Measuring the Top-Quark Event Properties at the LHC with CMS

A “Shopping List” with a particular Focus on Studies/Strategies with the Early Data

Tim Christiansen (CERN)

LHC Physics Workshop
Mumbai, India
21–27 October 2009

Outline

- Introduction
  - Top-Quark Physics: from the Tevatron to the LHC
  - First Pieces for first Data
- Top-Quark Pair Production
  - Emphasis on startup
  - Early Strategies for first Cross-Section Measurements
  - Brief Outlook onto 2nd-Generation Analyses
- Single Top-Quark Production
- Beyond the X-Section Measurements
  - m(ttbar) Spectra
  - Heavy-Flavour Content of ttbar Samples

} → En route to New Physics …

Brief Introduction:
Top Quarks and LHC's Perspective

The Top Quark

- The top quark (*t*) is by far the heaviest elementary particle known in the Standard Model (SM)
  - $m(\text{top}) = (171 \pm 1.1^{\text{stat.}} \pm 1.2^{\text{syst.}})\ \text{GeV}/c^2$ (Tevatron)
    - About 40× heavier than the second most massive quark (*b*)
    - About 2x as massive as the heaviest known Boson (*Z*)
- Weak-Isospin partner of the bottom quark (*b*)
  - Spin S=1/2, Ladung $Q = -2/3 \cdot Q(e)$, $I_3 = 1/2$
- Completes the SM picture of quarks and leptons:
  - The third generation

The 3 fermion generations:

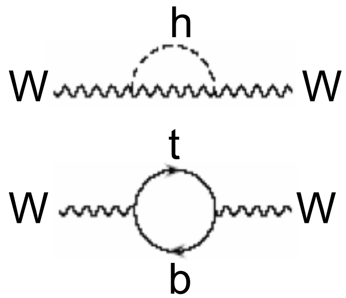
Quarks: $\begin{pmatrix} u \\ d \end{pmatrix} \quad \begin{pmatrix} c \\ s \end{pmatrix} \quad \begin{pmatrix} t \\ b \end{pmatrix}$

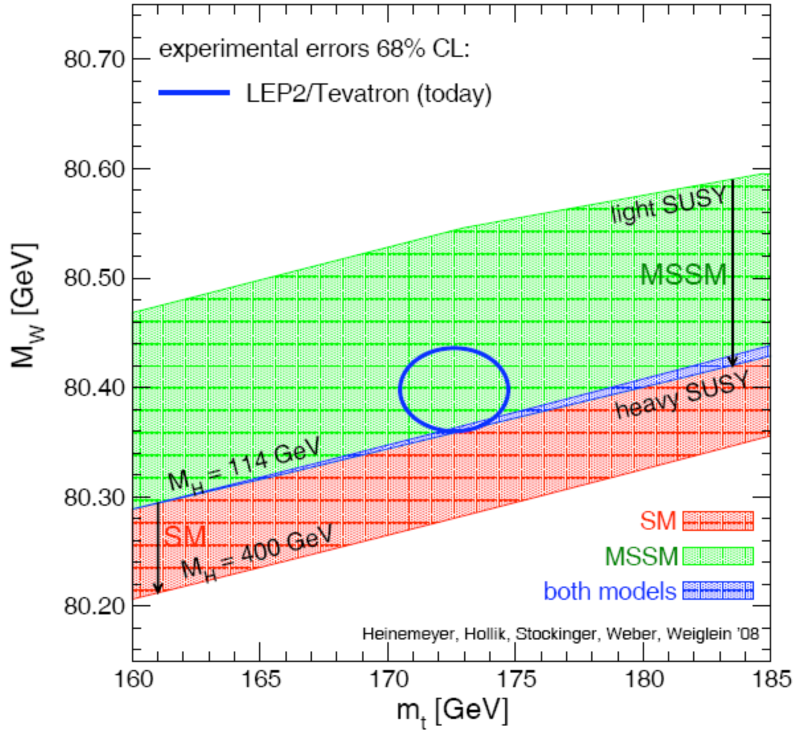
Leptons: $\begin{pmatrix} \nu_e \\ e \end{pmatrix} \quad \begin{pmatrix} \nu_\mu \\ \mu \end{pmatrix} \quad \begin{pmatrix} \nu_\tau \\ \tau \end{pmatrix}$

Interesting Physics with Top Quarks

- Mass
  - Very difficult measurement
  - Interesting not just per se, but as big component to EWK fits

    Virtual loops contributing to m_W



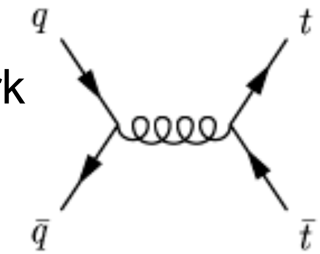


- Kinematical properties
  - Is there $X \rightarrow tt$?
  - W polarization
    - Fraction of longitudinally polarized Ws in top decays is ~70% according to SM
  - Spin correlations in production
- Rare decays
- Single-top
- More generally: the top quark is unusually heavy, maybe there is something different about it?

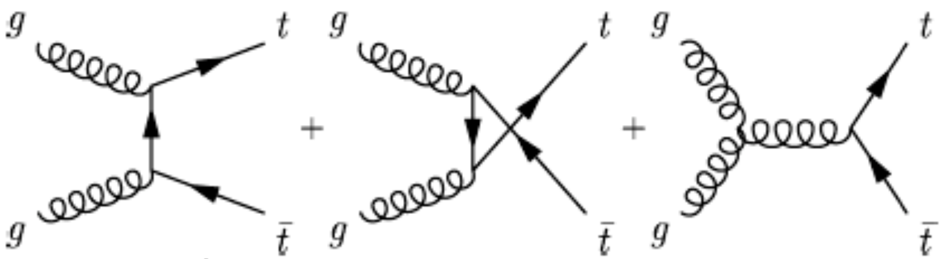
What we know about the Top Quark

- Practically everything that we (directly) know about the top quark comes from the Tevatron experiments DØ & CDF:
  - Tevatron: proton-antiproton collisions at nearly 2 TeV
  - Discovered 1994-5 with just “a handful” of candidates
  - Only few hundreds of top-quark events have been studied to date
- Production: predominately in top-antitop pairs (ttbar)

Quark-antiquark annihilation



← Dominant at the Tevatron

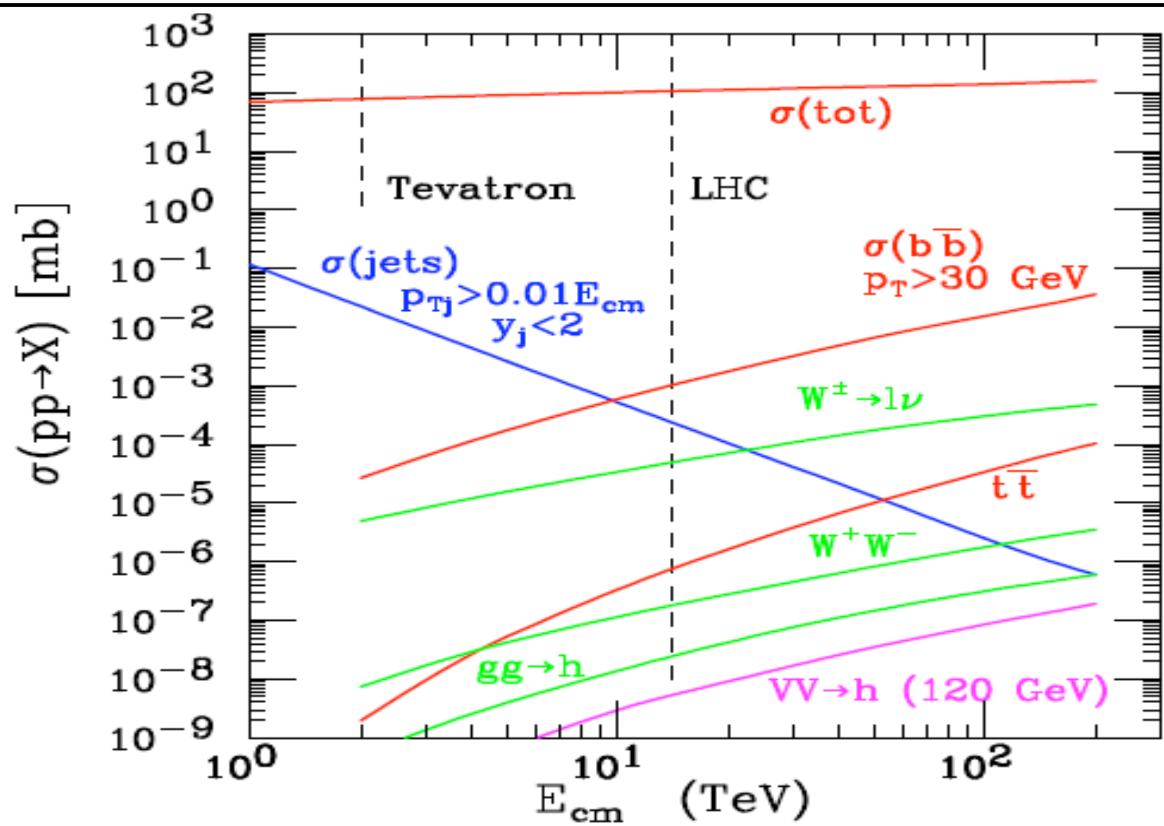


← Dominant at the LHC

Gluon-gluon processes

The Top Quark & the LHC

ttbar production cross sections as a function of collision energy √*s*

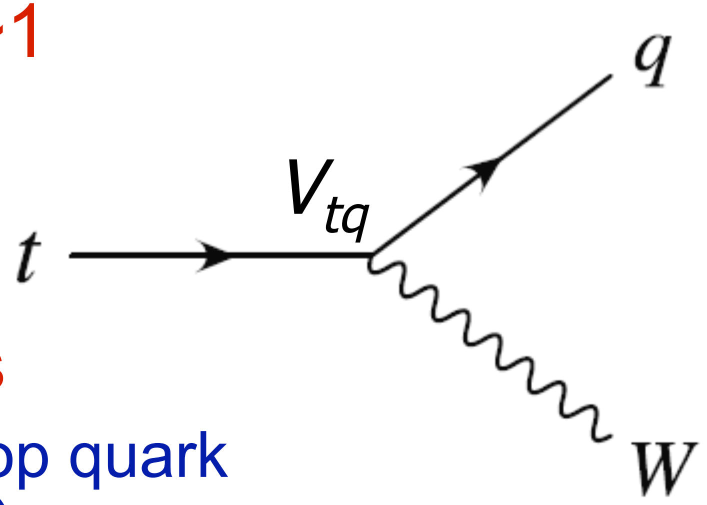


- At √*s* = 14 TeV, about 100× larger x-section* + much more luminosity than at the Tevatron:
  - The LHC is a **Top-Quark Factory**

* At √*s* = 10 TeV (7 TeV) still about 50× (20×) larger x-section (“LHC in 2010”)

Top-Quark Decay

- The CKM matrix element V_{tb}~1
  - t decays almost exclusively into a W boson and a b quark
- Top lifetime is so short, that it decays before it hadronizes
  - i.e. there is no “T meson”, the top quark decays freely (free quark decay)

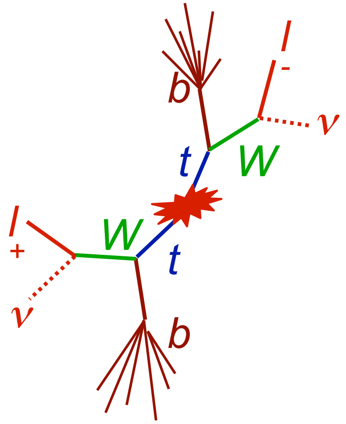


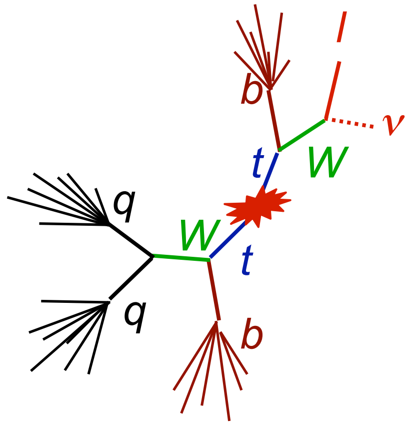
- The W boson is real, on-shell ($m_t > 2 \times m_W$)
  - W-decay $W^{\pm} \rightarrow \ell^{\pm}\nu$ (ℓ = lepton, i.e., $e^{\pm}$, $\mu^{\pm}$ oder $\tau^{\pm}$)
    - Branching ratio BR ~ 1/9 per lepton flavor
  - W-decay $W^{\pm} \rightarrow qq'$ (qq' = light quarks!)
    - BR ~ 2/3

