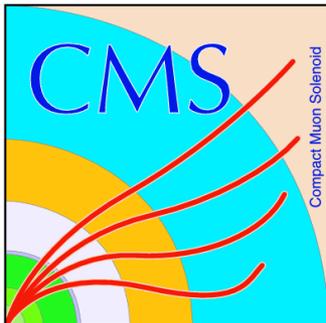


Boosted top quarks in physics analyses

Kevin Nash

On behalf of the CMS Collaboration

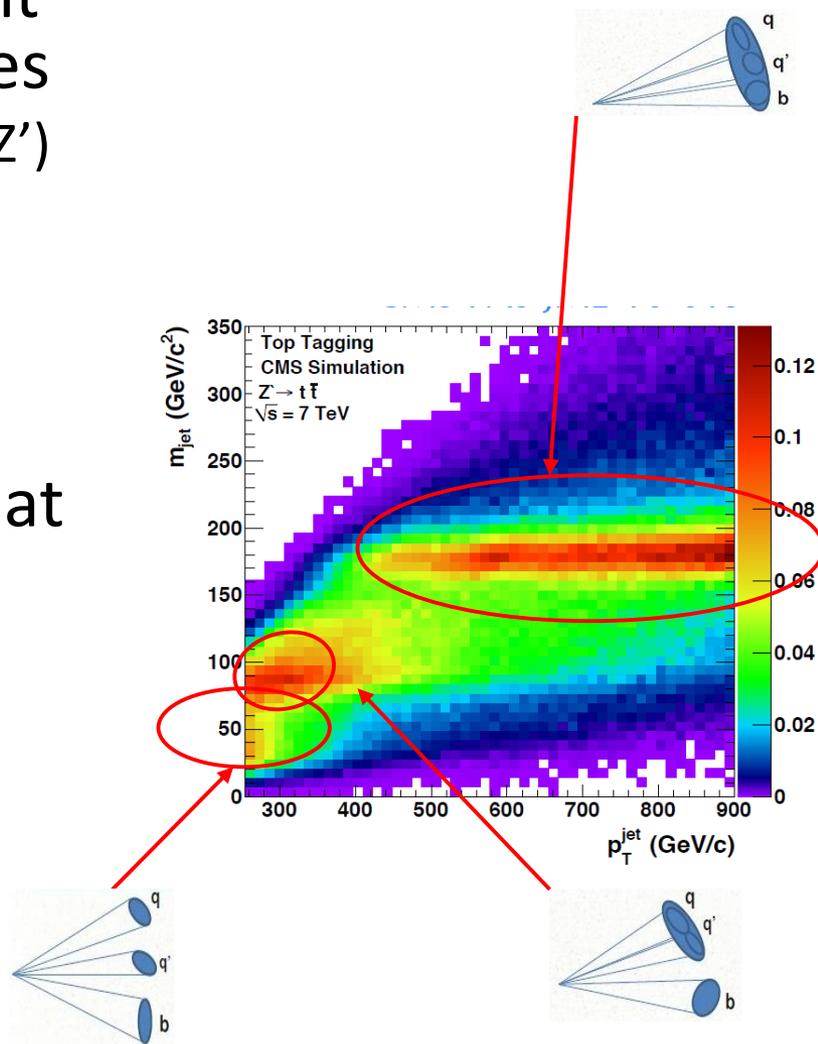


BOOST 2014



Introduction

- Top quarks play an important role for BSM physics searches
 - New massive gauge bosons (W' , Z')
 - Heavy quark partners (t' , b')
 - Kaluza-Klein excitations
 - SUSY
 - etc...
- Top quark daughters merge at high boost
 - Boosted top quark identification allows for sensitivity in very high resonant mass regions.
 - Hadronic top decay resolved as single jet $p_T \geq 400$ GeV



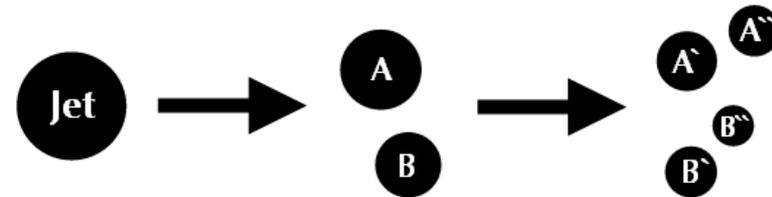
Outline

- Top tagging algorithms
 - CMS top tagger
 - HEP top tagger
- Physics analyses
 - Search for BSM tH production (B2G-14-002)
 - Search for BSM tb production (B2G-12-009)
 - Search for top squark pair production (SUS-13-015)
 - Search for $T_{5/3}$ pair production (B2G-12-012)
 - Search for T' pair production (B2G-12-015)
 - Search for BSM $t\bar{t}$ production (B2G-13-001)

Top-tagging algorithms

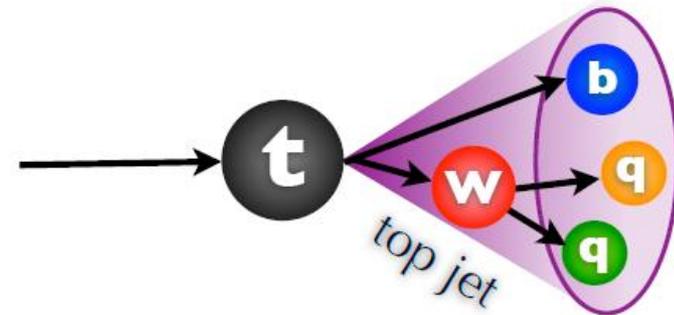
- CMS top-tagger

- Cluster jets using Cambridge Aachen ($R = 0.8$) sequential recombination algorithm
- Decluster twice to find up to four subjets
- Jet mass
 - $140 \text{ GeV} < M_{\text{Jet}} < 250 \text{ GeV}$
- Minimum pairwise mass
 - $M_{\text{min}} > 50 \text{ GeV}$
- Number of subjets
 - $N_{\text{sj}} > 2$



- N-subjettiness

- Use jet constituents to define how consistent the jet energy is with having N subjets



- Subjet b-tagging

- Search for a b-tag in the collection of subjets

Top-tagging algorithms

- HEP top-tagger

- Use large jets

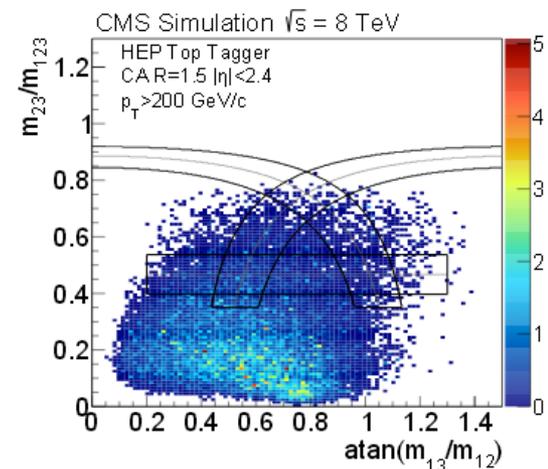
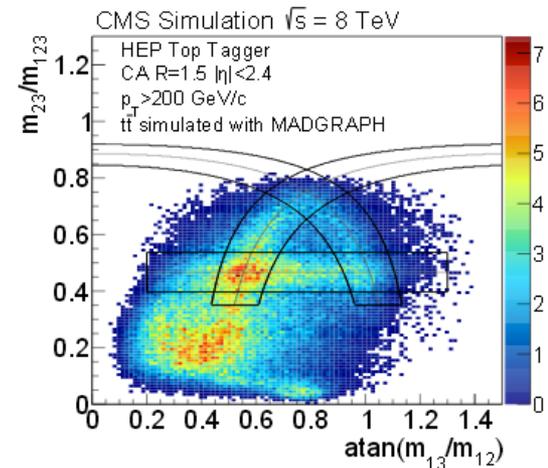
- Cambridge Aachen (R=1.5)

- Invariant subjet mass

- $140 < M_{123} < 250$
 - top mass selection

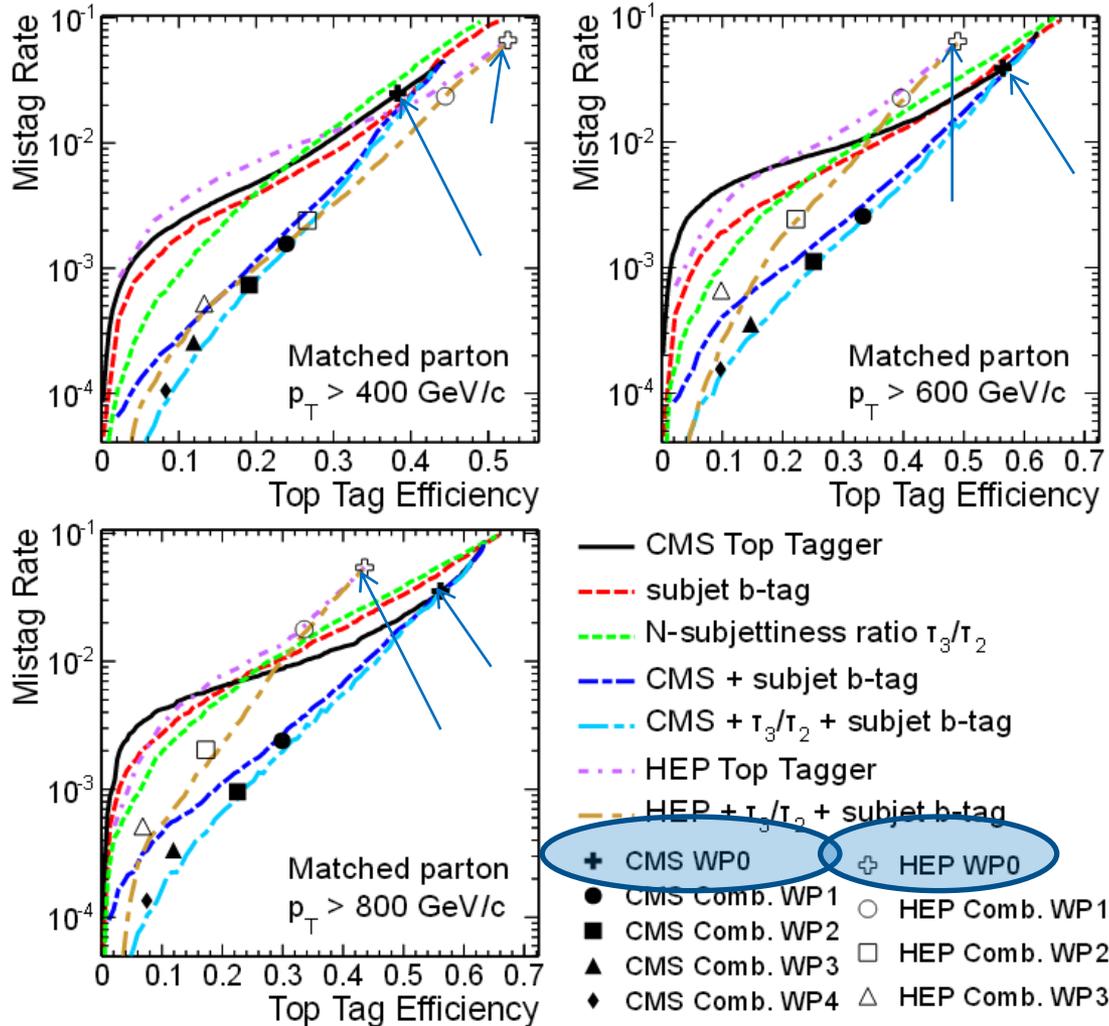
- Pairwise subjet mass ratios

- M_{12}, M_{23}, M_{13}
 - W mass selection



Top-tagging algorithms

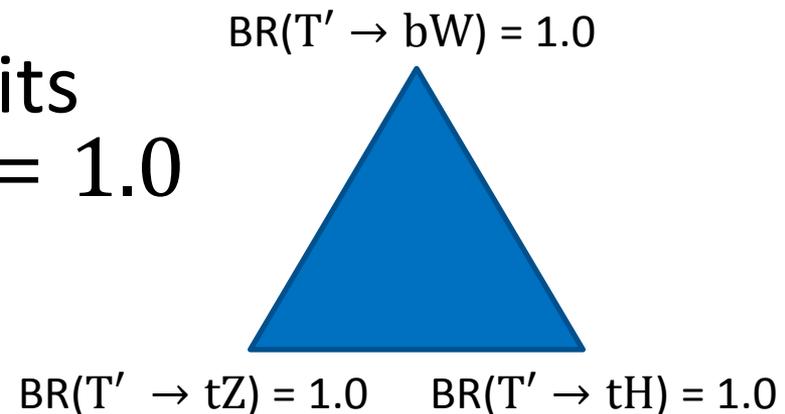
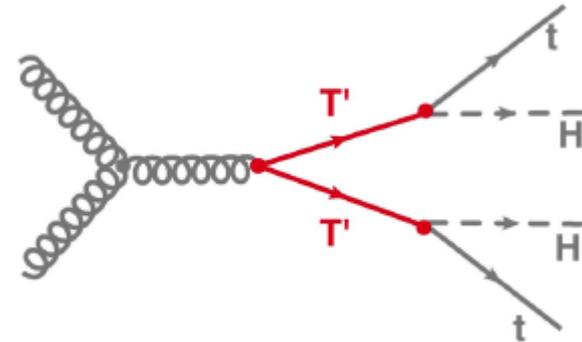
CMS Simulation, $\sqrt{s} = 8$ TeV



Search for tH resonances

Search for tH resonances

- Pair of heavy resonances decaying to a Higgs boson and top quark
 - $T'T' \rightarrow (tH) (tH)$
 - $t \rightarrow (Wb) \rightarrow$ All Hadronic
 - $H \rightarrow (bb)$
- Extract cross-section limits assuming $BR(T' \rightarrow tH) = 1.0$
- Scan T' BR space
 - $T' \rightarrow Wb$
 - $T' \rightarrow tZ$
 - $T' \rightarrow tH$



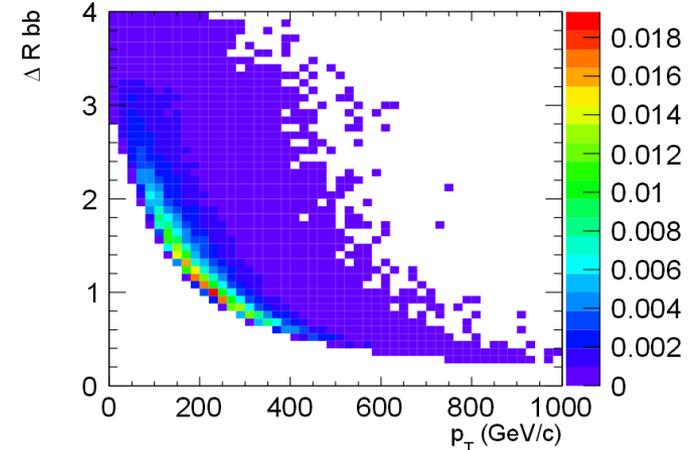
Search for tH resonances



- Merged tops
 - HEP top-tagger
 - At least one subjet b-tag
- Merged Higgs'
 - $H \rightarrow bb$ within single CA15 resolved jet
 - Two subjets, both subjet b-tagged
 - Di-subjet invariant mass larger than 60GeV

