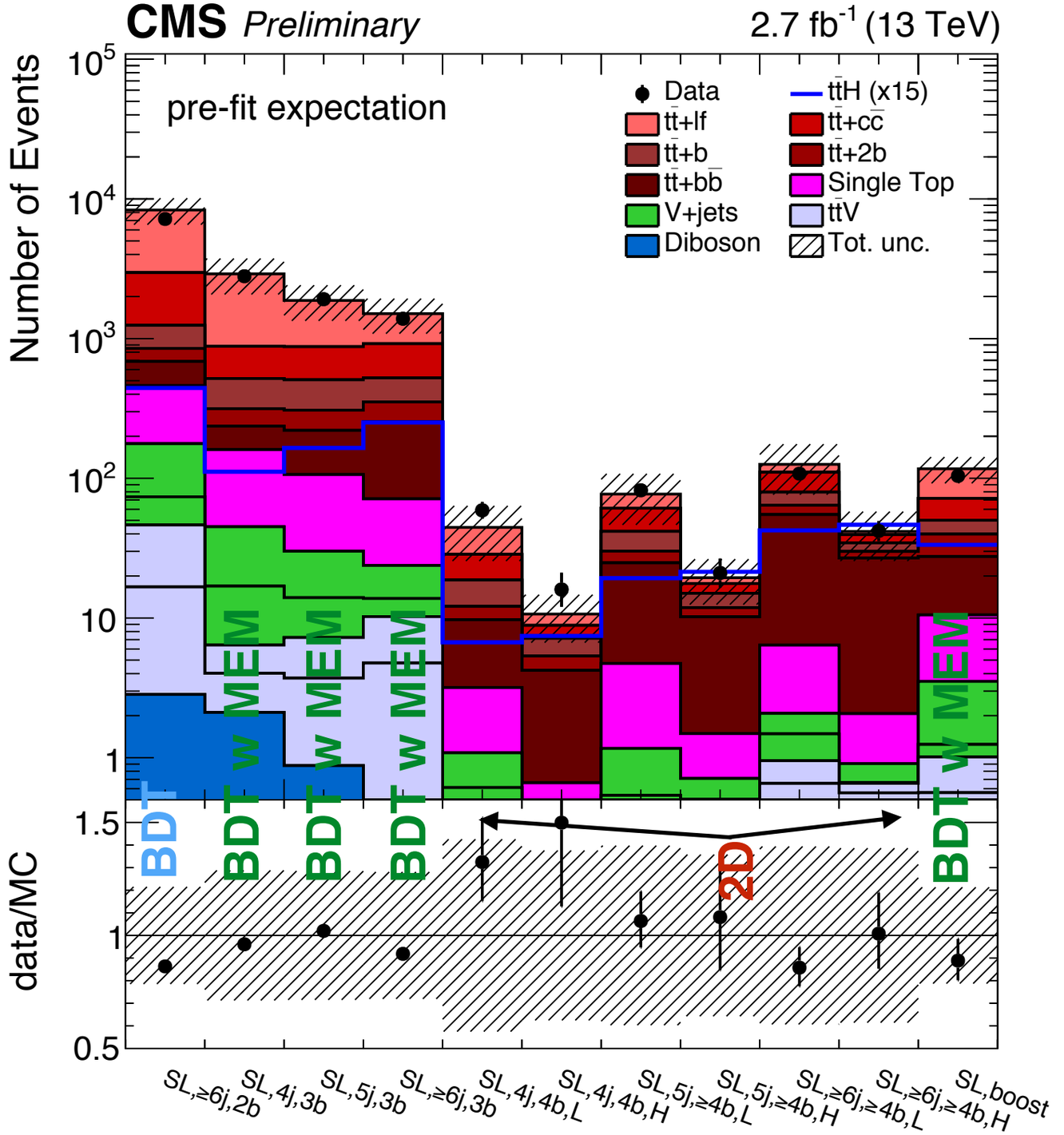
tt+H discrimination

- Purest category has 3.5% signal
- Need an MVA to separate out ttH(bb) from tt +bb
- **BDT** trained category-by-category to separate ttH(bb) from tt+jets: **need 500M+ fully simulated NLO events**
- **MEM@LO** using leptons and b-quark kinematics to separate ttH(bb) from tt+bb: **O(30) CPU-years to evaluate**



Combined discriminant

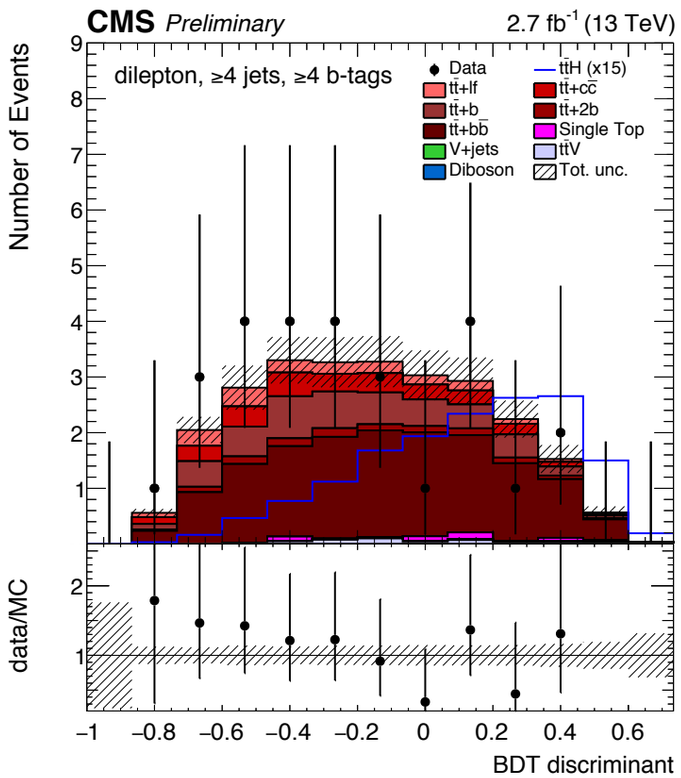
- Adopted combination scheme:
  - **BDT deployed in 2-tag events**: light flavour dominated, no MEM needed
  - **MEM included in BDT in 3-tag for events with high BLR**: can deploy both MEM & BDT, stats for training
  - **2D MEM vs. BDT in 4t:** low MC stats, but MEM applicable for all events
- Limit changes of ~5-10% expected from binning effects: statistical variability of the procedure
- Expected without combination (BDT only or MEM+BLR only) stand around 4.7… 5.0 depending on the used binning. **Combination improves noticeably!**

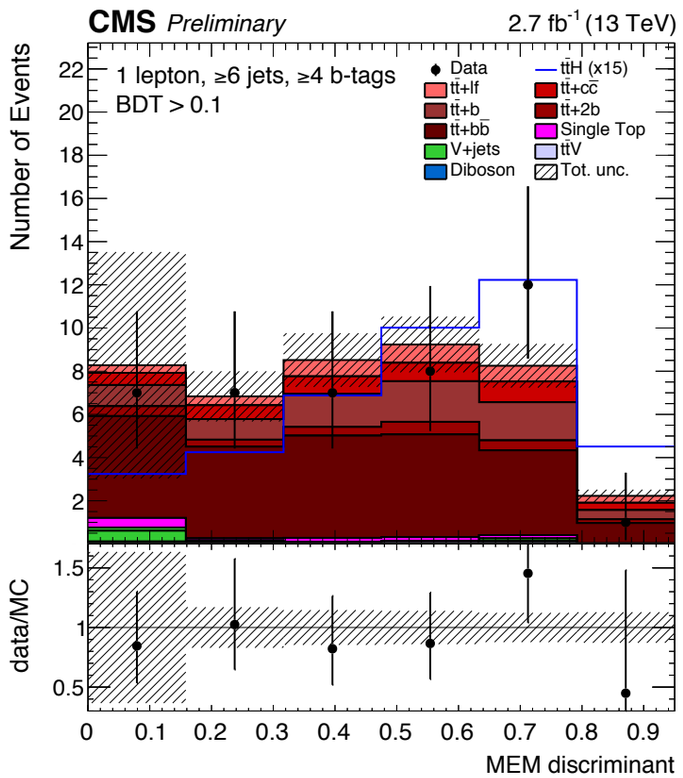


nota bene: MEM-only analysis possible but ~20% less sensitive

Limit procedure

- Fit the signal and background scale factors across all categories simultaneously
- Systematic uncertainties, MC stats as nuisance parameters, correlated between all categories as appropriate
- Quote median expected limit on $\mu=\sigma/\sigma_{SM}$ @ 95% CL





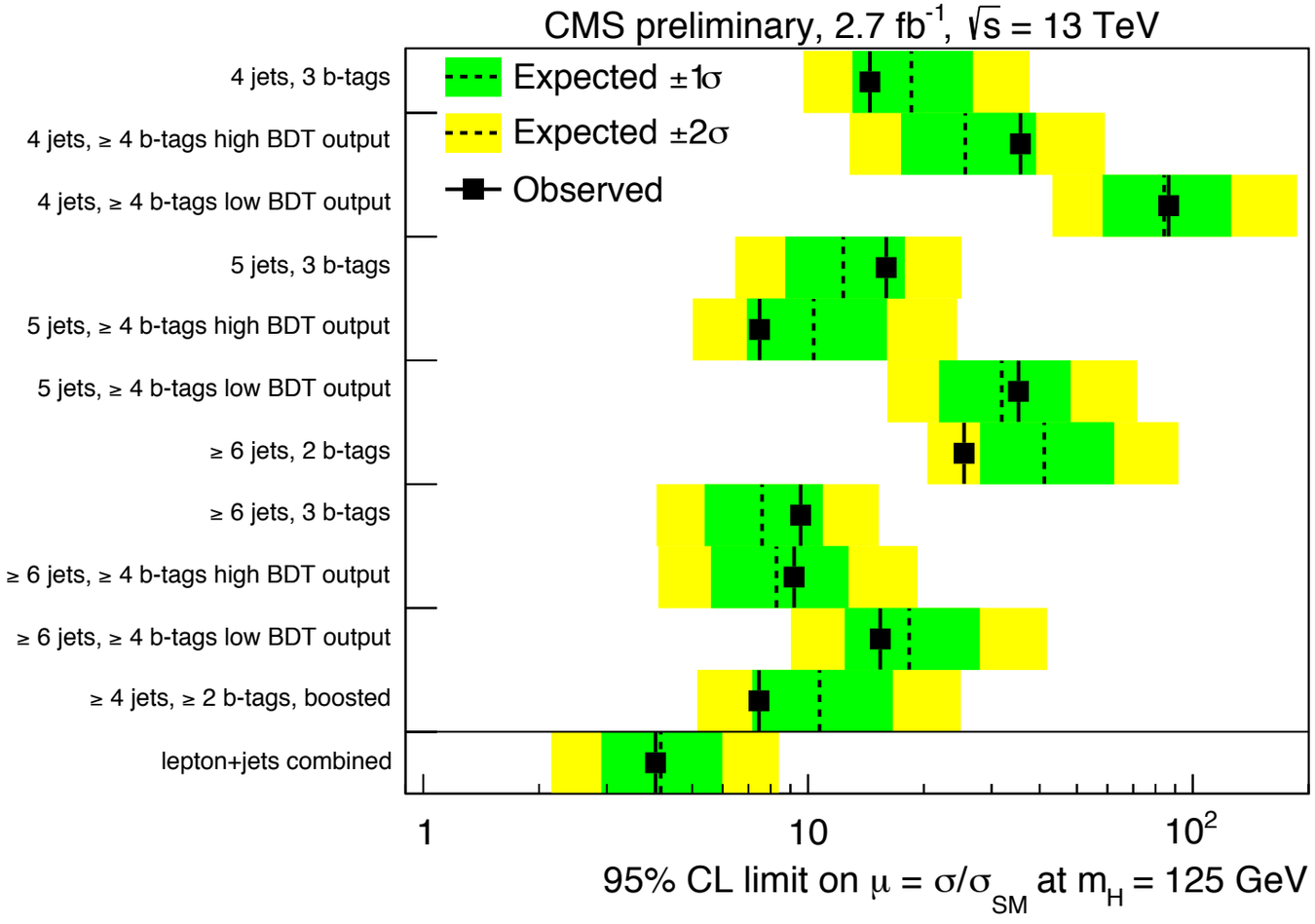
Systematic uncertainties

- b-tagger shape calibration
- rate and shape of tt+jets components (QCD scale)
  - 50% uncorrelated priors on tt+heavy
  - 30% on tt+light
- Rate of ttH from QCD scale
- jet energy scale, resolution (changes acceptance)