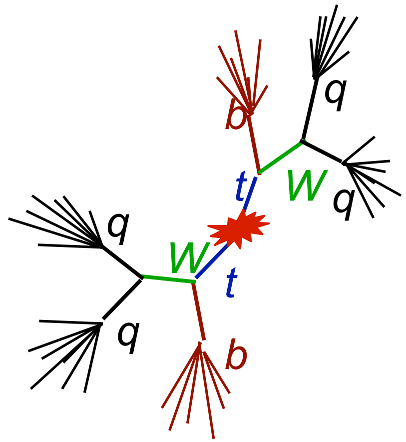
The 3 main Channels of ttbar Events

$$t\bar{t} \rightarrow W^+ b W^- \bar{b} \rightarrow \ldots$$

all-leptonic (dilepton)	semi-leptonic (l+jets)	all-hadronic





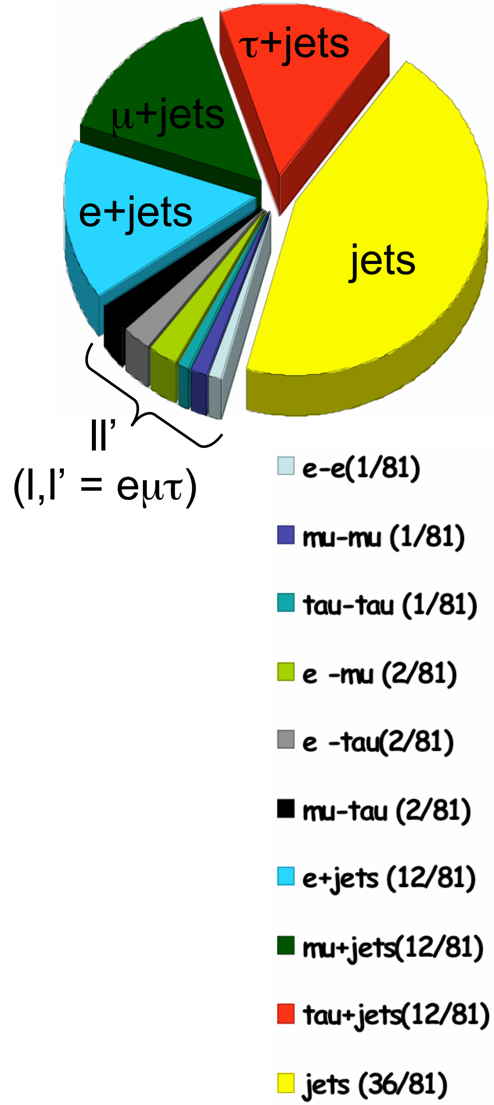


$$\rightarrow \bar{l}^+ \nu_l b l^- \bar{\nu}_l \bar{b}$$

$$\rightarrow q\bar{q}' b l^- \bar{\nu}_l \bar{b} + \bar{l}^+ \nu_l b q \bar{q}' \bar{b}$$

$$\rightarrow q\bar{q}' b q'' \bar{q}''' \bar{b}$$

10.3 % **43.5 %** **46.2 %**



Top at CMS

- The top-physics programme has distinct phases:
  - Establish tt at CMS
  - Basic studies of tt:
    - Is the selected ttbar sample consistent with ttbar hypothesis?
      - By construction, this is a broad search for BSM physics
      - Compare kinematical distributions with ttbar simulation (p_T, H_T, η, multiplicity, MET, p_T(tt), m(tt), angles, …)
        - Discrepancies? ⇒ compare different channels, take ratios for cancellation of systematics, …
        - N.B.: **need event selections that facilitate such comparisons**!
      - σ(tt) comes out of this programme
  - Top as calibration tool
    - The rich ttbar signature provides m_t & m_W (t & W are <u>on-shell</u> in ttbar→WbWb!) → constrained fits → jet-energy calibration!
    - Sel. w/o b-tagging → indep. measurement of $\varepsilon_{\text{b-tagging}}$
    - Test bench for missing transverse energy

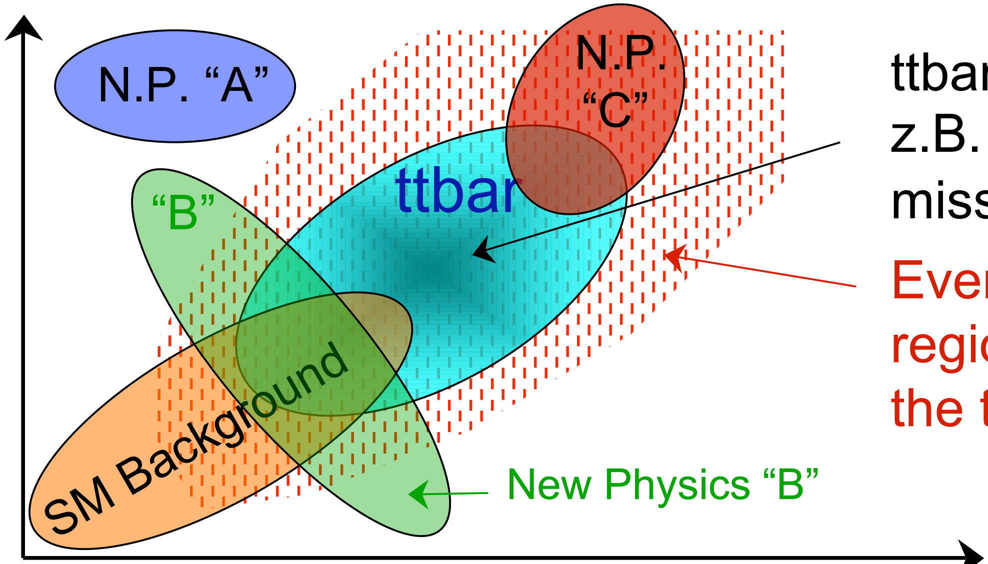
    } → Talk by J. D'Hondt
  - Detailed studies, single-top, etc. …

Establishing ttbar at CMS

- The mission of the LHC is the search for (and the study of) New Physics!
  - Already the first top studies should be seen in this light
    - The most exotic of the known particles (mass, free decay …)
      - Rich signature: ttbar requires the reconstruction of
        - All kinds of leptons (e,μ,τ)
        - Hadronic jets (from heavy and light quarks)
        - Missing transverse energy (neutrino(s))
        - *b*-tagging: identification of jets from heavy quarks
  - Demonstration of our understanding of the Detectors and of the Standard Model (SM)
  - Measure the ttbar cross-section
  - Measure kinematical properties
  - Some of those measurements can be seen as searches themselves:
    - There can be New Physics in the ttbar sample

Top Signals & New Physics?

- ttbar events live in a complex multidimensional region of the section space



ttbar "core region",
z.B. 1μ+4jets, 2 b-tags,
missing transverse Energy $\not{E}_T$

Event selection (incl. SM-Control regions, i.e., it is not limited to the ttbar core region)

- SM background (W+jets, QCD-multijets, Z+jets, …) lives in a different region, overlapping with ttbar
- Simplified: New Physics
  - may be completely disjoint from ttbar signal (A)
  - may have significant overlap with SM (and ttbar) (B)
  - may have an effect on the tail of the ttbar distribution (C)
- More direct: New Physics may also manifest themselves directly in the production (z.B. $X \rightarrow$ttbar) or in the top decay

Few Words on ttbar Challenges …

- The event selection includes a large SM control region
  - because the ttbar signature is not as clean as a resonance, for example
- We will need to demonstrate our understanding of the control regions before moving on to the ttbar measurement itself
- Usually, all SM non-ttbar backgrounds have lower jet multiplicities than ttbar

  $\Rightarrow$ ttbar analysis is an analysis as a function of N_{jets}, with and without b-tagging

Few Words on ttbar Challenges … 2

σ_{LHC}(ttbar) ≈ 100 x $\sigma_{Tevatron}$(ttbar) (√s = 14 TeV), so what's so hard about it?! …

- Compared to √s, the top it is not that heavy anymore…
  - QCD (multi-jet events) in l+jets is a challenge
    - One fake lepton or one b→l + many jets
  - At the Tevatron this is almost eliminated by MET cut of order 25 GeV
  - The typical MET (ν) of a ttbar event is 40 GeV @ LHC
  - At CMS, QCD events with 4 jets of E_T~30 GeV have typical reconstructed MET simlilar to the MET in ttbar!
  - Instead:
    - at the moment rely on high thresholds for jets and lepton p_T and very tight isolation (will loose quite a bit of efficiency)
    - find data-driven methods to measure QCD contribution (ongoing)
- ttbar goes up by x 100, but inclusive W only by x10 compared to Tevatron, that's great, no?

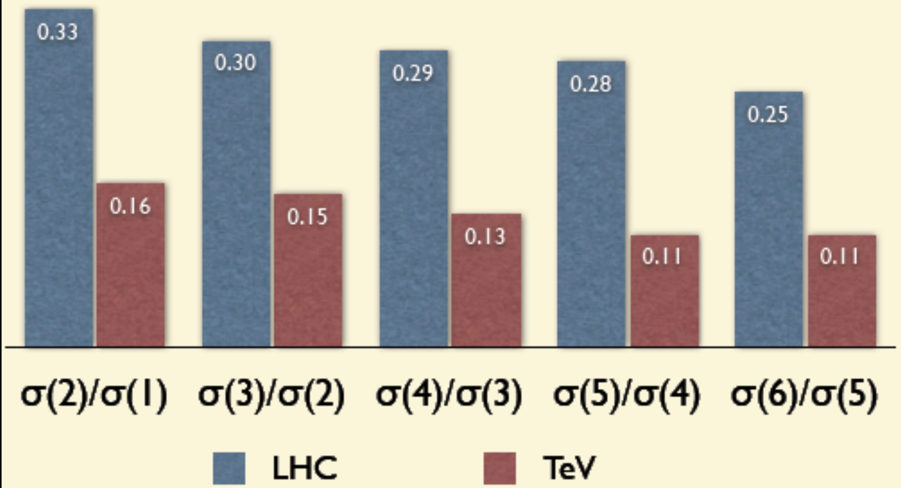
W+4 jets still dominant BG in l+jets!

(slide stolen form Michelangolo Mangano)

W+Multijet rates

σxB(W→eν)[pb]	N jet=1	N jet=2	N jet=3	N jet=4	N jet=5	N jet=6
LHC	3400	1130	340	100	28	7
Tevatron	230	37	5.7	0.75	0.08	0.009

E_T(jets) > 20 GeV , |η|<2.5 , ΔR>0.7

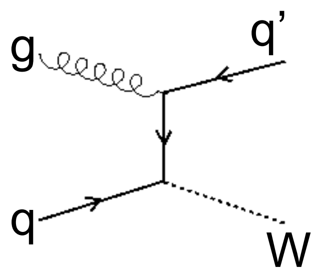


√s = 14 TeV

- Ratios almost constant over a large range of multiplicities
- O(α_s) at Tevatron, but much bigger at LHC

29

Lumi(*gq*) is high at LHC:



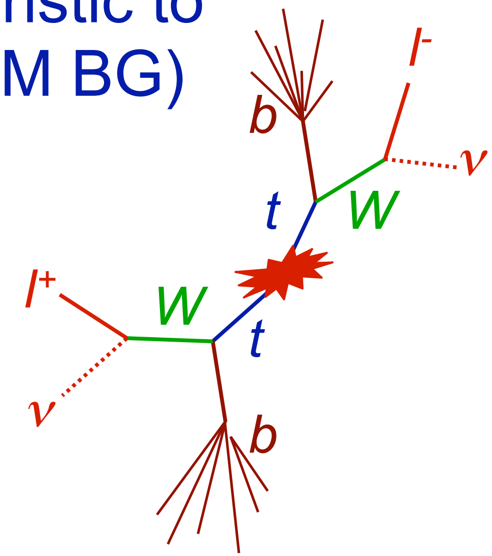
W+jets from higher orders

σ(W inclusive) increases by x10
σ(W+4 jets) increases by x100

First Signals & Cross Section in the Di-Lepton Channel

First ttbar Cross Section in the <u>Dilepton Channel</u>

- CMS Physics Analysis Summary TOP-09-002
  http://cms-physics.web.cern.ch/cms-physics/public/TOP-09-002-pas.pdf
- Analysis strategy:
  - Simple counting experiment
  - *N*(cand) after a selection that is characteristic to tt→lνblνb is compared to expected *N*(SM BG)
  - Any excess is ascribed to *N*(ttbar)
  - *N*(SM background) estimated from ttbar-depleted control samples



Signature: 2 prompt, "high"-p_T leptons, 2 (b-)jets, MET>0

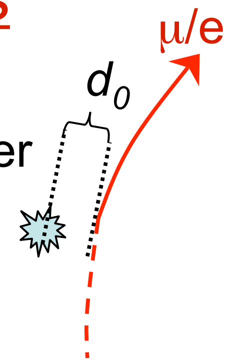
Background: Drell-Yan ($Z/\gamma^* \rightarrow l^+l^-$) + jets, W+jets, QCD multijet events (w/ fake prompt leptons), single-top

} DY+jets, W+jets, and QCD BG is hard to correctly simulate!

Dilpetons

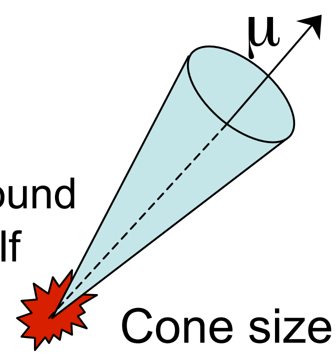
- Selection: 3 Channels <u>ee+jets</u>, <u>μμ+jets</u> and <u>eμ+jets</u>
  - Trigger:
    - single-μ p_T > 9 GeV for μμ
    - single-e p_T > 15 GeV for ee
    - .OR. of the above for eμ

Impact parameter
⊥ beam axis:



  - ≥ 2 isolated leptons (e or μ) with p_T > 20 GeV, $|\eta|$ < 2.4
    - Electrons: electron-ID, dedicated photon conversion removal, loose cut on the 2D impact parameter (IP) to select <u>prompt</u> electrons
    - Muons: Tracks in the muon system associated (and fitted together) with central tracks, muon-ID, loose IP cut
    - Isolation: $p_T(\mu/e) / [\sum_{\Delta R<0.3/4} E_T^{calo} + p_T(\mu/e)] > 0.9/0.8$,
      $p_T(\mu/e) / [\sum_{\Delta R<0.3} p_T^{track} + p_T(\mu/e)] > 0.9$

$\sum_{\Delta R<X}$ sums over track momenta (calo deposits) in the cone around the lepton, but excluding the track (energy) of the lepton itself

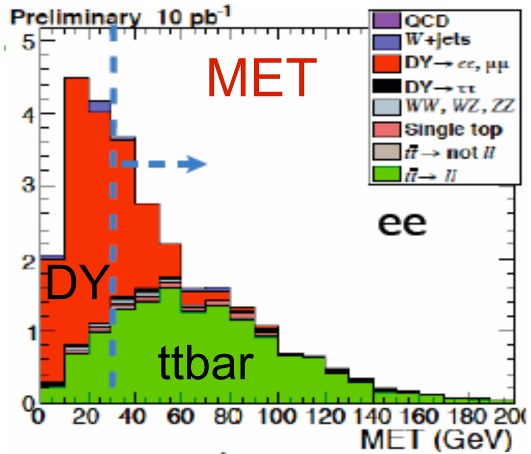


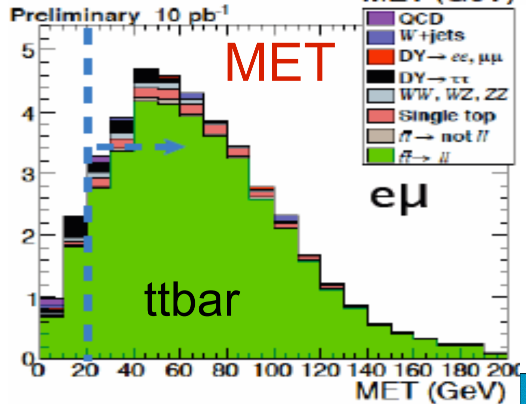
Cone size
$\Delta R = \sqrt{(\Delta\phi^2+\Delta\eta^2)}$

  - Opposite sign of the 2 lepton charges
  - Only 1 lepton pair per event:
    - In case of >2 of such leptons in the event, unambiguously associate them with 1 channel (sort by p_T & quality/iso) → no overlap

