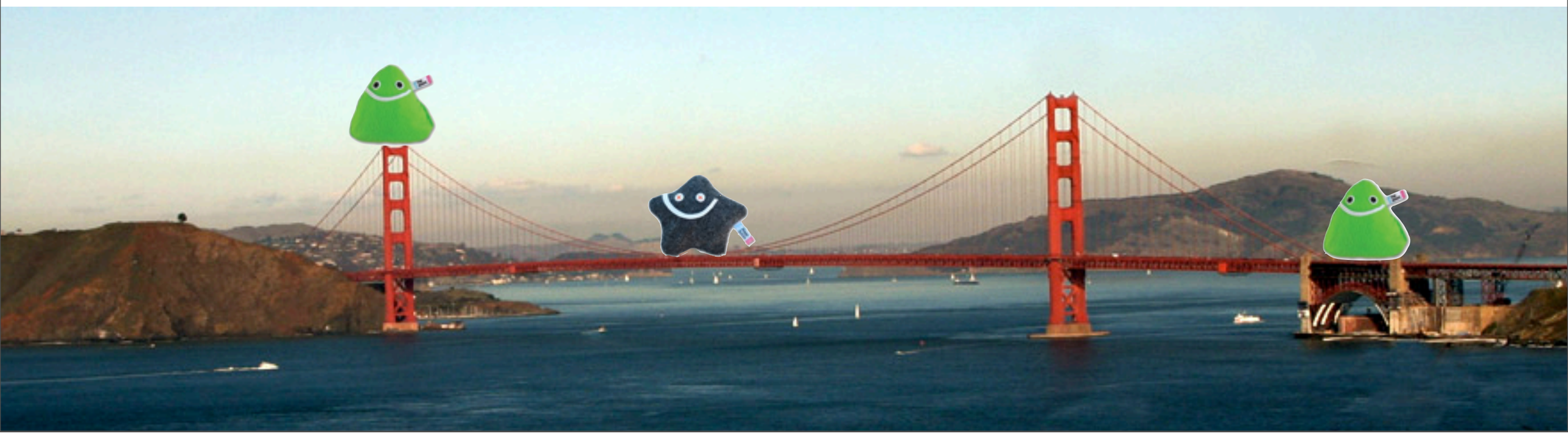
ETH zürich

Searches for ttH(bb) production

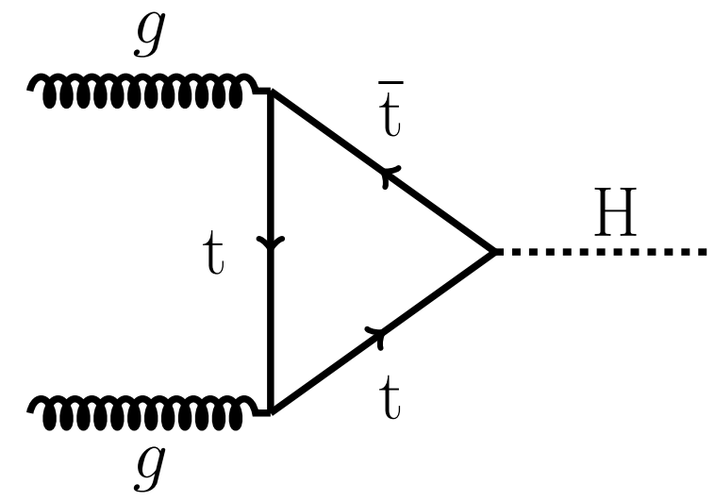
Gregor Kasieczka for the CMS Collaboration

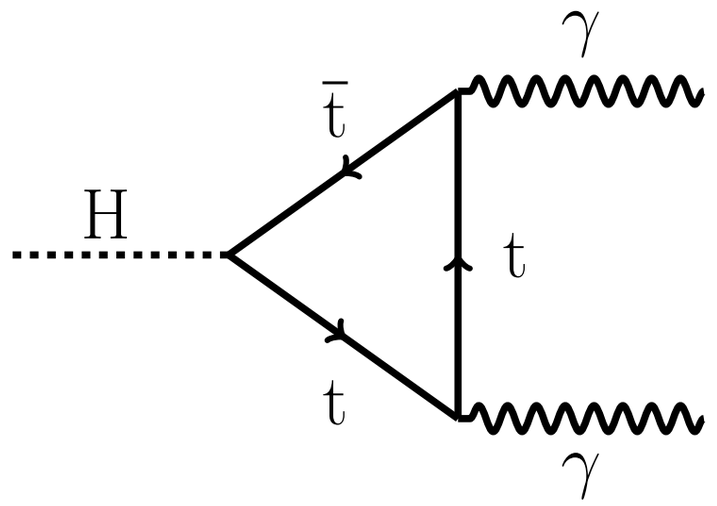
Higgs Couplings 2016



Why ttH?

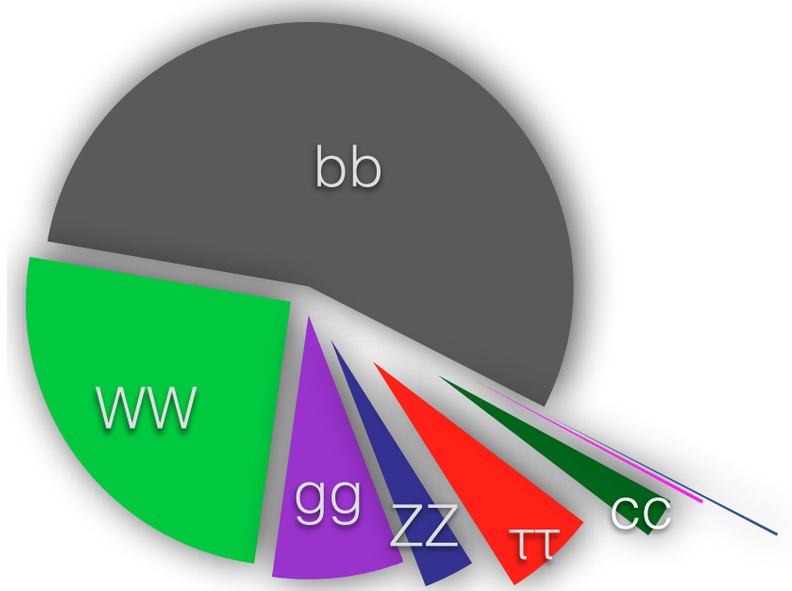
- Large top quark mass:
  - expect top quark Yukawa (Y_t) coupling to be of order one
  - too large to observe Higgs-to-top decay
- Y_t also contributes to
  - fermion loop in H production from gluons
  - H decay to photons
- But due to potential BSM physics contributing to the loop they can't be used to measure Y_t directly
- ttH allows direct measurement of Y_t
  - Deviation from SM prediction could hint at heavy top-partners

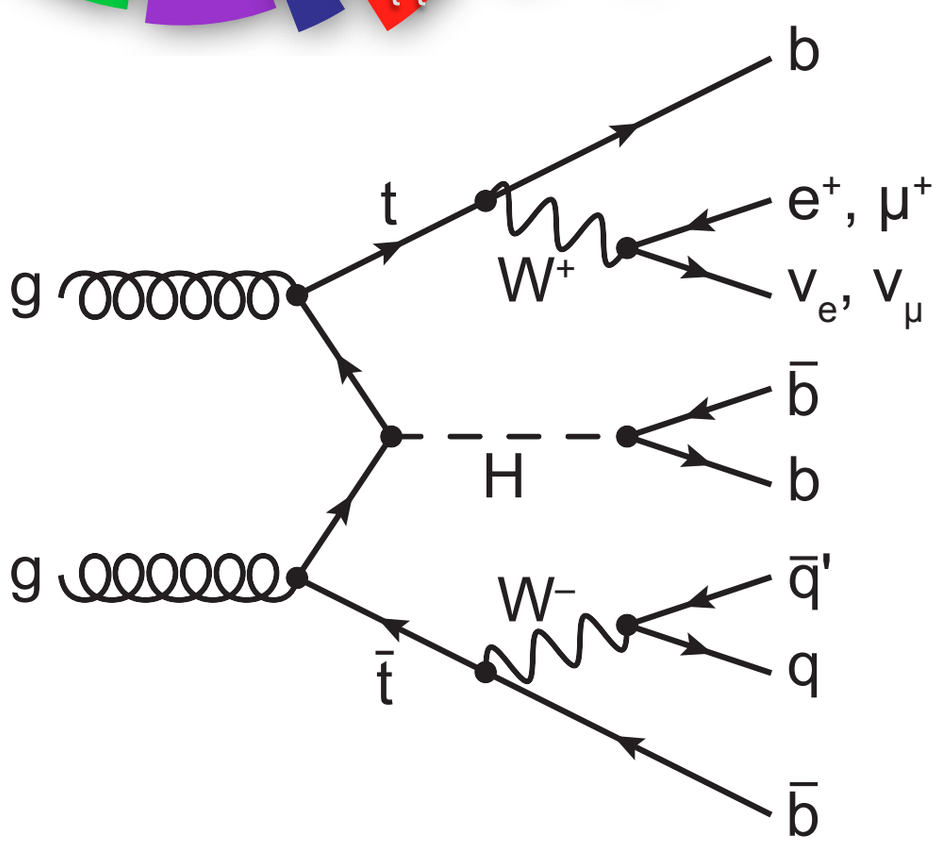




Why H(bb)?

- Advantages:
  - Highest branching fraction
  - Fermion-only production and decay
- Disadvantages:
  - complicated final state
  - large backgrounds
  - irreducible ttbb background has large theory uncertainties





Overview

- Up to **12.9 fb^{-1}** at 13 TeV
- Details: CMS PAS HIG-16-038
- Strategy:

I. Select events compatible with tt→dilepton / lepton+jets and H → bb decay

II. Categorize according to jet and b-tag multiplicity

III. Discriminate between signal/background with BDTs and MEM

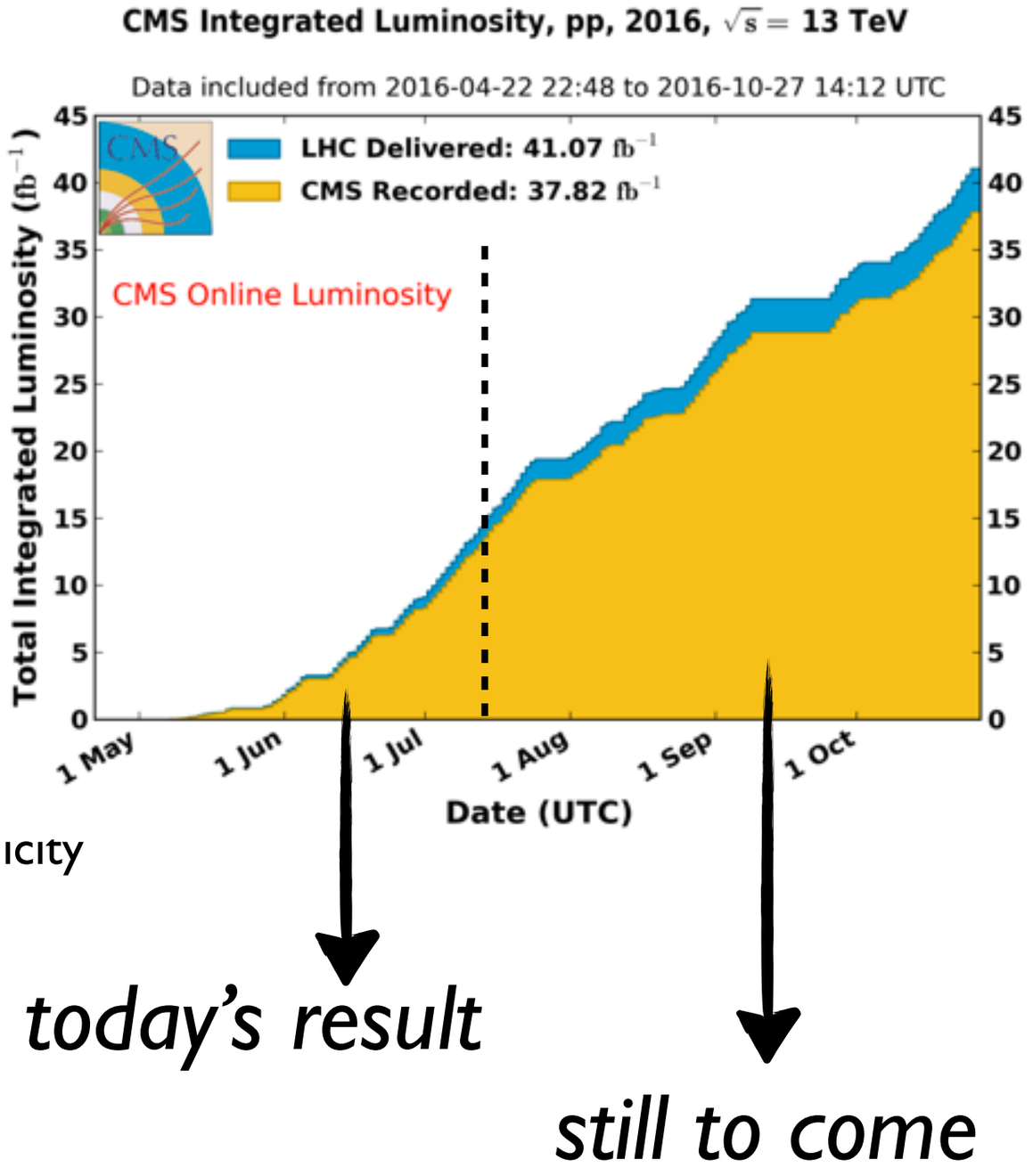
IV. Extract limit by combined fit to data in all categories

I: Event Selection

- Jets: anti-k_T (R=0.4) using CHS input
  - $\geq$ 4 with p_T > 30 GeV for Lepton+Jets
  - $\geq$ 2 with p_T > 30 GeV (p_T > 20 GeV subleading) for Dilepton
- Leptons:
  - Exactly one electron (p_T>30 GeV) OR muon (p_T>25 GeV) for Lepton+Jets
  - Exactly two opposite-sign leptons. Leading p_T > 25 GeV
    sub-leading p_T>15 GeV
- Corrections:
  - Pile up reweighting
  - Lepton scale factors
  - Jet energy scale and resolution corrections
  - b-tag discriminator

1: Samples

- Data:
  - Lepton+Jets: 12.9 fb^{-1}
  - Dilepton: 12.9 fb^{-1} (μμ, eμ) / 11.4 fb^{-1} (ee)
- Signal and tt background MC:
  - Powheg Box V2 + Pythia 8
  - New CMS tune: CUETP8M2T4
    - α_{ISR} = 0.118 and h_{damp}=272.2
    - Improves data/MC agreement for jet multiplicity

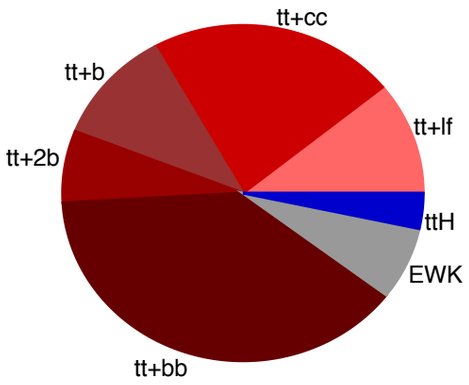


II: Categories

CMS *Simulation*

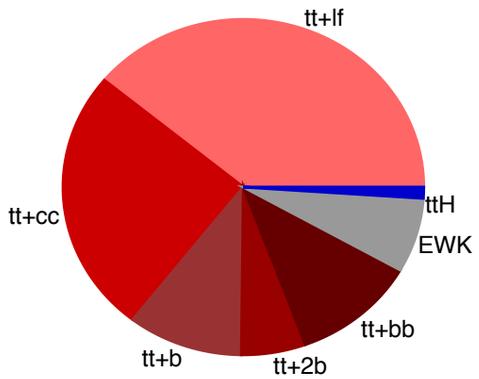
Lepton+Jets Channel

≥ 6 jet, ≥ 4 b-tags



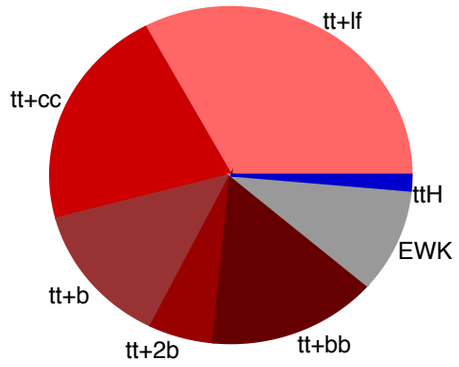
S/B=0.035, S/√B=0.973

≥ 6 jets, 3 b-tags



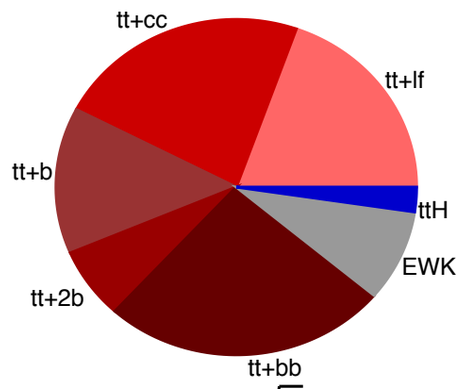
S/B=0.011, S/√B=0.895

4 jets, 4 b-tags



S/B=0.015, S/√B=0.242

5 jets, ≥ 4 b-tags

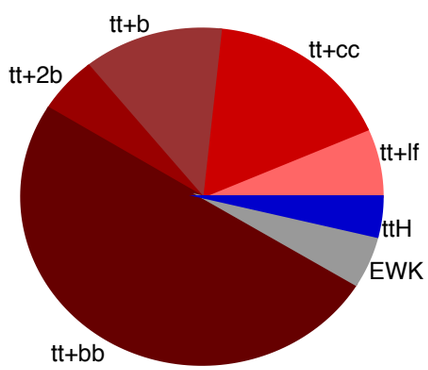


S/B=0.024, S/√B=0.532

CMS *Simulation*

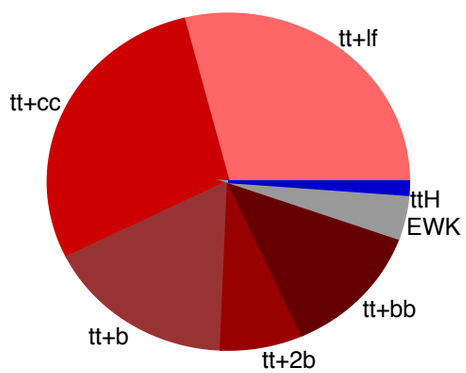
Dilepton Channel

≥ 4 jets, ≥ 4 b-tags



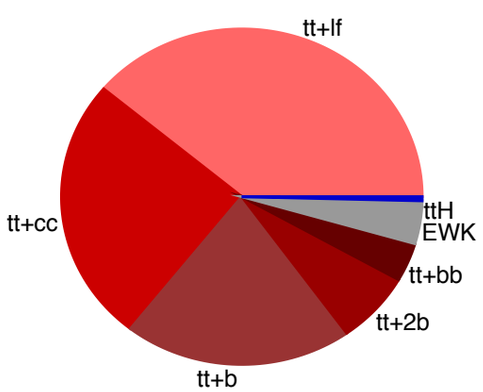
S/B=0.040, S/√B=0.417

≥ 4 jets, 3 b-tags



S/B=0.012, S/√B=0.453

3 jets, 3 b-tags



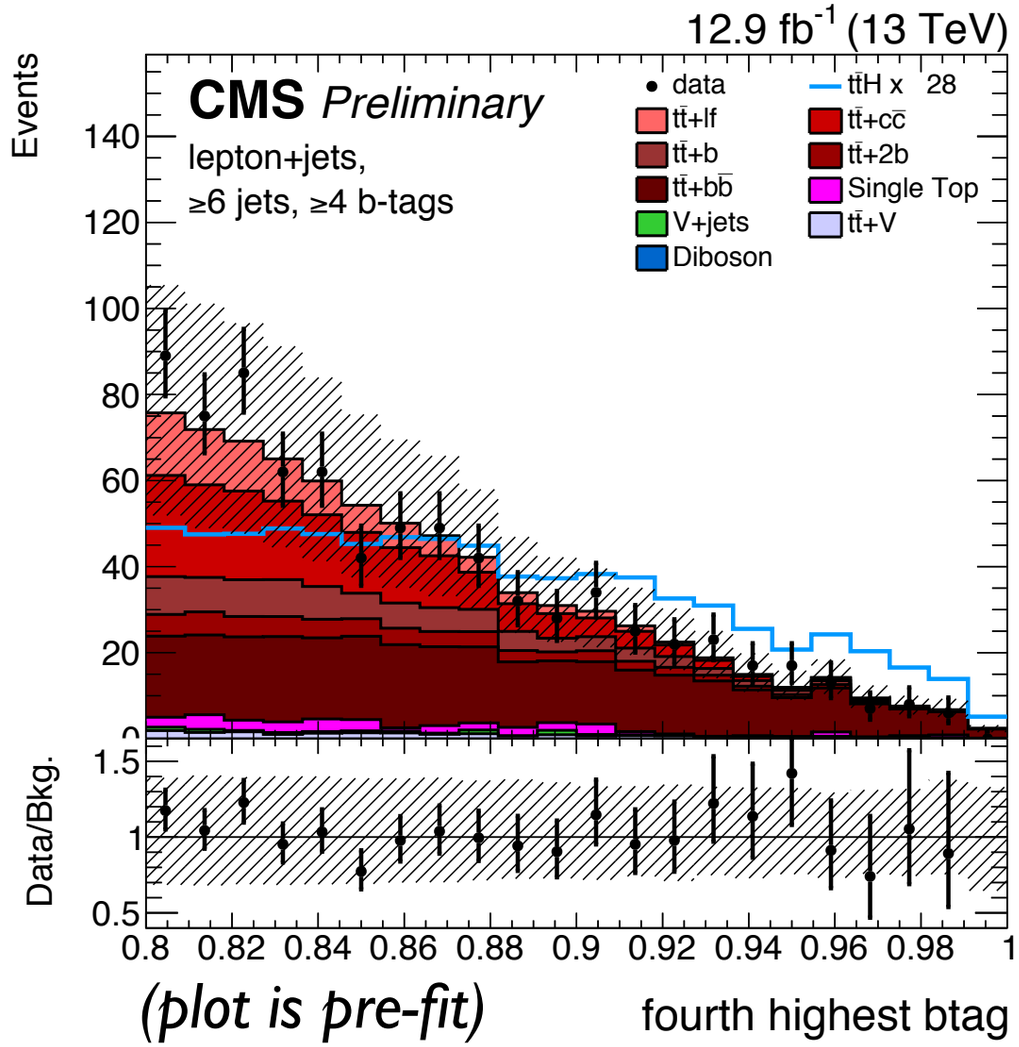
S/B=0.004, S/√B=0.084

III: Boosted Decision Tree

- Gradient boosting decision trees
- Train for ttH(bb) vs sum-of-backgrounds
- Individual training and variable selection for each jet/b-tag category
- Hyperparameters and selection of variables optimized using particle swarm algorithm
- Only consider well modeled variables

Example:
BDT input in Lepton+jets ≥6 jets ≥4 tags:

- best Higgs mass
- M_2(tag, tag) closest to 125 GeV
- M(jets, lepton, MET)
- 4th and 5th highest b-tag discriminator score
- ∑ p_T (jets, lepton, MET)



III: Matrix Element for ttH(bb) vs ttbb

Numerical integration

Momentum conservation

Resolution function (allow ISR)

$$w(\vec{y}|\mathcal{H}) = \sum_{i=1}^{N_a} \int \frac{\mathrm{d}x_a \mathrm{d}x_b}{2x_a x_b s} \int \prod_{k=1}^{8} \left(\frac{\mathrm{d}^3\vec{p}_k}{(2\pi)^3 2E_k} \right) (2\pi)^4 \delta^{(E,z)} \left(p_a + p_b - \sum_{k=1}^{8} p_k \right) \mathcal{R}^{(x,y)} \left(\vec{p}_{\mathrm{T}}, \sum_{k=1}^{8} p_k \right) \times$$

$$g(x_a, \mu_{\mathrm{F}}) g(x_b, \mu_{\mathrm{F}}) |\mathcal{M}_{\mathcal{H}}(p_a, p_b, p_1, \ldots, p_8)|^2 W(\vec{y}, \vec{p}) \,,$$

Parton density functions

LO Scattering amplitude (Open Loops)

Detector transfer function

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

Construct per-event signal/background probability
using full kinematic information in an analytical approach

Works best for final states with many reconstructed objects.

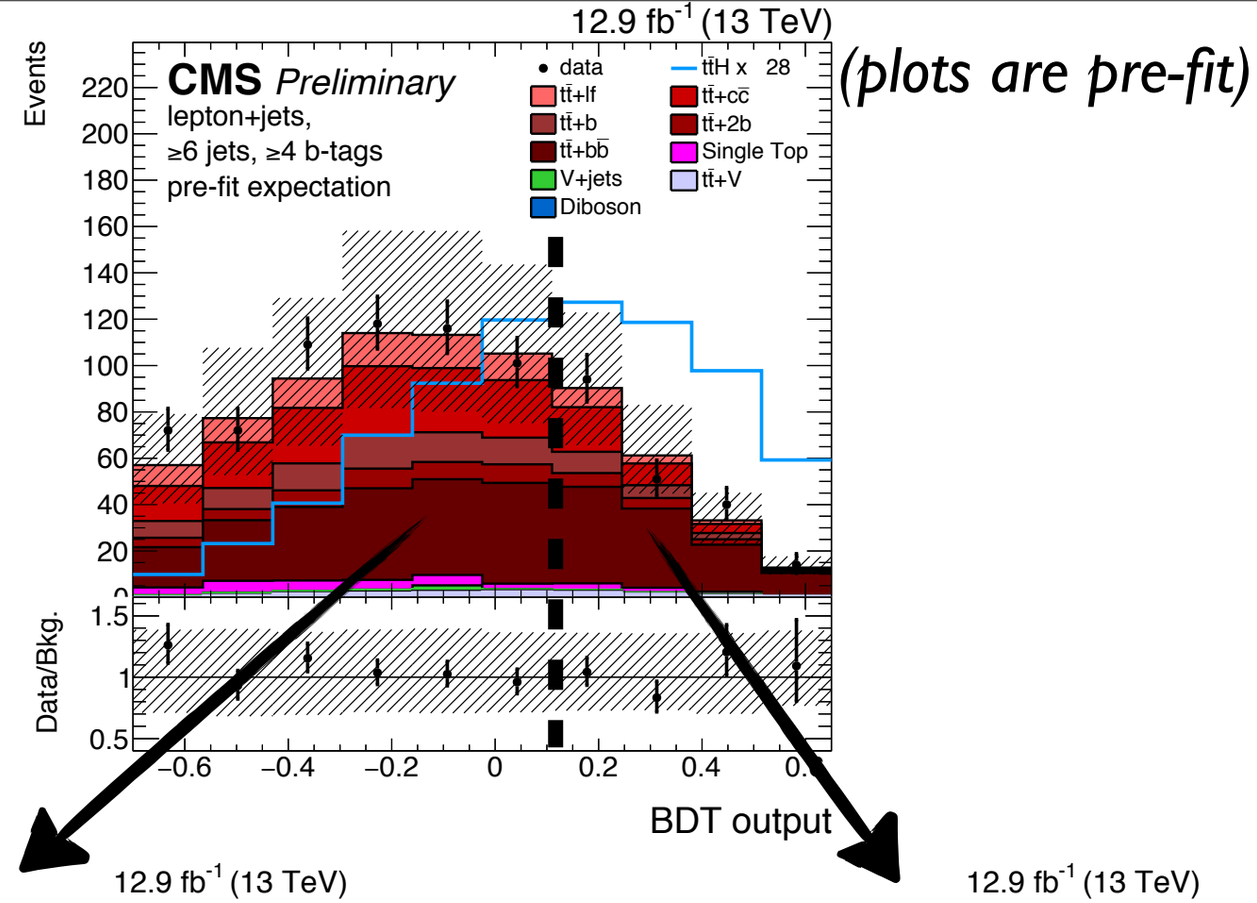
Possible but computationally expensive to *integrate* over missing jets

