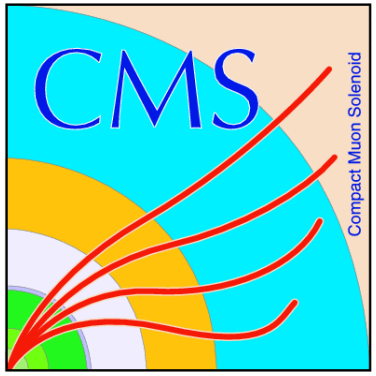


First observation and measurement of single top production in the tW channel in pp collisions

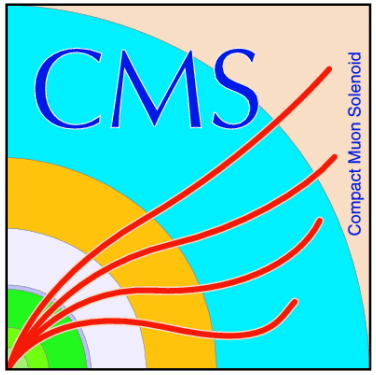
KU

Phil Baringer, University of Kansas
on behalf of the CMS collaboration

DPF 2013, August 13-17, University of California, Santa Cruz

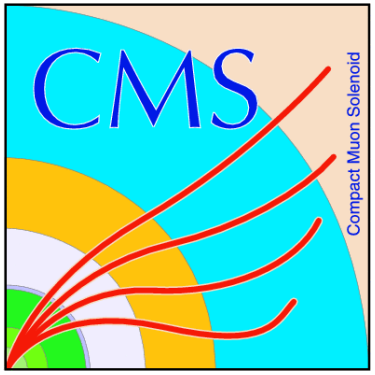
Outline

1. Introduction
2. The CMS detector
3. Multivariate analysis
4. Cross check analyses
5. Summary

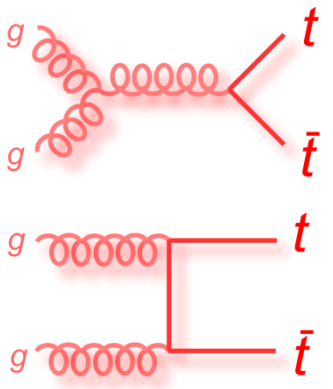


KU

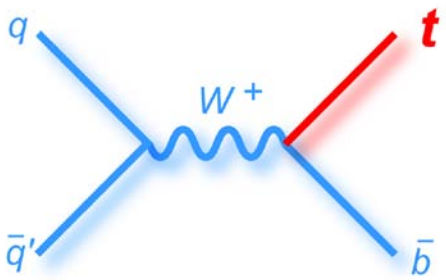
Introduction



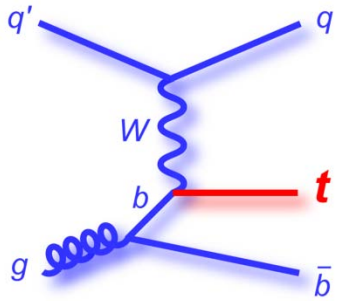
Electroweak production of the top quark



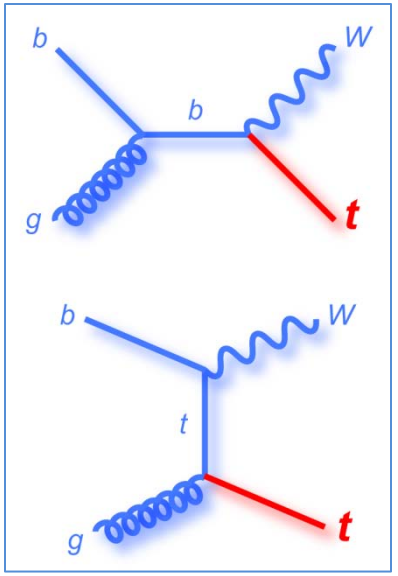
- At hadron colliders, top quarks are most commonly produced in pairs via the strong force
- Single top quarks can be produced via electroweak interactions



s-channel



t-channel



tW associated production

Single top physics

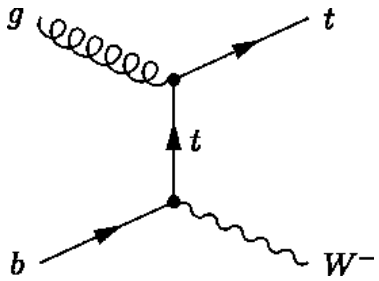
- Test of Standard Model; study W-t-b vertex
- Sensitive to new physics, with each channel sensitive to different non-Standard Model processes (like FCNC, extra generations)
- Cross sections (for top mass of 173 GeV):

N. Kidonakis arxiv.org/pdf/1205.3453v1 (2012); arxiv.org/pdf/0909.0037

***M. Czakon, P. Fiedler and A. Mitov arXiv.org/pdf/1303.6254**

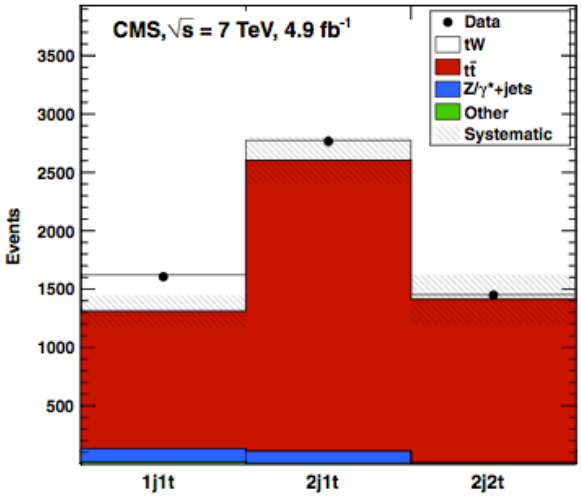
σ(pb)	t-channel	s-channel	tW	t-tbar
Tevatron (1.96 TeV)	2.08	1.05	0.22	7.08
LHC (7 TeV)	65.9	4.56	15.6	163
LHC (8 TeV)	87.2	5.55	22.2	245*

Prior tW results



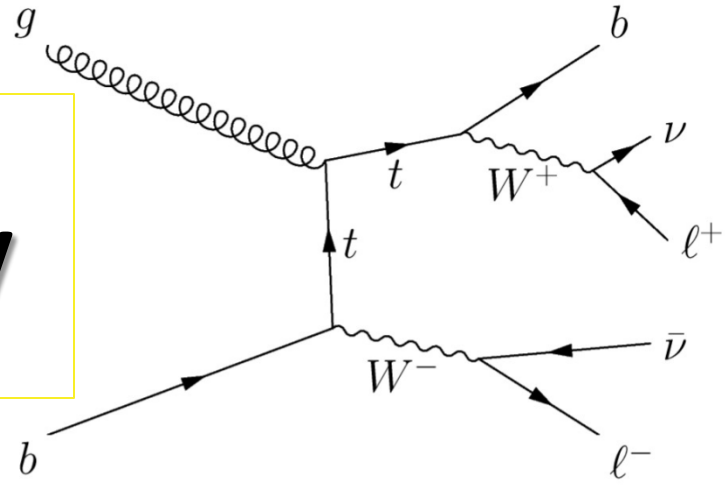
- Channel not accessible at the Tevatron
- Evidence for tW from CMS and ATLAS in the 2011, 7 TeV, LHC data sample

CMS: "Evidence for associated production of a single top quark and W boson in pp collisions at 7 TeV" Phys. Rev. Lett. 110, 022003 (2013) – 4.0 σ significance, corresponding to a cross section of 16 +5 -4 pb



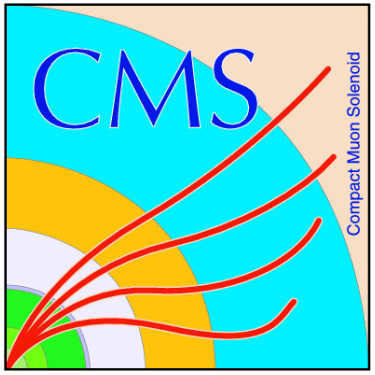
ATLAS: "Evidence for the associated production of a W boson and a top quark in ATLAS at √s=7 TeV," Phys.Lett. B716 (2012) 142 – 3.3 σ significance, corresponding to a cross section of 16.8±2.9±4.9 pb

tW Selection Strategy



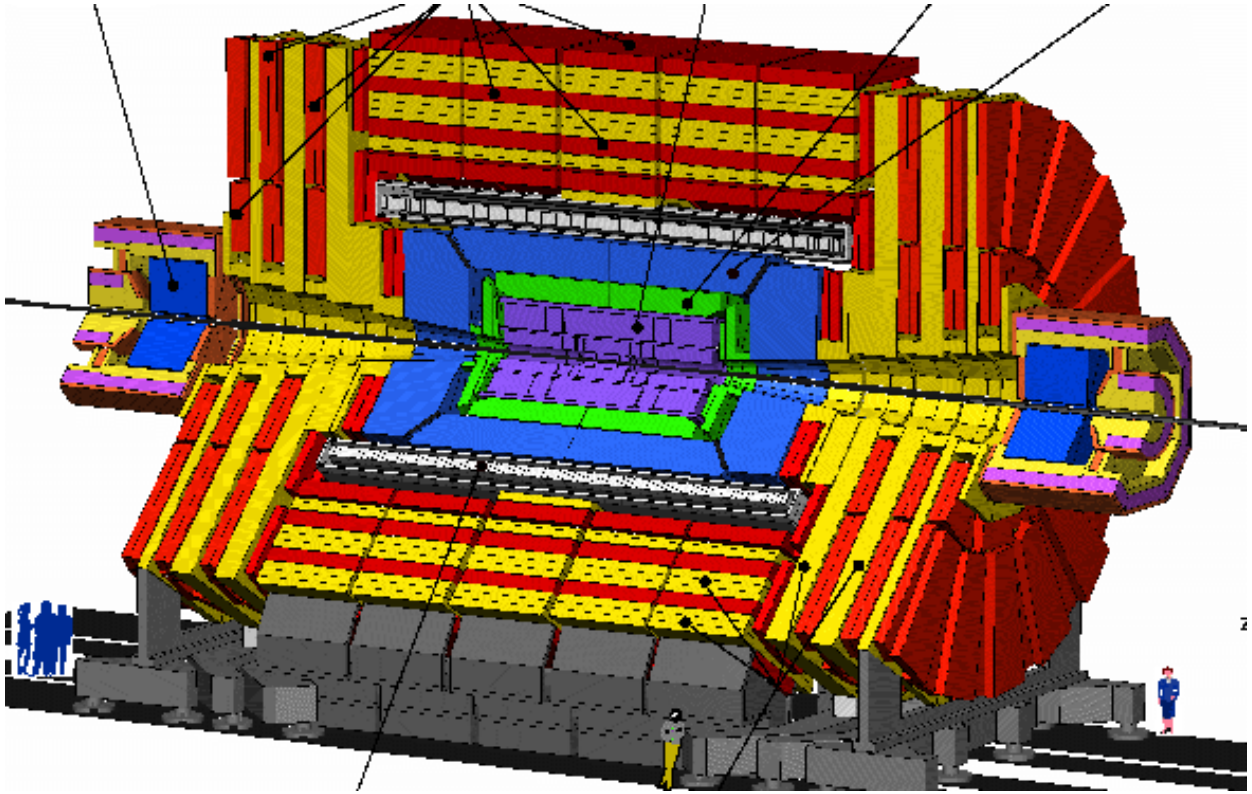
- Look for events where each W decays leptonically
- Charged lepton = electron or muon
- Final state has:
  - Two oppositely charged leptons
  - One jet from the *b* quark
  - Missing energy from the neutrinos

