The Standard Model production of four top quarks in pp collisions at $\sqrt{s} = 13$ TeV at CMS

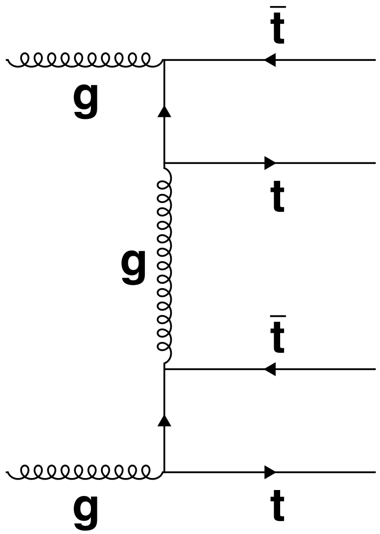
Lana Beck on behalf of the CMS Collaboration

20th Sept - TOP2016

CMS-PAS-TOP-16-016
http://cds.cern.ch/record/2205265
Made public 4th Aug 2016 – Presented at ICHEP 2016

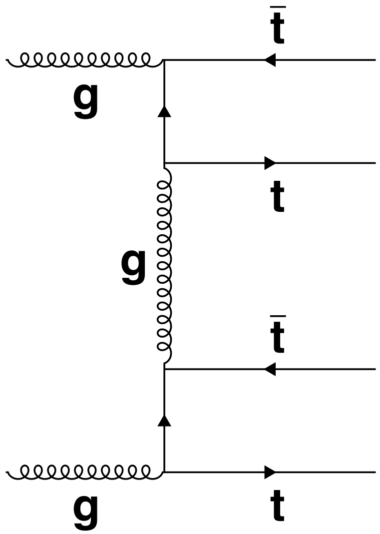
Motivation

- It is an extremely rare SM process
- Cross section of ≈ 9.2* fb at 13 TeV
- Cross section is enhanced in many BSM models [* arXiv:1405.0301]



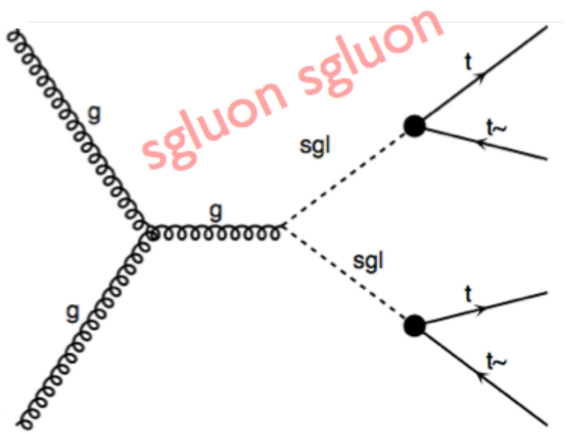
Motivation

- It is an extremely rare SM process
- Cross section of ≈ 9.2* fb at 13 TeV
- Cross section is enhanced in many BSM models [* arXiv:1405.0301]



Models with tttt signatures include:

- Supersymmetry
- New gauge forces
- Higgs boson or top quark compositeness
- Extra dimensions



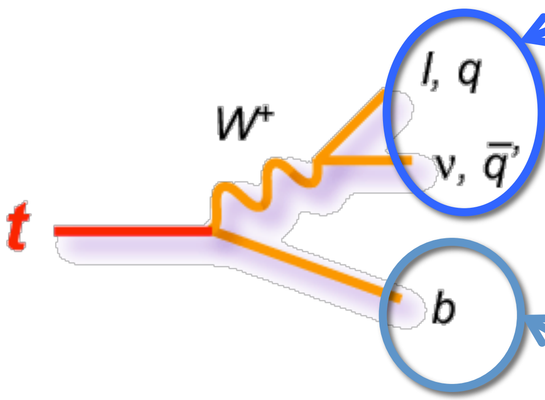
[arXiv:1212.3360]

- At $\sqrt{s} = 8$ TeV CMS set a 95% CL limit of 32 fb (25*σ_{obs}/σ_{SM}) where σ_{SM} was 1.3 fb

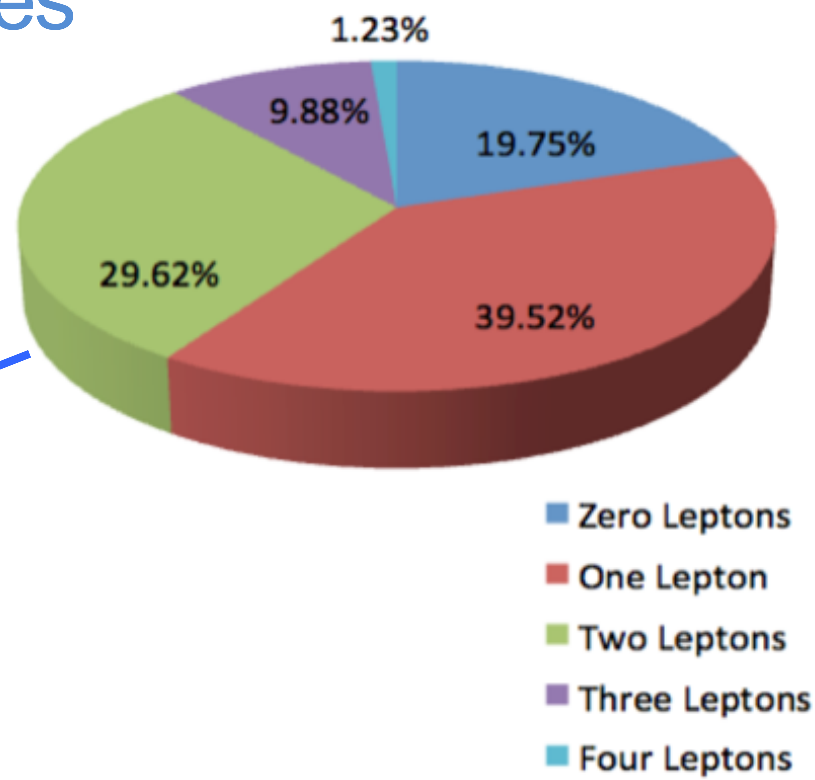
[arXiv:1409.7339]

Four Top Quark Final States

- Assume Standard Model branching ratio t⟶Wb



Four Top decay channels



- Single lepton + OS dilepton studied
- Final states with e / μ only
- ~40% branching fraction

b-quark identification is integral to top quark reconstruction

Strategy

Baseline selection - Select for $t\bar{t}$ + 2jets and require at least 2 b-tags

Strategy

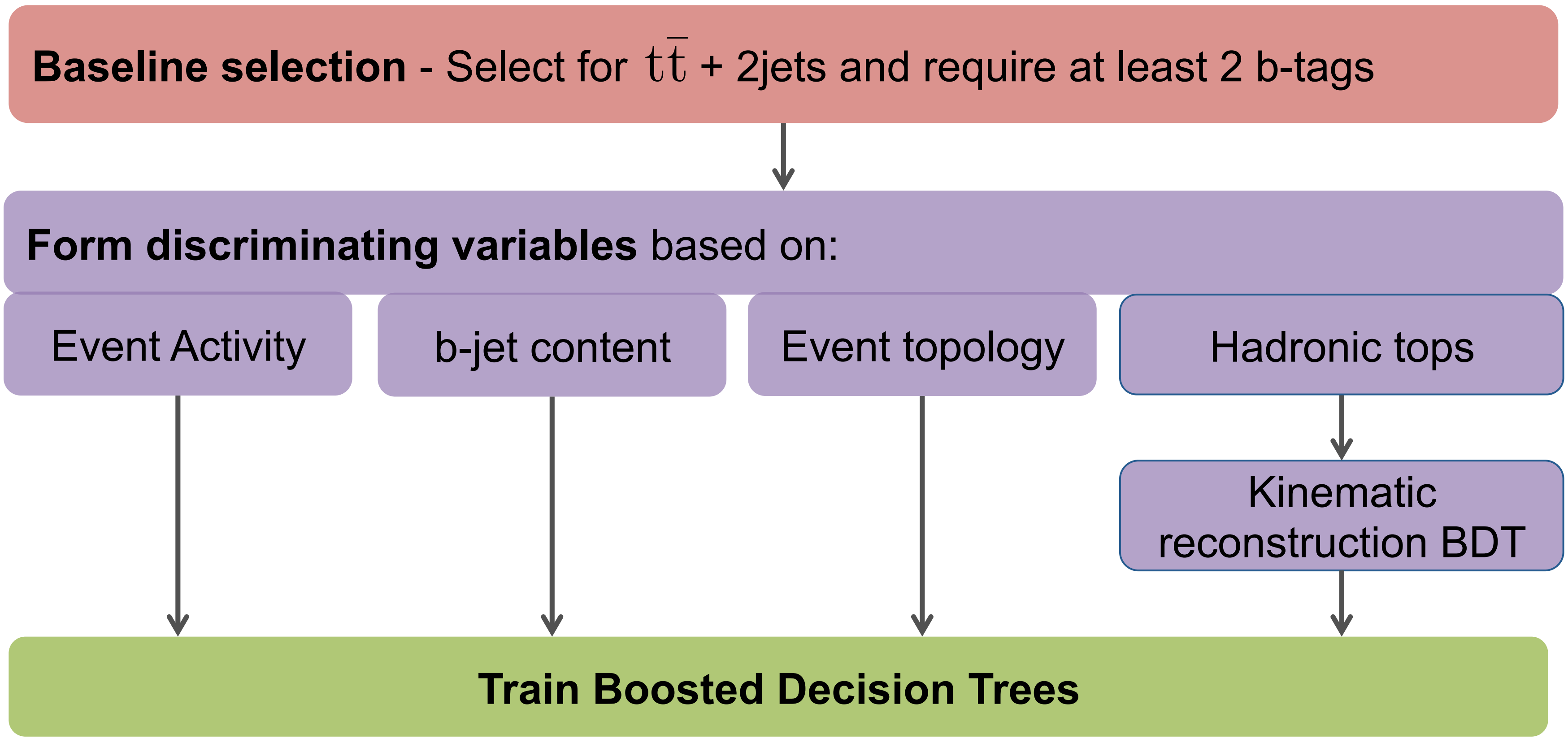
Baseline selection - Select for $t\bar{t}$ + 2jets and require at least 2 b-tags

