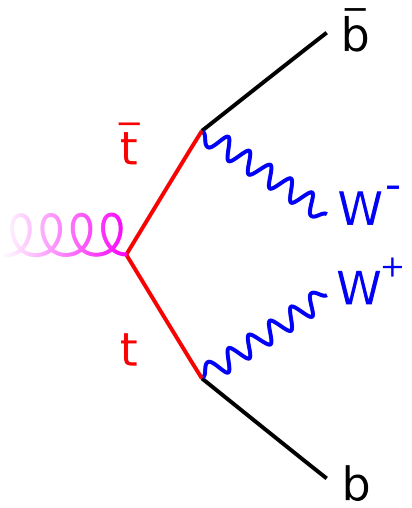
Measurement of the top quark pair production cross section at √s=7 TeV

Julien Caudron (UCLouvain), on behalf of CMS Collaboration

- Introduction
- Semi-leptonic:
  - strategy
  - analysis using b-tag
  - analysis without b-tag
- Dileptonic:
  - strategy
  - results
- Combination of semi-leptonic and dileptonic

DIS2011, Newport News
April 14th 2011

$t\bar{t}$ measurement at CMS



The $t\bar{t}$ pair production decays in 2 W bosons and 2 b quarks

The decays of the W bosons determine the final state:

- dileptonic, where the ee (1.9%), eμ (2.7%) and μμ (1.9%) final states are studied
- semi-leptonic, where the e+jets (17%) and μ+jets (17%) final states are studied

In 2010, CMS has collected an integrated luminosity of **35.9 pb^{-1}** of data at $\sqrt{s}$ = 7 TeV

First publication: A first measurement at 3.1 pb^{-1} in the dileptonic channel:

TOP-10-001: «*First Measurement of the Cross Section for Top-Quark Pair Production in Proton-Proton Collisions at $\sqrt{s}$ = 7 TeV*», Phys. Lett. B695 (2011) 424

Recent results, for 35.9 pb^{-1}:

- TOP-10-002: «*Measurement of the $t\bar{t}$ Pair production Cross Section at $\sqrt{s}$ = 7 TeV using the **Kinematic Properties** of **Lepton + Jets** Events*»
- TOP-10-003: «*Measurement of the $t\bar{t}$ Pair Production Cross Section at $\sqrt{s}$ = 7 TeV using **b-quark Jet Identification** Techniques in **Lepton + Jets** Events*»
- TOP-10-005: «*Measurement of the top-quark pair-production cross section in the **dilepton channel** at $\sqrt{s}$ = 7 TeV*»
- TOP-11-001: «*Combination of top pair production cross sections in pp collisions at $\sqrt{s}$ = 7 TeV and comparisons with theory*» (= **TOP-10-003** and **TOP-10-005** combined results)

Semi-leptonic: strategy

Selection:

one lepton (with second lepton veto)**:**
p_T **> 30 (20)** GeV/c, **|η| < 2.5 (2.1)** for e (μ),
Isolated in tracker and calorimeter

jets selection:
corrected Jet, **|η|<2.4**

p_T **> 30**

no b-tags analysis

p_T **> 25**

$\not{E}_T$ and b-tag selection:
$\not{E}_T$ > 20, Secondary Vertex
b-tag analysis

Main backgrounds :

- W→ℓν+Jets, divided in b-jets, c-jets and light-jets in b-tag analysis
- Z→ℓℓ+Jets, γ+Jets, Single top
- QCD (estimated from data)

Cross section measurement :

Obtained from fit of distributions:

for b-tag analysis:

Simultaneous fit of

- Secondary Vertex Mass (from tracks associated with the vertex with a pion mass assumption)
- Number of jets and b-tagged jets

for no b-tags analysis:

- $\not{E}_T$ for = 3 jets
- M3 for >= 4 jets (inv. mass of the 3 jets of highest $|\sum \vec{p}_T|$)

Semi-leptonic: b-tag analysis 1/2

Fit procedure :

For each process i, for each channel (lept) and jets / b-tagged jets bin (jets, tag) :

Scale factor for the simulated cross section

Nuisance parameter, where X can be: b-tag efficiency, mistag, JES, Q^2 scale

$$\mathcal{N}_i^{\text{pred}}(\text{lept},\text{jets},\text{tag}) = k_i \cdot N_i^{\text{MC}}(\text{lept},\text{jets},\text{tag}) \cdot \prod_X P_i^{\text{X}}(\text{lept},\text{jets},\text{tag}|\mathcal{R}_{\text{X}})$$

Number of predicted events

Number of events predicted by MC corrected for Data/MC discrepancies

Polynomial functions describing the effect of the nuisance parameter, obtained from MC

The fit is done simultaneously on i, lept, jets and tag, its parameters are k_i and R_x.

Constraints:

k is constrained around a normalization factor from simulation or from data driven estimate

Systematic uncertainties:

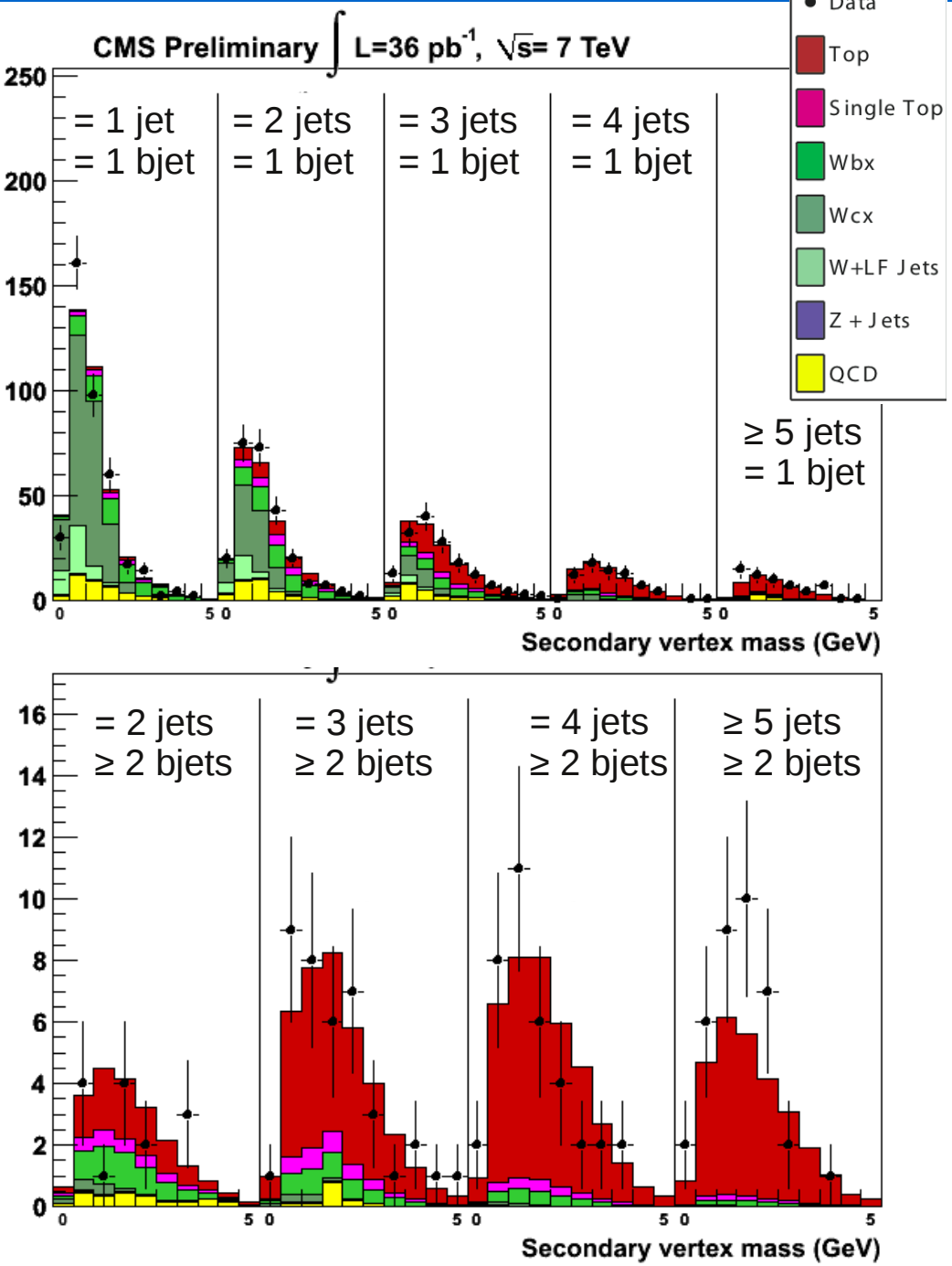
incorporated in the profile likelihood or treated additionally

Systematics for e+jets and μ+jets analysis :
(sources correlated)

Source	Uncertainty (%)
Systematic uncertainties	
Lepton ID/reco/trigger	3
Unclustered $E_{\text{T}}^{\text{miss}}$ resolution	< 1
$t\bar{t}$+ Jets Q^2-scale	2
ISR/FSR	2
ME to PS matching	2
PDF	3.4
Profile likelihood parameters	
Jet energy scale and resolution	7.0
b tag efficiency	7.5
W+Jets Q^2-scale	9.1
Combined	11.6

Semi-leptonic: b-tag analysis 2/2

Fitted secondary vertex mass in e+jets channel :



Results :

The fit returns values for:

- b-tag scale factor: 97.5 $^{+5}_{-4}$%

  (consistent with other b-tag studies)

- Jet Energy Scale: slightly harder than predicted

- scale factors w.r.t. the NLO prediction:

  for Wb: 1.9 $^{+0.6}_{-0.5}$

  for Wc: 1.4 ± 0.2

  (which are similar to recent observations at Tevatron)

- $k_{t\bar{t}}$, leading to the $t\bar{t}$ cross-section:

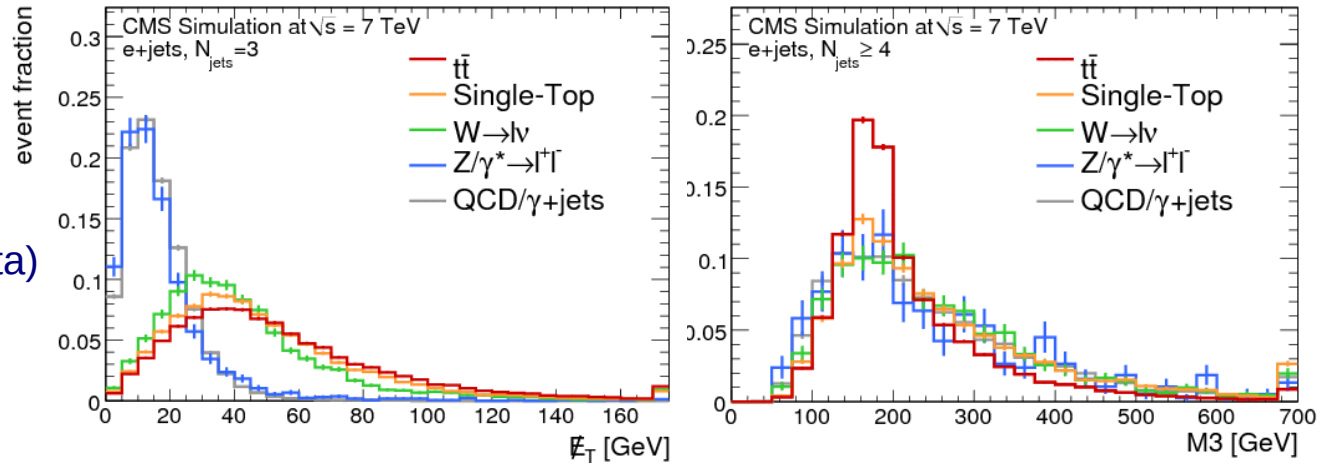
$\sigma_{t\bar{t}}$(semi-lept with b-tag) =
150 ± 9 (stat) ± 17 (syst) ± 6 (lumi) pb

Semi-leptonic: no b-tags analysis

Fit procedure on $\not\!\!E_T$ or M3 :
simultaneous fit of signal and background contamination
for both channels,
based on templates (from sim. and data)