The CMS Detector



KU

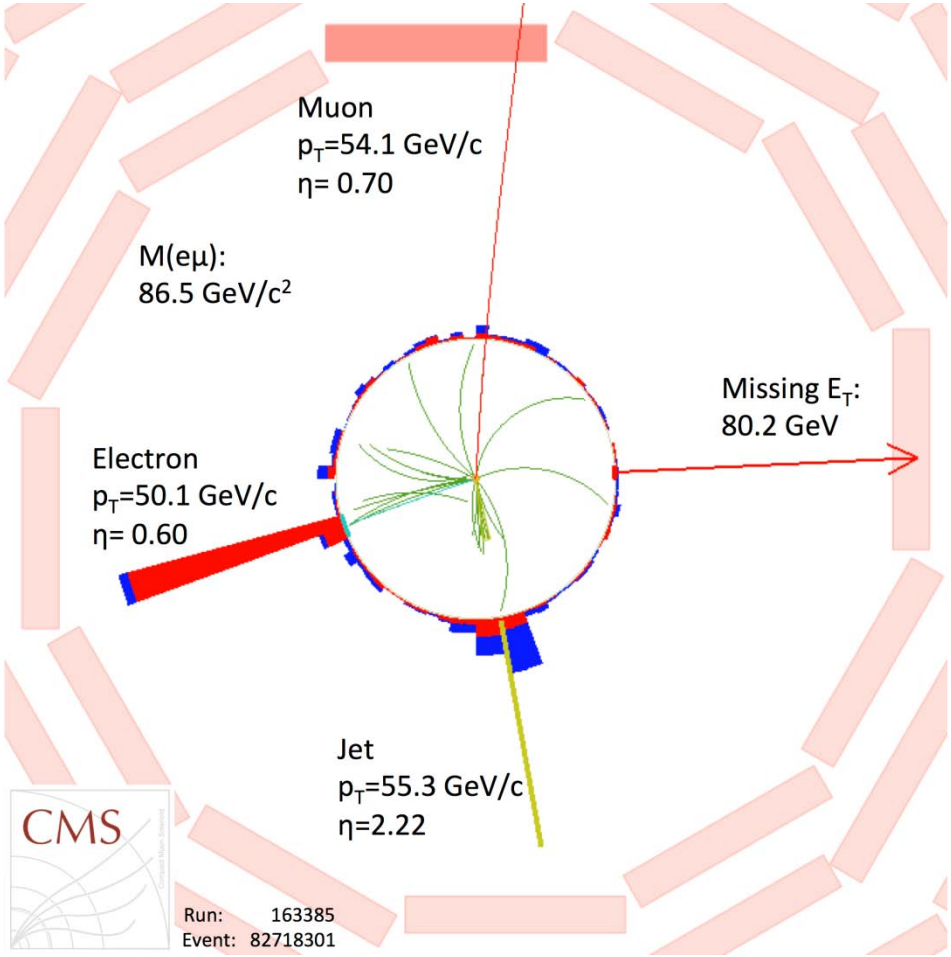
CMS-Compact Muon Solenoid



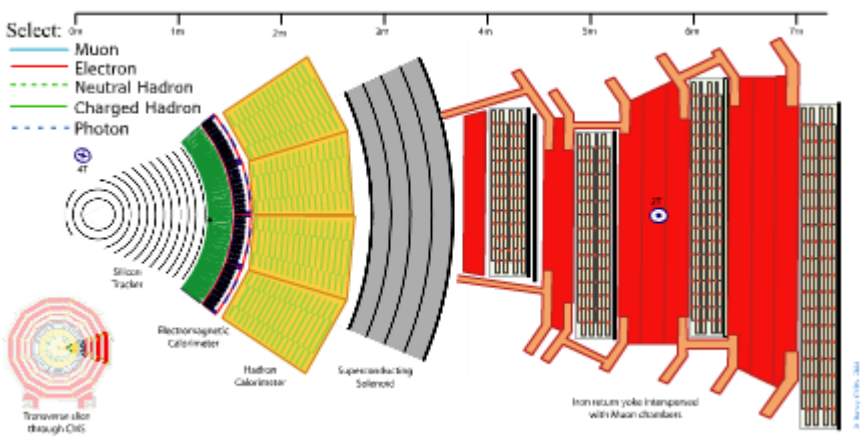
Physical Parameters:
12,500 tons
21 m long
15 m diameter

Over 100,000,000 individual detecting elements

Candidate tW Event from 7 TeV data

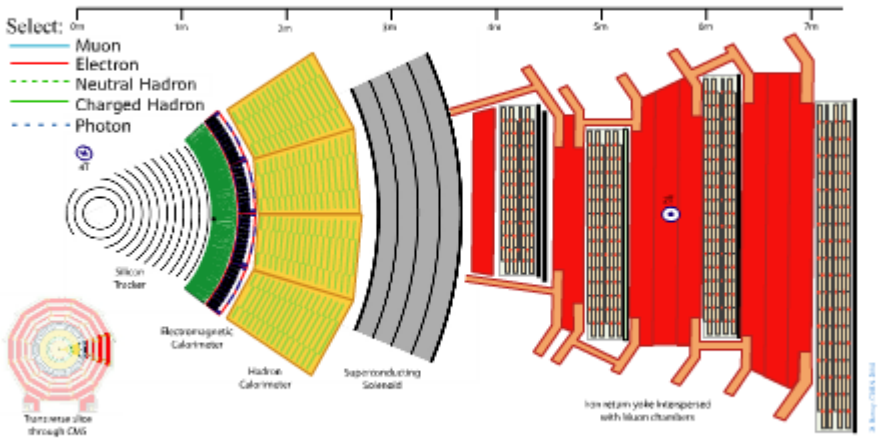


Electron Selection



- Electrons—energy deposits in EM calorimeter matched to hits in tracker
- Isolated
- coming from primary vertex
- p_T>20 GeV
- $|\eta|$<2.5
- "loose electrons": p_T>10 GeV, $|\eta|$<2.5

Muon Selection



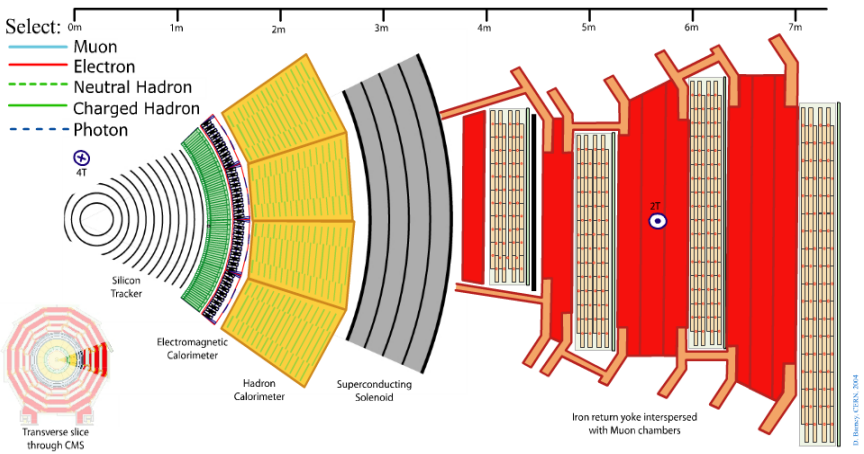
- Muons—reconstructed using particle flow algorithm
- Reconstructed in both tracker and muon system
- Isolated
- p_T>20 GeV
- $|\eta|$<2.4
- "loose muon": p_T>10 GeV, $|\eta|$<2.5

Jet Selection

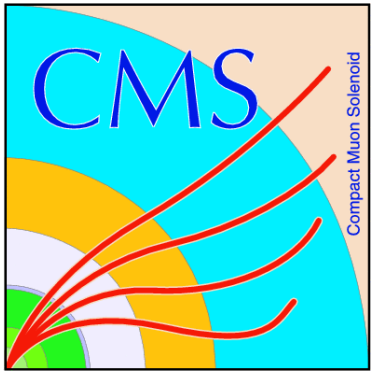
- Jets– particle flow, anti-k_T algorithm with resolution parameter size 0.5

M. Cacciari, G. P. Salam, and G. Soyez, "The Anti-k(t) jet clustering algorithm", JHEP **0804 (2008) 063, doi:10.1088/1126-6708/2008/04/063, arXiv:0802.1189.**

- Jet energy corrections applied
- "Tight jet":
  - Corrected p_T>30 GeV
  - $|\eta|$<2.4
- "Loose jet":
  - fails "tight" cuts
  - p_T>20 GeV
  - $|\eta|$<4.9