↓

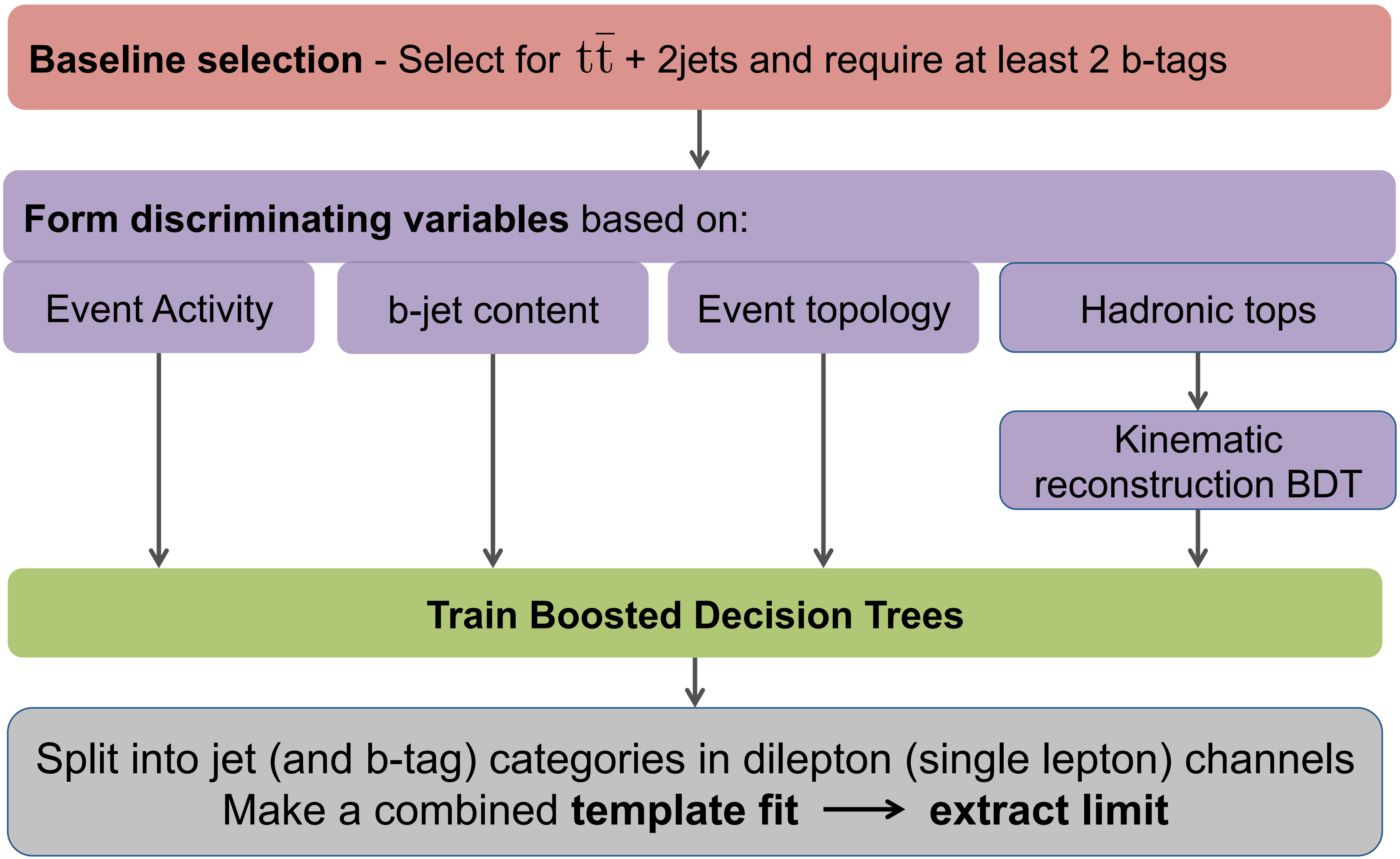
Form discriminating variables based on:

- Event Activity
- b-jet content
- Event topology
- Hadronic tops
  - ↓ Kinematic reconstruction BDT

Strategy

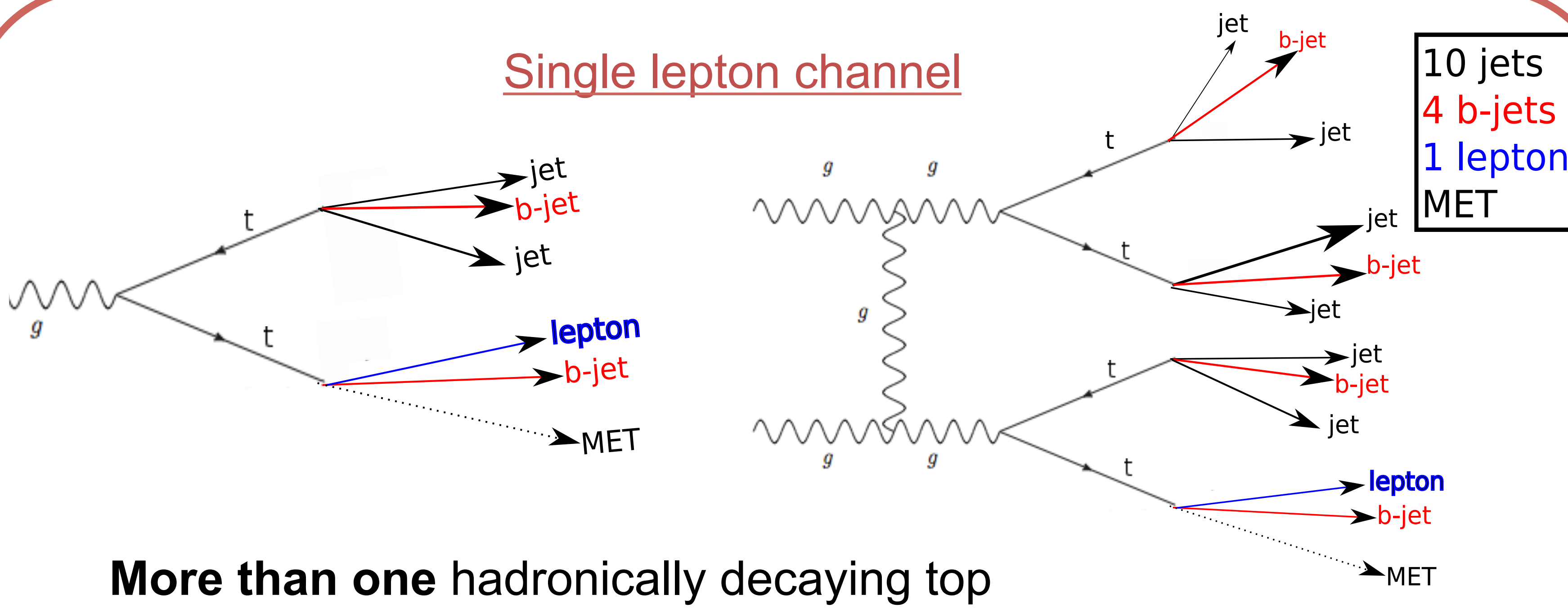

Baseline selection - Select for $t\bar{t}$ + 2jets and require at least 2 b-tags
Form discriminating variables based on:
Event Activity
b-jet content
Event topology
Hadronic tops
Kinematic reconstruction BDT
Train Boosted Decision Trees

Strategy


Baseline selection - Select for $t\bar{t}$ + 2jets and require at least 2 b-tags
Form discriminating variables based on:
Event Activity
b-jet content
Event topology
Hadronic tops
Kinematic reconstruction BDT
Train Boosted Decision Trees
Split into jet (and b-tag) categories in dilepton (single lepton) channels
Make a combined **template fit** ⟶ **extract limit**

Dominant background is ttbar production

Single lepton channel

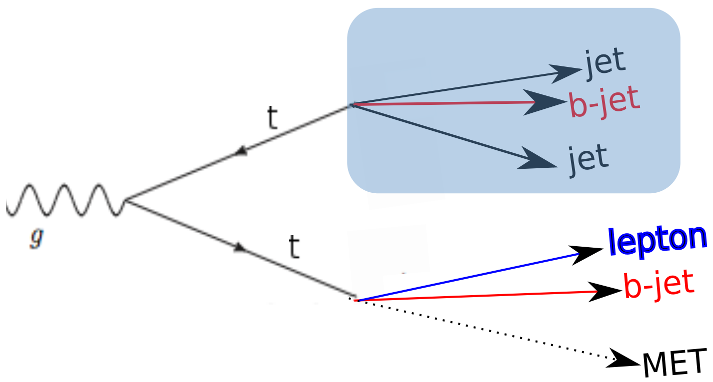


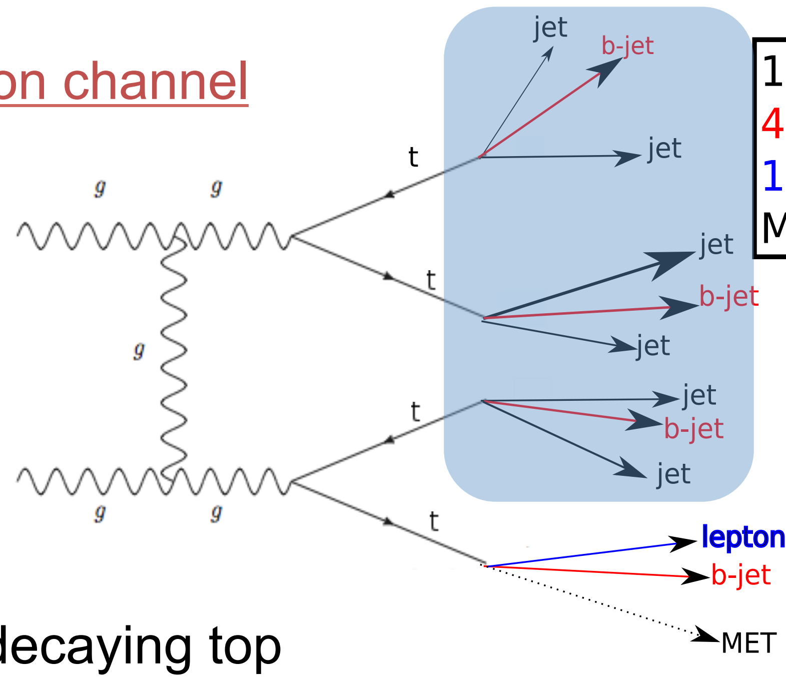
More than one hadronically decaying top quark in $t\bar{t}t\bar{t}$ events but not in $t\bar{t}$ events