Syst. uncert. : Effect on simulated templates when ±1σ variation
Dominant effect: JES
(followed by W+jets Q^2 scale)

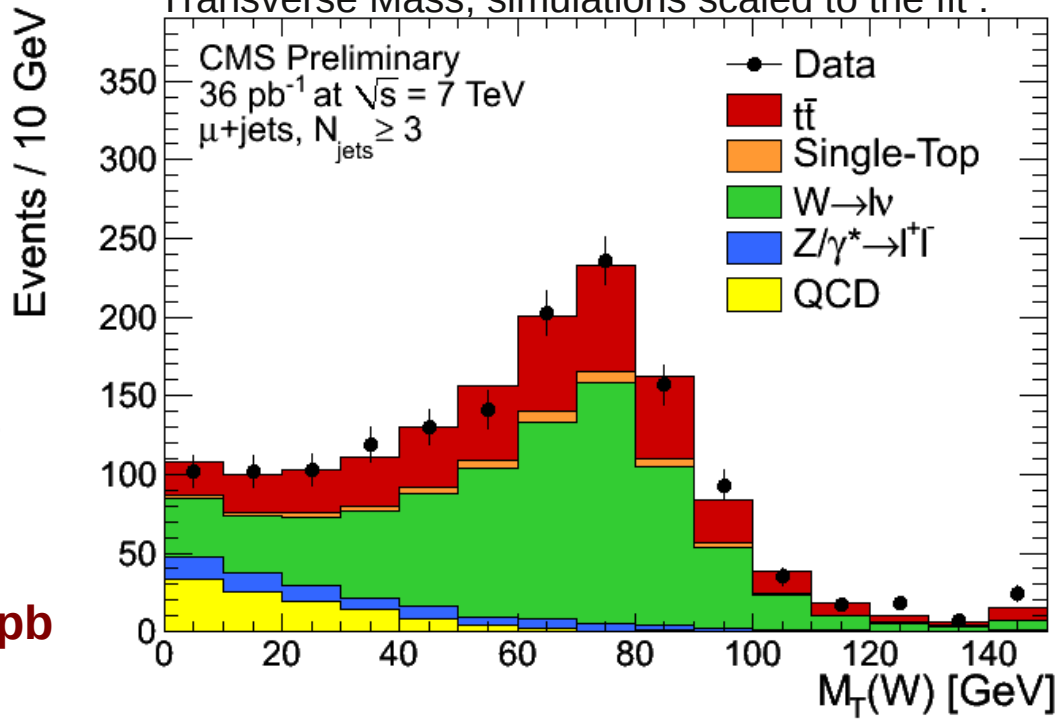
templates in e+jets :



Constraints:

- QCD normalization depending on the channel
- N(W+jets) / N(Z+jets) = theory ± 30%,
- Normalization factor w.r.t. theory are ≥0,
- Single Top normalization factor = 1.0 ± 0.3

Transverse Mass, simulations scaled to the fit :



→ **$\sigma_{t\bar{t}}$(semi-lept without b-tag) =**
173^{+39}_{-32} (stat+syst) ± 7 (lumi) pb

Dileptonic: strategy

Selection:

two opposite charge leptons:
p_T **> 20** GeV/c, **|η| < 2.5 (2.4)** for e (μ),
Isolated in tracker and calorimeter

↓

invariant mass selection:
$M_{\ell\ell}$ **> 12** GeV/c², $M_{\ell\ell} \notin$ **[91 ± 15]**

↓

jets selection:
corrected Jet, p_T **> 30 GeV/c, |η|<2.5**

=1 jet ≥2 jets

$\not{E}_T$ selection:
$\not{E}_T$ **> 50** GeV for ee/μμ
$M^e_T+M^\mu_T$ **> 130** GeV for eμ
1 jet no b-tags

$\not{E}_T$ selection:
$\not{E}_T$ **> 30** GeV for ee/μμ
2 jets no b-tags

↓

b-tag selection: Track Counting
at least 1 b-tagged jet **2 jets 1 b-tag**

Main backgrounds after leptonic selection :

- Drell-Yan → $\ell\ell$: main background, rejected by Z veto, jets and $\not{E}_T$, estimated from data
- W+Jets, semi-lept. $t\bar{t}$, QCD: from non-W/Z decays, estimated from data
- Single top tW, diboson, Z→ττ: small cross-sections, estimated from MC

Cross section measurement :

using a simple counting method:

$$\sigma^{t\bar{t}}_{data} = \sigma^{t\bar{t}}_{theory} \frac{N - B}{SF \cdot S_{exp}} = \frac{S_{obs}}{SF \cdot \mathcal{L} \cdot A}$$

SF: Data/MC scale factor,
S_{exp}: Expected signal evts from simulation

For each channel, for 2 jets no b-tags, 2 jets 1 b-tag and 1 jet no b-tags
→ 9 cross sections measurements, combined to obtain 1 dilepton cross section.

Dileptonic: results 1/2

Systematic uncertainties :

- For signal selection: cf. table
  (given in percent, the sources are uncorrelated, the channels are 100% (anti)correlated)
- For backgrounds:
  - extracted from the data-driven methods
  - assigned conservatively from the MC simulation
- For luminosity uncertainty: 4%

	$N_{jet} = 1$		$N_{jet} \geq 2$	
Source	$e^+e^- + \mu^+\mu^-$	$e^\pm\mu^\mp$	$e^+e^- + \mu^+\mu^-$	$e^\pm\mu^\mp$
Lepton selection	1.91/1.30	1.11	1.91/1.30	1.11
Energy scale	−3.0	−5.5	3.8	2.8
Lepton selection model	4.0	4.0	4.0	4.0
Branching ratio	1.7	1.7	1.7	1.7
Decay model	2.0	2.0	2.0	2.0
Event Q^2 scale	8.2	10	−2.3	−1.7
Top-quark mass	−2.9	−1.0	2.6	1.5
Jet and $\not\!\!E_T$ model	−3.0	−1.0	3.2	0.4
Shower model	1.0	3.3	−0.7	−0.7
Pileup	−2.0	−2.0	0.8	0.8
Subtotal (before tags)	11.2/11.1	13.1	8.0/7.9	6.2
b tagging ($\geq$ 1 b tag)			5.0	5.0
Subtotal with tags			9.5/9.4	8.0
Luminosity	4	4	4	4

Final state	e^+e^-	$\mu^+\mu^-$	$e^\pm\mu^\mp$
At least two jets, no b-tagging requirement			
Events in data	23	28	60
All backgrounds	5.5 ± 2.3	9.5 ± 4.3	6.7 ± 2.0
Cross section, pb	$189 \pm 52 \pm 29 \pm 8$	$159 \pm 45 \pm 39 \pm 6$	$160 \pm 23 \pm 12 \pm 6$
At least two jets, at least one b-jet			
Events in data	15	24	51
All backgrounds	2.3 ± 1.4	3.8 ± 2.0	3.0 ± 1.4
Cross section, pb	$150 \pm 46 \pm 22 \pm 6$	$186 \pm 45 \pm 25 \pm 7$	$156 \pm 23 \pm 13 \pm 6$
One jet, no b-tagging requirement			
Events in data	8	10	18
All backgrounds	2.1 ± 0.7	7.1 ± 4.3	4.9 ± 1.5
Cross section, pb	$282 \pm 135 \pm 45 \pm 11$	$107 \pm 119 \pm 163 \pm 4$	$200 \pm 65 \pm 35 \pm 8$

Results:

9 cross sections are obtained

2 jets 1 b-tag: more precise than without b-tag for the ee and μμ channels (but not for the eμ channel)

1 jet no b-tags: less precise but improves the combined result

Dileptonic: results 2/2

Combination method:

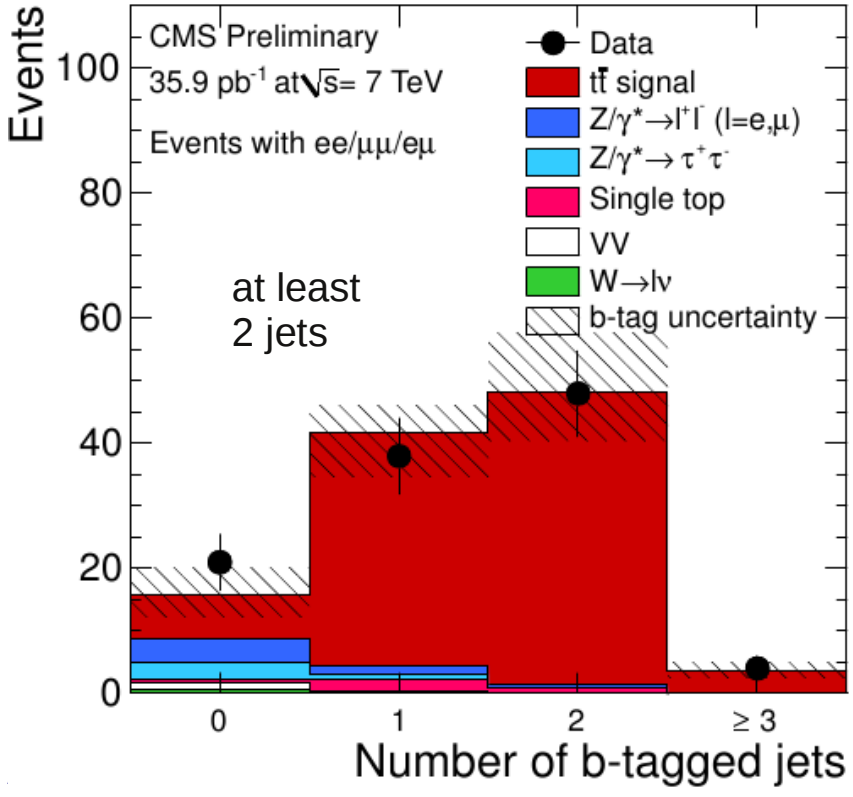
The 9 cross sections are combined using the BLUE technique (Best Linear Unbiased Estimator)

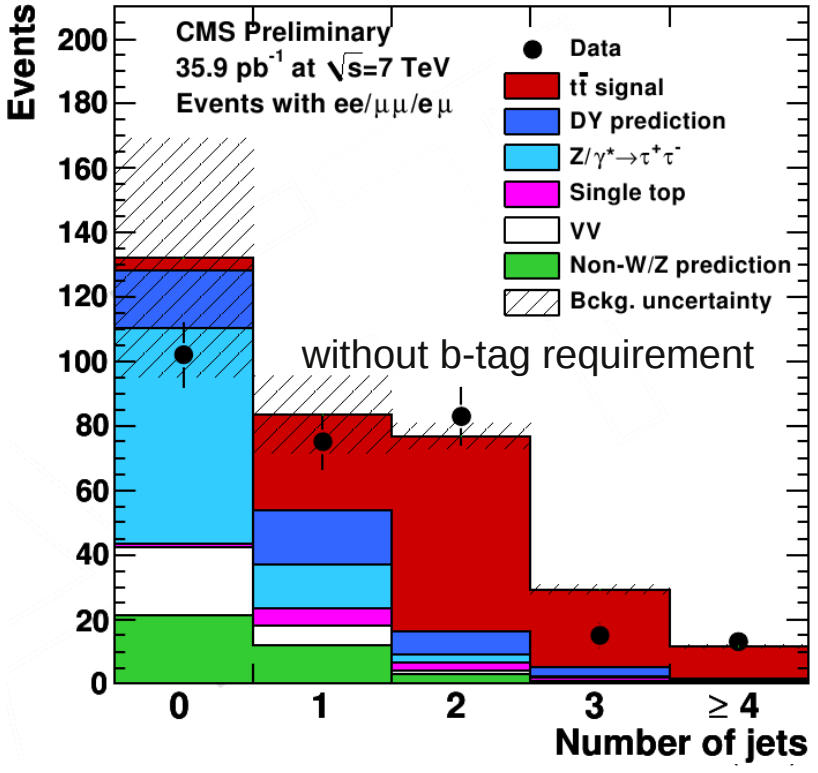
It takes into account the correlations between different contributions to the measurements

Results after combination:

$\sigma_{t\bar{t}}$**(dilepton) = 168 ± 18 (stat) ± 14 (syst) ± 7 (lumi) pb**

The luminosity uncertainty can be reduced using the normalisation w.r.t. to Z/γ* evts in the [76, 106] $M_{\ell\ell}$ window