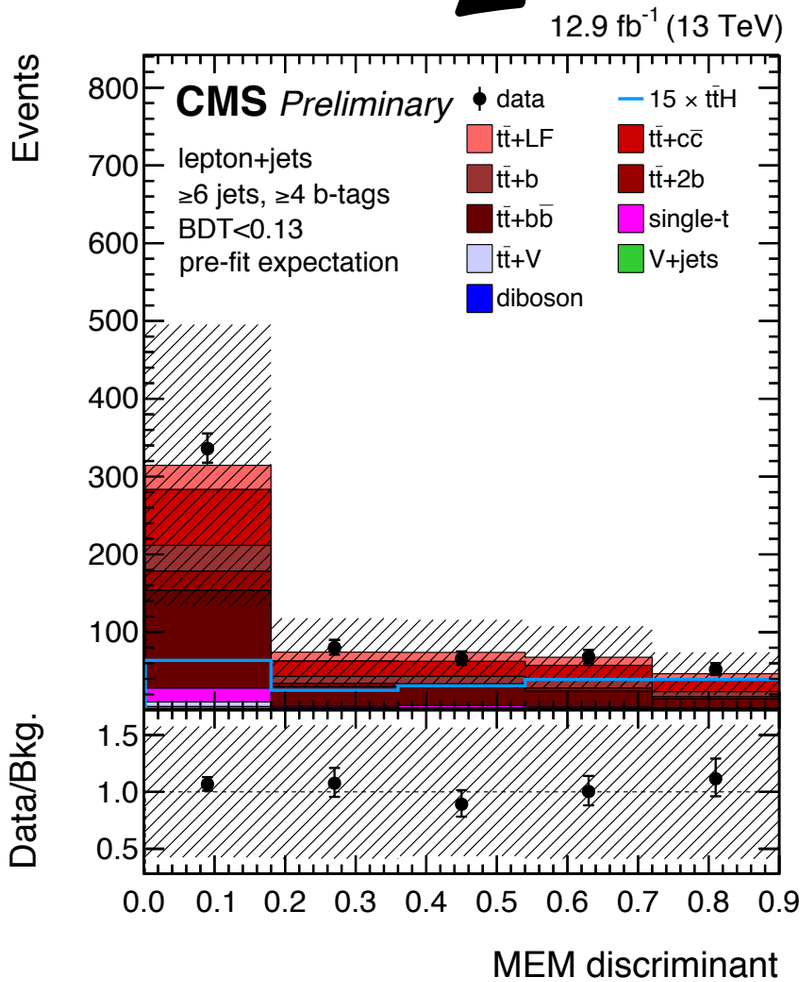
IV: Lepton+Jets fit

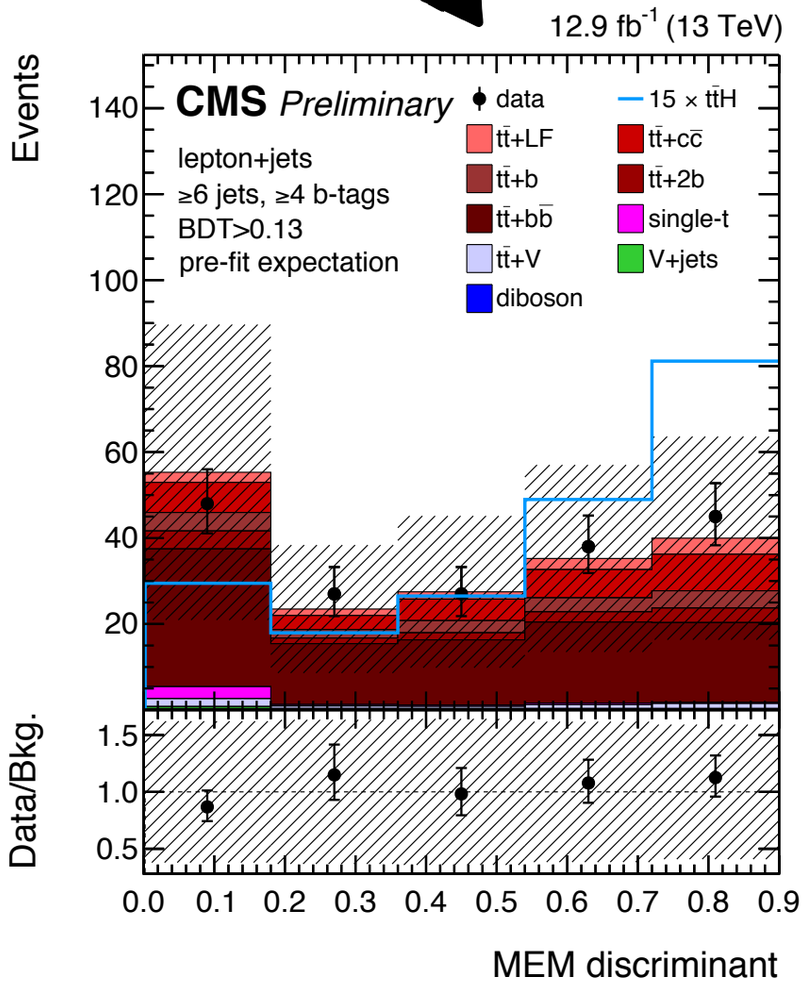
	4 jets	5 jets	≥6 jets
3 tags			low/high BDT
≥4 tags	low/high BDT	low/high BDT	low/high BDT

- Divide events by jet/tag multiplicity
- Subdivide by low/high BDT
- Fit MEM discriminant

(plots are pre-fit)



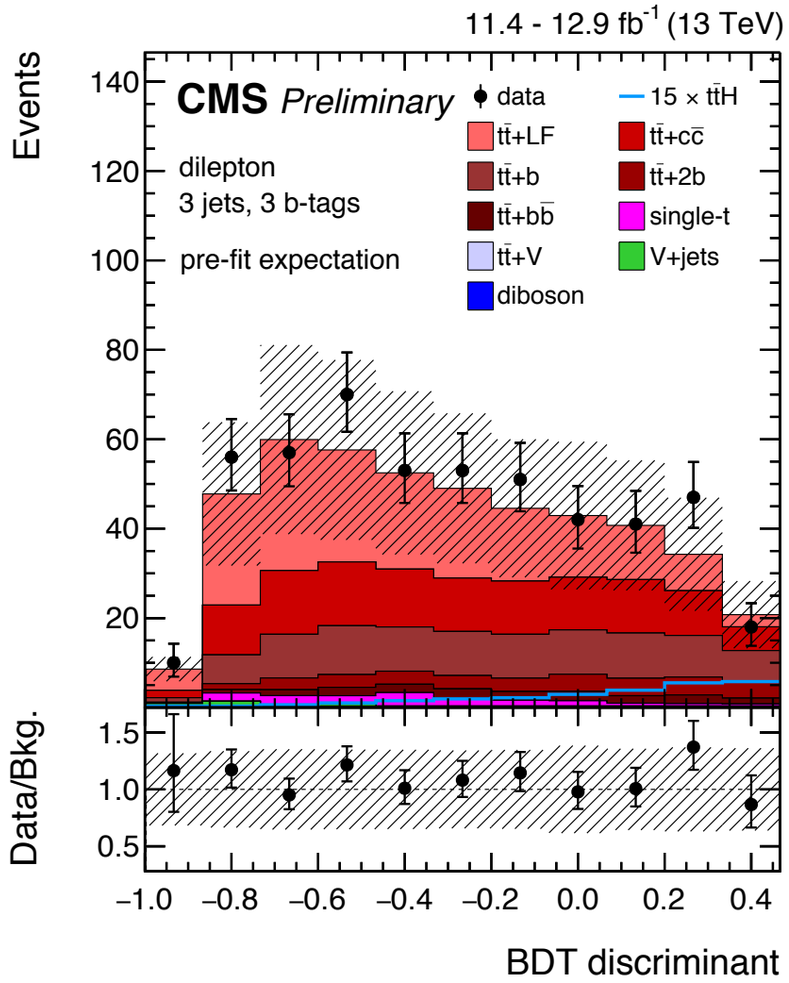


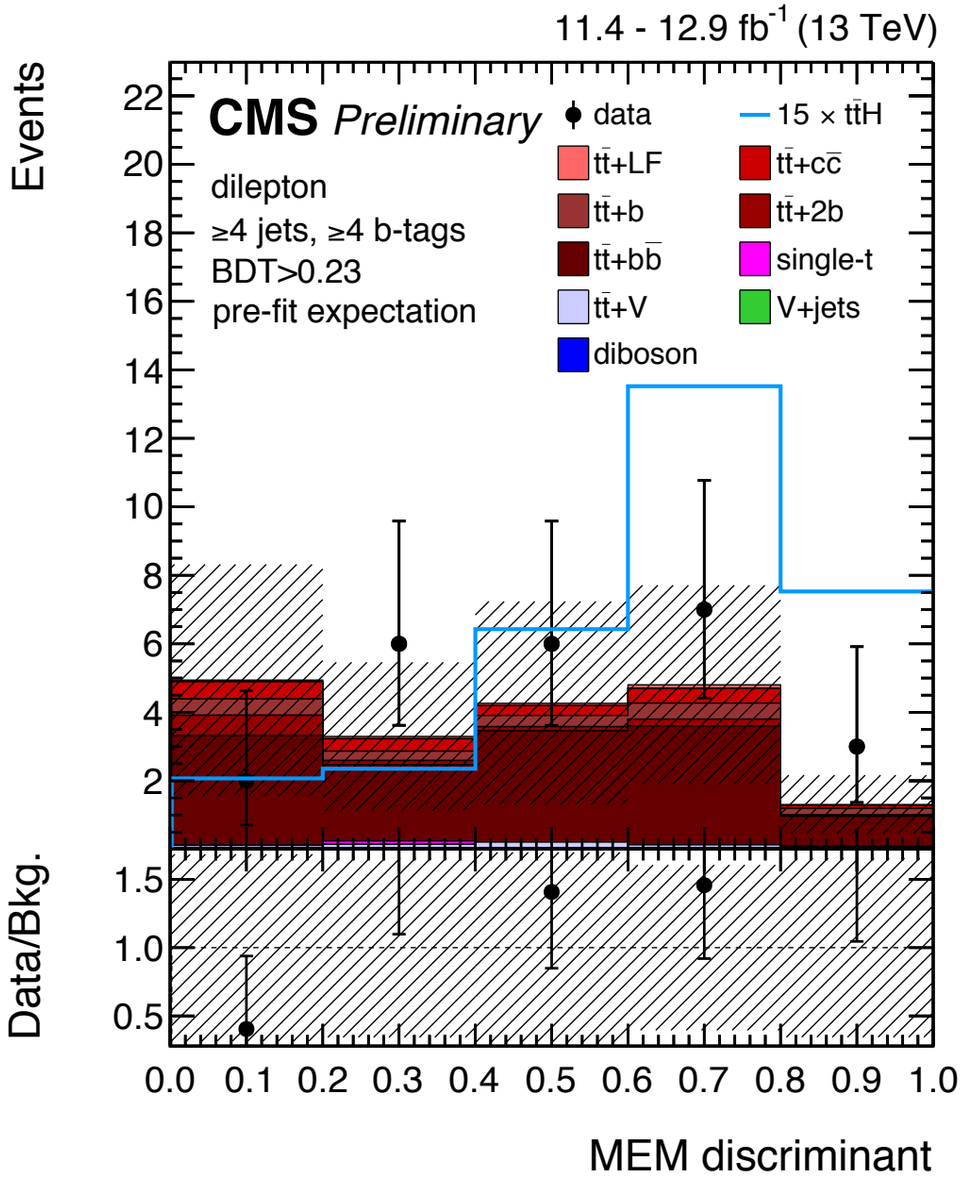


IV: Dilepton fit

(plots are pre-fit)

	3 jets	≥4 jets
3 tags	BDT only	low/high BDT
≥4 tags		low/high BDT



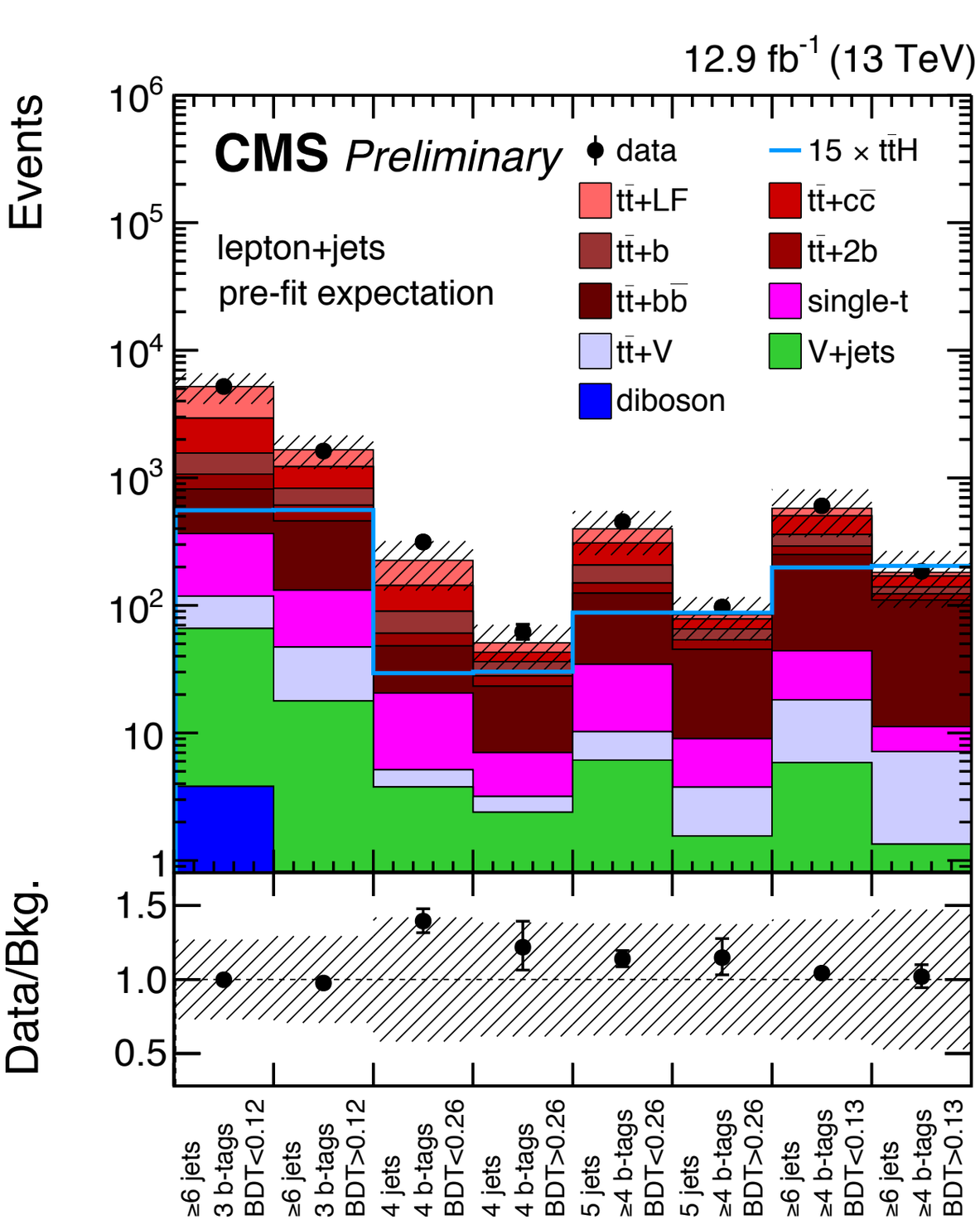


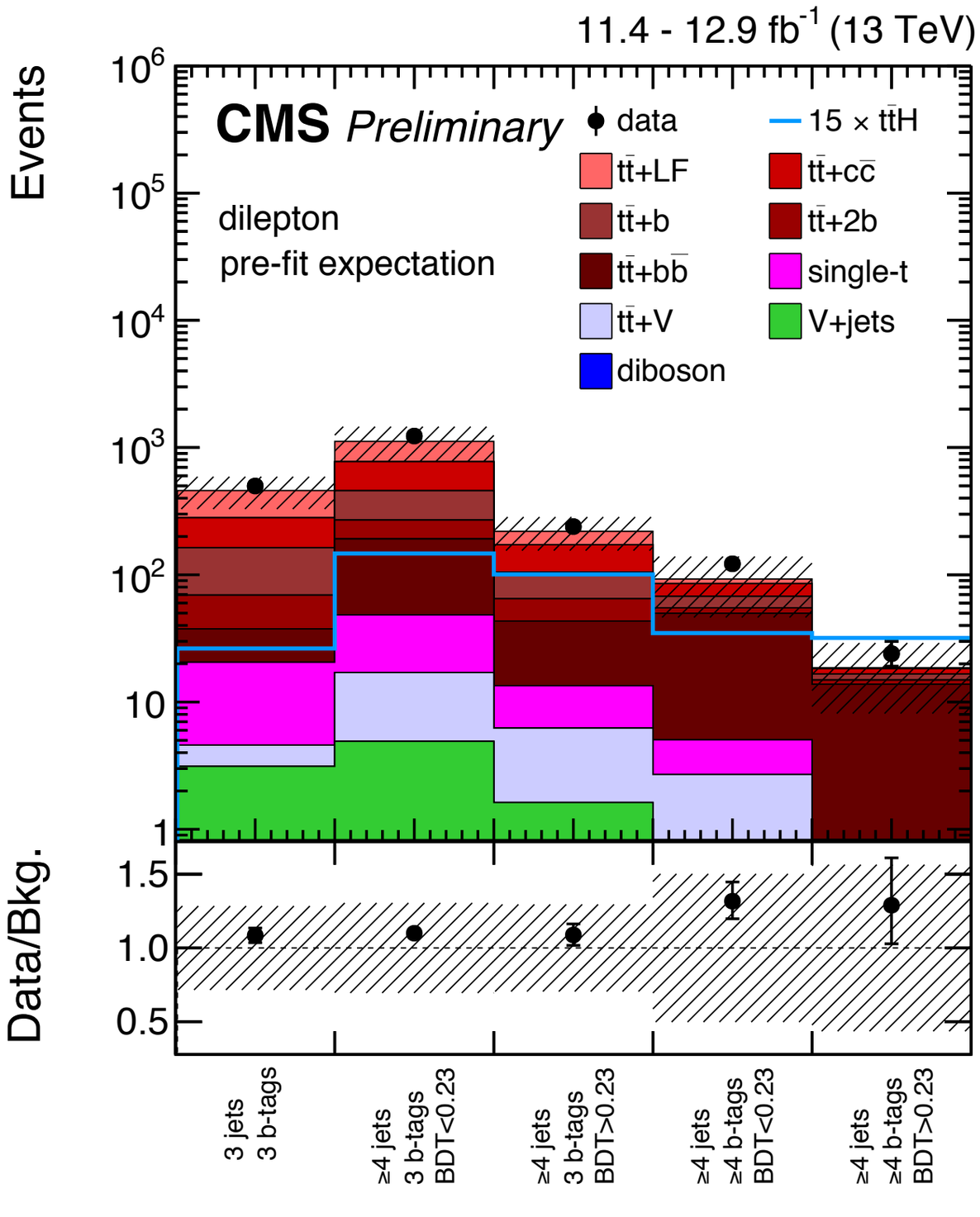
- Divide events by jet/tag multiplicity
- Subdivide by low/high BDT
- Fit MEM discriminant (except 3 jets/3 tags)

Systematics

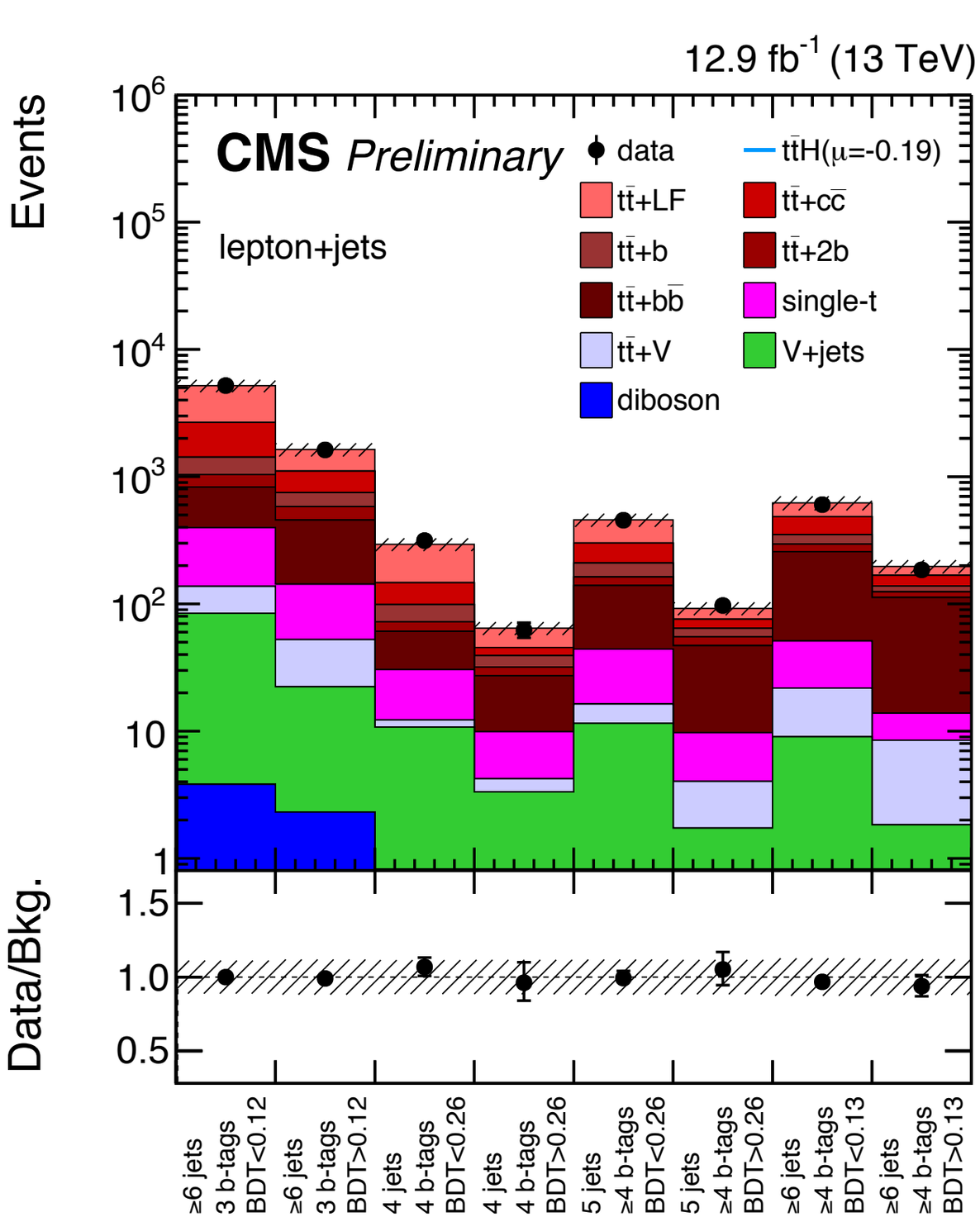
- Shape uncertainties
  - b-tag discriminator
  - JES and JER
  - Lepton ID/Isolation, Trigger efficiency, Pile up, Q^2 scale,
- Rate uncertainties
  - Normalisation of different processes
    Separate for each tt+heavy flavour channel
  - PDF for different production channels (gg, qq, qg, ttH)
  - Parton Shower
- MC bin-by-bin statistics