Dilepton channel: One hadronically decaying top quark reconstructable in $t\bar{t}t\bar{t}$ events, not in dileptonic $t\bar{t}$

Dominant background is ttbar production

Single lepton channel





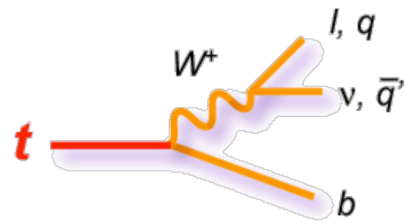
10 jets
4 b-jets
1 lepton
MET

More than one hadronically decaying top quark in $t\bar{t}t\bar{t}$ events but not in $t\bar{t}$ events

Dilepton channel: One hadronically decaying top quark reconstructable in $t\bar{t}t\bar{t}$ events, not in dileptonic $t\bar{t}$

2 x Boosted decision trees

Kinematic reconstruction of hadronic tops

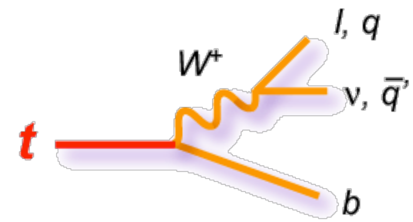


Looks at combinations of 3 jets (tri-jet)
~60% probability to select correct tri-jet
Two variables defined from this reco

Uses Angular variables
Invariant mass of trijet → top
Invariant mass of dijet → W

2 x Boosted decision trees

Kinematic reconstruction of hadronic tops



Looks at combinations of 3 jets (tri-jet)
~60% probability to select correct tri-jet
Two variables defined from this reco

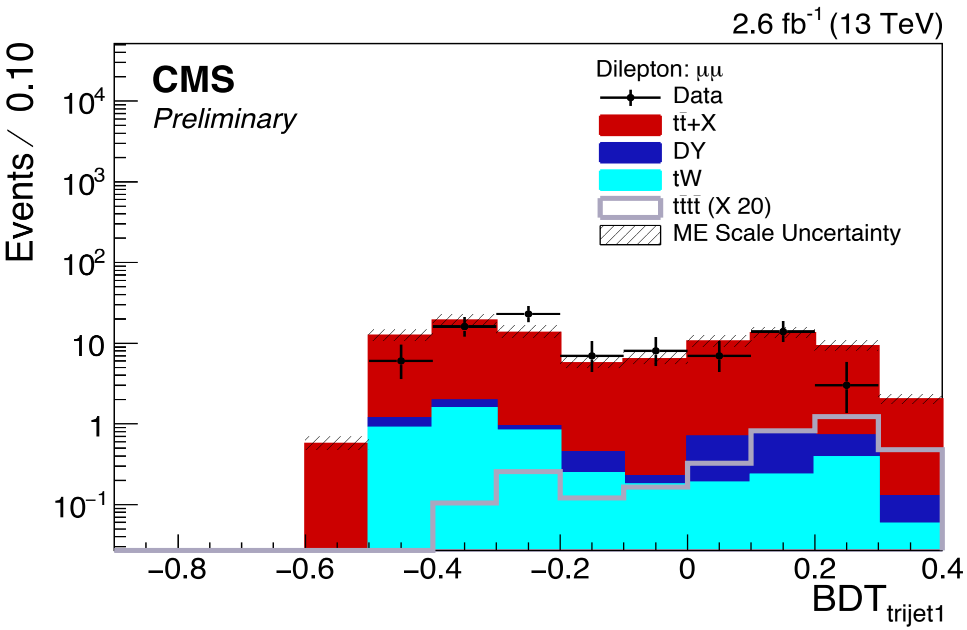
Uses Angular variables
Invariant mass of trijet → top
Invariant mass of dijet → W

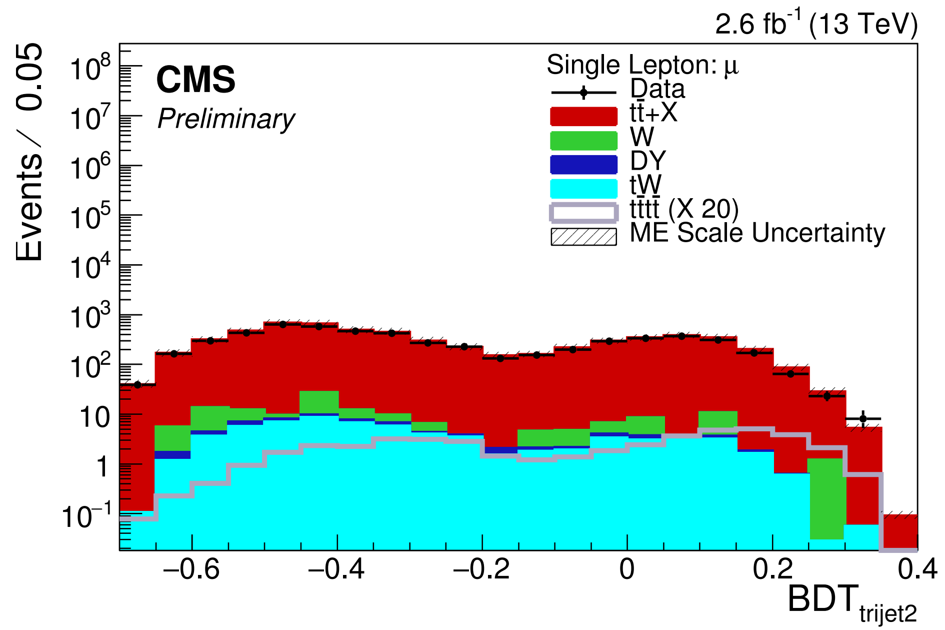
Event level BDT

Takes discriminating variables - including those defined from hadronic top reco

Tries to separate tttt signature from main ttbar background

Variables defined from hadronic top reco





$BDT_{trijet1}$: Highest ranked reconstructed top quark (dilepton)

$BDT_{trijet2}$: Highest ranked when the first highest ranked tri-jet removed (l+jets)

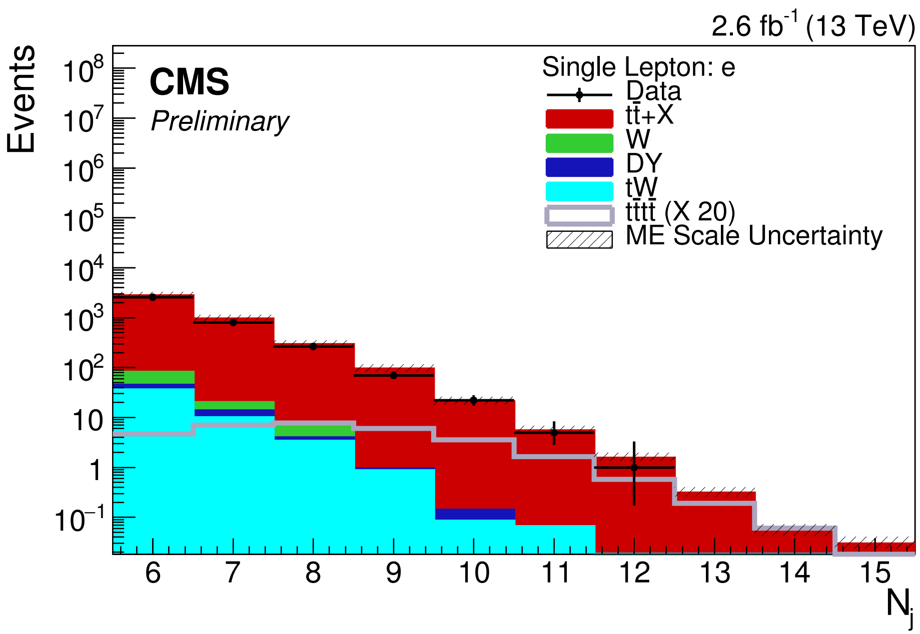
These variables are combined with other discriminating variables in the Event-level BDT.

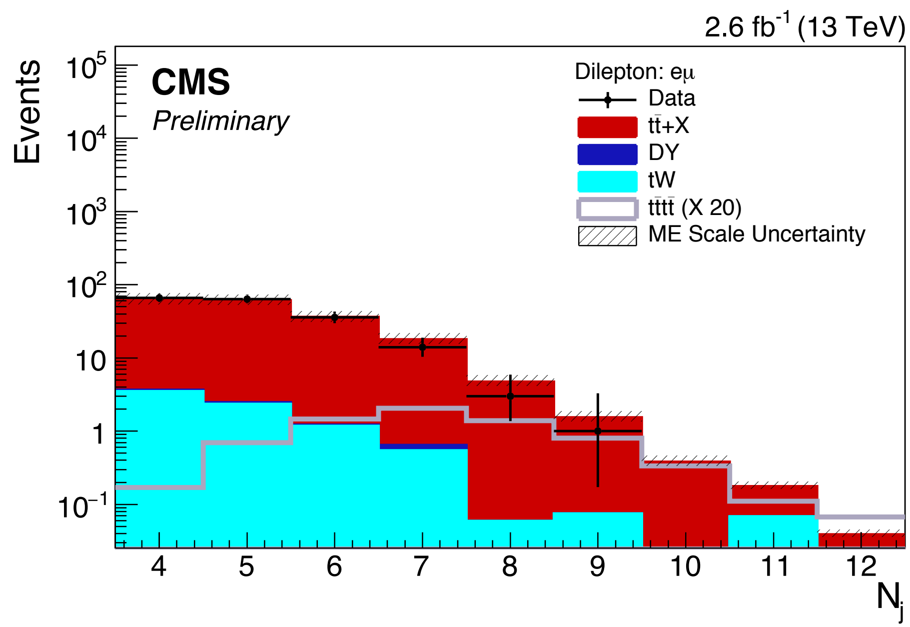
Event-level BDT training

Variables are selected to distinguish between ttbar and tttt events, including

- Number of jets, N_j
- Leading lepton p_T
- $BDT_{trijet1}$ or $BDT_{trijet2}$
- Sphericity