Combination

Combining the results, using the BLUE technique:

- dileptonic : 168 ± 18 (stat) ± 14 (syst) ± 7 (lumi) pb
- semi-leptonic : 150 ± 9 (stat) ± 17 (syst) ± 6 (lumi) pb

→ **$\sigma_{t\bar{t}}$(CMS) =**

158 ± 10 (unc.) ± 15 (cor.) ± 6 (lumi) pb

where unc. : statistical uncertainties + background modelling in the dileptonic analysis

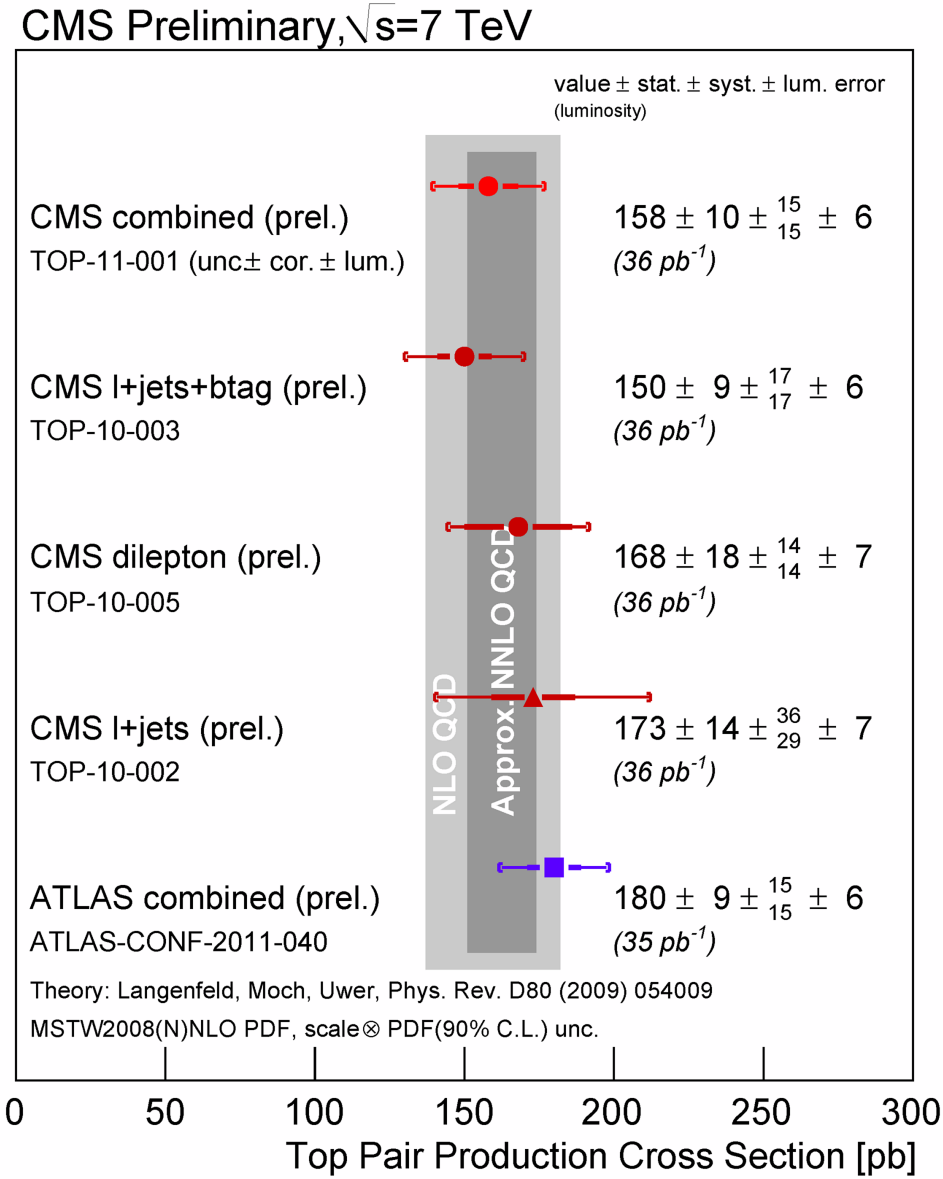
cor.: all the other uncertainties

This result is consistent with the theoretical $t\bar{t}$ cross section at approximate NNLO :

$\sigma_{t\bar{t}}$(HATHOR)= 164 $^{+5}_{-9}$ (scale) $^{+9}_{-9}$ (pdf) pb

$\sigma_{t\bar{t}}$(Kidonakis)= 163 $^{+7}_{-5}$ (scale) $^{+9}_{-9}$ (pdf) pb

and more precise than the NLO theory uncertainty

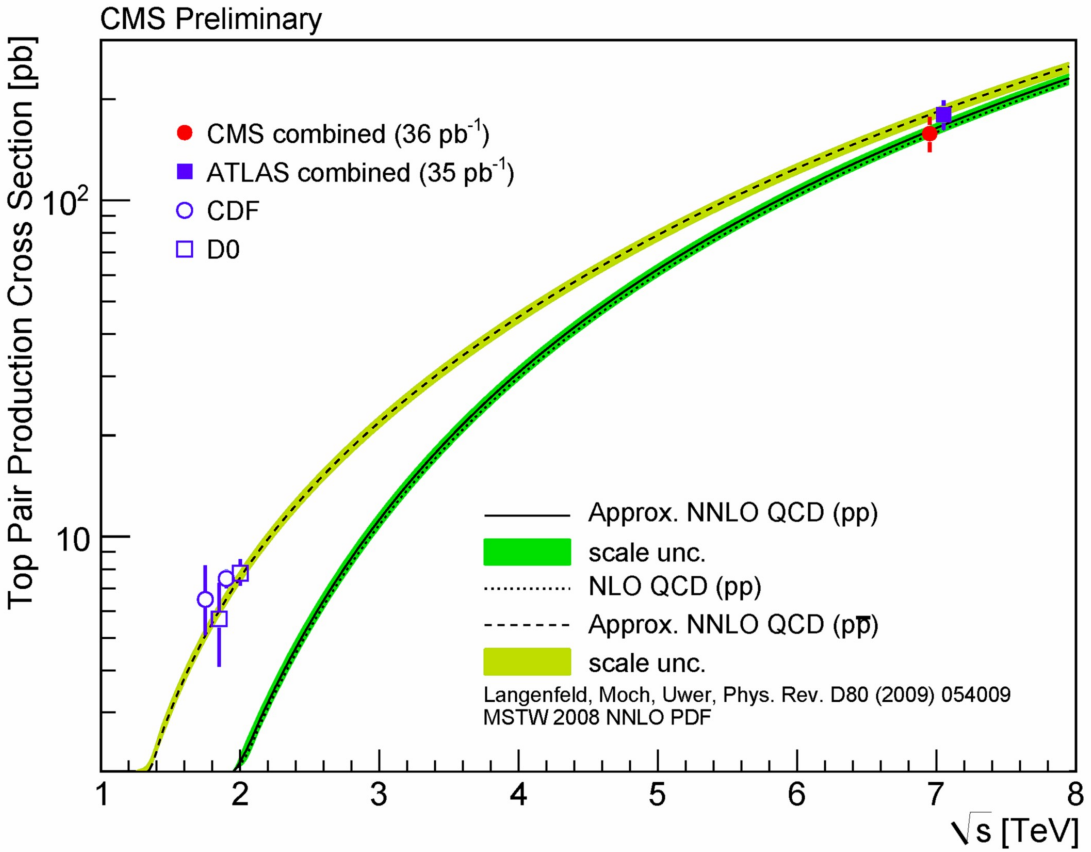


Conclusion

With **35.9 pb^{-1}** of data, harvested during the year 2010 **at √s = 7 TeV**, the CMS collaboration has extracted $t\bar{t}$ production cross section in the dileptonic and semi-leptonic channels.

Combining the dileptonic and semi-leptonic with b-tag using the BLUE technique, we obtain:

$\sigma_{t\bar{t}}$(CMS) = 158 ± 10 (unc) ± 15 (cor) ± 6 (lumi) pb



This result is in good agreement with the theory and fits nicely with previous data

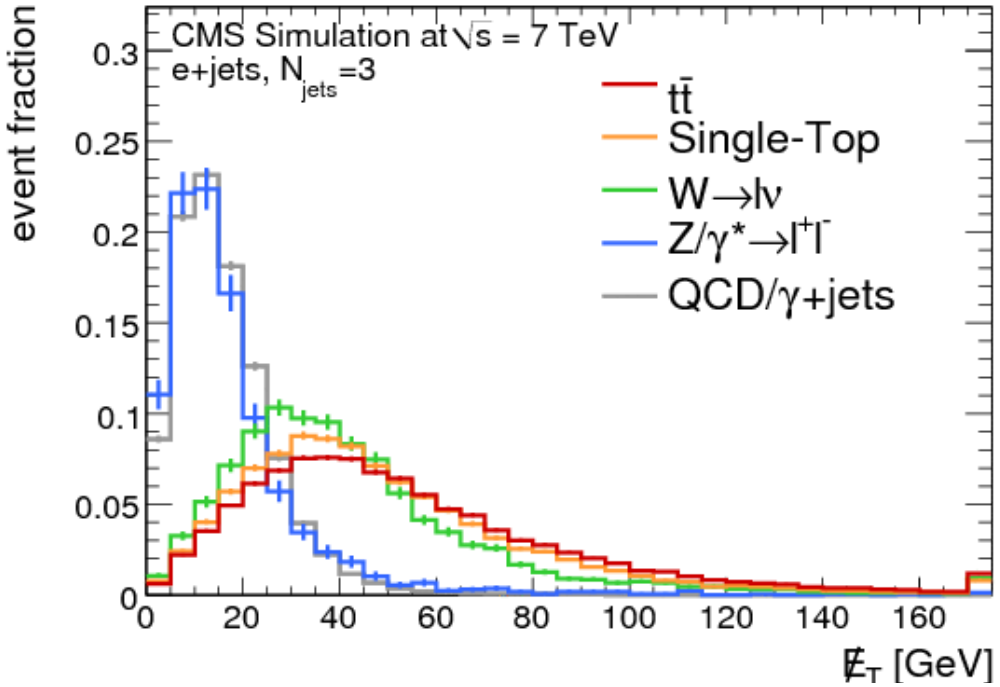
Currently, the uncertainties remain large to constrain the theory

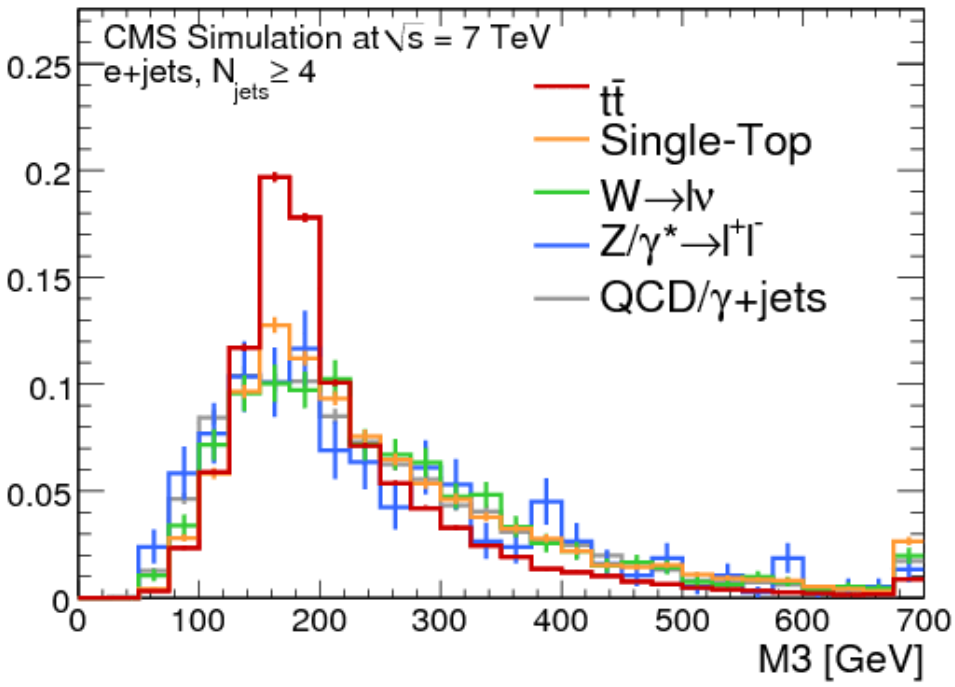
but :