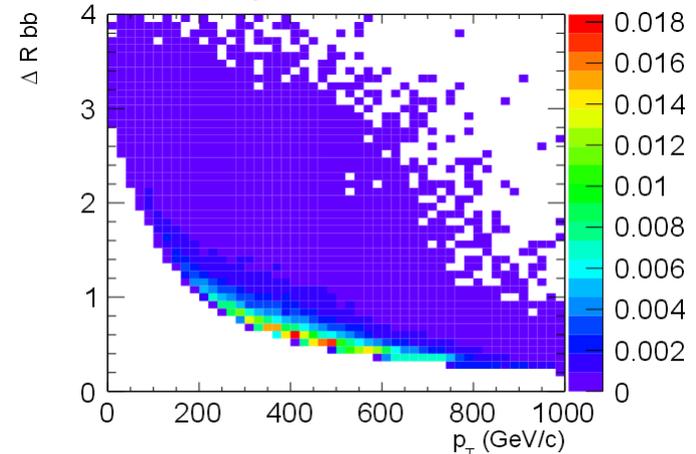
$M_{T'} = 500 \text{ GeV}$

CMS Simulation Preliminary



$M_{T'} = 1000 \text{ GeV}$

CMS Simulation Preliminary



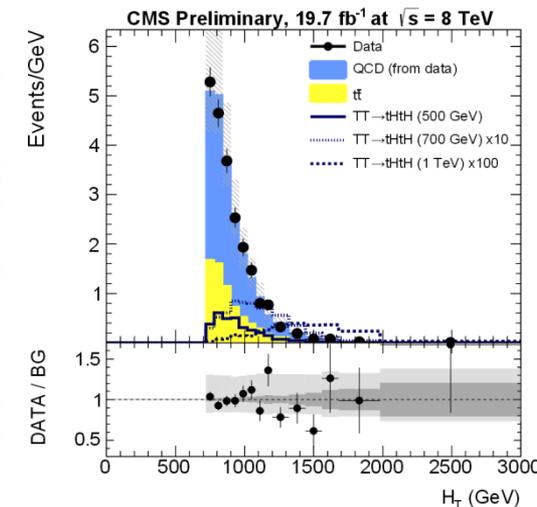
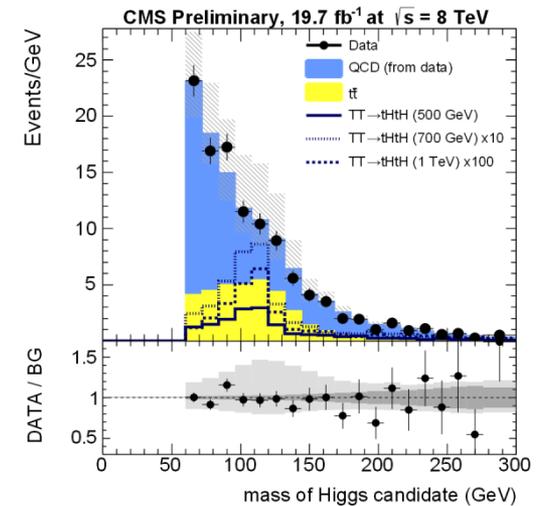
Search for tH resonances

- $t\bar{t}$ background estimate from Monte Carlo
- QCD background estimate from data
 - ABCD method
 - Higgs-tag and top-tag can define four regions
 - Three control regions
- Normalization: $N_D = \frac{N_C \times N_B}{N_A}$
- Shapes: Use B region
 - H_T and m_H

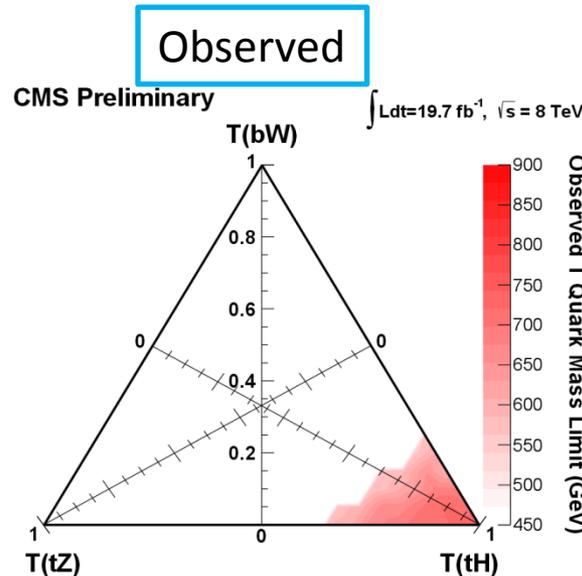
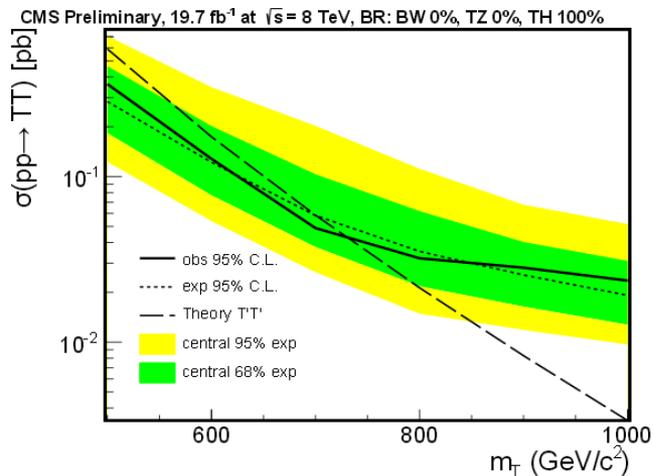
	Inv. H-tag	H-tag
Inv. top-tag	A	B
top-tag	C	D(Signal)

Search for tH resonances

- Exclude $M_{T'} < 747\text{GeV}$
 - Given $\text{BR}(T' \rightarrow tH) = 1.0$
- Scan BR space



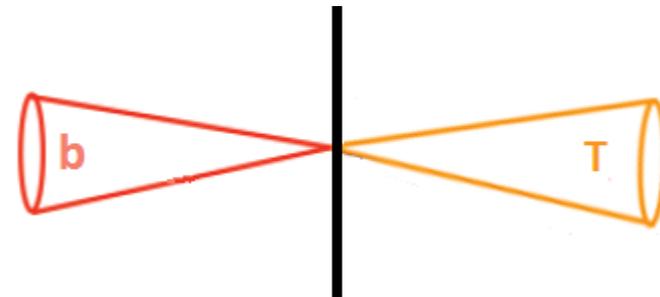
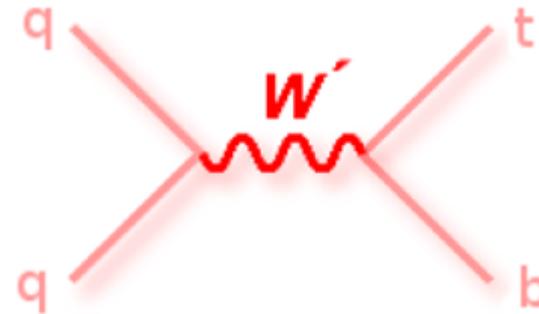
$\text{BR}(T' \rightarrow tH) = 1.0$



Search for tb resonances

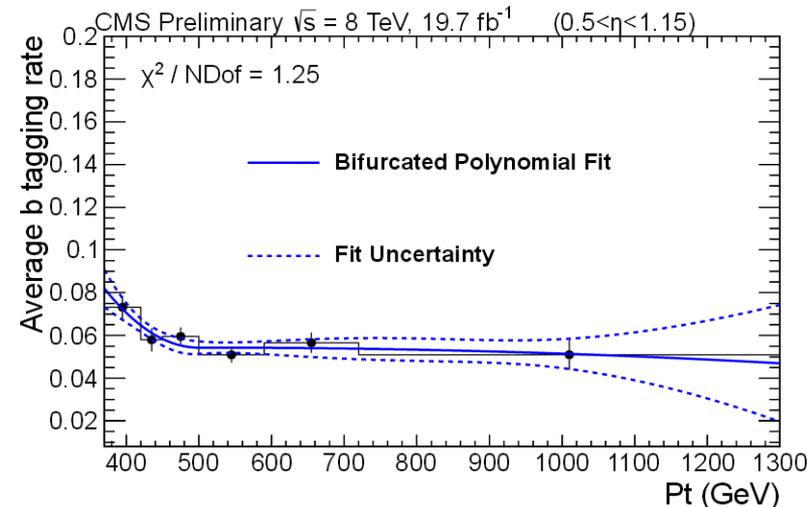
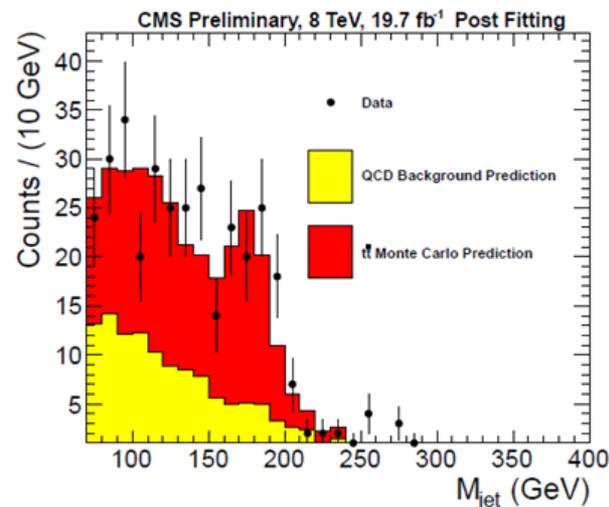
Search for tb resonances

- Search for heavy tb resonances
 - $tb \rightarrow (Wb)(b) \rightarrow$ All Hadronic
- Generic W' couplings
 - Right-handed, left-handed, mixed
 - Left handed and mixed include SM W interference
- Cross-section limits placed on right-handed W'
- Place limits on left- and right-handed coupling strengths
- CMS top-tagger
 - N-subjettiness
 - Subjet b-tagging
- b-tagging on opposite jet
- Dijet topology



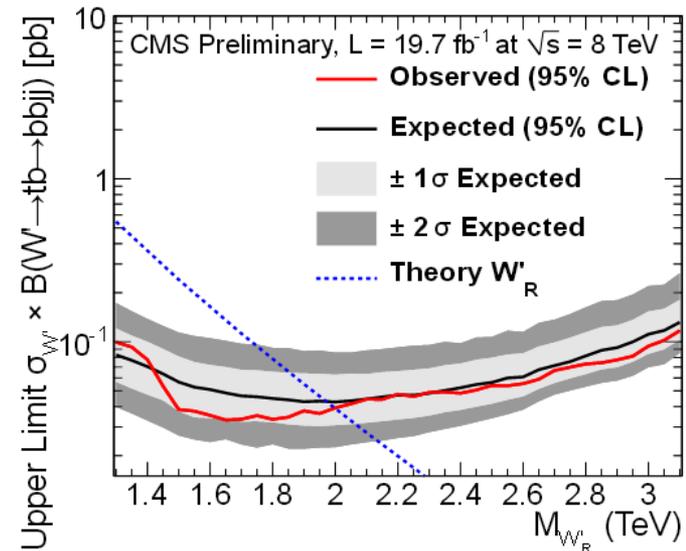
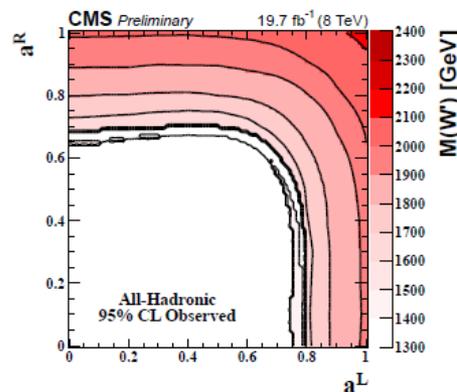
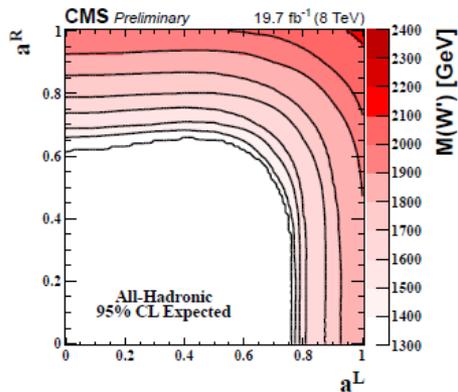
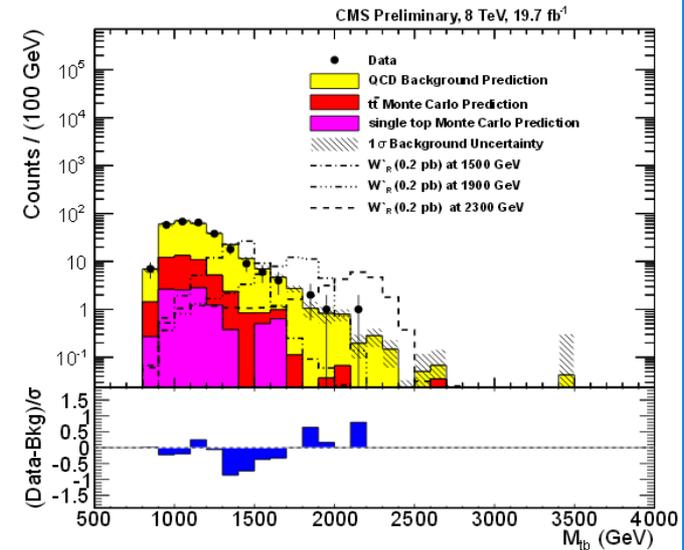
Search for tb resonances

- Background sources mainly $t\bar{t}$ and QCD
- $t\bar{t}$ estimated from Monte Carlo
 - Data-driven normalization measurement
 - Use b candidate mass control region
- QCD Shape and normalization from data
 - Find average b -tagging rate for QCD jets in a control region
 - Apply average b -tagging rate to signal region in lieu of b -tagging



Search for tb resonances

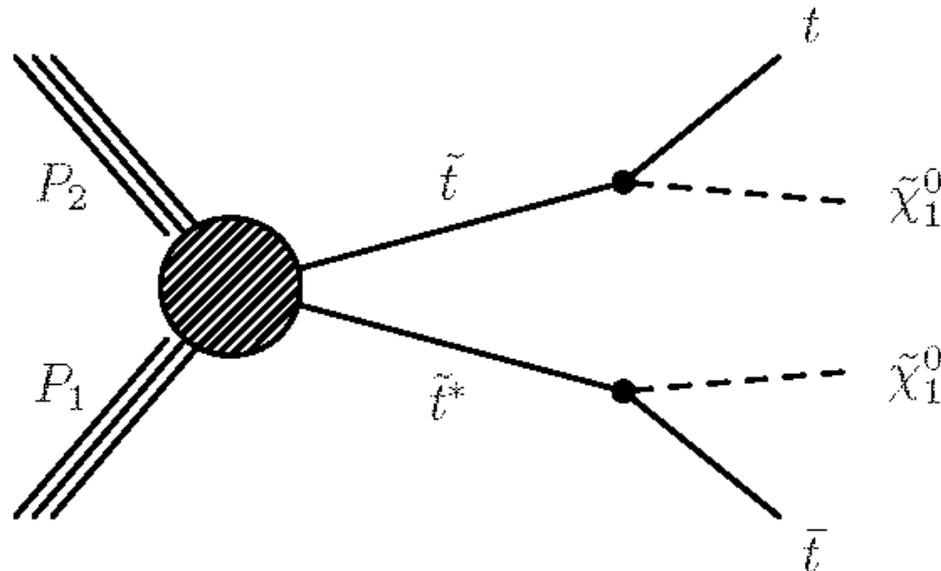
- Exclude $M_{W'_R} < 2.02$ TeV
- Place limits on left- and right-handed coupling strengths
- Combination with the semileptonic $W' \rightarrow tb$ decay channel underway (B2G-12-010)



Search for a top squark

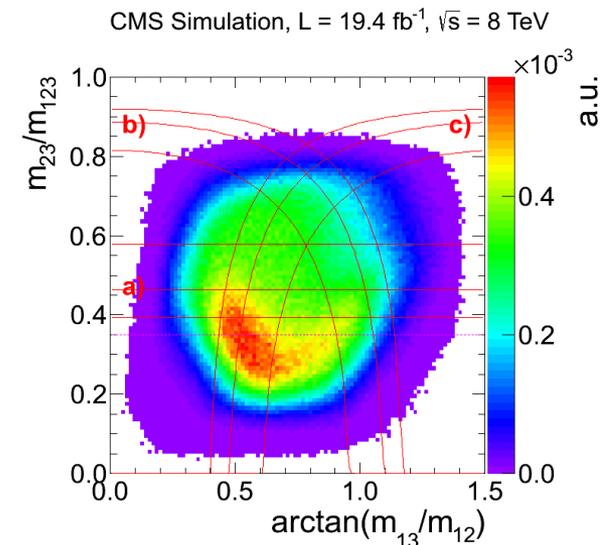
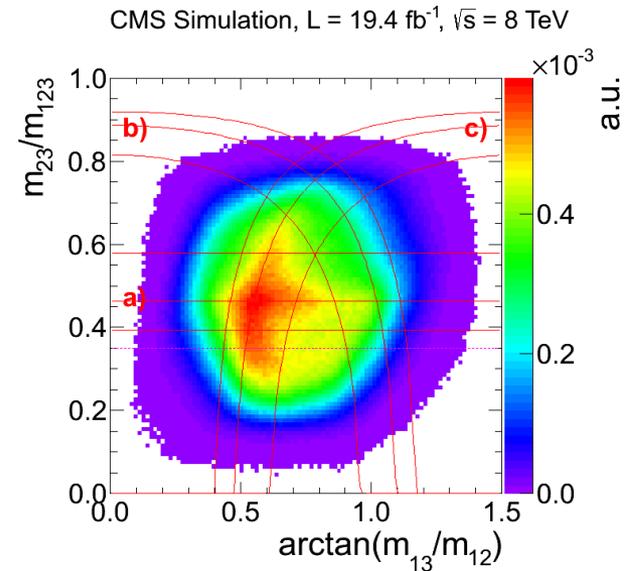
Search for a top squark