Limits

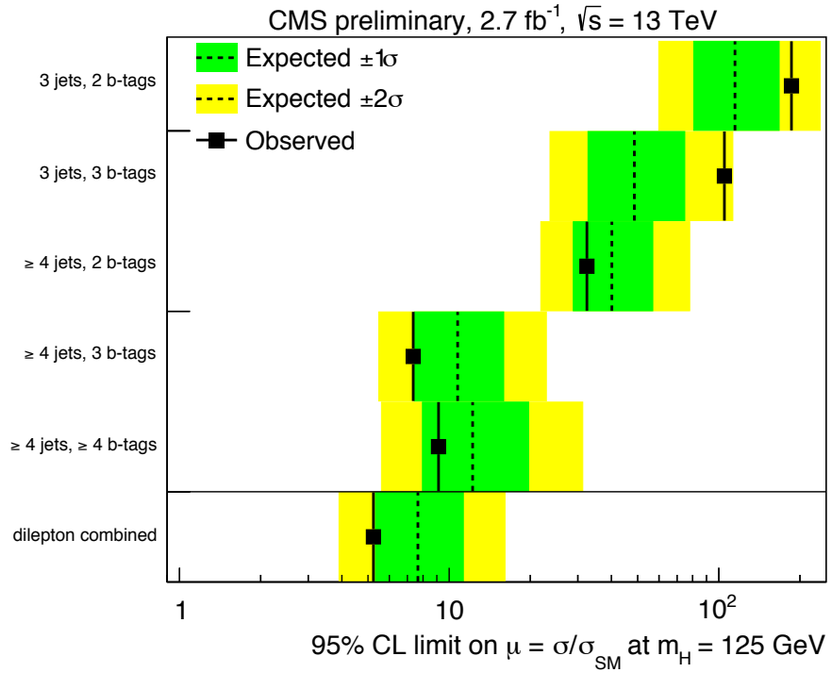
lepton+jets

Category	Observed	Expected
4 jets, 3 b-tags	14.5	$18.6^{+8.2}_{-5.5}$
4 jets, $\geq$ 4 b-tags high BDT output	35.7	$25.6^{+13.4}_{-8.1}$
4 jets, $\geq$ 4 b-tags low BDT output	86.6	$84.2^{+41.3}_{-25.8}$
5 jets, 3 b-tags	16.0	$12.3^{+5.5}_{-3.6}$
5 jets, $\geq$ 4 b-tags high BDT output	7.5	$10.3^{+5.6}_{-3.4}$
5 jets, $\geq$ 4 b-tags low BDT output	35.2	$31.9^{+16.1}_{-9.9}$
$\geq$ 6 jets, 2 b-tags	25.4	$41.1^{+21.1}_{-13.1}$
$\geq$ 6 jets, 3 b-tags	9.6	$7.6^{+3.3}_{-2.2}$
$\geq$ 6 jets, $\geq$ 4 b-tags high BDT output	9.2	$8.3^{+4.4}_{-2.7}$
$\geq$ 6 jets, $\geq$ 4 b-tags low BDT output	15.4	$18.3^{+9.6}_{-5.8}$
$\geq$ 4 jets, $\geq$ 2 b-tags, boosted	7.5	$10.7^{+5.9}_{-3.5}$
lepton+jets combined	4.0	$4.1^{+1.8}_{-1.2}$

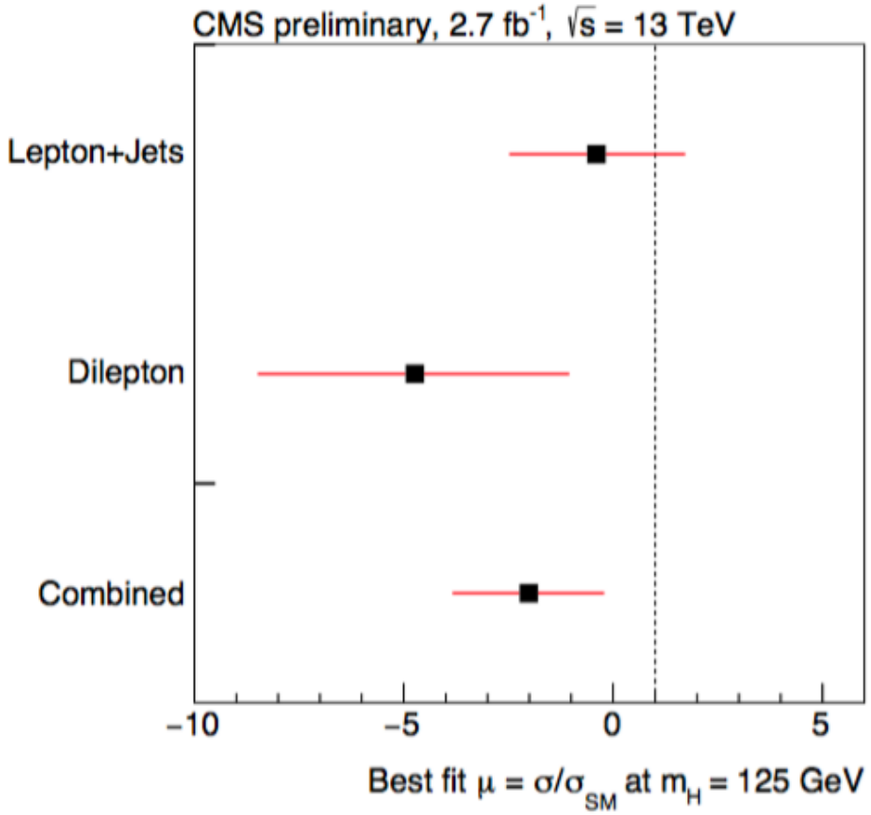


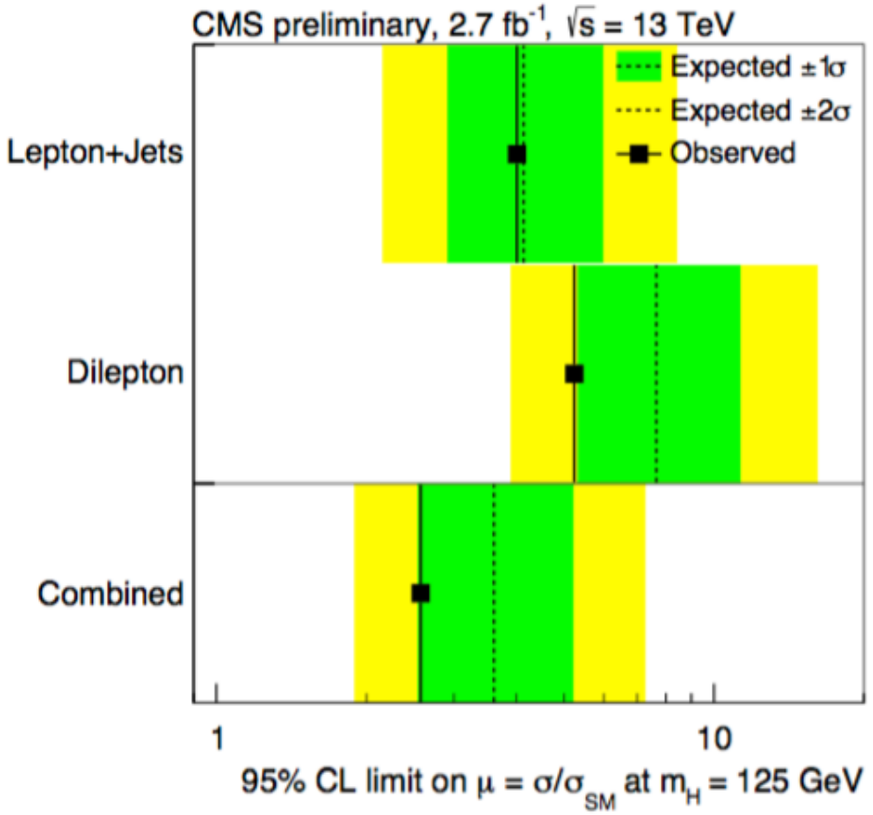
di-lepton

Category	Observed	Expected
3 jets, 2 b-tags	186.0	$114.8^{+52.6}_{-34.1}$
$\geq$ 3 jets, 3 b-tags	104.9	$48.6^{+26.2}_{-15.9}$
$\geq$ 4 jets, 2 b-tags	32.4	$40.1^{+16.8}_{-11.3}$
$\geq$ 4 jets, 3 b-tags	7.4	$10.8^{+5.2}_{-3.3}$
$\geq$ 4 jets, $\geq$ 4 b-tags	9.1	$12.2^{+7.5}_{-4.3}$
dilepton combined	5.2	$7.7^{+3.6}_{-2.3}$



Combined limits





Channel	Best-fit μ	Observed UL	Expected UL
Lepton+jets	$-0.4^{+1.0}_{-1.0}(\text{stat})^{+1.9}_{-1.9}(\text{syst})$	4.0	$4.1^{+1.8}_{-1.2}$
Dilepton	$-4.7^{+2.1}_{-1.9}(\text{stat})^{+3.0}_{-3.3}(\text{syst})$	5.2	$7.7^{+3.6}_{-2.3}$
Combined	$-2.0^{+0.9}_{-0.8}(\text{stat})^{+1.6}_{-1.6}(\text{syst})$	2.6	$3.6^{+1.6}_{-1.1}$

Summary

- **First results on ttH(bb) with 13 TeV data**

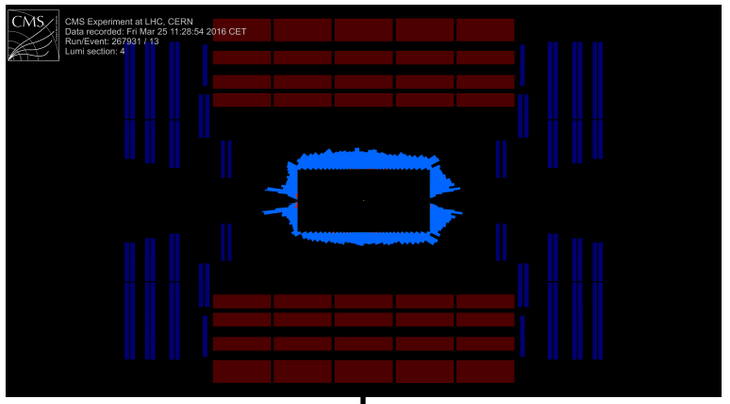
- Build on existing run I analyses, combining several techniques:

  - improved tt+bb discrimination, MEM deployed on additional expensive final states, BDT optimized