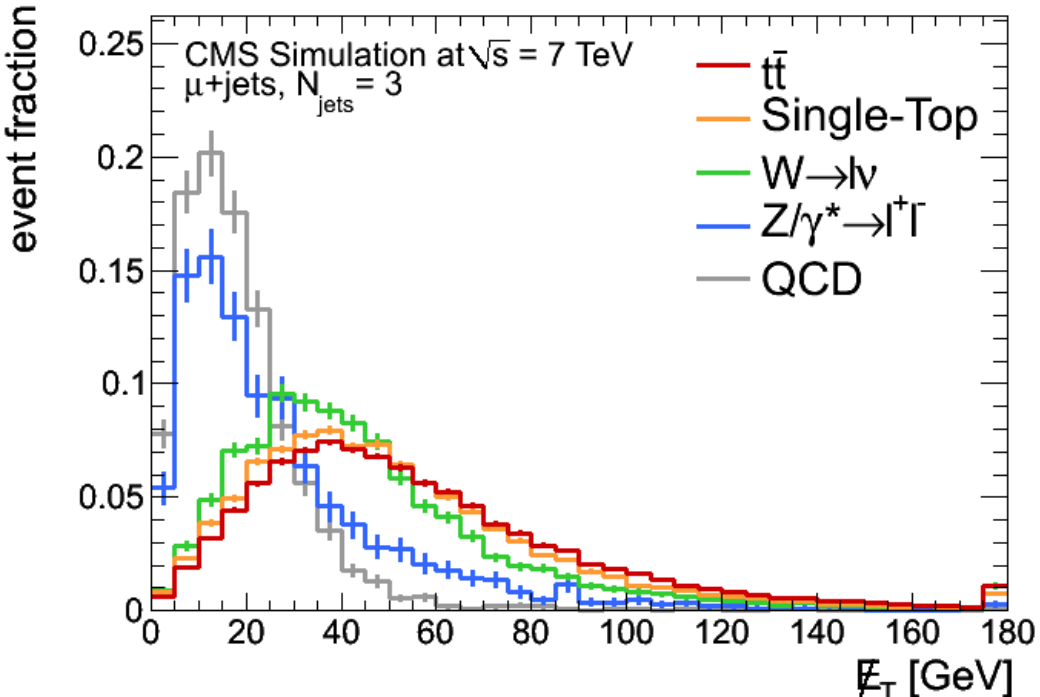
- impressive results after few tens of pb^{-1} (the current LHC regime is ~10pb^{-1} / 3 days of run)
- methods are ready and work well
- there is still room for improvement (syst., luminosity uncert., combination, ...)