- Search for top squark pair production
 - $\tilde{t}\tilde{t} \rightarrow (t\tilde{\chi}_1^0)(t\tilde{\chi}_1^0)$
 - $t \rightarrow (Wb) \rightarrow$ All Hadronic
 - $\tilde{\chi}_1^0$ (neutralino) detectable as MET



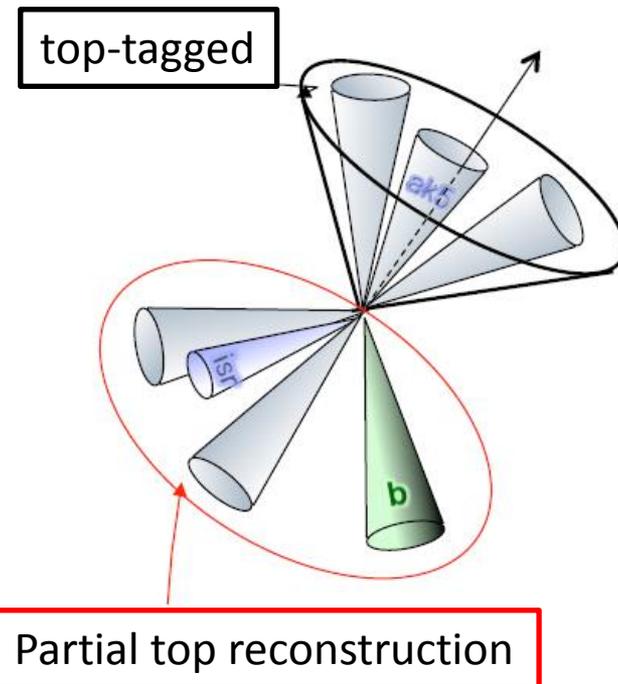
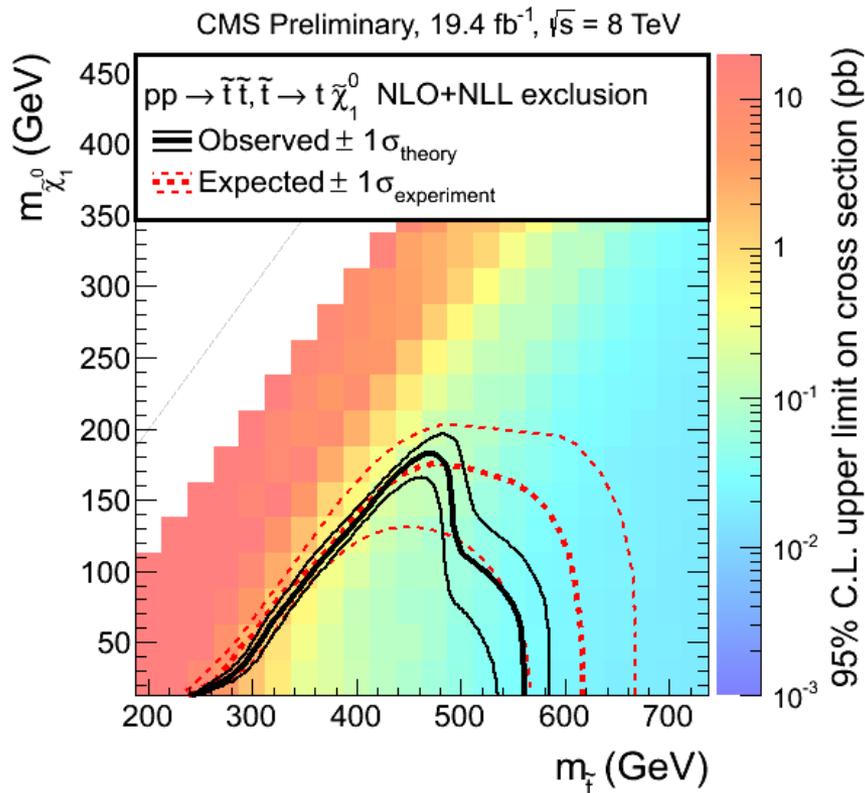
Search for a top squark

- HEP top-tagger (slightly modified)
 - Fully resolved AK5 jets instead of CA15 subjets
 - $\Delta R < 1.5$
 - Pass one of following
 - a.) $R_{\min} < m_{23}/m_{123} < R_{\max}$
 - b.) $R_{\min} < m_{12}/m_{123} < R_{\max}$
 - c.) $R_{\min} < m_{13}/m_{123} < R_{\max}$
 - $R_{\min} = 0.85 M_W/M_{\text{top}}$
 - $R_{\max} = 1.25 M_W/M_{\text{top}}$
 - $m_{23}/m_{123} > 0.35$
- Partial top reconstruction
 - Consider the remaining jets
 - Higher signal acceptance than full top tag



Search for a top squark

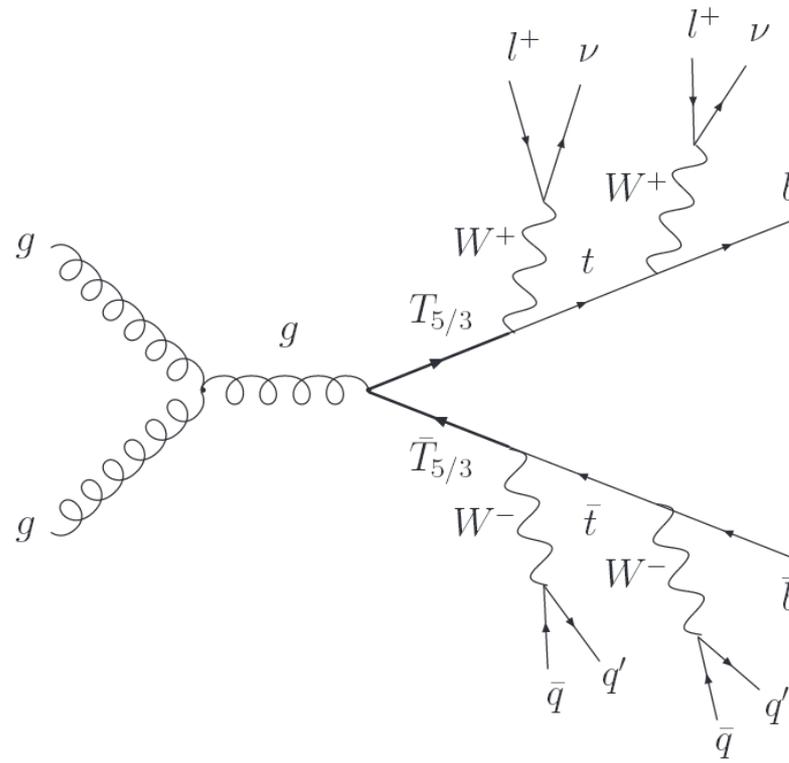
- Place limits in $m_{\tilde{t}}$, $m_{\tilde{\chi}_1^0}$ space



Search for a $T_{5/3}$ top partner

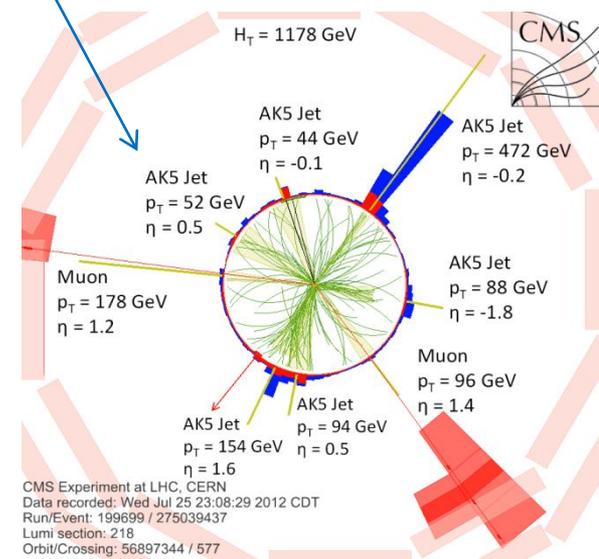
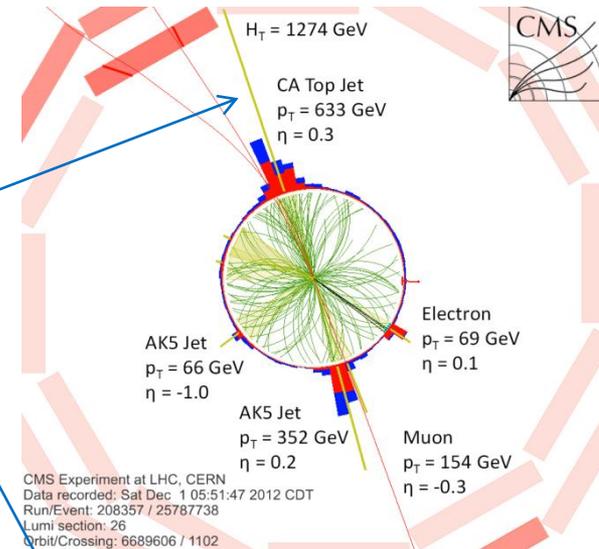
Search for a $T_{5/3}$ top partner

- Search for pair-produced $T_{5/3}$
 - $T_{5/3}T_{5/3} \rightarrow (W^+t)(W^-\bar{t})$
 - Same-sign dileptons



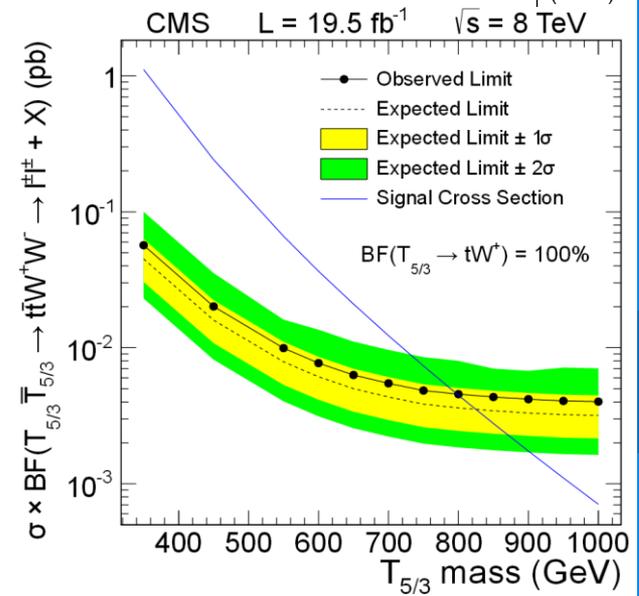
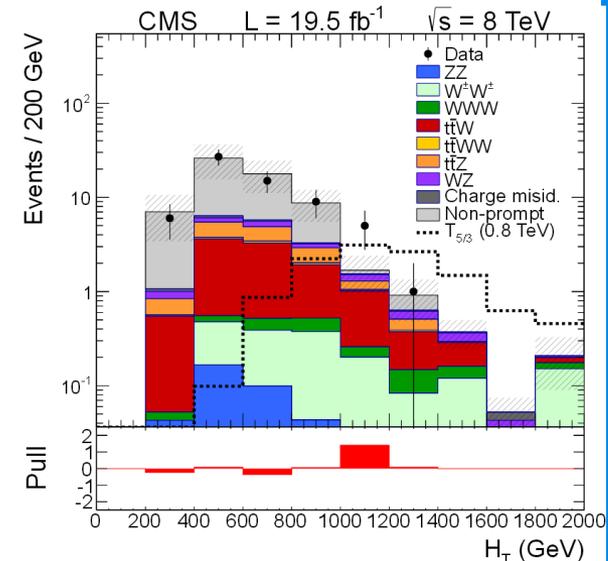
Search for a $T_{5/3}$ top partner

- Use CMS top-tagger and boosted W-tagging
 - Search for top-tagged jets
 - Search for W-tagged jets
 - Search for AK5 jets
 - $\Delta R > 0.8$ from CA8 tagged jets
- Require at least 5 “physics objects” over the two same-sign leptons
 - Leptons other than the same sign leptons = 1
 - AK5 jets = 1
 - CA8 W-tagged jet = 2
 - CA8 top-tagged jet = 3



Search for a $T_{5/3}$ top partner

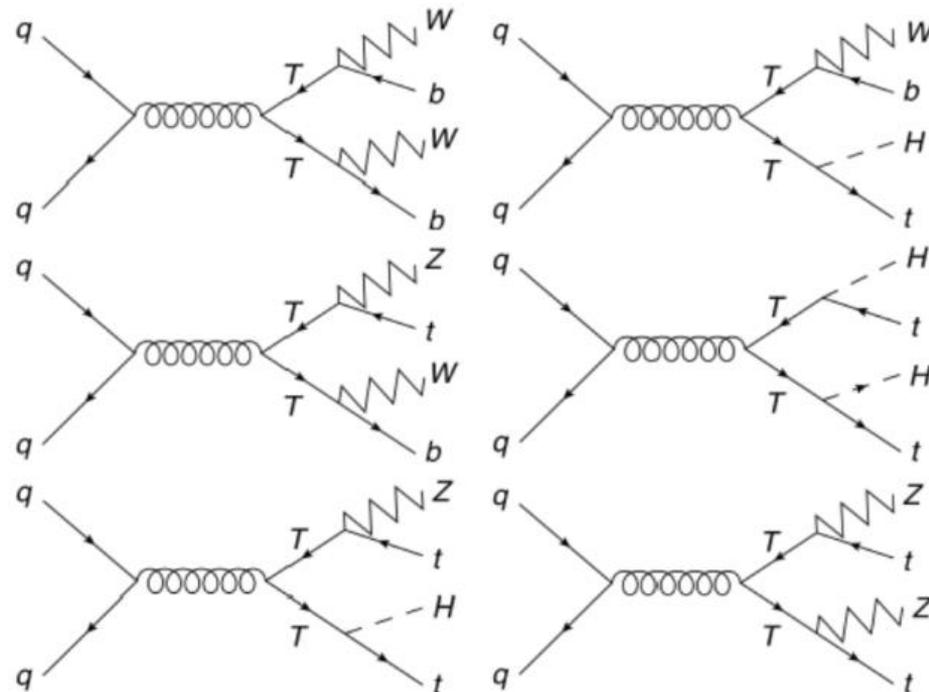
- Rare backgrounds
 - WZ, ZZ, $t\bar{t}W$ etc.
 - Opposite sign backgrounds
 - $t\bar{t}$, Drell-Yan
 - Charge mis-identification
 - Non-prompt background
 - Jets misidentified as leptons, conversions, etc.
 - data driven technique
- $T_{5/3}$ mass < 800 GeV