Dilepton Selection continued

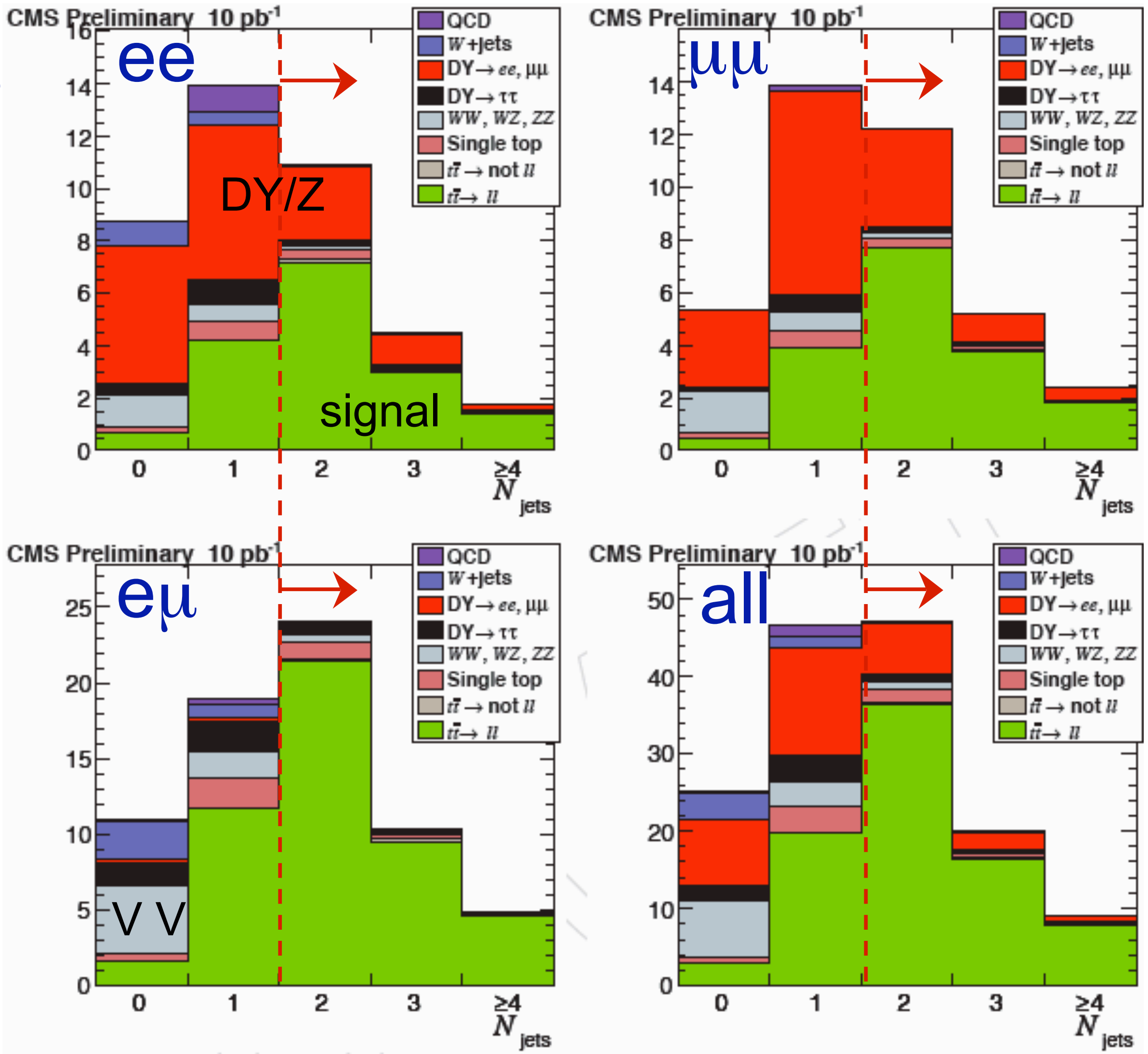
- Jets: R=0.5 SisCone*, calibrated (no flavor correction)
  $p_T > 30$ GeV, $|\eta| < 2.4$ *Seedless Infrared-safe Cone Algorithm
  - Discard jets closer than $\Delta R = 0.3$ from electron candidate
  - "default" is calo-jets; possible alternative: track-jets (studied also) or track-corr. jets, Particle Flow …
  - Ask for $N_{jet} \geq 2$
- *Z*-veto for $\mu\mu$, ee channels:
  - $|m(ll)-M_Z| < 15$ GeV and
  - missing-E_T:
    - MET (corrected for jets and muons)
    - MET > 20 (30) GeV for eμ (ee & $\mu\mu$)
- N(jet) ≥ 2 is the signal sample
- N(jet) = 0 or 1 as control sample





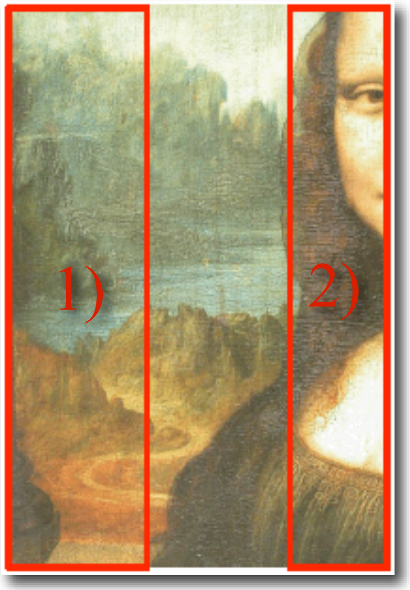
Dileptons

- Event yields:
  - Jet multiplicity for the 3 channels & combined



Background Estimation

- BUT: Even though we may qualitatively understand the background distributions, their <u>normalization</u> (= effective cross-section after all cuts) has <u>LARGE uncertainties</u> if taken from simulation only!
  - Estimate main BG sources from the data themselves:



Strategy:

1) Study how a background-dominated region looks
2) Use this knowledge to estimate the background under the signal

Background Determination

- Data-driven method for Drell-Yan (DY) estimation:
  - Method to predict DY→ee,μμ (not ττ)

$$N_{DY}^{out\ (est)} = \frac{N_{DY\ DATA}^{in}}{N_{DY\ MC}^{in}} \cdot N_{DY\ MC}^{out}$$

Estimate DY outside of Z-veto region (near m_Z) by counting events inside, take $R_{out/in} = N^{out} / N^{in}$ from simulation

  - Subtract non-DY contribution near m_Z using eμ data:

$$N_{DY}^{out\ (est)} = (N_{ll\ DATA}^{in} - \overbrace{k \cdot N_{e\mu\ DATA}^{in}}) \cdot R_{out/in} \qquad k_{\mu\mu} = \frac{1}{2}\sqrt{\frac{n_{\mu\mu}^{obs}}{n_{ee}^{obs}}}$$

k_{ee}= 1/$k_{\mu\mu}$

  - Main assumption: Trust MC prediction of $R_{out/in}$ more than MC prediction of σ, N(jet), or MET

Final State	nJets	$R_{out/in}$	$N_{out\ (est)}$	$N_{DY}^{out\ (true)} + N_{ZZ}^{out\ (true)}$
ee	≥ 2	0.100 ± 0.010	4.23 ± 0.41	4.03 ± 0.37
$\mu\mu$	≥ 2	0.105 ± 0.009	5.33 ± 0.46	5.13 ± 0.4

Background from Fake Leptons

- Here, "fake" means mis-identified prompt leptons originating from jets (e.g. W+jets, QCD)
- Use the fake rate FR:
  - FR = N(fake l passes all cuts) / N(fake l passes loose ID/isolation)
  - loose ≡ loosely isolated, loose lepton-ID, no IP cut
  - Measure FR=FR(p_T,$|\eta|$) from inclusive multi-jet sample (here ~ all leptons are "fakes")
  - Prediction: N(evts w/ loose e/μ) × $FR_{e/\mu}$ × (1-$FR_{e/\mu}$)

e.g.: ee

	$N_{\text{jets}} = 0$	$N_{\text{jets}} = 1$	$N_{\text{jets}} \geq 2$
ttdil (observed)	0.7 ± 0.1	4.2 ± 0.1	11.6 ± 0.2
ttdil (predicted)	0.012 ± 0.001	0.051 ± 0.003	0.112 ± 0.004
W+jets (observed)	0.9 ± 0.2	0.5 ± 0.2	0.2 ± 0.1
W+jets (predicted)	1.08 ± 0.04	0.43 ± 0.03	0.22 ± 0.02

← **spillage** (can be corrected for)

← **fakes**

Dileptons: Expected Precision & Systematics

- Generally **small backgrounds**, dominating BGs (DY for ll, part of WZ,ZZ) and fake leptons (e.g. W+jets) are under control and can be estimated from the data (single-top from simulation, take ~large uncertainty)
- Note: General flow/method (for BG estimates) independent of jet type (calo, tracks, JPT, PF) or MET (calo, track-MET)
- For 10/pb at 10 TeV:
  - lept-ID & isolation from T&P
  - varied ±10% JES
  - Theory: Comparison MG w/ Pythia & MC@NLO
  - Residual BGs (tW, part of V V, DY→ττ): assign 50%
  - Lumi-uncertainty quoted separately
- Combined:
  $\delta\sigma/\sigma \approx 15\%^{(stat.)} \pm 10\%^{(syst.)} \pm 10\%^{(lumi)}$

Event yields for 10 pb^{-1} @ 10 TeV

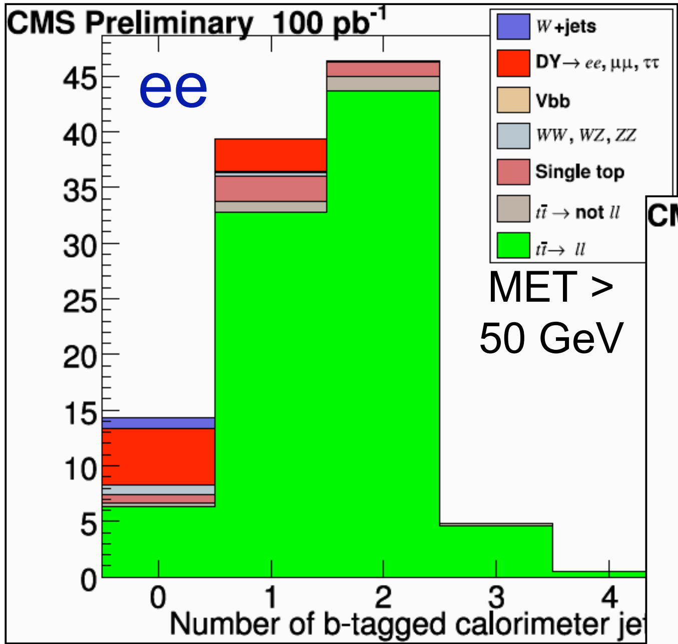
Data sample	Main selection e^+e^-	$\mu^+\mu^-$	$e^\pm\mu^\mp$
$t\bar{t} \to \ell\ell$	11.6 ± 0.2	13.2 ± 0.2	35.6 ± 0.4
other $t\bar{t}$	0.21 ± 0.03	0.04 ± 0.01	0.46 ± 0.04
Single top	0.46 ± 0.03	0.56 ± 0.03	1.40 ± 0.06
$WW/WZ/ZZ$	0.26 ± 0.02	0.33 ± 0.03	0.71 ± 0.05
DY→ $\tau\tau$ + jets	0.3 ± 0.1	0.3 ± 0.1	0.7 ± 0.2
DY→ $ee/\mu\mu$ + jets	4.1 ± 0.4	5.3 ± 0.4	0.08 ± 0.05
W + jets	0.2 ± 0.1	< 0.1	0.3 ± 0.1
QCD	< 1	< 0.4	< 0.4
Total backgrounds	5.5 ± 0.4	6.6 ± 0.4	3.7 ± 0.2
Data driven fakes	1.1 ± 0.6	0.8 ± 0.4	2.5 ± 1.2
Data driven DY	4.0 ± 1.3	5.1 ± 1.6	

Relative uncertainties (in %)

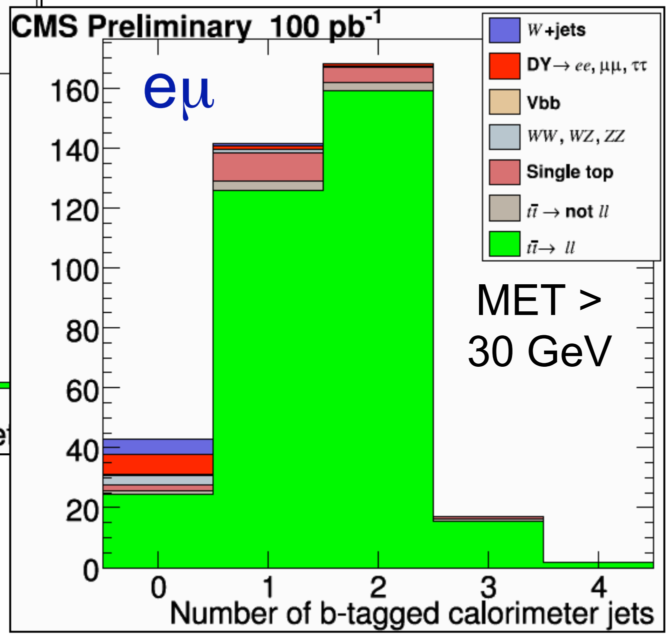
Source	e^+e^- and $\mu^+\mu^-$	$e^\pm\mu^\mp$
Statistical	25	18
Lepton ID	5	5
Lepton isolation	3	3
Jet energy scale	8	5
Theory	4	4
DY→ $ee, \mu\mu$ method	10	
Fake leptons method	4	4
Residual background	5	4
Integrated luminosity	10	10

Dilepton Channel ++

- So what comes next?
  - Is the selected sample really the SM ttbar? First handle: Heavy-flavor content (expect ~2 b-jets) with b-tagging



Requires an <u>independent</u> measurement of ε(b-tag) and fake b-tag rate ε(fake)



simple tagger example:

≥ 2 of the tracks associated with the jet have a significant displacement from the primary vertex of the event