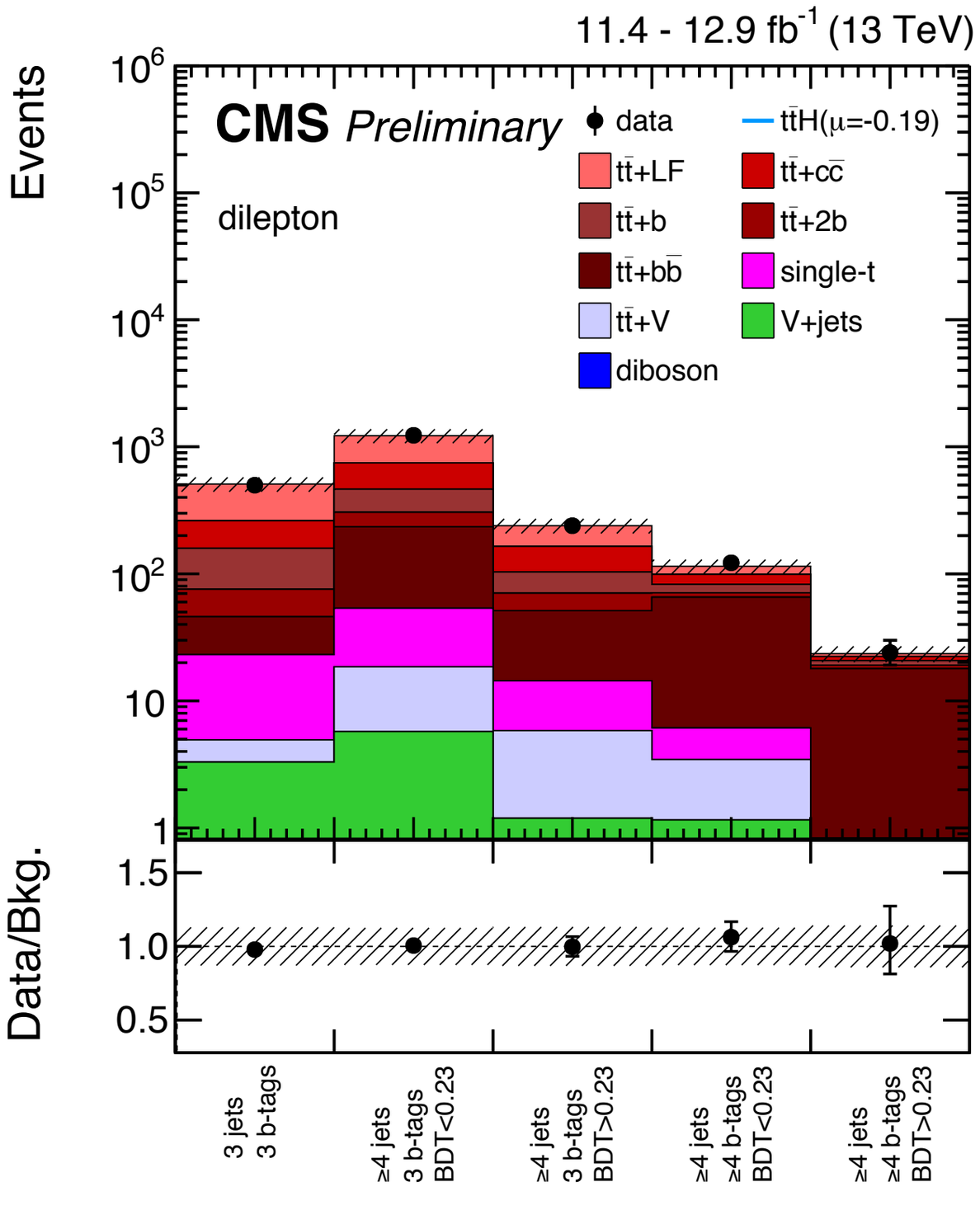
Yields - Pre Fit





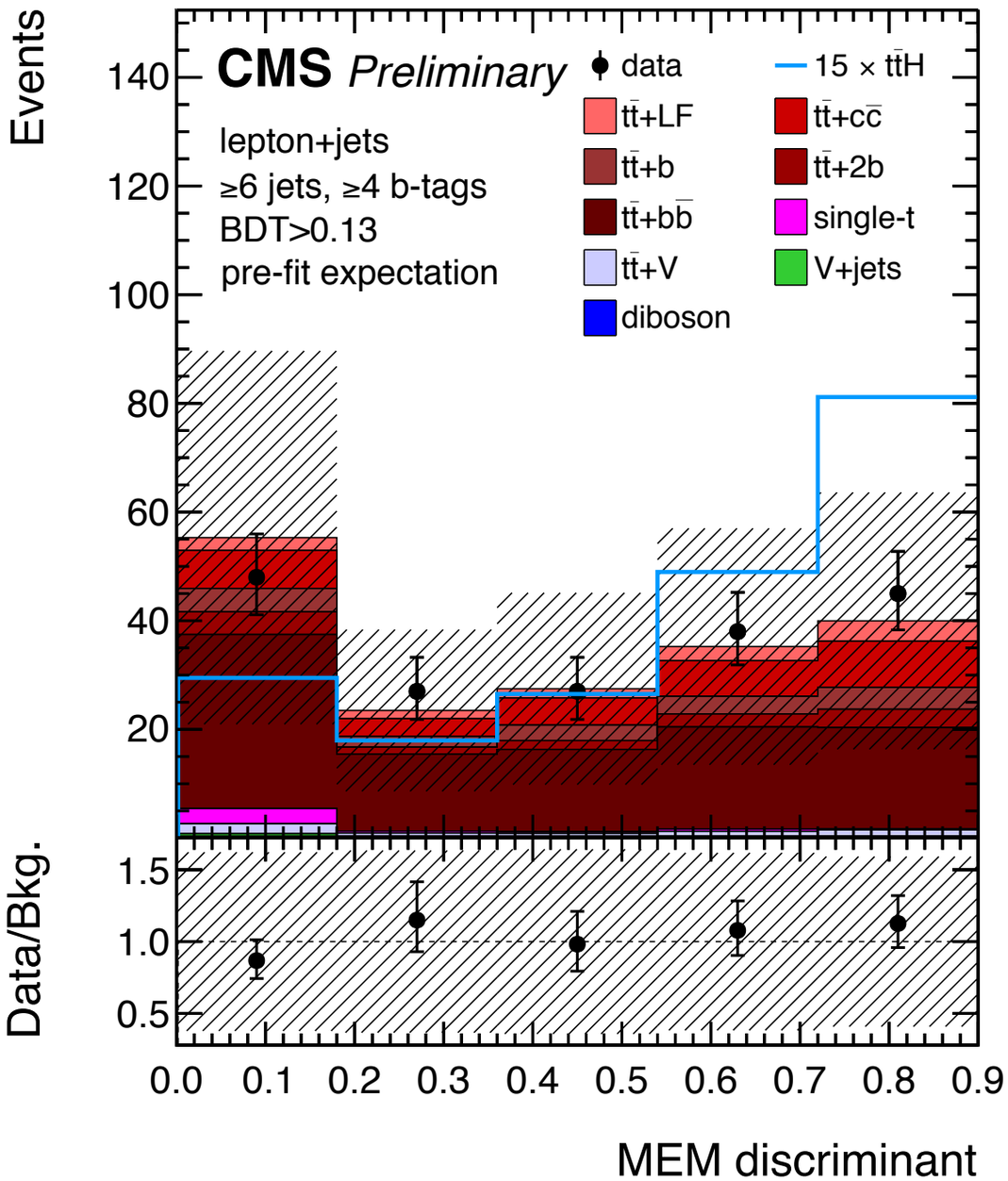
Yields - Post Fit



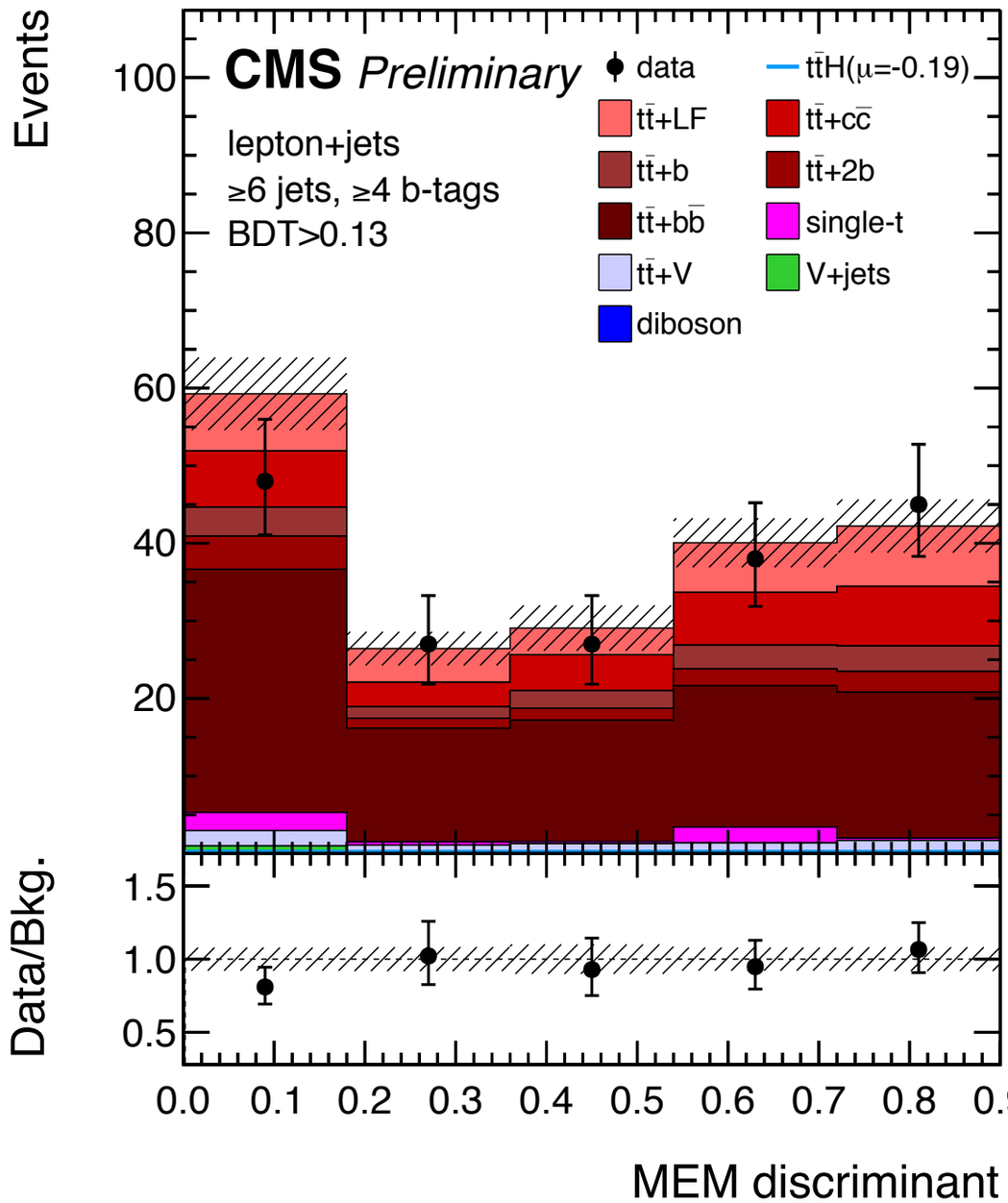


Lepton + Jets j6 t4, high BDT

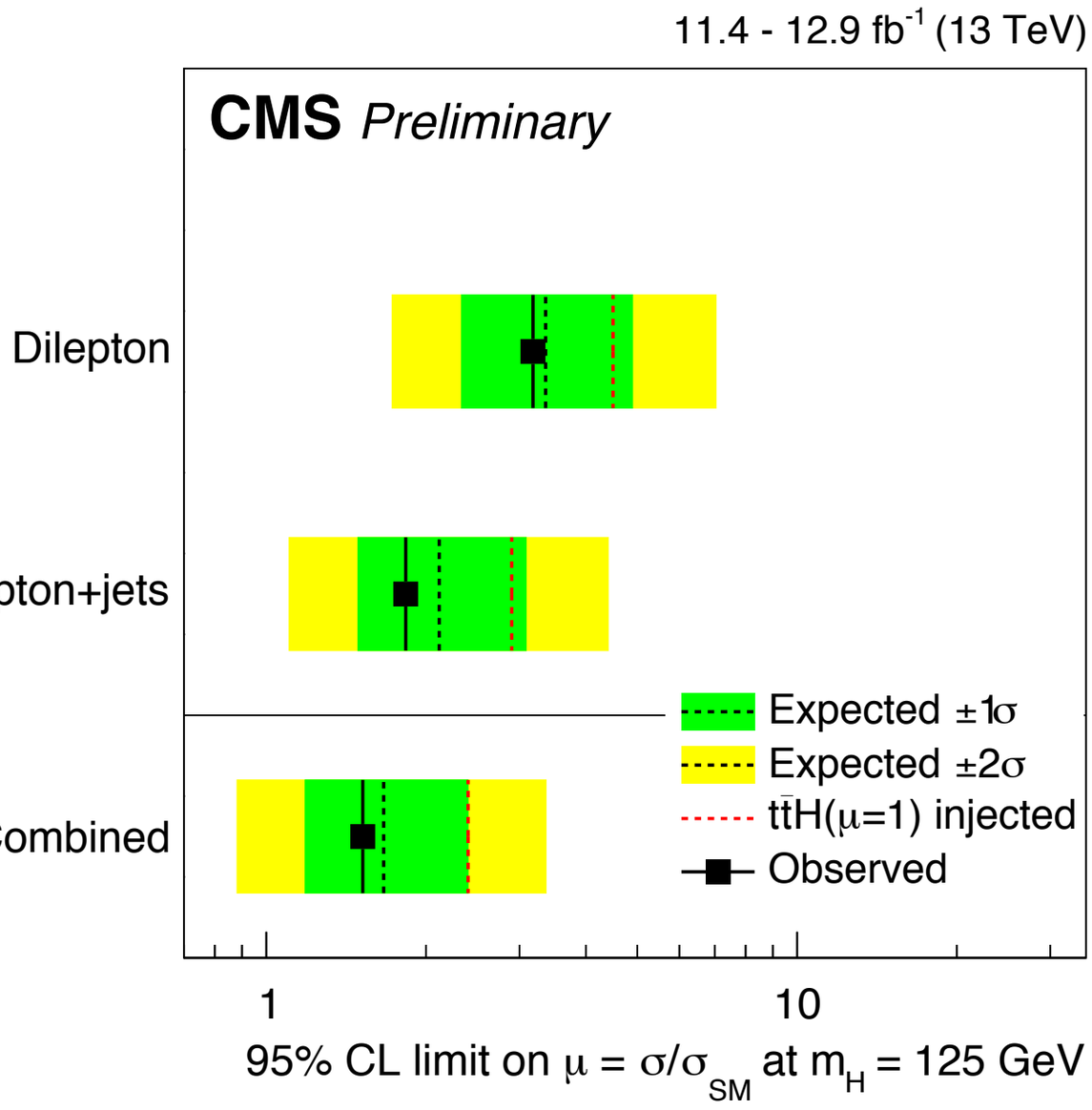
pre-fit

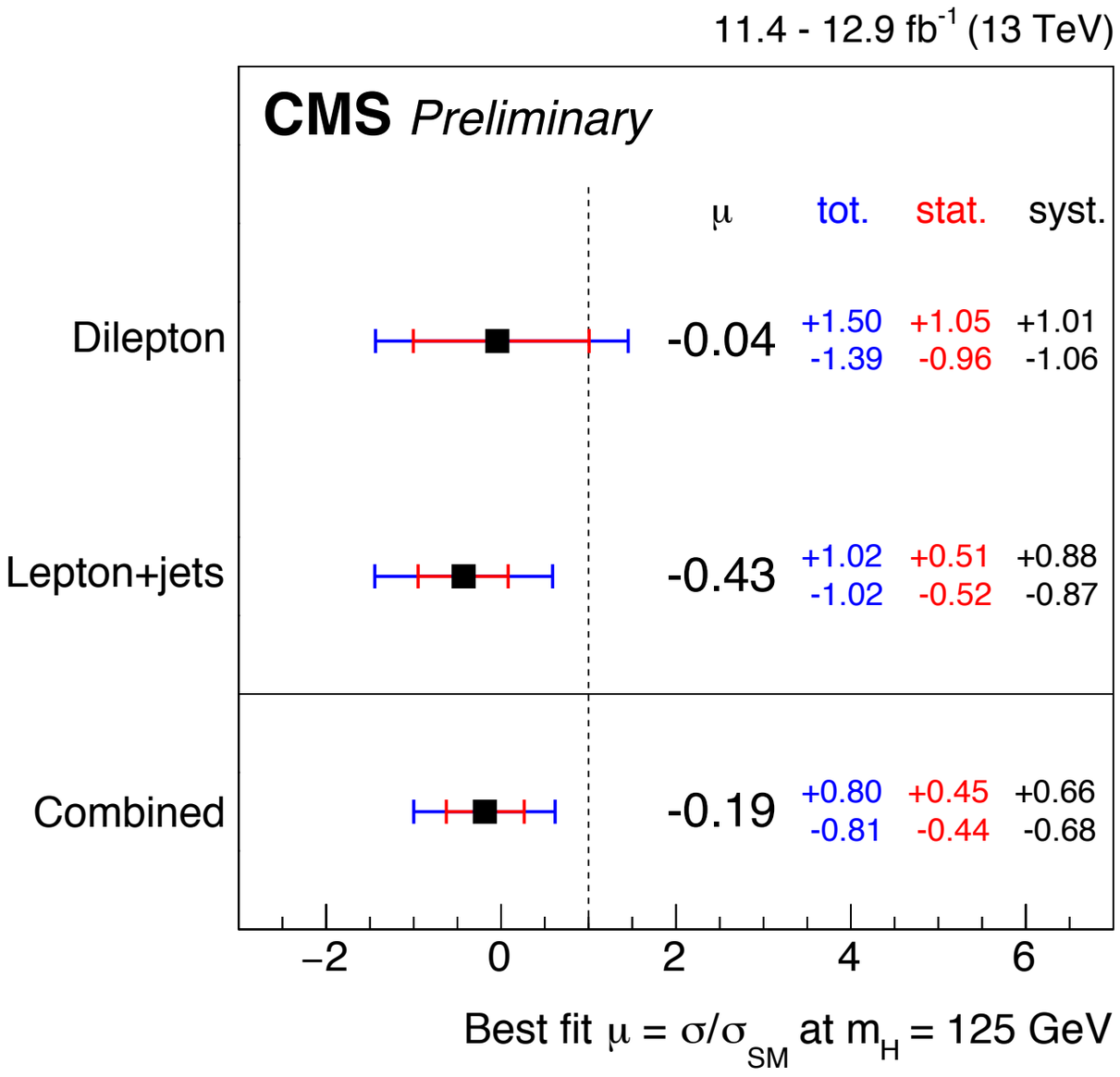


post-fit

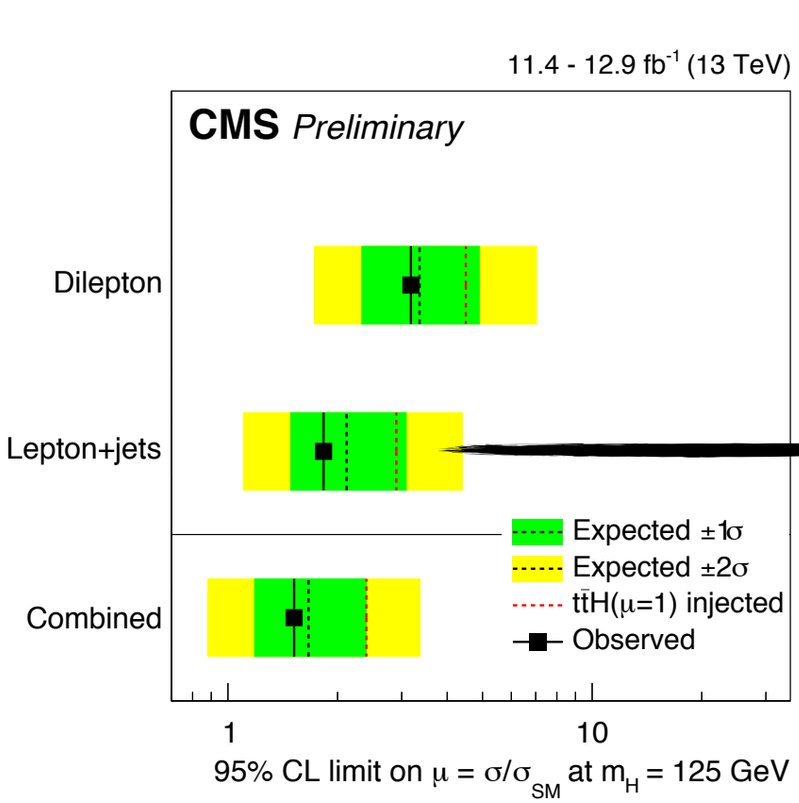


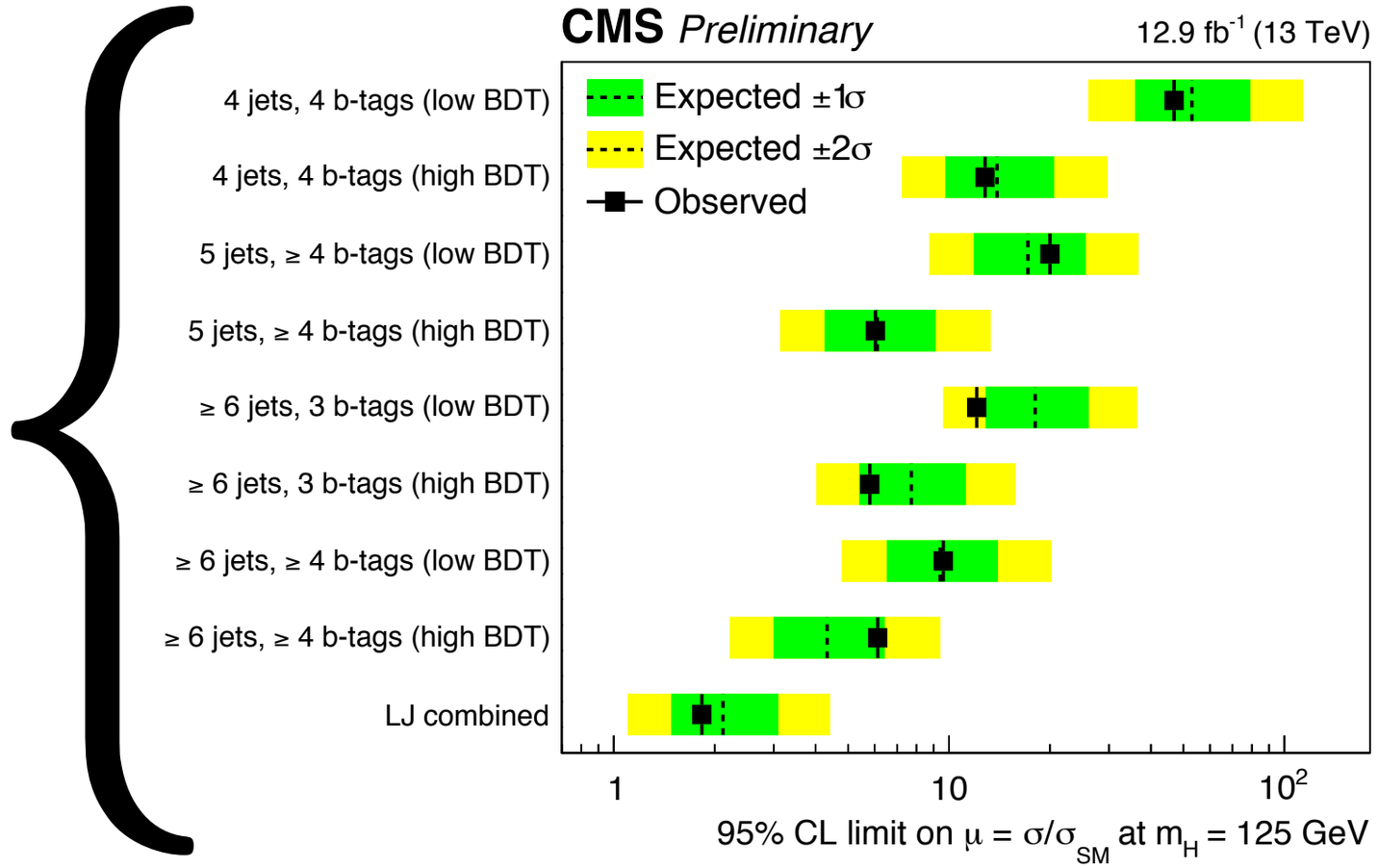
Results


11.4 - 12.9 fb-1 (13 TeV)
CMS Preliminary
Dilepton
Lepton+jets
Combined
Expected ±1σ
Expected ±2σ
tt̄H(μ=1) injected
Observed
1
10
95% CL limit on μ = σ/σSM at mH = 125 GeV


11.4 - 12.9 fb-1 (13 TeV)
CMS Preliminary
μ tot. stat. syst.
Dilepton
-0.04 +1.50 -1.39 +1.05 -0.96 +1.01 -1.06
Lepton+jets
-0.43 +1.02 -1.02 +0.51 -0.52 +0.88 -0.87
Combined
-0.19 +0.80 -0.81 +0.45 -0.44 +0.66 -0.68
−2
0
2
4
6
Best fit μ = σ/σSM at mH = 125 GeV

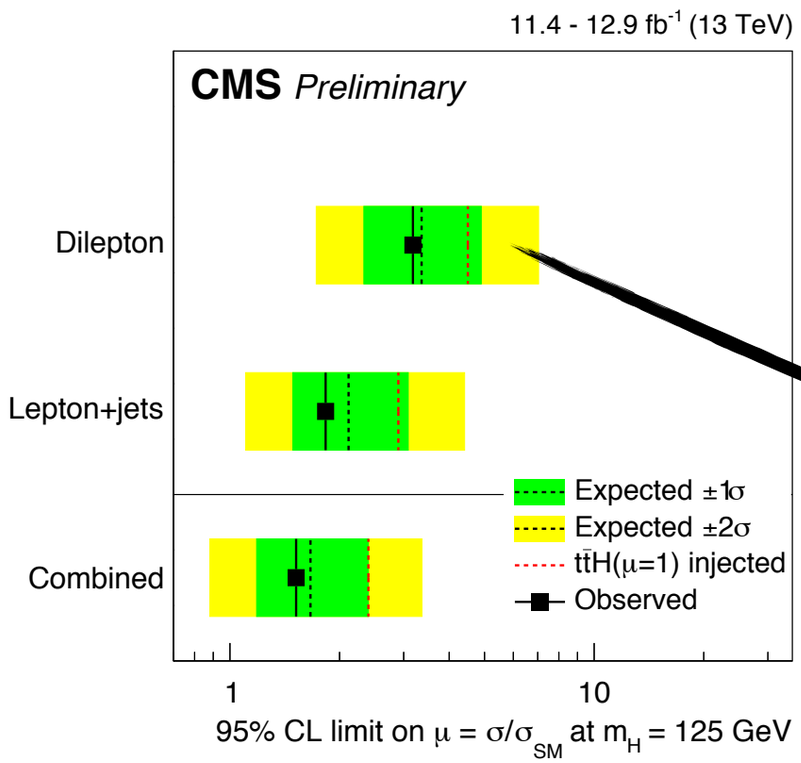
Lepton+Jets per Category

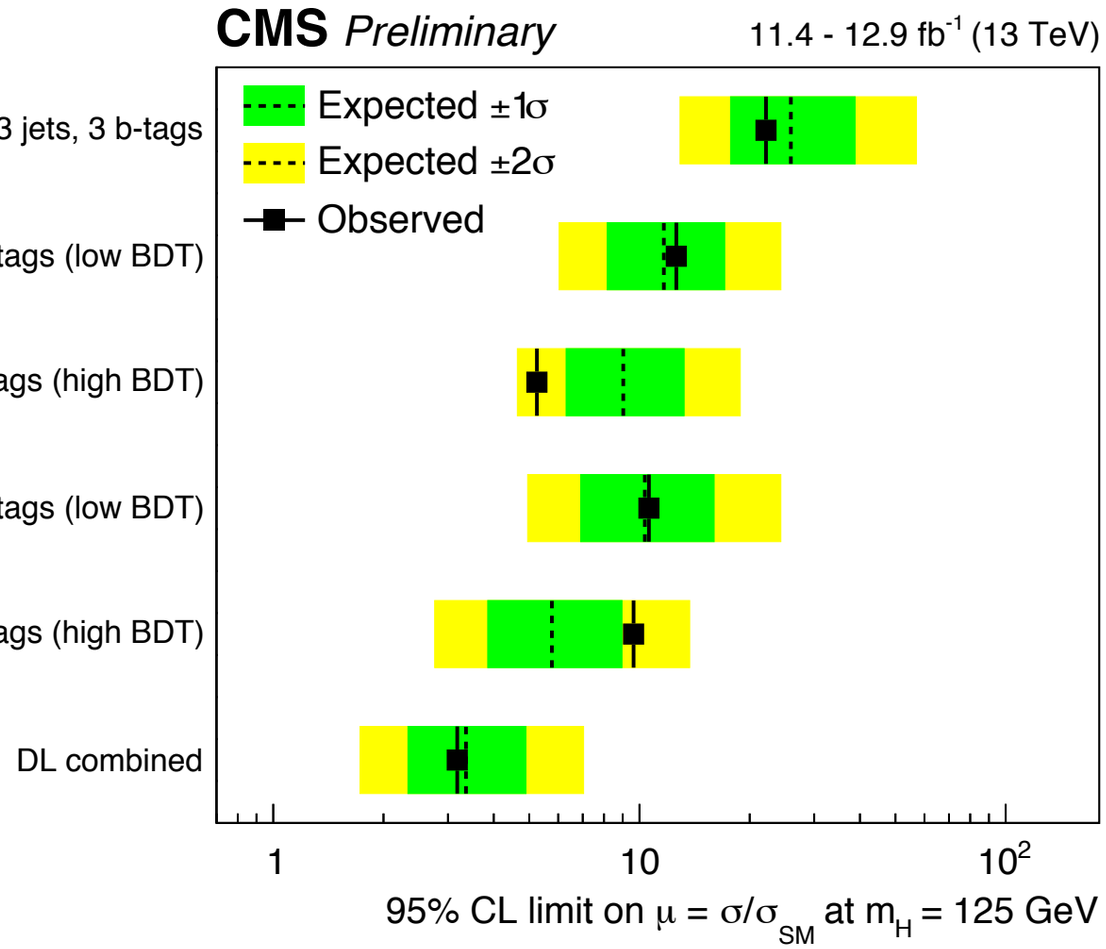




- High-purity bins provide highest sensitivity (as constructed)
- Most sensitive single category shows upward fluctuation in signal region

Dilepton per Category

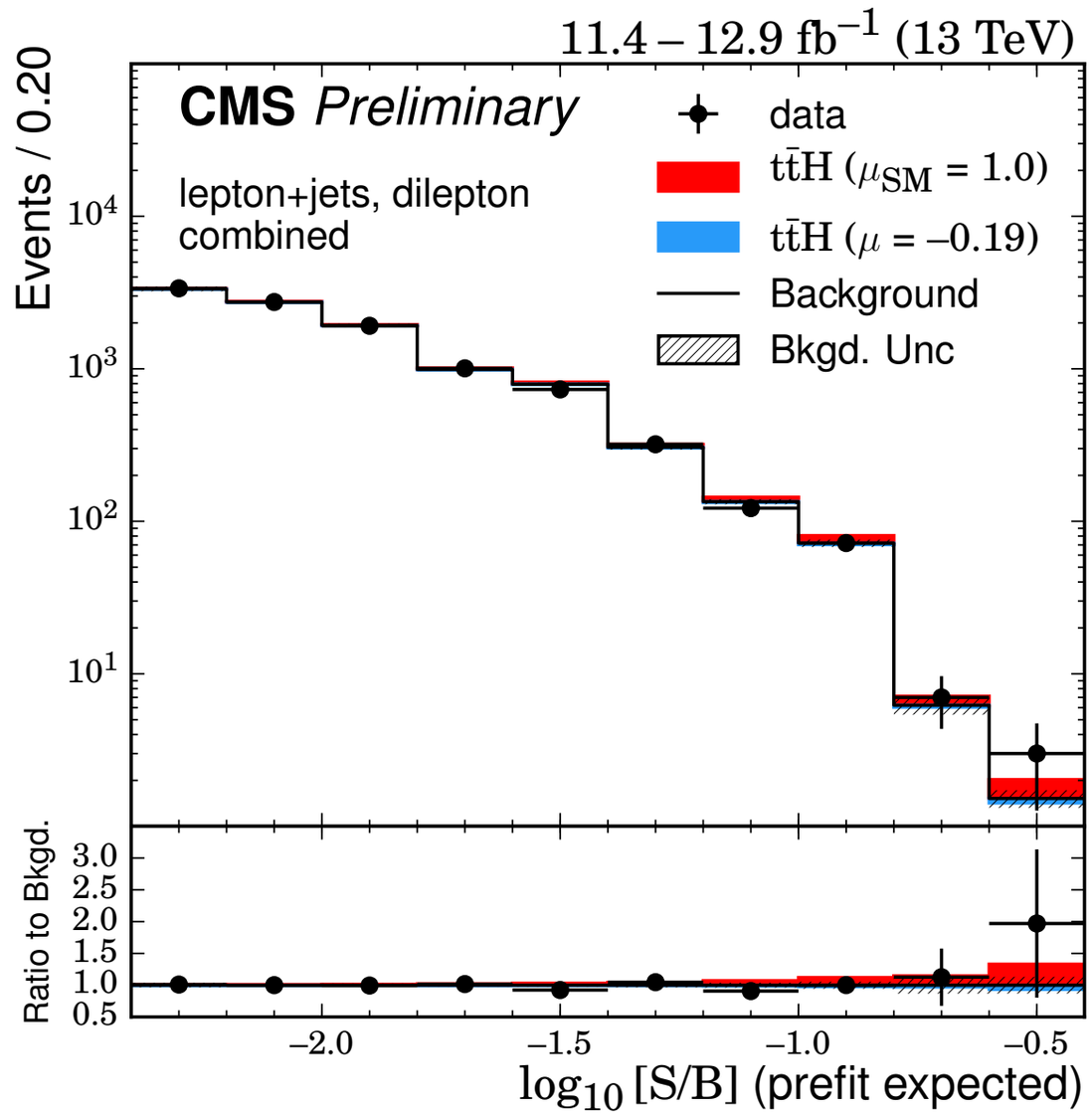




- High-purity bins provide highest sensitivity (as constructed)
- Most sensitive single category shows upward fluctuation in signal region

Conclusions

- Update of HIG-16-004 to up to 12.9 fb^{-1}
- Simple combination of BDT and MEM discrimination techniques
- Observed (expected) upper limit of $\mu < 1.5$ (1.7) at the 95% confidence level, and a best fit value of $\mu = -0.19^{+0.80}_{-0.81}$
- Now we need to get back and analyze the 40 fb^{-1} on tape
  - Revisit boosted events
  - More complex classifier combination
  - Higher statistics ttbb sample
  - Stay tuned..



Thank you!

Bonus Slides

Particle Swarm

- See: *Particle swarm optimization*,
  J. Kennedy, R. Eberhart
  Proceedings of the IEEE International Conference on Neural Networks, 1995.

- Optimization algorithm

- Different BDT setting (i.e. tree structure and variables) form the search-space

- A specific setting corresponds to one point in this search space

- **Algorithm**:

  - Create swarm of candidate BDTs

  - Each BDT is initialized with a random set of input variables and position in parameter-space

  - Do N iterations

    - Repeatedly train/test at current position.

    - Vary input variables to maximize ROC while KS>threshold

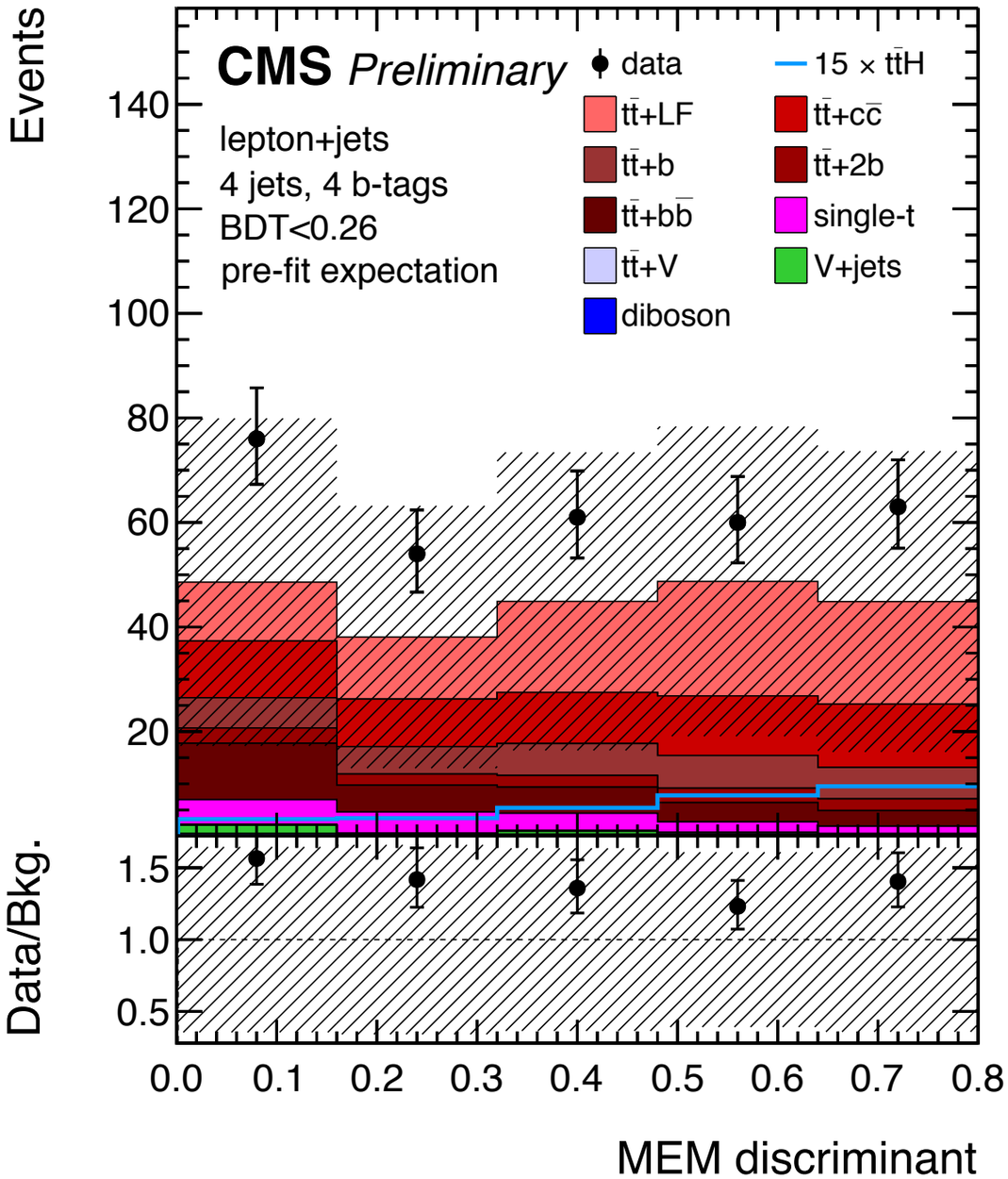
    - Then the BDTs move to new positions, based on their own and swarms best previous positions