Search for a T' top partner

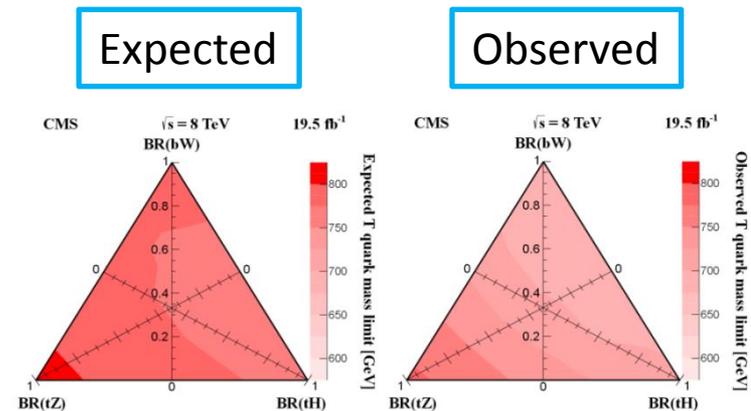
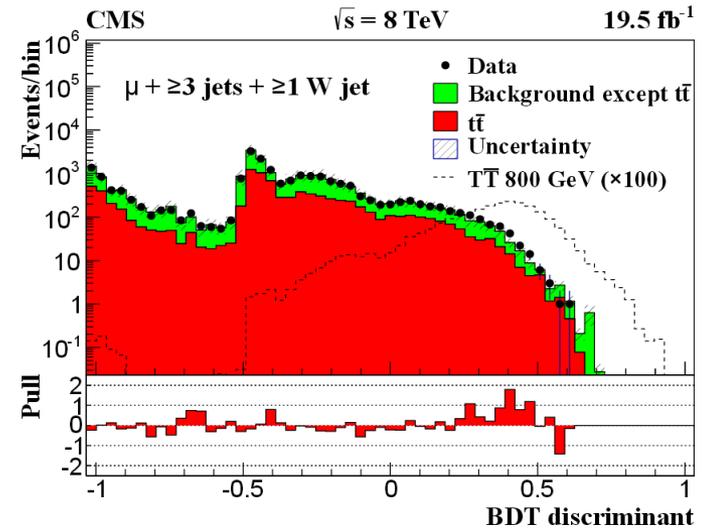
Search for a T' top partner

- Pair produced T'
 - $T'T' \rightarrow (XY)(X'Y')$
- Scan T' BR space
 - $T' \rightarrow Wb$
 - $T' \rightarrow tZ$
 - $T' \rightarrow tH$
- Combination of lepton+jets and multilepton channels
- CMS top-tagger
- Boosted W-tagging



Search for a T' top partner

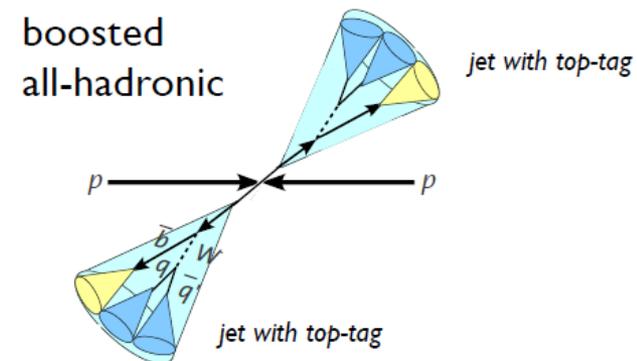
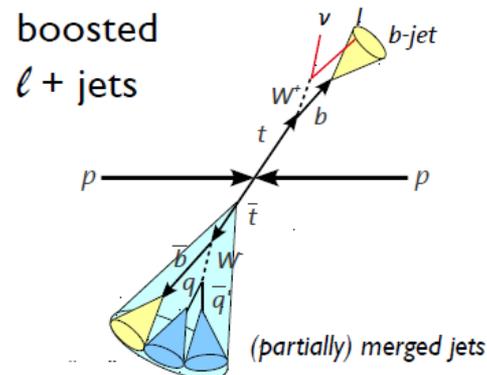
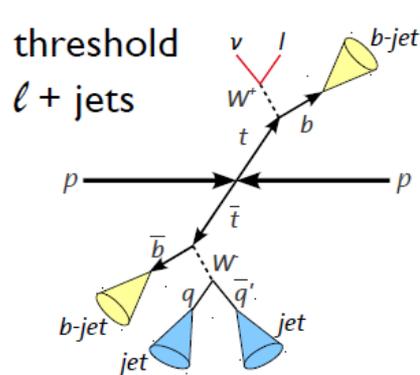
- Background estimate from Monte Carlo
 - Primarily $t\bar{t}$, W+Jets
 - Derive W+Jets k-factors from data
 - Corrections to SHyFT k-factors
- Boosted Decision Tree (BDT)
 - Number of top-jets used as BDT input
 - Along with NJets, Nhtags, MET, lepton p_T etc.
- T' mass lower limit between 687 and 782 GeV



Search for BSM $t\bar{t}$ resonances

Search for BSM $t\bar{t}$ resonances

- Search for heavy $t\bar{t}$ resonances
 - $t\bar{t} \rightarrow (Wb)(Wb)$
 - Narrow Z'
 - Wide Z'
 - KK gluon
- Combination of
 - Semileptonic channel (B2G-12-006)
 - All-hadronic (B2G-12-005)



Search for BSM $t\bar{t}$ resonances

- All-hadronic

- Use CMS top-tagger
- $t\bar{t}$ background from Monte Carlo
- NTMJ (QCD) background from data

- Predict parameter (F)
- Using top-mistagging rate (r)
- Weight pre top-tagged sample (N)

$$F_k = \sum_{i=1}^{N_{jets}} N_{k,i} \times r_i,$$

- Anti-tagged jet

- Invert substructure selection $M_{\min} < 30$ GeV
- Keep top mass window

- Probe jet

- Top-tag for mistag rate

- Top-candidate mass not correctly predicted

- Randomly drawn from QCD Monte Carlo distribution (keep top mass window)

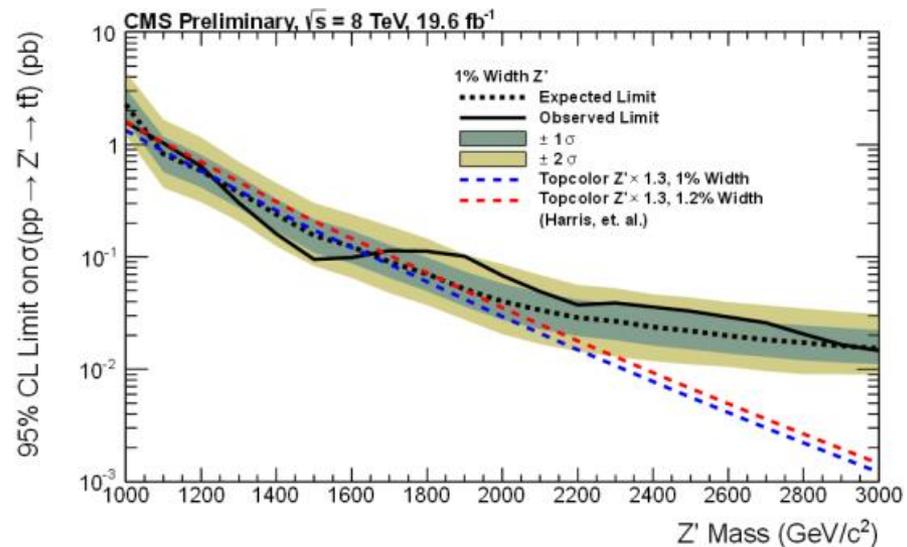
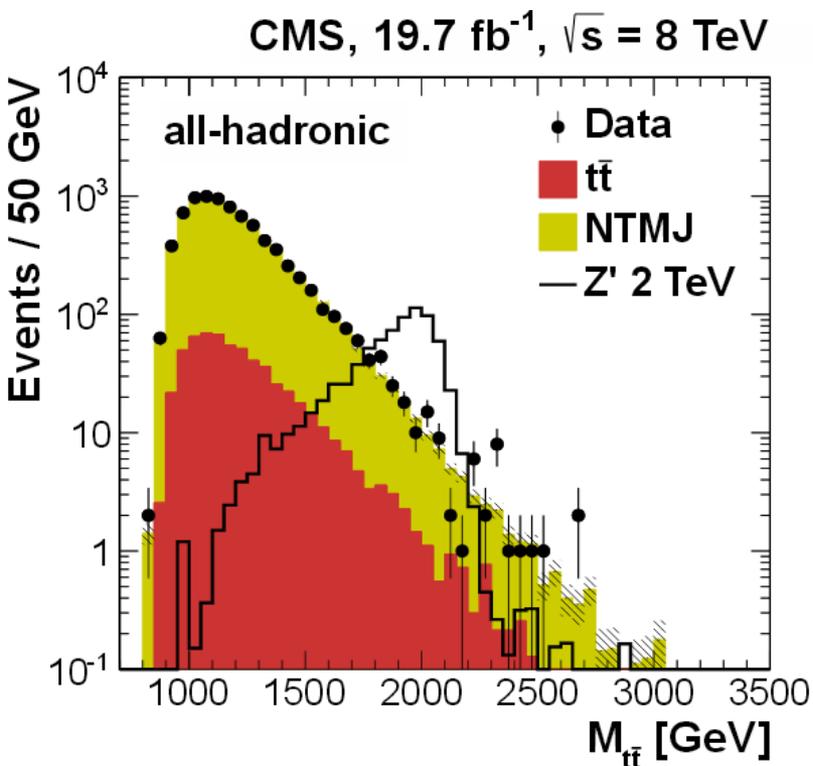
- $t\bar{t}$ background subtracted from top-mistagging rate

- Take into account correlated and uncorrelated statistical uncertainties

$$\sigma(F_k) = \sqrt{\sum_{i=1}^{N_{jets}} \left((N_{k,i} \times \sigma(r_i))^2 + \left(\sqrt{N_{k,i}} \times r_i \right)^2 \right)^2},$$

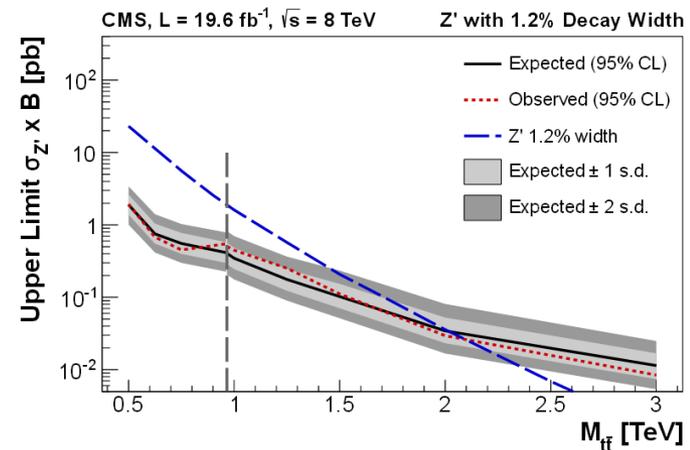
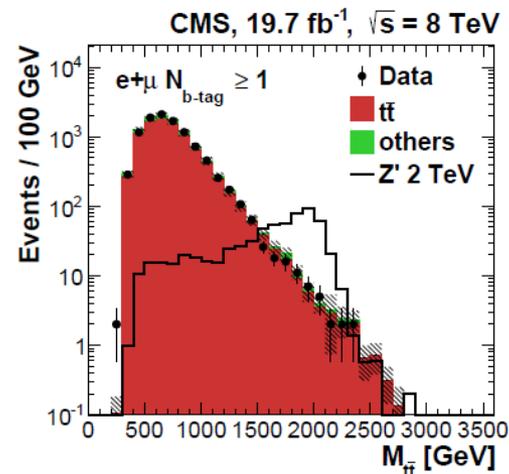
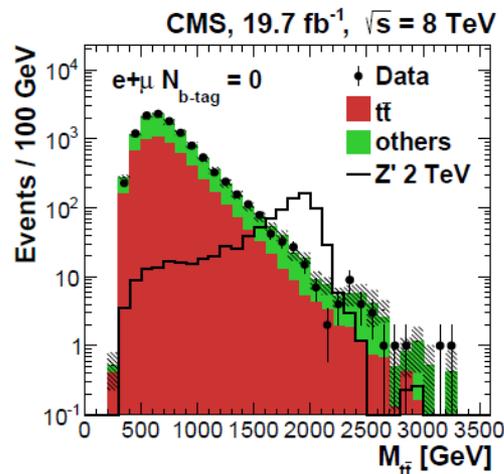
Search for BSM $t\bar{t}$ resonances

- Exclude Narrow Z' mass < 1.7 TeV



Search for BSM $t\bar{t}$ resonances

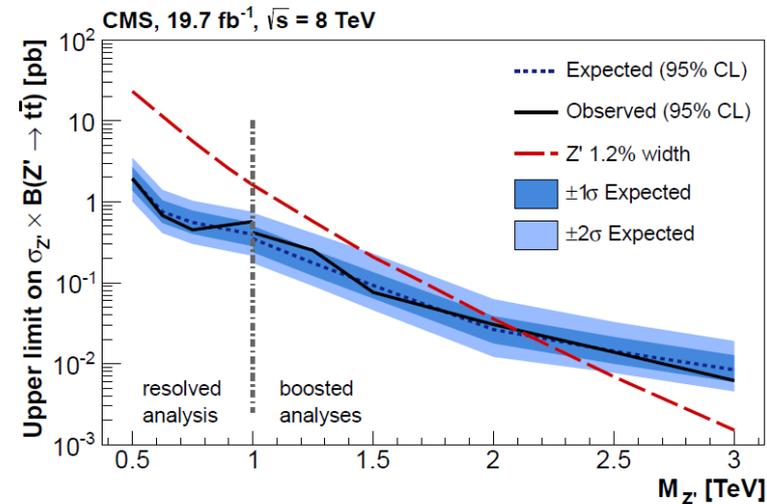
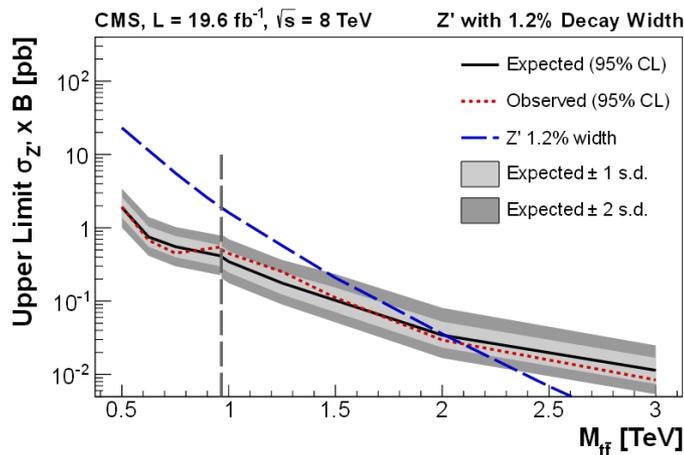
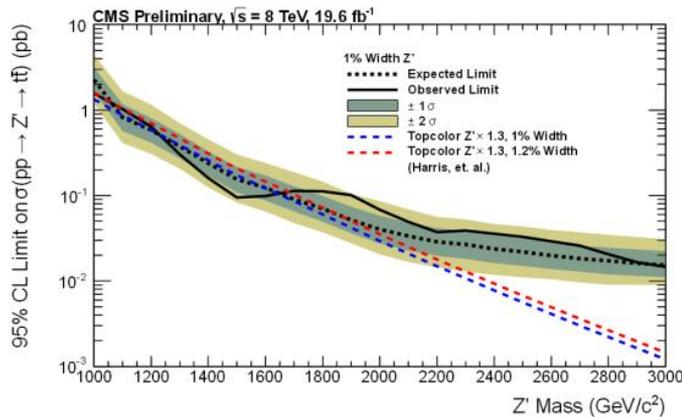
- Boosted Semileptonic
 - Lepton isolation breaks down
 - Boosted semileptonic top selection
 - $\Delta R(\text{lepton, closest jet}) > 0.5$
OR
 $p_T^{\text{rel}}(\text{lepton, closest jet}) > 25 \text{ GeV}$
 - Background estimate from Monte Carlo
 - Primarily $t\bar{t}$, W+Jets
 - Exclude Narrow Z' mass $< 2.1 \text{ TeV}$