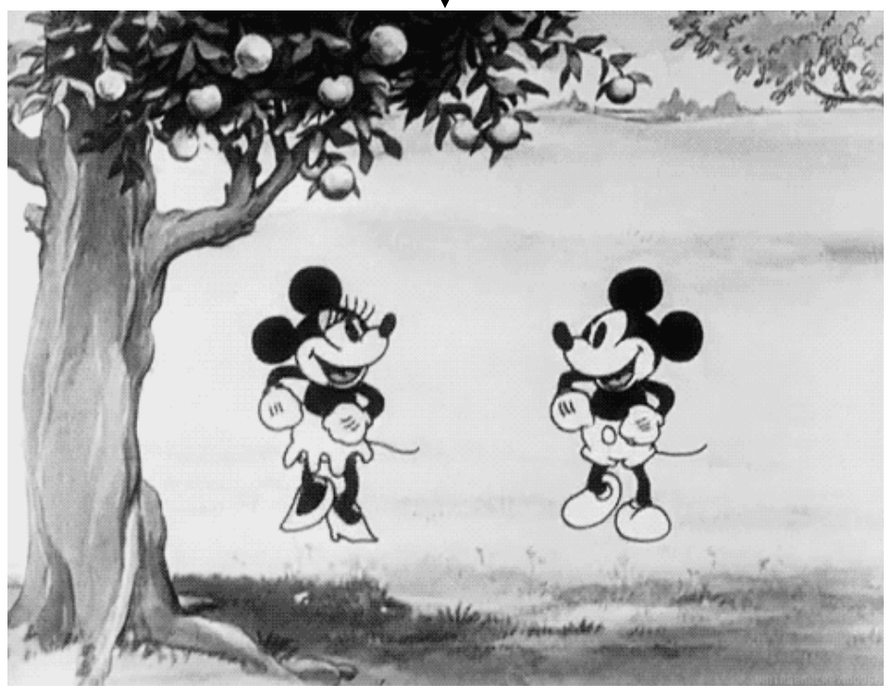
  - MEM+BDT combination methods explored

  - set up first boosted analysis

- Ready for 2016 data taking with 30/fb expected



Channel	Best-fit μ	Observed UL	Expected UL
Lepton+jets	$-0.4^{+1.0}_{-1.0}(\text{stat})^{+1.9}_{-1.9}(\text{syst})$	4.0	$4.1^{+1.8}_{-1.2}$
Dilepton	$-4.7^{+2.1}_{-1.9}(\text{stat})^{+3.0}_{-3.3}(\text{syst})$	5.2	$7.7^{+3.6}_{-2.3}$
Combined	$-2.0^{+0.9}_{-0.8}(\text{stat})^{+1.6}_{-1.6}(\text{syst})$	2.6	$3.6^{+1.6}_{-1.1}$



Many thanks to the organizers for a wonderful conference!

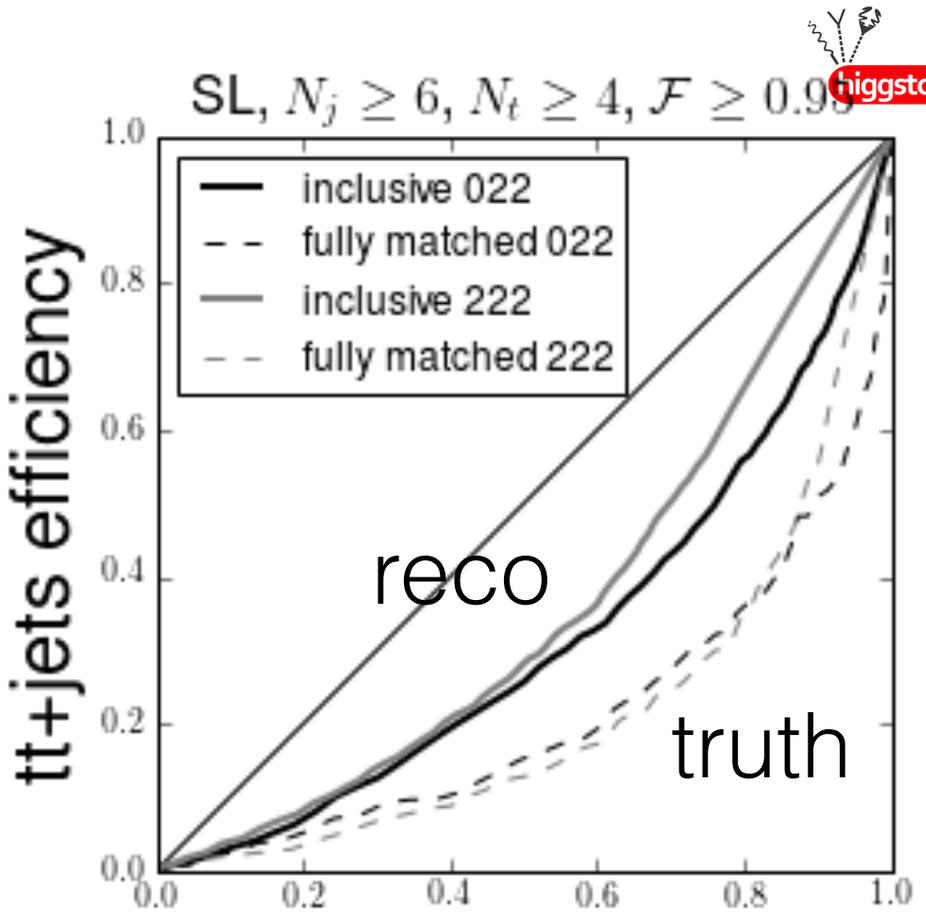
Bonus slides

ETH Institute for Particle Physics

MEM

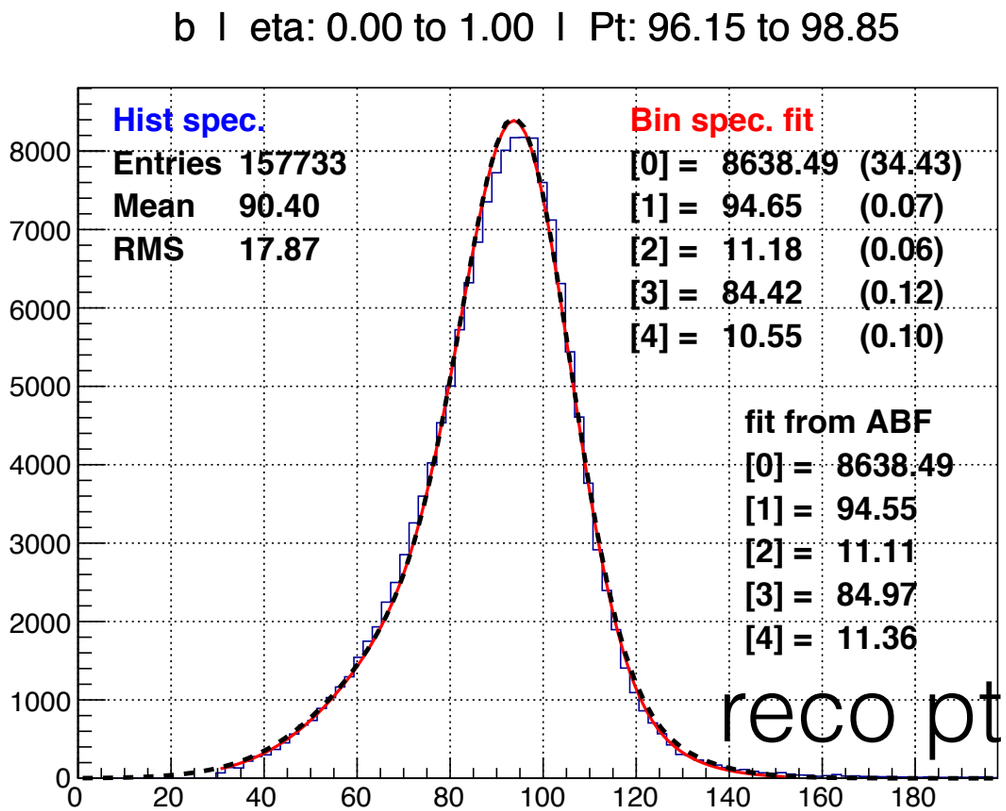
$$P_{\mathrm{s/b}} = \frac{w(\vec{y}|\mathrm{t\bar{t}H})}{w(\vec{y}|\mathrm{t\bar{t}H}) + k_{\mathrm{s/b}} w(\vec{y}|\mathrm{t\bar{t}+b\bar{b}})}$$

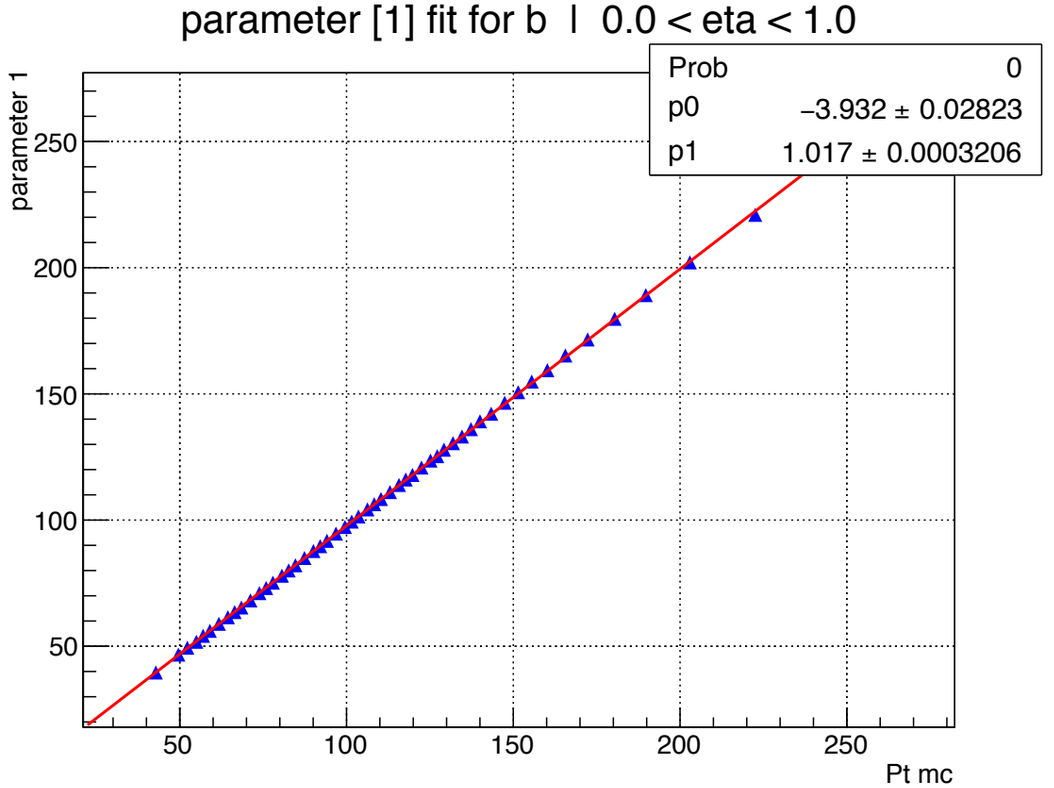
- Based on Run I approach, new validated code
- event interpretation × detector description
- Uses LO hard scattering from OpenLoops, gluon fusion only
- consider only b-quarks, integrate over light quarks
- Unresolved events from incompatible kinematics, resolve with BDT



Transfer functions

- Derived from ttH(bb) simulation
- double-gaussian fit in bins of generated pt, parameter dependence on pt polynomial





$$F(p_T|p_{T,\text{gen}}) = p_0\left[0.7\exp\left(\frac{p_T - p_1}{p_2}\right)^2 + 0.3\exp\left(\frac{p_T - p_3}{p_2 + p_4}\right)^2\right]$$

BDT base variables (SL)

- **Inputs:** O(30) baseline + 8 (boosted) + 2 (MEM, BLR) variables, optimized from Run I
- **Object and event kinematics:** pt of 4 hardest jets, sum of all jet transverse momenta and energies, MET, invariant mass of jet+lepton+MET, M3 (inv. mass of 3 hardest jets)
- **angular:** mean, min ΔR; mean η of (b-tagged) jets, ΔR(lep, closest jet), inv. mass of two closest (ΔR) b-tagged jets, inv. mass of b-tagged jets closest to 125 GeV
- **event shape:** Fox-Wolfram moments 0-3
- **best Higgs mass, geom. mean of top-higgs rapidity differences:** χ2-based reco of leptonic & hadronic top, Higgs in 6j4t events
- In case of 5 jets, promote highest jet with pt>20, similarly in 3t, promote jet with 4th highest CSV

Table 8: BDT input variable assignment per category in the lepton+jets channel.