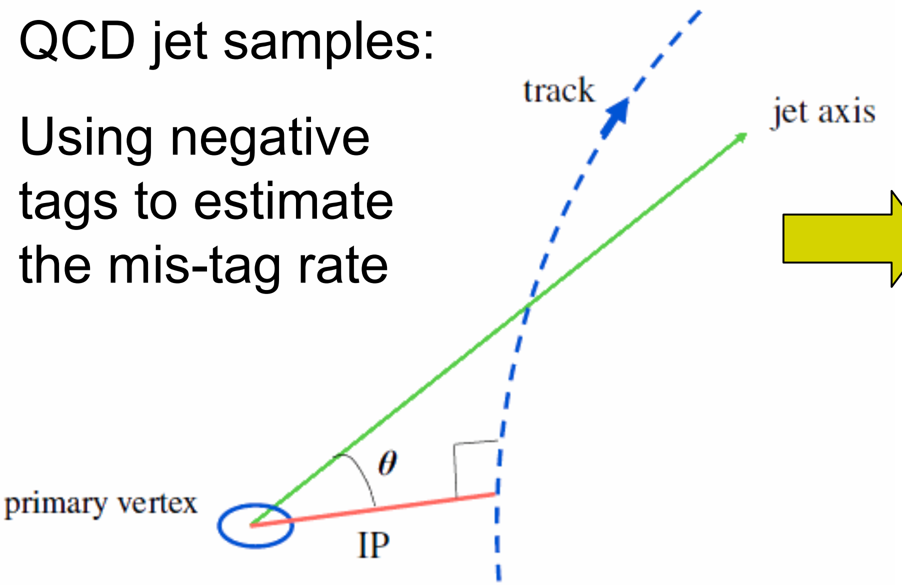
→ Comparison with SM prediction

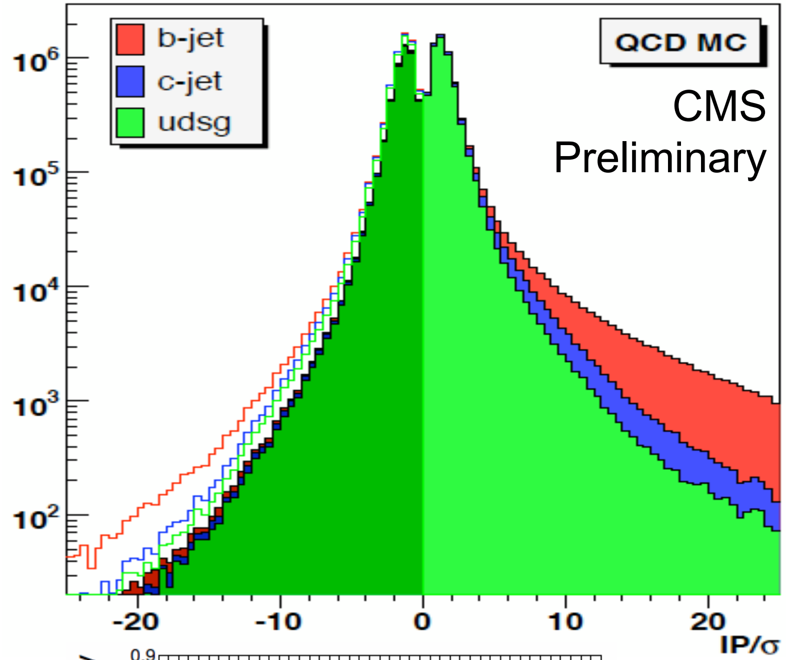
→ pure ttbar sample, but additional syst.

Measuring the B-tagging Performance

QCD jet samples:

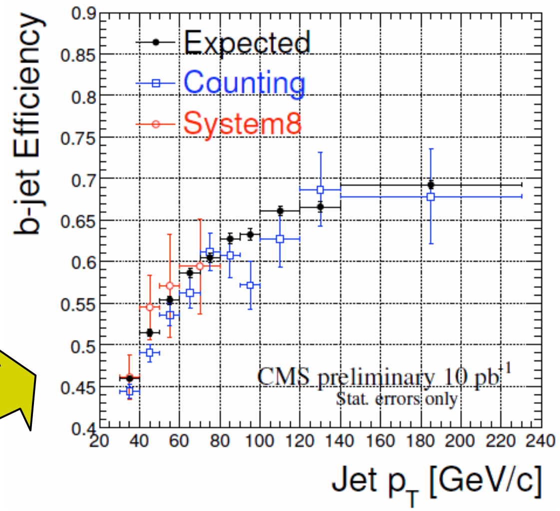
Using negative tags to estimate the mis-tag rate





long-lived particles (K^0_s, Λ), photon conversions, Bremsstrahlung, hadronic interactions w/ material → -/+ asymmetry also for udsg jets (→ corr. from MC)

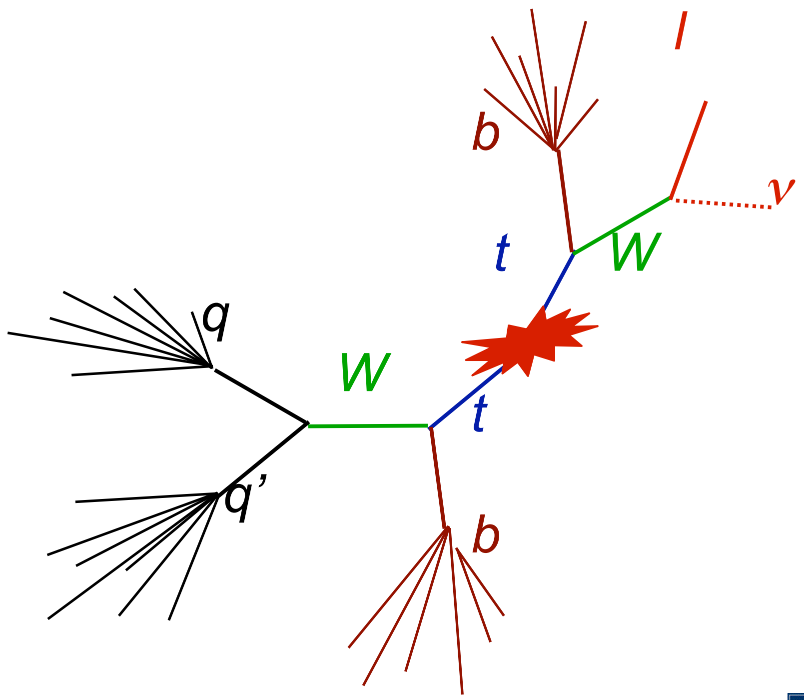
Efficiency also from QCD-jet samples:
$p_T^{rel}(\mu)$ with muons in jets
(or later system-8 …)



The Semi-Leptonic Channels: l+jets

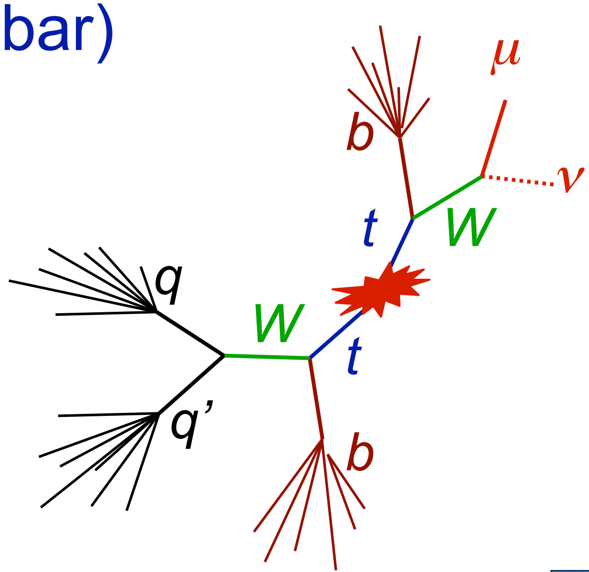
Next: Semileptonic Channel l+jets

- 1 “highly energetic” lepton
- ≥ 4 jets (2 b-jets, 2 light-quark jets)
- Hadronic leg t→Wb→qq’b is fully reconstructable
- Larger BR than di-lepton channel



ttbar x-section in μ+jets

- CMS PAS TOP-09-003:
  http://cms-physics.web.cern.ch/cms-physics/public/TOP-09-003-pas.pdf
- Analysis strategy:
  - Simple counting experiment, robust selection
  - N(cand) after a selection that is characteristic to tt→lνbqqb is compared to expected N(SM backgd.)
  - Any excess is ascribed to N(ttbar)
  - N(SM BG) is estimated from ttbar-depleted control sample

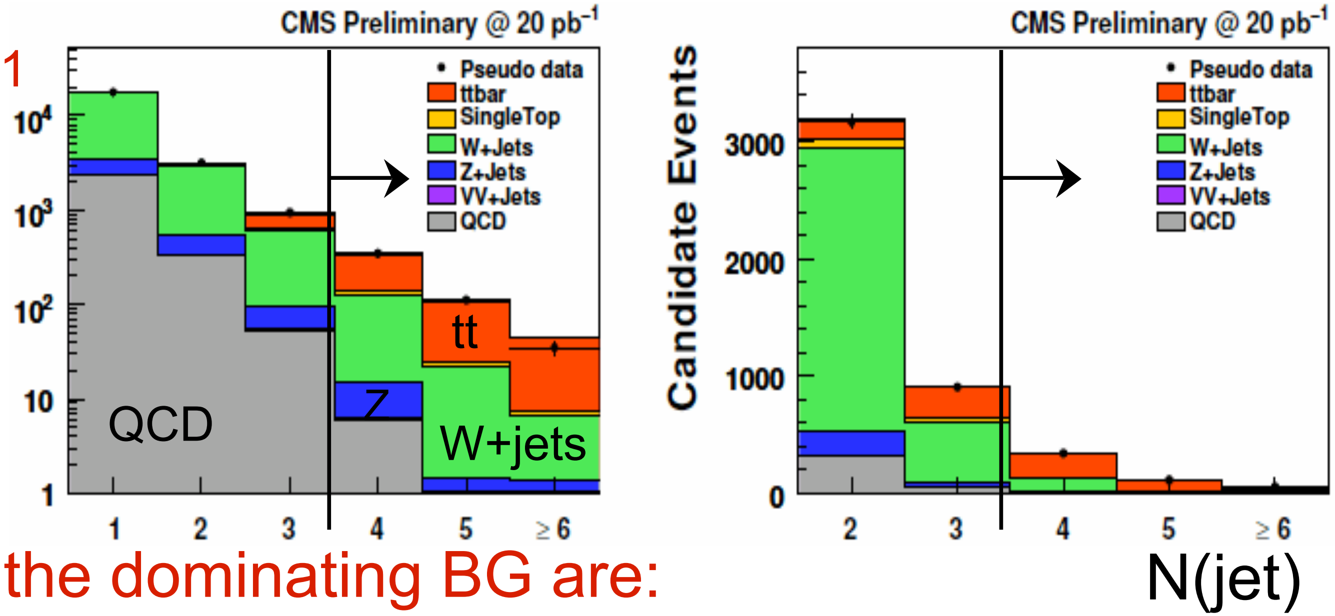


μ+jets Selection

- Lepton selection (similar to di-lepton analysis)
  - Trigger: single muon, $p_T > 9$ GeV
  - Exactly 1 isolated global muon:
    - $|\eta(\mu)| < 2.1$, $p_T > 20$ GeV
    - 2D distance from primary vertex: $d_0 < 200\ \mu$m
    - Quality, e.g.
      - χ^2 of track fit, minimum number of hits in the tracker
      - m.i.p. compatibility in E/HCAL, and
      - rel-Isolation: $[\sum_{\Delta R<0.3} E_T^{calo} + \sum_{\Delta R<0.3} p_T^{tracks}] / p_T(\mu) < 0.05$
  - Veto on any additional isolated lepton (μ or e) over threshold (loose ID & isolation)
- Jets: $|\eta| < 2.4$, $p_T > 30$ GeV, SisCone R=0.5
- Count N(jet), require N(jet) ≥ 4 for singal sample

μ+jets

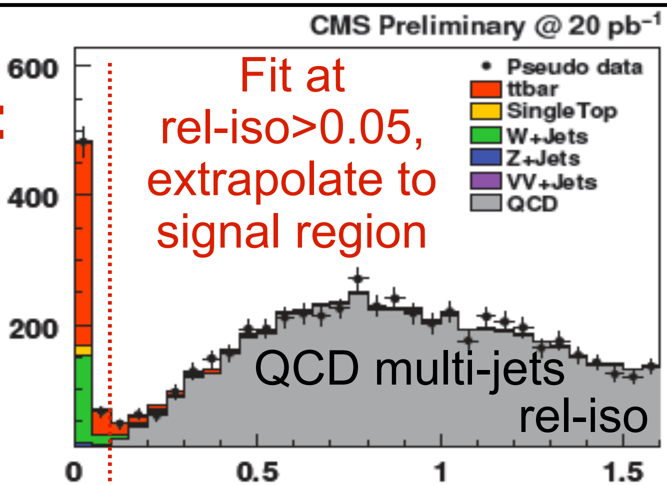
N(jet) for 20 pb^{-1} at 10 TeV:

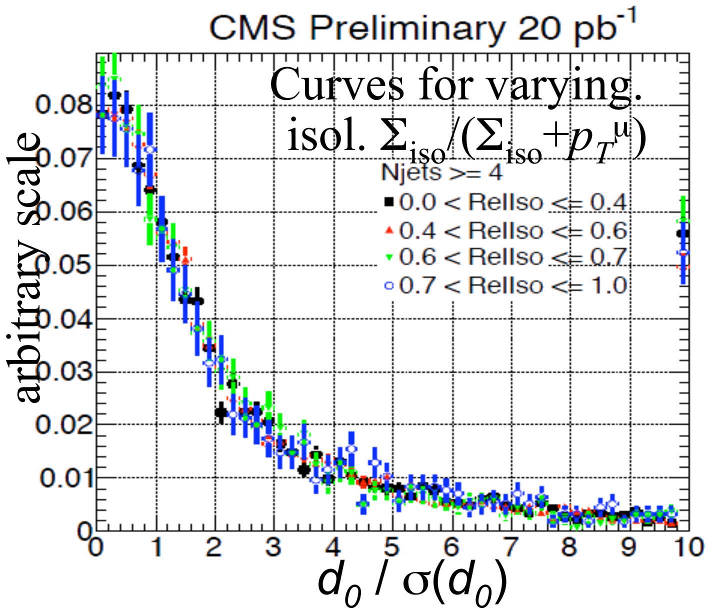


- For *N*(jet) ≥ 4, the dominating BG are:
  - W/Z+jets
    - Normalization from template fit
  - QCD multi-jet events (w/ 1 fake isolated lepton)
    - Looks small, but could be larger real life (have means to adjust)
    - Use data-driven method for subtraction
  - Take others (e.g. single-top) from simulation, assign conservative X% uncertainty

QCD Multi-Jet BG: e.g., ABCD or Fit

1D fit:





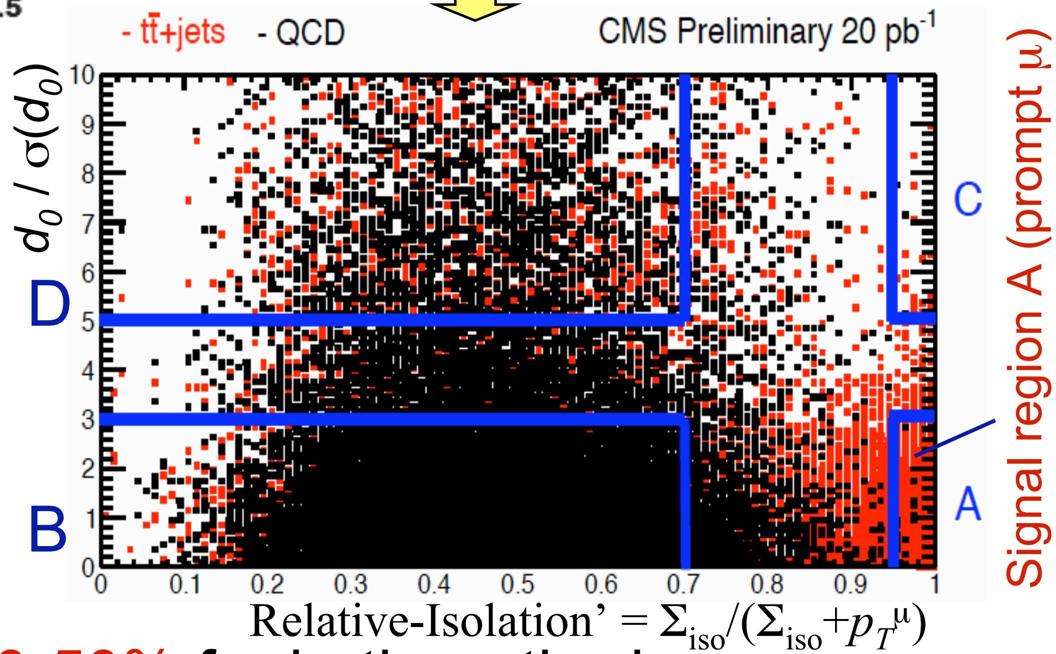
or "ABCD Method":

2 variables independent for QCD BG:

$\Rightarrow N_A^{QCD} = (N_B^{QCD} \times N_C^{QCD}) / N_D^{QCD}$

BG dominating in regions B,C,D

$\Rightarrow N_A^{QCD} \approx (N_B^{total} \times N_C^{total}) / N_D^{total}$

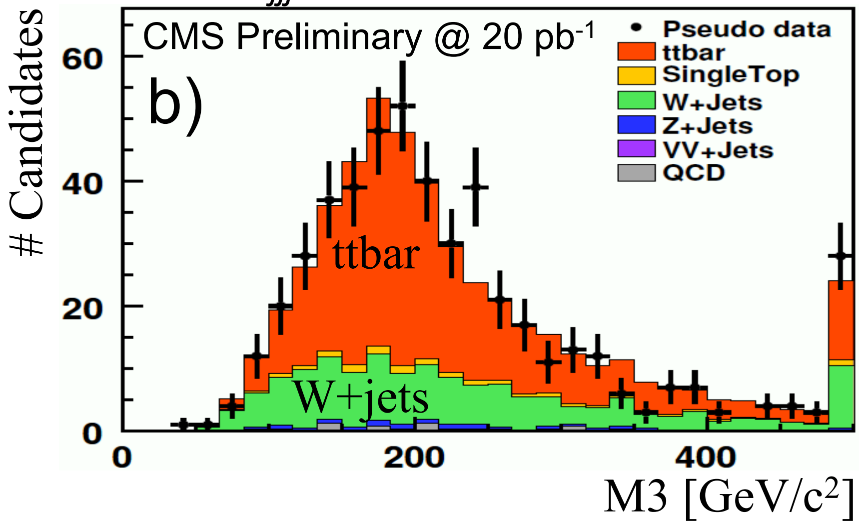


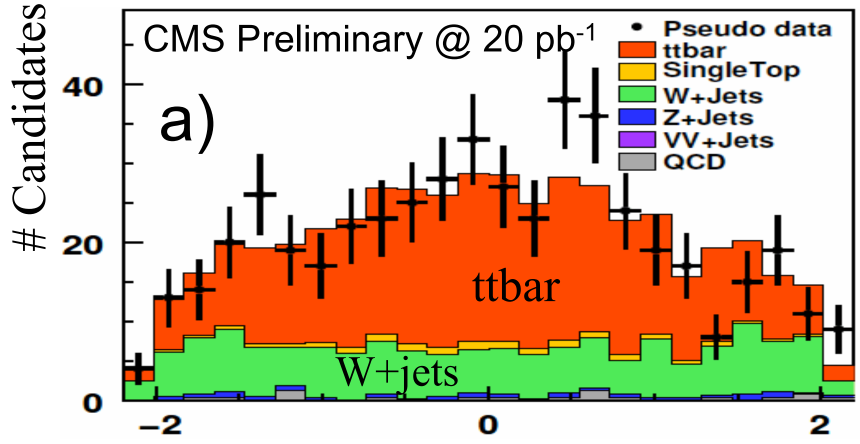
$\delta N_{QCD}/N_{QCD} \approx$ 30-50% for both methods

Statistical Subtraction of W/Z + jets

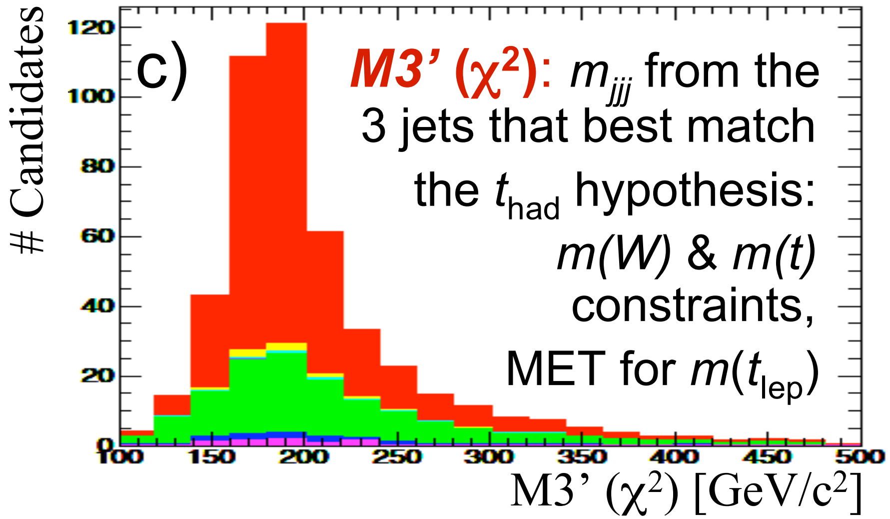
- For x-section extraction, perform **1D template fit** to variable *X* with templates:
  - signal: shape from MC
  - W/Z+jets (similar shape)
  - QCD: shape & normalization from data
  - single-top: shape & normalization from MC
- Studied 3 choices for *X*:

M3: m_{jjj} for 3*j* with max $p_T(jjj)$

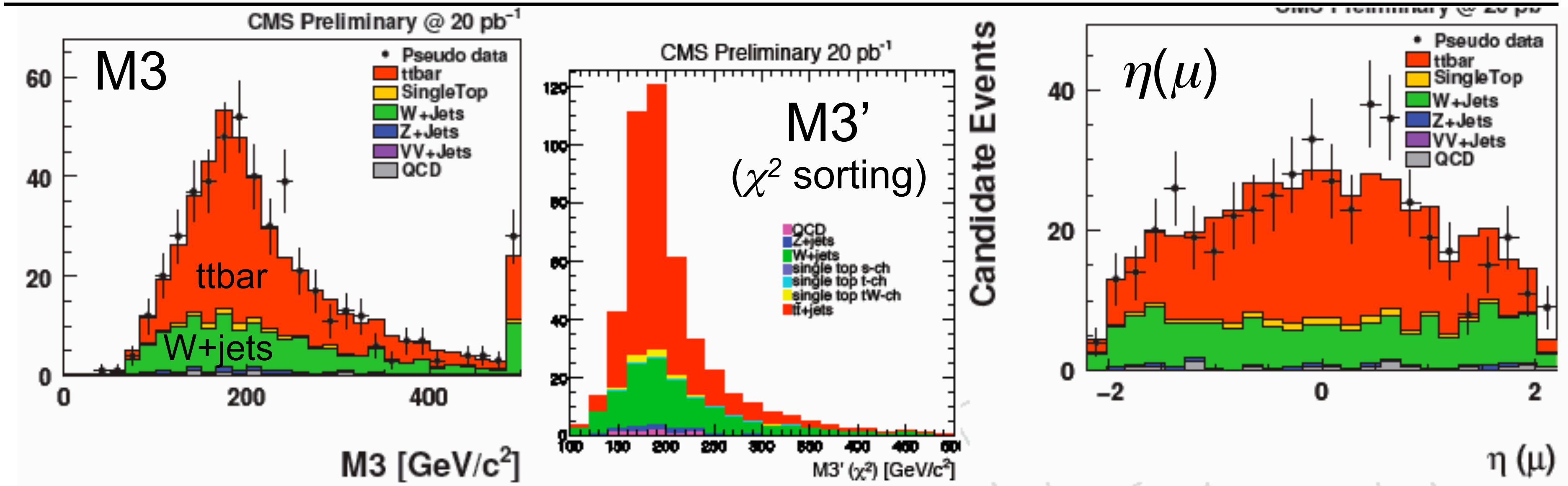




Pseudorapidy $\eta(\mu) = -\ln[\tan(\theta/2)]$



Template Fit → σ(ttbar)



$\delta\sigma/\sigma \approx 16\%^{(stat)} + 24\%^{(syst)}$

$\delta\sigma/\sigma \approx 18\%^{(stat)} + 19\%^{(syst)}$

$\delta\sigma/\sigma \approx 12\%^{(stat)} + 25\%^{(syst)}$

- All 3 variables give somewhat comparable results
- $|\eta(\mu)|$ with smallest syst. uncertainties (the fact that ttbar is produced more centrally than W/Z+jets comes mainly from the mass difference ⇒ least sensitive to JES)

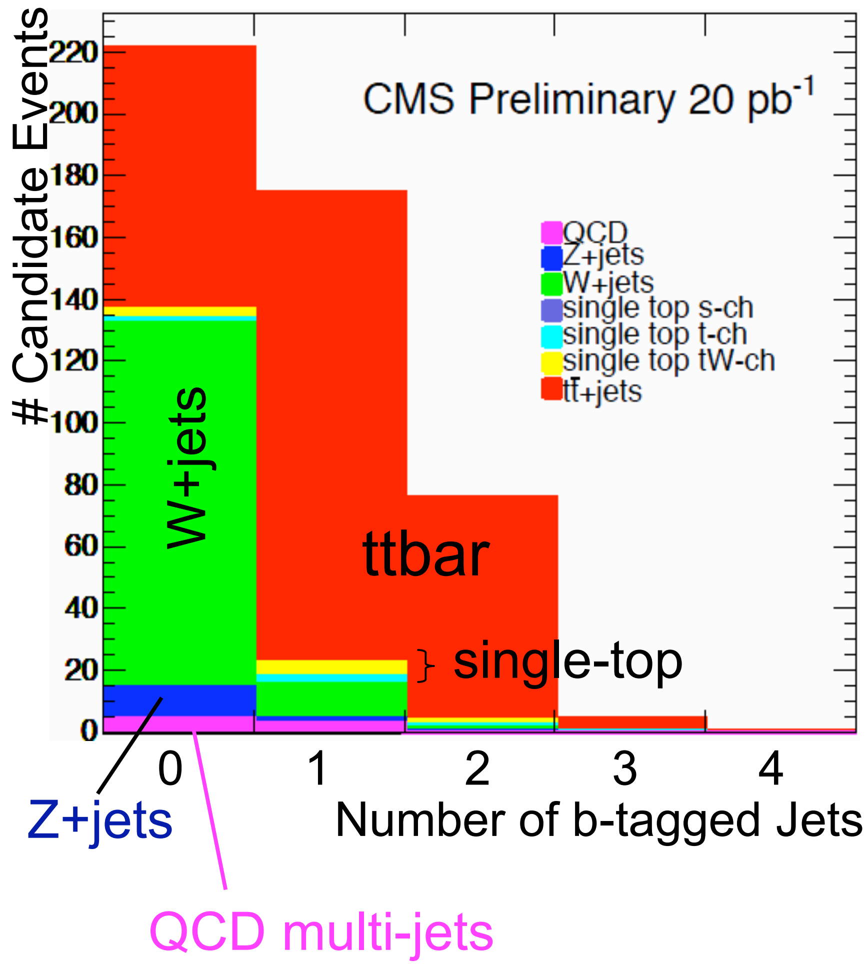
μ+jets Fit Systematics

Source for 20 pb^{-1} at 10 TeV	Uncertainty [%]		
	Fit to $\eta(\mu)$	Fit to M3	Fit to M3′
Statistical Uncertainty (20 pb^{-1})	17.7	16.3	11.5
Jet Energy Scale	16.7	15.1	19
$t\bar{t}$ MC Generator	1.9	14.9	14
$t\bar{t}$ ISR/FSR	3.3	7.7	2
W+jets Factorization scale	4.4	4.7	4
W+jets Matching threshold	5.5	2.8	4
Single Top Shape	0.1	0.8	1
PDF Uncertainty	5.0	5.0	5.0
Total Systematic Error	19.2	23.8	25.0
Luminosity Error	10.0	10.0	10.0

- JES from ±10% variation of jet energies
- Theory: from syst. samples (comparison of different Monte-Carlo generators)
- PDF uncertainty from CTEC6.6 using re-weighting methods implemented in the LHAPDF package
- Quote luminosity-uncertainty separately