Search for BSM $t\bar{t}$ resonances

• Combination

- Exclude Narrow Z' mass < 2.1 TeV



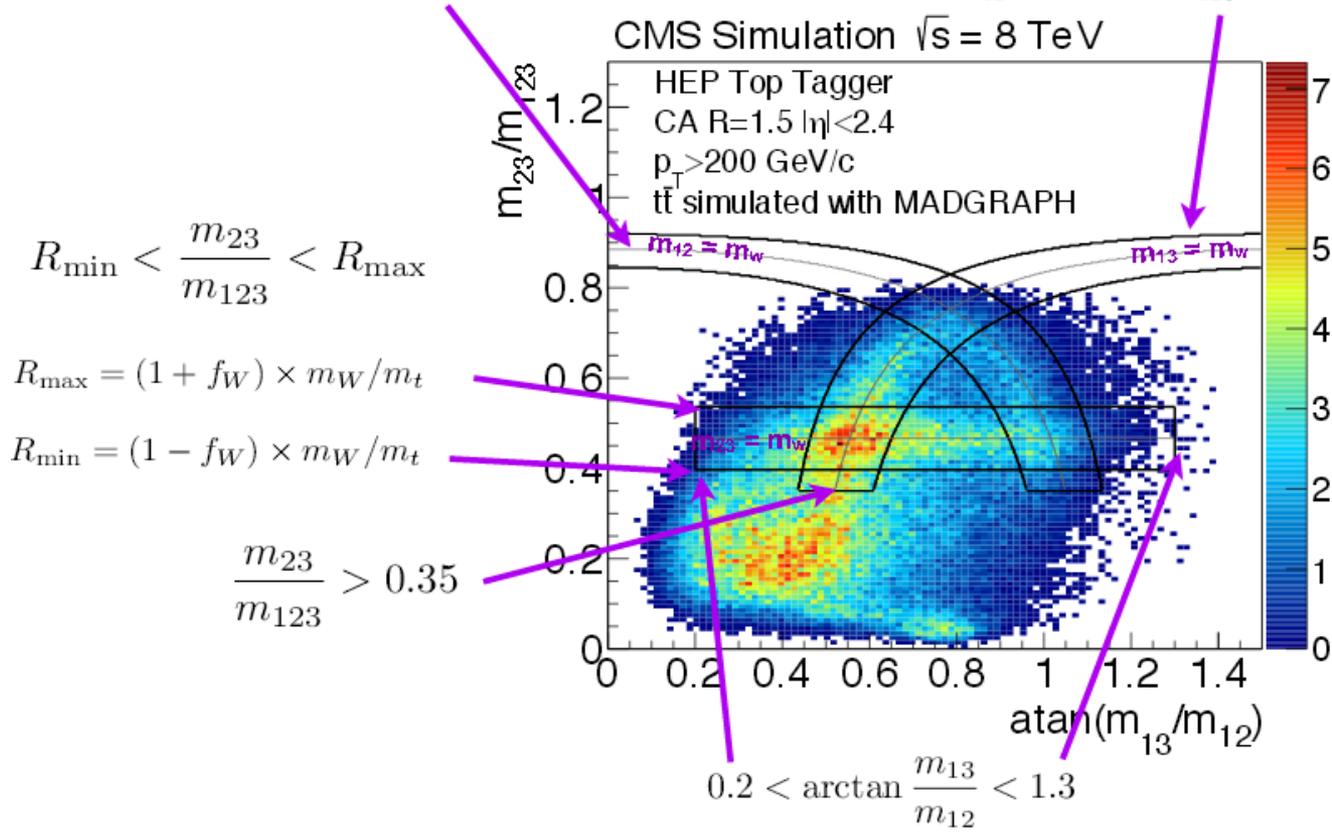
Summary

- Searches for BSM signals decaying to top quarks have been performed
- Boosted top-tagging techniques used to place limits in the high resonant mass regime
- Constantly evolving techniques enhance sensitivity further than ever before
- No excesses found at 8TeV, but techniques will prove invaluable at 13TeV

BACKUP

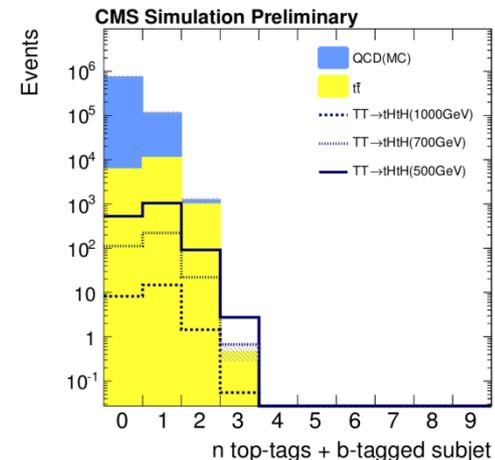
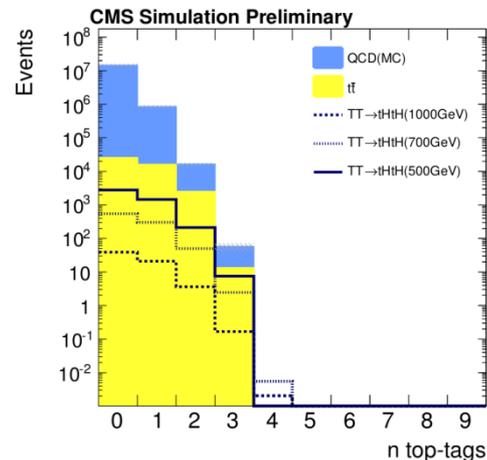
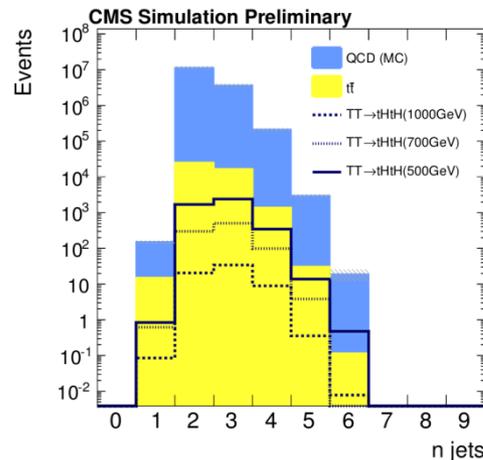
HEP top-tagger

$$R_{\min}^2 \left(1 + \left(\frac{m_{12}}{m_{13}}\right)^2\right) < 1 - \left(\frac{m_{23}}{m_{123}}\right)^2 < R_{\max}^2 \left(1 + \left(\frac{m_{12}}{m_{13}}\right)^2\right) \quad R_{\min}^2 \left(1 + \left(\frac{m_{13}}{m_{12}}\right)^2\right) < 1 - \left(\frac{m_{23}}{m_{123}}\right)^2 < R_{\max}^2 \left(1 + \left(\frac{m_{13}}{m_{12}}\right)^2\right)$$



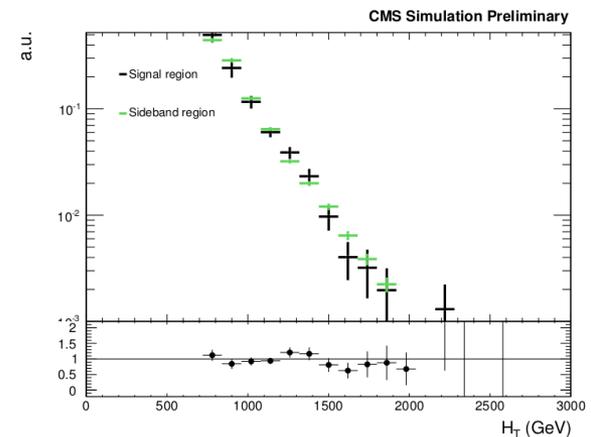
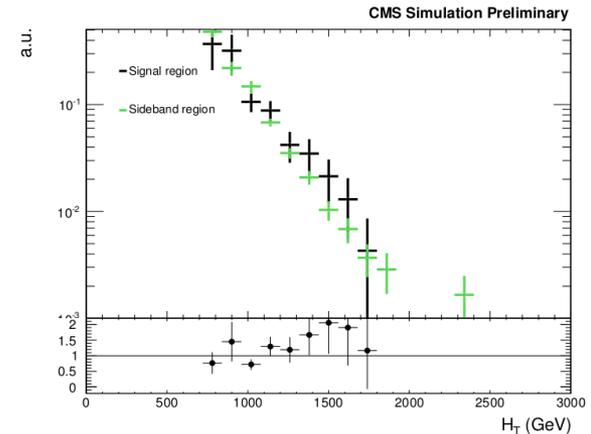
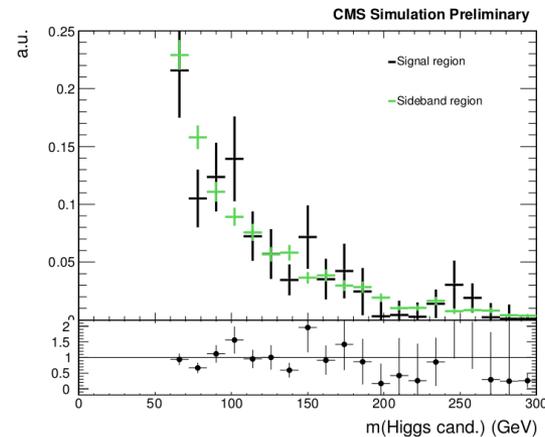
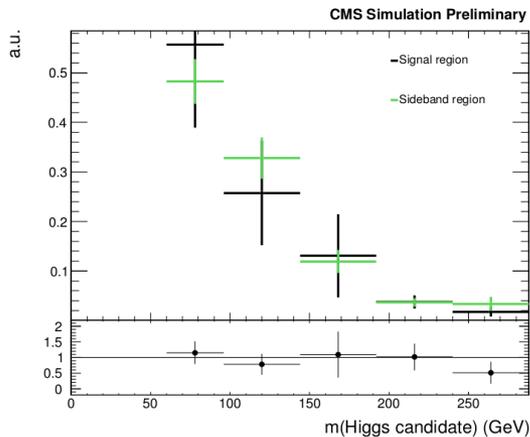
Search for tH resonances

- Selection cut flow
 - Number of Jets ($pt > 150$)
 - Number of Jets HTT ($pt > 200$)
 - Number of Jets HTT + SJBtag



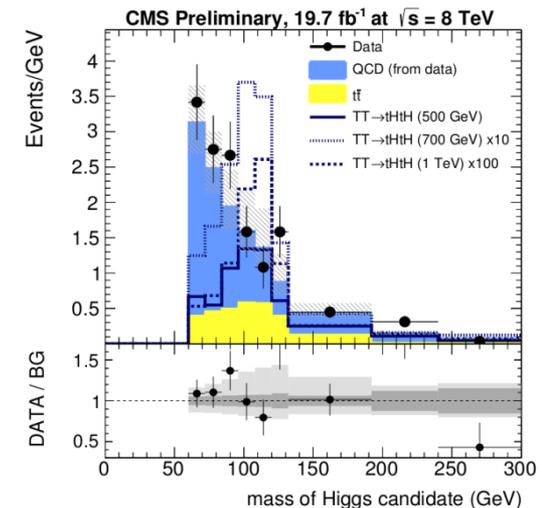
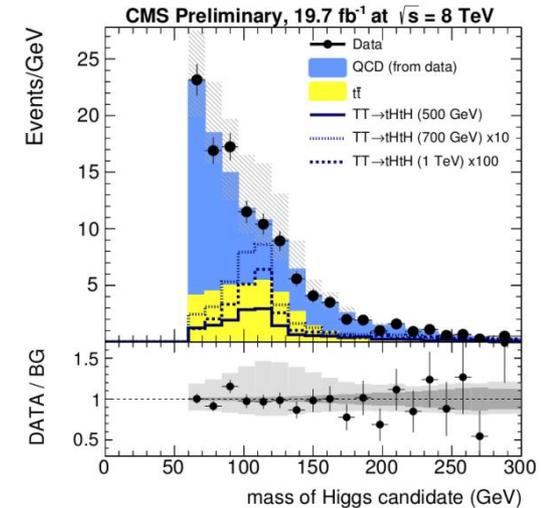
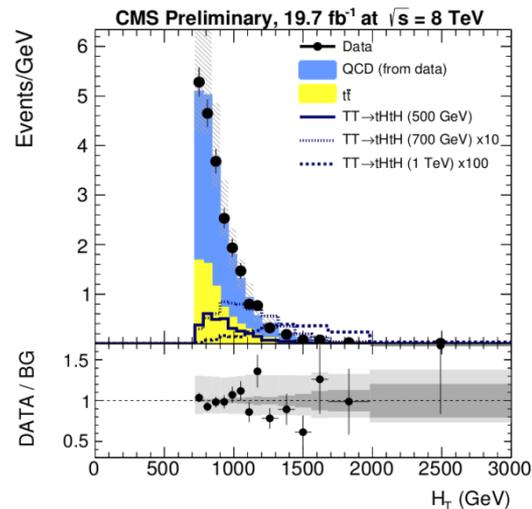
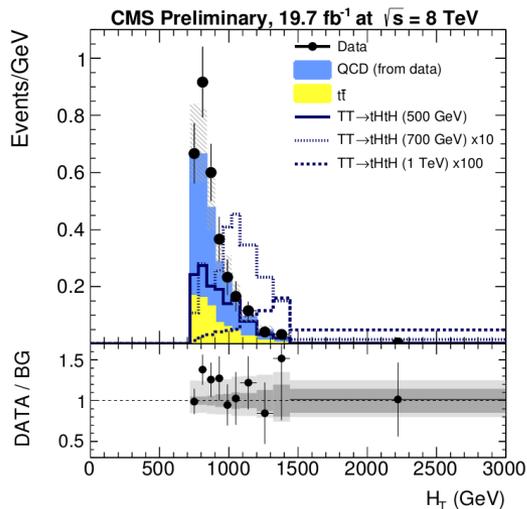
Search for tH resonances

- Closure of B region shapes
 - For H_T and m_H shapes
 - Single and multi-higgs channels



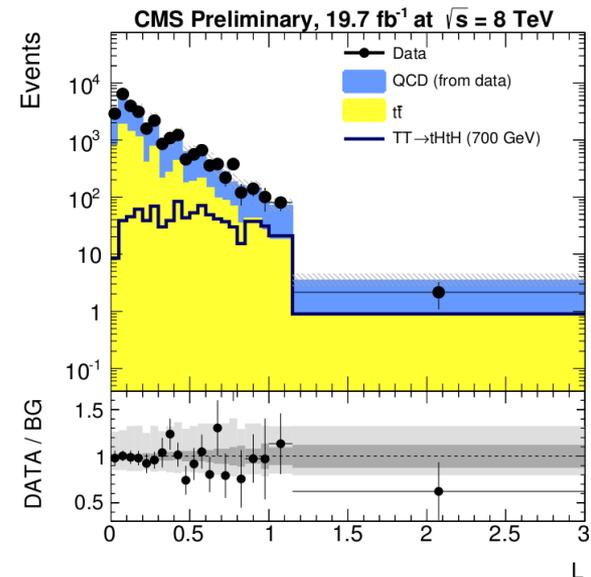
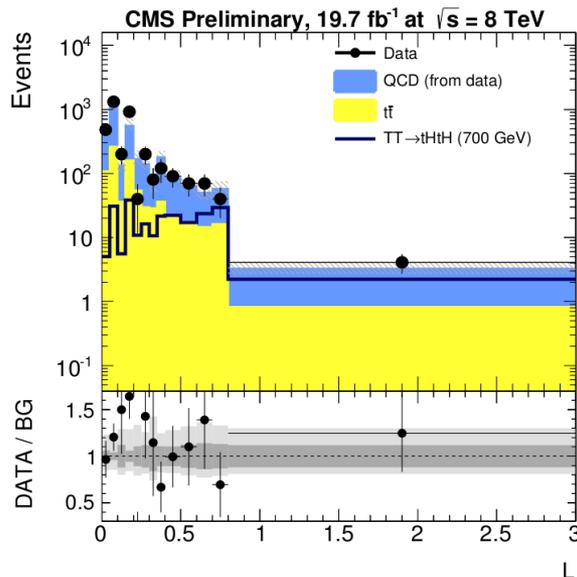
Search for tH resonances