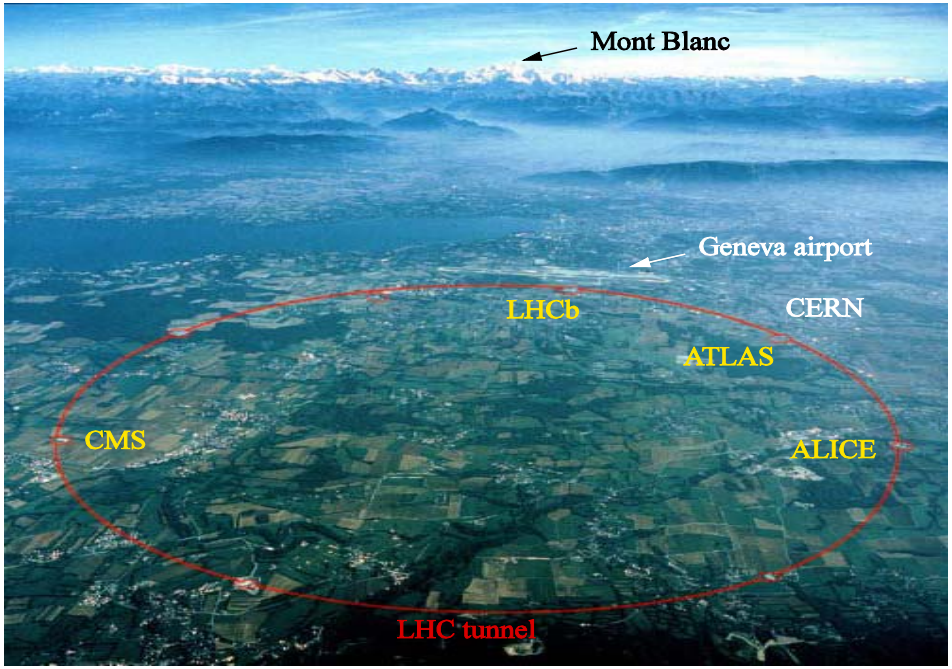


Multivariate Analysis



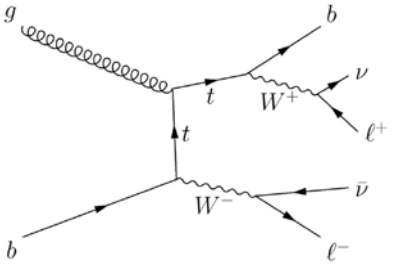
KU

Data Selection



- pp collisions at $\sqrt{s}$ =8 TeV at the Large Hadron Collider
- Integrated luminosity corresponding to 12.2 fb^{-1}
- Triggers requiring two leptons (e or μ), one with p_T>17 GeV, the second with p_T>8 GeV

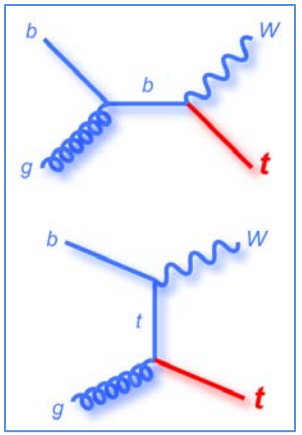
Event selection

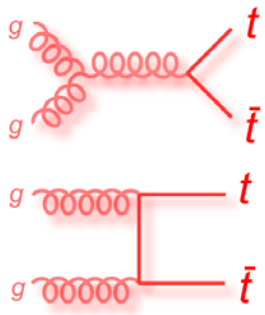


- Well-reconstructed primary vertex (at least 4 tracks, $|z|<24$ cm, $\rho<2.0$ cm)
- Two oppositely charged leptons
- No additional loose leptons
- Invariant mass of two leptons (m_{ll})>20 GeV (to remove low mass Z^*/γ events)
- In *ee* and $\mu\mu$ events:
  - Remove events with $81<m_{ll}<101$ GeV (to reduce *Z*+jets, *ZZ* and *WZ* backgrounds)
  - Require transverse missing energy of at least 50 GeV

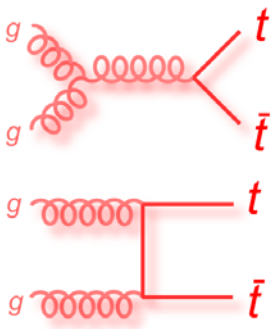
Signal and Control Regions

- Signal region has exactly one tight jet that is *b*-tagged (1j1t)
  - *b*-tagging done with a multivariate algorithm using tracking information
- Control regions (dominated by $t\bar{t}$)
  - Exactly two tight jets, one of the *b*-tagged (2j1t)
  - Exactly two tight jets, both of them *b*-tagged (2j2t)





Backgrounds



- Dominant background is $t\bar{t}$ (~75%)
  - Boosted decision tree (BDT) is trained to distinguish tW from $t\bar{t}$
  - Shapes of BDT discriminant plots in signal and control regions used to determine cross section and significance
- Next largest background is *Z*+jets (~5%)
  - Remaining events after cuts estimated from simulation
  - Corrections made to simulation using *Z*+jets enriched data sample

Boosted Decision Tree (BDT)

- TMVA package used
- Trained on Monte Carlo: 200k dilepton events from tW and 200k top pair events
- POWHEG generator used for signal, MADGRAPH for top pair background
- Full detector response simulated with GEANT4
- 13 kinematic variables used to discriminate signal and background

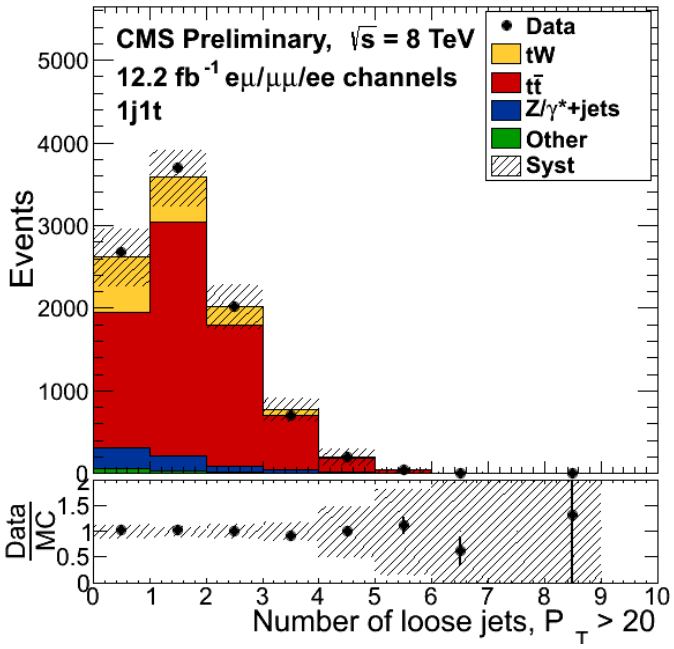
Kinematic variables for BDT

- Data/MC agreement checked in several control regions (2j1t, 2j2t, 2j0t, 1j0t)

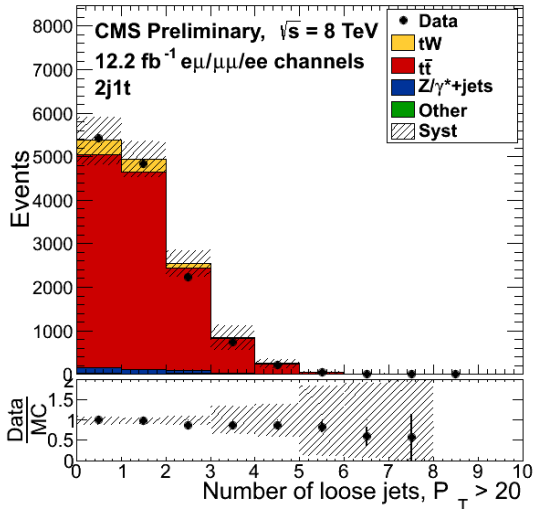
Variable	Description
Nloosejets	Number of loose jets, $p_T > 20$ GeV, $\|\eta\| < 4.9$
NloosejetsCentral	Number of loose jets, $p_T > 20$ GeV, $\|\eta\| < 2.4$
NbtaggedLoosejets	Number of loose jets, $p_T > 20$ GeV, CSVM btagged
$p_{T,sys}$	Vector sum of p_T of leptons, jet, and E_T^{miss}
H_T	Scalar sum of p_T of leptons, jet, and E_T^{miss}
Jet p_T	p_T of the leading, tight, b-tagged jet
Loose jet p_T	p_T of leading loose jet, defined as 0 for events with no loose jet present
$p_{T,sys}/H_T$	Ratio of $p_{T,sys}$ to H_T for the event
Msys	Invariant mass of the combination of the leptons, jet, and E_T^{miss}
centralityJLL	Centrality of jet and leptons
$H_{T,leptons}/H_T$	Ratio of scalar sum of p_T of the leptons to the H_T of full system
p_T-jll	Vector sum of p_T of jet and leptons
E_T^{miss}	Missing transverse energy in the event