Outlook for μ+jets

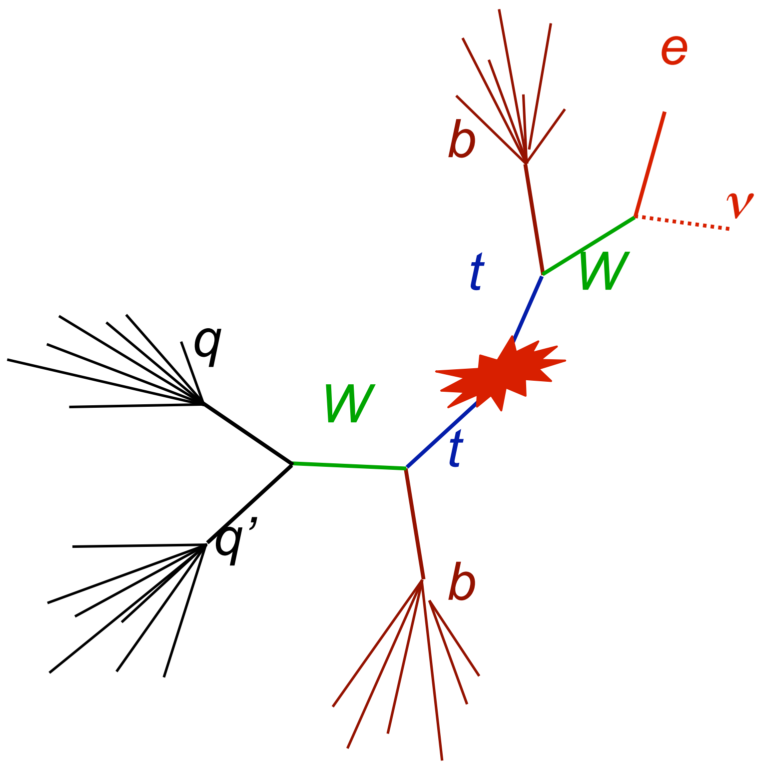
- Ask for N>0 b-tagged jets in the event (in addition to previous selection):
  - Here with a simple secondary-vertex tagger
  - Requires a good understanding of the heavy-flavor content in W+jets events (t.b. measured in control sample $N_{jet}<4$!?)
  - Eff. & fake rate from jet samples



→ pure selection vs. additional systematic uncertainty

ttbar cross-section in e+jets

- CMS PAS TOP-09-004
  - http://cms-physics.web.cern.ch/cms-physics/public/TOP-09-004-pas.pdf
- Similar approach as for μ+jets:

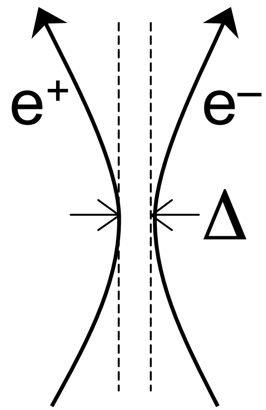


e+jets Selection

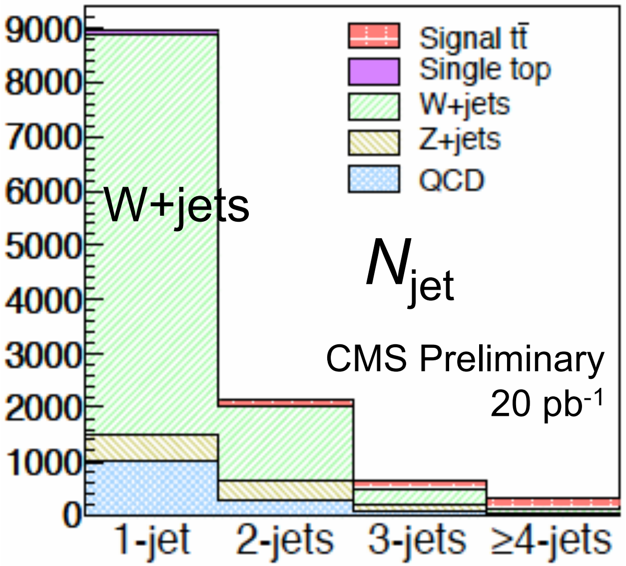
- Trigger: single electron with p_T>15 GeV
- Exactly 1 electron:
  - E_T > 30 GeV, $|\eta_e|$ < 2.5 (excluding 1.442<$|\eta_e|$<1.560 between barrel and endcap region)
  - $|d_0|$ < 200μm
  - rel-iso < 0.1, tight electron-ID (shower shape, H/E, etc. …)
- Veto additional isolated muons (p_T > 20 GeV)
- Jets:
  - As in other analyses: SisCone0.5, calo-jets, $|\eta_j|$<2.4
  - p_T > 30 GeV
  - Electron-jet cleaning: ΔR(e,j)>0.3
- Additional Z-veto: 2 options (will pick 1):
  - Option A: Reject events with additional loose electrons
  - Option α: Allow additional loose e, but reject 76<M_{ee}<106 GeV

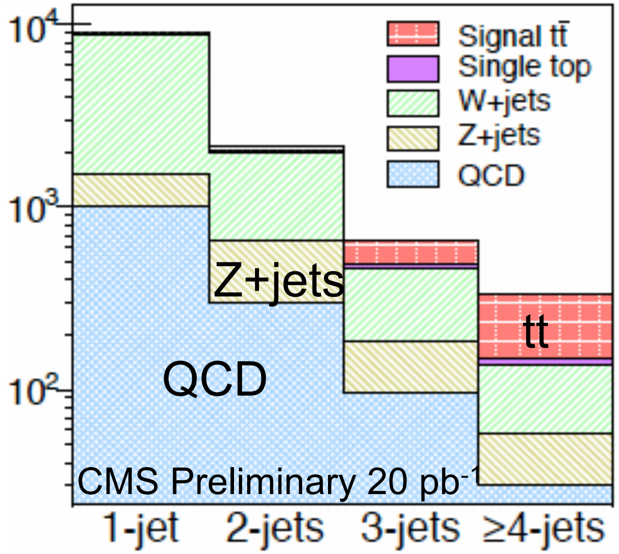
Additional rejection of QCD multi-jets

- Here, present 2 alternatives for further QCD rejection:
  - A: $|\eta_e| < 1.442$ barrel only:
    - Most material is in the forward region ⇒ removing the forward region reduces background from photon conversions
  - B:
    - MET > 20 GeV
    - Explicit photon-conversion removal:
      - Look for track pairs of opposite curvature
      - If $\Delta < 0.04$ cm & one of the tracks is within R=0.3 of the electron candidate, then reject the event



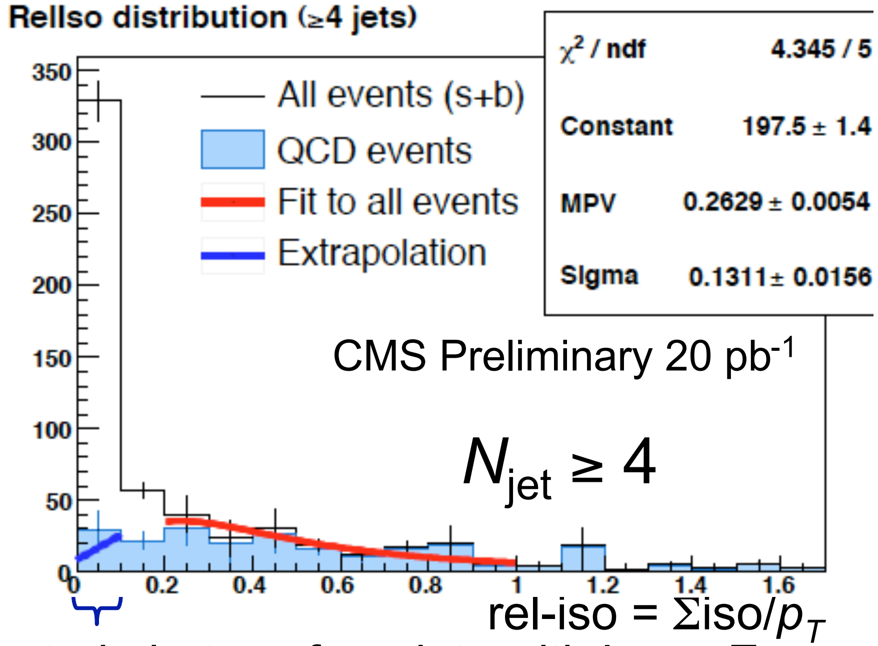
e+jets





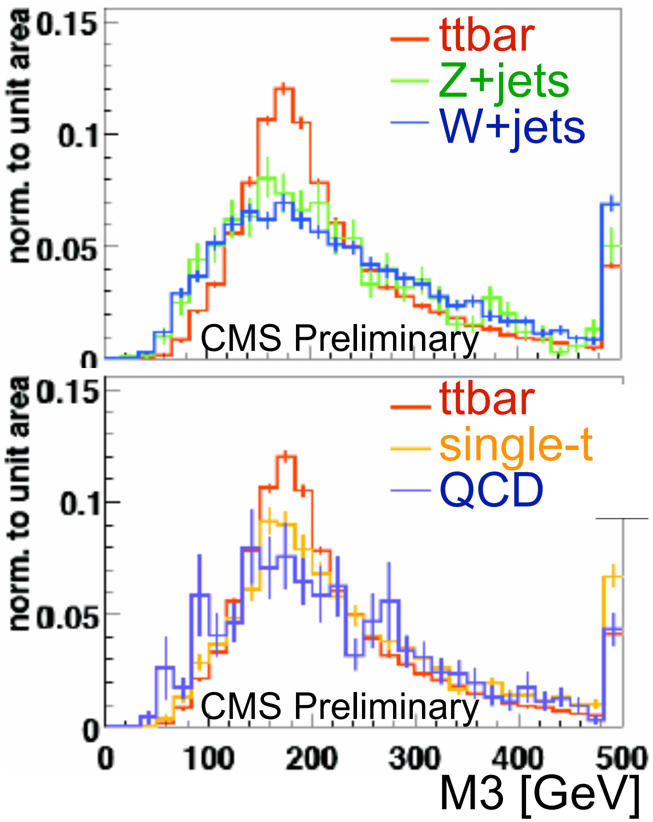
- B: Backgrounds for $N_{jet} \geq 4$:
  - QCD multi-jet events:
    - Estimate from rel-iso fit: OK to ~ 50%

Fit in all jet bins:



Beware: converted photons from jets with large E-fraction in π^0 can really look like an isolated electron → enforce conversion killer, maybe add alternative method/check for QCD estimation (e.g. w/ M_T)? …

e+jets Results



M3 = mass of the 3-jet system with largest p_T(jjj)

- M3: a simple "reconstruction" of the hadronically decaying top
- Statistical separation using template distributions → σ(ttbar)

Expected uncertainties:

for 20 pb^{-1}	Relative Systematic Uncertainty
Jet Energy Scale	15%
MC Generator	10%
$t\bar{t}$ ISR/FSR uncertainty	3%
W+jets MC Factorization Scale	1%
W+jets MC Matching threshold	5%
Shape uncertainty of Single Top	1%
Shape uncertainty of QCD	2%
PDF uncertainty	5%
Total	20%

Luminosity uncertainty (10%) quoted separately …

Few words of care …

- The analyses shown are complete in the sense that all base-ingredients have been worked out at least once, however:
  - Collision data will likely ask for new strategies, modifications (MET, primary vertex, calibrations, …)
    - Also alternative choices, advanced reconstruction tools …
  - So far, it was said that lepton efficiencies (trigger, reconstruction, isolation, quality cuts …) will come from *Tag & Probe* with Z→ll data
    - Has yet to be exercised in the context of ttbar analyses (e.g. corrections required due to high multiplicity in ttbar)
  - Need a luminosity estimate to measure σ(ttbar)
    - tools for lumi calculation; need to exercise them (in conjunction with bad run/lumi-segment list)
    - if problems with lumi numbers arriving in time, could also measure σ(ttbar)/σ(Z) (and take σ(Z) from theory) … t.b.i.

Then, there is b-tagging: A shopping list

- Next: want to define a strategy for cross-section measurement with b-tagging (lepton and/or life-time)
  - Natural order:
    - First observe/measure σ(tt) with kinematical analysis (w/o b-tags)
    - Then, verify compatibility with SM: e.g. differential x-sections, but also N(b-tag) distribution ($\rightarrow$ establishing b-tagging)
    - Finally: Include b-tagging in x-section measurement
  - Currently, we are after the strategy t.b. used with first data
    - I.e. fine tuning of the details (algorithm improvement, optimization of working points or selection) are of 2nd priority
    - Take existing b-tag algorithms + selection from kinematical analysis, $\rightarrow$ put them together
  - Goal: complete strategy before 1st collisions
  - Assume that b-tag eff. measured via System8/p_T^{rel} in QCD bbbar
    - How to apply tag-efficiency to ttbar sample? (there are indications that there are some subtleties involved …)

B-tagging II: Ingredients

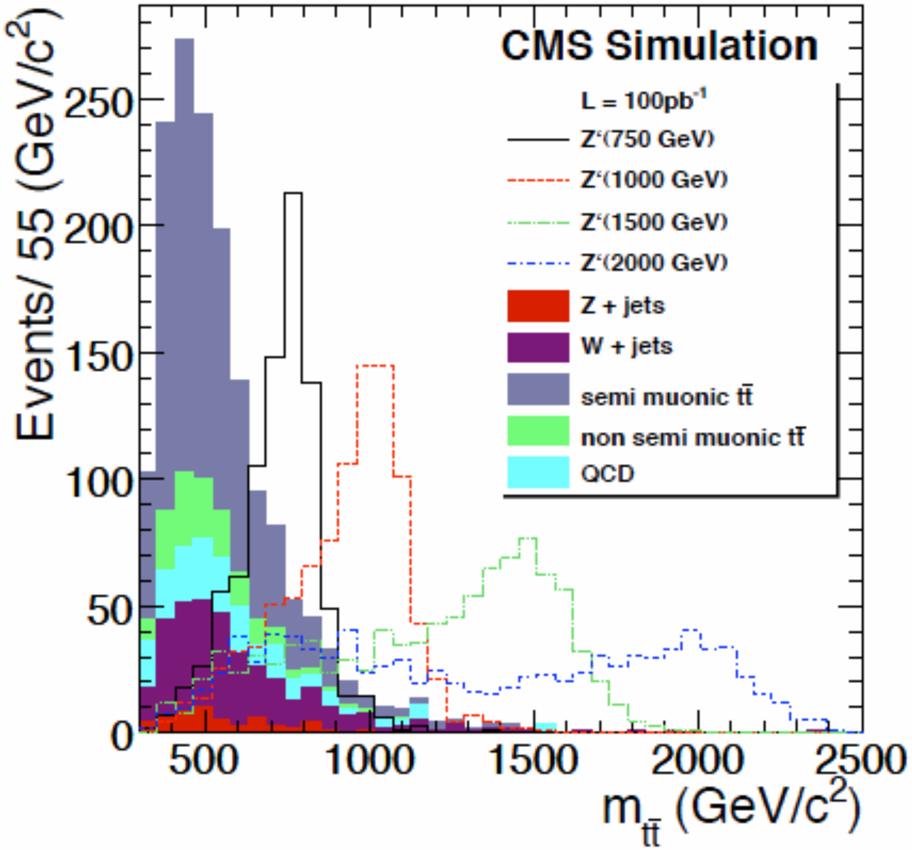
- for $N_{jet} \geq 4$ bin (in e/μ+jets):
  - N(ttbar) = N(obs.) - N(W+light) - N(W+heavy) - N(QCD) - N(VV) - N(singl-top) - N(Z)
- This is really a measurement of the background!
- N_{jet} = 1,2,3 is depleted in ttbar → control samples
  - W/Z+light: these are fake tags; take fake-tag parameterization from b-tag group, apply to W+light
    - Does it work? Is FR=FR(p_T,η,N_{tracks} …)? parameterization adequate?
  - W/Z+heavy: Wbbbar, Wccbar, Wc (heavy-flavor MC tools)
    - Need σ(W+heavy)/σ(W+light)
      - Simply take it from MC?
      - Or measure in N_{jet}=1,2, extrapolate to $N_{jet} \geq 4$? associated syst.?
  - QCD: have data-driven methods for analysis w/o b-tag, do they still apply?
  - Single-top: directly from MC with uncertainty from x-section?

