Measurement of the single top production with the CMS detector

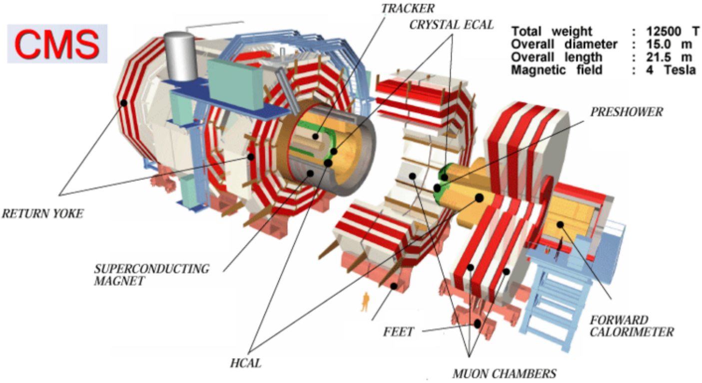
Andrey Popov
On behalf of the CMS collaboration

QFTHEP 2011, Sochi, Russia
26.09.2011

Outline

- Introduction
  - CMS detector
  - Physical motivation for single top studies
  - Single top processes
- Single top t- & tW-channels
  - General description of the process
  - Event selection
  - Data-driven estimations
  - Analysis procedure
  - Systematics
  - Results
- Conclusions

CMS detector

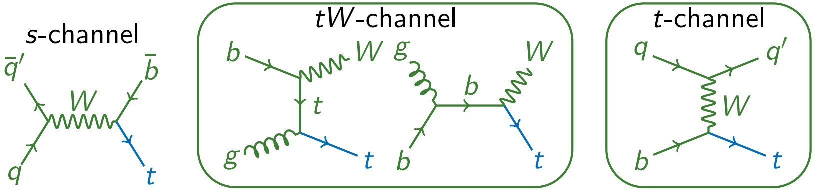


Why is single top interesting?

Single top probes the W-t-b interaction and can be sensitive to

- additional non-standard particles
  - extra quarks
  - extra gauge bosons
  - additional scalar bosons
- modified t-quark interactions
  - FCNC contribution
  - right-handed interactions

Single top processes

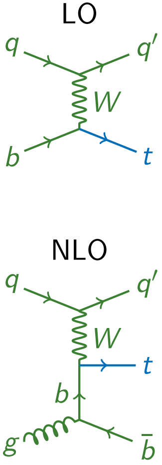


Different rates and different kinematical regions

x-sec, pb $m_t = 173$ GeV	s-channel	tW-channel	t-channel	$t\bar{t}$
TEVATRON, $p\bar{p}$ @ 1.96 TeV	1.04	0.22	2.08	7.2
LHC, pp @ 7 TeV	4.59	15.6	63.2	165

N. Kidonakis, Phys. Rev. D **81**, 054028 (2010), Phys. Rev. D **82**, 054018 (2010), Phys. Rev. D **83**, 091503(R) (2011), arXiv:0909.0037v1

Single top t-channel



- Analysis focused on leptonic decays of W: $t \to Wb \to l\nu b$
- Signal signature:
  - exactly one isolated lepton,
  - forward light-quark jet,
  - central b-jet from top decay,
  - additional b-jet with small p_T is outside acceptance
- Main backgrounds:
  - $t\bar{t}$: similar kinematics, high rate
  - $W(\to l\nu) + \text{jets}$
  - QCD: extreme kinematical region, data-driven estimation needed

Analysis strategy

Two independent and complementary analyses performed:

- "2D" Exploits signal-specific angular properties performing a 2D fit in the corresponding variables
- "BDT" Multivariate approach with boosted decision trees. Combines many distinctive variables into a single powerful discriminator

Results in t-channel presented here correspond to 36 pb^{-1} collected in pp run of 2010. The updated results are in preparation

Details in Phys. Rev. Lett. **107**, 091802 (2011) and arXiv:1106.3052

Event selection (1/2)