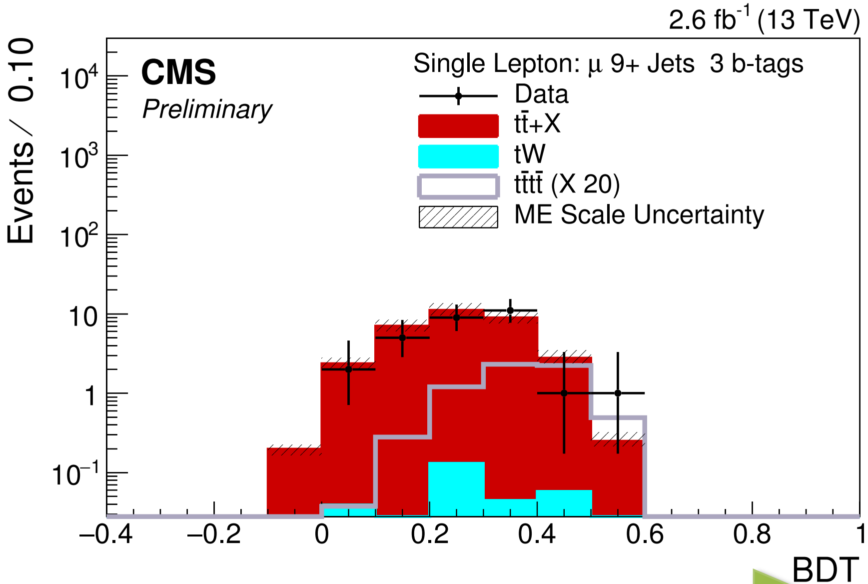
More variables in backup





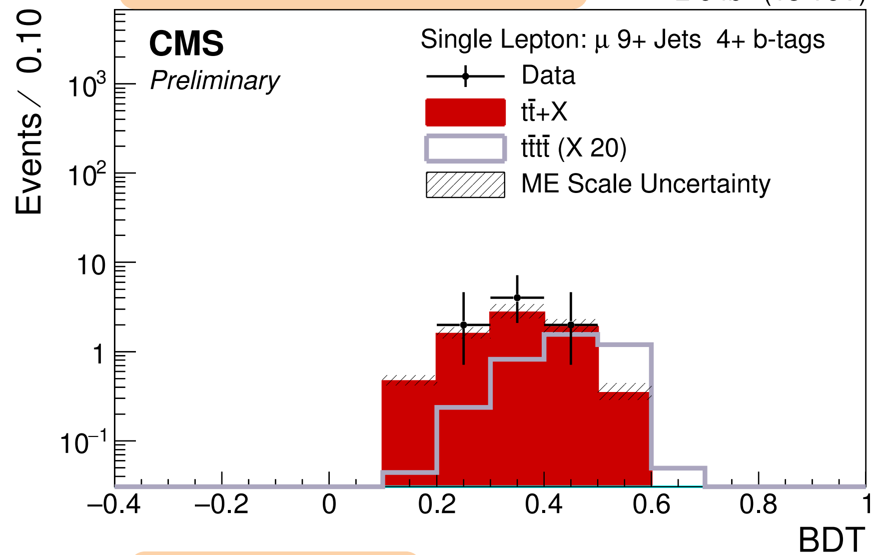
Separate trainings are performed for mu+jets and e+jets channels
Combined training for the dilepton channel due to limited MC statistics

BDT output distributions



More tttt signal-like

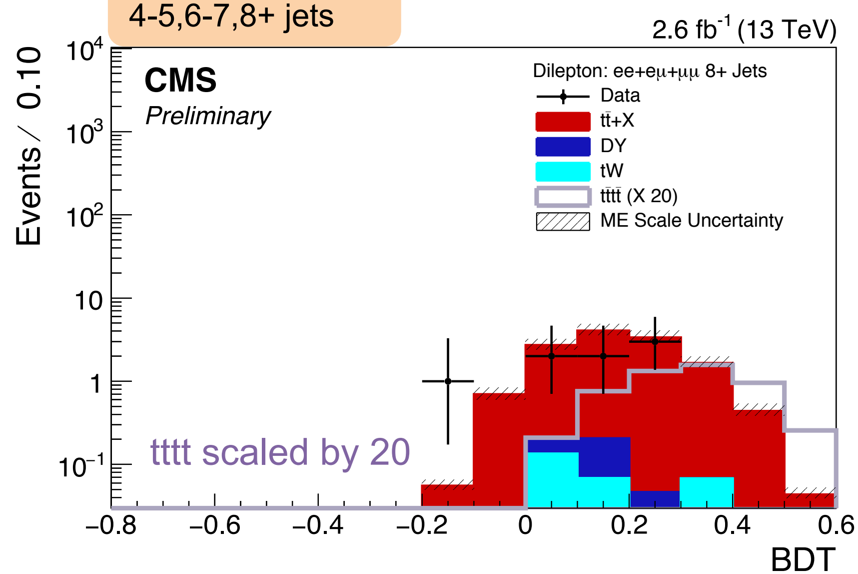
Single lepton channel:
6,7,8,9+ jets and 2,3,4+ b-tags



Splitting BDT output into categories has been shown to improve expected limit.

- Lower jet and b-tag categories constrain ttbar so effectively acting as control regions
- Higher jet and b-tag categories are more sensitive to signal

Dilepton channel:
4-5,6-7,8+ jets



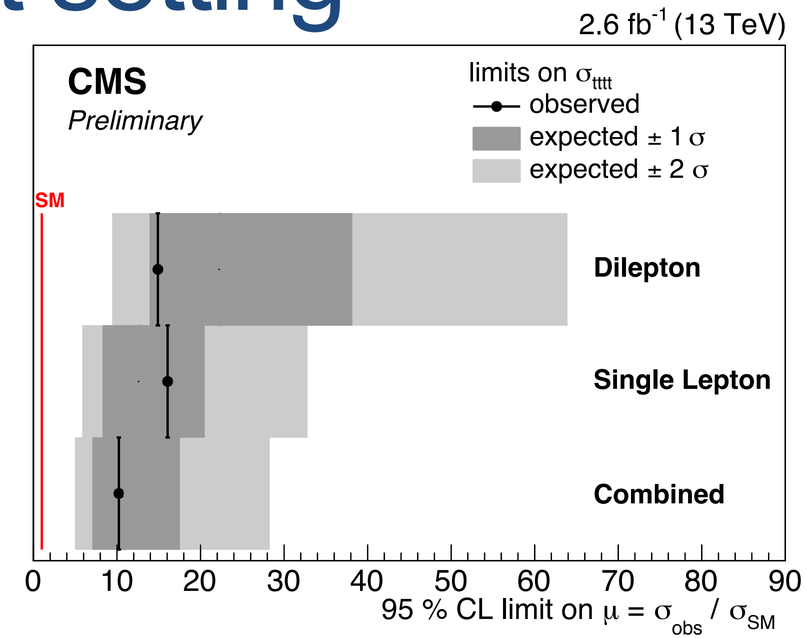
Template Fit and Limit setting

95% CL upper limit on four top production cross with 2.6 fb^{-1} of data assuming σ_{SM}^{tttt} = 9.2 fb :

10.2 obs ($10.8^{+6.7}_{-3.8}$ exp) * SM

94 fb obs (99.4^{+62}_{-35} fb exp)

ME Scale uncertainty is leading systematic, full systematics in backup



Other results

	Final State	√s TeV	Lumi fb^{-1}	Expected cross section fb	Observed cross section fb
CMS-Top-13-012	Lepton + jets	8 *	20	32 ± 17	32.5
ATLAS-CONF-2016-020	Lepton + jets	13	3.5	147 +212 -110	193
CMS-SUS-15-008	SS dilepton	13	2.1	120 ± 58	120
CMS-SUS-16-020	SS dilepton	13	12.9 **	46 +12 -15	57