Beyond the Measurement of the Inclusive Cross Section

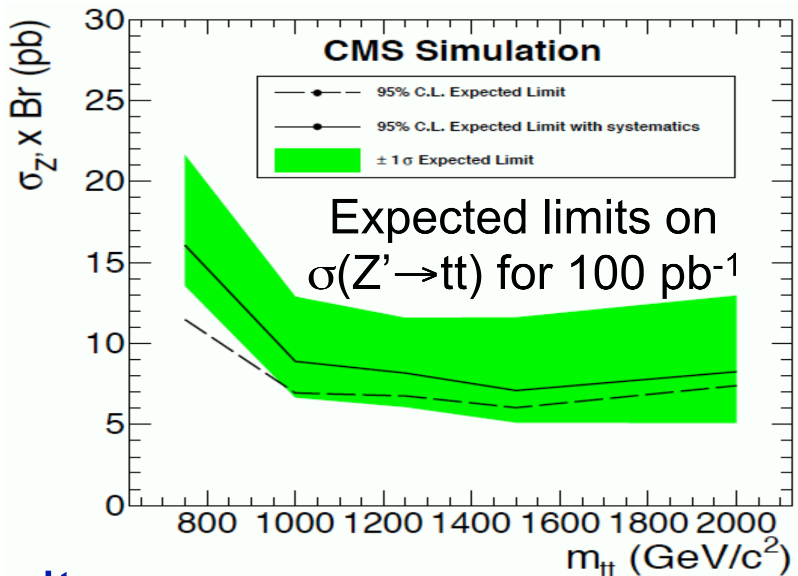
New Physics in the Top Sector?

- With little more data, study the selected samples in more detail, e.g. the m(ttbar) spectrum

e.g. μ+jets channel:



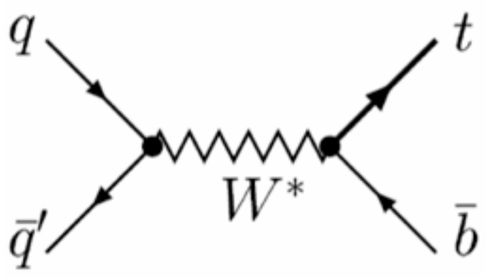
Requires dedicated isolation for high-p_T top quarks, or even dedicated top-tagging, dedicated jet algorithms …



more from Jorgen D'Hondt …

Single-Top

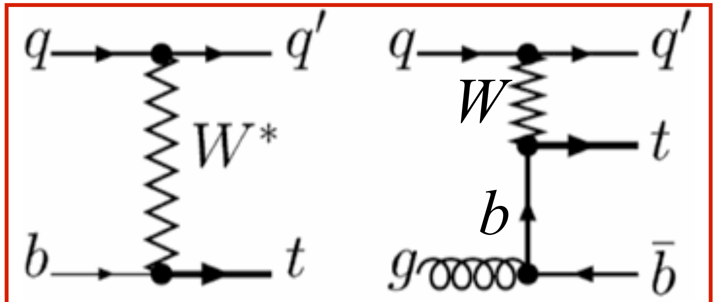
Electroweak Production: Single-Top



s-channel

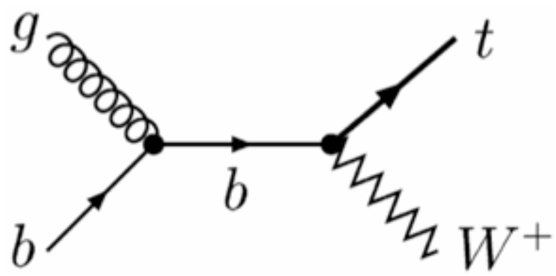
σ = 5 pb

(at $\sqrt{s}$ = 10 TeV)



t-channel

σ = 130 pb



tW-channel

σ = 29 pb

- Tevatron: recent 5σ observation in s+t channels (~ 1+2 pb)
  - tW channel negligible at Tevatron
- LHC @ 10 TeV: t-channel is dominant

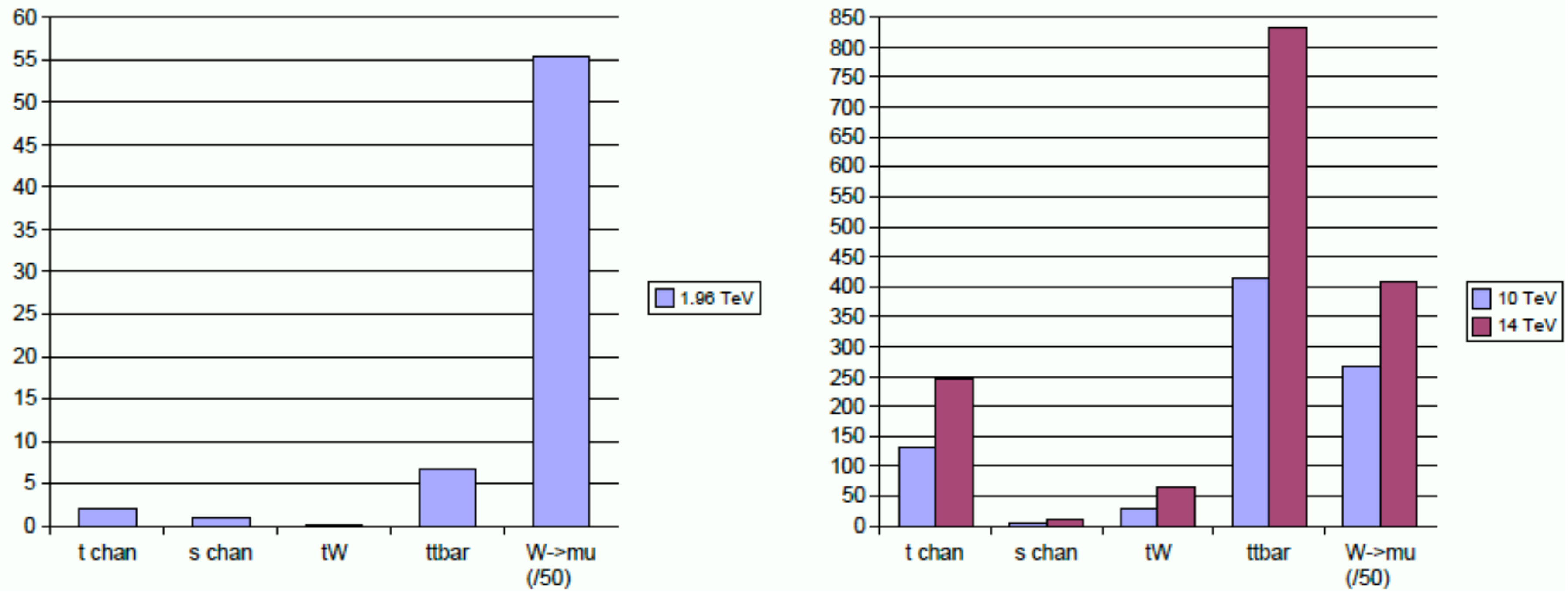
Route:
- Confirm observation
- Competitive constraints on $|V_{tb}|$
- FCNC and anonymous Wtb couplings
- all channels …

CMS PAS TOP-09-005:

http://cms-physics.web.cern.ch/cms-physics/public/TOP-09-005-pas.pdf

From the Tevatron to the LHC

Inclusive cross sections [pb]

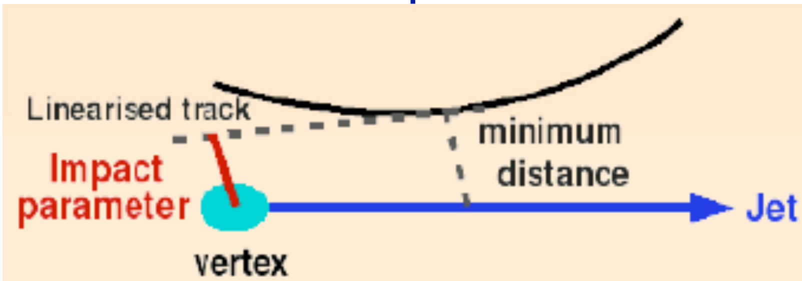


LHC analysis will be quite different:

W+jets will be less of a concern than ttbar

t-Channel: Selection

- Single-muon, extra-lepton veto, isolated
  - exactly 1 μ: $p_T > 20$ GeV, $|\eta| < 2.1$
  - no isolated electron
  - relative isolation: $p_T/(p_T$+tkIso+caloIso)>0.95
- Two jets, well separated from the muon candidate
  - IterativeCone R=0.5
  - $p_T > 20$ GeV, $|\eta| < 2.1$
  - $\Delta R(\mu,\text{jet}) > 0.3$ (otherwise reject event)
- One b-jet
  - ≥ 2 tracks in the jet with significant distance from interaction point



  - 1 jet passing a tight selection
  - no other jet that passes even light tag selection
- On-shell W boson
  - $M_T > 50$ GeV

$$M_T = \sqrt{(p_{T,\mu} + p_{T,\nu})^2 - (p_{x,\mu} + p_{x,\nu})^2 - (p_{y,\mu} + p_{y,\nu})^2}$$

Background from QCD Multi-Jet Events

- How to estimate BG with "fake prompt/isolated" muons (QCD multi-jet events with muons from b/c decays in jets) from signal-like events (prompt isolate muons from W decays):
  - Parameterize the shape of the transverse mass distribution M_T: $F(M_T) = a \times S(M_T) + b \times B(M_T)$

$S(M_T)$: signal-like events (e.g. s.t., $W \to \mu\nu$)

e.g. from ctrl. sample:
i) W-enriched sample, w/o b-tagging, or
ii) Z-enriched sample with 2 μ,
$76 < m_{\mu\mu} < 106$ GeV, 2 jets, w/o b-tagging
Scale μ-momenta by M_W/M_Z
then treat one of the muons as neutrino by adding its p_T to MET

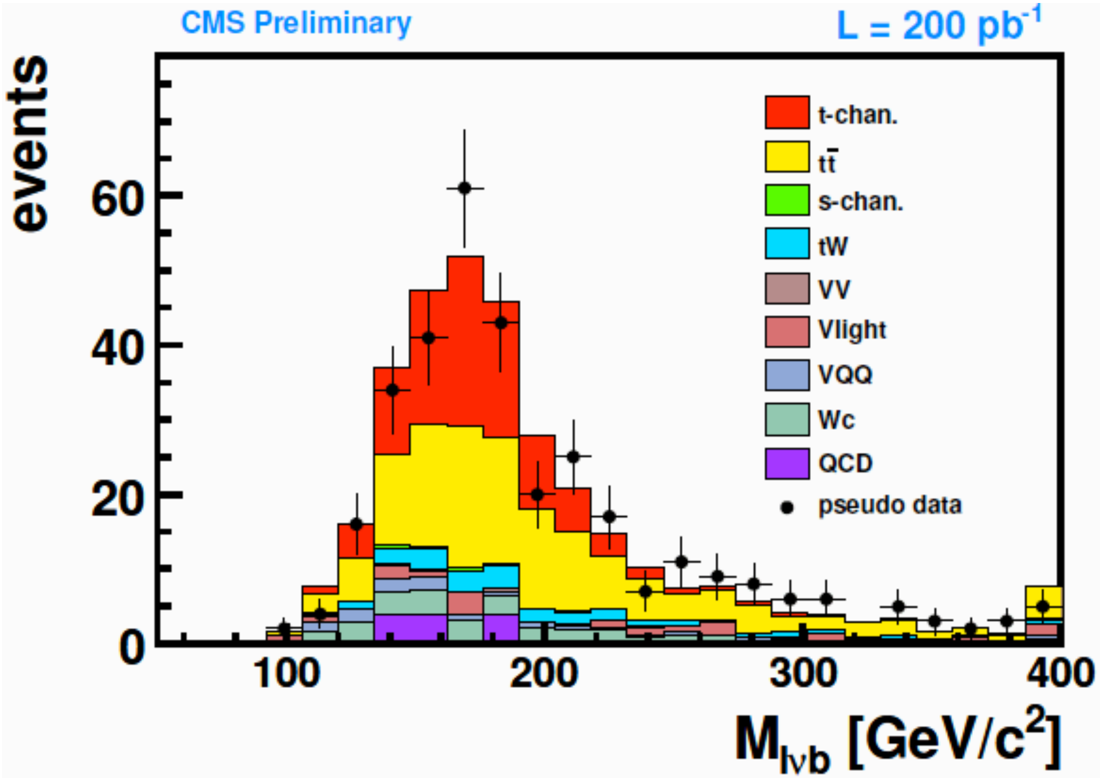
$B(M_T)$: background shape

e.g. from ctrl. sample: anti-isolation, no b-tagging cut

⇨ estimate 22 for 12±7 expected (3 MC events!), assigned 45% uncert.