- Signal region H_T and m_H shapes



Search for tH resonances

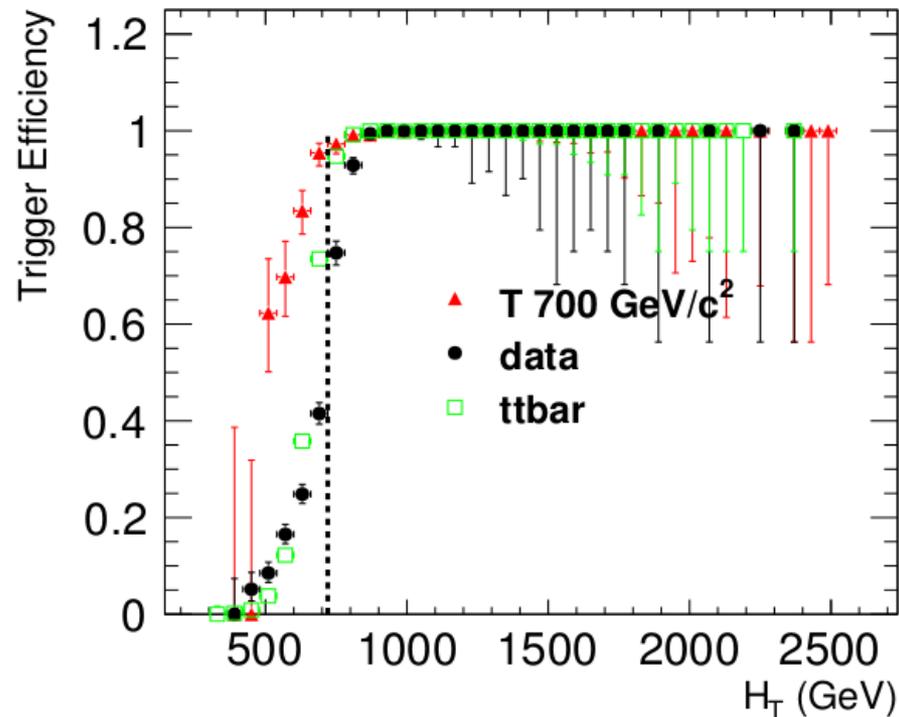
- Signal region L shapes 700GeV



Search for tH resonances

- Trigger efficiency HT750

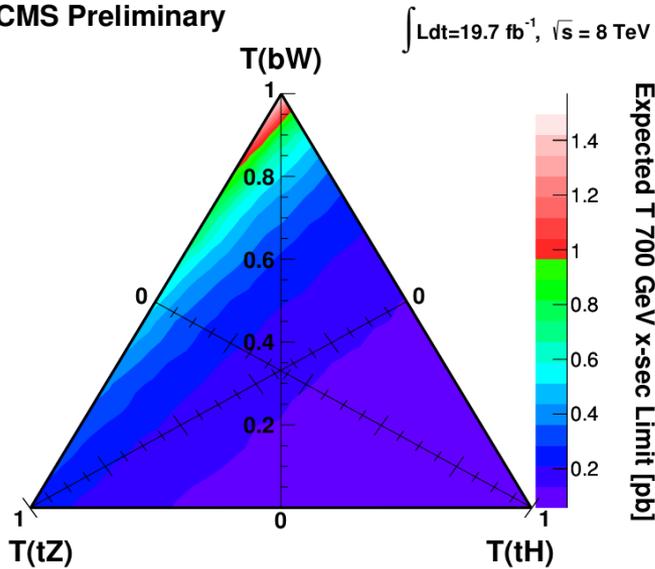
CMS Preliminary, 19.7 fb⁻¹ at $\sqrt{s} = 8$ TeV



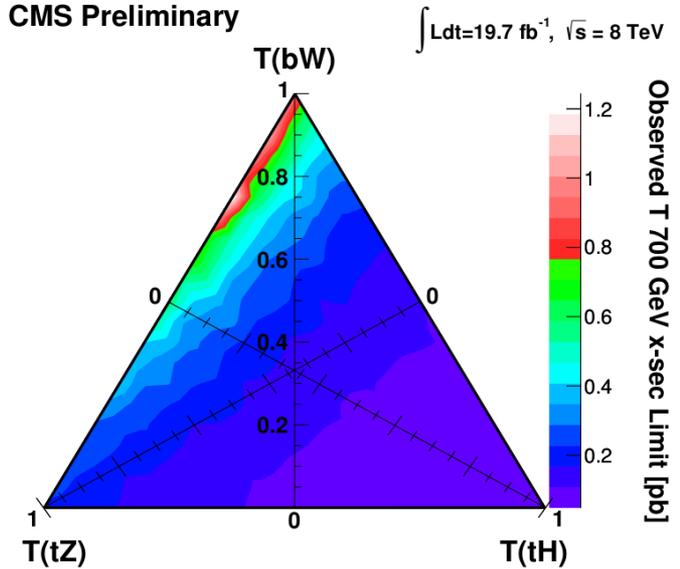
Search for tH resonances

- Cross section exclusion 700GeV

CMS Preliminary

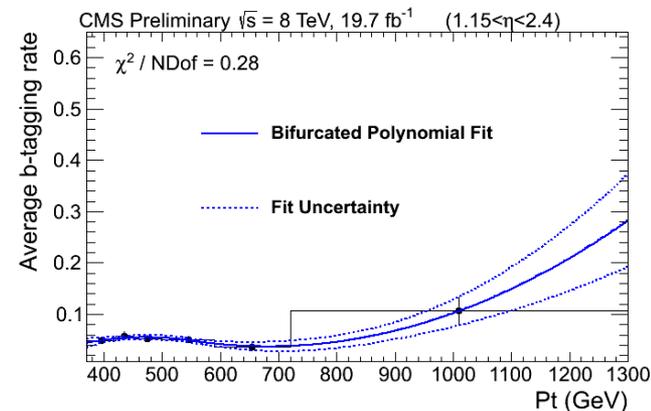
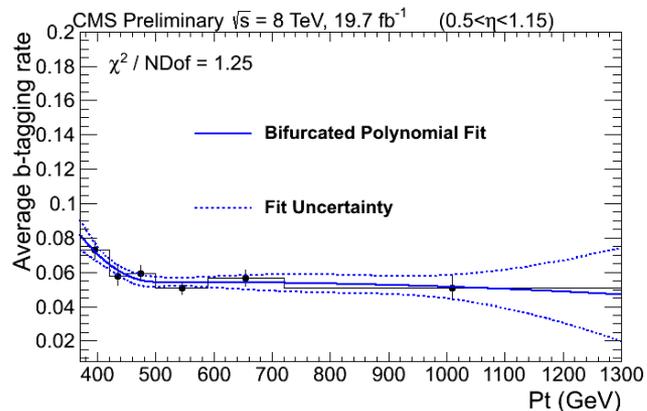
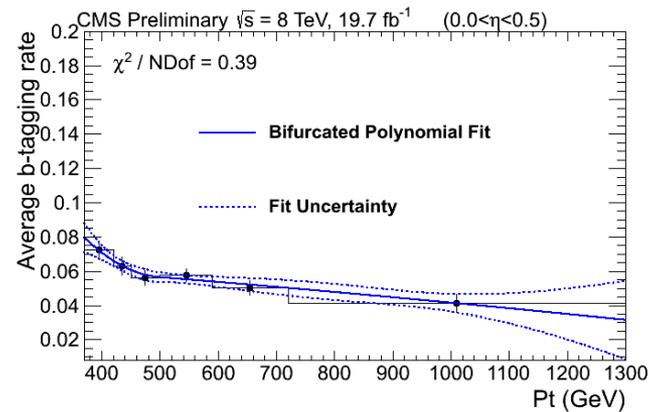


CMS Preliminary



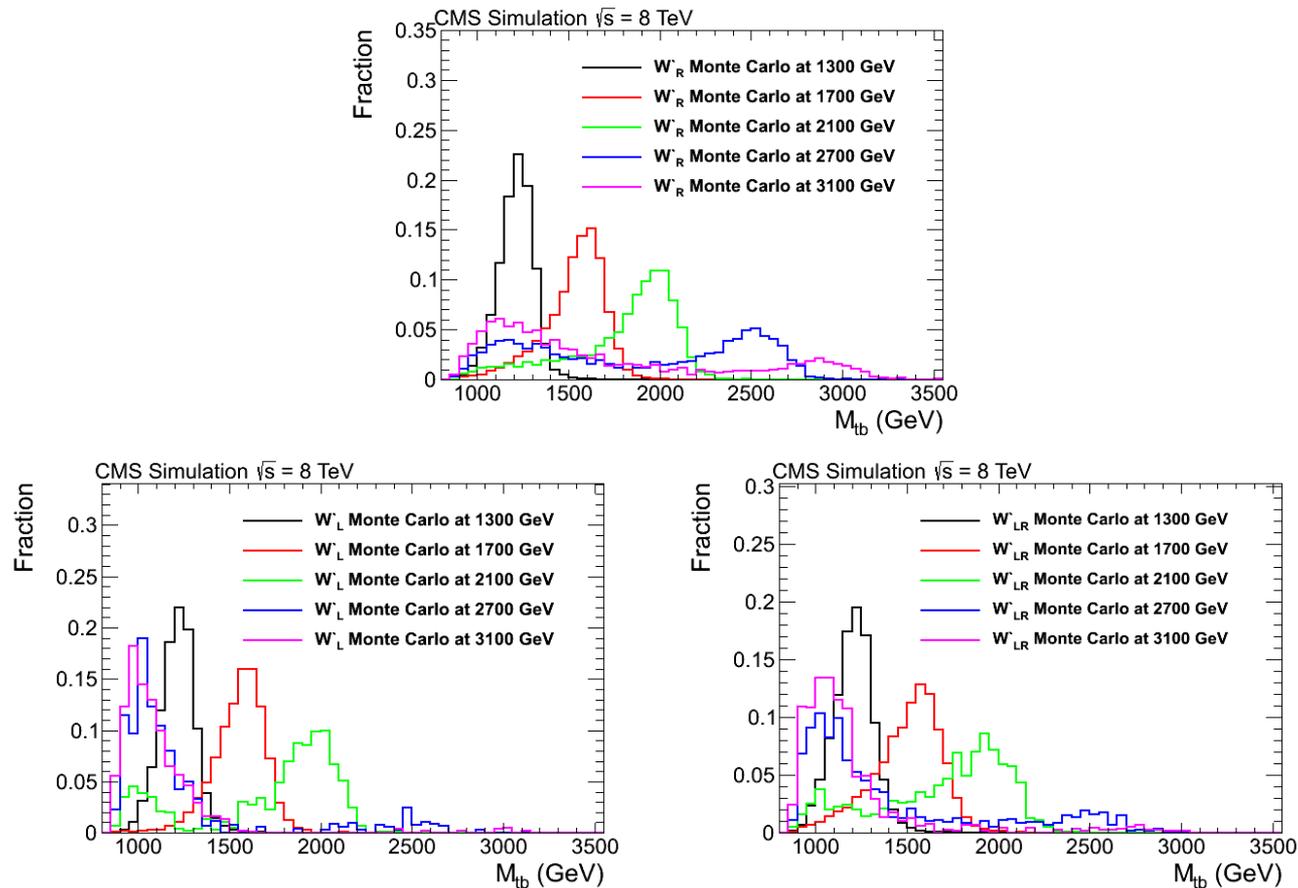
Search for a W' Boson

- Average b-tagging rate regions



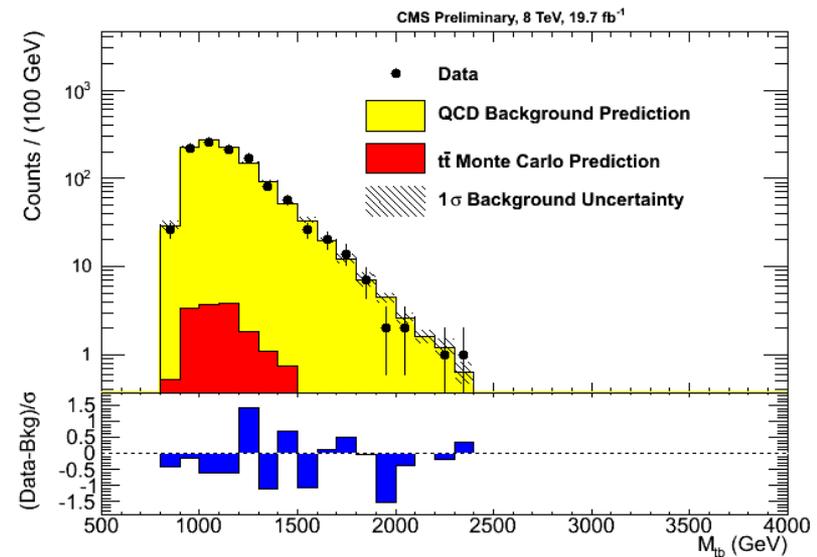
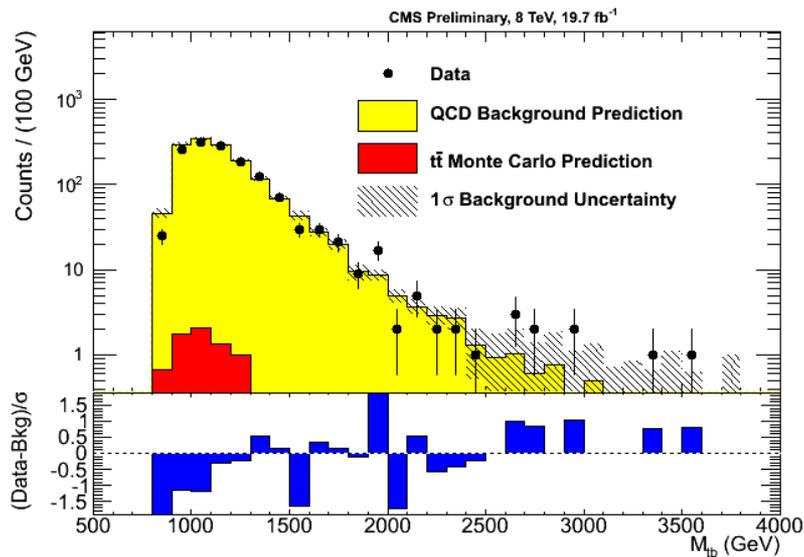
Search for a W' Boson