Final Discriminators

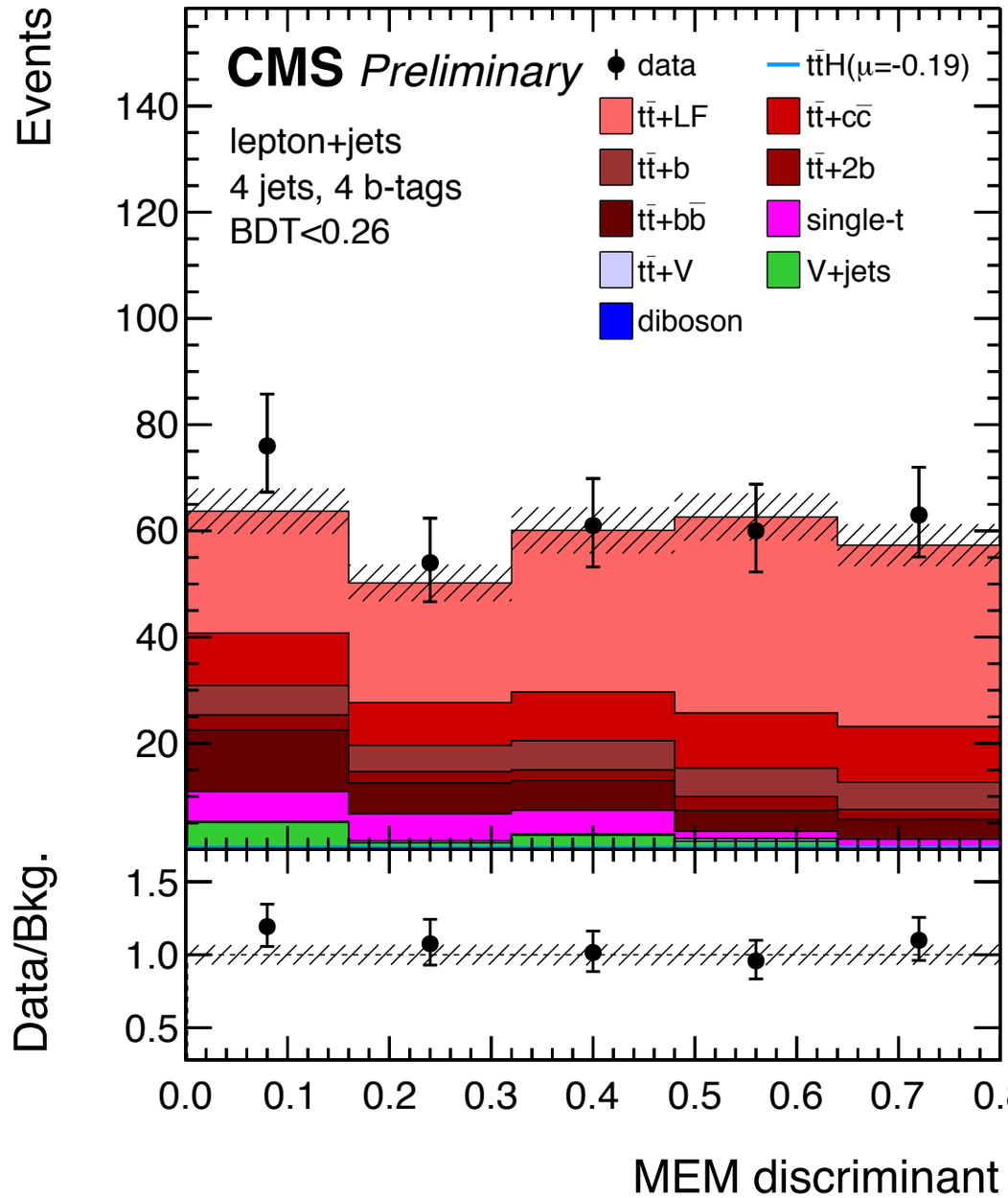
Lepton+Jets

Lepton + Jets j4 t4, low BDT

pre-fit

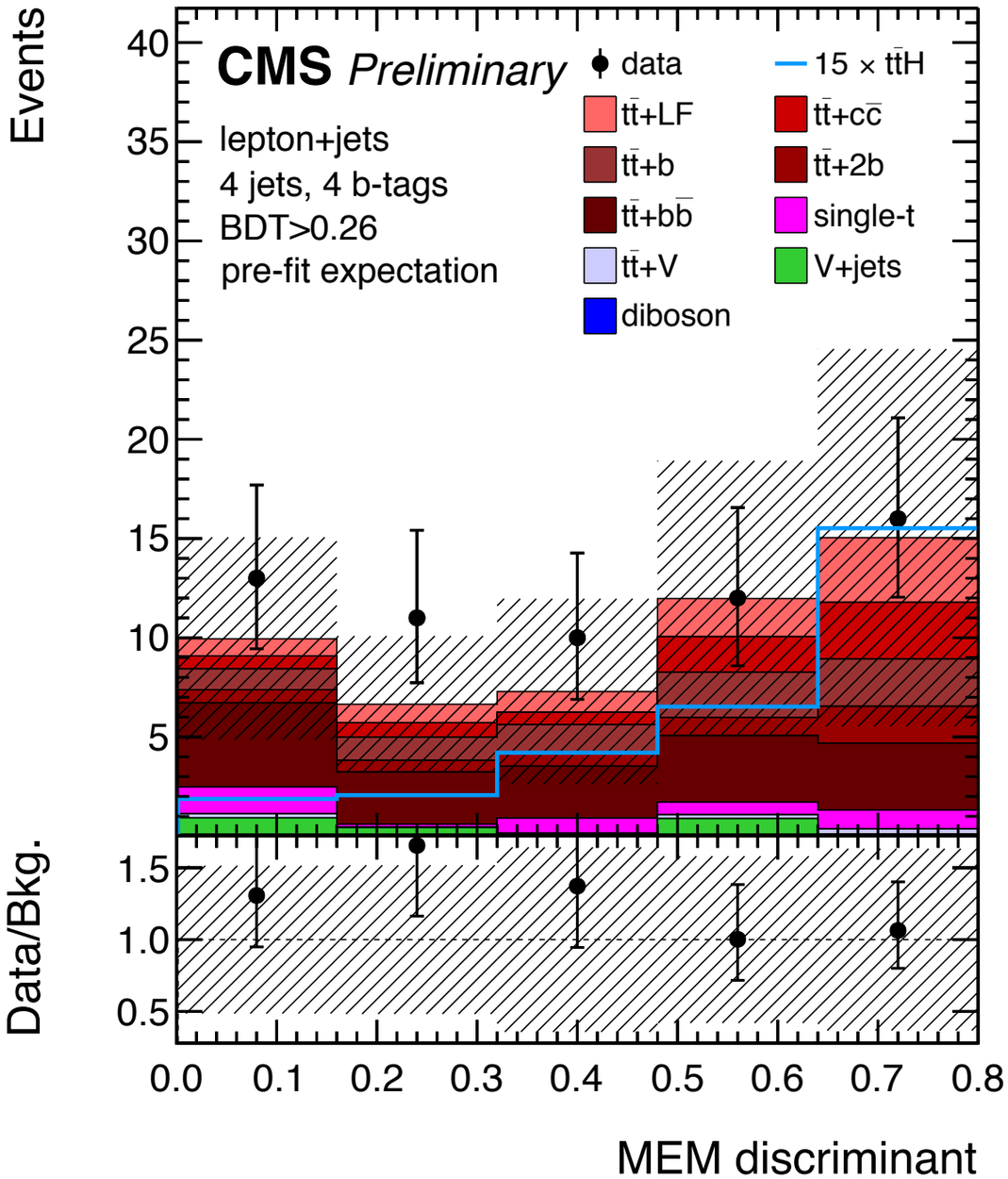


post-fit

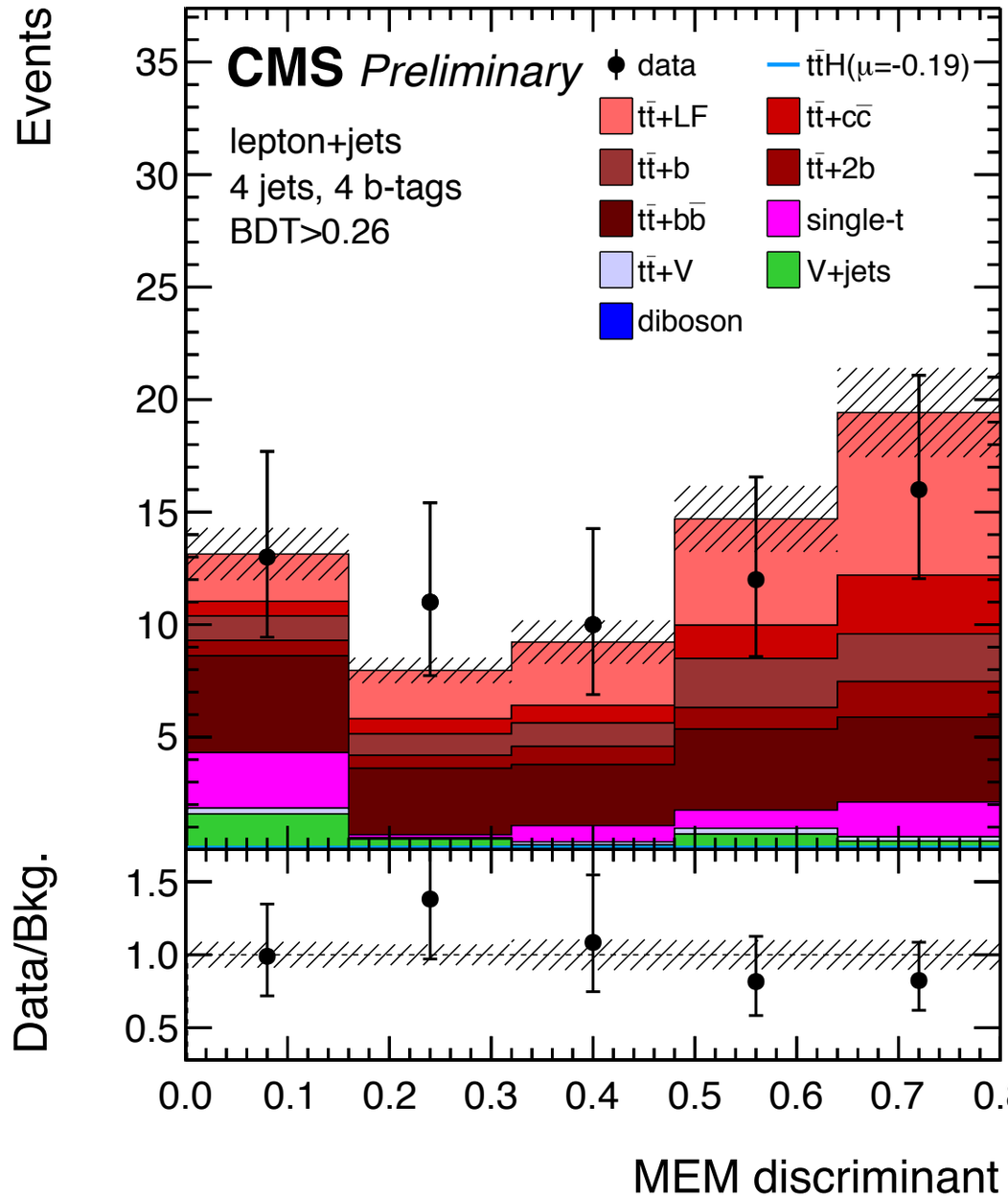


Lepton + Jets j4 t4, high BDT

pre-fit

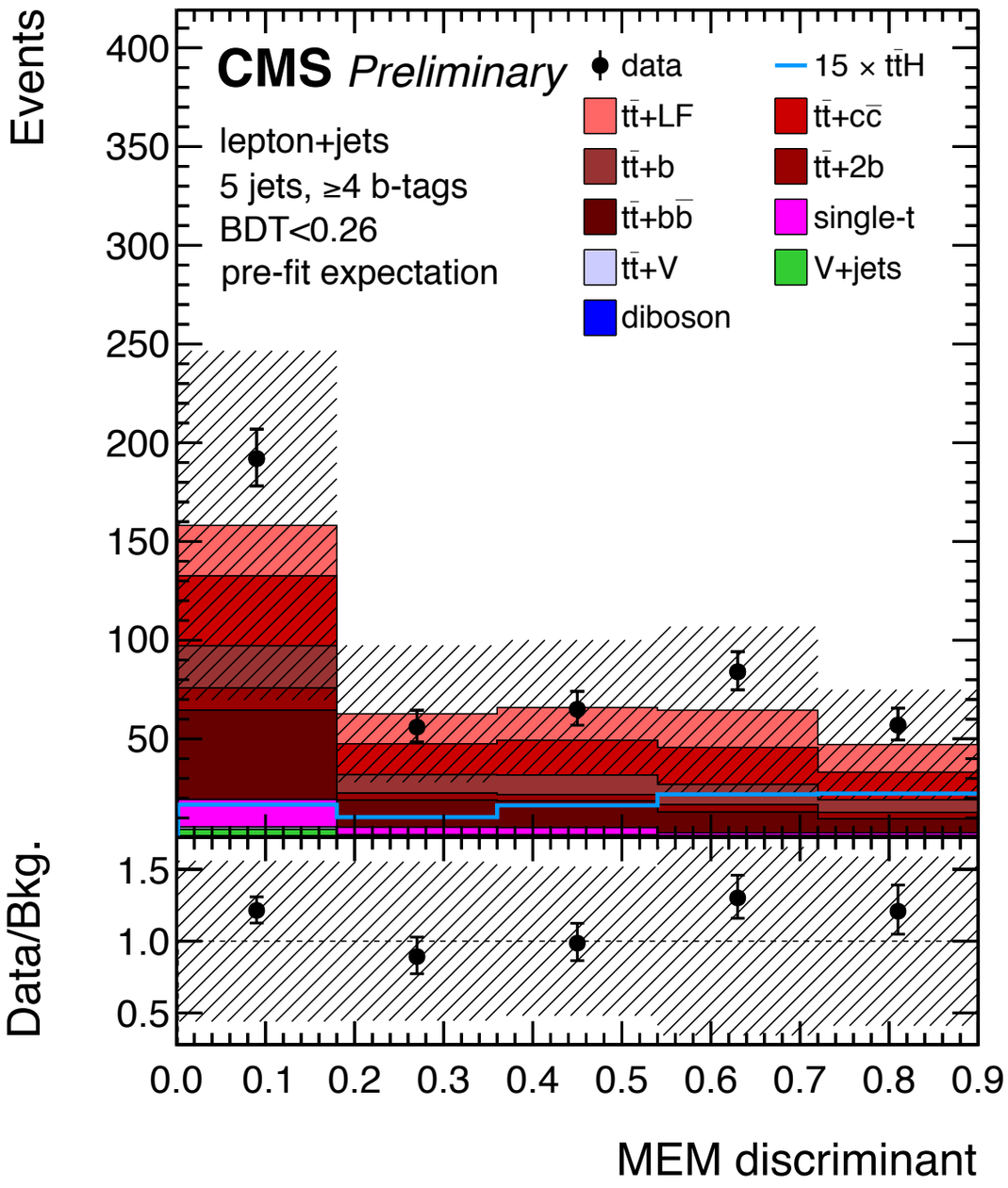


post-fit

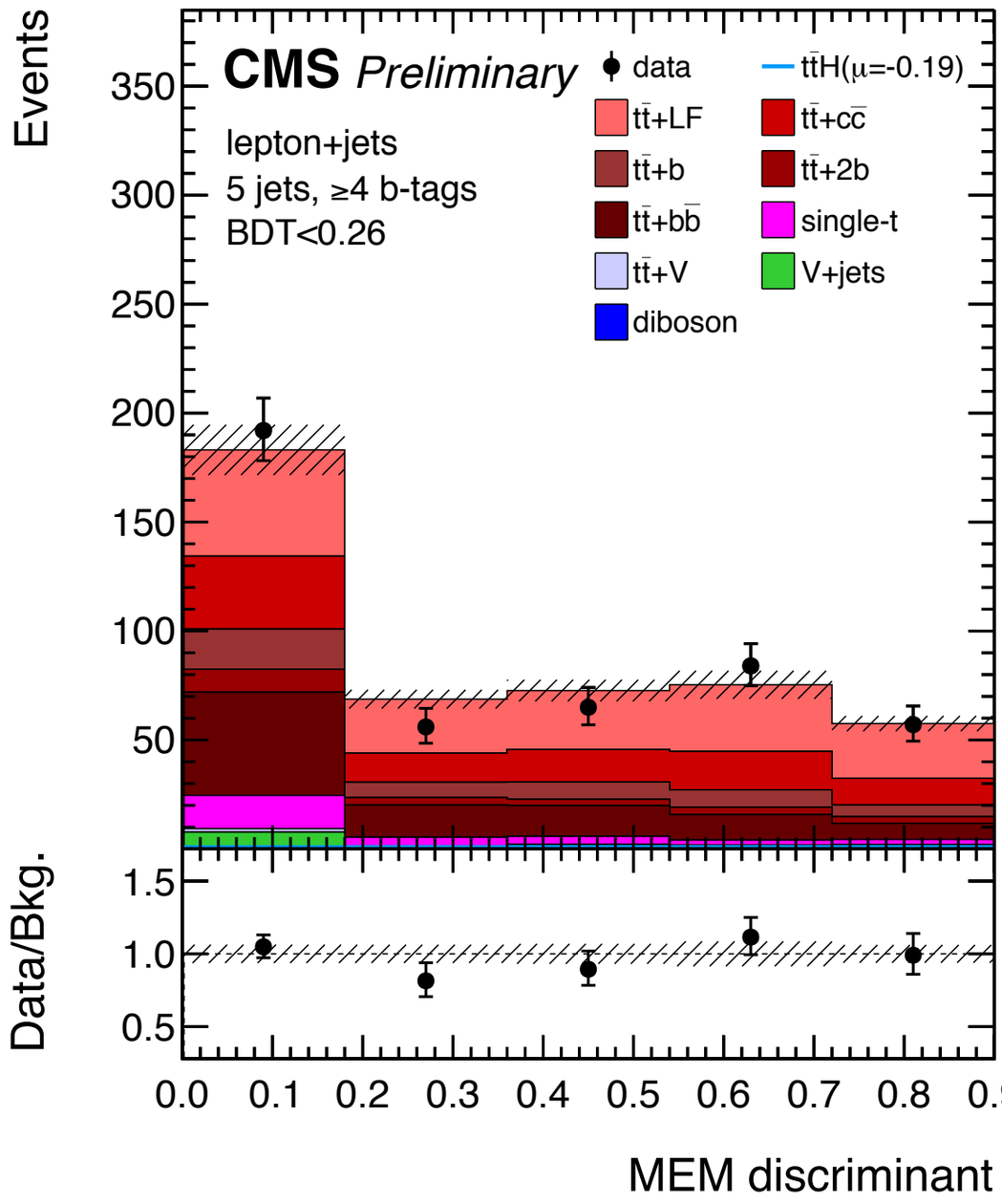


Lepton + Jets j5 t4, low BDT

pre-fit

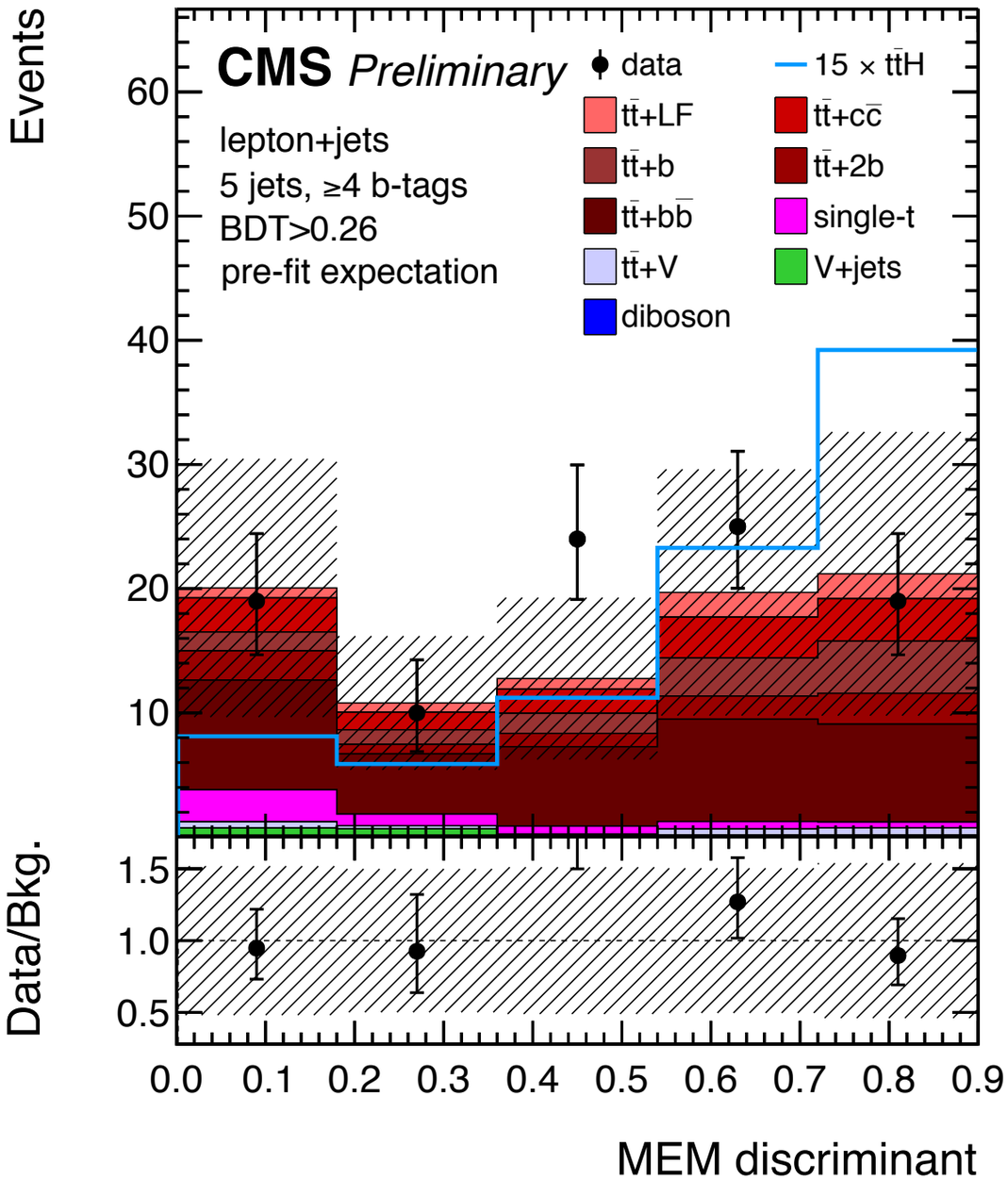


post-fit

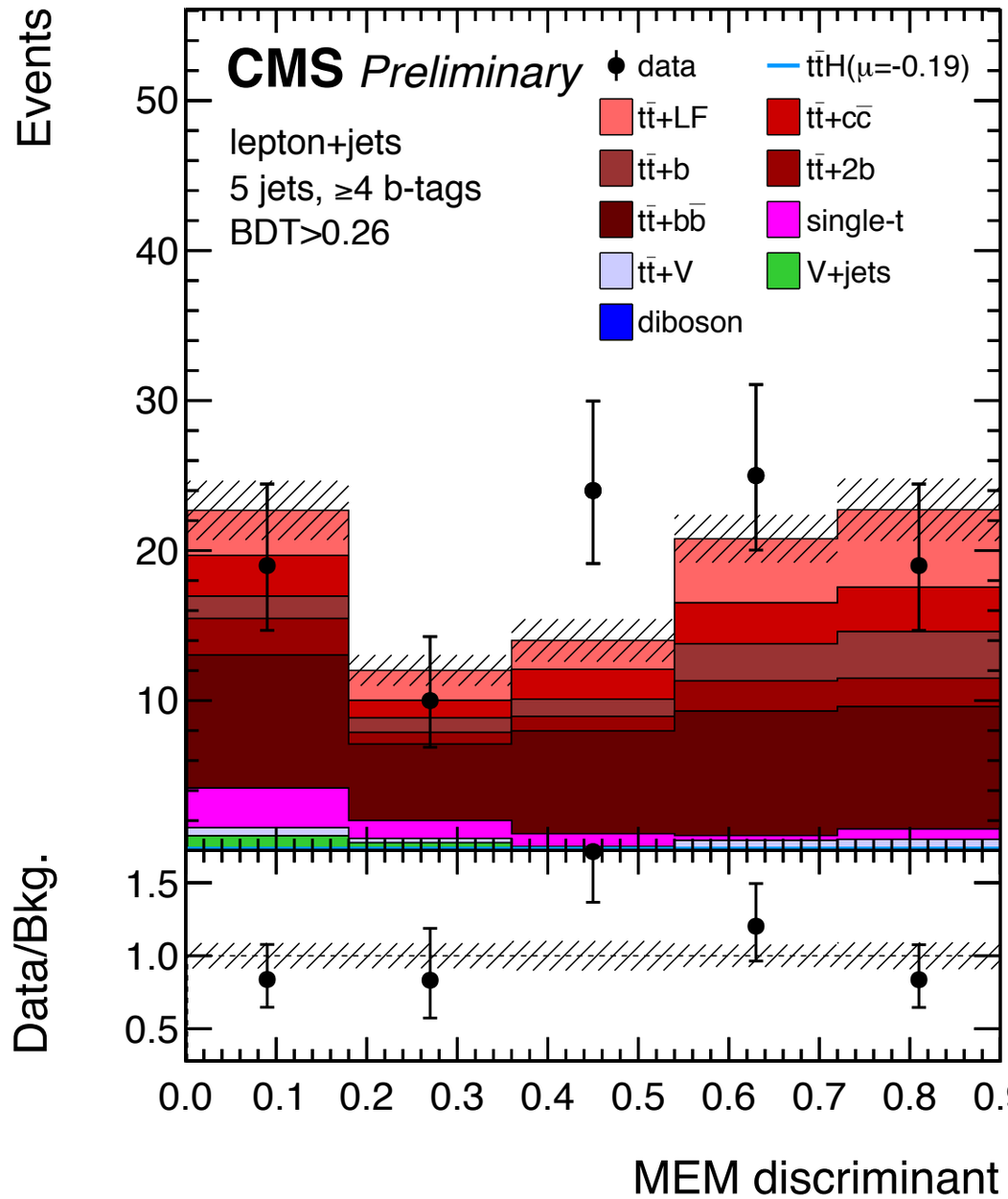


Lepton + Jets j5 t4, high BDT

pre-fit

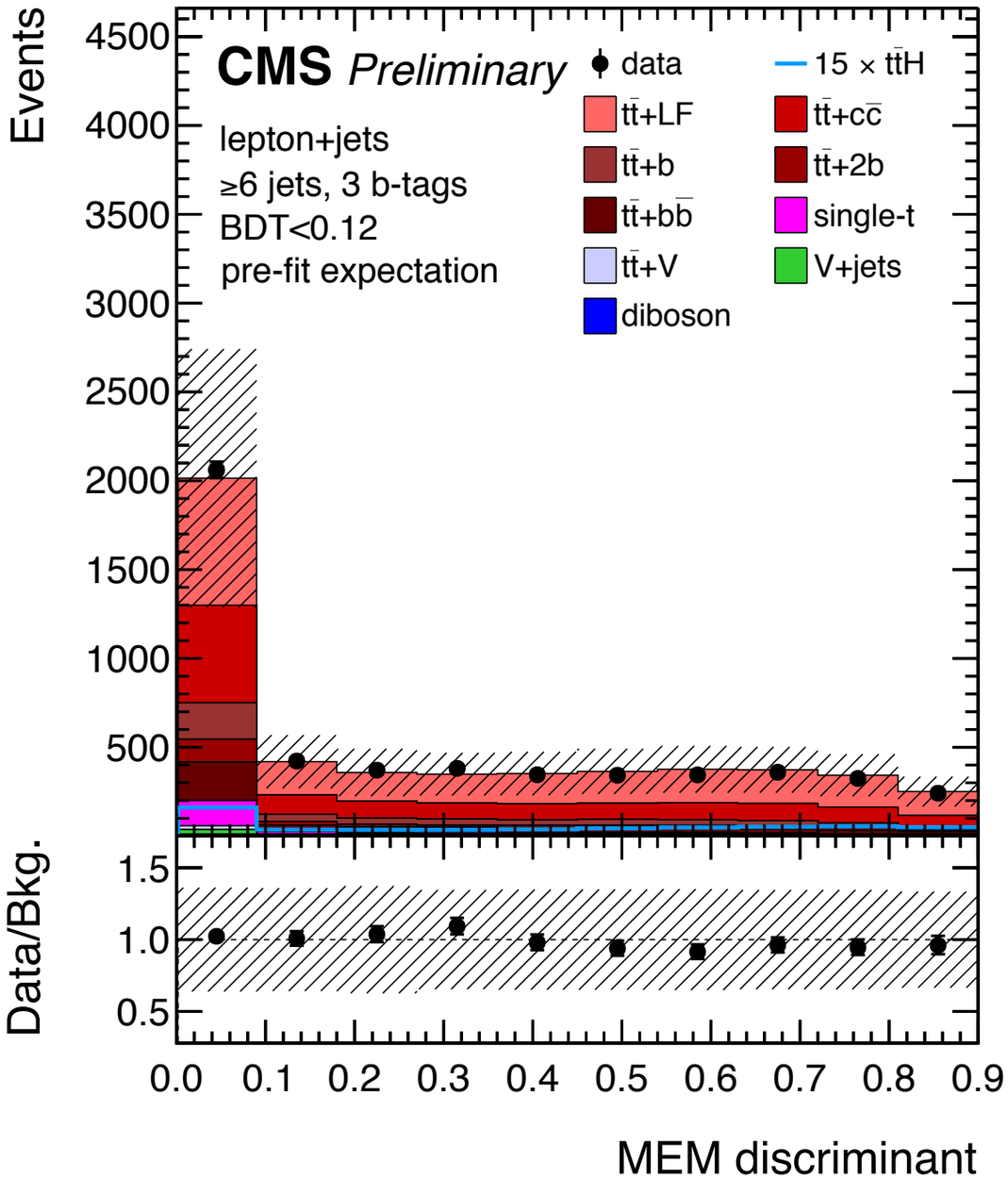


post-fit

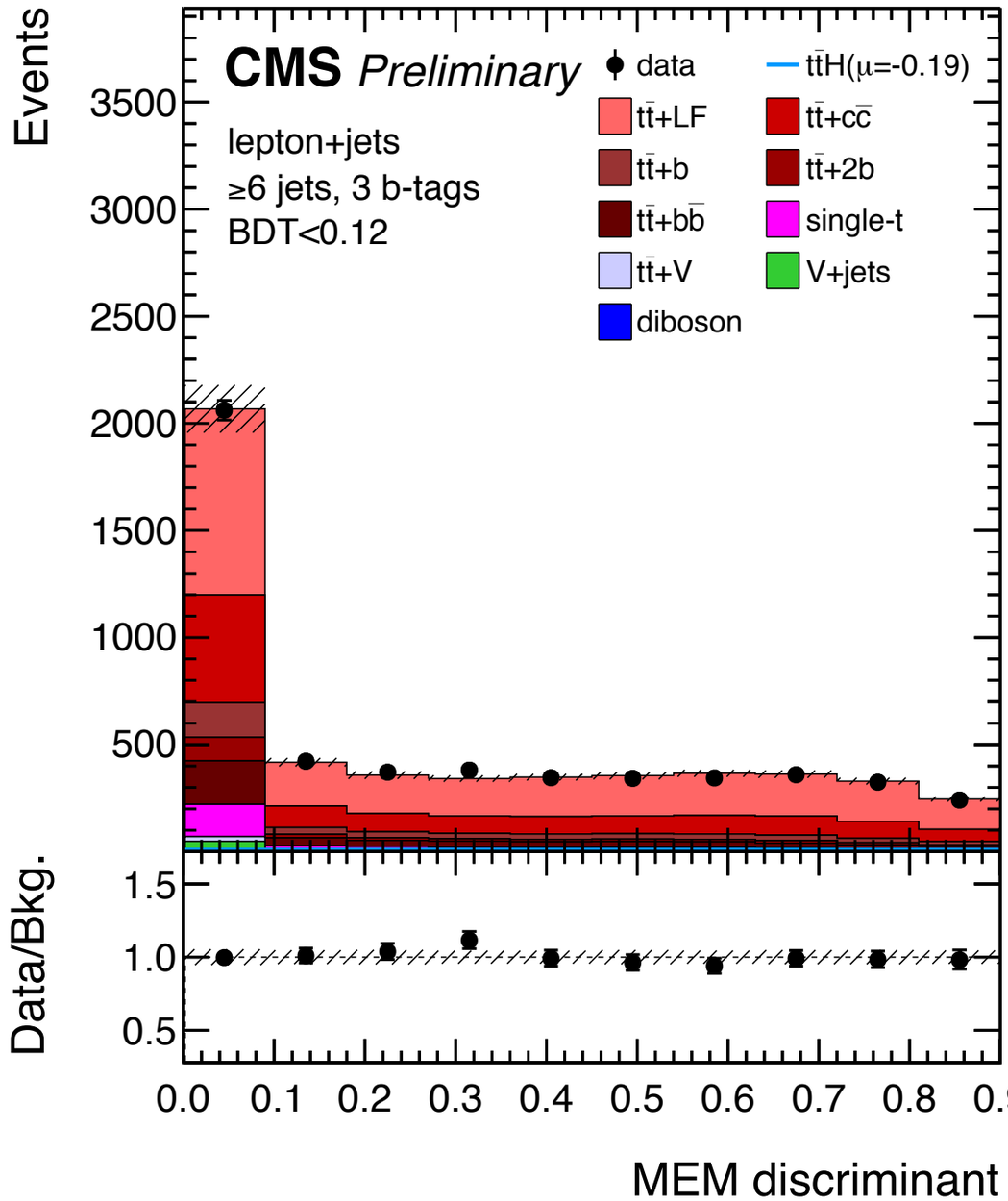


Lepton + Jets j6 t3, low BDT

pre-fit

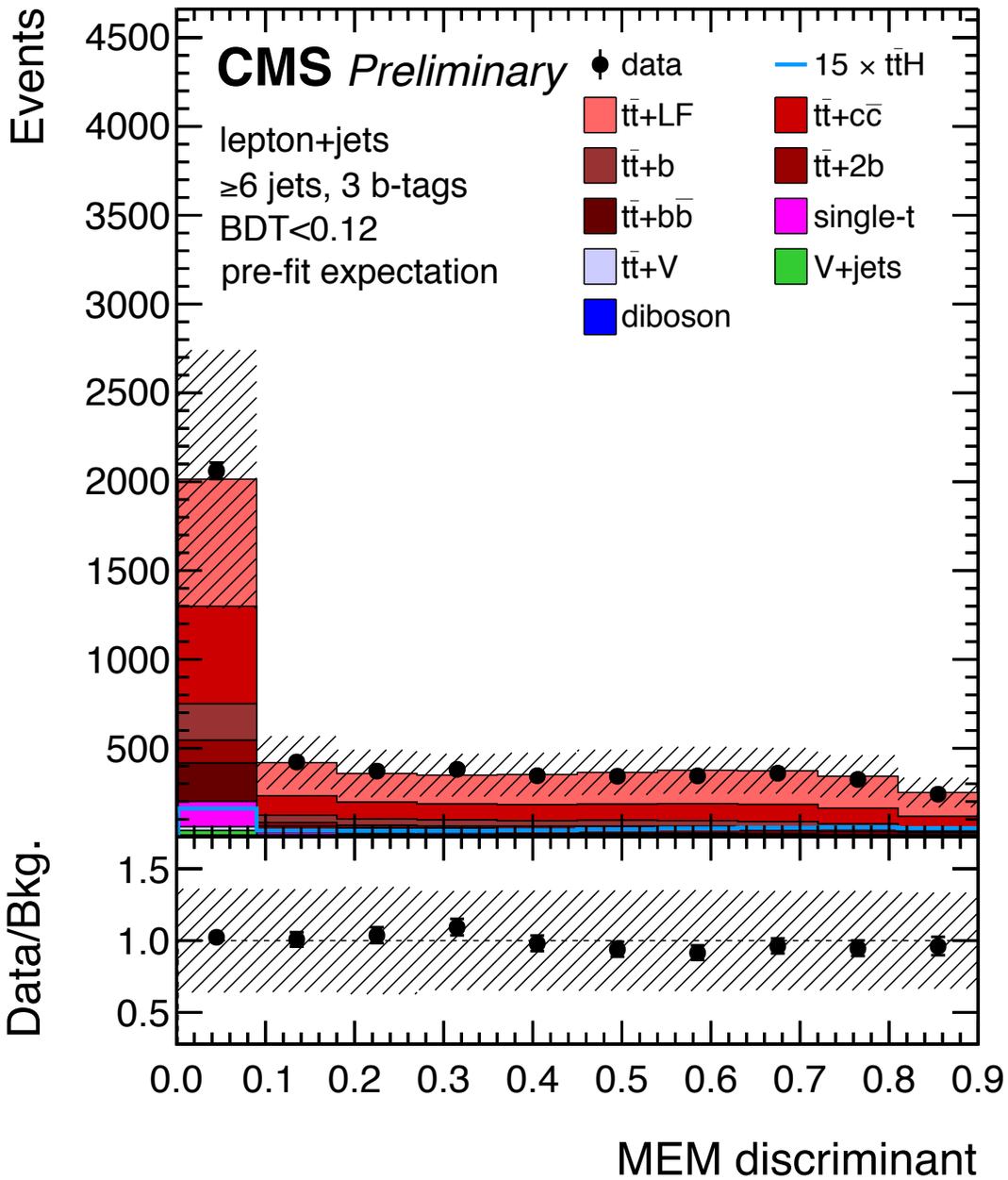


post-fit

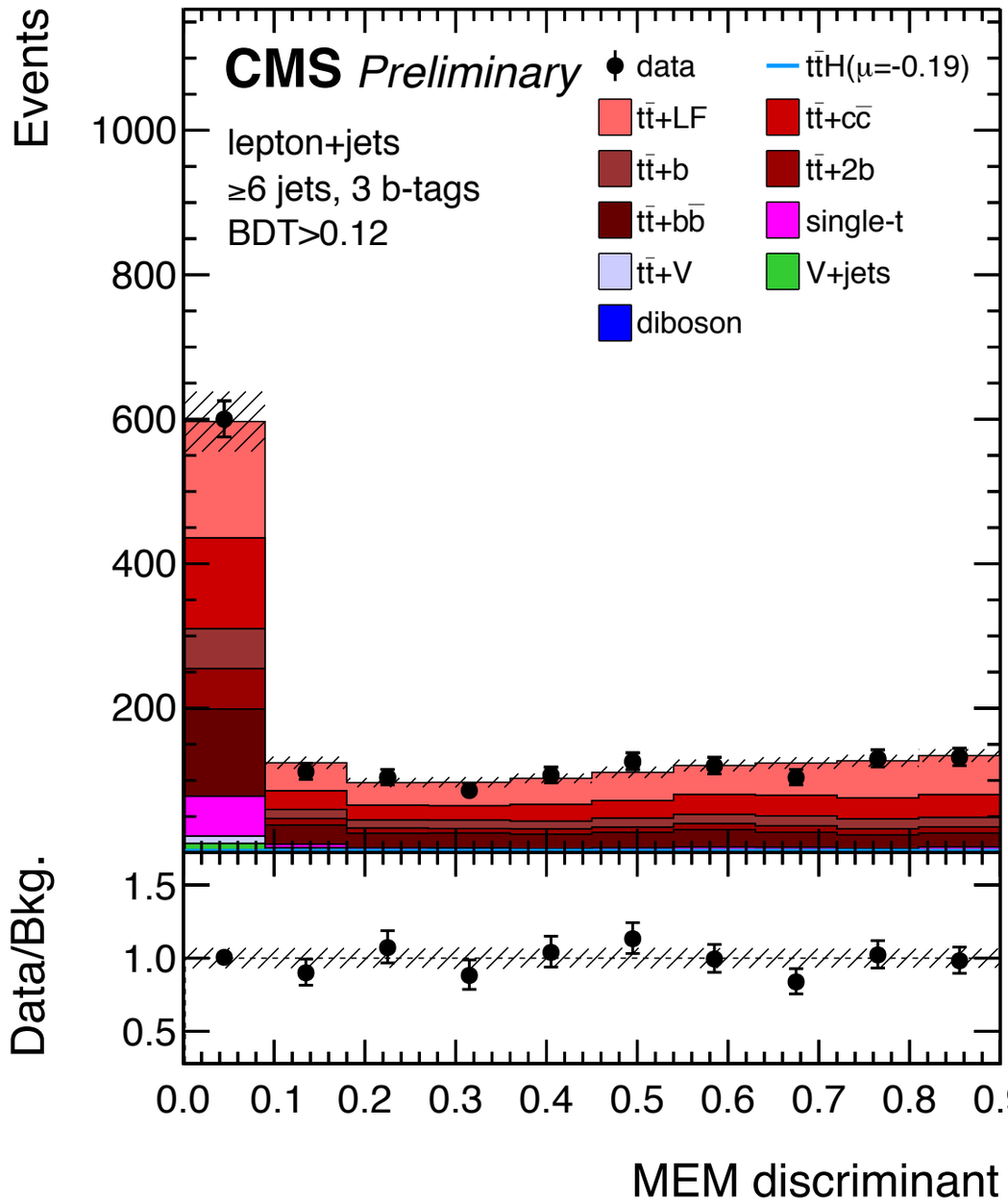


Lepton + Jets j6 t3, high BDT

pre-fit

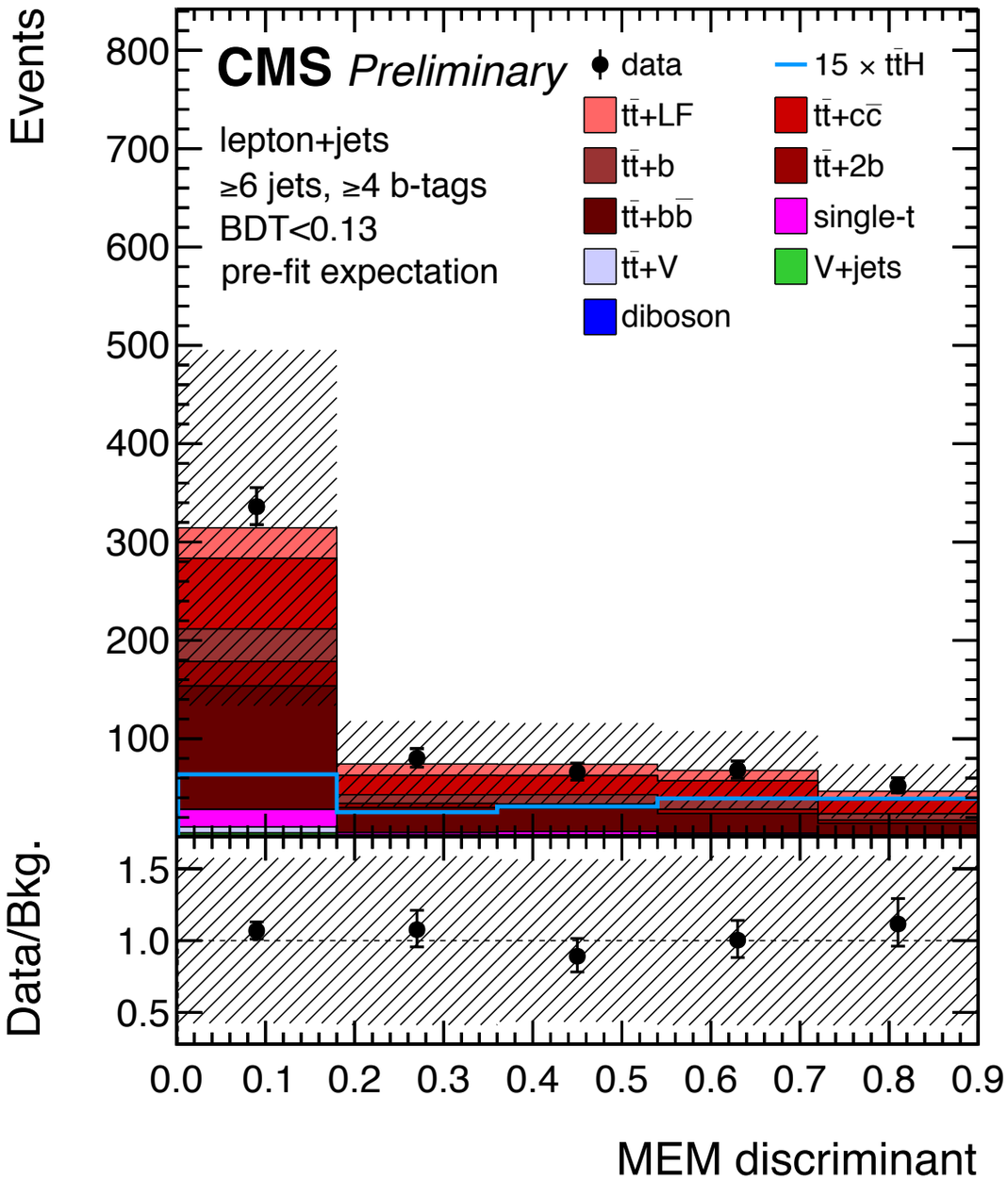


post-fit

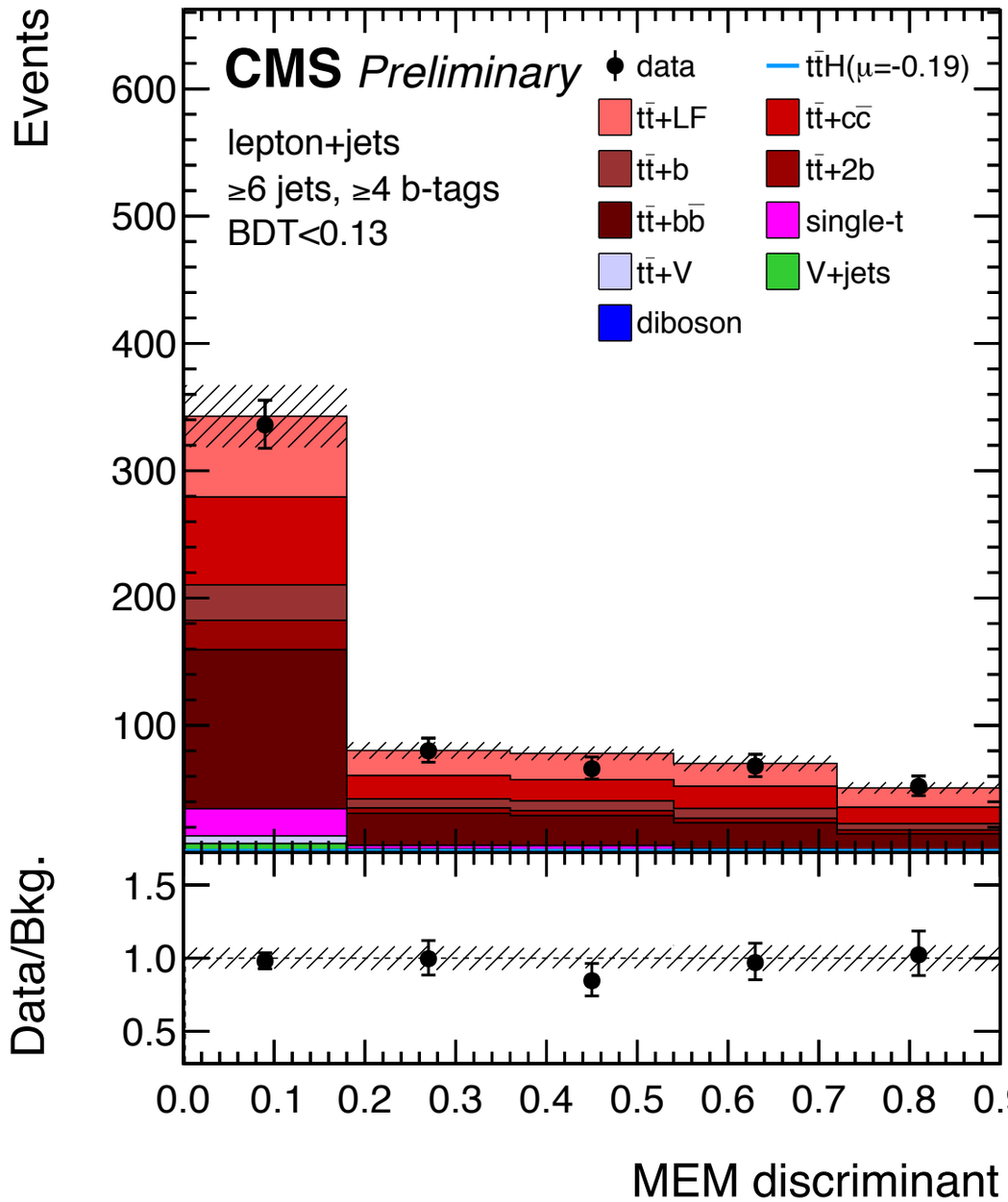


Lepton + Jets j6 t4, low BDT