Candidate Events:

- Reconstructed top mass $M(l\nu b)$:



for 200 pb^{-1}
@ 10 TeV

MET → $p_T(\nu)$ $p_z(\nu)$ from $M_W = M(l\nu)$ constraint

Will give a first indication that selected sample is indeed from top-quarks

Extracting N(t-channel single-top)

- Smoking gun:
  - V–A structure of the weak interaction

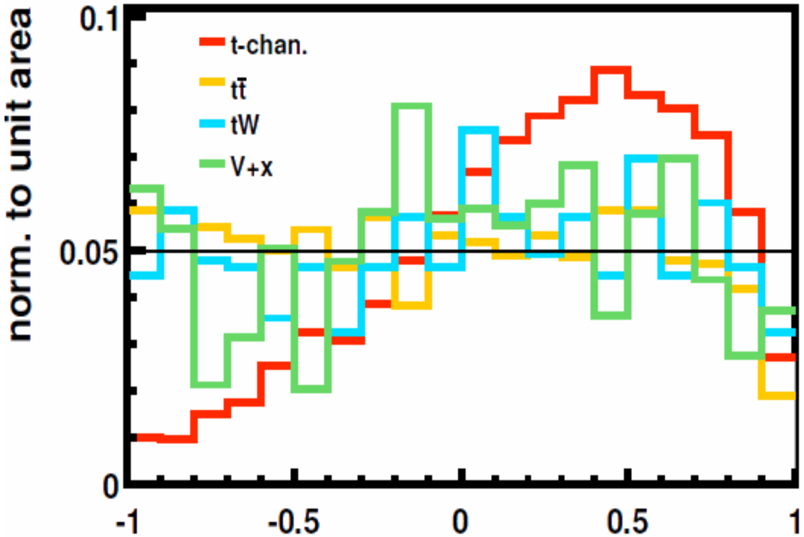
    ⇒ ~ 100% top-quark polarization w.r.t. to the spin axis

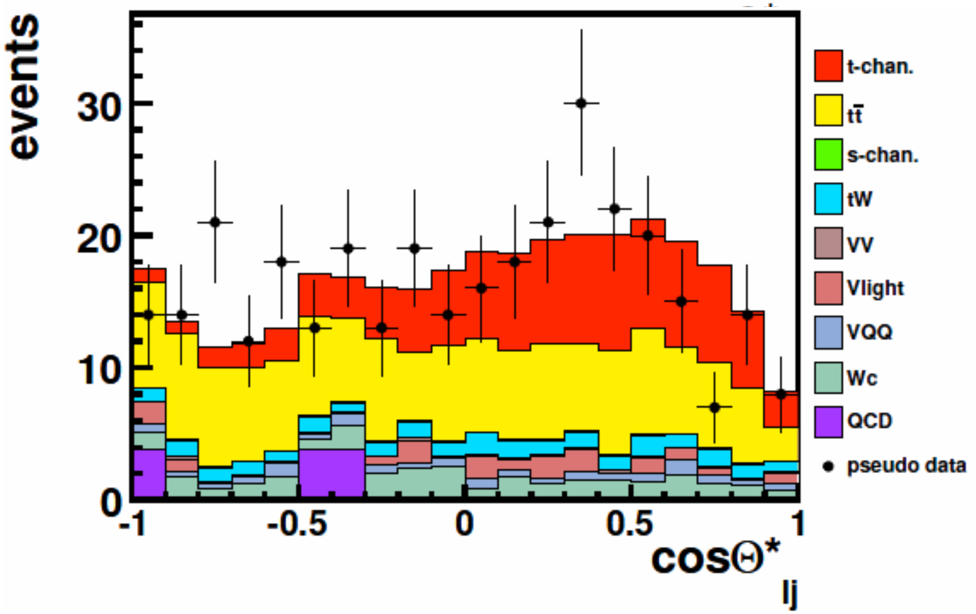
$$\frac{1}{\Gamma}\frac{d\Gamma}{d\cos\theta^*_{lj}} = \frac{1}{2}(1 + \underbrace{A}\cos\theta^*_{lj})$$

1 for $l^\pm$

θ^*_{lj}: angle btwn. direction of the outgoing lepton and the spin axis (≈ un-tagged-jet axis) in the top rest frame

➡ Fit to cos θ_{lj}*





Single-Top Results

- Expected sensitivity for 200 pb^{-1}

Source of uncertainty	$\Delta\sigma$ [%]	Expected sensitivity
statistical	$\pm$ 35	2.8σ
b tagging	$\pm$ 7.3	2.7σ
mistag	$\pm$ 0.4	2.7σ
JES	$\pm$ 5.5	2.7σ
MET	$\pm$ 9.9	2.7σ
PDF	$\pm$ 5.5	2.7σ
total	$\pm$ 39	2.7σ

This, with a **simple and robust** selection and **not** a complex multi-variate analysis, and with 200 pb^{-1} only!

Future options: e.g. the use of the charge asymmetry (more u than d quarks in protons), …

Conclusions

- There is plenty of top signal expected in the data of the first LHC run
  - Possibly even at 7 TeV: ttbar cross-section is ~ 40% compared to 10 TeV
- ttbar is a milestone of the LHC physics program
  - understanding of the detector, reconstruction, and the Standard Model
  - Important background for searches
  - calibration tool: jets, b-tagging
- Focused on first robust analyses for cross-section measurements
  - Estimation of backgrounds from the data themselves

Summary continued

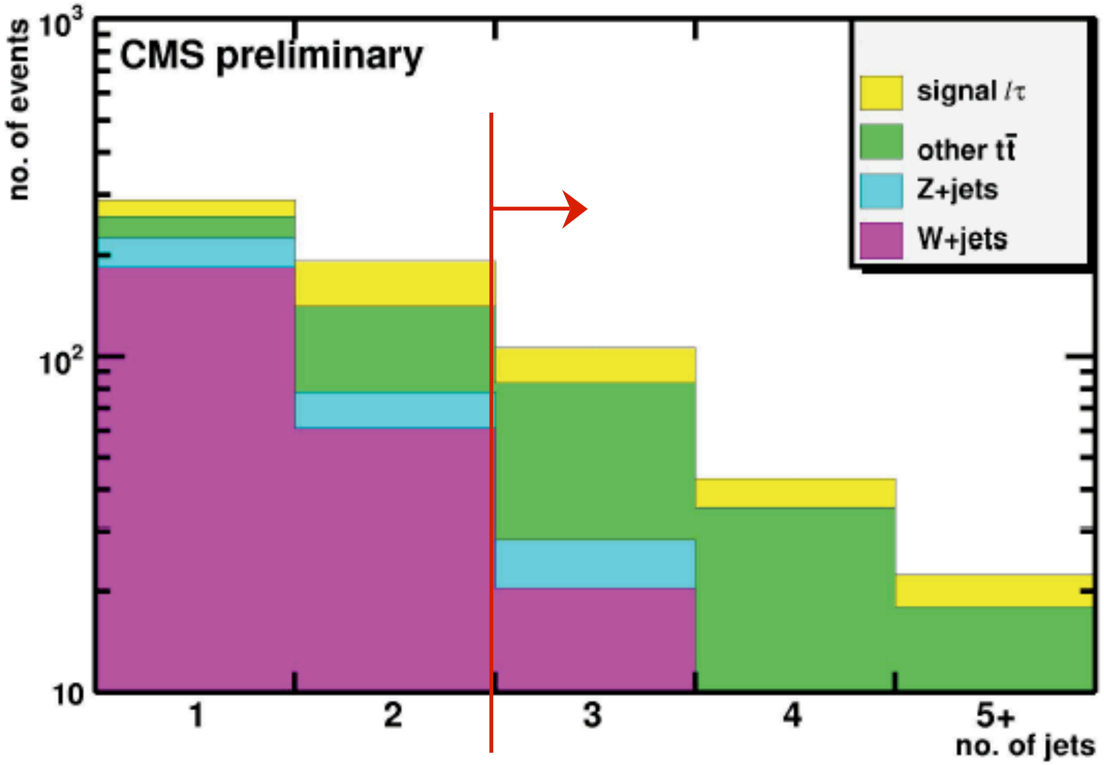
- Of course, this is not the full top-physics program
  - Once the first top-quark signals are established, need to look at the samples from all angles: is there more to them? compatibility with the standard model expectations, new physics …
    - Heavy flavor content
    - differential cross sections, e.g. m(ttbar), p_T, $|\eta|$ …
    - comparison of different channels (also hadronic ttbar!)
    - V_{tb} in production and/or decay, rare decays
    - ttbar with taus (see CMS PAS TOP-08-004 & backup)
    - Fully-hadronic channel tt→jets
    - …
  - Top-quark properties
    - measurement of the top-quark mass m(top)
    - spin correlation, W polarization in ttbar events
    - …

Thanks for the invitation and the wonderful organization of the LHC Physics Workshop!

Backup

τ Di-lepton Channels: τ+e & τ+μ with Jets

- Isolated lepton (μ or e), p_T > 20 GeV
- Tau as narrow jet with leading track p_T > 20 GeV
- Count other jets with E_T > 30 GeV, and $|\eta|$ < 2.4
- Missing transverse energy MET > 60 GeV



Expectations for 100 pb^{-1}

Fake-Tau Estimation

- Same approach as for electron fakes
  - "Measure" τ-fake rate in QCD events
  - Apply to Monte-Carlo W+jets sample to predict l+ τ rate
  - Compare to l+τ rate "measured" in W+jets MC

Good to about 30 %

Table 3: Expected number of τ-fake events (from "data" and from MC expectations) in $\mathcal{L}$ = 100 pb^{-1}. Uncertainties are statistical only.

Method		τ-fakes from "data"	expected from event selection
1- or 3-prongs	γ+jets	523±8	438±20
	"all" jets	542±9	
	leading jet	639±10	
	next-to-leading	498±9	
	back-to-back	577±10	
1-prong	γ+jets	308±5	278±16
	"all" jets	361±7	
	leading jet	417±7	
	next-to-leading	341±7	
	back-to-back	385±7	

S/B ≈ 0.4 for 1-prong decyas

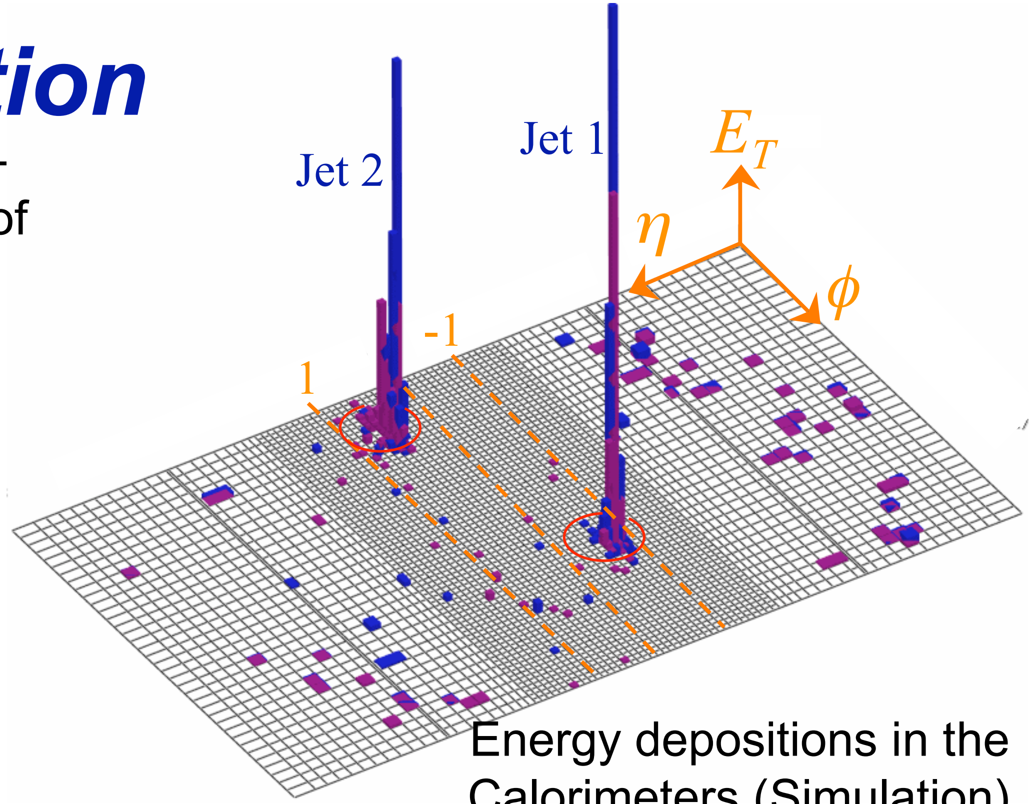
(CMS PAS TOP-08-004)

Jets Reconstrcution

Jets are the experimental signature of quarks & gluons, registered as collimated sprays of particles

Jet algorithms: mathematic rule how these particle/energy sprays are clustered and summed

The “raw” jet parameters need to be calibrated, so that the jets (energy, direction) “best correspond” to the originating quarks/guons



Energy depositions in the Calorimeters (Simulation)

Limitations / necessary corrections:

Detector noise, pile-up (signals from other collisions), underlying event, detector response (inhomogeneity, particle- and energy-dependent response), particles missed by the jet algorithm (magnetic field!), different showering/fragmentation of gluons, light and heavy quarks, muon corrections, etc. ..

→ Calibration with collision data, but this needs some time and patience …

LHC Expectations for the upcoming Run

http://lhc-commissioning.web.cern.ch/lhc-commissioning/luminosity/09-10-lumi-estimate.htm

Month	Comment	Turn around time	Energy [TeV]	Max number bunches	Protons/Bunch	% nom. intensity	Min beta*	Peak Luminosity $cm^{-2}s^{-1}$	Integrated Luminosity	events/X
1	Beam commissioning								First collisions	
2	Pilot physics, partial squeeze, gentle increase in bunch intensity, avaialbility low	Long	3.5	43	3×10^{10}		4 m	8.6×10^{29}	$100 - 200\ nb^{-1}$	
3		5	3.5	43	5×10^{10}		4 m	2.4×10^{30}	$\sim 1\ pb^{-1}$	
4		5	3.5	156	5×10^{10}	2.5	2 m	1.7×10^{31}	$\sim 9\ pb^{-1}$	
5a	No crossing angle - could at this stage push intensity see 5b	5	3.5	156	7×10^{10}	3.4	2 m	3.4×10^{31}	$\sim 18\ pb^{-1}$	0.8
5b	No crossing angle - squeezing to beta* = 1m at this stage would double these lumi numbers (and the pile-up)	5	3.5	156	10×10^{10}	4.8	2 m	6.9×10^{31}	$\sim 36\ pb^{-1}$	1.6
6	Possible shift to higher energy - would anticipate ~4 weeks to reestablish physics follow by a fairly gentle increase back up in intensity.	Would aim to first provide a period of physics at the higher energy (4.5 TeV, say) without crossing angle, this could be followed by a move to 50 ns with a limited number of bunches. Note that the total intensity limit will go down with the move to higher energy.								
7	4 - 5 TeV (5 TeV luminosity quoted - doesn't make too much difference). No crossing angle.	5	4 -5	156	7×10^{10}	3.4	2 m	4.9×10^{31}	$\sim 26\ pb^{-1}$	
8	50 ns - nominal crossing angle - aperture restricts squeezing further - note limited complement of bunches.	5	4 -5	144	7×10^{10}	3.1	2 m	4.4×10^{31}	$\sim 23\ pb^{-1}$	
9	50 ns	5	4 -5	288	7×10^{10}	6.2	2 m	8.8×10^{31}	$\sim 46\ pb^{-1}$	
10	50 ns*	5	4 -5	432	7×10^{10}	9.4	2 m	1.3×10^{32}	$\sim 69\ pb^{-1}$	
(11)	50 ns*	5	4 -5	432	9×10^{10}	11.5*	2 m	2.1×10^{32}	$\sim 110\ pb^{-1}$	