* σ_{SM}^{tttt} = 1.3 fb at 8TeV

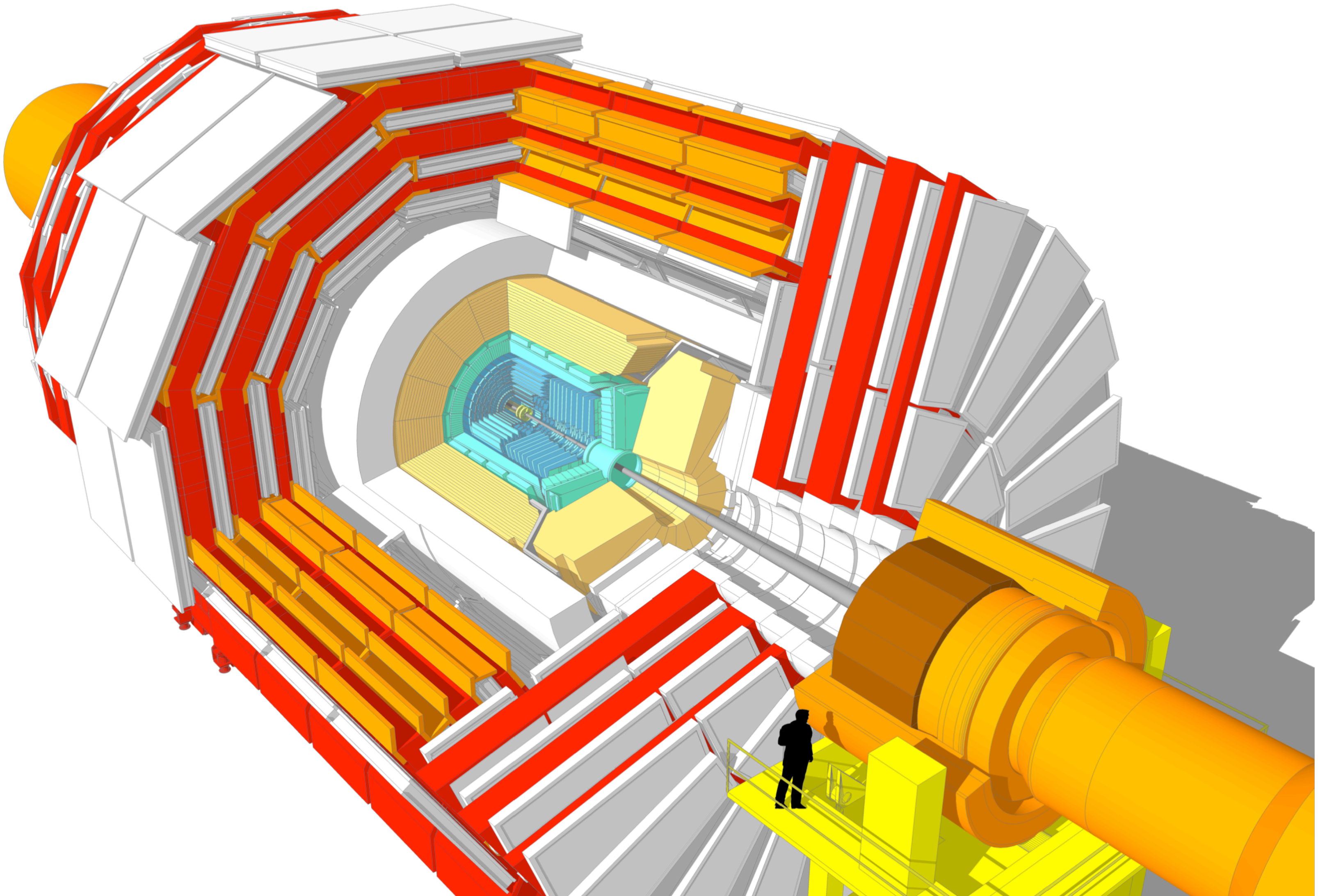
** ICHEP 2016 dataset

Conclusions

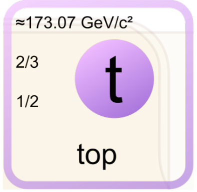
- Four top production is a rare standard model process
- Some BSM models can enhance the tttt cross section
- Identifying b-quarks + reconstructing top quarks is integral to this analysis
- Boosted Decision Tree algorithms are used to separate signal and background
- At 13 TeV we set a 95% CL limit of $\mathbf{10.2}$ x SM with an expected limit of $\mathbf{10.8^{+6.7}_{-3.8}}$
- $\mathbf{94}$ fb obs ($\mathbf{99.4^{+62}_{-35}}$ fb exp)

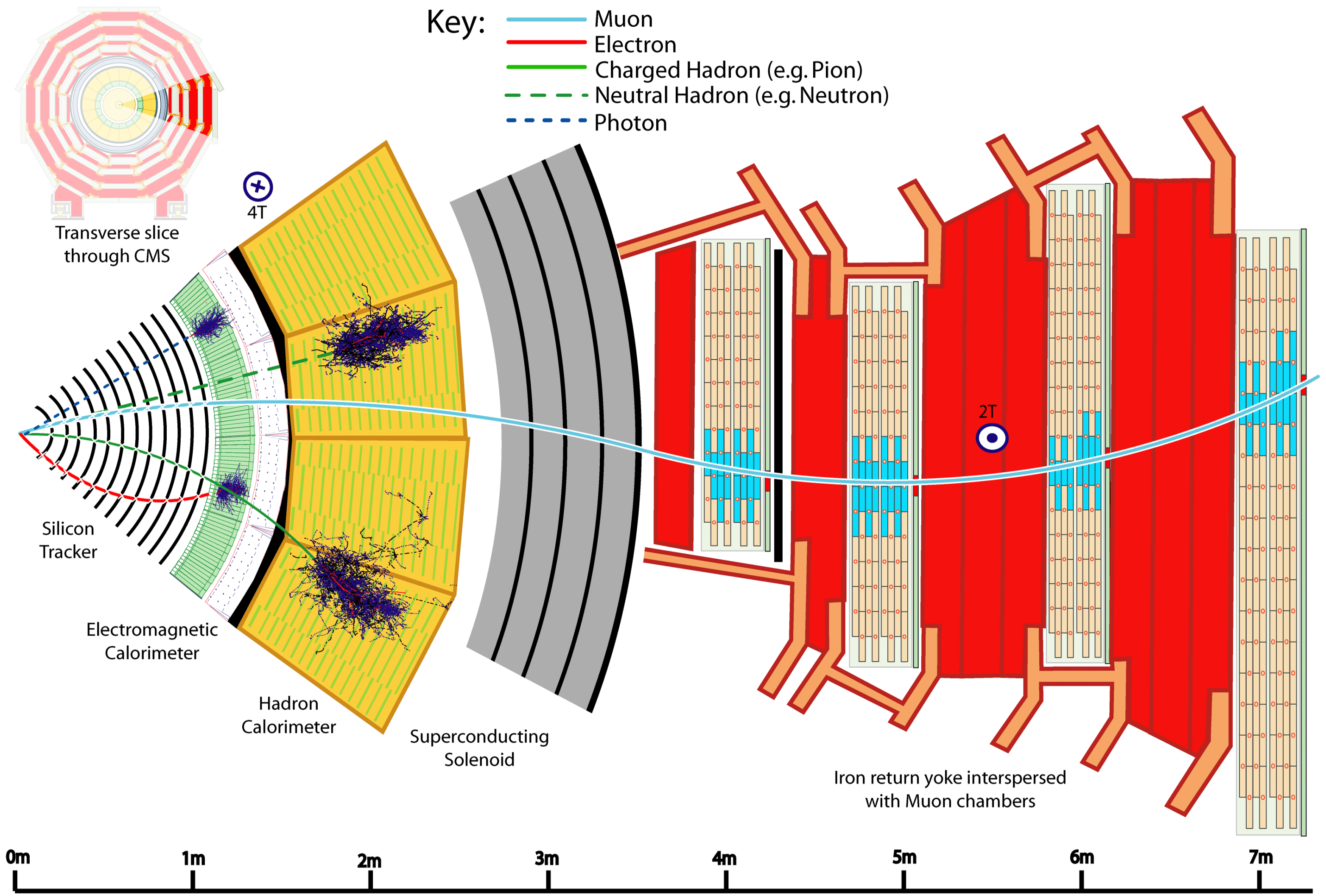
BACKUP

Detector

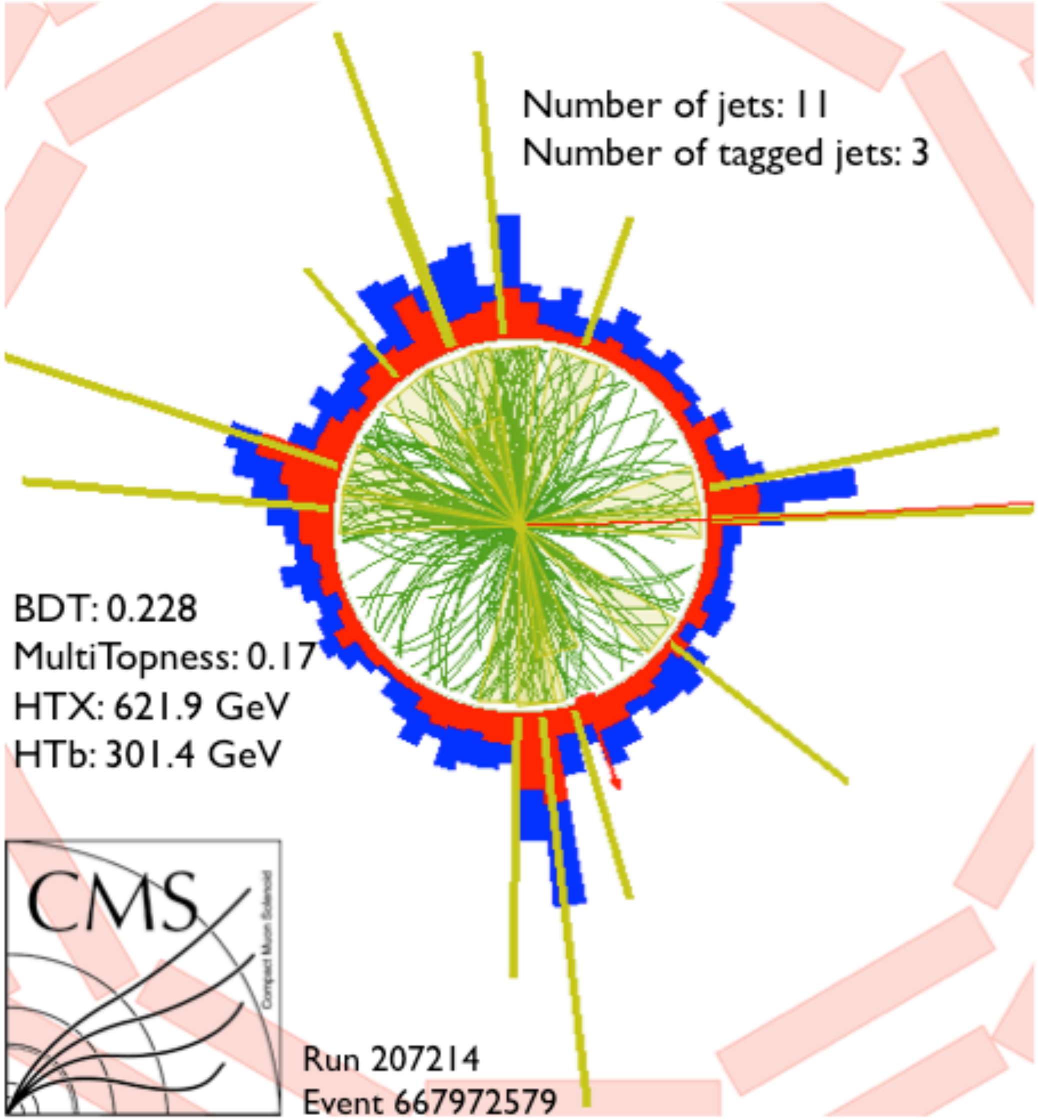


CMS detector





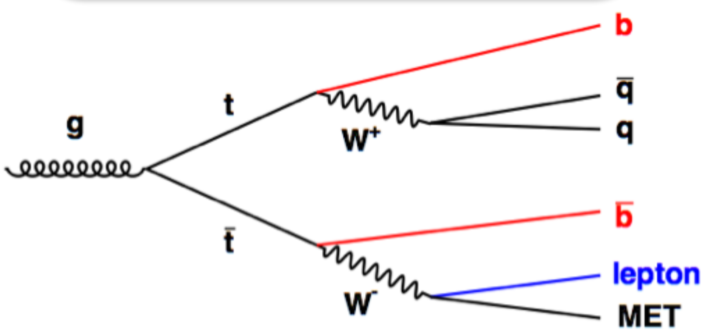
Four top quark production candidate at 8Tev