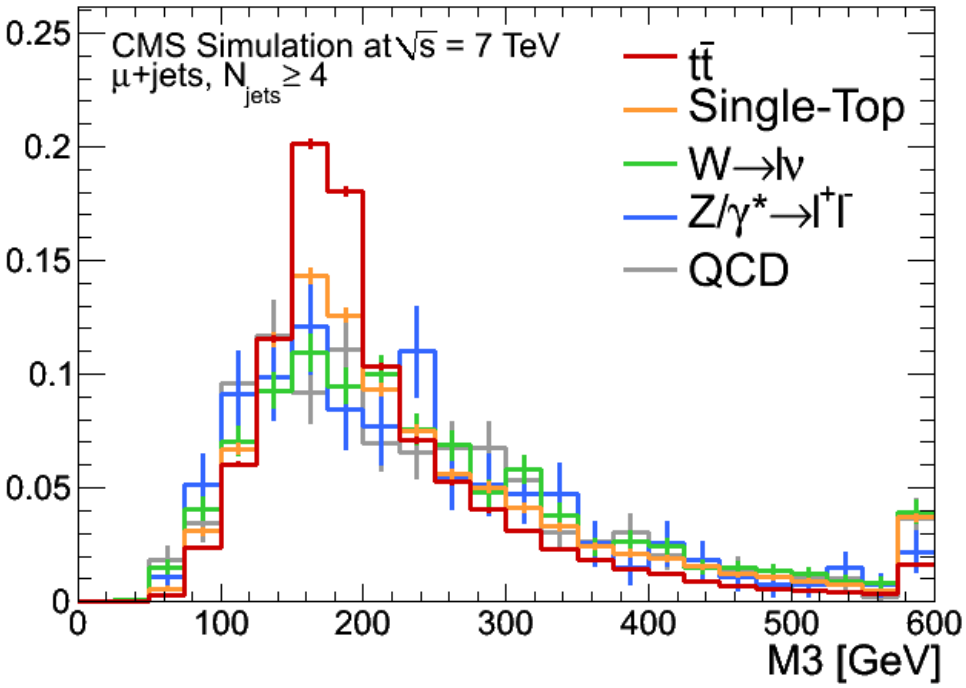
Backup slides

Semi-leptonic: no btag, templates

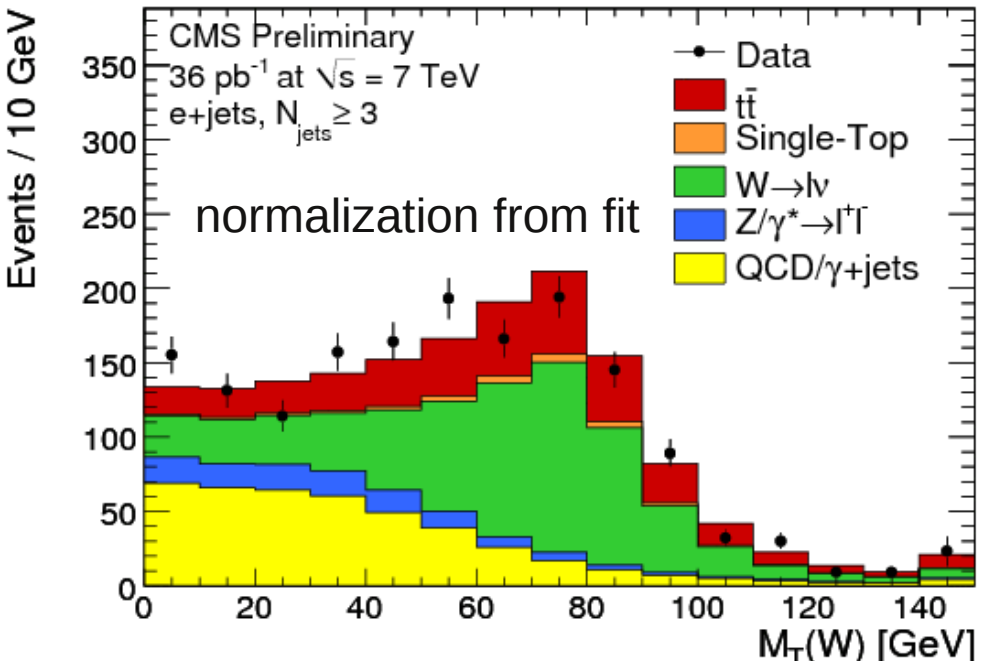


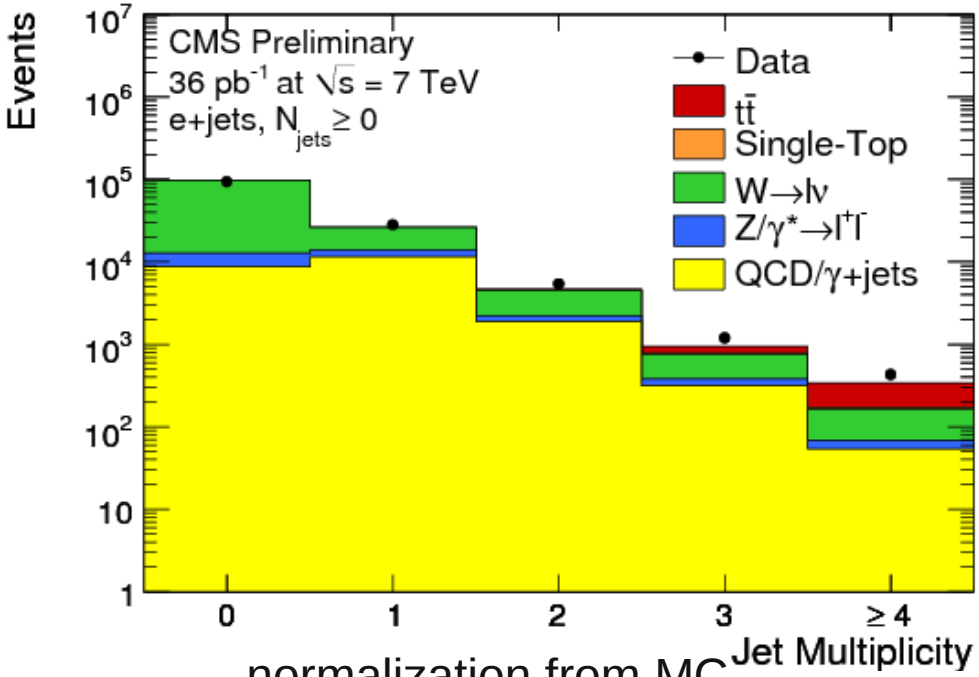




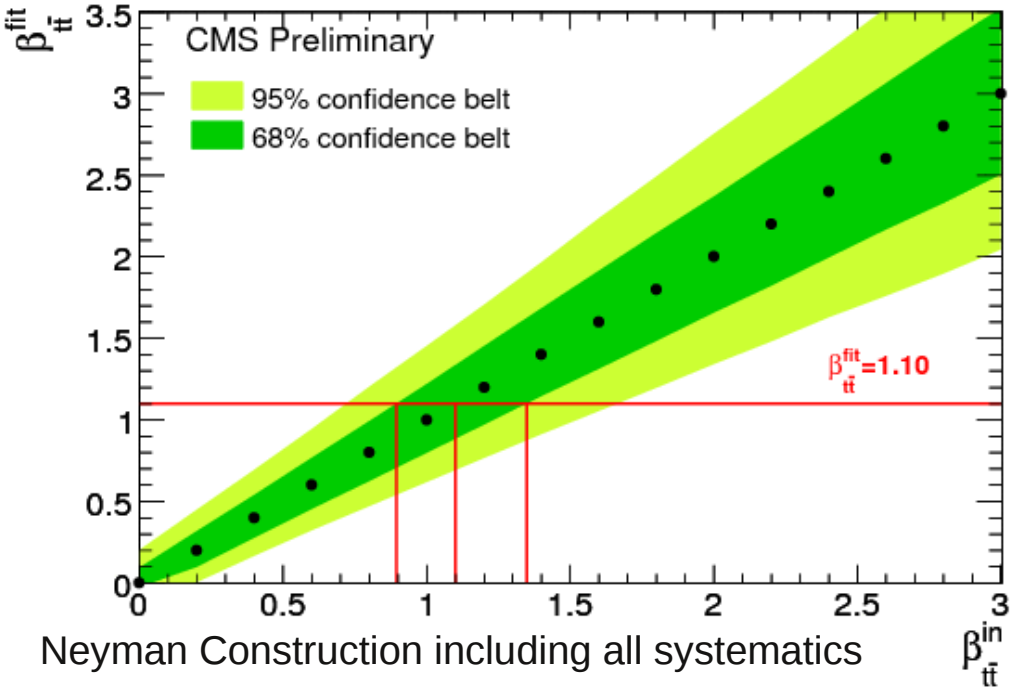


Semi-leptonic: no btag

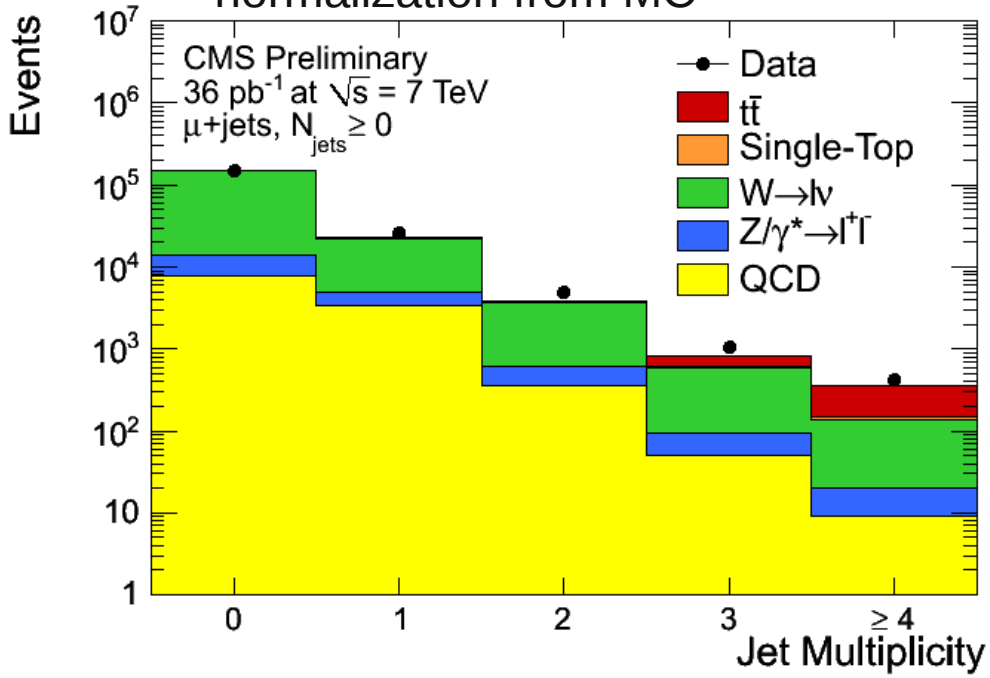




normalization from MC



Neyman Construction including all systematics



Semi-leptonic: no btag

event yields from simulation :

	electron+jets		muon+jets	
	$N_{jet} = 3$	$N_{jet} \geq 4$	$N_{jet} = 3$	$N_{jet} \geq 4$
$t\bar{t}$	157 ± 25	168 ± 27	197 ± 31	211 ± 33
single top	22 ± 1	8.8 ± 0.6	30 ± 1	11 ± 1
W+jets	374 ± 27	94 ± 7	486 ± 34	115 ± 9
Z+jets	66 ± 5	15 ± 1	46 ± 3	11 ± 1
QCD	314 ± 19	53 ± 8	49 ± 3	8.9 ± 1.0
sum MC	934 ± 55	339 ± 32	807 ± 53	358 ± 37
Data	1183	428	1064	423

result of the fit for e+jets and μ+jets:

	$\beta_{t\bar{t}}$	N_{ST}	N_{W+jets}	N_{Z+jets}	N_{QCD} e+jets	N_{QCD} μ+jets
predicted	1.00	72 ± 4	1069 ± 77	138 ± 10	367 ± 27	58 ± 4
fitted	1.10	76 ± 22	1475 ± 86	184 ± 51	440 ± 44	113 ± 31

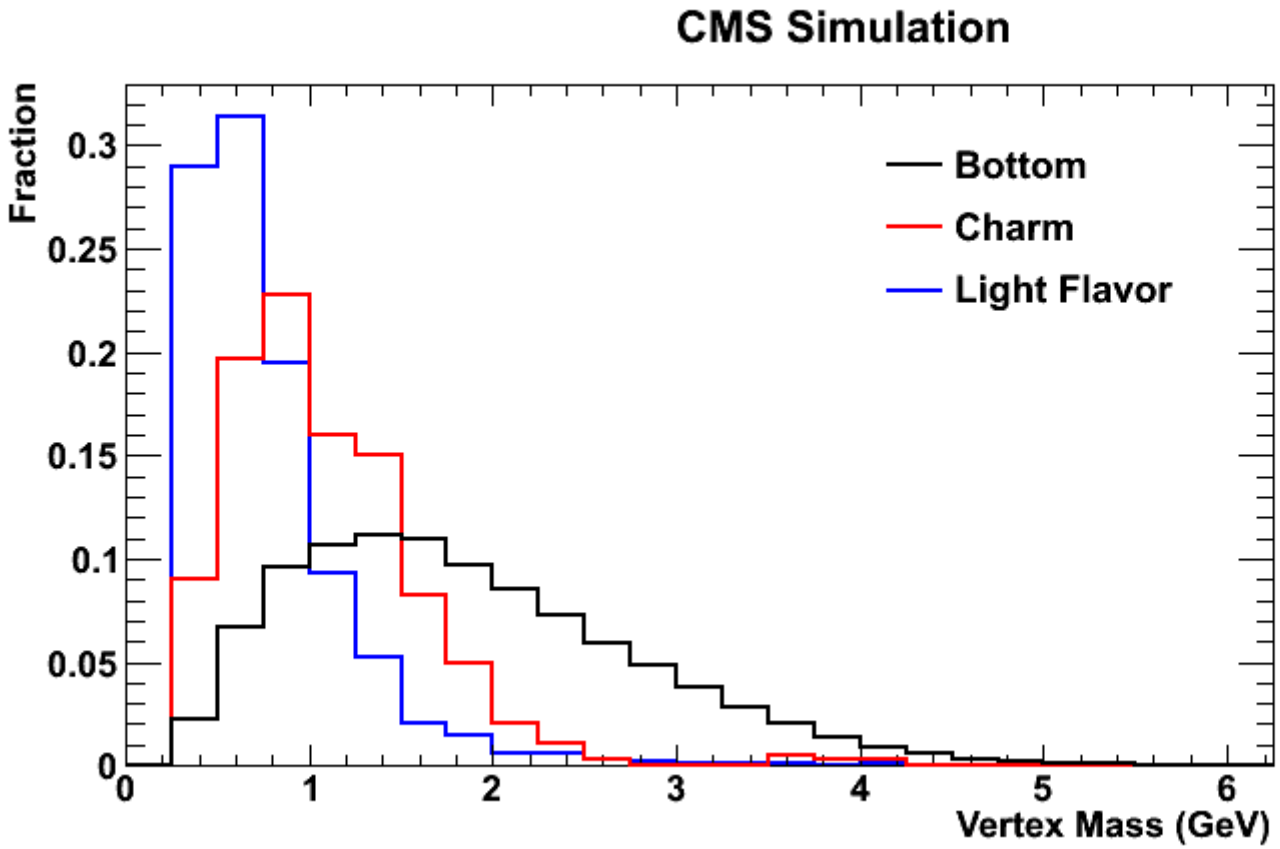
electrons + jets only: $\sigma_{t\bar{t}} = 180^{+45}_{-38}(\text{stat.} + \text{syst.}) \pm 7(\text{lumi.})\,\text{pb}$

muons + jets only: $\sigma_{t\bar{t}} = 168^{+42}_{-35}(\text{stat.} + \text{syst.}) \pm 7(\text{lumi.})\,\text{pb}$

combination : $\sigma_{t\bar{t}} = 173^{+39}_{-32}(\text{stat} + \text{syst}) \pm 7(\text{lumi})\,\text{pb}$

Semi-leptonic: btag

template:

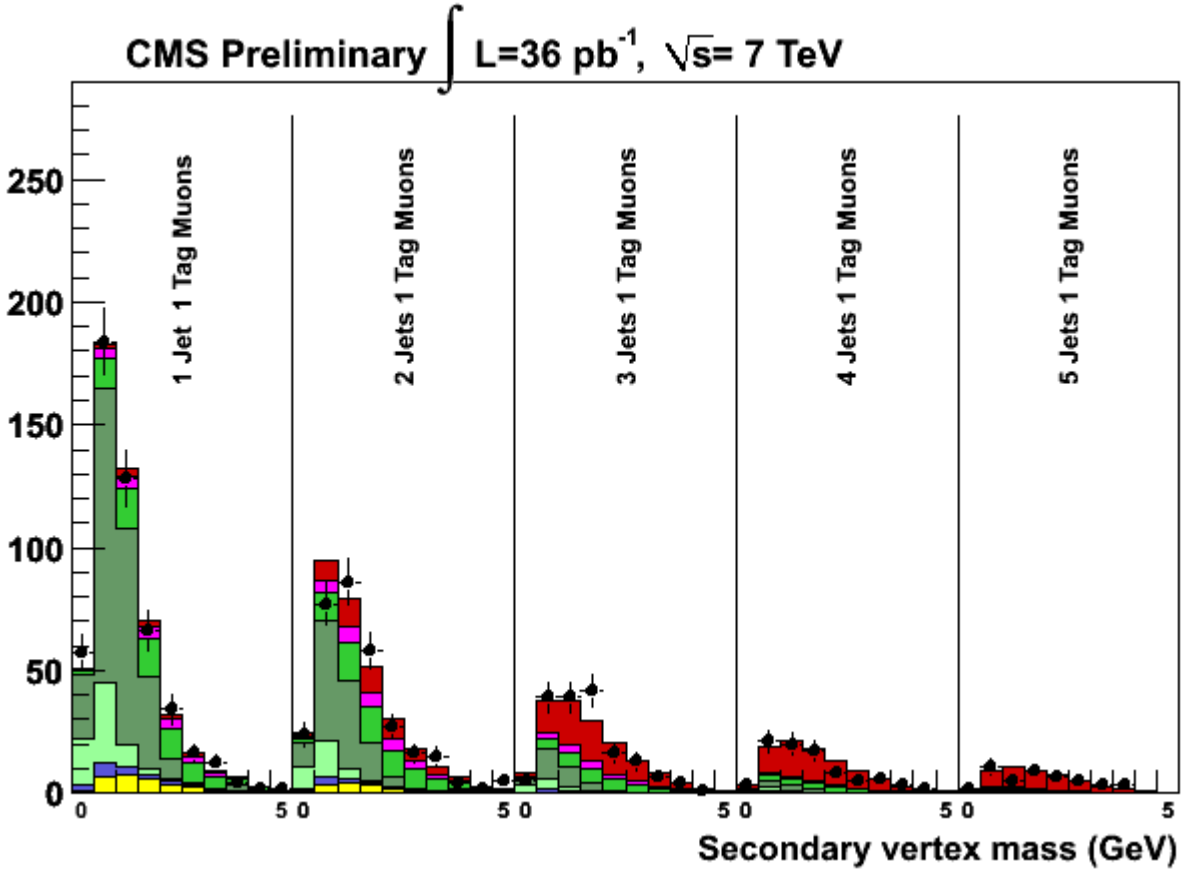


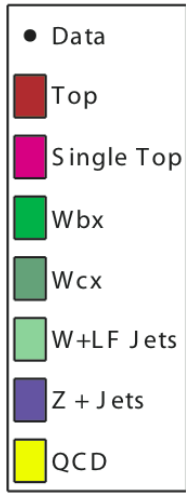
electrons + jets only: $\sigma_{t\bar{t}} = 158 \pm 14\ (\text{stat.}) \pm 19\ (\text{syst.}) \pm 6\ (\text{lum.})\ \text{pb}$