Very similar in 2D and BDT analyses

Leptons

- Exactly one "tight" muon (electron) with
  - $p_\mathrm{T}\,(E_\mathrm{T}) > 20\,(30)$ GeV/c, $|\eta| < 2.1\,(2.5)$
  - relative isolation $I_\mathrm{r} = I_\mathrm{abs}/p_\mathrm{T}\,(E_\mathrm{T}) < 0.05\,(0.1)$, absolute isolation I_abs is sum of p_T in cone of radius 0.3 around the lepton, excluding the lepton itself
  - no jet within cone of radius 0.3 around the lepton
- Veto additional "loose" leptons with
  - $p_\mathrm{T}\,(E_\mathrm{T}) > 10\,(15)$ GeV
  - $I_\mathrm{r} < 0.2$

Event selection (2/2)

Jets

- Exactly two anti-k_{T} 0.5 jets with $p_{\mathrm{T}} > 30$ GeV$/c$, $|\eta| < 5$
- Exactly one b-tagged jet (The b-tagging algorithm exploits the impact parameter of the tracks associated with the jet)
- (2D analysis only) Exactly one b-vetoed jet
- (BDT analysis only) $\Delta\phi(j_1, j_2) < 3$: excludes back-to-back W+jets events poorly modeled with Pythia D6T tune

Invariant mass

- $M_{\mathrm{T}} > 40\,(50)$ GeV$/c$ in muon (electron) channel, where

$$M_{\mathrm{T}} = \sqrt{(p_{\mathrm{T}}(l) + p_{\mathrm{T}}(\nu))^2 - (p_x(l) + p_x(\nu))^2 - (p_y(l) + p_y(\nu))^2}$$

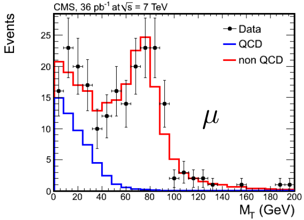
Data-driven QCD estimation

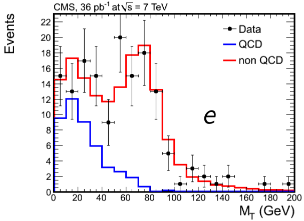
The analysis probes for a very specific kinematical region populated by the tails of QCD distribution only

- Amount of QCD estimated with maximum likelihood fit to M_T:

$$F(M_T) = a \cdot S(M_T) + b \cdot B(M_T)$$

- Shape of non-QCD $S(M_T)$ taken from simulation
- Shape of QCD $B(M_T)$ taken from orthogonal data sample with inverted cut in I_r

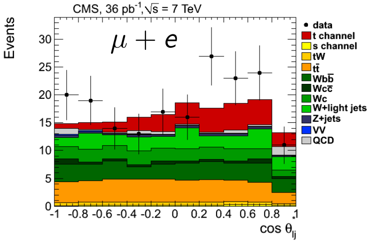


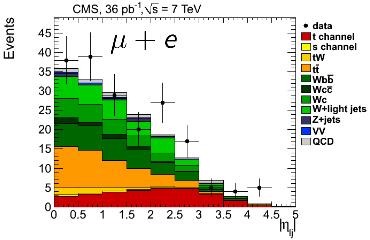


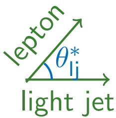
Event yields after selection

Process	2D, μ channel	2D, e channel	BDT, μ channel	BDT, e channel
single top, t channel	17.6 ± 0.7 (†)	11.2 ± 0.4 (†)	17.6 ± 0.7 (†)	10.7 ± 0.5 (†)
single top, s channel	0.9 ± 0.3	0.6 ± 0.2	1.4 ± 0.5	1.0 ± 0.3
single top, tW	3.1 ± 0.9	2.4 ± 0.7	3.8 ± 1.1	< 0.1
WW	0.29 ± 0.09	0.23 ± 0.07	0.32 ± 0.10	0.23 ± 0.07
WZ	0.24 ± 0.07	0.17 ± 0.05	0.33 ± 0.10	1.5 ± 0.4
ZZ	0.018 ± 0.005	0.011 ± 0.003	0.020 ± 0.006	< 0.1
$W+$ light partons	18.2 ± 5.5	11.6 ± 2.3	8.4 ± 4.2	7.0 ± 3.5
$Z + X$	1.7 ± 0.5	1.6 ± 0.3	0.7 ± 0.2	0.05 ± 0.03
QCD	0.6 ± 0.3	$2.6^{+3.4}_{-2.6}$	4.9 ± 2.5	5.3 ± 5.3
$VQ\bar{Q}$	20.4 ± 10.2	14.1 ± 7.1	17.6 ± 8.8	11.7 ± 5.8
Wc	$12.9\,^{+12.9}_{-6.5}$	$9.4\,^{+9.4}_{-4.7}$	$9.2\,^{+9.2}_{-4.6}$	$5.9\,^{+5.9}_{-2.9}$
$t\bar{t}$	20.3 ± 3.6	15.6 ± 2.8	34.9 ± 4.9	22.9 ± 3.2
Total background	78.6 ± 15.2	58.4 ± 11.0	82.4 ± 13.1	55.9 ± 10.2
Signal + background	96.2 ± 15.3	69.6 ± 11.0	100.0 ± 13.2	66.6 ± 10.2
Data	112	72	139	82