pre-fit

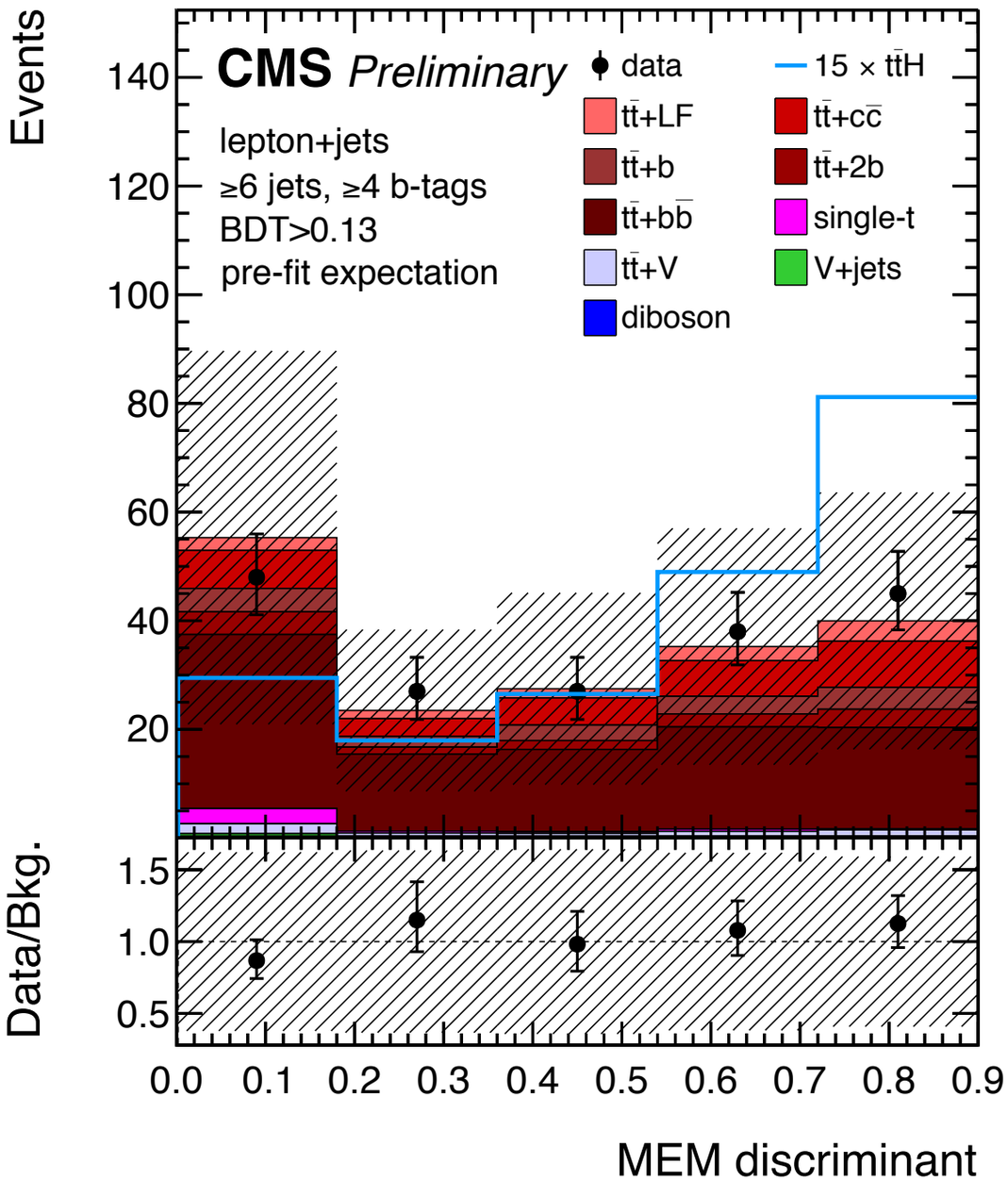


post-fit

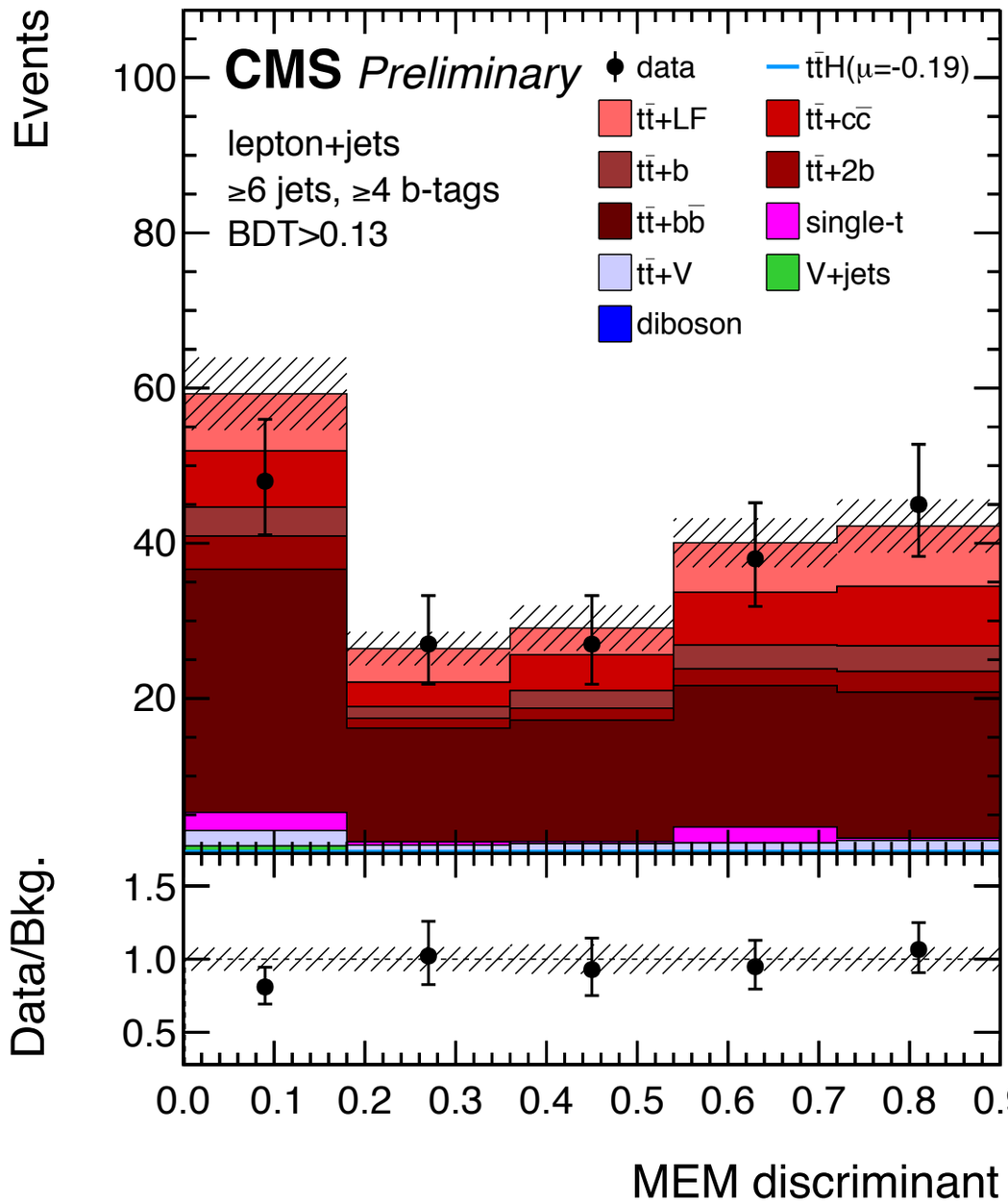


Lepton + Jets j6 t4, high BDT

pre-fit



post-fit

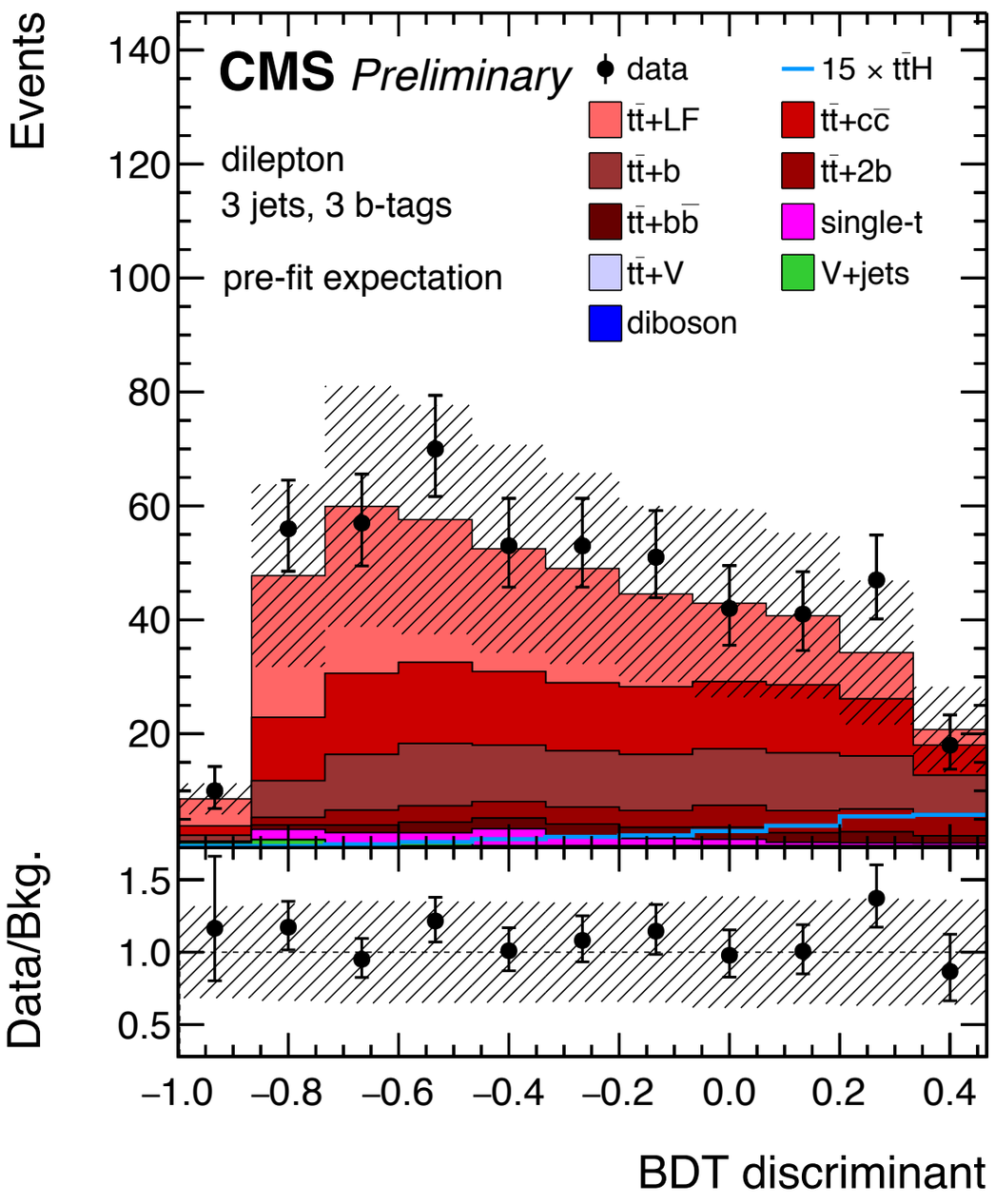


Final Discriminators

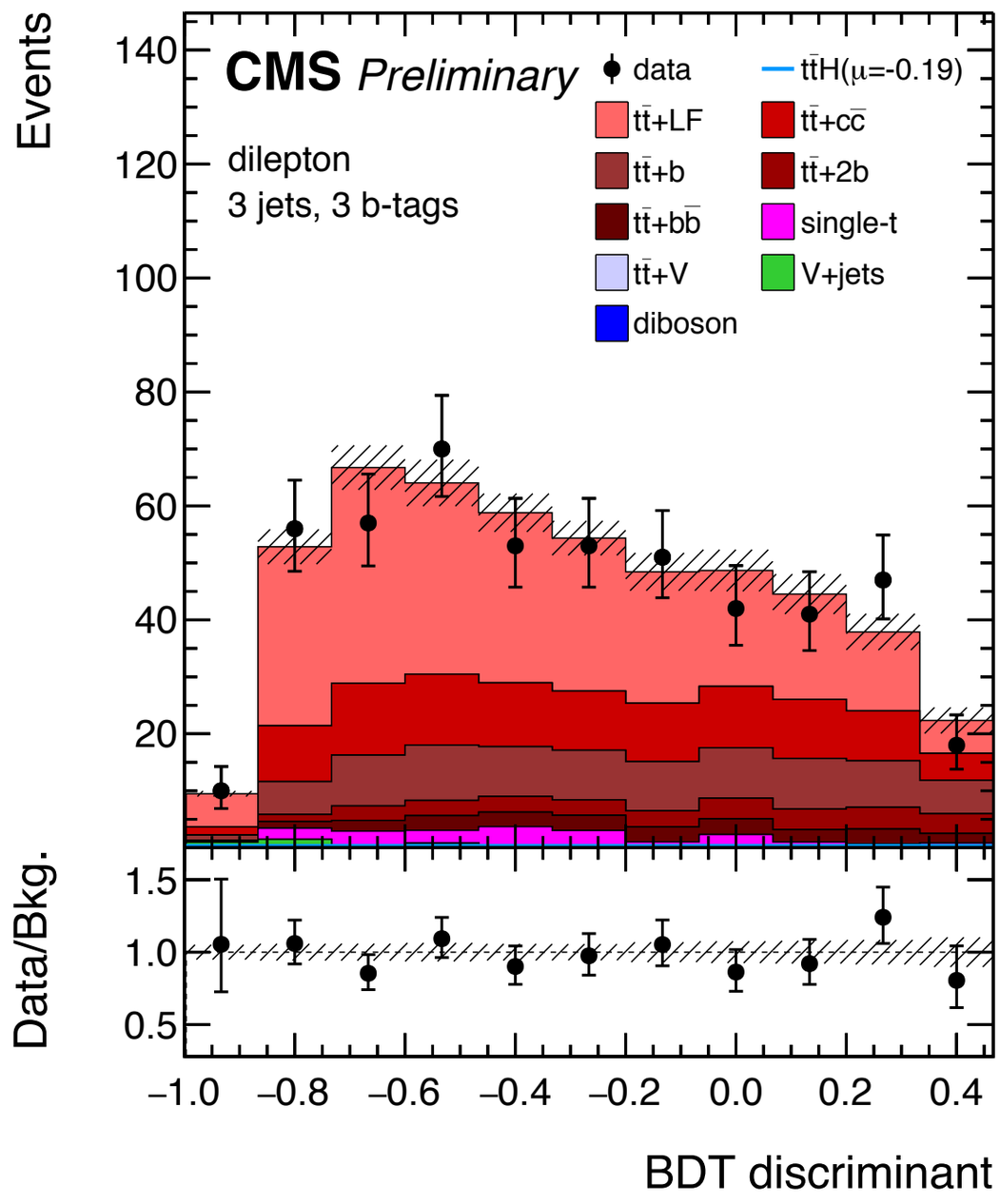
Dilepton

Dilepton j3 t3

pre-fit

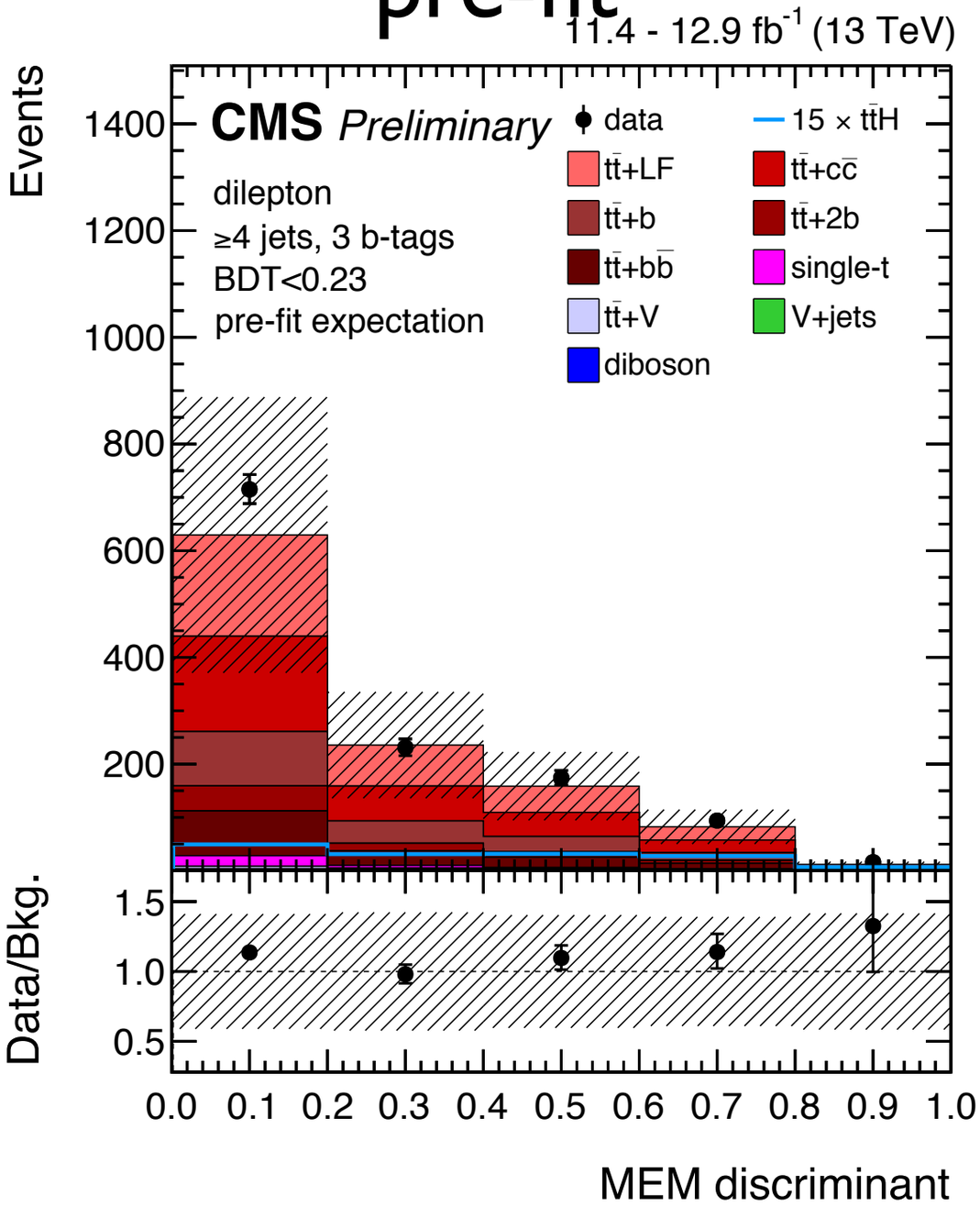


post-fit

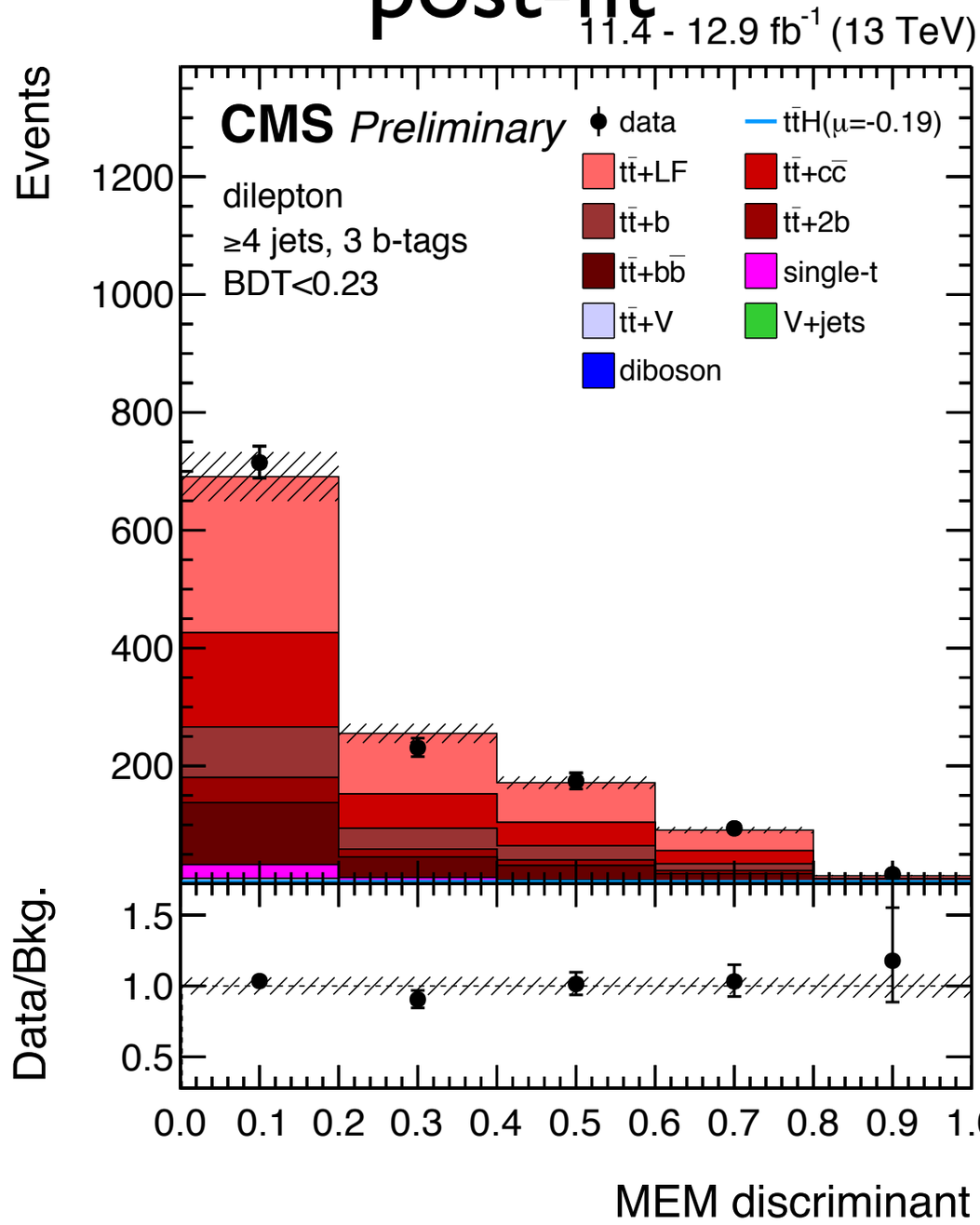


Dilepton j4 t3, low BDT

pre-fit

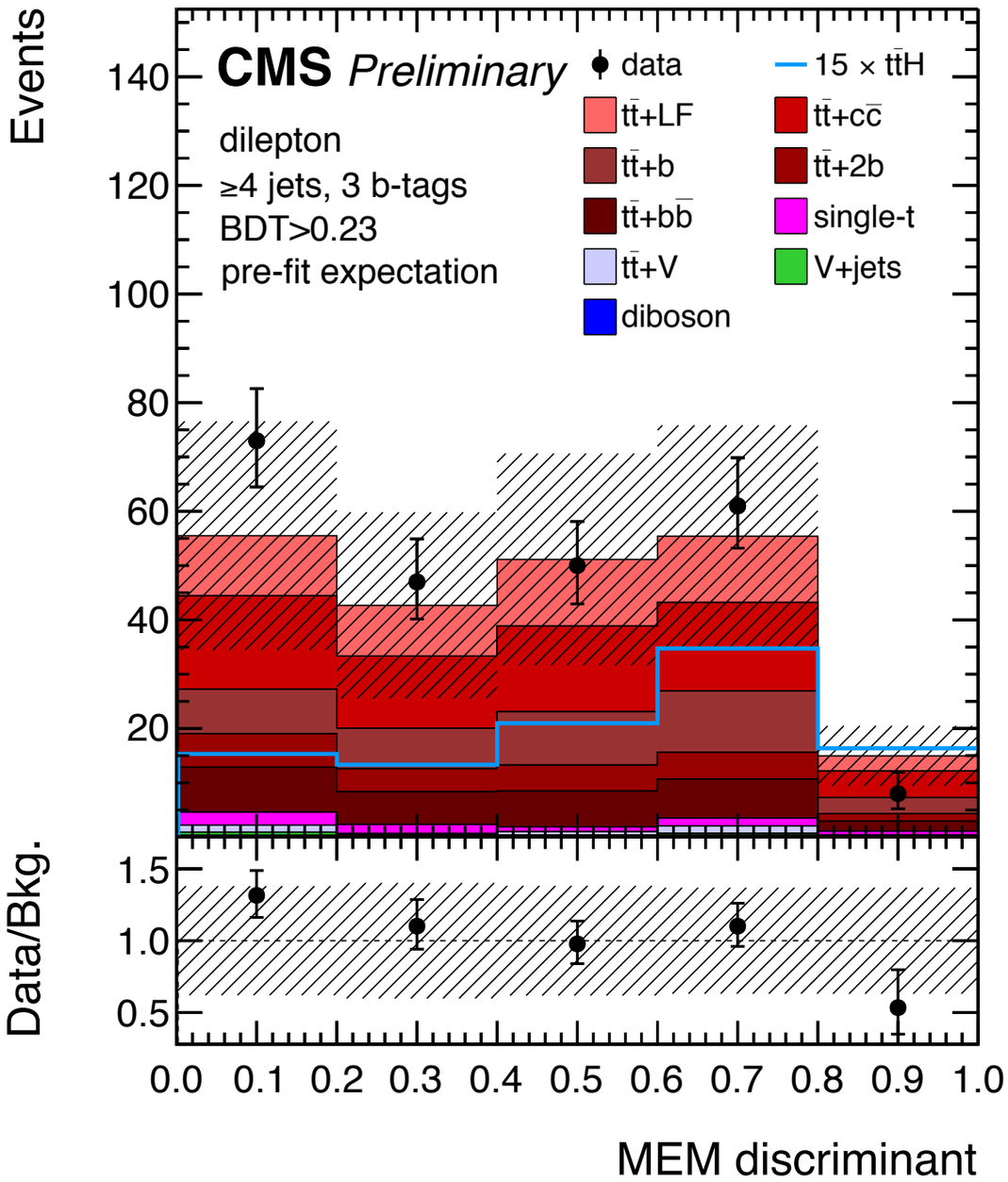


post-fit

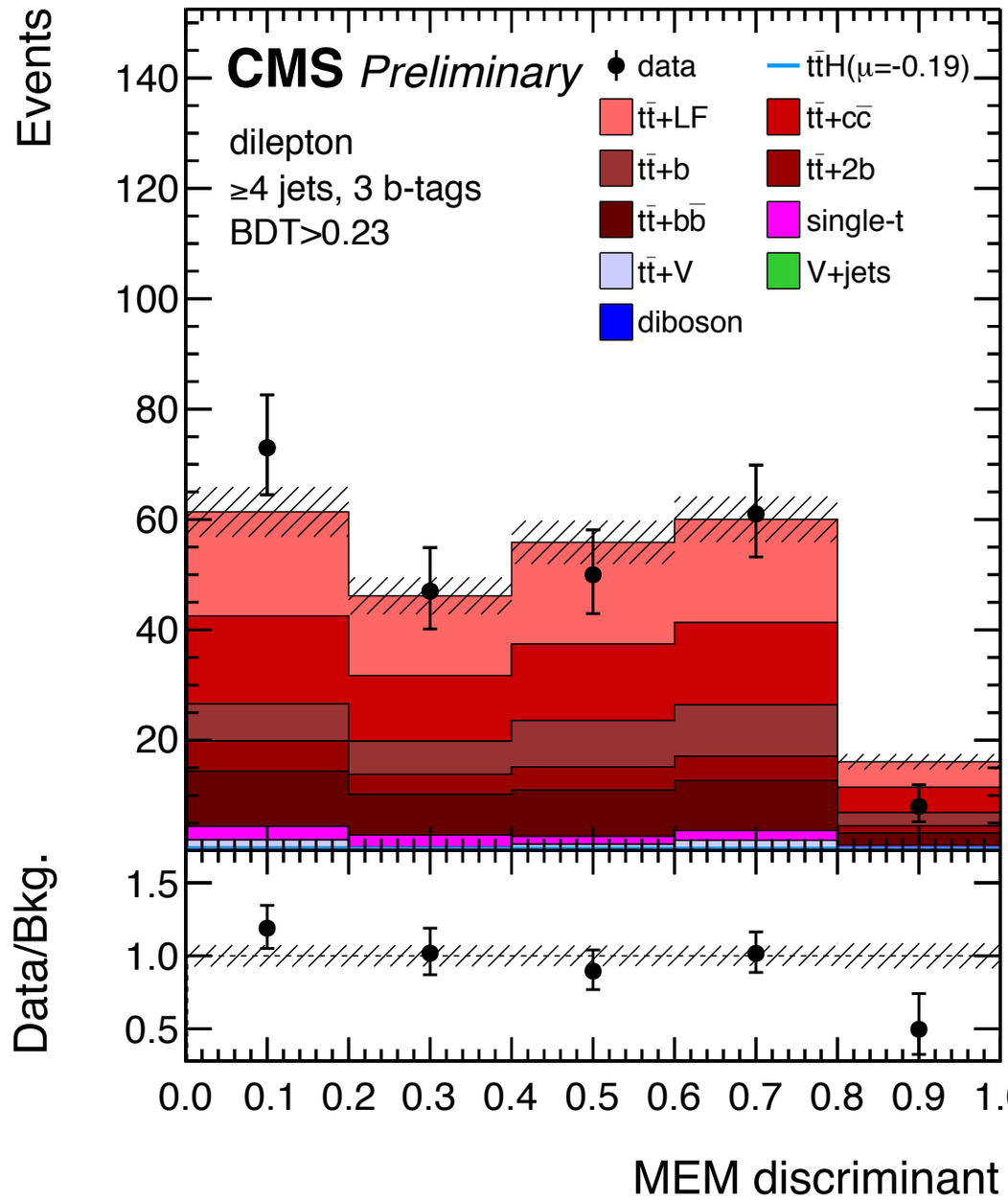


Dilepton j4 t3, high BDT

pre-fit

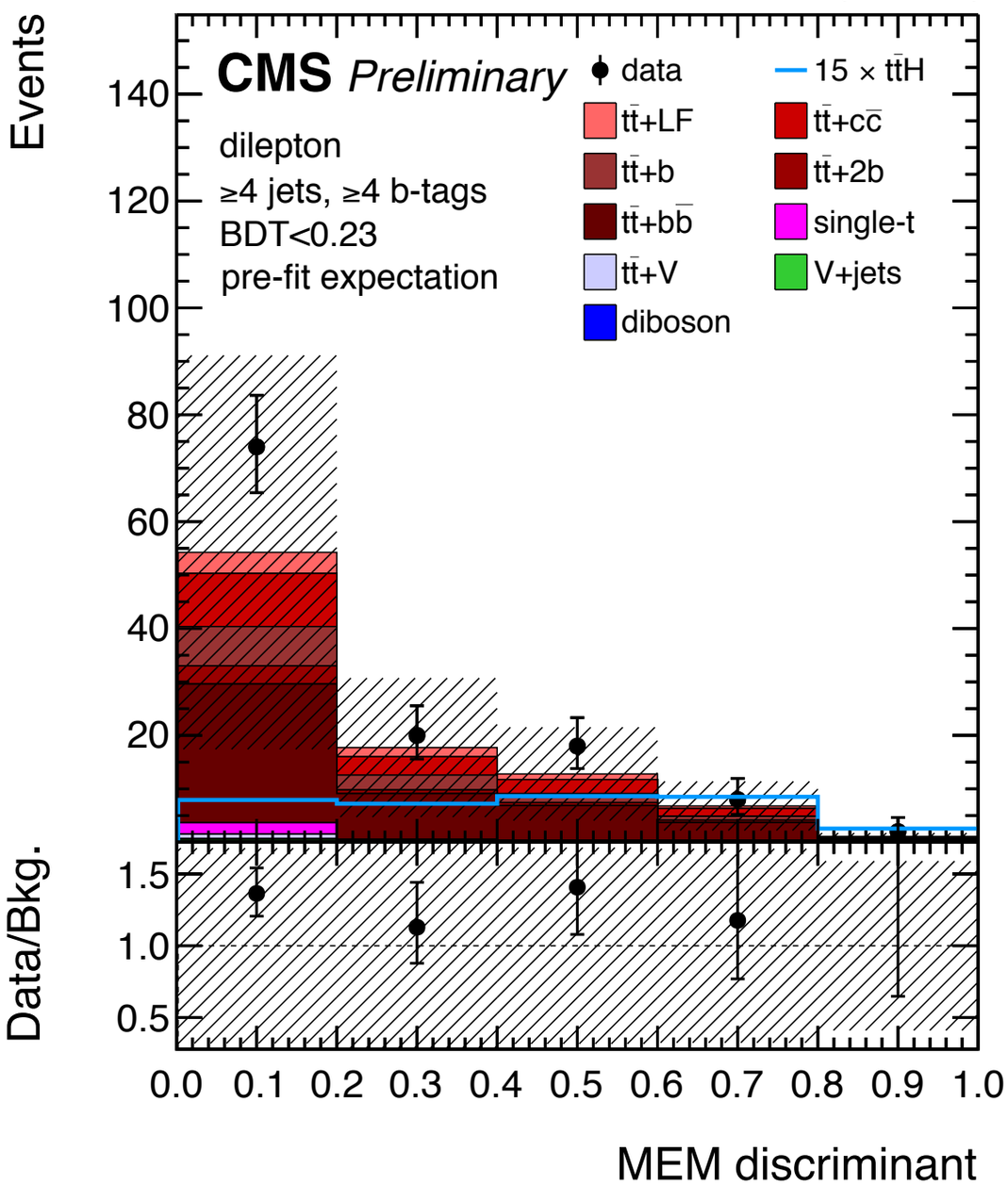


post-fit

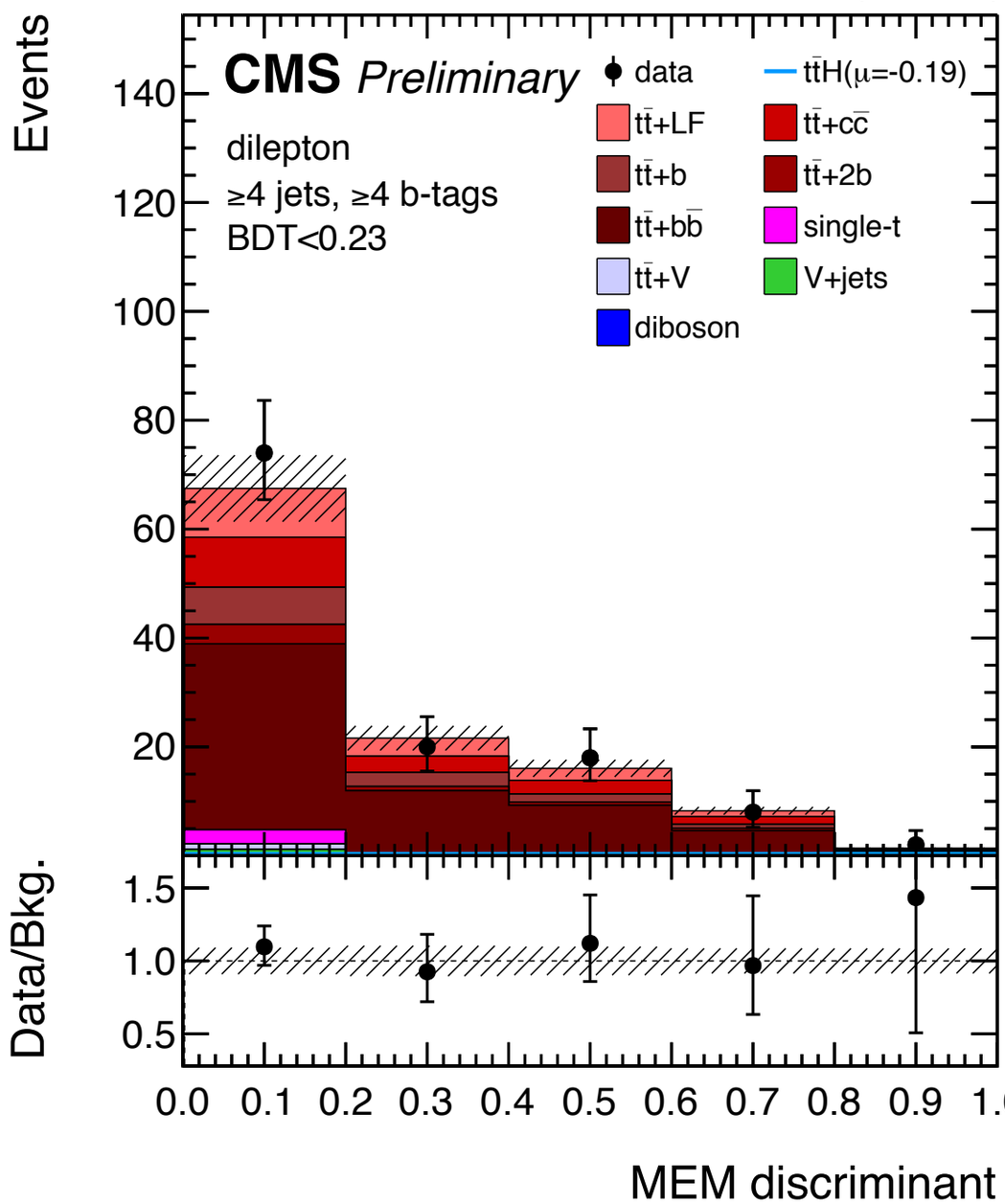


Dilepton j4 t4, low BDT

pre-fit

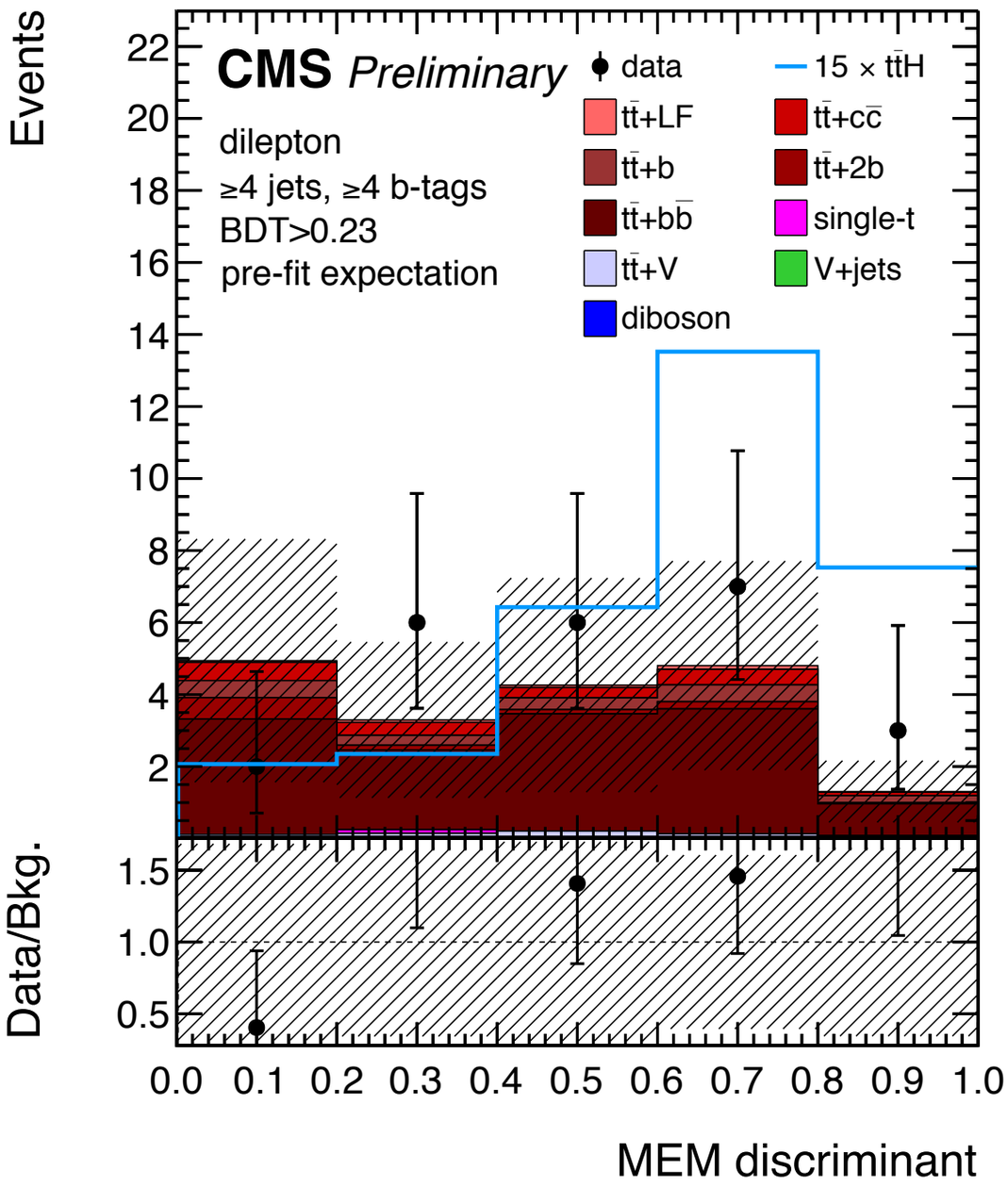


post-fit

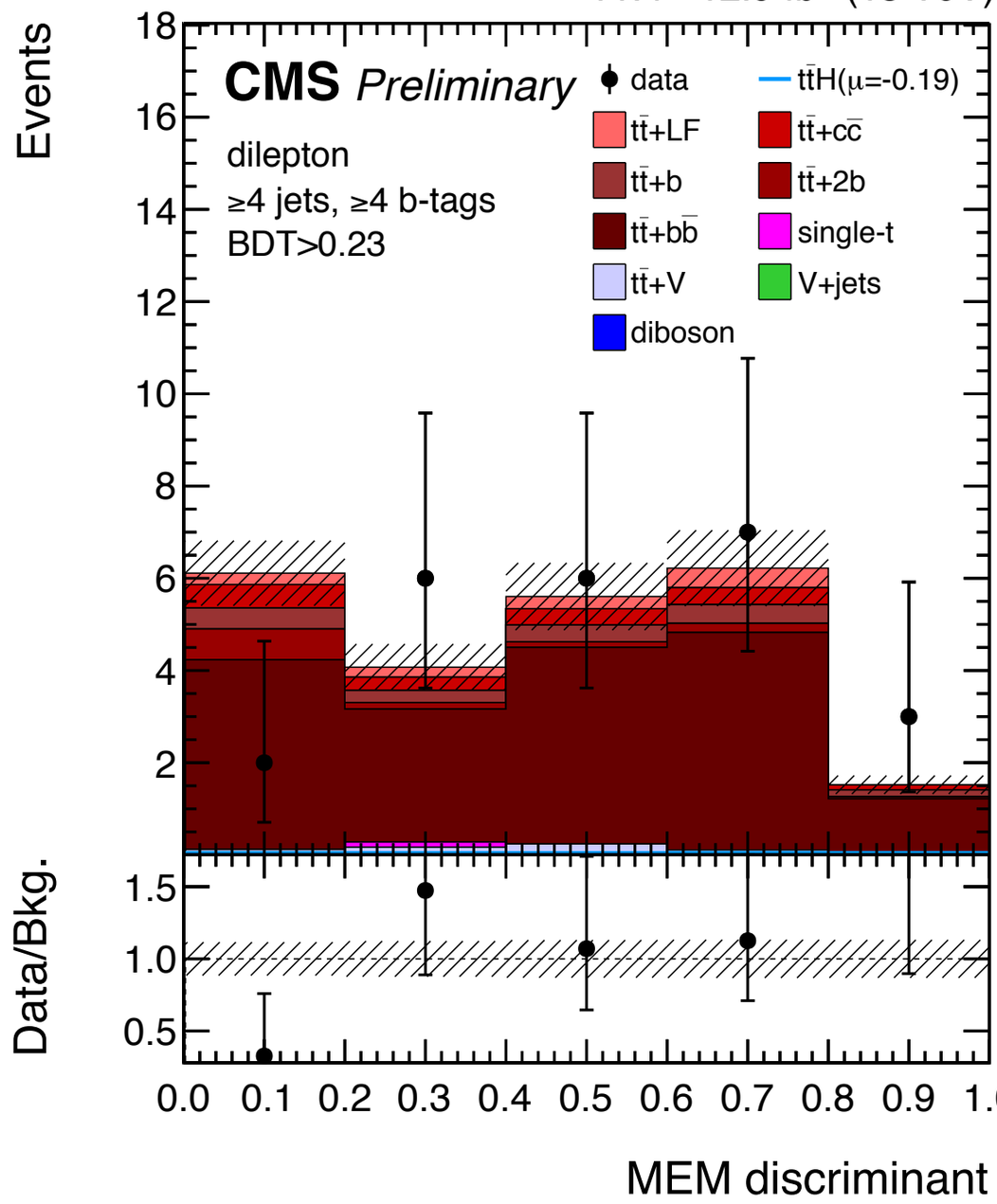


Dilepton j4 t4, high BDT

pre-fit



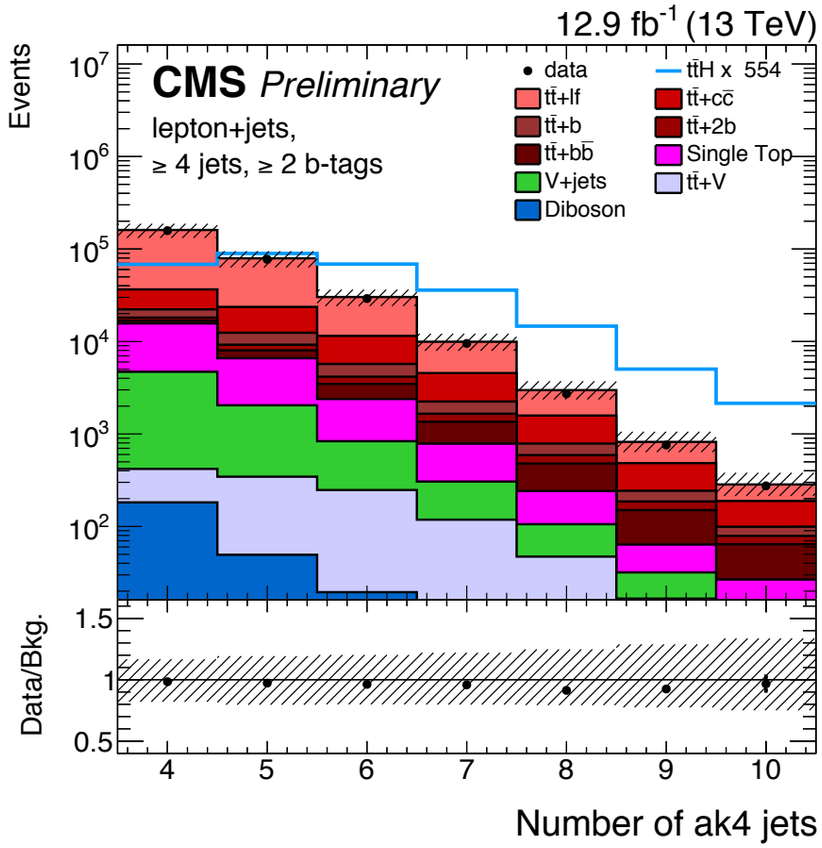
post-fit

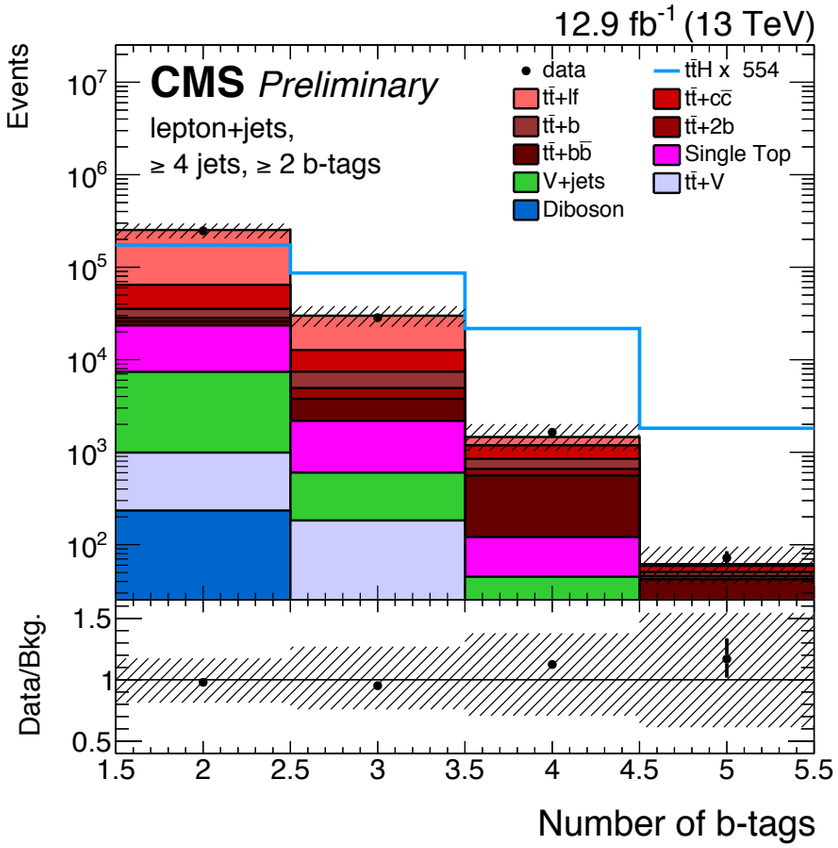