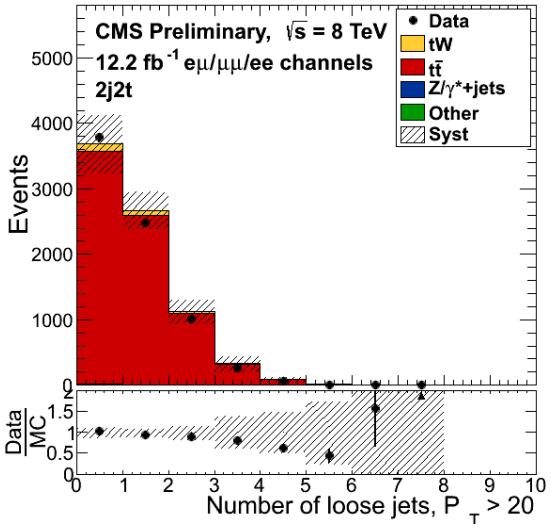
Number of loose jets

- "Loose jet" p_T>20 GeV, $|\eta|$<4.9, fails "tight" cuts



Signal region

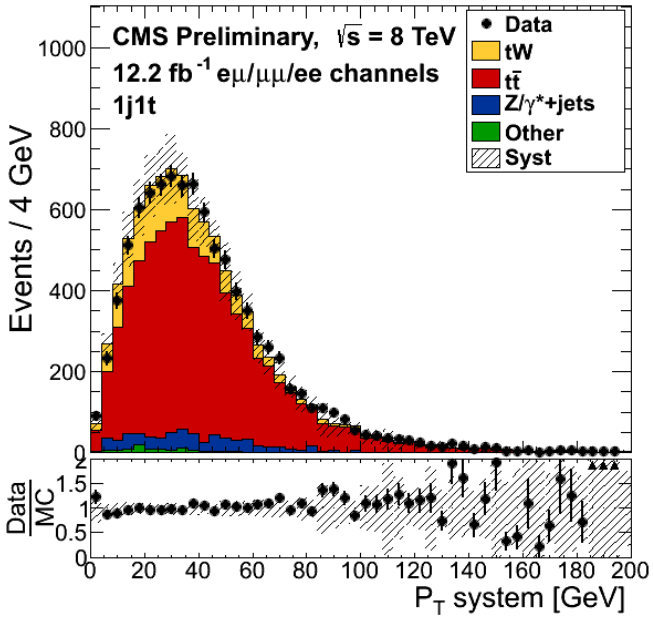




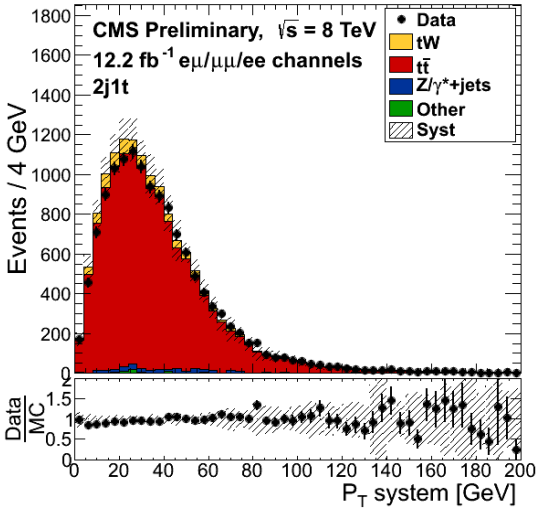
Control regions

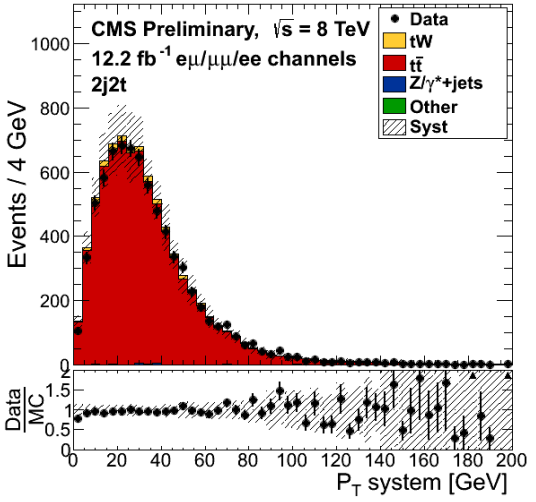
p_T of the system

- Vector sum of p_T's of leptons, jet and missing E_T



Signal region

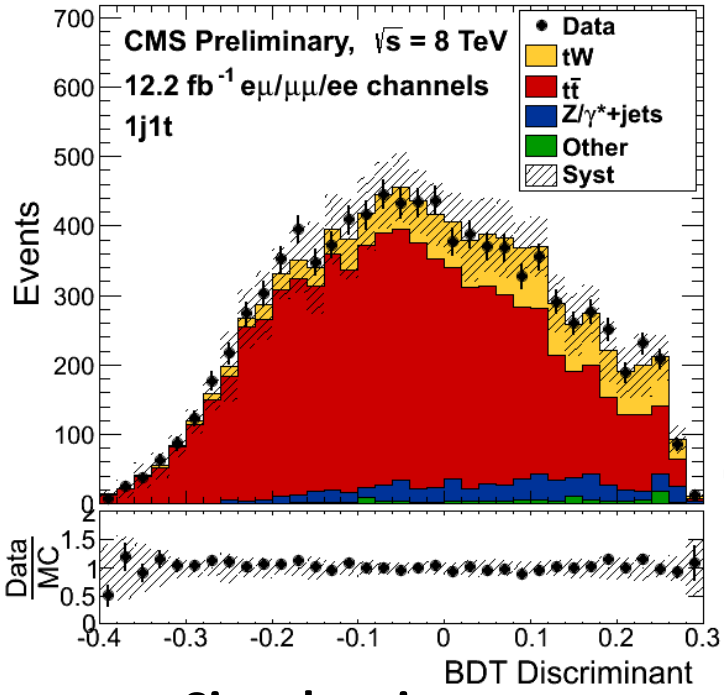




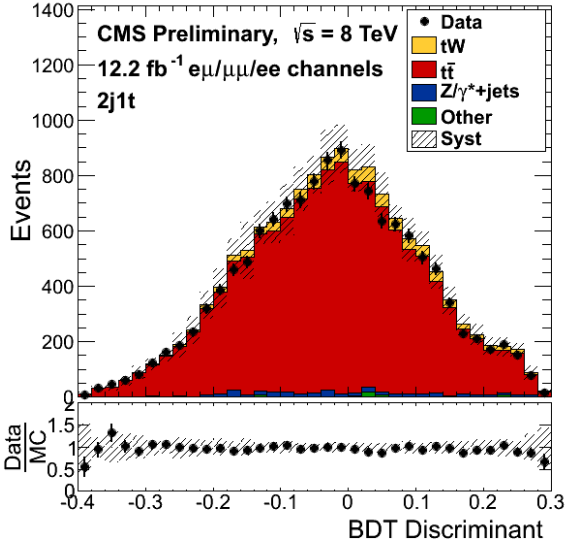
Control regions

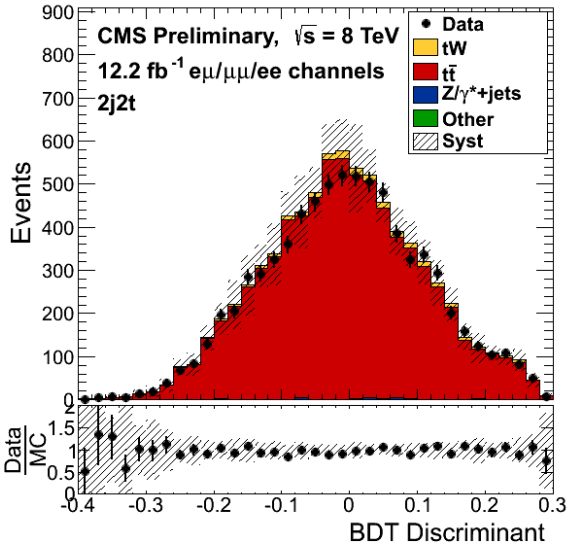
BDT discriminant

- Signal-like events positive, background-like negative
- Binned likelihood fit done simultaneously for all channels (*ee*, *eμ*, *μμ*), all regions (1j1t, 2j1t, 2j2t)



Signal region





Control regions

Systematics affecting rate only

- Luminosity --4.4% uncertainty on CMS measurement
- Lepton efficiency -- 1.9-2.3% uncertainty from varying trigger, reconstruction and identification efficiencies
- $t\bar{t}$ cross section -- 6.8% uncertainty in CMS measured value

Systematics giving shape variations

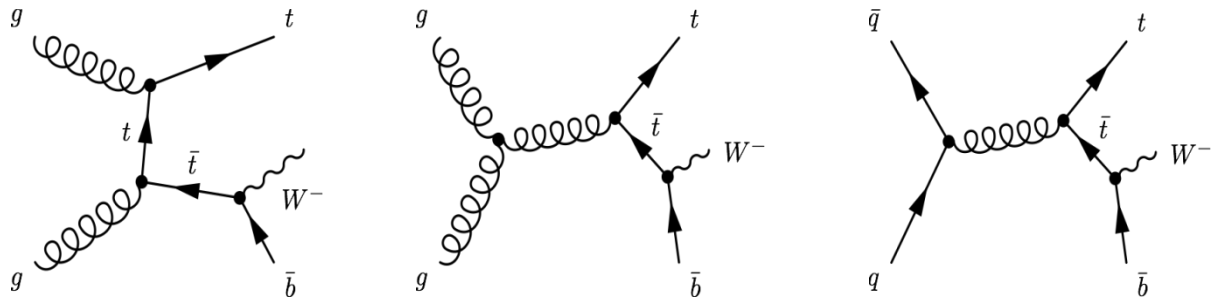
- pile-up multiplicity
- jet energy scale
- jet energy resolution
- b-tagging data/MC scale factors
- missing energy modeling
- *Z* + jets scale factors
- PDF uncertainties
- statistics of simulated data
- theory uncertainties (see next two slides)

Dominant systematics

Table shows systematic uncertainties extracted by fixing sources one at a time and measuring the difference in the cross section uncertainty

Systematic Uncertainty	$\Delta\sigma$ (pb)	$\frac{\Delta\sigma}{\sigma}$
ME/PS matching thresholds	3.25	14%
Q^2 scale	2.68	11%
Top quark mass	2.28	10%
Statistical	2.13	9%
Luminosity	1.13	5%
JES	0.91	4%
$t\bar{t}$ cross section	0.87	4%
Z+jet data/MC scale factor	0.56	2%
tW DR/DS scheme	0.45	2%
PDF	0.33	1%
Lepton identification	0.31	1%
JER	0.27	1%
B-tagging data/MC scale factor	0.20	$< 1\%$
$t\bar{t}$ Spin Correlations	0.12	$< 1\%$
Top Pt Reweighting	0.12	$< 1\%$
Event pile up	0.11	$< 1\%$
$E_{\mathrm{T}}^{\mathrm{miss}}$ modeling	0.07	$< 1\%$
Lepton energy scale	0.02	$< 1\%$
Total	5.58	24%

Theory uncertainties



tW mixes with top pair production at NLO. The above diagrams are removed from the signal definition when doubly resonant

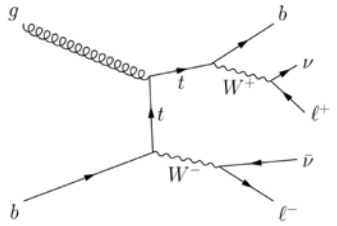
These uncertainties are externalized in the significance calculation, which gives a more conservative error determination