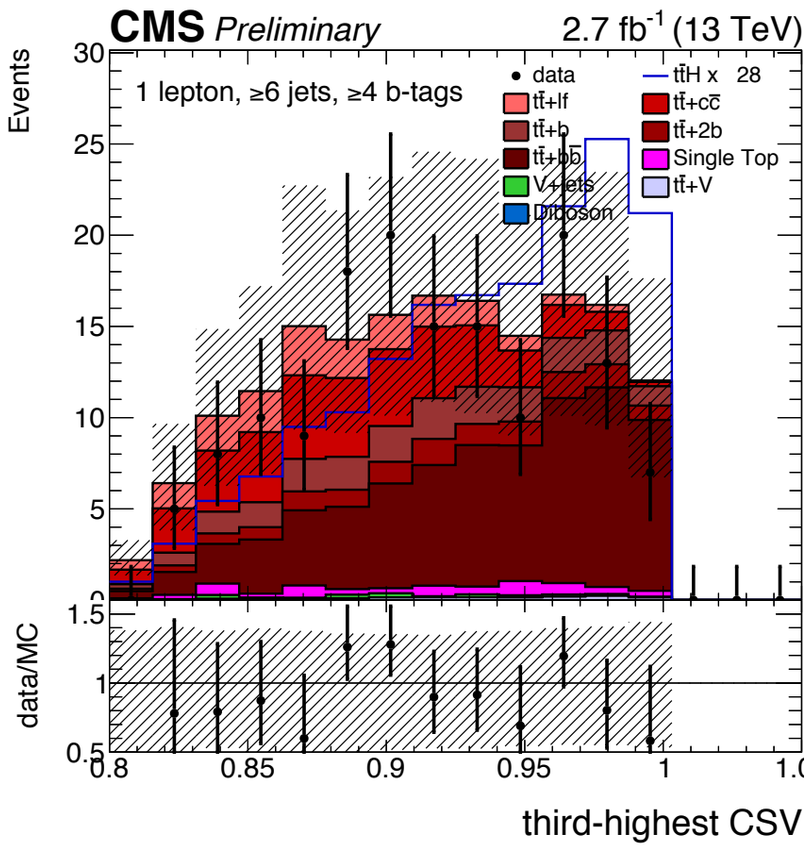
≥ 4 jets, ≥ 2 b-tags boosted	4 jets, 3 b-tags	4 jets, ≥ 4 b-tags
avg ΔR(tag,tag) τ_2/τ_1 of Higgs cand. third-highest CSV fourth-highest CSV $\Delta\eta$(top,Higgs) aplanarity m(Higgs, di-filterjet) min ΔR(tag,tag) avg CSV (all) MEM discriminator (using subjets) b-tagging likelihood ratio	H_1 b-tagging likelihood ratio $\sum p_T$(jets,leptons,MET) MEM discriminator avg CSV (tags) avg CSV (all) jet 2 p_T jet 4 p_T	closest tagged dijet mass b-tagging likelihood ratio $\sum p_T$(jets,lepton,MET) avg ΔR(tag,tag) H_3 jet 1 p_T
	5 jets, 3 b-tags	5 jets, ≥ 4 b-tags
	MEM discriminator avg ΔR(tag,tag) min ΔR(lepton,jet) b-tagging likelihood ratio fourth-highest CSV H_1 dev from avg CSV (tags) avg $\Delta\eta$(jet,jet) avg CSV (tags) avg CSV (all) max $\Delta\lvert\eta\rvert$ (tag, avg jet $\lvert\eta\rvert$)	b-tagging likelihood ratio jet 3 p_T tagged dijet mass closest to 125 avg $\Delta\eta$(jet,jet) avg ΔR(tag,tag) H_1 fifth-highest CSV $(\Sigma$ jet $p_T)/(\Sigma$ jet E)
≥ 6 jets, 2 tags	≥ 6 jets, 3 tags	≥ 6 jets, ≥ 4 tags
avg $\Delta\eta$(tag,tag) avg ΔR(tag,tag) ΔR(jet1, jet2) b-tagging likelihood ratio max $\Delta\lvert\eta\rvert$ (tag, avg tags $\lvert\eta\rvert$) third-highest CSV sphericity fourth-highest CSV max $\Delta\lvert\eta\rvert$ (tag, avg jet $\lvert\eta\rvert$) min ΔR(tag,tag)	b-tagging likelihood ratio $\sqrt{\Delta\eta(t^{lep}, bb) \times \Delta\eta(t^{had}, bb)}$ HT MEM discriminator $\sum p_T$(jets,lepton,MET) H_1 fourth-highest CSV avg CSV (tags) max $\Delta\lvert\eta\rvert$ (tag, avg jet $\lvert\eta\rvert$)	$\sum p_T$(jets,leptons,MET) H_3 best Higgs mass b-tagging likelihood ratio tagged dijet mass closest to 125 fifth-highest CSV $(\Sigma$ jet $p_T)/(\Sigma$ jet E) jet 4 p_T sphericity max $\Delta\lvert\eta\rvert$ (tag, avg tag $\lvert\eta\rvert$) second-highest CSV

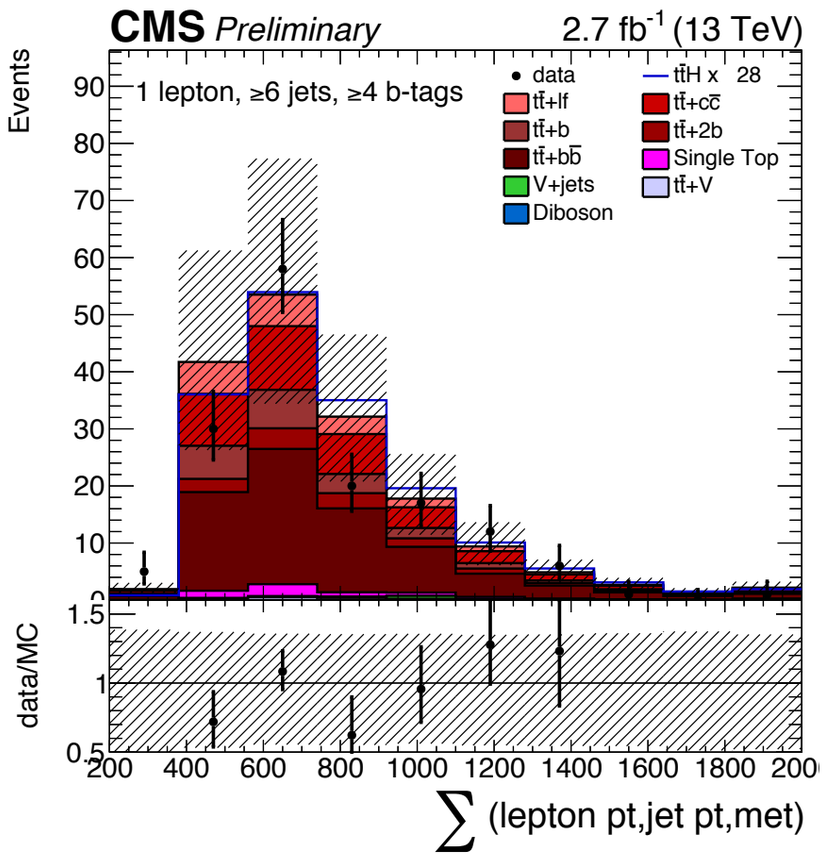
DRAFT

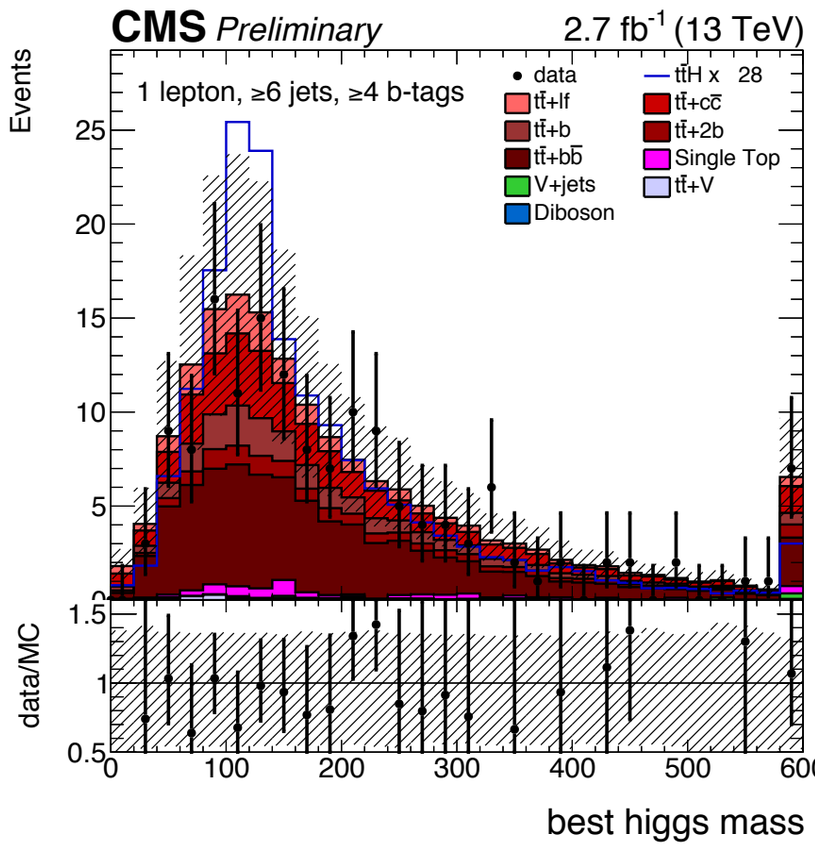
ETH Institute for Particle Physics

higgstools

BDT variables in 6j4t







- Example BDT input variables in the SL channel

Dilepton BDT

- O(20) variables

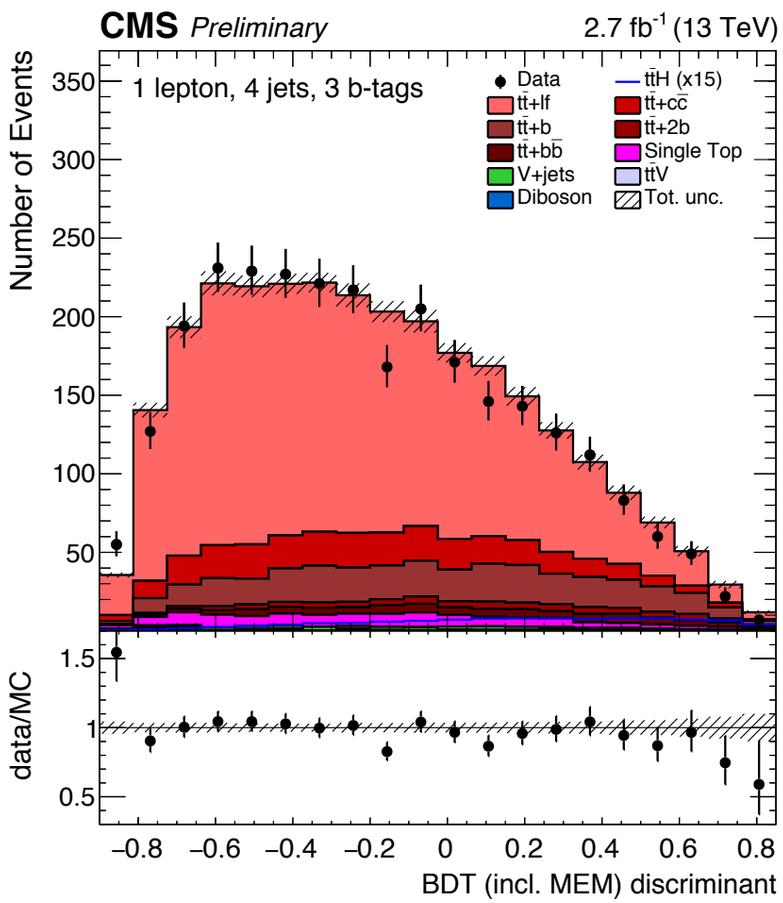
- **Object and event kinematics:** sum pt(jets, leptons), centrality(b-tags), inv. mass of jet pair closest to 125 GeV, mass, sum pt of closest b-tag pair, median mass of all jet pairs, mass, sum pt(jet, tag) with min ΔR