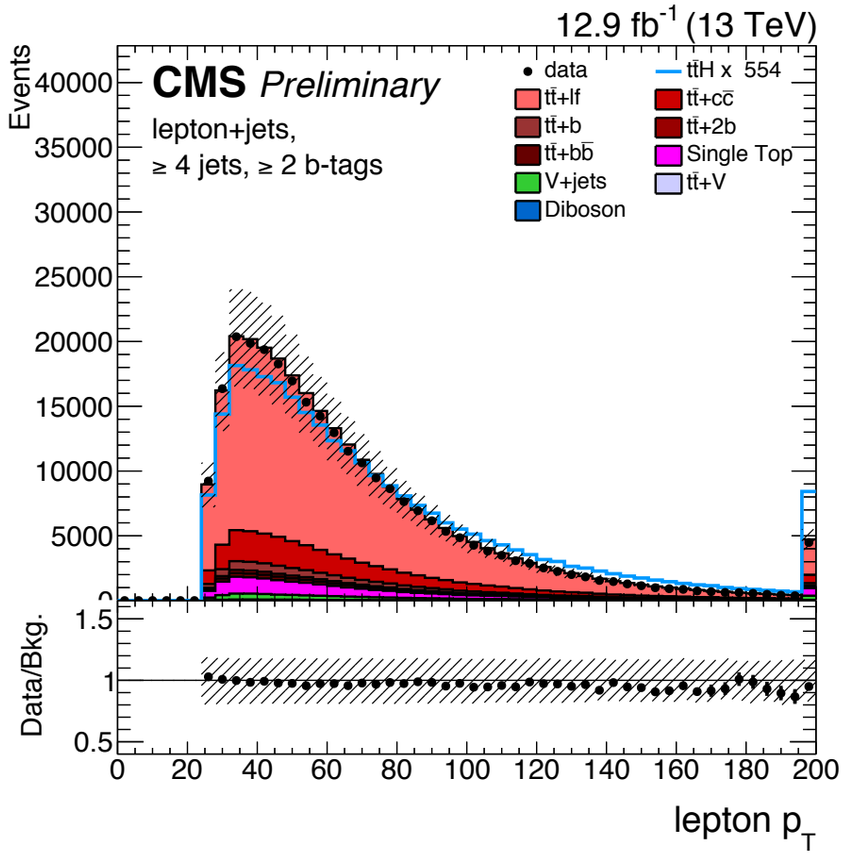
Extra Variables

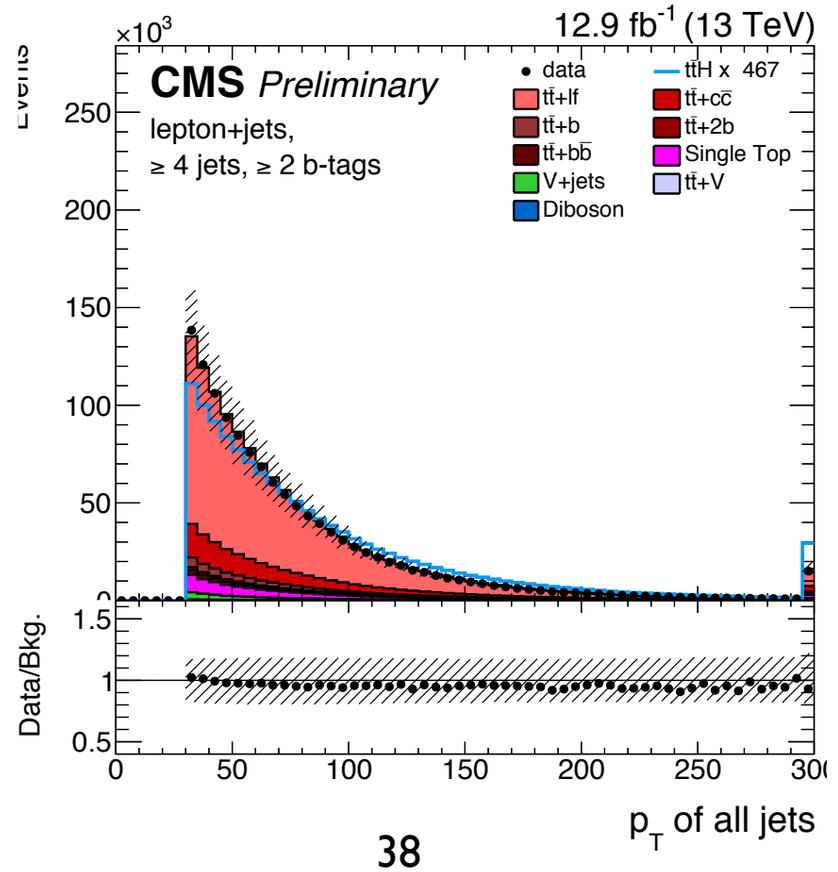
Lepton+Jets

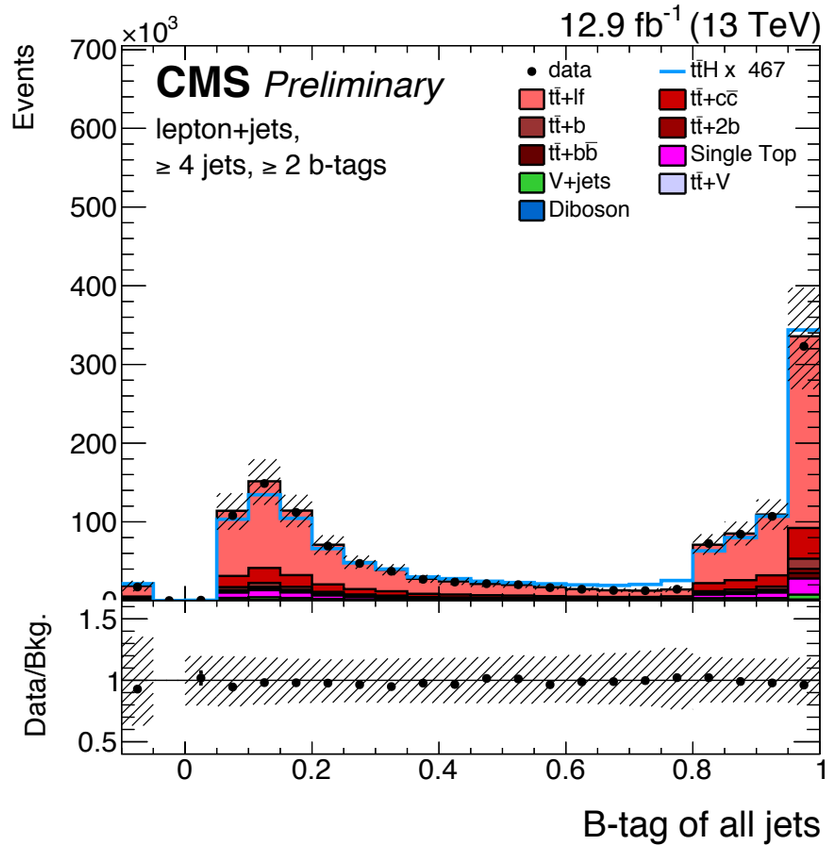
Lepton+Jets Basics



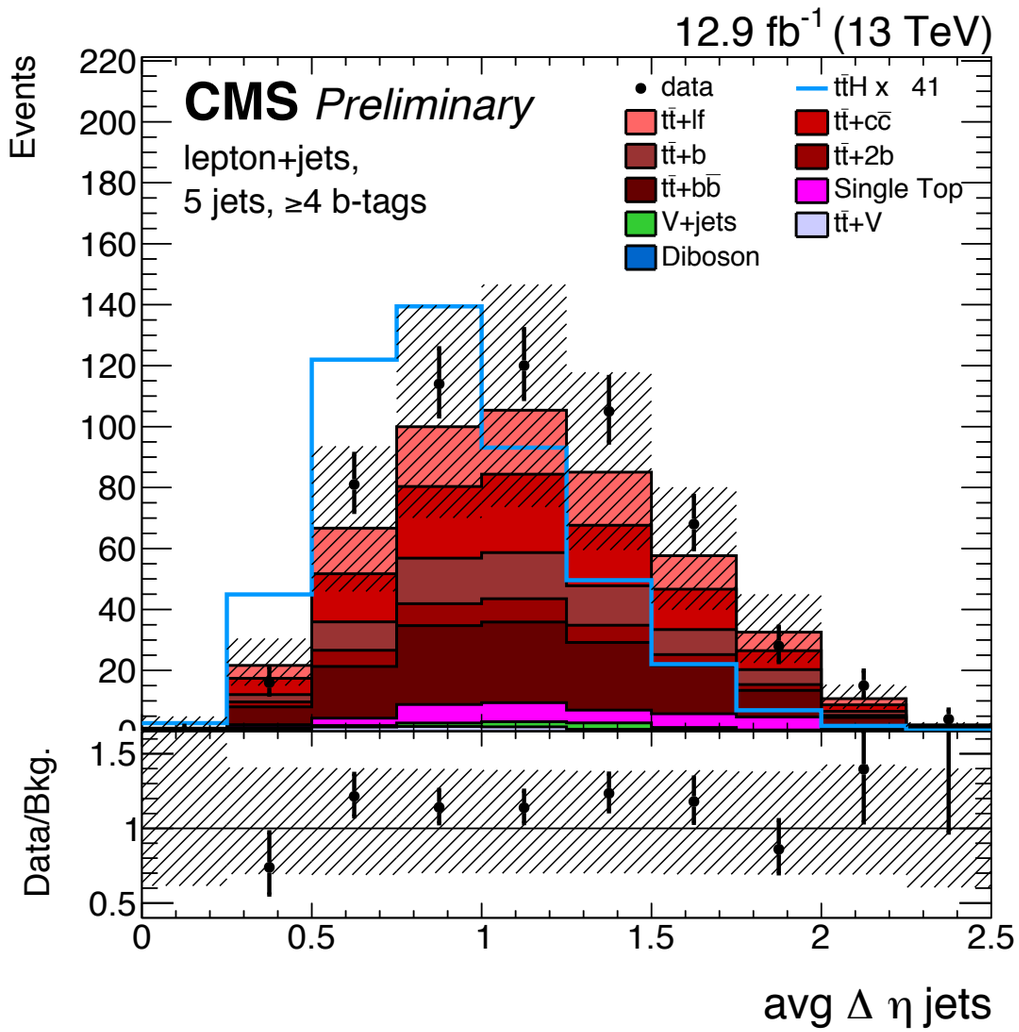


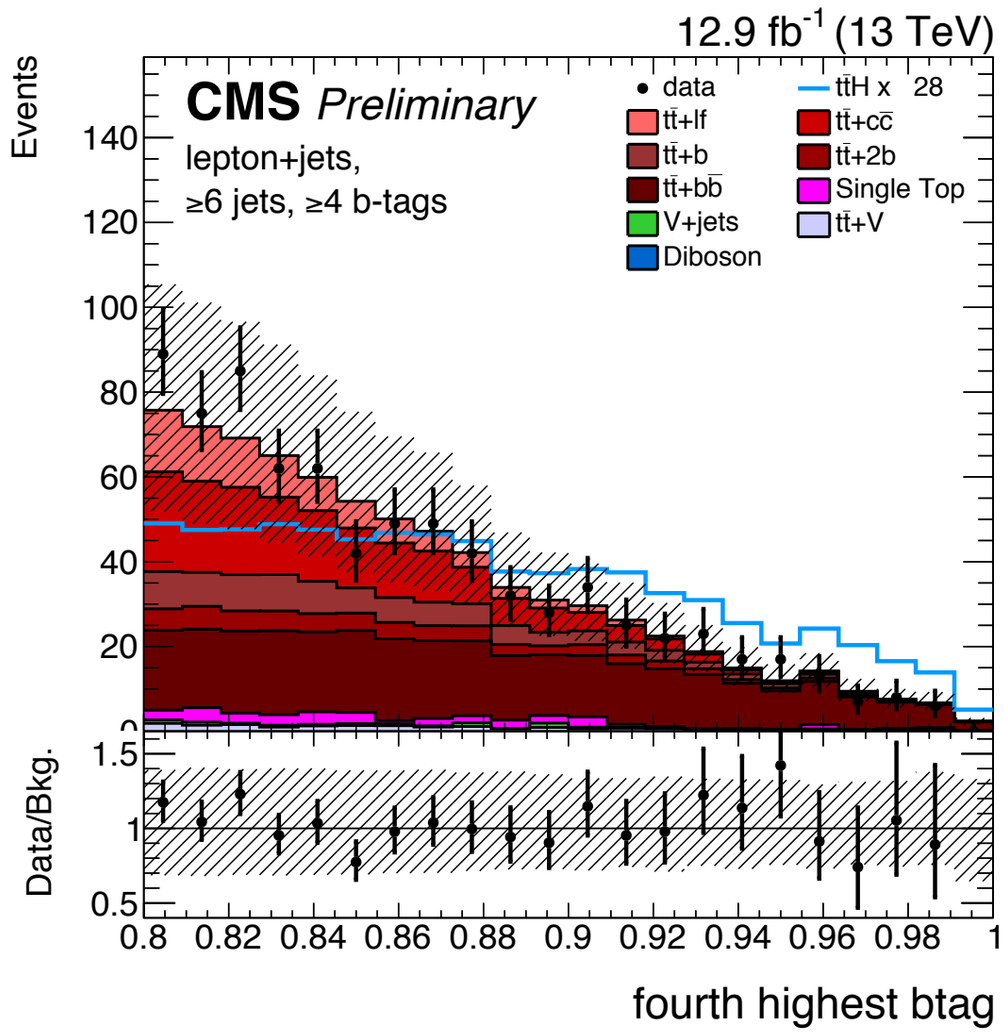


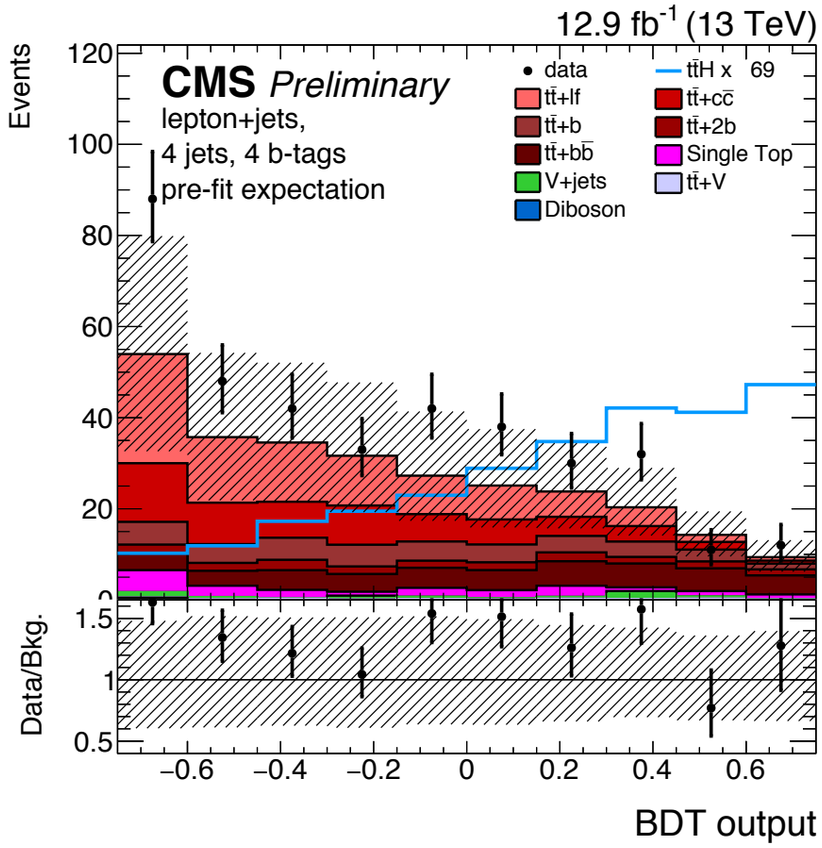


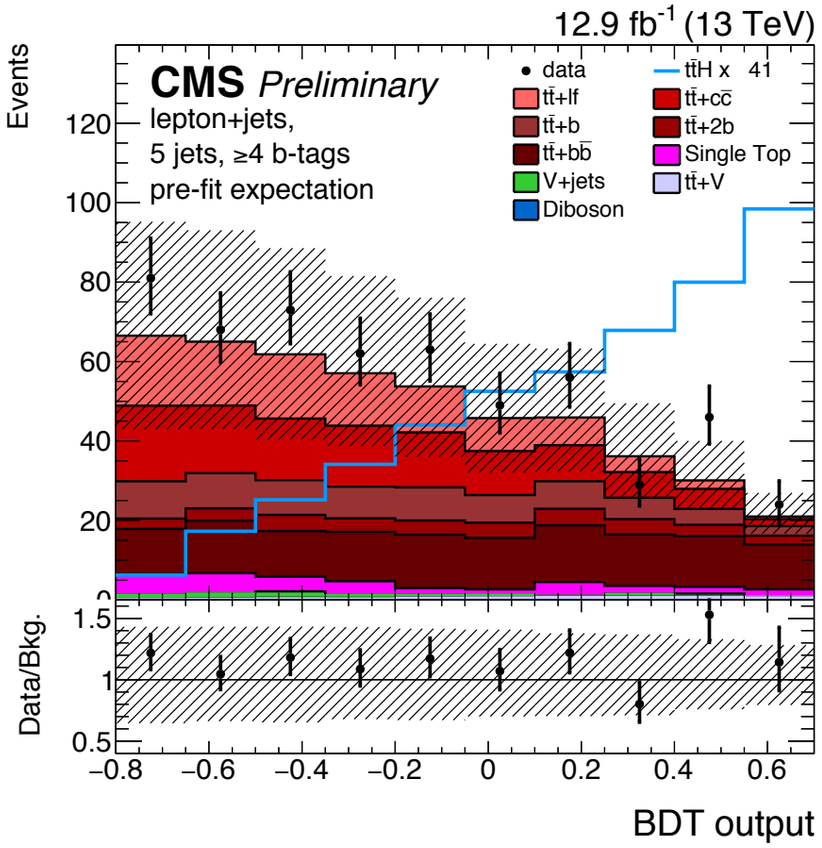


Lepton+Jets
BDT Inputs

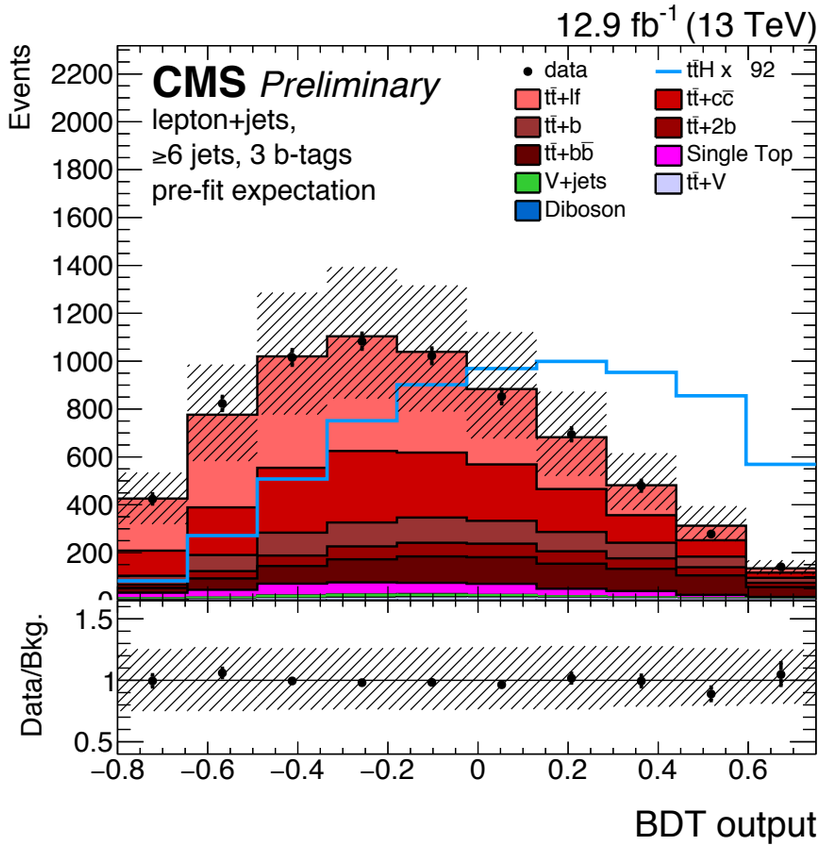


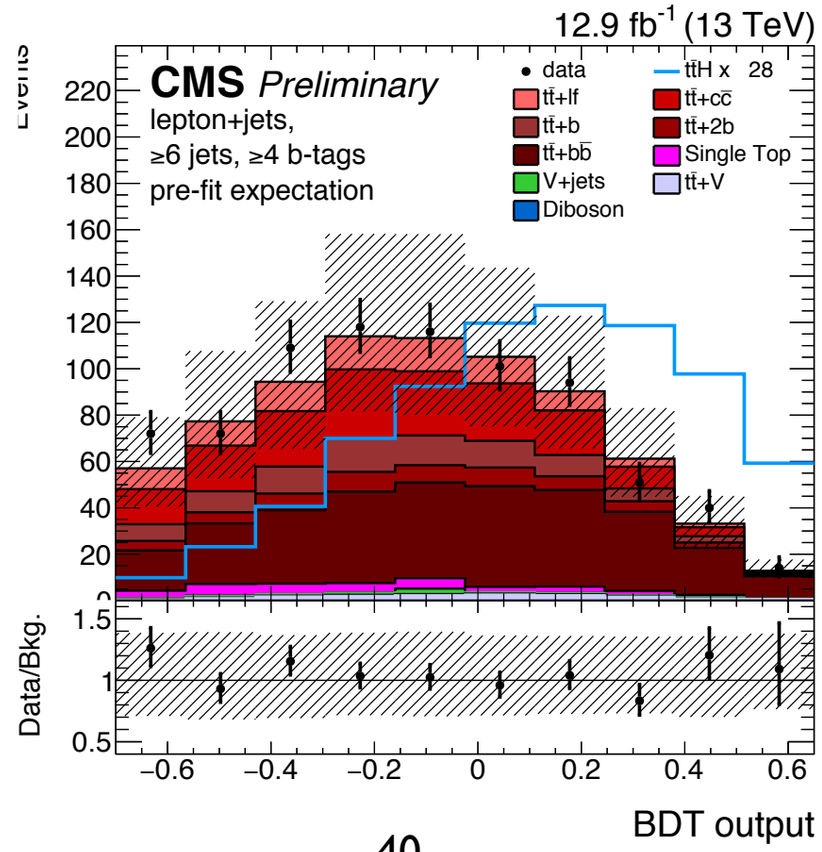






Lepton+Jets BDT Output





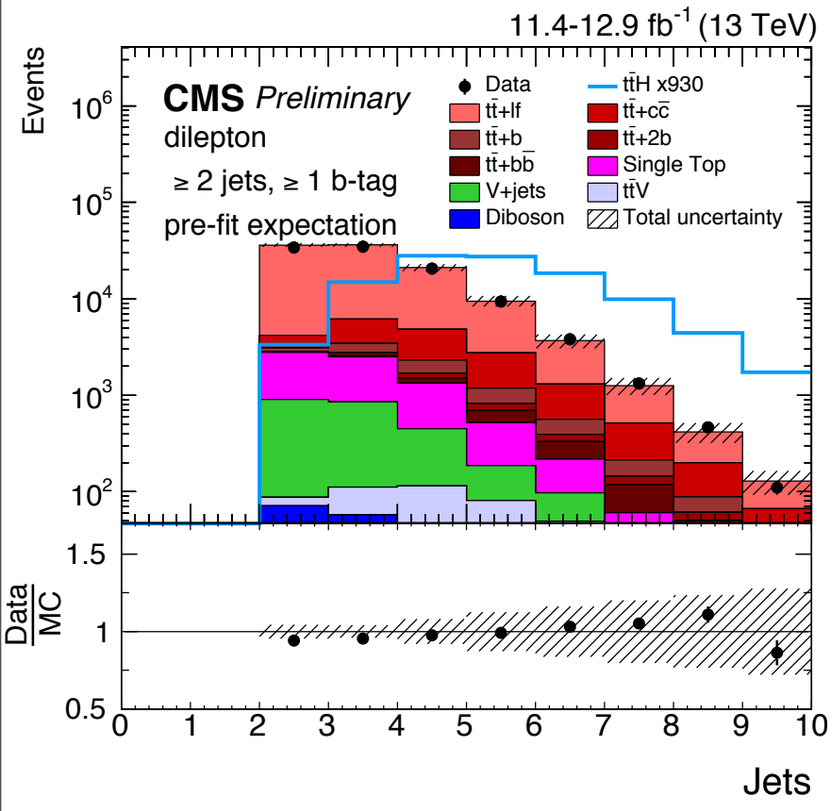
List of BDT Inputs

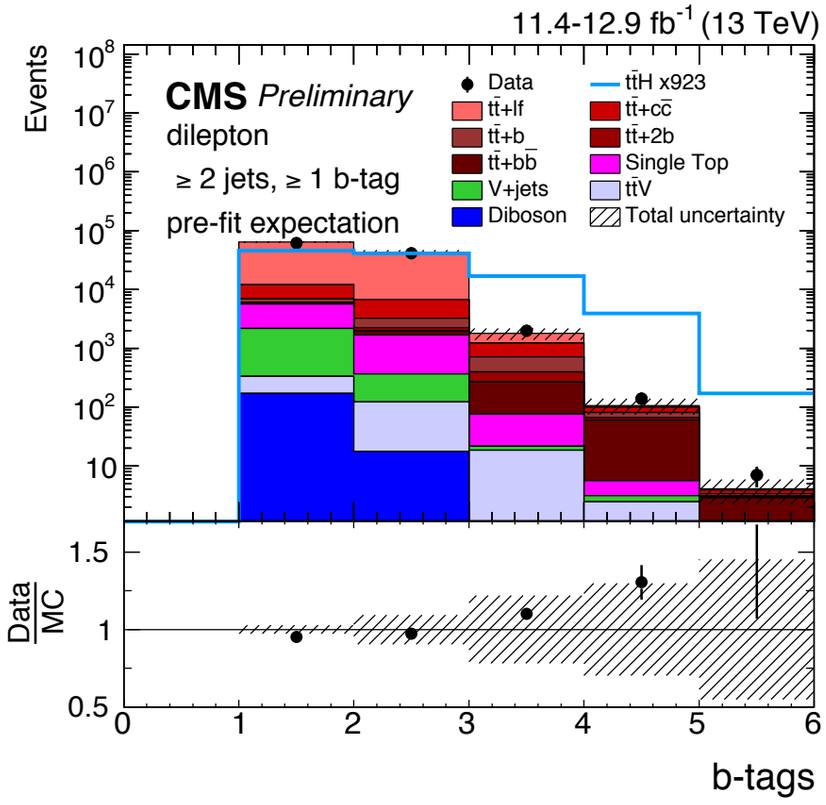
4 jets, 4 tags	5 jets, ≥ 4 tags
$\sum p_{\mathrm{T}}(\mathrm{jets}, \mathrm{lepton}, \mathrm{MET})$ avg. CSVv2 of b-tagged jets aplanarity H_3 $(\Sigma p_{\mathrm{T}}(\mathrm{jet}))/(\Sigma E(\mathrm{jet}))$ M_2 of min $\Delta\mathrm{R}(\mathrm{tag}, \mathrm{tag})$	avg. $\Delta\eta(\mathrm{jet}, \mathrm{jet})$ HT avg. CSVv2 of b-tagged jets $M_2(\mathrm{tag}, \mathrm{tag})$ closest to 125 M_3 $\sum p_{\mathrm{T}}(\mathrm{jets}, \mathrm{lepton}, \mathrm{MET})$ M_2 of min $\Delta\mathrm{R}(\mathrm{tag}, \mathrm{tag})$ aplanarity avg. $\Delta\mathrm{R}(\mathrm{tag}, \mathrm{tag})$
≥ 6 jets, 3 tags	≥ 6 jets, ≥ 4 tags
aplanarity $\sqrt{\Delta\eta(\mathrm{t}^{\mathrm{lep}}, \mathrm{bb}) \times \Delta\eta(\mathrm{t}^{\mathrm{had}}, \mathrm{bb})}$ $(\Sigma p_{\mathrm{T}}(\mathrm{jet}))/(\Sigma E(\mathrm{jet}))$ min $\Delta\mathrm{R}(\mathrm{tag}, \mathrm{tag})$ 2nd moment of b-tagged jets' CSVv2 $\sum p_{\mathrm{T}}(\mathrm{jets}, \mathrm{lepton}, \mathrm{MET})$ b-tagging likelihood ratio	best Higgs mass $M_2(\mathrm{tag}, \mathrm{tag})$ closest to 125 $M(\mathrm{jets}, \mathrm{lepton}, \mathrm{MET})$ 4th highest CSVv2 $\sum p_{\mathrm{T}}(\mathrm{jets}, \mathrm{lepton}, \mathrm{MET})$ 5th highest CSVv2