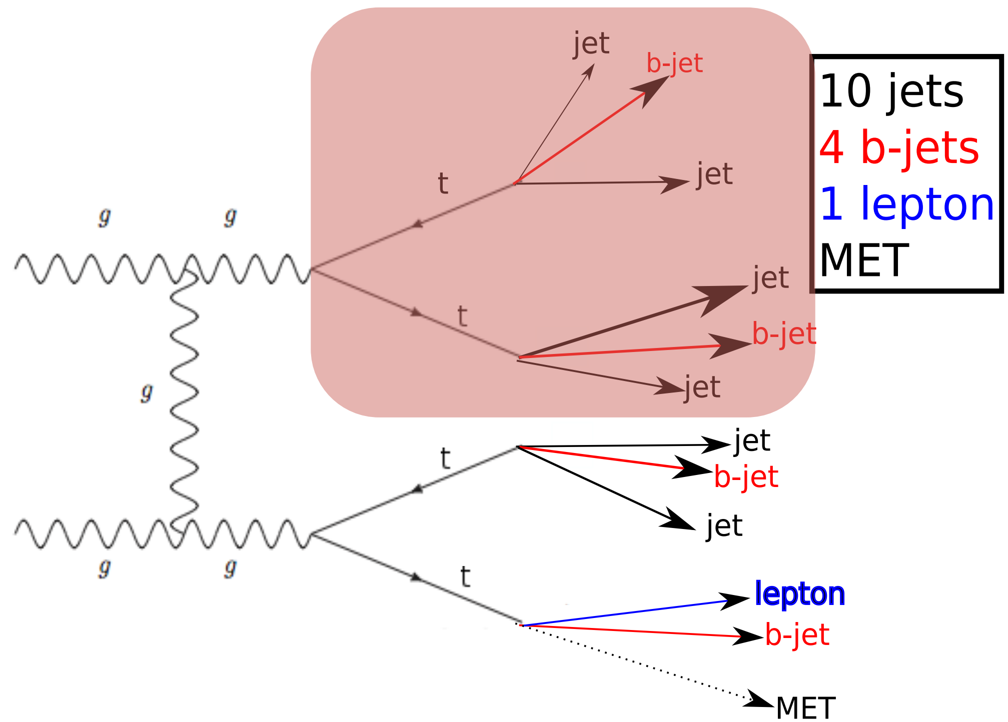
Single lepton channel

Selection

- Exactly 1 tight mu/e
- Loose lepton veto
- ≥ 6 jets
- ≥ 2 b-tagged jets

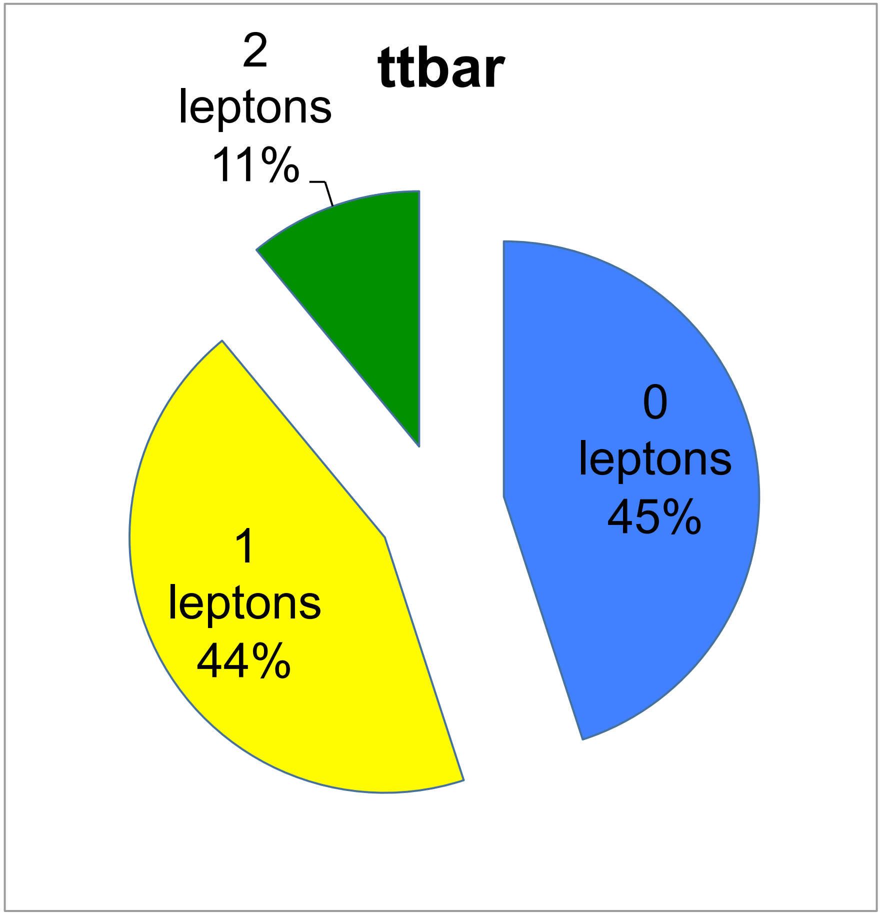


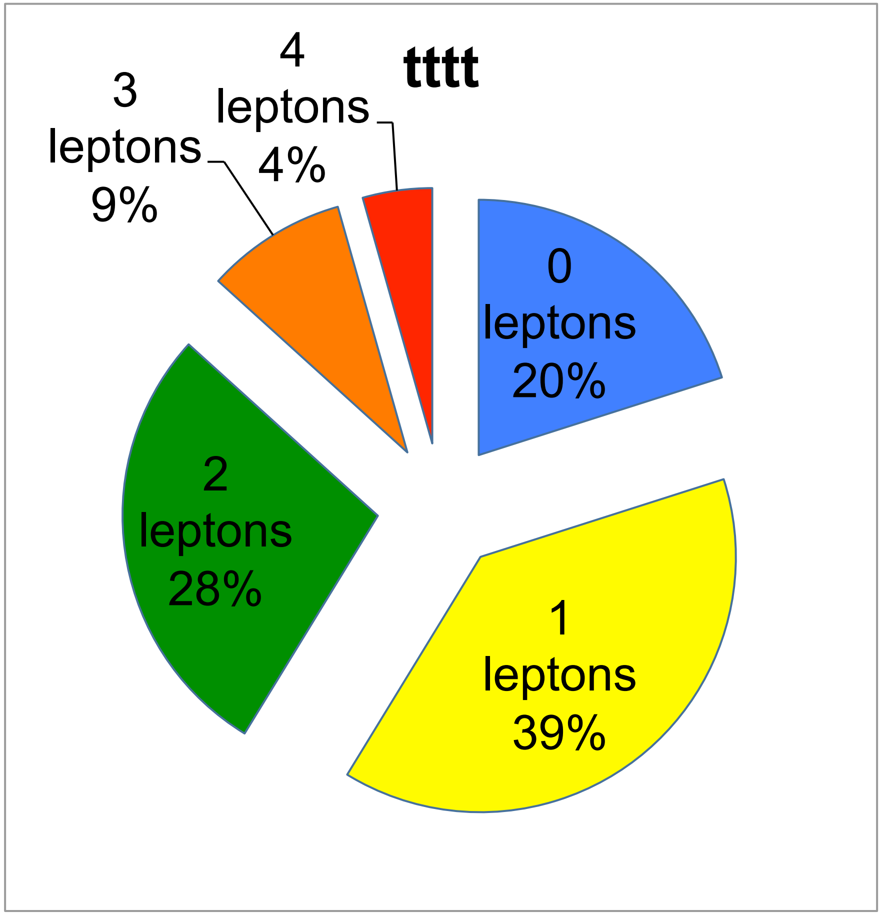
Main background, ttbar production



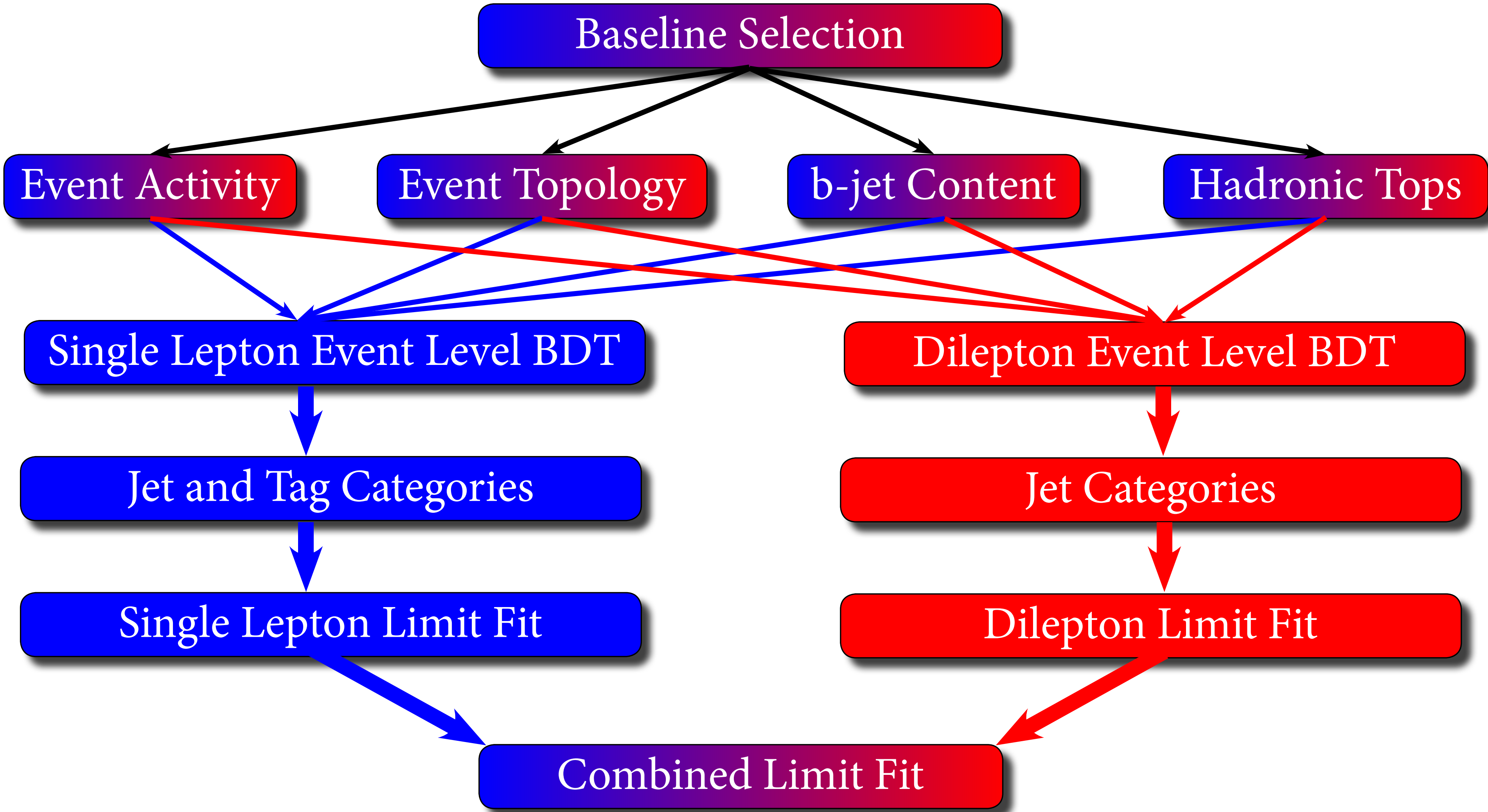
Should be possible to reconstruct more than one hadronically decaying top quark in signal events but not in background