- **angular:** mean ΔR of all b-tags; min ΔR(closest b-tags, jets), max Δη(furthest b-tags), twist angle

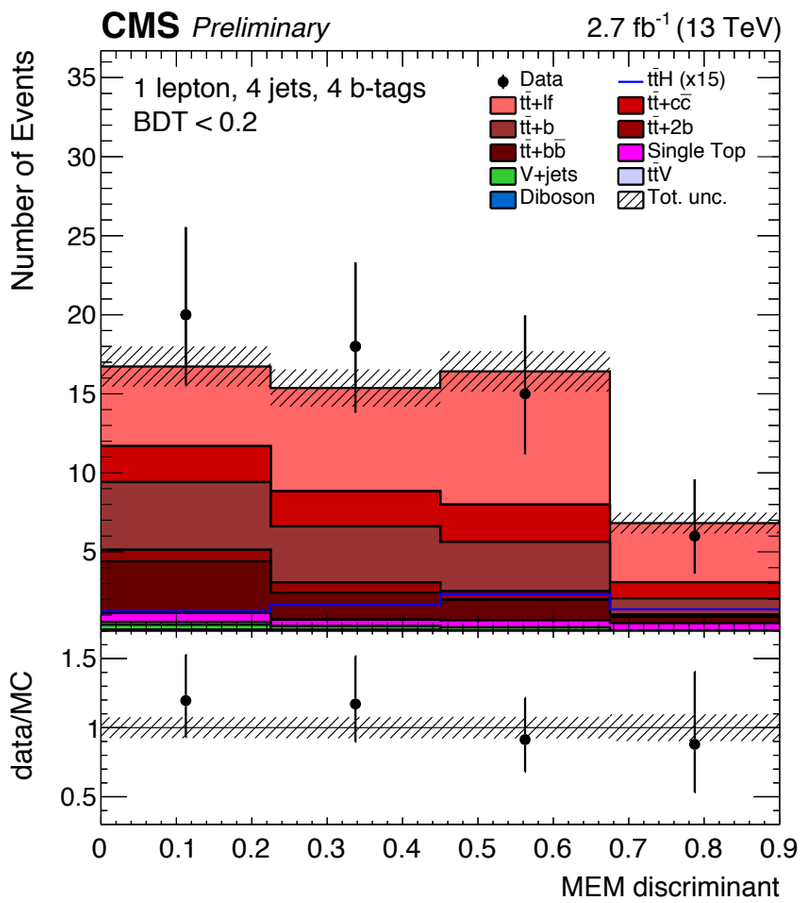
- **Event properties:** number of jets, average b-discriminant for (b-tagged) jets, Fox-Wolfram moments