- Largest systematic uncertainties come from theory uncertainties:
  - Jet-parton matching thresholds in top pair simulations
  - Q^2 scale
  - Top mass uncertainty
  - DR (diagram removal)/DS (diagram subtraction) scheme for separating higher order top pair and tW diagrams
  - top pair spin correlations
  - top p_T reweighting

Statistical analysis

- Simultaneous binned likelihood fit of BDT distribution for all channels (*ee*, *eμ*, *μμ*), all regions (1j1t, 2j1t, 2j2t)
- Expected yield in bin *i*: $\lambda_i = \mu S_i + \sum_k B_{k,i}$
- Templates for signal and background taken from Monte Carlo
- Nuisance parameter , θ, introduced for each independent source of systematic uncertainty that changes the template
- Theory-based values fixed at central value in likelihood fit, uncertainties included in pseudo-experiments
- Test statistic for pseudo-experiments: $q_0 = \frac{\delta}{\delta\mu} L\left(\mu = 0, \hat{\theta}_0 \mid \text{data}\right)$, where μ is signal strength
- Evaluate for background-only and signal + background hypotheses
- Profile likelihood fit to get cross section and 68% confidence level, signal and background rates allowed to float

Results

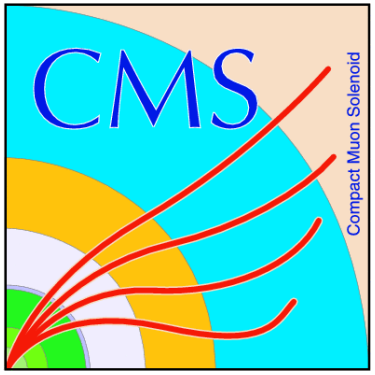


- An excess of events is observed compared to a background-only hypothesis based on fitting the shape of the BDT discriminant
- Observed significance=6.0σ
- Expected significance from MC=5.4σ
- Measured tW cross section: 23.4 +5.5-5.4 pb
- Standard model: 22.2 ± 0.6 (scale) ± 1.4 (PDF) pb
  N. Kidonakis arxiv.org/pdf/1205.3453v1 (2012)
- CKM $|V_{tb}|$ matrix element (assume $|V_{tb}| >> |V_{td}|, |V_{ts}|$):

  $|V_{tb}| = 1.03 \pm 0.12(\text{exp}) \pm 0.04(\text{th})$

  $|V_{tb}| > 0.78$ at 95% C.L., when constrained to be ≤ 1

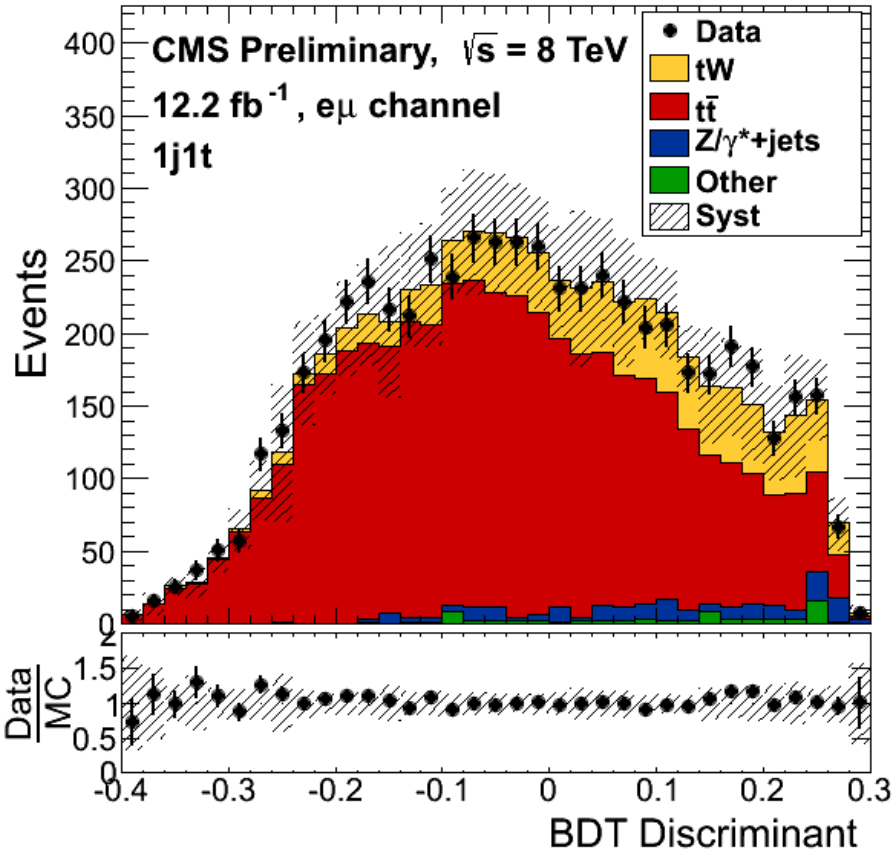
Cross check analyses



KU

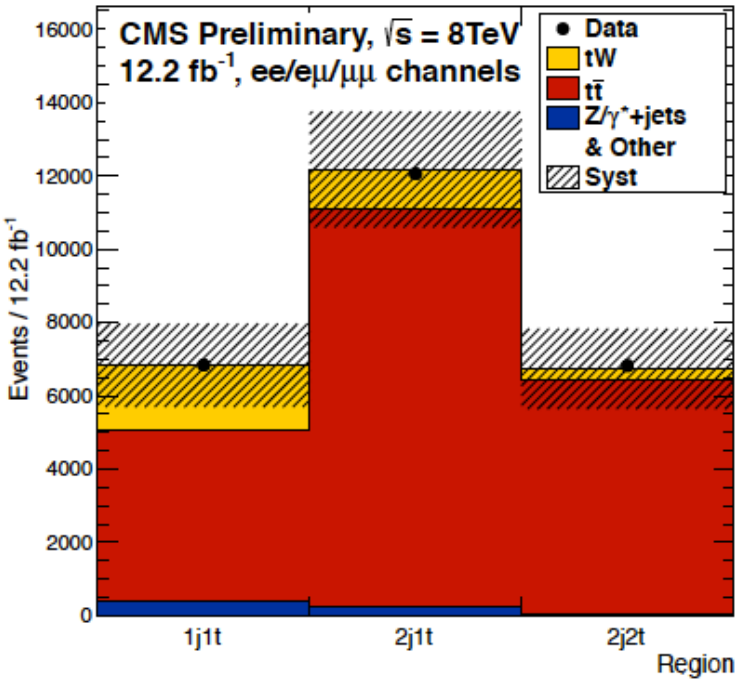
Lepton channel checks

- $e\mu$, ee and $\mu\mu$ channels give consistent results
- $e\mu$ channel alone has:
  - Observed significance=5.7σ
  - Expected significance from MC=4.4σ
  - Measured tW cross section: 29.0 +6.2-6.1 pb



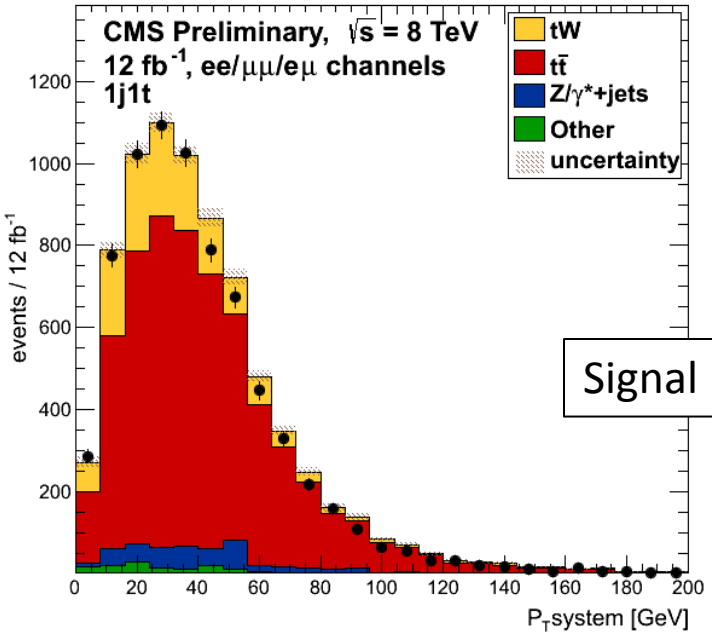
Cut and count

- Independent cut-and-count analysis gives observed significance of 3.6 σ
- Additional cuts relative to the BDT analysis:
  - Veto events with loose b-tagged jets
  - Ht>160GeV (*eμ* final state only)
- Fit to event counts only in each region

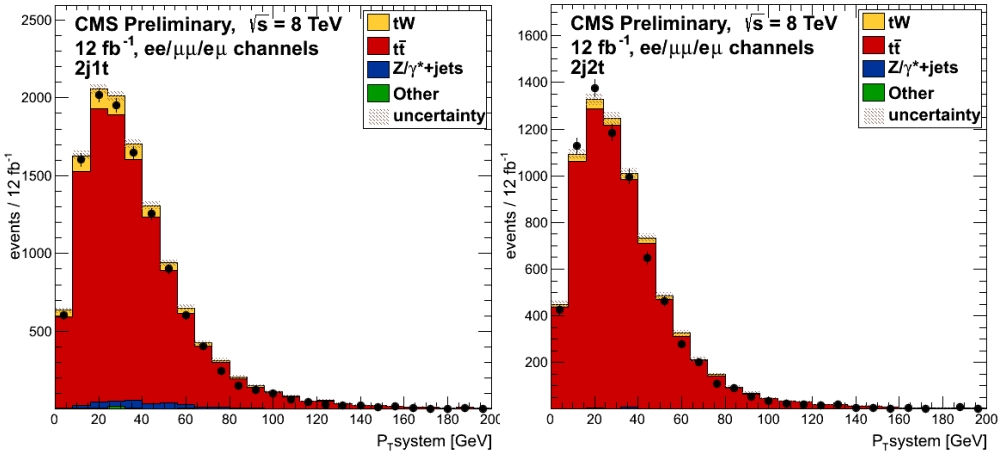


p_T system Fit

- Same selection as the cut-and-count
- Instead of fitting numbers of events, fit the p_T of system distribution (vector sum of p_T's of leptons, jet and missing E_T)
- gives observed significance of 4.0 σ