2D analysis variables







$\cos\theta^*_{lj}$

- Due to $V - A$ structure of weak interaction t-quark is almost 100% left-handed polarized
- This feature propagates to signal asymmetry in distribution over $\cos\theta^*_{lj}$ (angle calculated in t-quark rest frame)

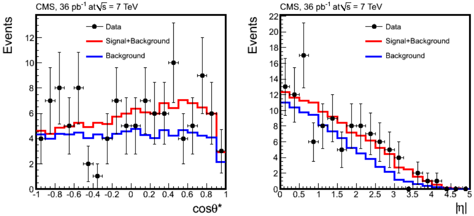
η_{lj}

- Light-quark jet recoiling against much more massive t-quark has non-central pseudorapidity distribution

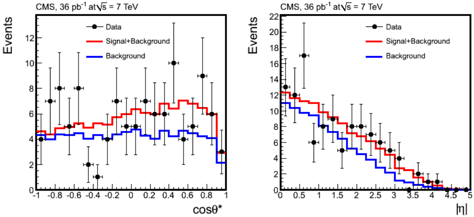
2D signal extraction

The x-section is determined with extended maximum likelihood fit to $(\cos\theta^*_{lj},\ |\eta_{lj}|)$ distribution

μ



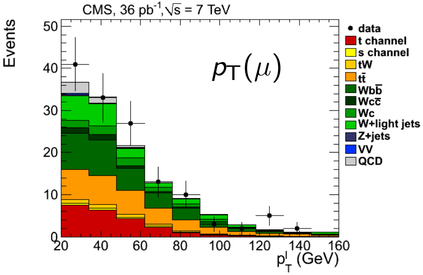
e



BDT analysis variables

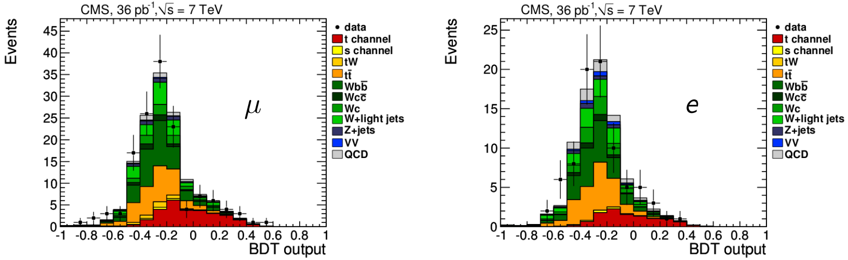
- 37 variables used in total. They reflect different event properties:
  - kinematics and properties of the lepton and jets
  - correlations between these objects
  - properties of derived objects (t-quark, W, etc.)
  - angular distributions between the original and derived objects
  - global event characteristics (sphericity, total transverse energy, etc.)