lepton+jets categories

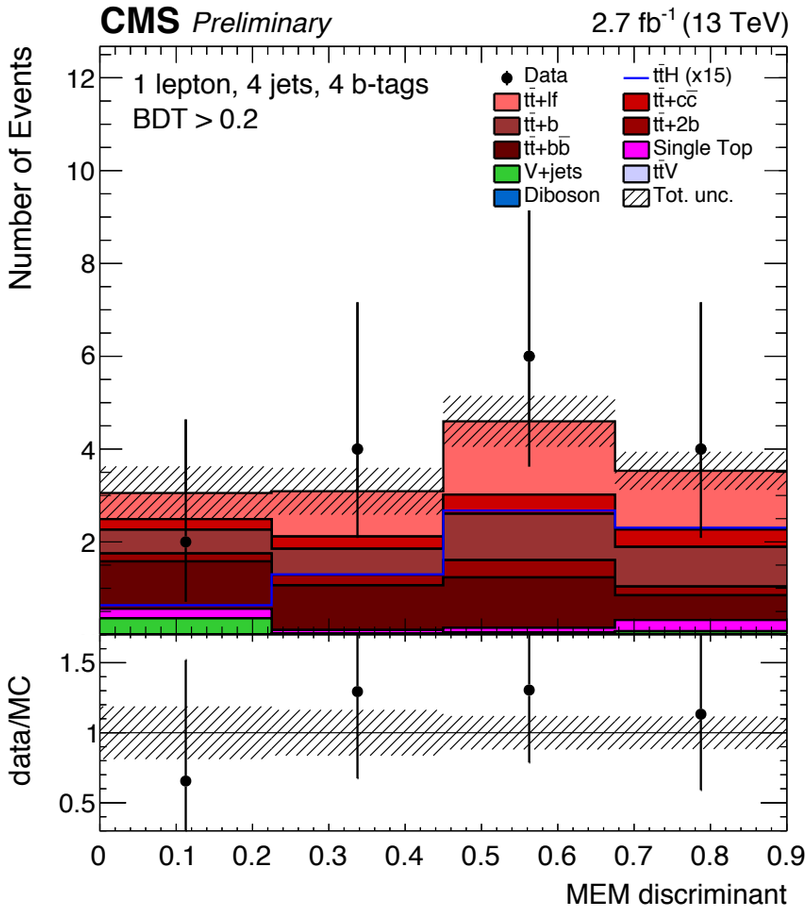
4j3t



4j4t BDT low

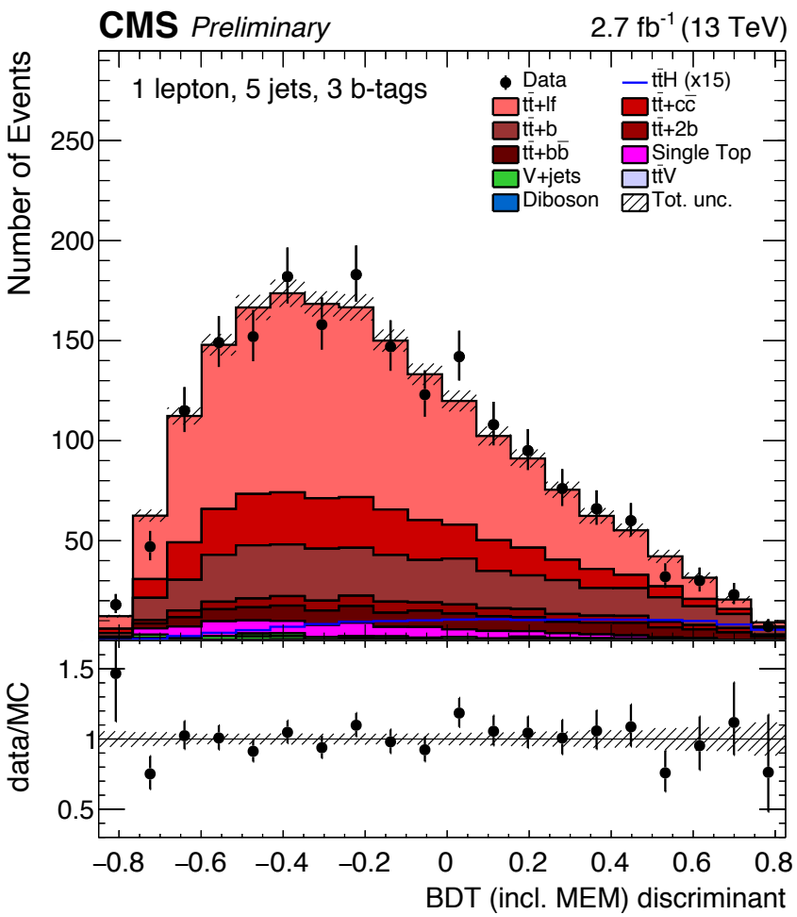


4j4t BDT high

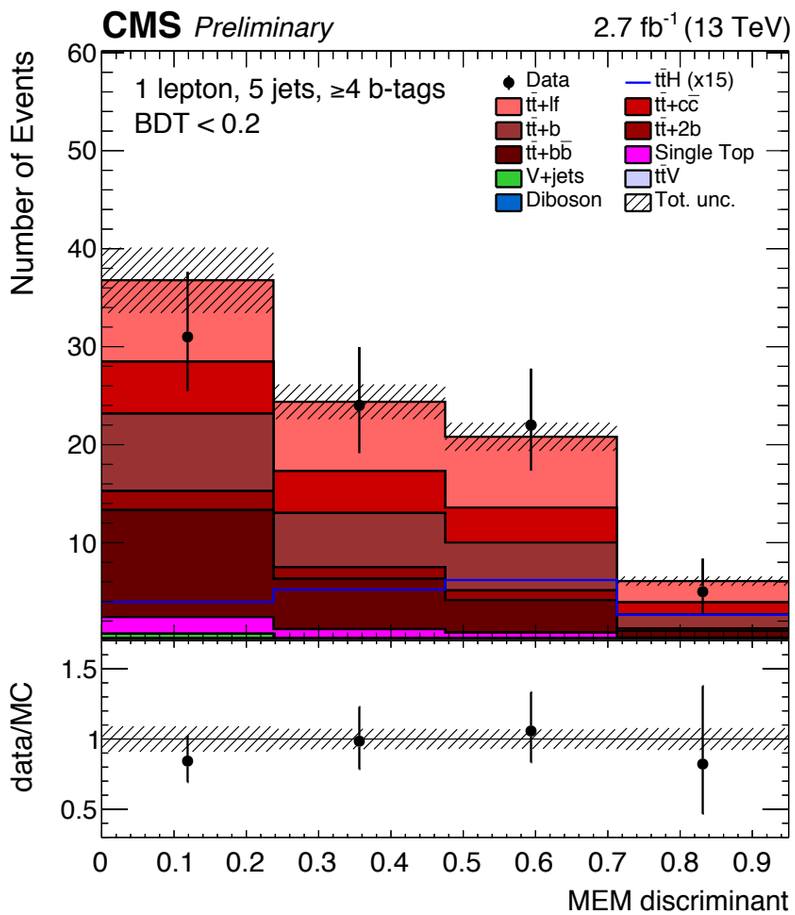


lepton+jets categories

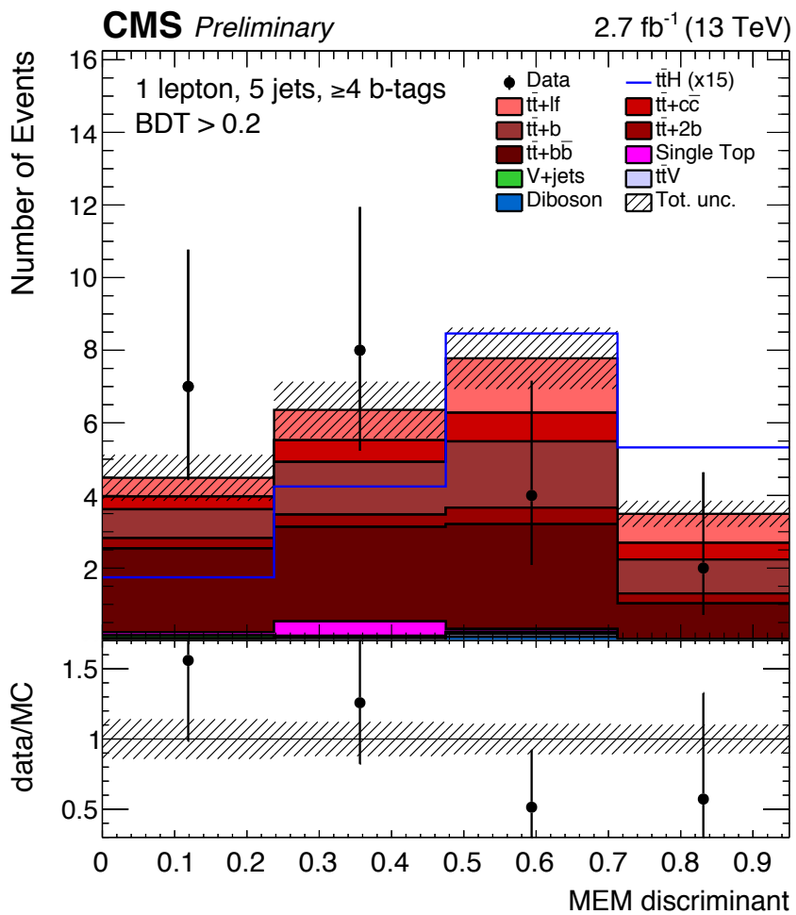
5j3t



5j4t BDT low

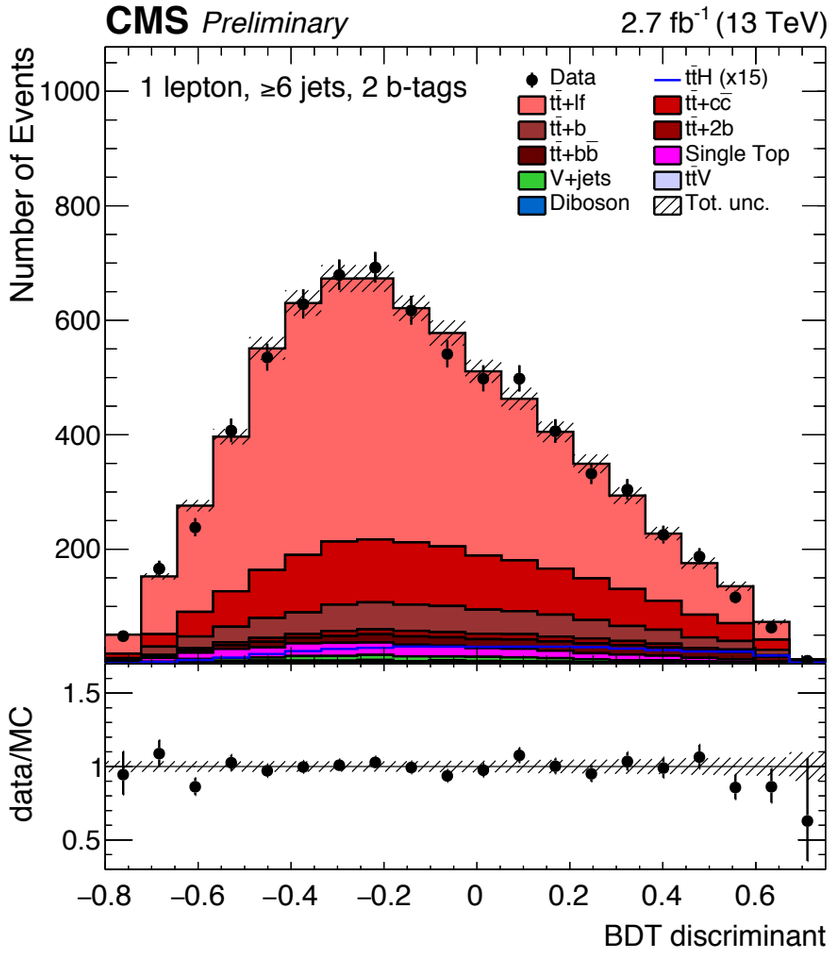


5j4t BDT high

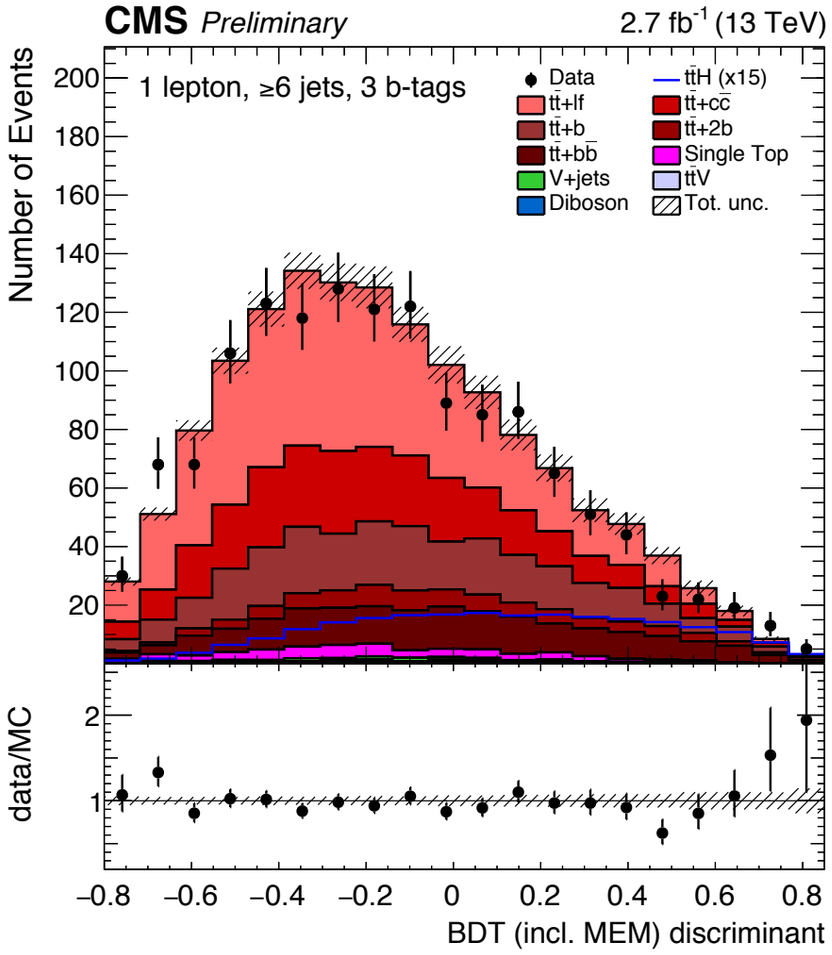


lepton+jets categories

6j2t BDT low

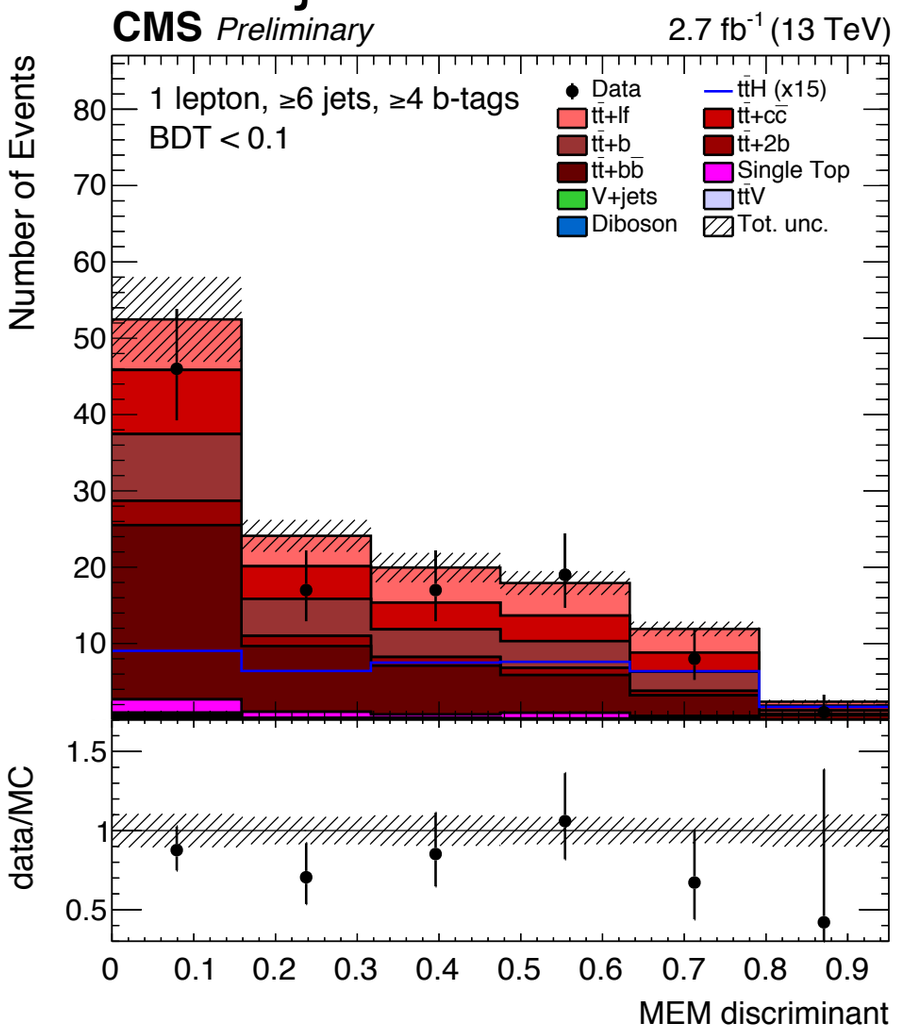


6j3t BDT high

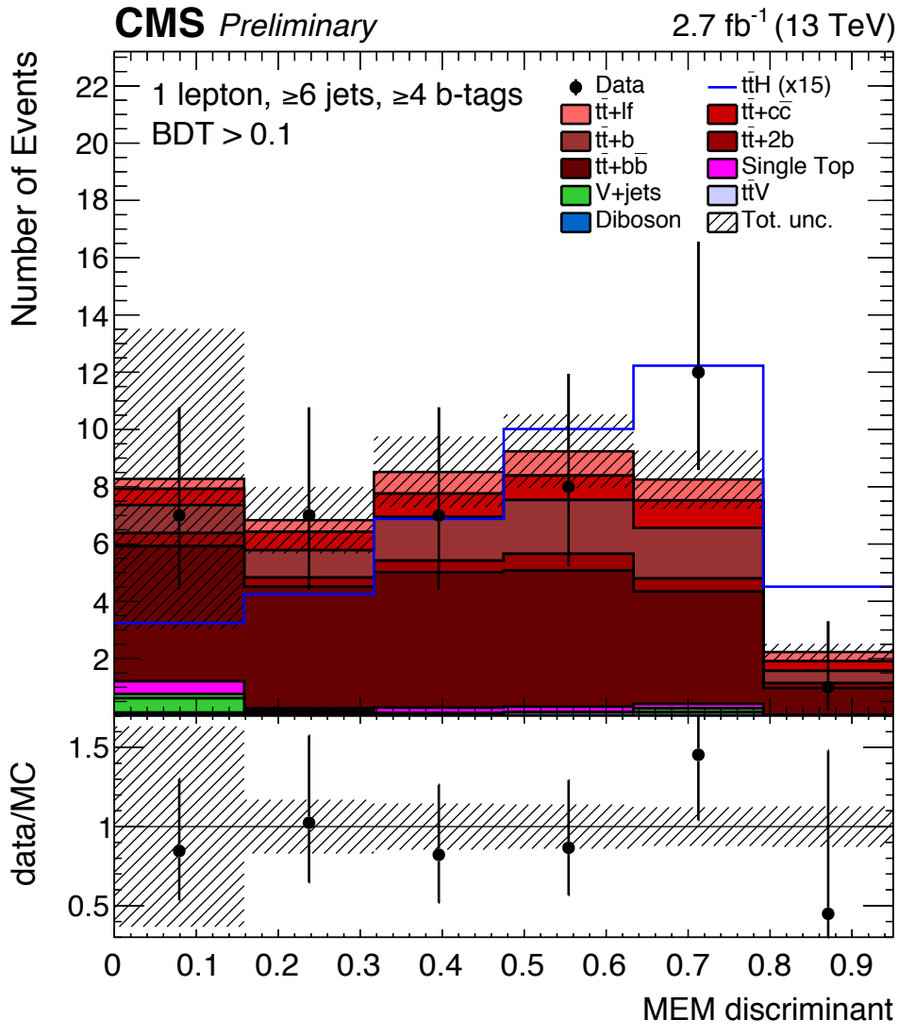


lepton+jets categories

6j4t BDT low

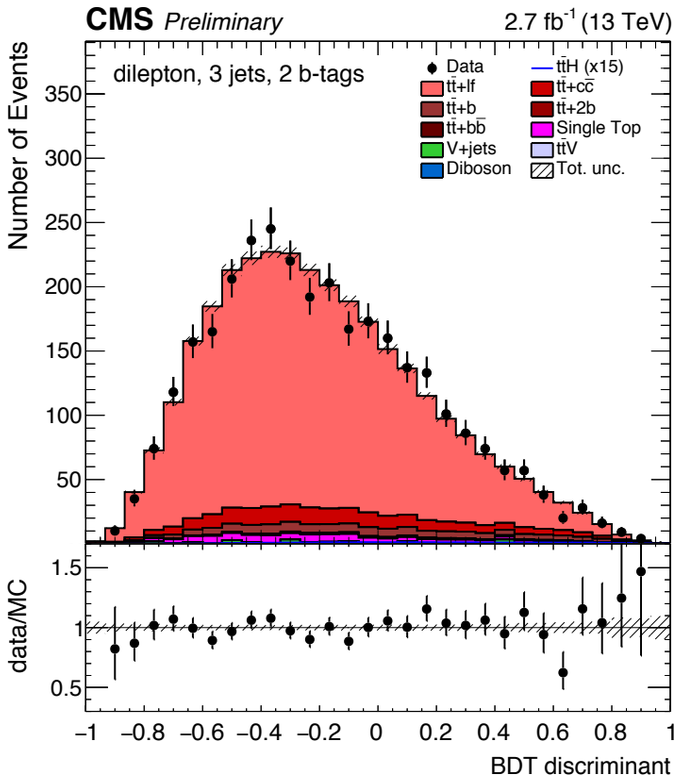


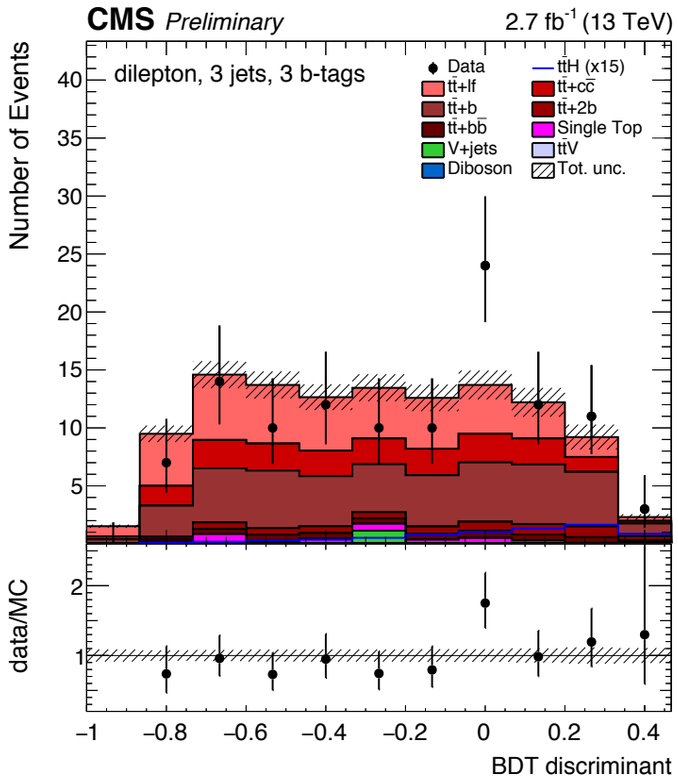
6j4t BDT high

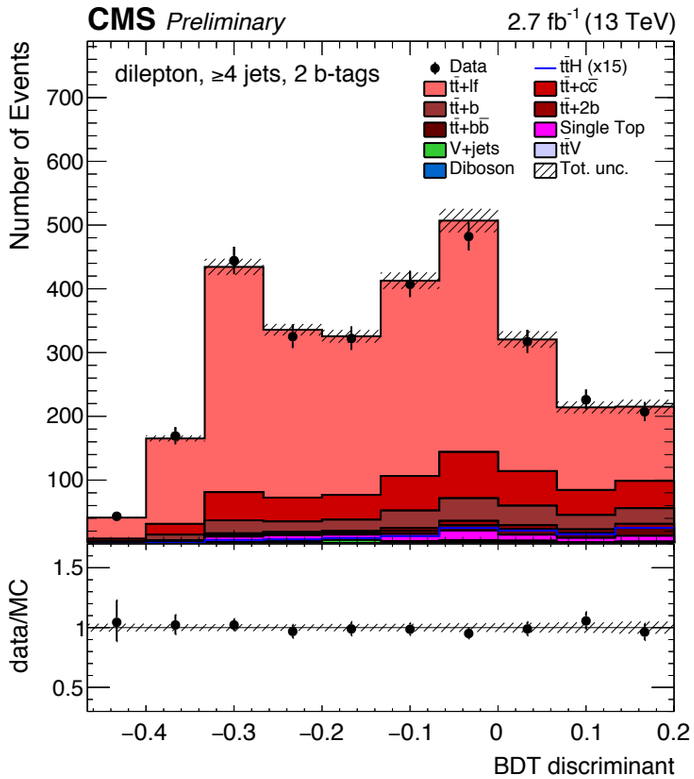


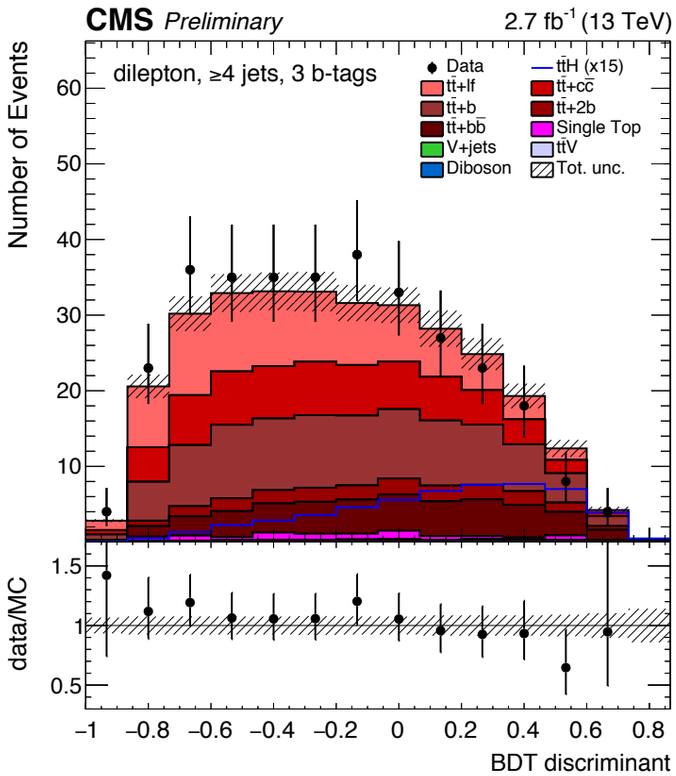
ETH Institute for Particle Physics

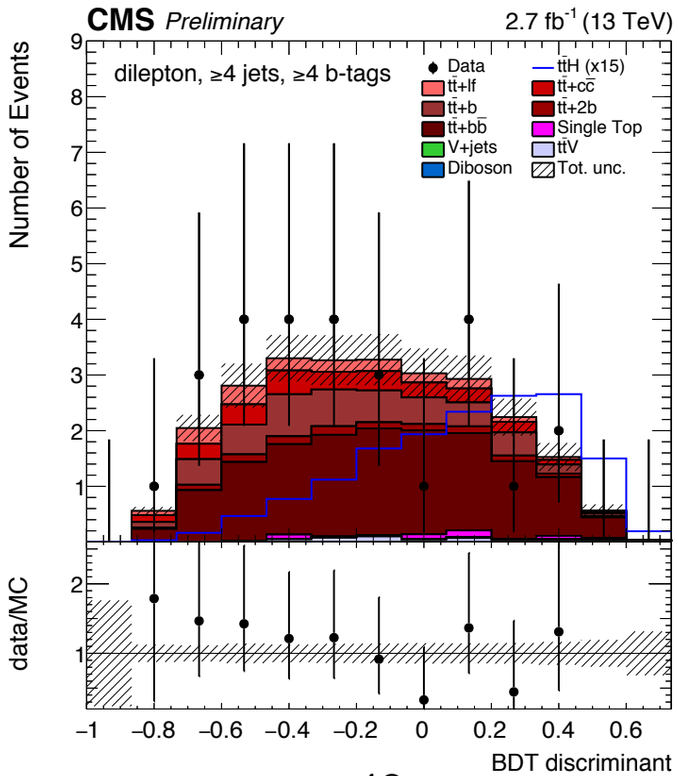
dilepton categories





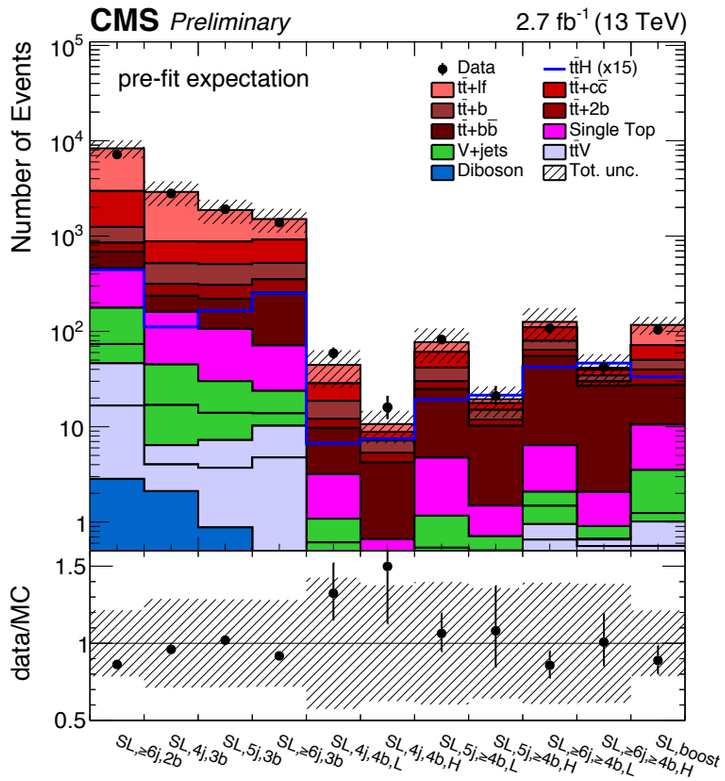


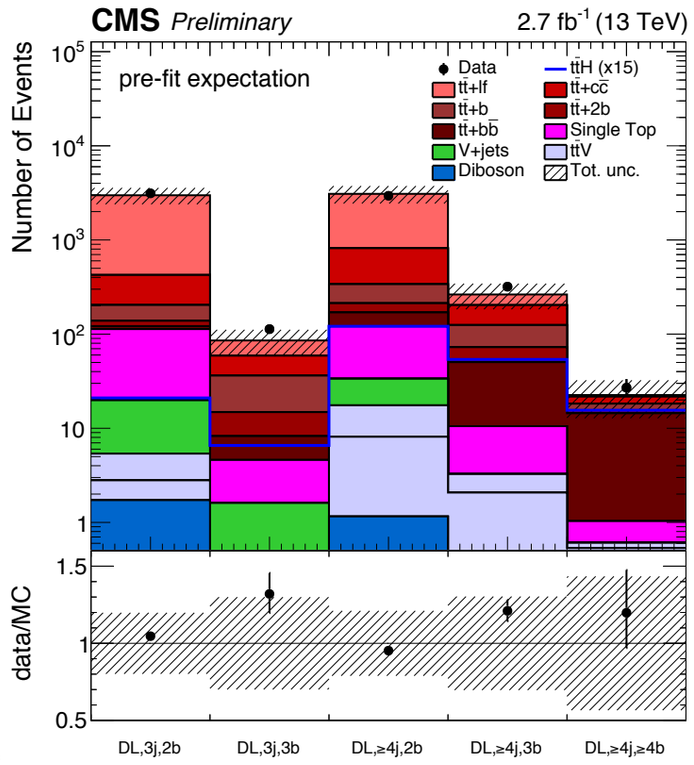




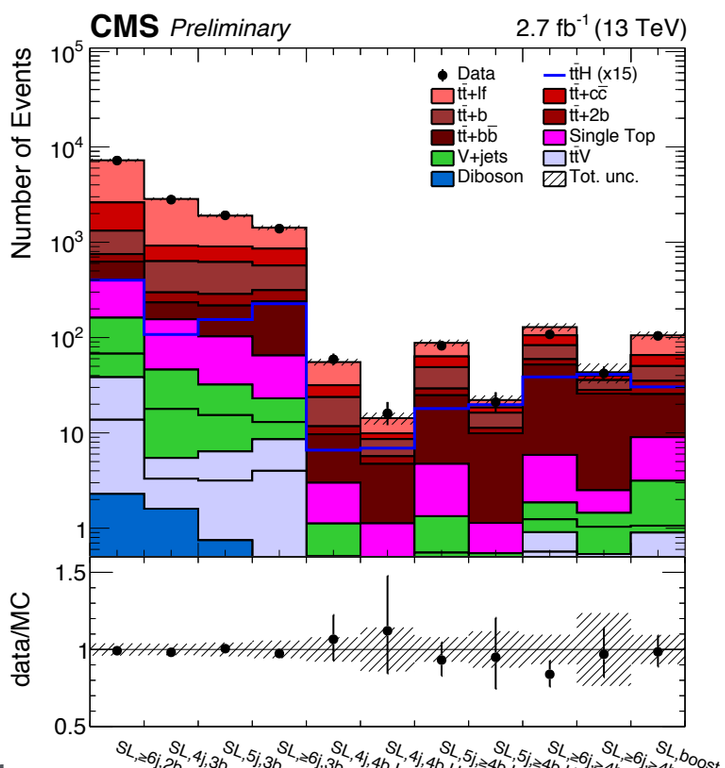
Pre-fit vs. post-fit yields

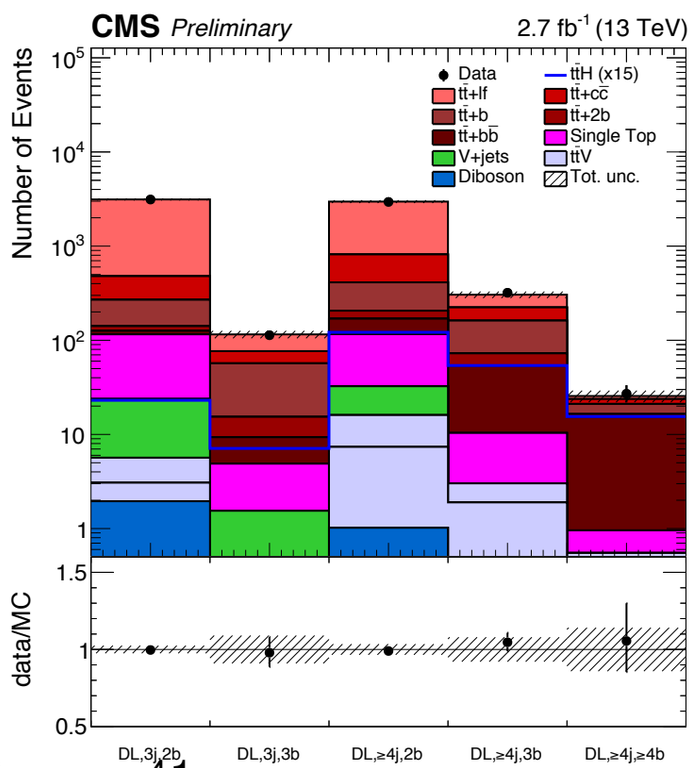
prefit





postfit

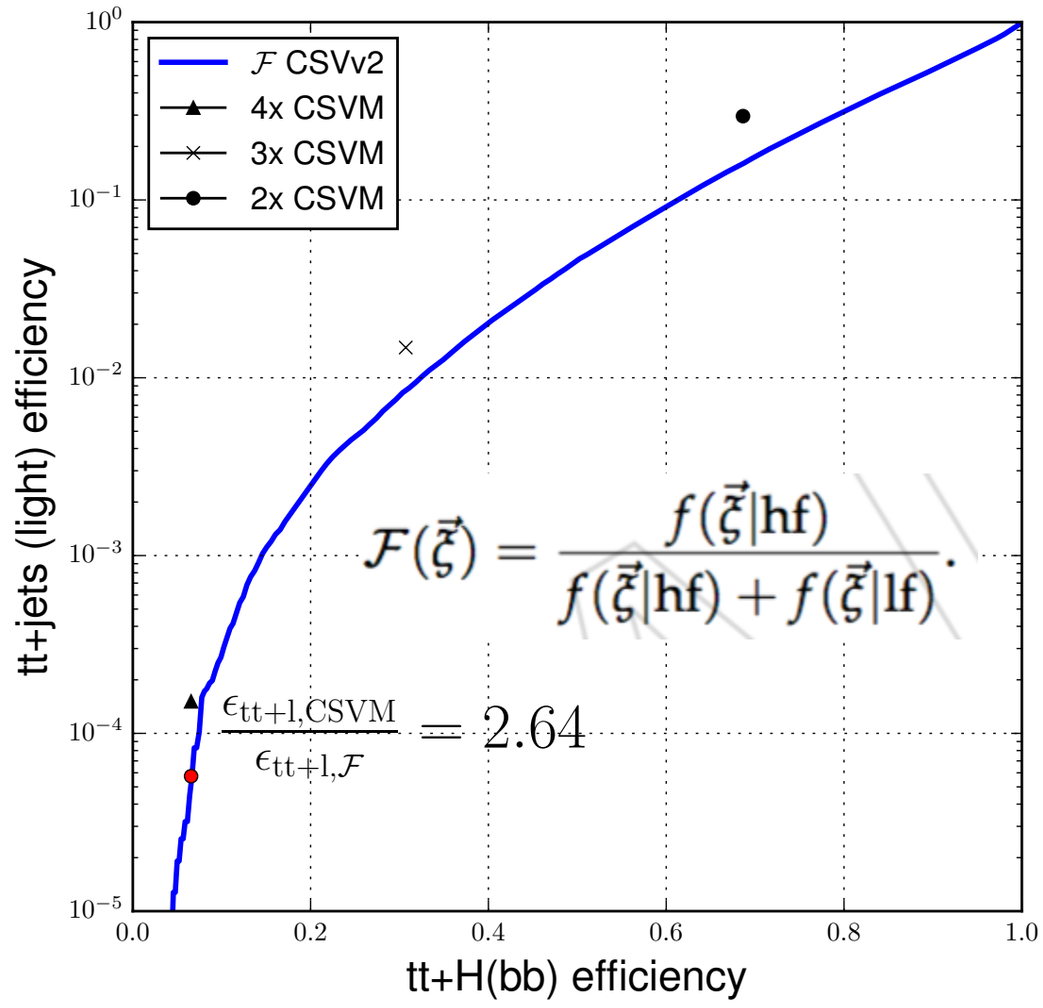




- Fit is able to accommodate observed yields without overt tension

b-tagging likelihood

- Separate tt+bb (4b -quark) from tt+lf (2 b-quark)