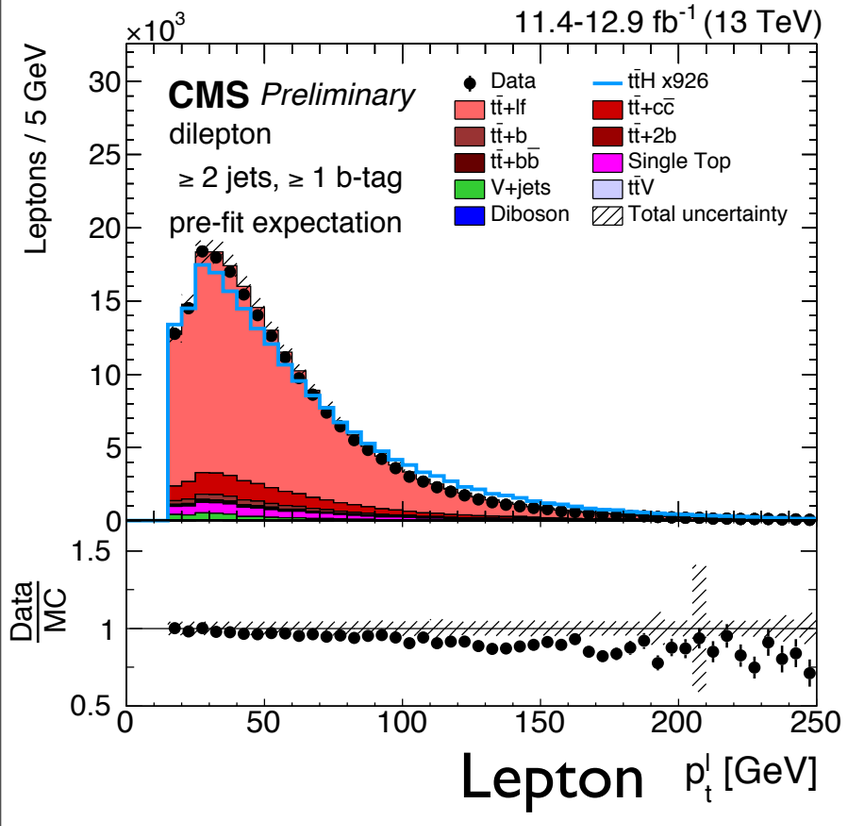
Extra Variables

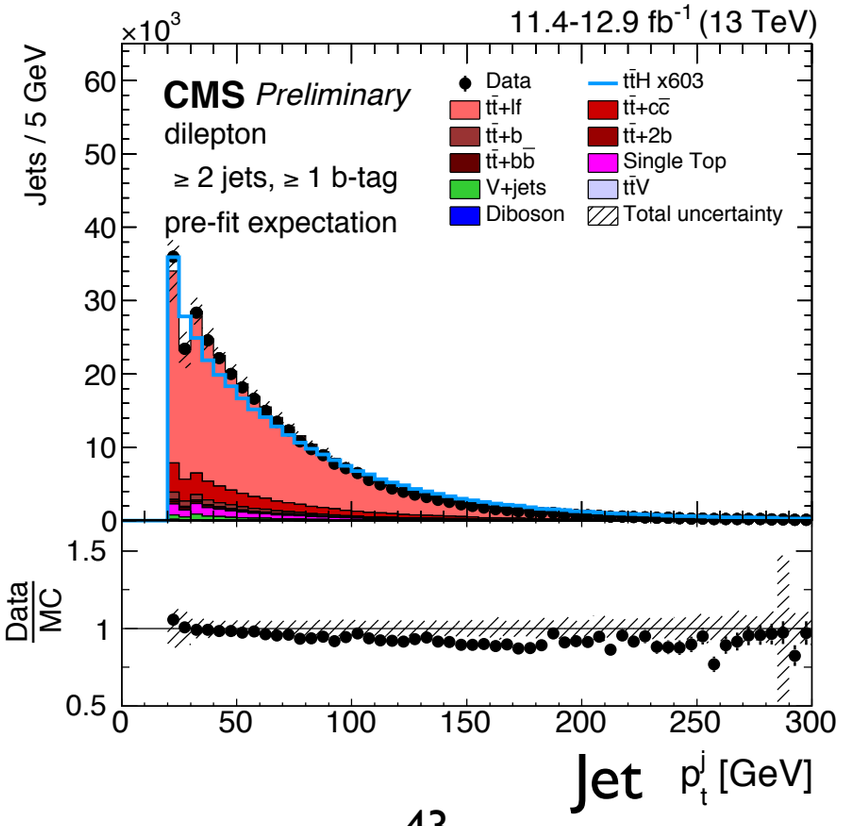
Dilepton

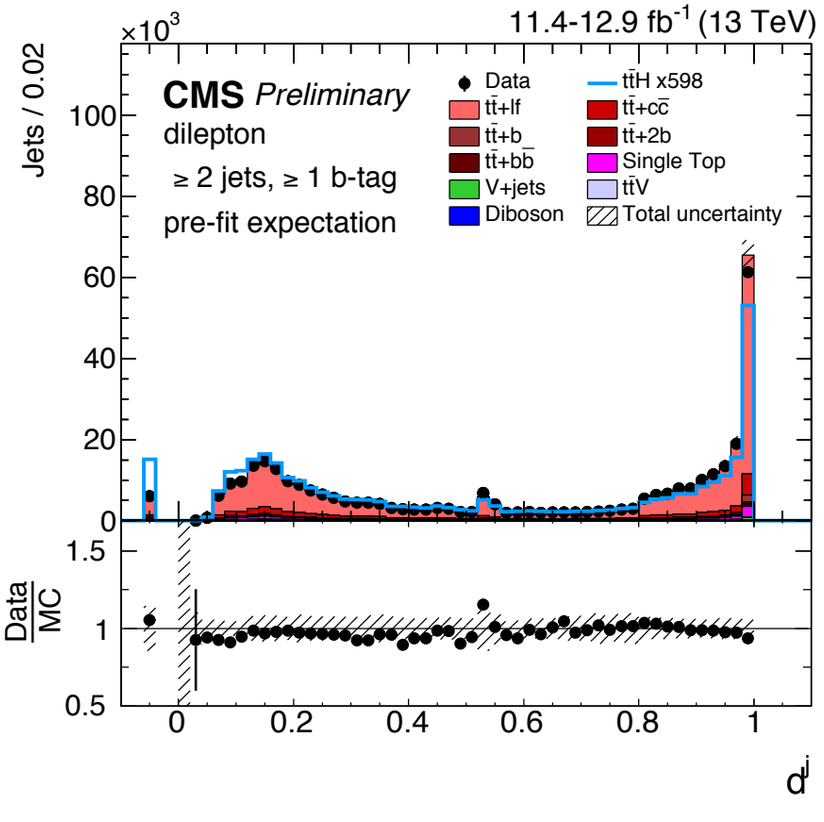
Dilepton Basics



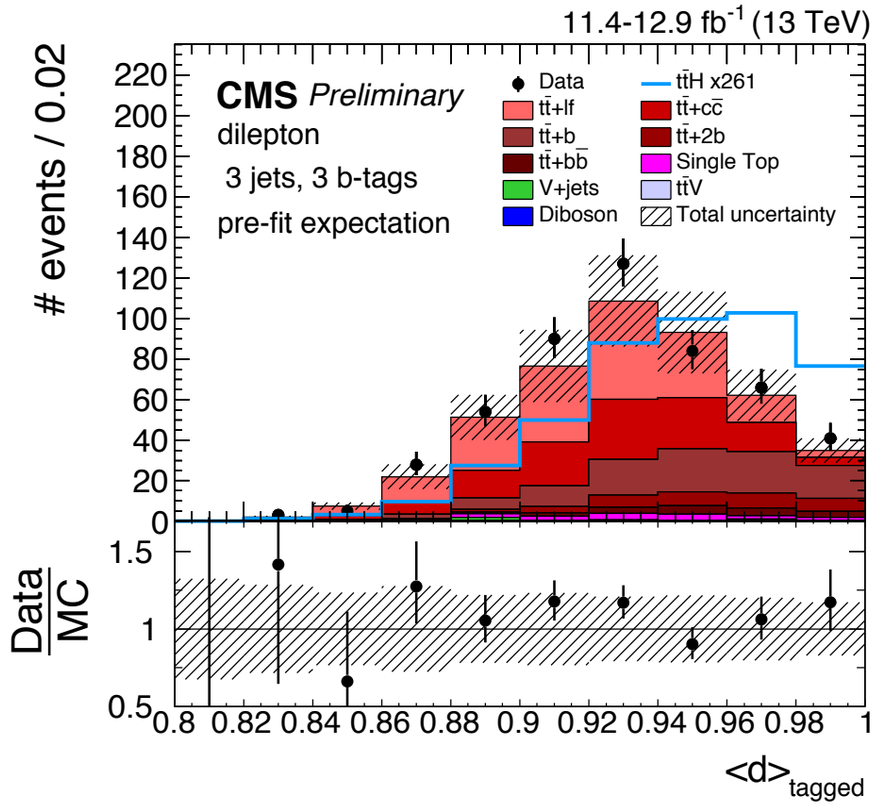


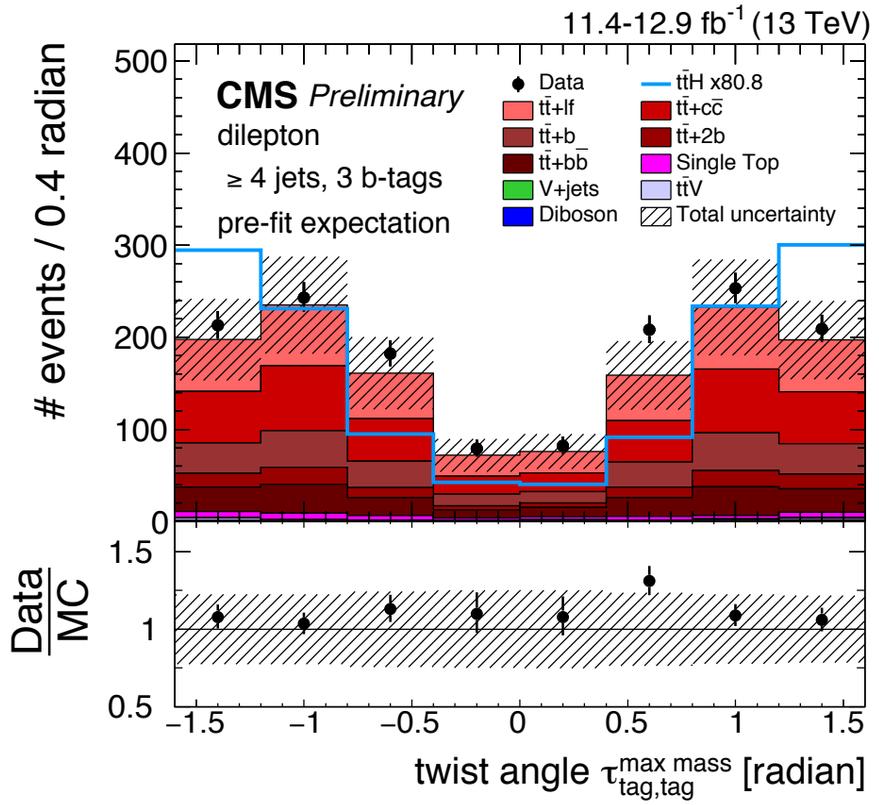


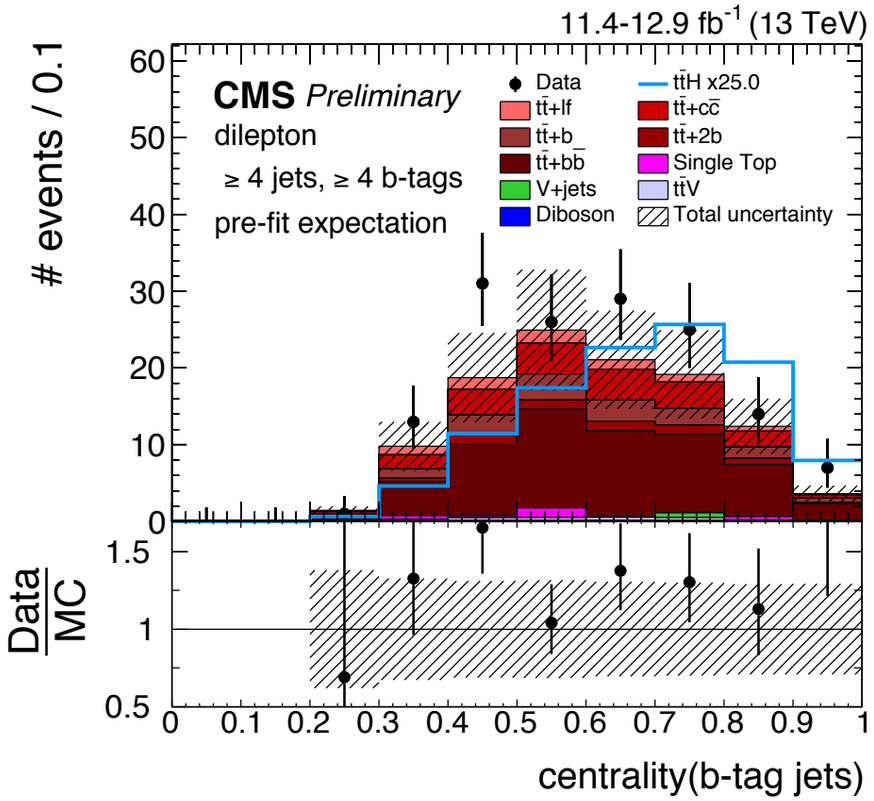




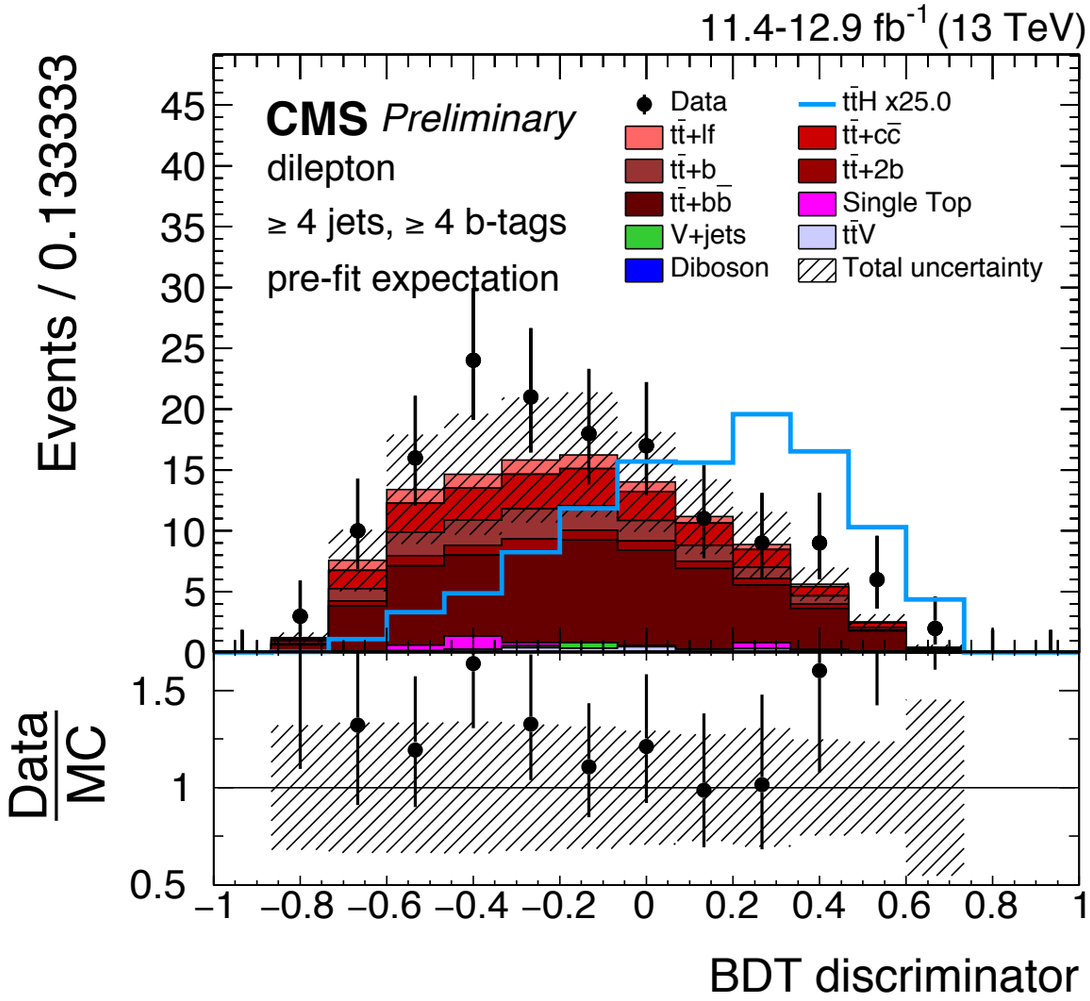
Dilepton BDT Input

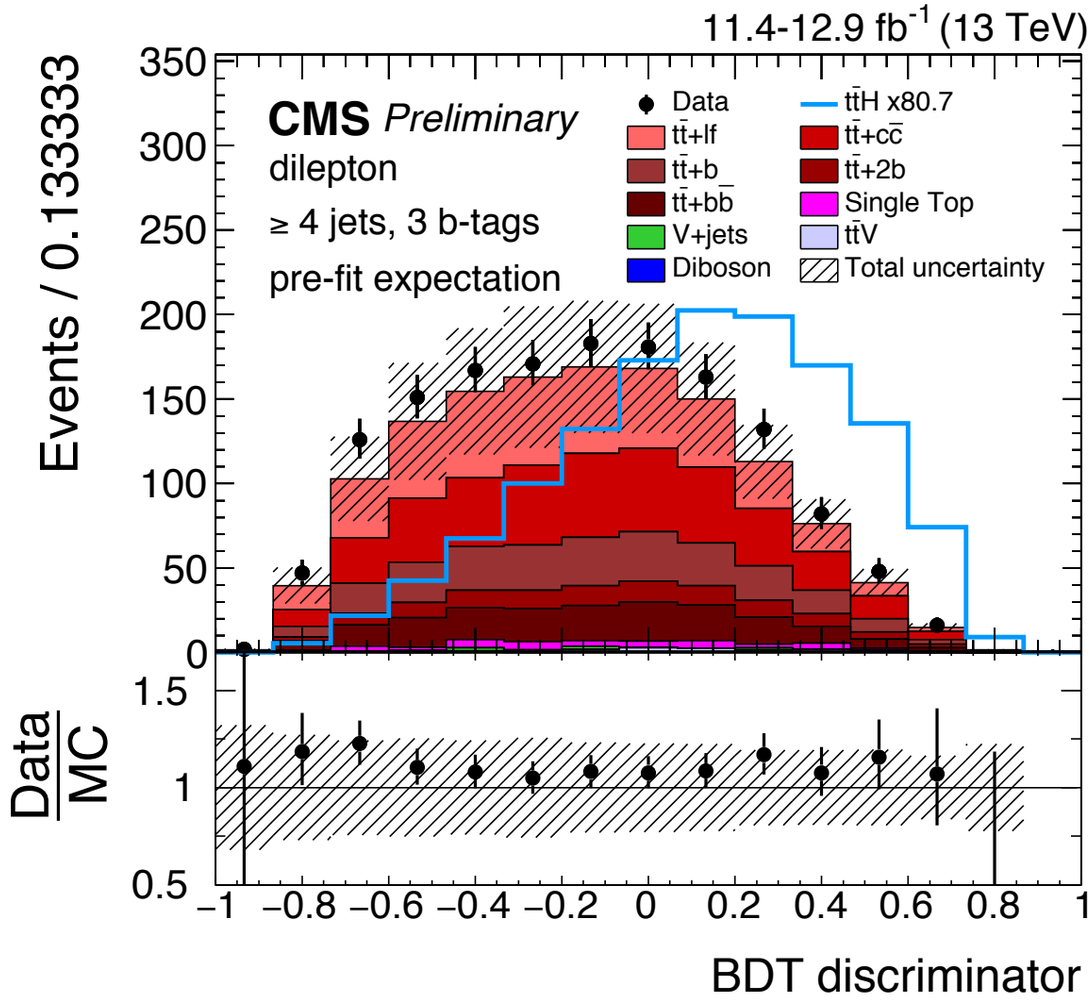






Dilepton
BDT Ouput





List of BDT Inputs

3 jets, 3 tags	$\geq$4 jets, 3 tags	$\geq$4 jets, $\geq$4 tags
$\langle d \rangle_{\text{tagged}}$	Centrality(jets & leptons)	Centrality(jets & leptons)
$H_1(\text{jets})$	C(jets)	Centrality(tags)
$M^{\text{bj}}_{\text{higgs-like}}$	$H_2(\text{tags})$	H_T^{tags}
$M^{\text{max mass}}_{\text{tag,tag}}$	$M^{\text{jj}}_{\text{higgs-like}}$	$M^{\text{jj}}_{\text{higgs-like}}$
min $\Delta R_{\text{tag,tag}}$	$M^{\text{max}p_T}_{\text{jet,jet,jet}}$	min $\Delta R_{\text{jet,jet}}$
max $\Delta\eta_{\text{jet,jet}}$	$M^{\text{min}\Delta R}_{\text{tag,tag}}$	$M^{\text{min}\Delta R}_{jet,tag}$
min $\Delta R_{\text{jet,jet}}$	min $\Delta R_{\text{tag,tag}}$	$M^{\text{max mass}}_{\text{tag,tag}}$
$\sum p_{T\text{jets,leptons}}$	max $\Delta\eta_{\text{tag,tag}}$	$M^{\text{min}\Delta R}_{\text{tag,tag}}$
$\text{H}_4/\text{H}_0(\text{tags})$	$\tau^{\text{max mass}}_{\text{tag,tag}}$	max $\Delta\eta_{\text{jet,jet}}$
		max $\Delta\eta_{\text{tag,tag}}$
		median $M_{\text{jet,jet}}$

Systematics

Source	Type	Remarks
Luminosity	rate	Signal and all backgrounds
Lepton ID/Iso	shape	Signal and all backgrounds
Trigger efficiency	shape	Signal and all backgrounds
Pileup	shape	Signal and all backgrounds
Jet energy scale	shape	Signal and all backgrounds
Jet energy resolution	shape	Signal and all backgrounds
b-tag HF fraction	shape	Signal and all backgrounds
b-tag HF stats (linear)	shape	Signal and all backgrounds
b-tag HF stats (quadratic)	shape	Signal and all backgrounds
b-tag LF fraction	shape	Signal and all backgrounds
b-tag LF stats (linear)	shape	Signal and all backgrounds
b-tag LF stats (quadratic)	shape	Signal and all backgrounds
b-tag charm (linear)	shape	Signal and all backgrounds
b-tag charm (quadratic)	shape	Signal and all backgrounds
QCD scale ($t\bar{t}H$)	rate	Scale uncertainty of NLO $t\bar{t}H$ prediction
QCD scale ($t\bar{t}$)	rate	Scale uncertainty of NLO $t\bar{t}$ prediction
QCD scale ($t\bar{t}$+HF)	rate	Additional 50% rate uncertainty of $t\bar{t}$+HF predictions
QCD scale (t)	rate	Scale uncertainty of NLO single t prediction
QCD scale (V)	rate	Scale uncertainty of NNLO W and Z prediction
QCD scale (VV)	rate	Scale uncertainty of NLO diboson prediction
pdf (gg)	rate	PDF uncertainty for gg initiated processes except $t\bar{t}H$
pdf (gg $t\bar{t}H$)	rate	PDF uncertainty for $t\bar{t}H$
pdf ($q\bar{q}$)	rate	PDF uncertainty of $q\bar{q}$ initiated processes ($t\bar{t}$ W, W, Z)
pdf (qg)	rate	PDF uncertainty of qg initiated processes (single t)
Q^2 scale ($t\bar{t}$)	shape	Renormalization and factorization scale uncertainties of the $t\bar{t}$ ME generator, independent for additional jet flavors
PS Scale ($t\bar{t}$)	rate	Renormalization and factorization scale uncertainties of the parton shower (for $t\bar{t}$ events), independent for additional jet flavors
Bin-by-bin statistics	shape	statistical uncertainty of the signal and background prediction due to the limited sample size

Pre-fit Shape impact on lepton+jets j6 t3

Process	$t\bar{t}$ rate up/down [%]	$t\bar{t}H$ rate up/down [%]
Jet energy scale	$+12.6/-11.8$	$+8.4/-8.0$
Jet energy resolution	$+0.2/-0.3$	$-0.0/-0.1$
Pile-up	$+0.1/-0.1$	$-0.2/+0.1$
Electron efficiency	$+0.5/-0.5$	$+0.5/-0.5$
Muon efficiency	$+0.4/-0.4$	$+0.4/-0.4$
Electron trigger efficiency	$+1.2/-1.2$	$+1.3/-1.3$
Muon trigger efficiency	$+0.8/-0.8$	$+0.9/-0.9$
b-Tag HF contamination	$-9.4/+9.8$	$-2.6/+2.8$
b-Tag HF stats (linear)	$-3.1/+3.3$	$-2.5/+2.7$
b-Tag HF stats (quadratic)	$+2.6/-2.4$	$+2.4/-2.2$
b-Tag LF contamination	$+7.1/-5.2$	$+5.8/-4.5$
b-Tag LF stats (linear)	$-2.0/+4.4$	$+0.5/+1.5$
b-Tag LF stats (quadratic)	$+2.1/+0.2$	$+1.5/+0.5$
b-Tag charm Uncertainty (linear)	$-11.1/+14.9$	$-3.1/+4.1$
b-Tag charm Uncertainty (quadratic)	$+0.5/-0.5$	$-0.0/+0.0$
Q^2 scale ($t\bar{t}$+LF)	$-6.2/+7.5$	$-$
Q^2 scale ($t\bar{t}$+b)	$-1.7/+2.0$	$-$
Q^2 scale ($t\bar{t}$+2b)	$-1.1/+1.4$	$-$
Q^2 scale ($t\bar{t}$+$b\bar{b}$)	$-2.0/+2.5$	$-$
Q^2 scale ($t\bar{t}$+$c\bar{c}$)	$-4.3/+5.4$	$-$
PS scale ($t\bar{t}$+LF)	$+4.8/-9.0$	$-$
PS scale ($t\bar{t}$+b)	$-0.9/+0.7$	$-$
PS scale ($t\bar{t}$+2b)	$-0.8/+0.9$	$-$
PS scale ($t\bar{t}$+$b\bar{b}$)	$-1.5/+2.7$	$-$
PS scale ($t\bar{t}$+$c\bar{c}$)	$-3.9/+3.0$	$-$