- The most sensitive variables:
  - lepton's p_T
  - $\hat{s}$ (invariant mass of system $W + j + j$)
  - dijet p_T
  - p_T of the most b-tagged jet
  - reconstructed t-quark mass



BDT signal extraction

The x-section is determined from a binned likelihood fit to BDT output within a Bayesian approach. The systematics is treated as nuisance parameters, they are marginalized out through Markov chain Monte Carlo (MCMC)



Systematics

Impact estimated through pseudo-experiments

Main sources of systematics:

- *b*-tagging
- Q^2 scaling
- jet energy scale

uncertainty	**correlation**	impact on **2D** $-$	**2D** $+$	**BDT** $-$	**BDT** $+$
statistical only	60	52		39	
shared shape/rate uncertainties:					
ISR/FSR for $t\bar{t}$	100	-1.0	$+1.5$	< 0.2	< 0.2
Q^2 for $t\bar{t}$	100	$+3.5$	-3.5	$+0.3$	-0.4
Q^2 for V+jets	100	$+5.7$	-12.0	$+2.6$	-4.5
Jet energy scale	100	-8.8	$+3.6$	-5.1	$+1.2$
b tagging efficiency	100	-19.6	$+19.8$	-15.2	$+14.6$
MET (uncl. energy)	100	-5.7	$+3.7$	-3.9	-0.5
shared rate-only uncertainties:					
$t\bar{t}$ ($\pm$14%)	100	$+2.0$	-1.9	$+0.5$	-0.6
single top s ($\pm$30%)	100	-0.4	$+0.5$	-0.4	$+0.4$
single top tW ($\pm$30%)	100	$+1.1$	-1.0	< 0.2	< 0.2
$Wb\bar{b}$, $Wc\bar{c}$ ($\pm$50%)	100	-3.0	$+2.9$	$+1.7$	-1.9
Wc ($^{+100\%}_{-50\%}$)	100	-3.0	$+6.1$	-2.4	$+4.4$
Z+jets ($\pm$30%)	100	-0.6	$+0.7$	$+0.4$	-0.2
electron QCD (BDT: $\pm$100%, 2D: $^{+130\%}_{-100\%}$)	50	$+2.9$	-3.7	-1.7	$+1.7$
muon QCD (BDT: $\pm$50%, 2D: $\pm$50%)	50	< 0.2	< 0.2	-2.1	$+2.1$
signal model	100	-5.0	$+5.0$	-4.0	$+4.0$
BDT-only uncertainties:					
electron efficiency ($\pm$5%)	0	—	—	-1.4	$+1.4$
muon efficiency ($\pm$5%)	0	—	—	-3.6	$+3.5$
V+jets ($\pm$50%)	0	—	—	-1.5	< 0.2
2D-only uncertainties:					
muon W+light ($\pm$30%)	0	-1.4	$+1.4$	—	—
electron W+light ($\pm$20%)	0	-0.6	$+0.7$	—	—
W+light model uncertainties	0	-5.4	$+5.4$	—	—

Combination and results in t-channel

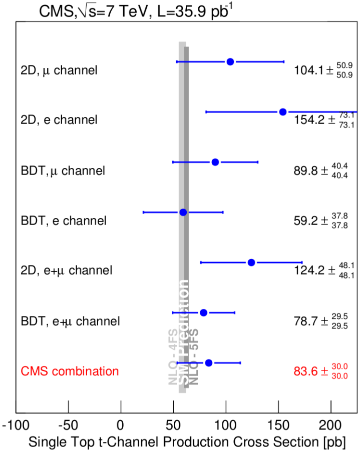
- 2D and BDT analyses combined through BLUE method
- The combined x-section:

$$83.6 \pm 29.8\,(\text{stat.}+\text{syst.}) \pm 3.3\,(\text{lumi.})\ \text{pb}$$

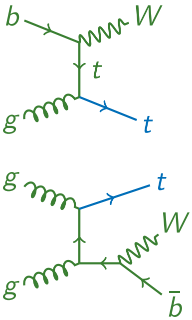
- Stat. significance is $3.7\,(3.5)\,\sigma$ w.r.t. expected significance $2.1^{+1.0}_{-1.1}$ $(2.9^{+1.0}_{-0.9})$ σ for 2D (BDT)
- 2D (BDT) sets 95% CL lower limit

$$|V_{tb}| > 0.62\,(0.68),$$

where $|V_{td}|, |V_{ts}| \ll |V_{tb}|$ and $|V_{tb}| \in [0, 1]$ assumed



Single top tW-channel



- Analysis exploits leptonic decays of W:
  $tW \to WbW \to l\nu bl'\nu'$
- Signal signature:
  - exactly two isolated leptons of opposite charge
  - exactly one b-jet within acceptance
  - large enough $\not{E}_{\mathrm{T}}$ due to neutrinos
- Main backgrounds:
  - $t\bar{t}$: very similar kinematics (interference at NLO!), high rate
  - $Z/\gamma^*(\to ll) + \text{jets}$