- Sum over permutations of associating 4 (2) b-quarks to jets, defined for ≥4 jet events
- Constructed from CSVv2 distributions extracted from tt+jets single-leptonic events
- Observe ~2x better tt+light rejection than flat N CSVM cut
- **Use blr cut to choose 4-b compatible events in 3-tag categories to deploy MEM**
- **Possible training input to BDT**



$$f(\vec{\xi}|\mathrm{t\bar{t}} + \mathrm{hf}) = \sum_{i_1} \sum_{i_2 \neq i_1} \cdots \sum_{i_6 \neq i_1,\ldots,i_5} \left\{ \prod_{k \in \{i_1,i_2,i_3,i_4\}} f_{\mathrm{hf}}(\xi_k) \prod_{m \in \{i_5,i_6\}} f_{\mathrm{lf}}(\xi_m) \right\},$$

"heavy" "light"

event yields

	6j2t	4j3t	5j3t	6j3t	boosted
$t\bar{t}$+lf	3989.4 ± 18.1	988.5 ± 8.6	520.4 ± 6.3	308.1 ± 4.8	52.2 ± 2.0
$t\bar{t} + c\bar{c}$	1378.3 ± 11.4	213.8 ± 4.0	225.2 ± 4.3	230.7 ± 4.3	26.8 ± 1.4
$t\bar{t}$+b	349.8 ± 5.9	129.9 ± 3.2	121.4 ± 2.9	112.2 ± 2.9	11.7 ± 0.9
$t\bar{t}$+2b	149.2 ± 3.7	47.7 ± 1.8	52.3 ± 1.9	64.9 ± 2.1	12.6 ± 1.0
$t\bar{t} + b\bar{b}$	223.1 ± 5.1	46.6 ± 1.8	77.5 ± 2.5	135.2 ± 3.5	16.9 ± 1.1
Single Top	220.8 ± 8.7	59.5 ± 4.2	38.6 ± 3.3	32.1 ± 3.1	5.6 ± 1.3
V+jets	92.8 ± 13.4	18.9 ± 4.4	11.5 ± 5.6	15.5 ± 6.2	5.0 ± 1.0