- Signal shapes



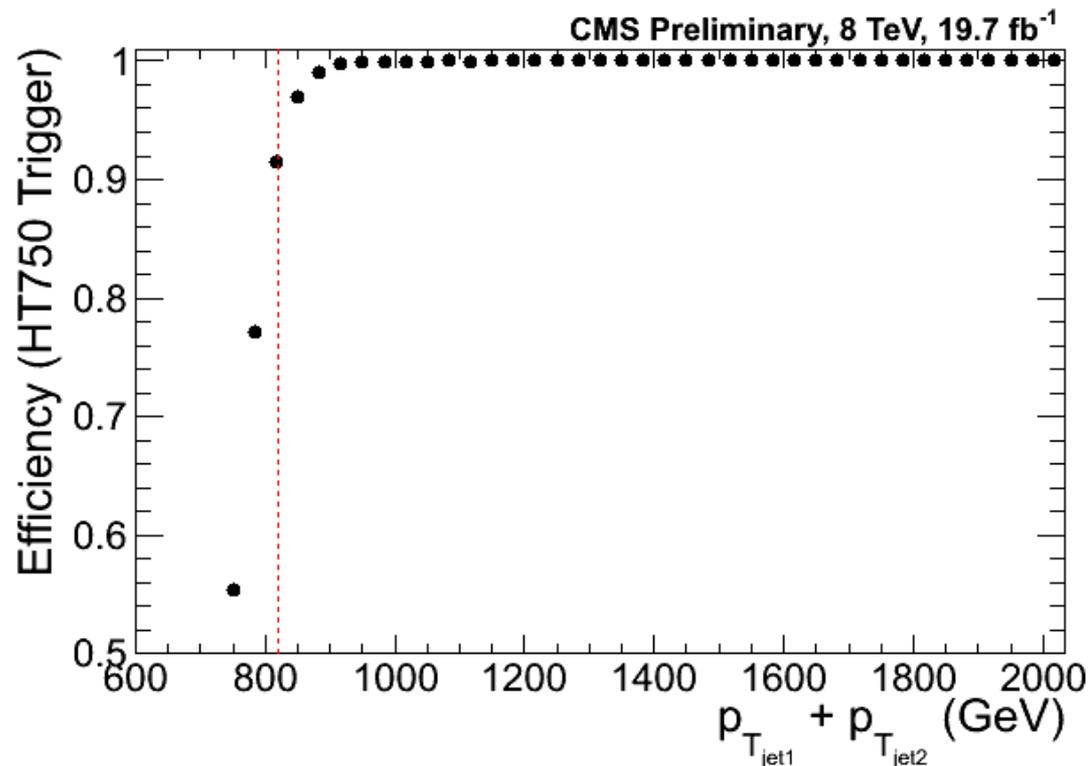
Search for a W' Boson

- Closure regions



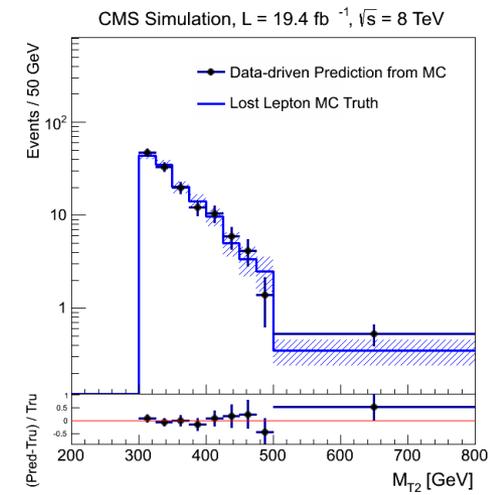
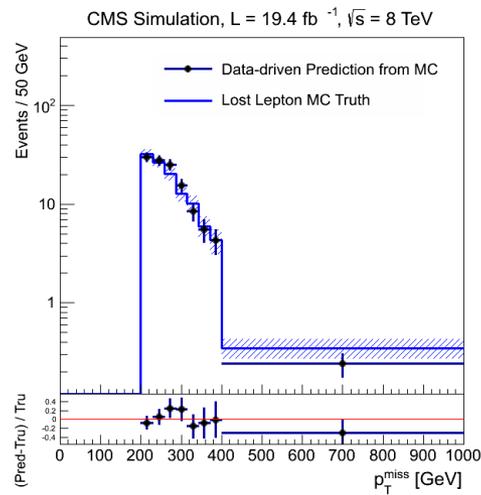
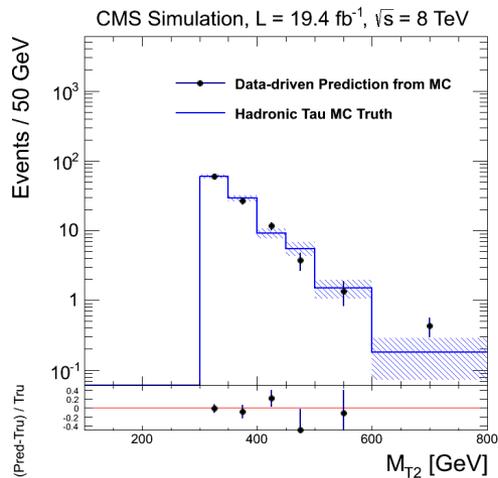
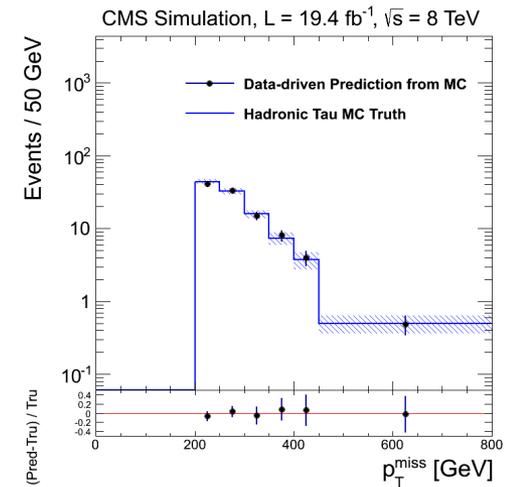
Search for a W' Boson

- Trigger efficiency HT750



Search for a top squark

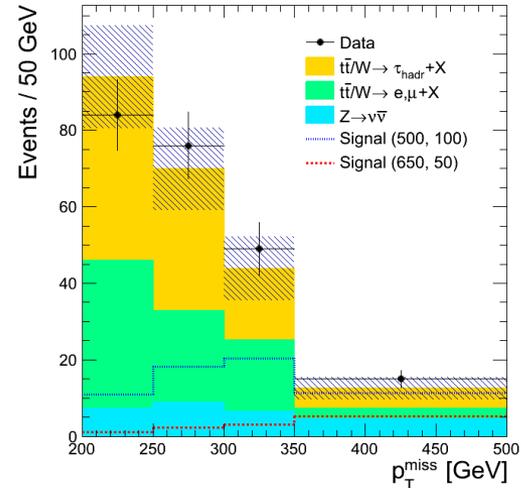
- Lost lepton and tau background closure test



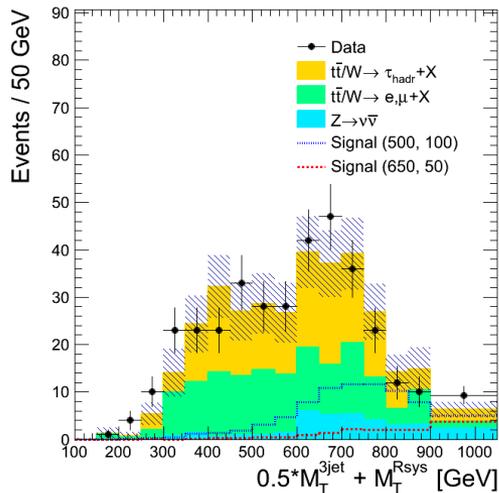
Search for a top squark

- Kinematic variables in signal region
 - Baseline cuts

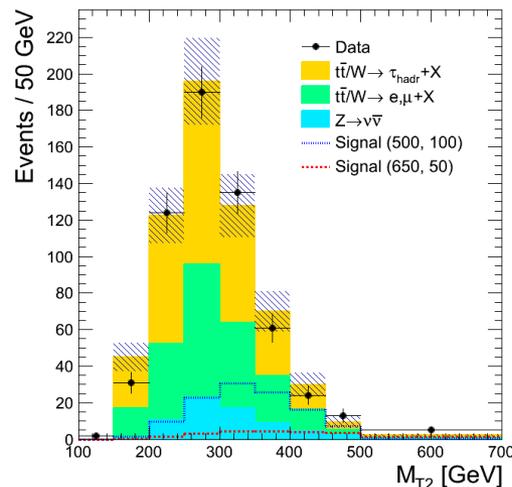
CMS Preliminary, $L = 19.4 \text{ fb}^{-1}$, $\sqrt{s} = 8 \text{ TeV}$



CMS Preliminary, $L = 19.4 \text{ fb}^{-1}$, $\sqrt{s} = 8 \text{ TeV}$



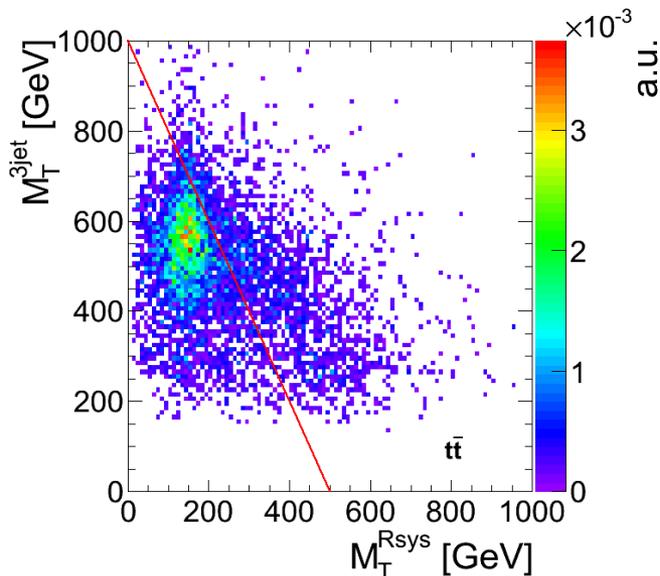
CMS Preliminary, $L = 19.4 \text{ fb}^{-1}$, $\sqrt{s} = 8 \text{ TeV}$



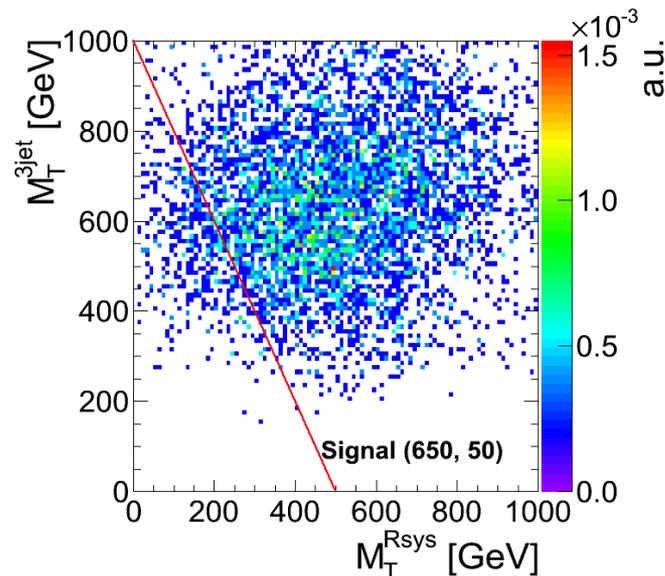
Search for a top squark

- $t\bar{t}$ discriminating selection

CMS Simulation, $L = 19.4 \text{ fb}^{-1}$, $\sqrt{s} = 8 \text{ TeV}$

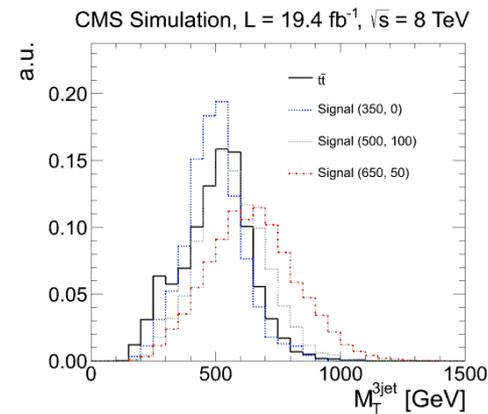
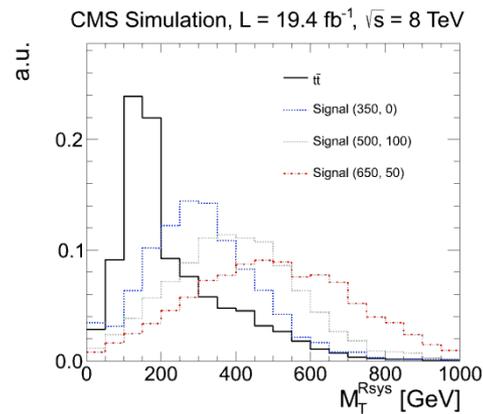
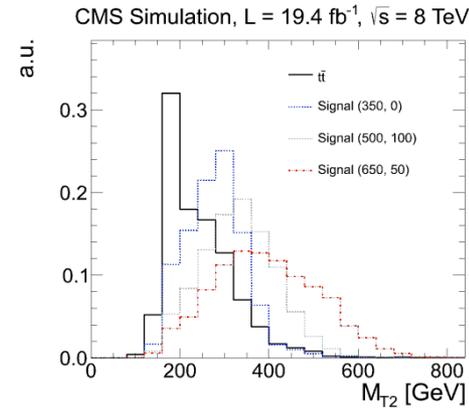
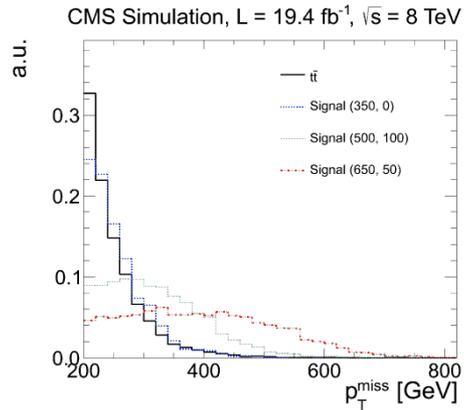


CMS Simulation, $L = 19.4 \text{ fb}^{-1}$, $\sqrt{s} = 8 \text{ TeV}$



Search for a top squark

- $t\bar{t}$ discriminating selection



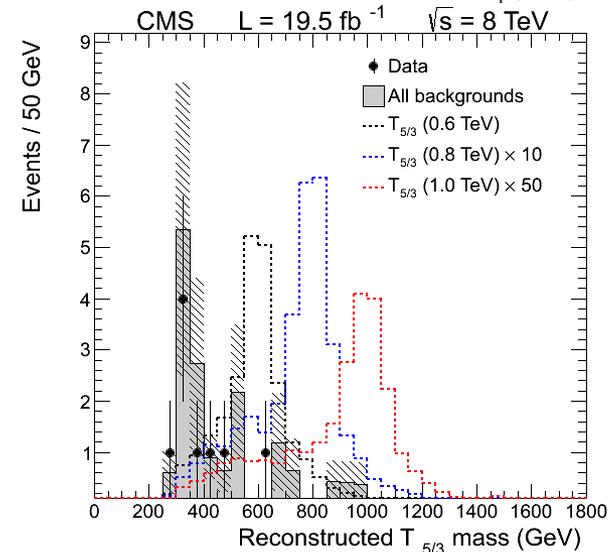
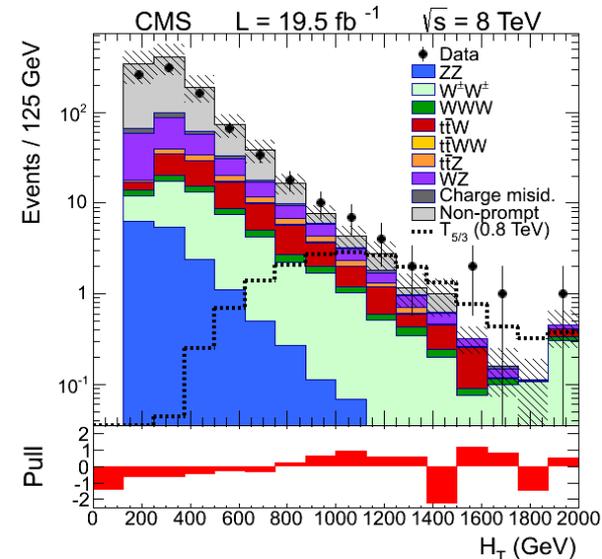
Search for a top squark

- Cutflow

	$p_T^{\text{miss}} > 200 \text{ GeV},$ $N_{\text{b-jets}} \geq 1$	$p_T^{\text{miss}} > 350 \text{ GeV},$ $N_{\text{b-jets}} \geq 1$	$p_T^{\text{miss}} > 200 \text{ GeV},$ $N_{\text{b-jets}} \geq 2$	$p_T^{\text{miss}} > 350 \text{ GeV},$ $N_{\text{b-jets}} \geq 2$
$t\bar{t}$	153.8 ± 5.7	18.9 ± 2.0	63.4 ± 3.7	6.3 ± 1.2
$W \rightarrow \ell\nu$	22.9 ± 2.9	5.8 ± 1.4	3.9 ± 1.2	1.1 ± 0.6
$Z \rightarrow \nu\bar{\nu}$	25.0 ± 1.2	8.4 ± 0.6	4.6 ± 0.5	1.3 ± 0.2
QCD	1.1 ± 0.6	$0.0^{+0.5}_{-0.0}$	$0.0^{+0.5}_{-0.0}$	$0.0^{+0.5}_{-0.0}$
single top	17.5 ± 3.9	5.2 ± 2.1	7.0 ± 2.5	1.8 ± 1.2
$t\bar{t}Z$	7.8 ± 0.4	2.3 ± 0.2	4.2 ± 0.3	1.4 ± 0.2
$t\bar{t}W$	2.4 ± 0.2	0.3 ± 0.1	1.1 ± 0.2	0.1 ± 0.1
ZZ	0.8 ± 0.2	0.3 ± 0.1	0.2 ± 0.1	$0.0^{+0.1}_{-0.0}$
WZ	0.5 ± 0.2	0.1 ± 0.1	0.1 ± 0.1	$0.0^{+0.1}_{-0.0}$
WW	0.8 ± 0.3	0.1 ± 0.1	0.3 ± 0.2	$0.0^{+0.2}_{-0.0}$
Total (no QCD)	231.5 ± 7.6	41.2 ± 3.3	84.7 ± 4.6	12.0 ± 1.8
Data	254	45	83	15
Signal (350, 0)	162.8 ± 7.2	11.3 ± 1.9	84.4 ± 5.2	7.5 ± 1.5
Signal (500, 100)	83.2 ± 1.7	33.7 ± 1.1	48.1 ± 1.3	19.8 ± 0.8
Signal (650, 50)	22.4 ± 0.4	15.8 ± 0.3	13.1 ± 0.3	9.3 ± 0.2