Analysis strategy

- Three leptonic final states: $ee, \mu\mu, e\mu$
- Pure counting experiment
- Rates for $t\bar{t}$ and $Z/\gamma^*(\rightarrow ll) +$ jets backgrounds estimated from data

Analysis in tW-channel is based on the integrated luminosity of 2.1 fb^{-1} recorded during the first data-taking period of 2011

Details in CMS PAS TOP-11-022

Event selection (1/3)

Leptons

- Exactly two "tight" leptons of opposite charge with
  - $p_{\mathrm{T}} > 20$ GeV/c, $|\eta| < 2.4\,(2.5)$ for a muon (electron)
  - relative isolation $I_{\mathrm{r}} < 0.15$
- Veto additional "loose" muons (electrons) with
  - $p_{\mathrm{T}} > 10\,(15)$ GeV/c, $|\eta| < 2.5$
  - $I_{\mathrm{r}} < 0.2$
- Reject events with leptons of the same flavor and dilepton mass
  - $81 < m_{ll} < 101$ GeV/c^2: Z veto
  - $m_{ll} < 20$ GeV/c^2: poor data-MC agreement
- $\not\!\!E_{\mathrm{T}} > 30$ GeV in ee and $\mu\mu$ channels (to reduce $Z/\gamma^*(\rightarrow ll)+$ jets)

Event selection (2/3)

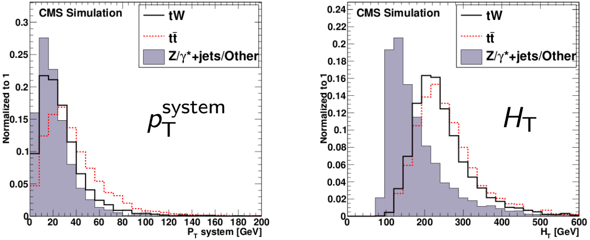
Jets

- Anti-k_T algorithm with cones 0.5
- Veto "tight" leptons inside cone 0.3 around a jet
- Exactly one jet with $p_T > 30$ GeV/c, $|\eta| < 2.4$
- The jet must be b-tagged (the b-tagging algorithm exploits reconstructed secondary vertex)
- Veto additional b-tagged jets with $p_T > 20$ GeV/c

Event selection (3/3)

Transverse balance

- p_T of system $l_1 + l_2 + j + \not{E}_T$ must be less than 60 GeV/c
- In $e\mu$ channel $H_T = p_T(l_1) + p_T(l_2) + p_T(j) + \not{E}_T > 160$ GeV



$e\mu$ final state, events passing leptonic step of selection and containing exactly one jet

Data-driven estimation of Z/γ^* and $t\bar{t}$

- $Z/\gamma^* +$ jets normalization estimated with the events failing Z veto:

$$N_{ll,\text{out}}^{\text{estimated}} = \frac{N_{ll,\text{out}}^{\text{MC}}}{N_{ll,\text{in}}^{\text{MC}}} \cdot \left(N_{ll,\text{in}}^{\text{observed}} - \frac{1}{2} k \cdot N_{e\mu,\text{in}}^{\text{observed}} \right)$$

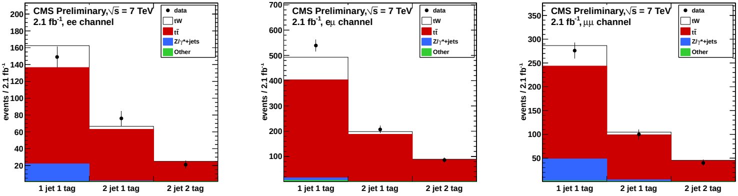
  - $N_{e\mu,\text{in}}^{\text{observed}}$ accounts for non-peaking backgrounds (e.g. $t\bar{t}$)
  - factor k is responsible for difference in e/μ acceptance (taken from data from numbers of near-peak ee and $\mu\mu$ events after leptonic selection only)
- $t\bar{t}$ rate is related to one of the largest sources of uncertainty
  - two control regions are defined: 2 jets, 1 tag and 2 jets, 2 tags; they are by far dominated by $t\bar{t}$
  - numbers of events in these regions are fed into the statistical procedure constraining $t\bar{t}$ normalization and b-tagging efficiency