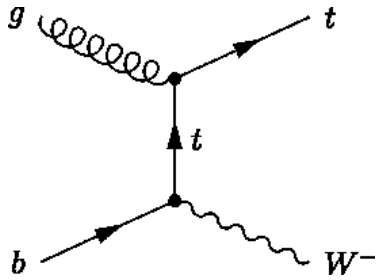


Signal region



Background regions

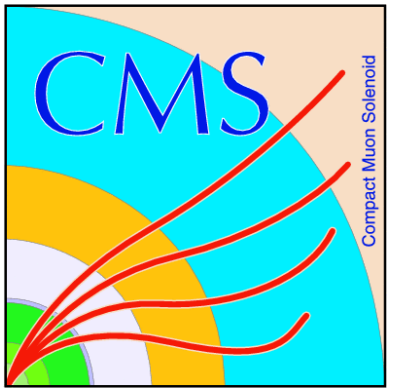
Summary



- single top tW associated production, has been observed in the dilepton channel at > 5.0 σ significance
- CMS used 12.2 fb^{-1} of pp collisions data at 8 TeV in this analysis
- Multivariate analysis using kinematic variables in a boosted decision tree (BDT) used to separate tW signal from top pair background
- Binned likelihood fit to the BDT discriminant used to measure significance and cross section
- Signal region (1j1t) and control regions (2j1t, 2j2t) used in fit
- Excess of events above background-only hypothesis is **6.0 σ**
- Measured tW cross section: **23.4 +5.5-5.4 pb**

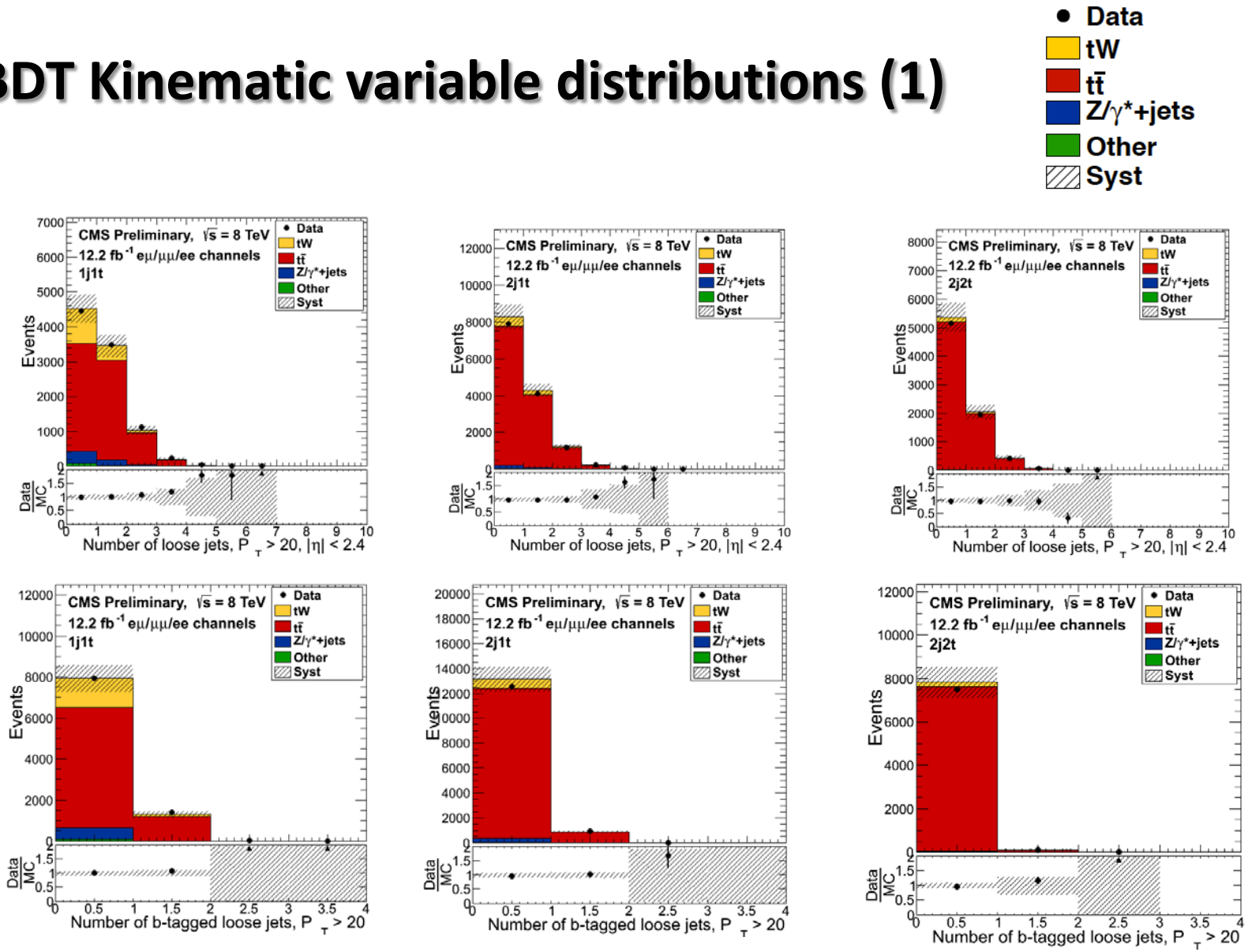
CMS public results links

- https://twiki.cern.ch/twiki/bin/view/CMSPublic/PhysicsResultsTOP12040
- http://cds.cern.ch/record/1563135