$t\bar{t}$+V	46.2 ± 1.5	3.8 ± 0.3	6.8 ± 0.5	10.2 ± 0.7	2.1 ± 0.3
Diboson	0.2 ± 0.1	0.5 ± 0.3	0.4 ± 0.3	0.4 ± 0.2	0.0 ± 0.0
Total bkg	6449.7 ± 28.1	1509.3 ± 12.0	1054.0 ± 10.9	909.4 ± 10.7	133.0 ± 3.4
$t\bar{t}H$	26.3 ± 0.2	4.6 ± 0.1	7.4 ± 0.1	12.0 ± 0.1	2.5 ± 0.1
Data	6161.0 ± 78.5	1422.0 ± 37.7	996.0 ± 31.6	894.0 ± 29.9	122.0 ± 11.0

	4j4tl	4j4th	5j4tl	5j4th	6j4tl	6j4th
$t\bar{t}$+lf	4.9 ± 0.6	1.0 ± 0.3	6.7 ± 0.7	0.9 ± 0.3	5.7 ± 0.6	1.8 ± 0.6
$t\bar{t} + c\bar{c}$	3.1 ± 0.4	0.4 ± 0.2	5.5 ± 0.6	1.0 ± 0.2	12.3 ± 1.0	2.9 ± 0.5
$t\bar{t}$+b	2.5 ± 0.4	0.6 ± 0.2	5.4 ± 0.6	1.0 ± 0.2	7.0 ± 0.7	1.7 ± 0.3
$t\bar{t}$+2b	0.6 ± 0.2	0.3 ± 0.1	2.7 ± 0.5	0.6 ± 0.2	4.7 ± 0.5	1.3 ± 0.3
$t\bar{t} + b\bar{b}$	2.4 ± 0.4	1.0 ± 0.2	10.9 ± 0.8	3.7 ± 0.5	26.5 ± 1.4	9.7 ± 0.8
Single Top	0.2 ± 0.1	0.2 ± 0.1	1.8 ± 0.7	1.0 ± 0.5	2.3 ± 0.7	0.1 ± 0.1
V+jets	0.0 ± 0.0	0.0 ± 0.0	0.0 ± 0.0	0.0 ± 0.0	0.5 ± 0.3	0.1 ± 0.1
$t\bar{t}$+V	0.1 ± 0.1	0.0 ± 0.0	0.4 ± 0.1	0.2 ± 0.1	1.2 ± 0.2	0.4 ± 0.1
Diboson	0.0 ± 0.0	0.0 ± 0.0	0.0 ± 0.0	0.0 ± 0.0	0.0 ± 0.0	0.4 ± 0.4
Total bkg	13.9 ± 1.0	3.6 ± 0.5	33.5 ± 1.6	8.5 ± 0.9	60.2 ± 2.1	18.2 ± 1.2
$t\bar{t}H$	0.2 ± 0.0	0.2 ± 0.0	0.7 ± 0.0	0.7 ± 0.0	1.5 ± 0.0	1.6 ± 0.0
Data	23.0 ± 4.8	5.0 ± 2.2	41.0 ± 6.4	6.0 ± 2.4	60.0 ± 7.7	12.0 ± 3.5