Event yields after selection (1/2)

process	ee channel	$e\mu$ channel	$\mu\mu$ channel
Signal region (1jet, 1tag)			
tW	24.7±0.9	88 ±2	39±1
$t\bar{t}$	110±4	372±8	174±5
$Z/\gamma^{\star}$(data-driven)	20.7 ± 3.9	10 ± 2	45.7 ± 6.1
other	1.0 ±0.2	5 ±1	2.1 ±0.2
all background	132 ± 4	387 ± 9	222 ± 8
data	149	539	276
Background region A (2jets, 1tag)			
$t\bar{t}$	53±3	169±5	81±4
tW, $Z/\gamma^{\star}$, other	5.1±0.8	11.0±0.6	9±1
data	76	208	100
Background region B (2jets, 2tag)			
$t\bar{t}$	23 ±2	73±4	37±3
tW, $Z/\gamma^{\star}$, other	1.2±0.5	2.2±0.2	1.2±0.2
data	21	86	40

Event yields after selection (2/2)

Event yields in signal and the two control regions



- $Z/\gamma^* +$ jets scaled to data-driven estimate
- $t\bar{t}$ scaled to result of the statistical procedure

Systematics

Systematic ($ee/e\mu/\mu\mu$) [%]	signal tW	$t\bar{t}$	$Z/\gamma^{\star}$	other
Luminosity	4.5	4.5	4.5	4.5
Pile-up multiplicity	0.48/0.55/0.73	-	-	$\star$
Trigger Efficiency	1.5	1.5	1.5	1.5
Muon reconstruction and identification	- /1/1	- /1/1	- /1/1	- /1/1
Electron reconstruction and identification	2/2/ -	2/2/ -	2/2/ -	2/2/ -
JES	$^{-2.5}_{+1.6}/^{-2.4}_{+0.1}/^{-0.6}_{+1.0}$	$^{-5.6}_{+4.4}/^{-6.0}_{+4.7}/^{-5.9}_{+2.3}$	-	$\star$
JER	1.1/0.5/0.4	3.1/3.9/4.4	-	$\star$
B-tagging	$^{-9.5}_{+10}/^{-9.8}_{+9.8}/^{-9.5}_{+10}$	$^{-8.5}_{+10}/^{-11}_{+10}/^{-9.1}_{+11}$	-	$\star$
Factorization/Normalization Scale (Q^2)	7.7/6/10	7.7/11/12	-	$\star$
ME/PS matching thresholds	-	5.7/0.7/2.3	-	$\star$
ISR/FSR	-	8.9/7.3/7.3	-	$\star$
DR/DS scheme	8.2/9.1/6.6	-	-	$\star$
$E_{\mathrm{T}}^{\mathrm{miss}}$ modeling	2.3/0.9/0.9	-	-	$\star$
PDF uncertainties	4.5/4.5/4.5	-	-	$\star$
Background Normalization	-	15/15/15	50/ 22/ 50	$\star$
Simulation statistics	3.5/1.9/2.7	-	-	17/21/11

"-" – systematics doesn't apply, "$\star$" – negligible

Results

- Measured x-section:

$$22^{+9}_{-7}\,(\text{stat.}+\text{syst.})\ \text{pb}$$

- Observed significance is $2.7\,\sigma$ consistent with the expected significance $1.8 \pm 0.9\,\sigma$

Conclusions

- Single top in the *t*-channel has been searched for with two complementary analyses: one exploiting two characteristic angular variables, another one using a multivariate technique
- Both *t*-channel analyses (re)found an evidence of the signal providing compatible results
- The analyses were combined to obtain a more precise measurement of *t*-channel x-section
- The x-section was translated into the lower limit on $|V_{tb}|$
- The tW-channel x-section was measured with a counting experiment
- All the results are in good agreement with the SM expectations

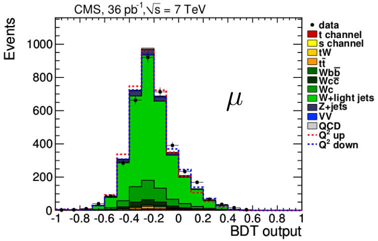
Thank you for your attention!