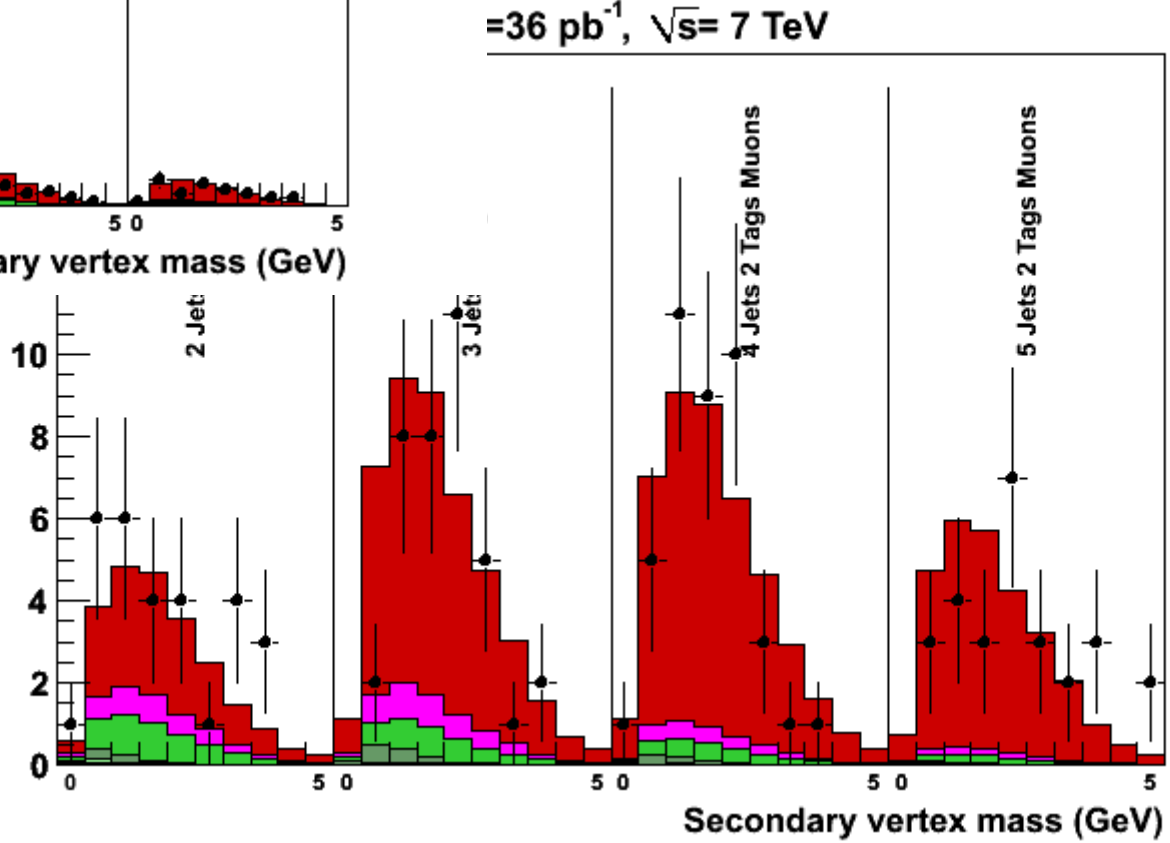
muons + jets only: $\sigma_{t\bar{t}} = 145 \pm 12\ (\text{stat.}) \pm 18\ (\text{syst.}) \pm 6\ (\text{lum.})\ \text{pb};$

combination : $\sigma_{t\bar{t}} = 150 \pm 9\ (\text{stat.}) \pm 17\ (\text{syst.}) \pm 6\ (\text{lum.})\ \text{pb}.$

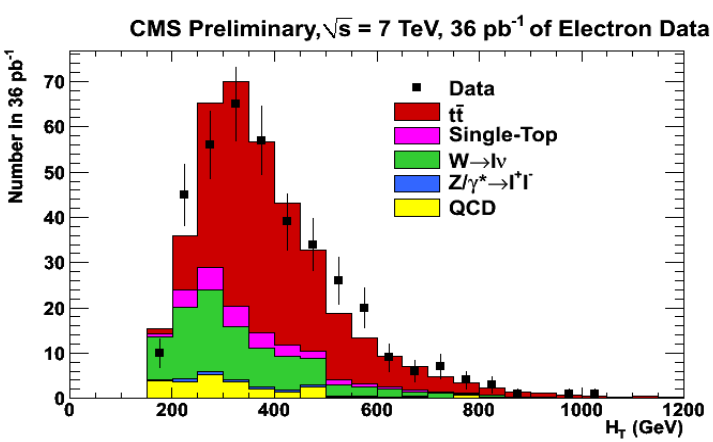
Semi-leptonic: btag

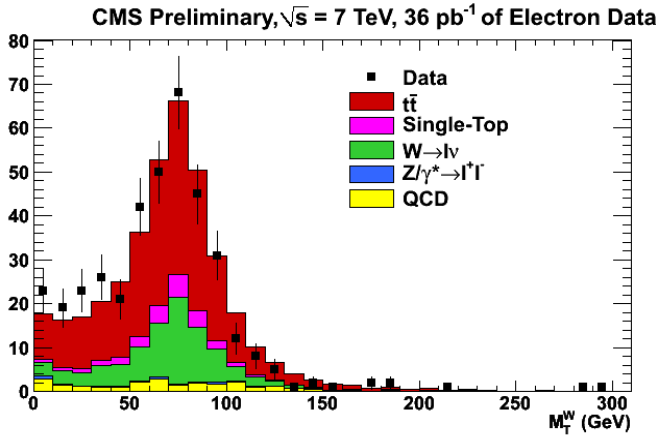


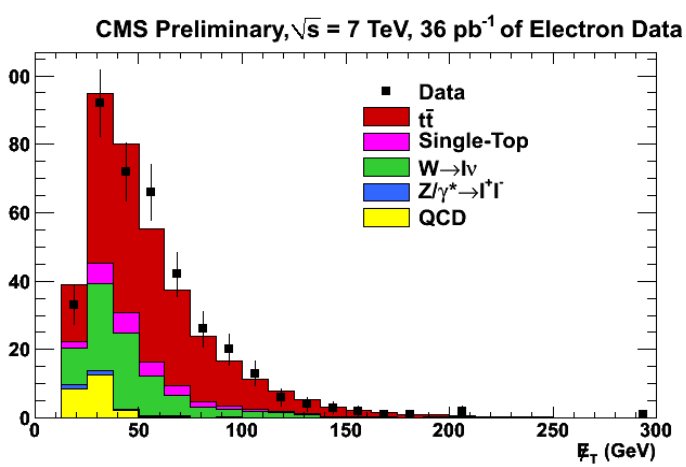


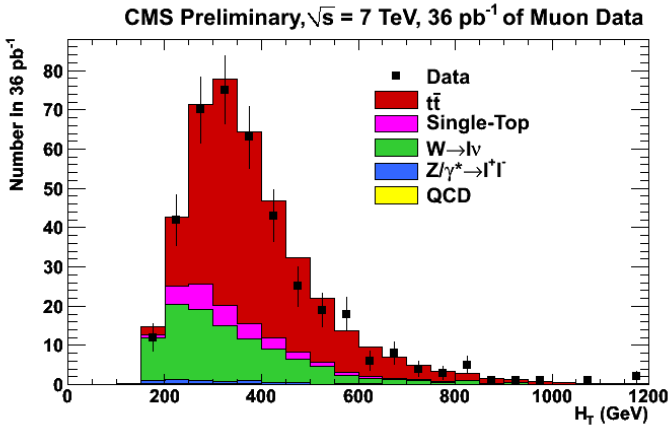


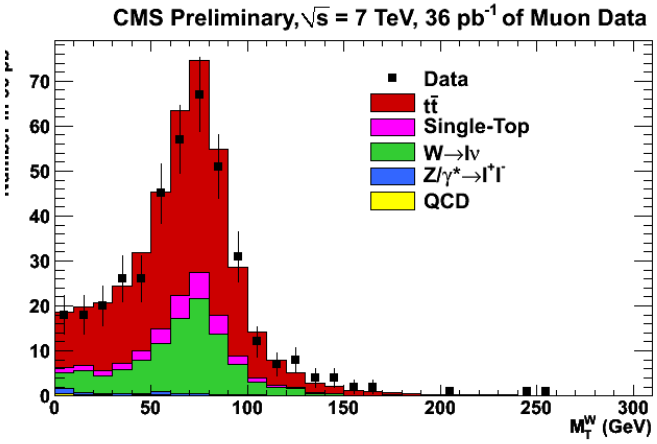
Semi-leptonic: btag

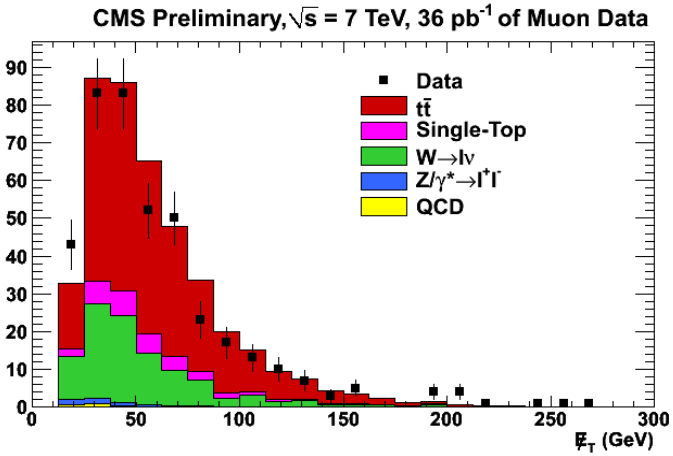








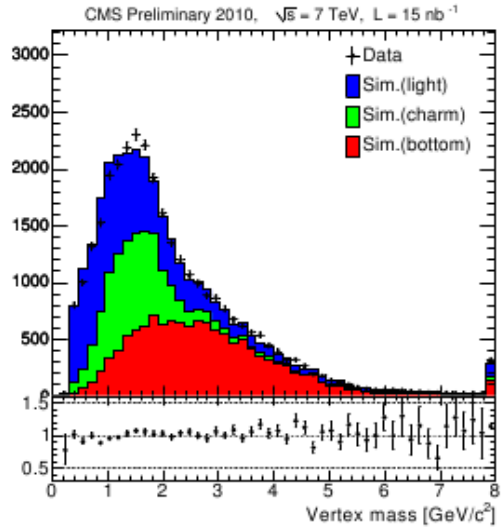


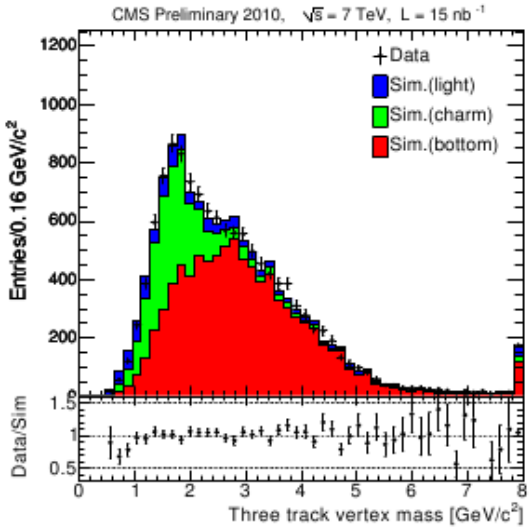


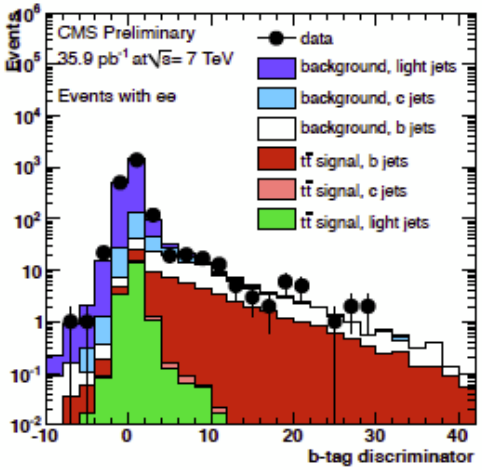
b-tag

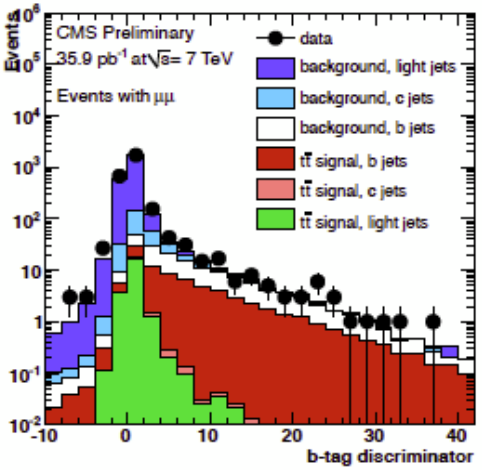
Secondary Vertex (BVT-10-001)