Multi-top production





Strategy


Baseline Selection
Event Activity
Event Topology
b-jet Content
Hadronic Tops
Single Lepton Event Level BDT
Dilepton Event Level BDT
Jet and Tag Categories
Jet Categories
Single Lepton Limit Fit
Dilepton Limit Fit
Combined Limit Fit

Distinguishing variables between good and bad tri-jet combinations

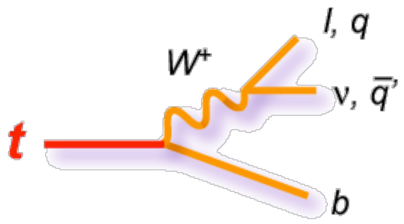
Good tri-jet defined as all 3 jets originating from a top quark and bad tri-jet being any other combination

Tri-jet invariant mass	good tri-jets should have an invariant mass close to top mass
Di-jet invariant mass	formed from two jets with smallest ΔR separation and should have mass close to the W mass
p_TRat	ratio of vectorial p_T to the scalar sum of the p_T of the jets in the tri-jet
$\Delta\phi_{T\text{-}W}$	$\Delta\phi$ between tri-jet and di-jet system used in invariant mass variable
$\Delta\phi_{T\text{-}b}$	$\Delta\phi$ between tri-jet and jet not included in di-jet system
CSV_b	CSV b-tagging discriminator for jet not used in di-jet

Template Fit and Limit setting

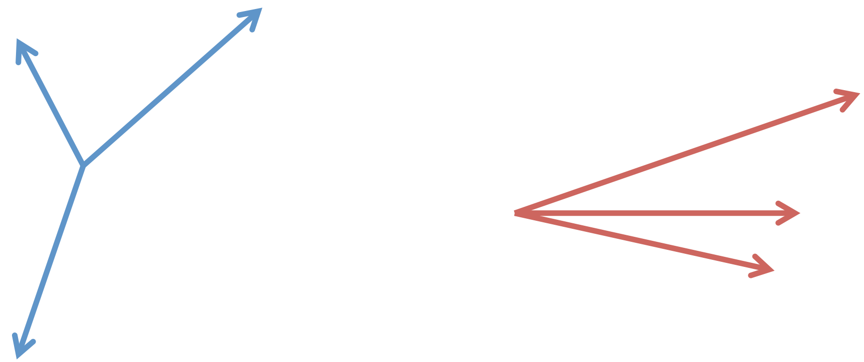
- 95% CL upper limit on four top production cross section using the Asymptotic CL_S method.

Channel	Expected Limit (x $\sigma^{SM}_{t\bar{t}t\bar{t}}$)	Observed Limit (x $\sigma^{SM}_{t\bar{t}t\bar{t}}$)
Single Lepton	$12.7^{+7.8}_{-4.4}$	16.1
Dilepton	$22.3^{+16.2}_{-8.4}$	14.9
Combined	$10.8^{+6.7}_{-3.8}$	10.2

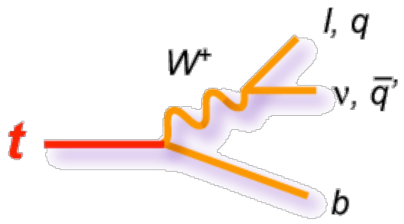


Reconstruction of top quarks

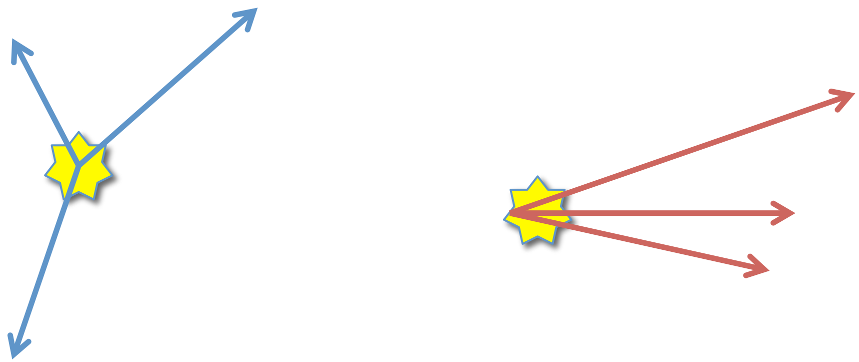
Look at every combination of jets in the event



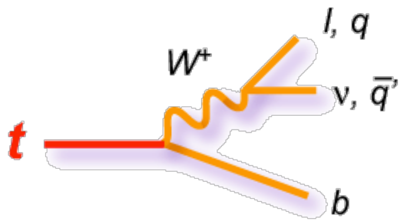
Reconstruction of top quarks



Look at every combination of jets in the event



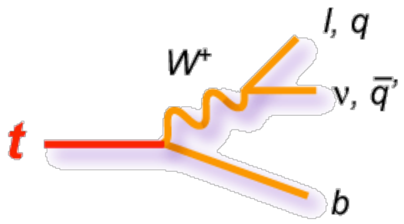
	Bad Combination	Good Combination
Invariant Mass of 3 jets	Not close to top	Close to top mass



Reconstruction of top quarks

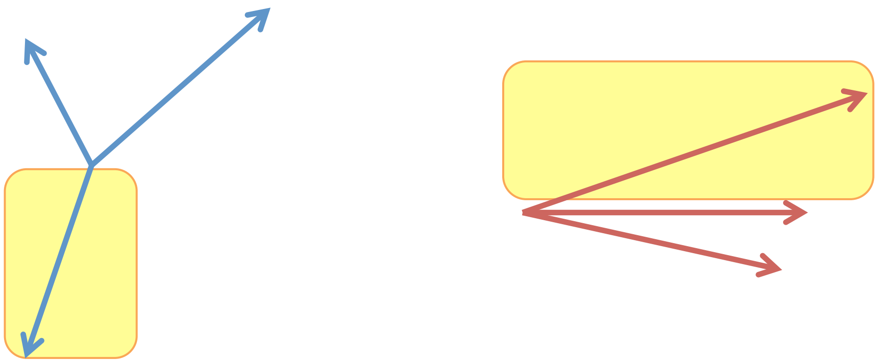
Look at every combination of jets in the event

	Bad Combination	Good Combination
Invariant Mass of 3 jets	Not close to top	Close to top mass
Inv. Mass of 2 closest jets	Not close to W	Close to W mass

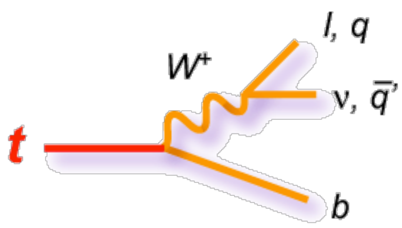


Reconstruction of top quarks

Look at every combination of jets in the event

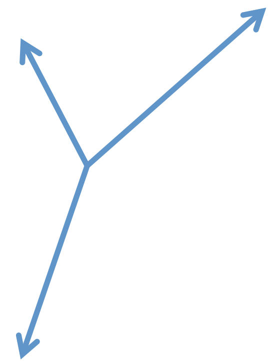


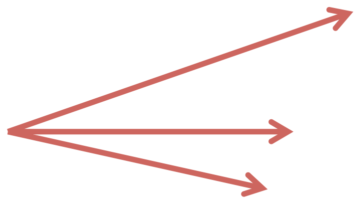
	Bad Combination	Good Combination
Invariant Mass of 3 jets	Not close to top	Close to top mass
Inv. Mass of 2 closest jets	Not close to W	Close to W mass
3rd jet btagged?	No	Yes



Reconstruction of top quarks

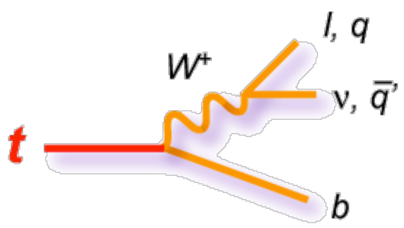
Look at every combination of jets in the event



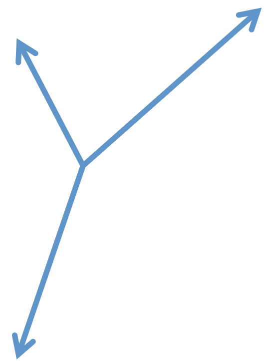


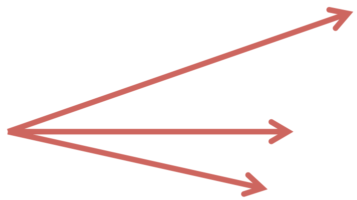
	Bad Combination	Good Combination
Invariant Mass of 3 jets	Not close to top	Close to top mass
Inv. Mass of 2 closest jets	Not close to W	Close to W mass
3rd jet btagged?	No	Yes
$\Sigma\vec{p}_T/\Sigma p_T$	0.1	0.9