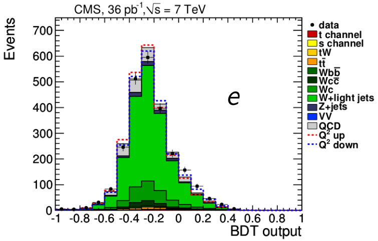
BACKUP SLIDES

Particle flow reconstruction

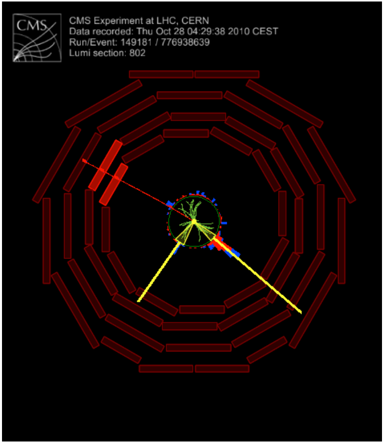
- Attempts to reconstruct every particle on individual basis
- Better performance thanks to using all the appropriate subdetector systems (e.g. jet performance can profit a lot from using tracker info)
- Every track or calorimeter energy deposit is guaranteed to be associated with one particle at maximum $\Rightarrow$ no double counting

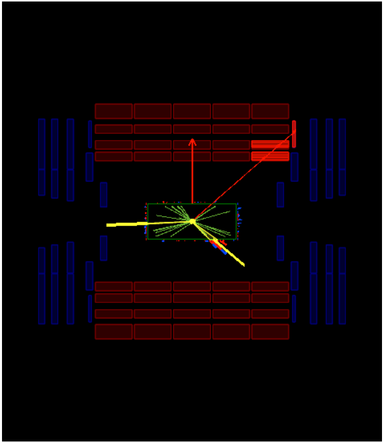
BDT validation in W-enriched (zero tag) sample



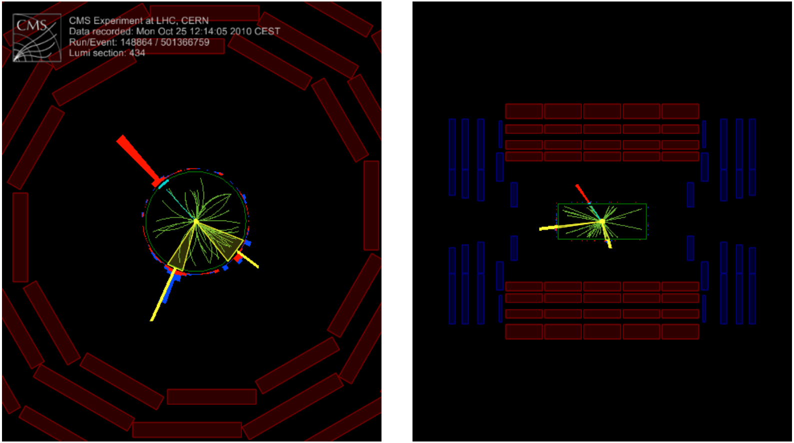


Golden candidate of t-channel process, muon



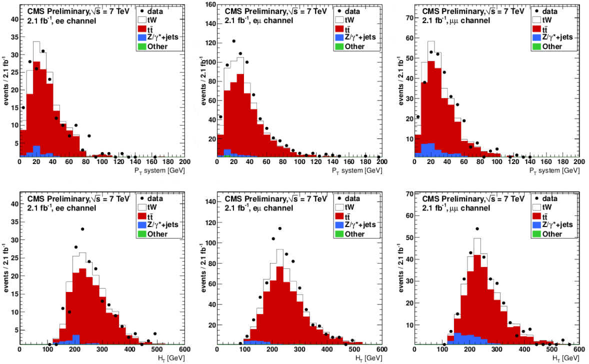


Golden candidate of t-channel process, electron



tW-analysis: p_T^{system} and H_T distributions

p_T^{system} (up) and H_T (down) in ee-, $e\mu$-, and $\mu\mu$-channels (from left to right)



Full selection except for requirements on these two variables