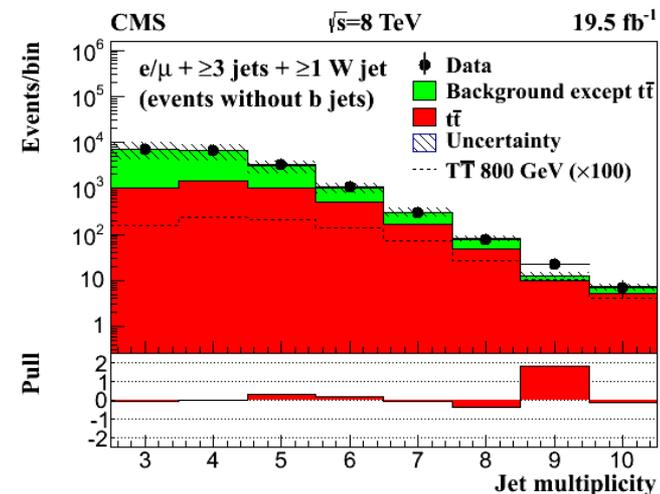
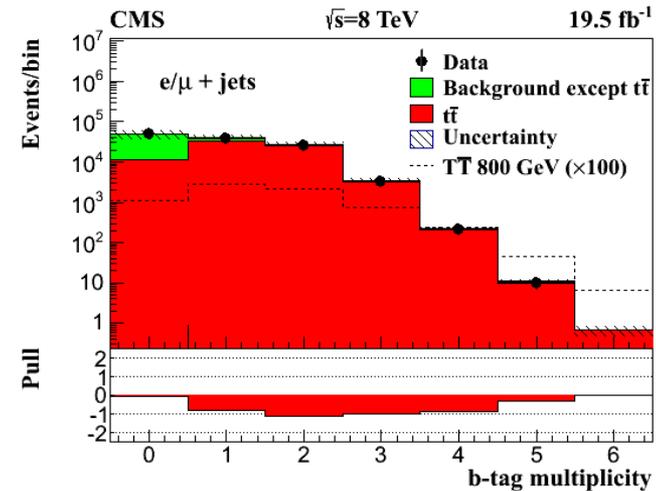
Search $T_{5/3}$ top partner

- H_T after SS selection, Z/quarkonia lepton invariant mass veto and at least 2 jets
- $M_{T_{5/3}}$ for all channels in the signal region



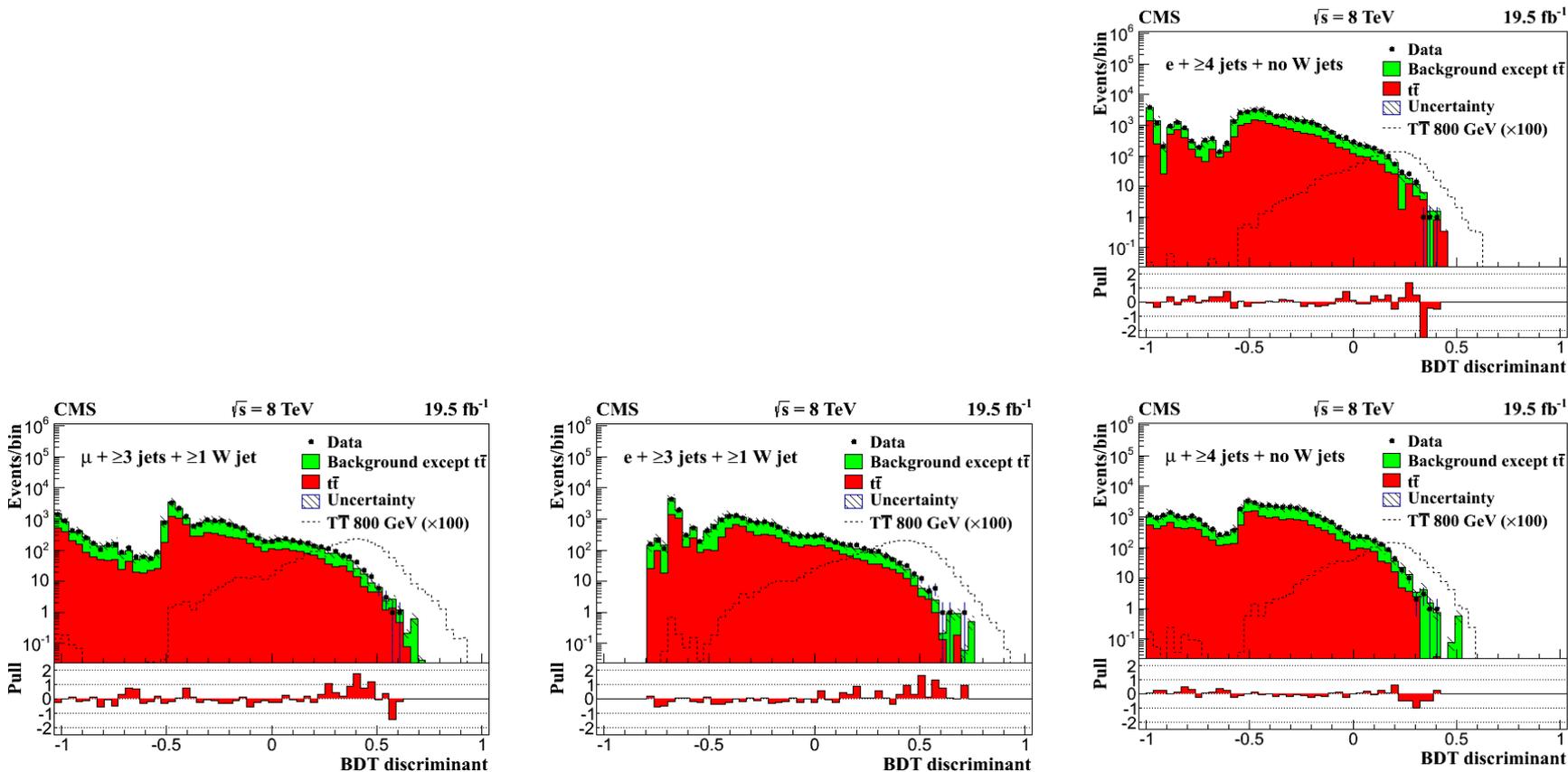
Search $T_{2/3}$ top partner

- b-tag and jet multiplicities



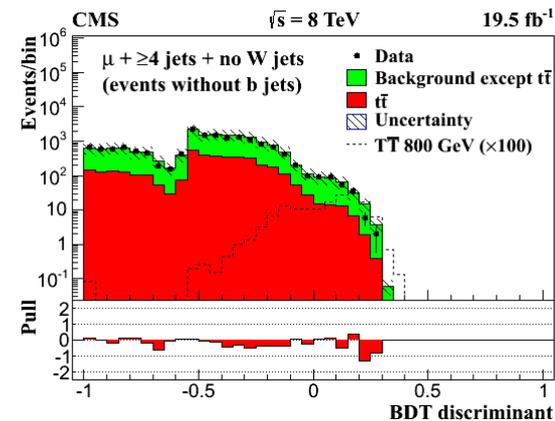
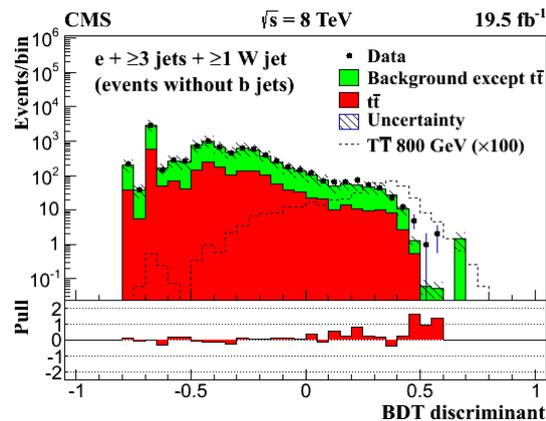
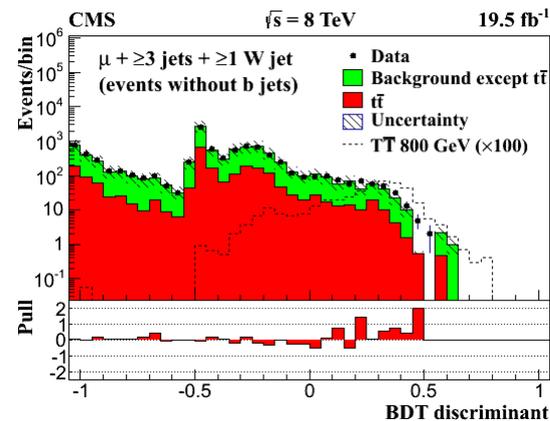
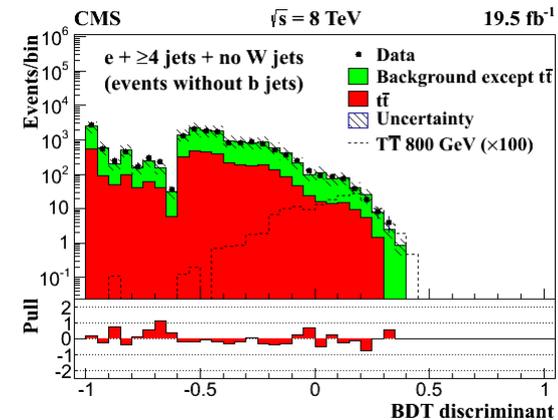
Search $T_{2/3}$ top partner

- BDT discriminant for various channels



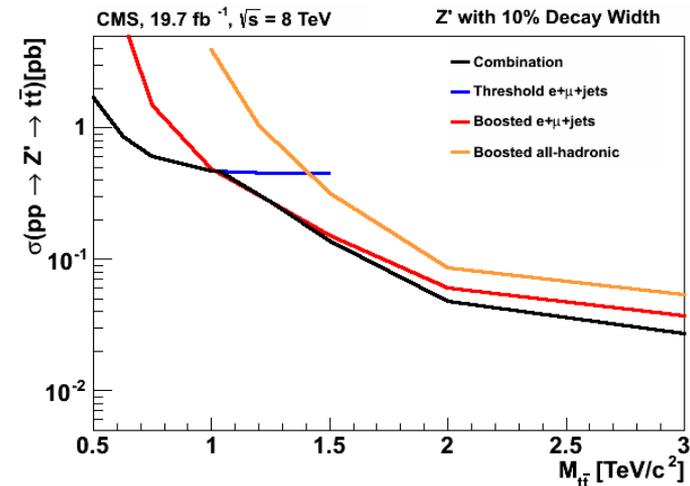
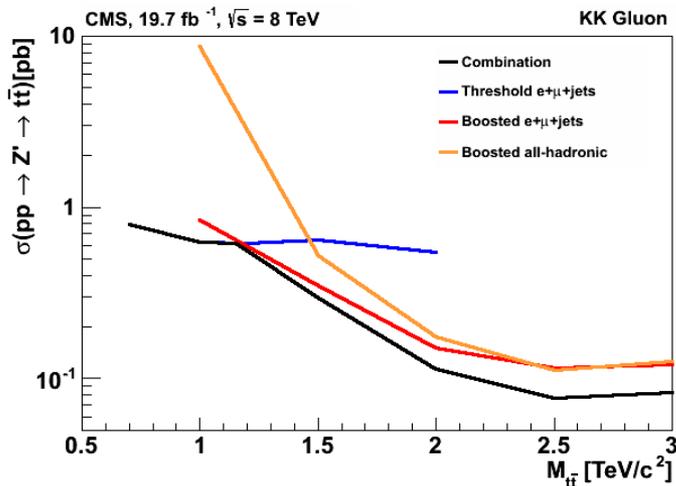
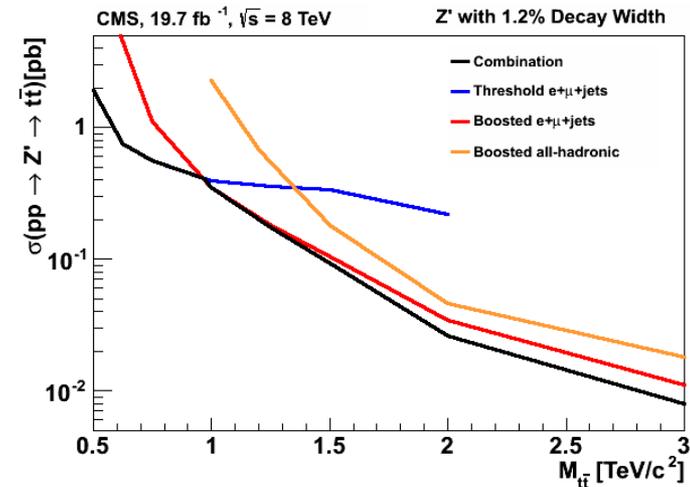
Search $T_{2/3}$ top partner

- BDT discriminant for various channels
 - Zero b-tags



Search for BSM $t\bar{t}$ resonances

- Limit comparison for various signal hypotheses



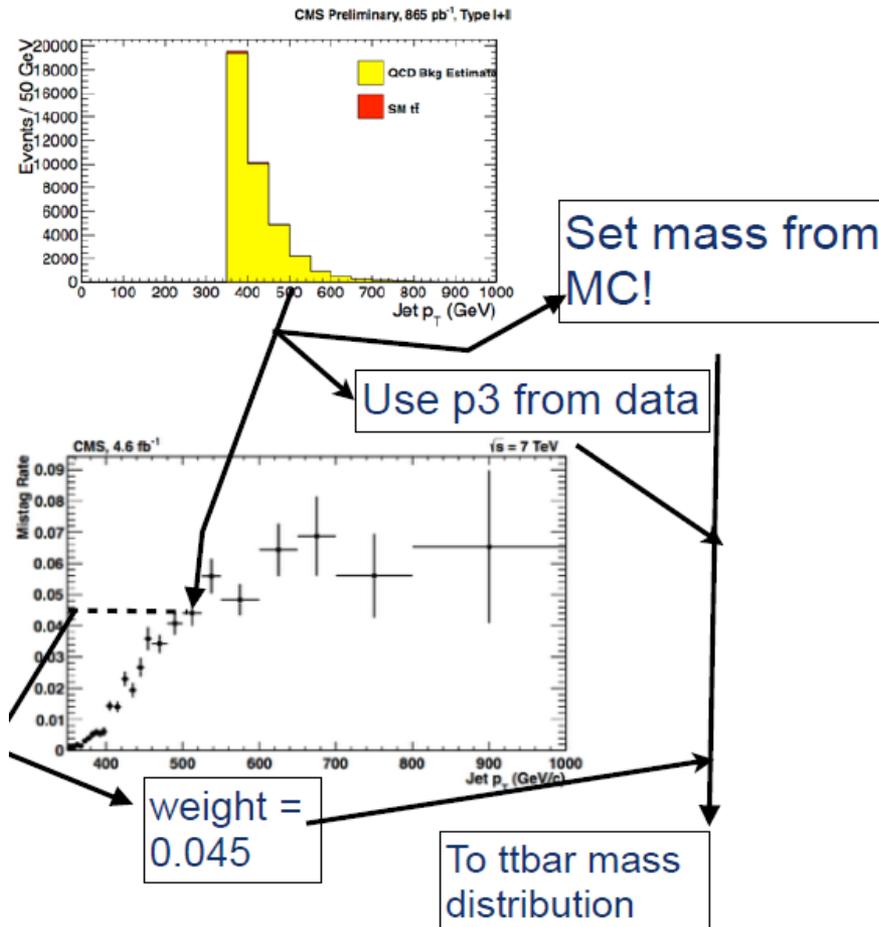
Search for BSM $t\bar{t}$ resonances

- Signal exclusion

Model	Observed Limit	Expected Limit
$Z', \Gamma_{Z'} / M_{Z'} = 1.2\%$	2.1 TeV	2.1 TeV
$Z', \Gamma_{Z'} / M_{Z'} = 10\%$	2.7 TeV	2.6 TeV
RS KK gluon	2.5 TeV	2.4 TeV

Search for BSM $t\bar{t}$ resonances

- Top-mistagging rate
 - 7 TeV



N-subjettiness

- Variables τ_N describe how consistent the jet energy is with having N subjets
 - Cut on τ_3/τ_2

$$\tau_N = \frac{1}{d_0} \sum_k p_{T,k} \min \{ \Delta R_{1,k}, \Delta R_{2,k}, \dots, \Delta R_{N,k} \}$$

Top-Tagging Scale Factor

- Use simulation for $t\bar{t}$ and Signal
- Need to extract Monte Carlo to data scale factor for top-tagging.
- We investigate this using a highly pure sample of semileptonic $t\bar{t}$
 - Documented in JME-13-007

Top-Tagging Scale Factor

Plots from
JME-13-007