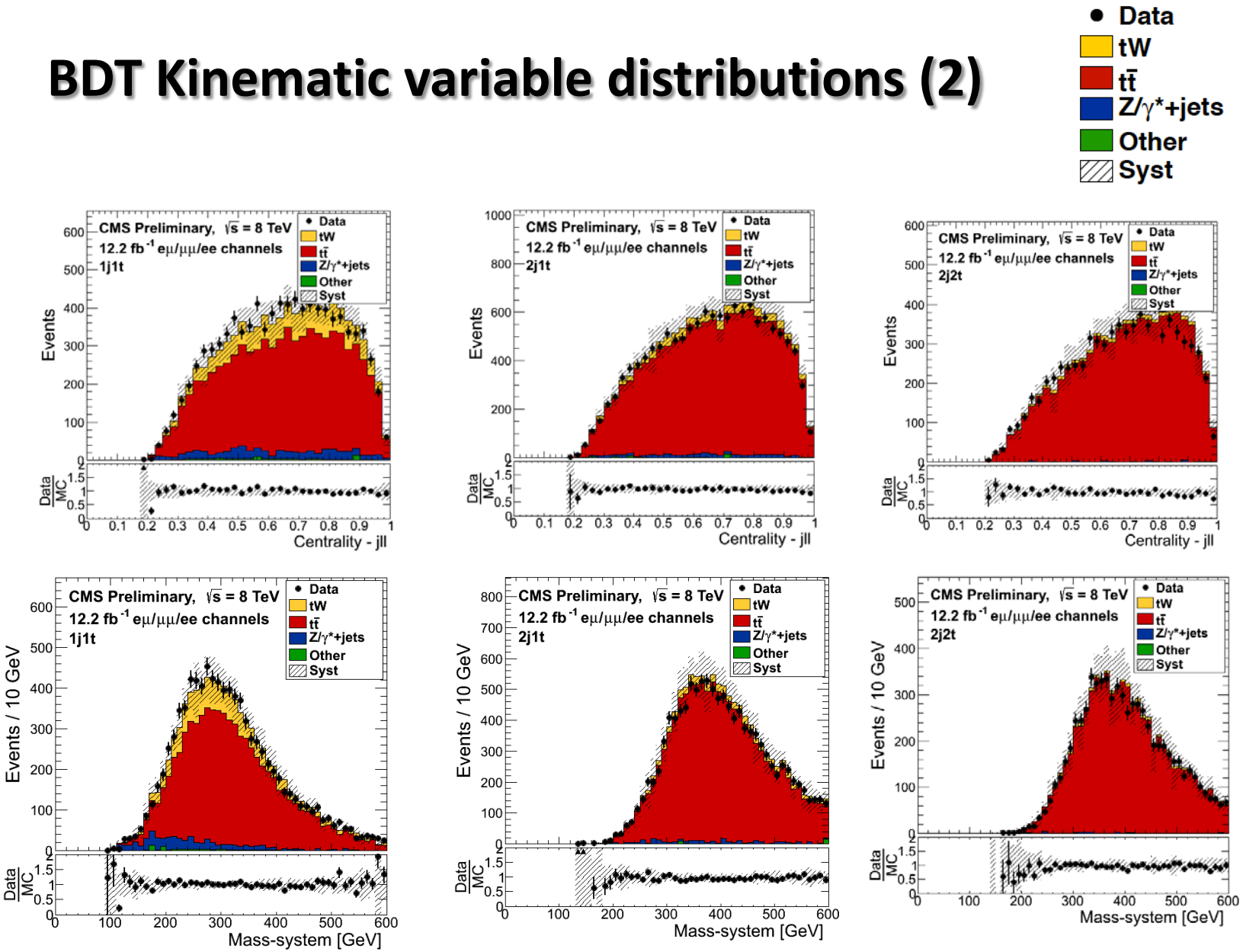


Backup

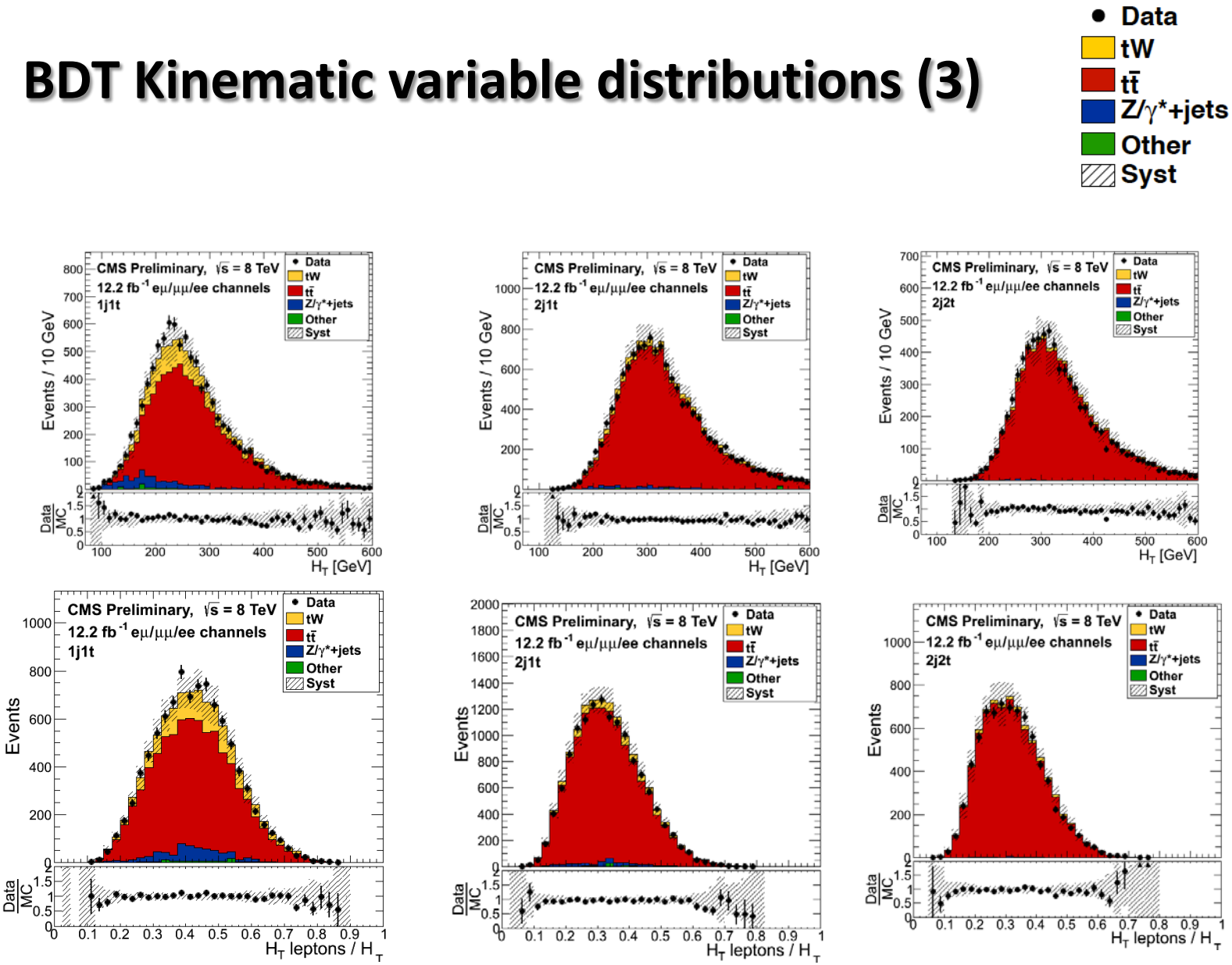
BDT Kinematic variable distributions (1)



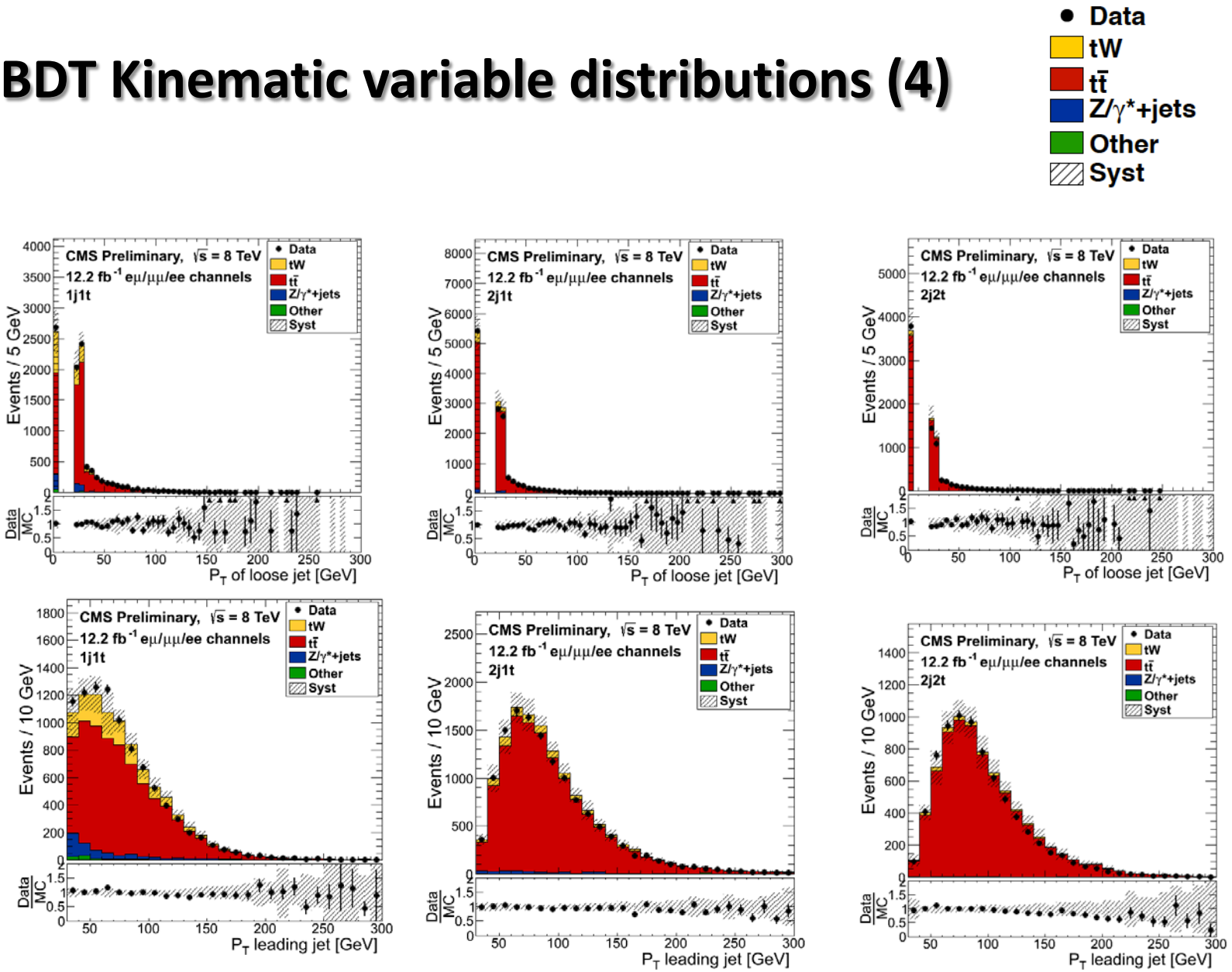
BDT Kinematic variable distributions (2)



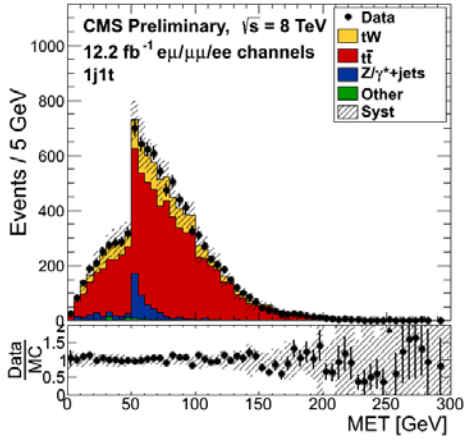
BDT Kinematic variable distributions (3)

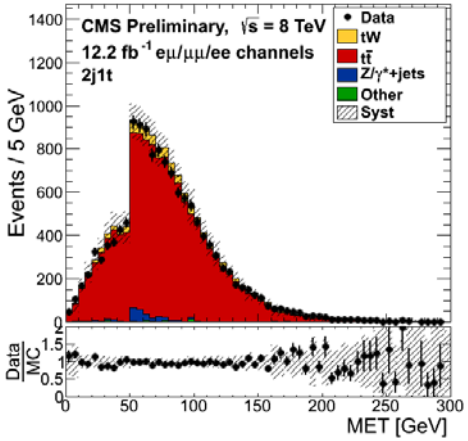


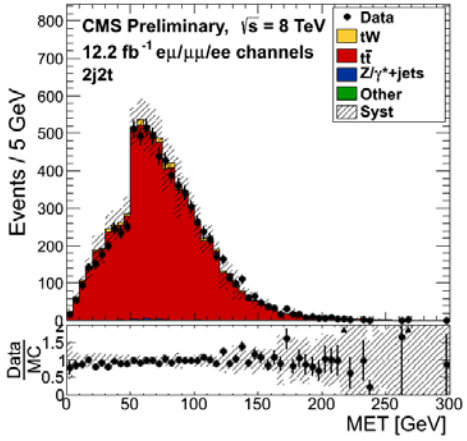
BDT Kinematic variable distributions (4)

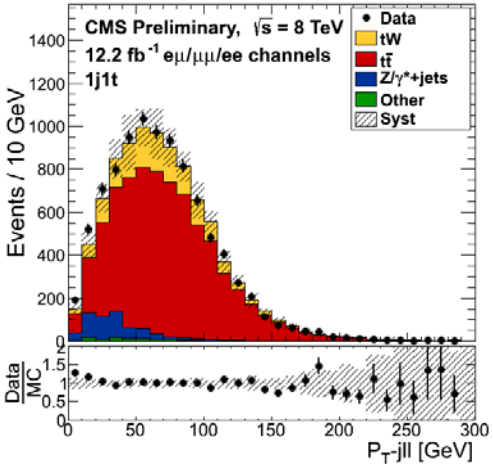


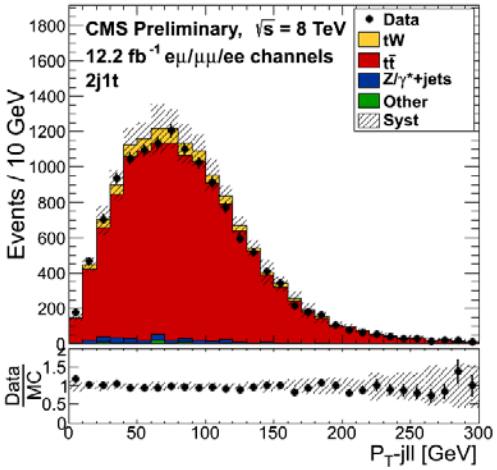
BDT Kinematic variable distributions (5)