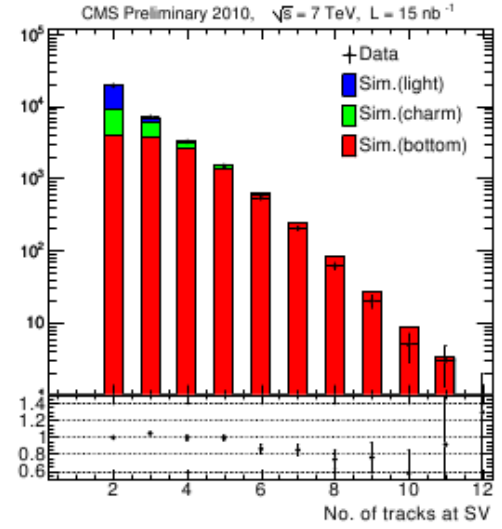
Track Counting: at least 1 jet and 2 oppositely charged leptons, from simulation

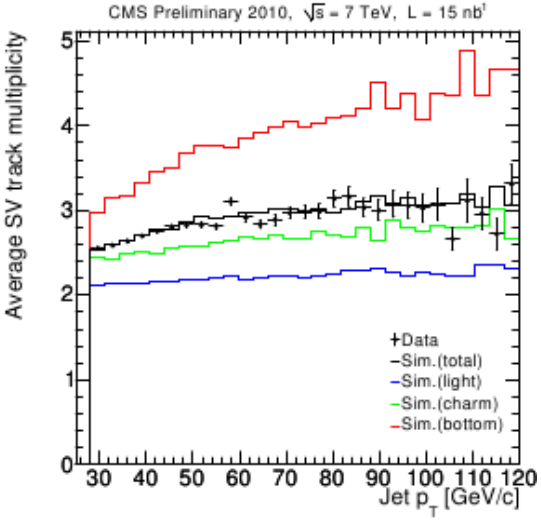


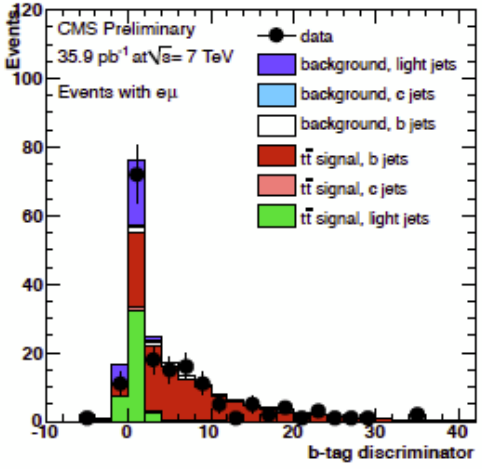


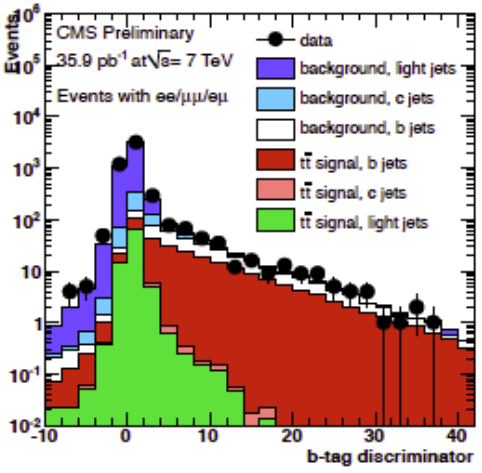








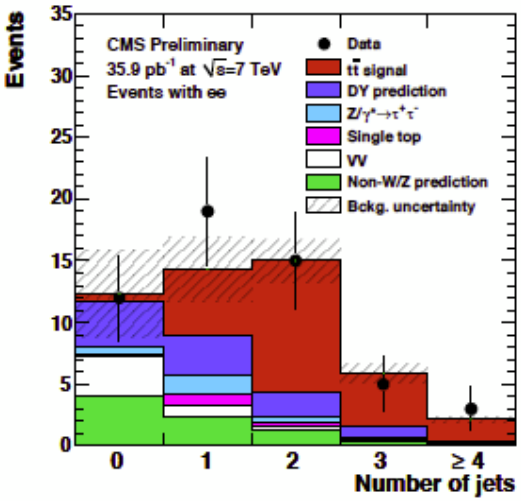


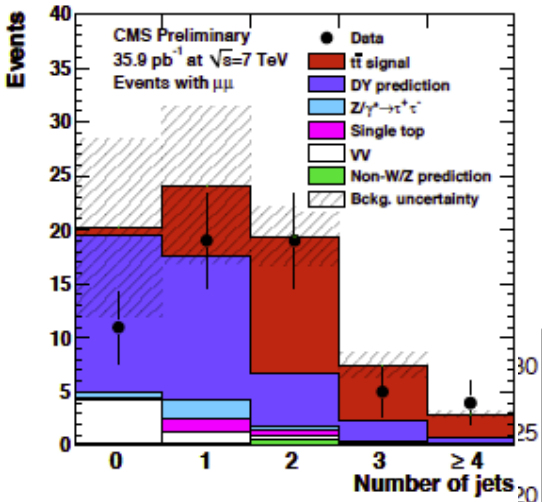


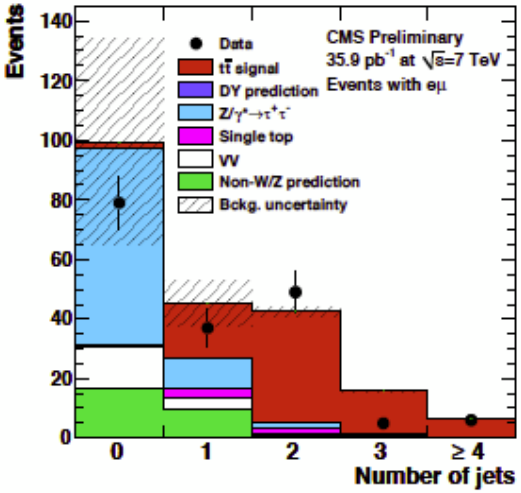
From data: SF in the fit obtained from QCD di-jet events

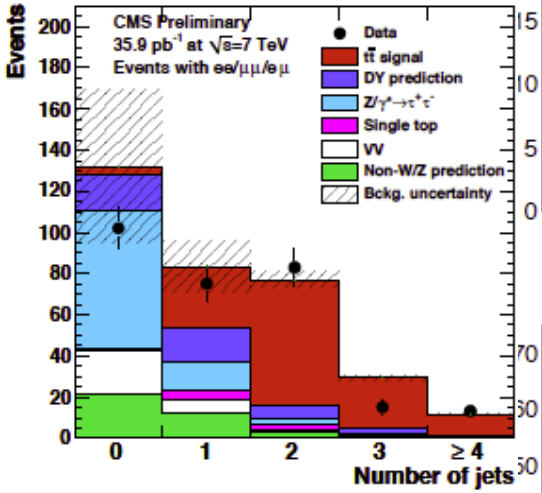
From data: SF extracted from 2tags bin / 1 tag bin, corrected for non-top BG

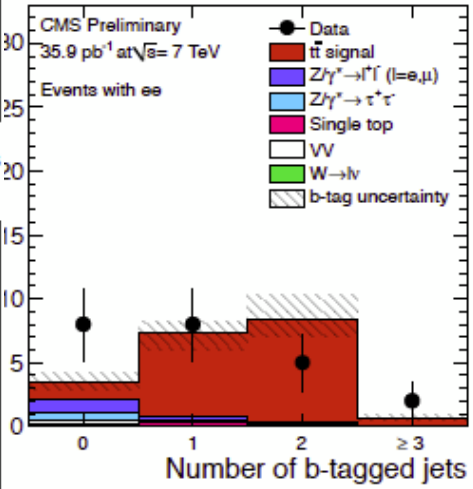
Dileptonic

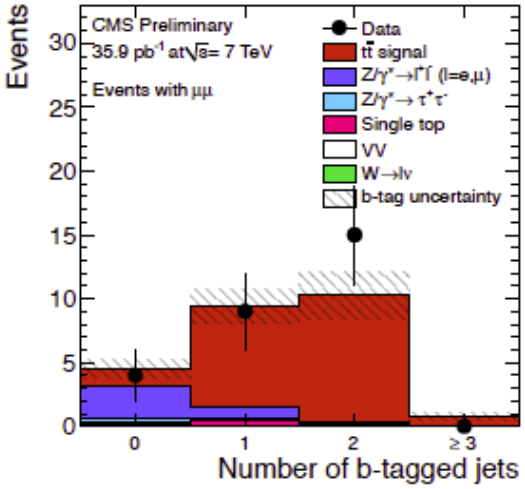

CMS Preliminary
35.9 pb-1 at √s=7 TeV
Events with ee
Data
tt̄ signal
DY prediction
Z/γ*→τ+τ-
Single top
VV
Non-W/Z prediction
Bckg. uncertainty
Events
Number of jets

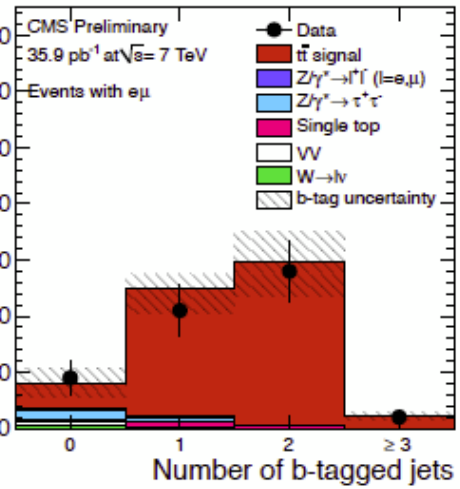

CMS Preliminary
35.9 pb-1 at √s=7 TeV
Events with μμ
Data
tt̄ signal
DY prediction
Z/γ*→τ+τ-
Single top
VV
Non-W/Z prediction
Bckg. uncertainty
Events
Number of jets

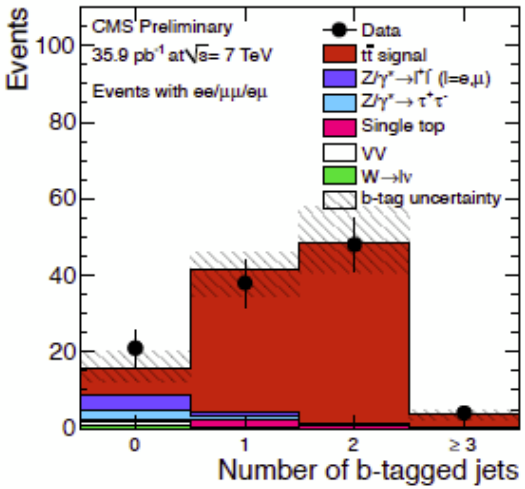

CMS Preliminary
35.9 pb-1 at √s=7 TeV
Events with eμ
Data
tt̄ signal
DY prediction
Z/γ*→τ+τ-
Single top
VV
Non-W/Z prediction
Bckg. uncertainty
Events
Number of jets


CMS Preliminary
35.9 pb-1 at √s=7 TeV
Events with ee/μμ/eμ
Data
tt̄ signal
DY prediction
Z/γ*→τ+τ-
Single top
VV
Non-W/Z prediction
Bckg. uncertainty
Events
Number of jets

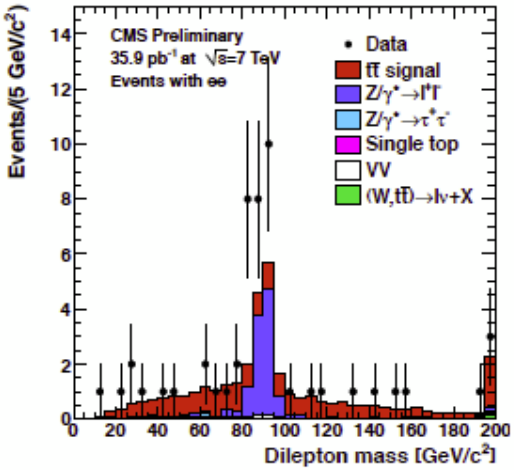

CMS Preliminary
35.9 pb-1 at √s= 7 TeV
Events with ee
Data
tt̄ signal
Z/γ*→l+l- (l=e,μ)
Z/γ*→τ+τ-
Single top
VV
W→lν
b-tag uncertainty
Number of b-tagged jets