Reconstruction of top quarks



Look at every combination of jets in the event





	Bad Combination	Good Combination
Invariant Mass of 3 jets	Not close to top	Close to top mass
Inv. Mass of 2 closest jets	Not close to W	Close to W mass
3rd jet btagged?	No	Yes
$\Sigma \vec{p}_T / \Sigma p_T$	0.1	0.9

Boosted Decision Tree

This method was used at 8 TeV in the search for four top quark production and is replicated at 13 TeV

Single lepton selection

- Trigger based on the presence of at least one isolated muon (electron) with $p_T \geq 18$ (23) GeV
- Exactly one μ/e
- Veto on additional loose leptons
- >= 6 jets
- >= 2 b-tags (CSVM)

Single lepton BDT variables

- N_j
- $BDT_{trijet2}$
- $H_T{}^b$
- $H_T{}^{Rat}$
- p^l1
- $N_j{}^W$
- $N_{tags}{}^L$
- $N_{tags}{}^M$
- $M_{RE}{}^H$
- $p_{T trijet1}$
- $H_T{}^X$
- T

Dilepton selection

- Triggers based on the appropriate lepton flavour for each of the final states. The e or μ ranked leading in pT, was required to have $pT \geq 17$ GeV and the second lepton was required to pass $pT \geq 12$ GeV for an electron $pT \geq 8$ GeV for a muon.
- Exactly two OS $\mu\mu$, ee, μe
- >= 6 jets
- >= 2 b-tags (CSVM)
- Transverse hadronic energy, HT>500

Dilepton BDT variables

- N_j
- $BDT_{trijet1}$
- H_T^b
- H_T^{Rat}
- p^l1
- N_j^W
- N_{tags}^L
- N_{tags}^M
- ΔR_{bb}
- η^{l1}
- H_T^{2M}
- ΔR_{ll}
- C
- p_T^{Jet3}
- p_T^{Jet4}
- S

Description of variables used – Event Activity

- Jet multiplicity (N_j), the number of selected jets present in the event. In Fig. 4 the N_j distributions in data and simulation are shown for the single lepton and dilepton channels.
- Weighted jet multiplicity (N_j^W). This observable takes into account the jet multiplicity while weighting with the p_T spectrum of the jets, and is defined as

$$N_j^W = \frac{\int_{30}^{125} N_j\left(E_{threshold}\right) * E_{threshold}\, dE_{threshold}}{\int_{30}^{125} E_{threshold} dE_{threshold}} = \frac{\Sigma_0^{N_j} N_j\left(E_{threshold}\right) \left[E^2_{threshold}/2\right]^{E_{lower}}_{E_{upper}}}{\left[E^2_{threshold}/2\right]^{125}_{30}}, \tag{1}$$

where $E_{threshold}$ is the p_T threshold above which a jet is counted. $N_j\left(E_{threshold}\right)$ are the number of jets above the energy threshold, E_{lower} is the set of $\left[30, E_j\right]$ and E_{upper} is the set of $\left[E_j, 125\right]$, where E_j are the energies of each jet in ascending order. This equation can be solved analytically as shown in the right part of Eq. 1. The numerator can be split into a sum of integrals in which the number of jets are constant in each of the integral ranges. It has higher values for events with many high-p_T jets than for events with the same number of jets where only few are high p_T and the rest are close to the selection threshold.
- Transverse b jet energy (H_T^b). H_T^b is the scalar sum of the p_T of all b tagged jets in the event, identified using the medium working point.
- Transverse hadronic energy ratio (H_T^{Rat}), this is the ratio of the H_T of the four (two) leading jets in the event in the single lepton(dilepton) channel to the H_T of the other jets in the event.
- Transverse hadronic energy in excess of the two b jets (H_T^{2M}), this is the H_T in the event minus the E_T of the two selected b jets.

Description of variables used – Event Activity

- Third and fourth jet transverse momentum (p_T^{Jet3} and p_T^{Jet4}) are the p_T values of the 3rd and 4th largest p_T jets in the event.
- Reduced event, hadronic invariant mass (M_{RE}^{H}), defined as the invariant mass of the system comprising all the jets in the reduced event, where the reduced event is constructed by subtracting the jets contained in the highest ranking trijet as defined by the kinematic reconstruction procedure described in 5.1. In $t\bar{t}$ events, the reduced event will typically only contain jets arising from initial and final state radiation. Conversely, a reduced $t\bar{t}t\bar{t}$ event could contain up to 2 hadronic top quarks and as a result numerous energetic jets.
- Reduced event H_T, H_T^X, defined as the H_T of all jets in the event excluding those contained in the highest ranking trijet.

Description of variables used – Event topology

- Event sphericity (S), the sphericity of a group of objects [37] is defined as $S^{\alpha\beta} = \frac{\sum_i p_i^\alpha p_i^\beta}{\sum_i |\vec{p}_i|^2}$, and calculated by creating the tensor $S^{\alpha\beta}$ from the three principal components of their momenta where α and β refer to the principal momenta of the i-th particle. Solving the eigenvalue problem of the tensor one constructs the sphericity, $S = \frac{3}{2}(\lambda_2 + \lambda_3)$, where λ_2 and λ_3 are there two smallest eigenvalues of $S^{\alpha\beta}$. The sphericity of all reconstructed objects in an event for signal $t\bar{t}t\bar{t}$ events should differentiate from $t\bar{t}$ events as for the same amount of energy present in the event, the $t\bar{t}$ will be less spherically distributed as it must be boosted off of extra energy in the event from sources such as ISR. The distribution of the event sphericity is shown in Fig. 5.
- Hadronic centrality ($C = H_T/H$), the centrality of a group of objects is defined as the scalar sum of the transverse energies of those objects divided by the scalar sum of their energies. C is defined as the centrality of all jets in the event.
- T is a thrust observable, defined as

$$T = \frac{\sum_i p^i_{jet} |cos\Delta\phi \left(\hat{p}_{trijet1}, p^i_{jet}\right)|}{\sum_i p^i_{jet}}, \tag{2}$$

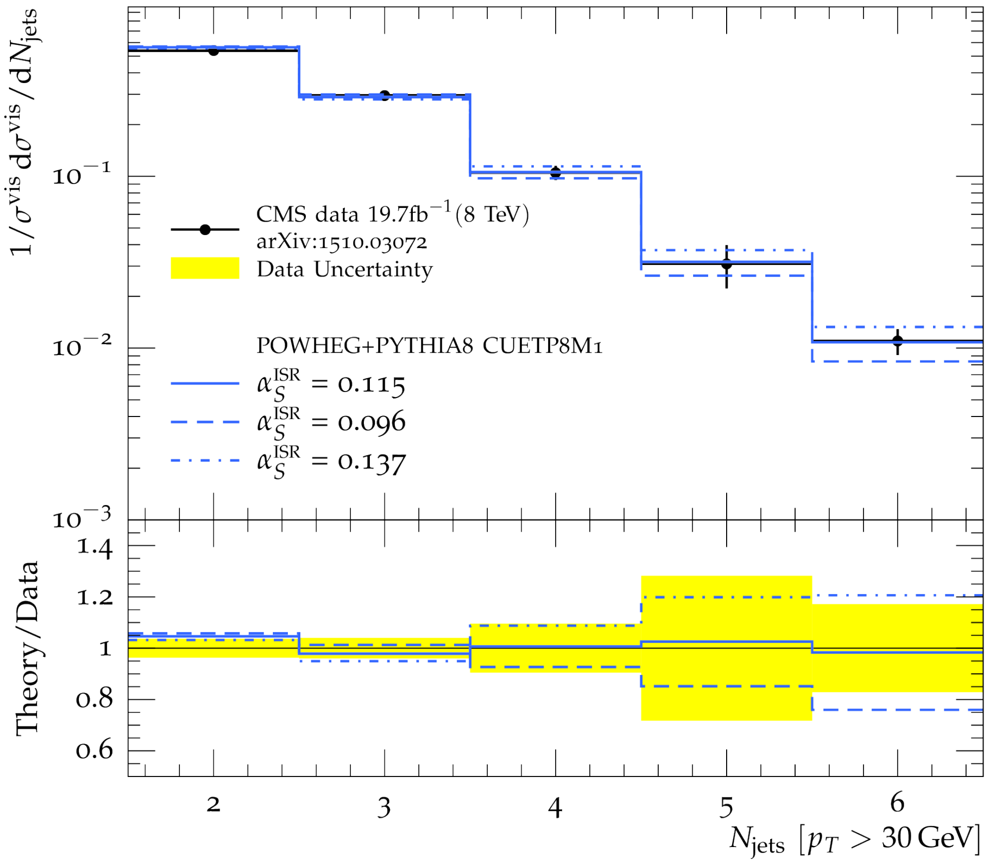
and is defined as the cosine of the angle in the transverse plane of the normalised p_T vector of the first reconstructed hadronic trijet, $\hat{p}_{trijet1}$, taken with the remaining jets in the event which are expected to be more aligned along the direction (or anti-direction) of the reconstructed hadronic top quark in $t\bar{t}$ events. The distributions for T are shown in Fig. 6.

Description of variables used – b-jet content

Signal $t\bar{t}t\bar{t}$ events are expected to have four b quarks in the final state compared to two for the main $t\bar{t}$ background. Therefore the presence of more than two b tagged jets is a source of discriminating power. The number of b quark jets identified with the loose, medium and tight b tagger can be used to separate $t\bar{t}t\bar{t}$ signal and $t\bar{t}$.

- The number of loose working point b-tags (N^{L}_{tags})
- The number of medium working point b-tags (N^{M}_{tags})
- The number of tight working point b-tags (N^{T}_{tags})

Tuning high jet multiplicity



Scale factor derived between optimal tune of alphaS = 0.115 and the alphaS value used in simulation, alphaS = 0.137

http://cms-results.web.cern.ch/cms-results/public-results/publications/TOP-12-041/index.html

Systematics uncertainties

Normalisation

- Luminosity
- MC cross sections
- Lepton SF
- alphaS tune

Shape

- ME scale on ttbar and tttt
- Hadronisation scale on ttbar
- JES
- JER
- tt generator
- b-tag SF for heavy flavour
- b-tag SF for light flavour
- Heavy flavour reweighting
- PU reweighting

→ using ttbb/ttjj fraction from https://cds.cern.ch/record/2202803