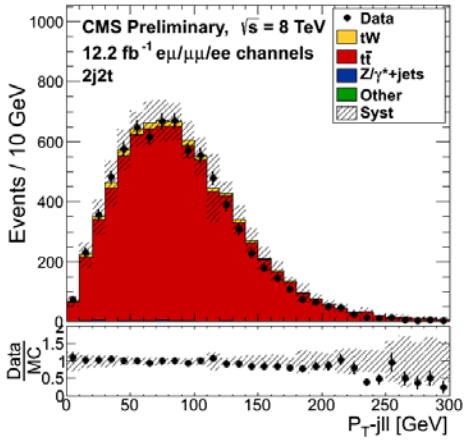




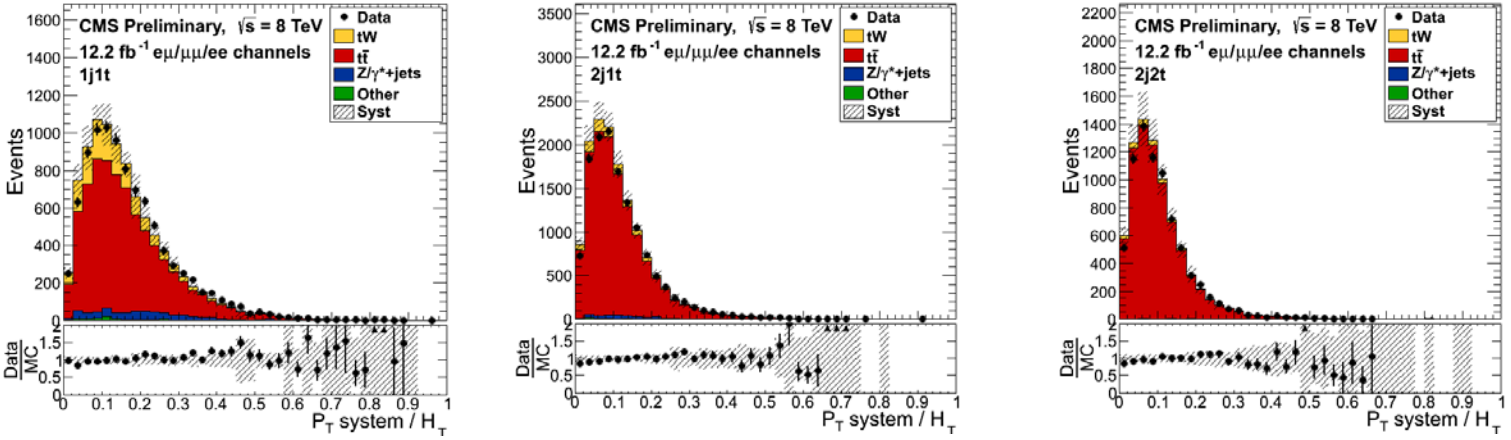






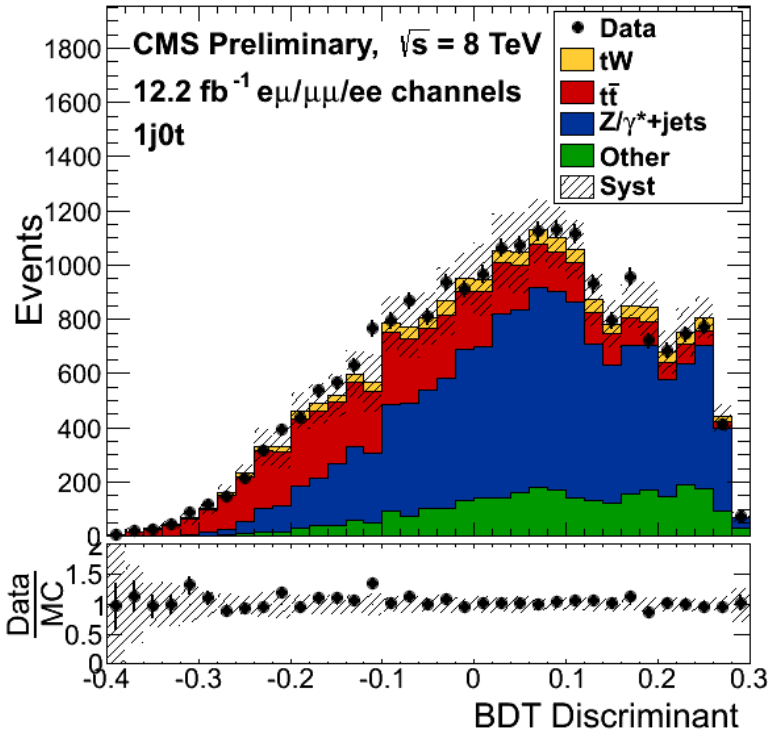


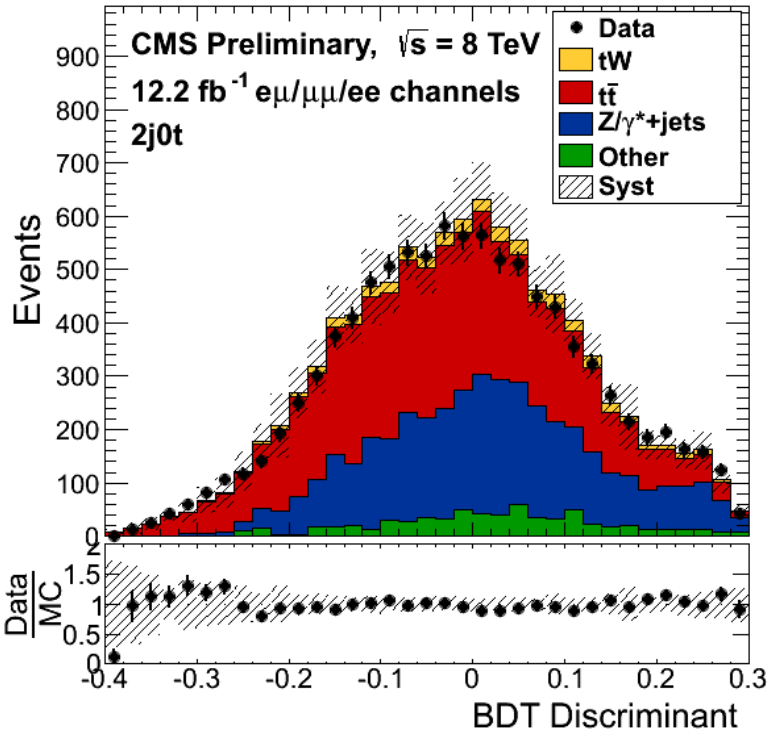
BDT Kinematic variable distributions (6)



Zero tag control regions

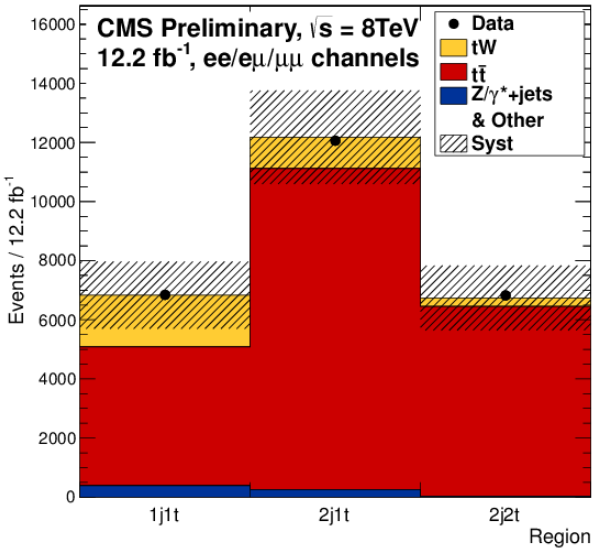
- BDT discriminant for events with no *b*-tagged jets



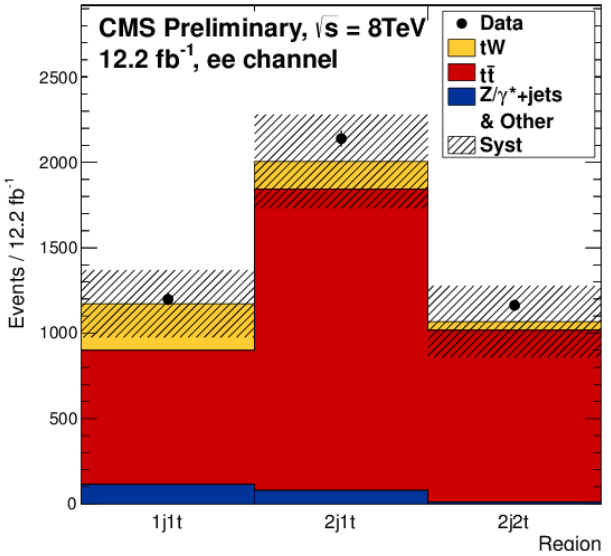


Event yields in signal and control regions

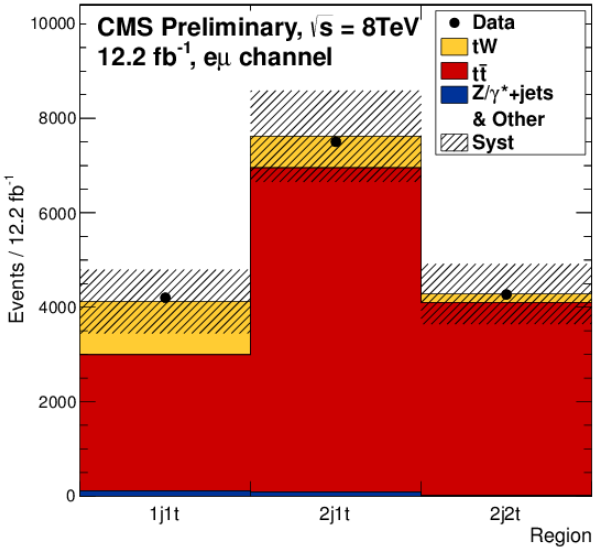
all



ee



eμ



μμ