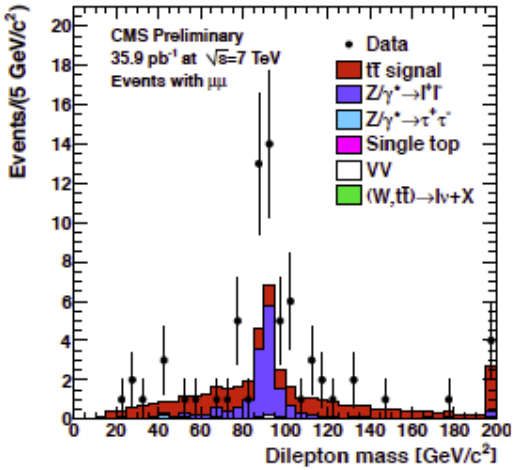

CMS Preliminary
35.9 pb-1 at √s= 7 TeV
Events with μμ
Data
tt̄ signal
Z/γ*→l+l- (l=e,μ)
Z/γ*→τ+τ-
Single top
VV
W→lν
b-tag uncertainty
Events
Number of b-tagged jets

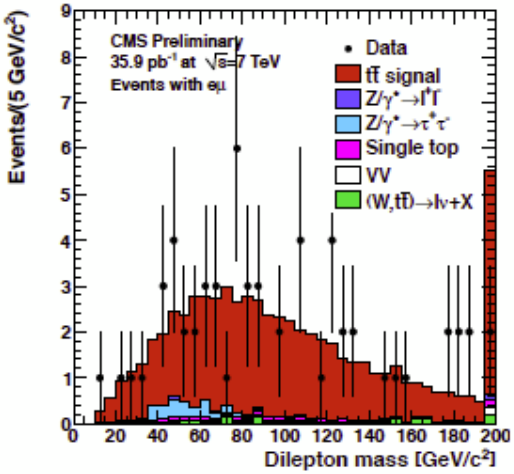

CMS Preliminary
35.9 pb-1 at √s= 7 TeV
Events with eμ
Data
tt̄ signal
Z/γ*→l+l- (l=e,μ)
Z/γ*→τ+τ-
Single top
VV
W→lν
b-tag uncertainty
Number of b-tagged jets

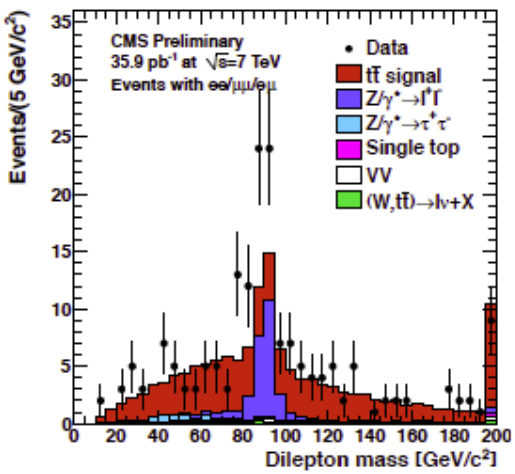

CMS Preliminary
35.9 pb-1 at √s= 7 TeV
Events with ee/μμ/eμ
Data
tt̄ signal
Z/γ*→l+l- (l=e,μ)
Z/γ*→τ+τ-
Single top
VV
W→lν
b-tag uncertainty
Events
Number of b-tagged jets

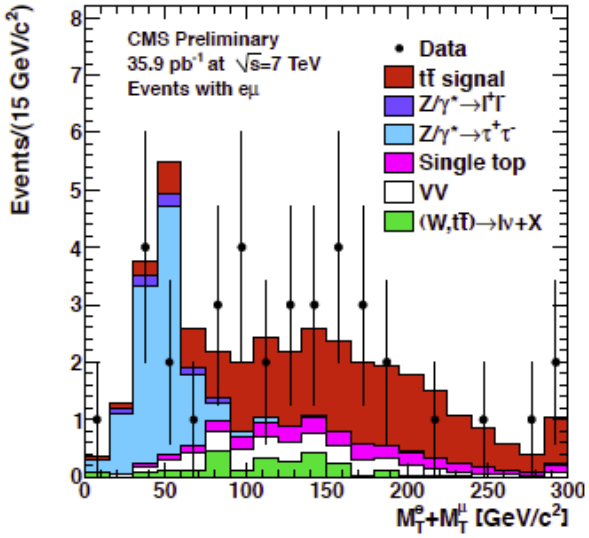
onic

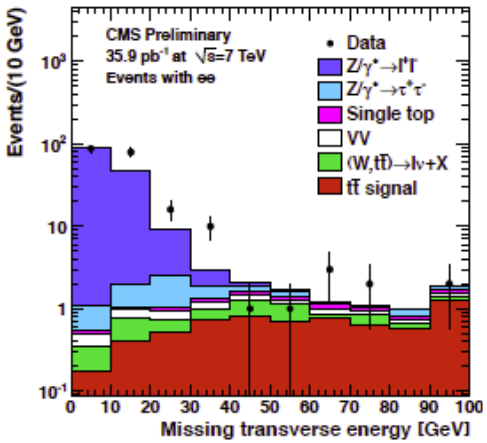

CMS Preliminary
35.9 pb⁻¹ at √s=7 TeV
Events with ee
Data
tt̄ signal
Z/γ*→l⁺l⁻
Z/γ*→τ⁺τ⁻
Single top
VV
(W,tt̄)→lν+X
Events/(5 GeV/c²)
Dilepton mass [GeV/c²]

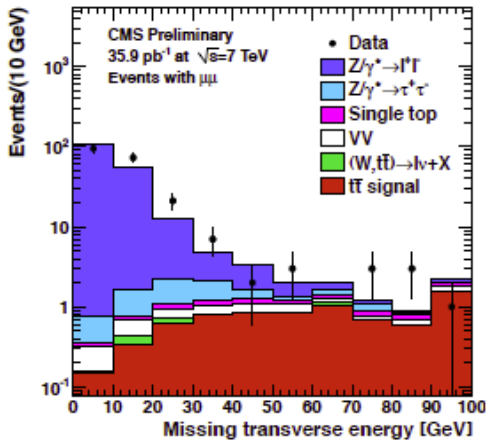

CMS Preliminary
35.9 pb⁻¹ at √s=7 TeV
Events with μμ
Data
tt̄ signal
Z/γ*→l⁺l⁻
Z/γ*→τ⁺τ⁻
Single top
VV
(W,tt̄)→lν+X
Events/(5 GeV/c²)
Dilepton mass [GeV/c²]


CMS Preliminary
35.9 pb⁻¹ at √s=7 TeV
Events with eμ
Data
tt̄ signal
Z/γ*→l⁺l⁻
Z/γ*→τ⁺τ⁻
Single top
VV
(W,tt̄)→lν+X
Events/(5 GeV/c²)
Dilepton mass [GeV/c²]


CMS Preliminary
35.9 pb⁻¹ at √s=7 TeV
Events with ee/μμ/eμ
Data
tt̄ signal
Z/γ*→l⁺l⁻
Z/γ*→τ⁺τ⁻
Single top
VV
(W,tt̄)→lν+X
Events/(5 GeV/c²)
Dilepton mass [GeV/c²]


CMS Preliminary
35.9 pb⁻¹ at √s=7 TeV
Events with eμ
Data
tt̄ signal
Z/γ*→l⁺l⁻
Z/γ*→τ⁺τ⁻
Single top
VV
(W,tt̄)→lν+X
Events/(15 GeV/c²)
M_T^e+M_T^μ [GeV/c²]


CMS Preliminary
35.9 pb⁻¹ at √s=7 TeV
Events with ee
Data
Z/γ*→l⁺l⁻
Z/γ*→τ⁺τ⁻
Single top
VV
(W,tt̄)→lν+X
tt̄ signal
Events/(10 GeV)
Missing transverse energy [GeV]


CMS Preliminary
35.9 pb⁻¹ at √s=7 TeV
Events with μμ
Data
Z/γ*→l⁺l⁻
Z/γ*→τ⁺τ⁻
Single top
VV
(W,tt̄)→lν+X
tt̄ signal
Events/(10 GeV)
Missing transverse energy [GeV]

BLUE Method

The combined measurement $\hat{m}$, calculated from individual measurements m_i, is given by:

$$\hat{m} = \frac{\sum_{ij} H_{ij} m_j}{\sum_{ij} H_{ij}},$$

With **H** being the inverted error matrix E. The latest is given by:

$$\mathbf{E} = \sum_{k=1}^{K} E^{(k)},$$

with the contribution from the systematic source k being:

$$E^{(k)} = r_{ij}^{(k)} \delta_i^{(k)} \delta_j^{(k)},$$

with r_{ij}^k the correlation factors (equal to 1 for $i = j$) and δ_i^k the uncertainties.

For each measurement, a weight can be calculated such that

$$\hat{m} = \sum_i w_i m_i,$$

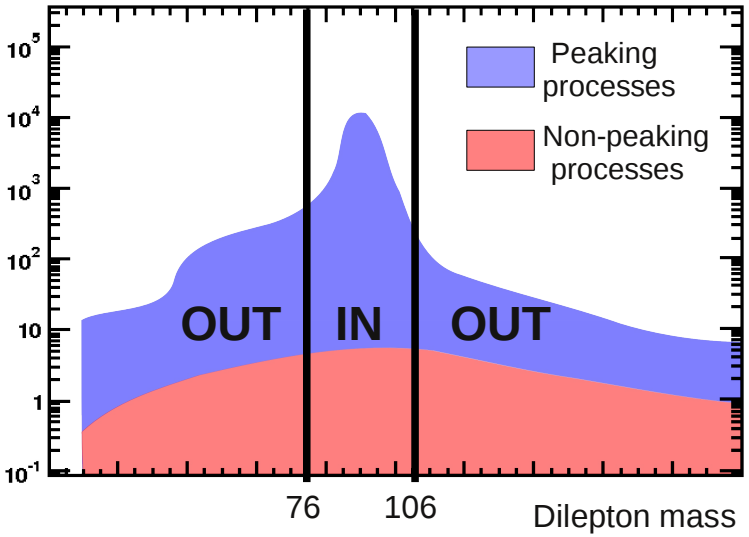
with

$$w_i = \frac{w_i^{raw}}{\sum_i w_i^{raw}}, \quad w_i^{raw} = \frac{\sum_j H_{ij}}{\sum_{ij} H_{ij}}.$$

More details in:
L. Lyons, D. Gibaut, and P. Clifford, "How to combine correlated estimates of a single physical quantity"
Nucl. Instrum. Meth. A270 (1988) 110
doi:10.1016/0168-9002(88)90018-6

Dileptonic: data-driven methods

1) Drell-Yan :



Estimation of the contribution outside the Z-veto region using the number of events inside the Z-veto region

$$N_{out}^{e^+e^-,exp} = \underbrace{R_{out/in}^{e^+e^-}}_{\text{from simulation}} \left(N_{in}^{e^+e^-} - \underbrace{0.5 N_{in}^{e^\pm\mu^\mp} k_{ee}}_{\text{non-peaking processes}} \right)$$

from simulation,
$R_{out/in} = N_{out}(DY\ MC) / N_{in}(DY\ MC)$

non-peaking processes from eμ channel data corrected for lepton eff ($k_{\ell\ell}$)

Similar for the μμ channel

Done with and w/o b-tag / Pile-up, for != jet multiplicity, $\not{E}_T$

2) Non W/Z decays lepton ("fakes") :

To estimate the ratio of Non W/Z decays leptons passing the standard lepton selection (FR), an inclusive QCD sample has been used.

FR (p_T, η) = n / (n + $\bar{n}$) where n is the number of leptons passing the tight selection

$\bar{n}$ is the number of leptons passing the loose selection and not the tight

2 "fakes":

$$N_{nn}^{QCD} = \sum_{i,j} \frac{FR_i FR_j}{(1-FR_i)(1-FR_j)} N_{\bar{n}\bar{n}}^{ij}$$

1 "fake" + 1 from W/Z boson (corrected from QCD/signal spillage):

$$N_{nn}^{Wj} = \sum_{i,j} \frac{FR_j}{(1-FR_j)} N_{n\bar{n}}^{ij} - 2N_{nn}^{QCD} - \Delta_{signal}$$

Done with and w/o b-tag, for != jet multiplicity, $\not{E}_T$,and tested for != loose selection in FR definition