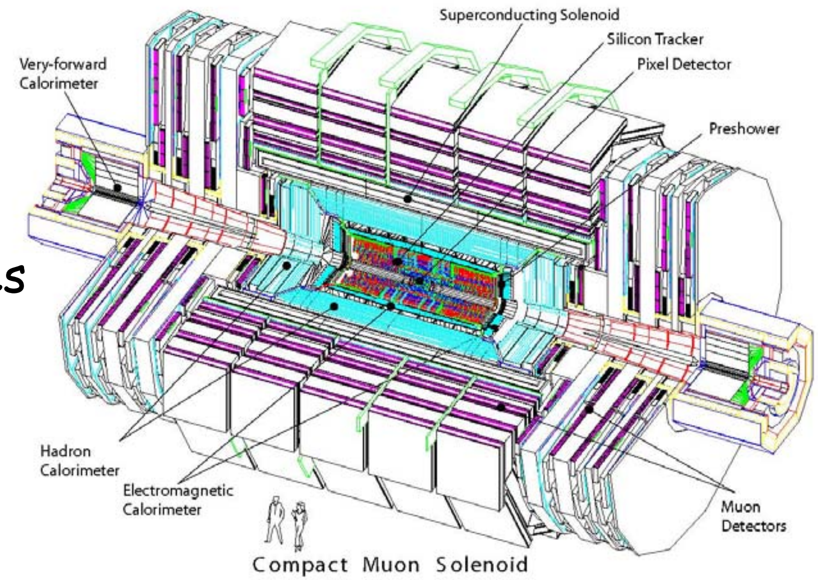
Early top Physics at LHC

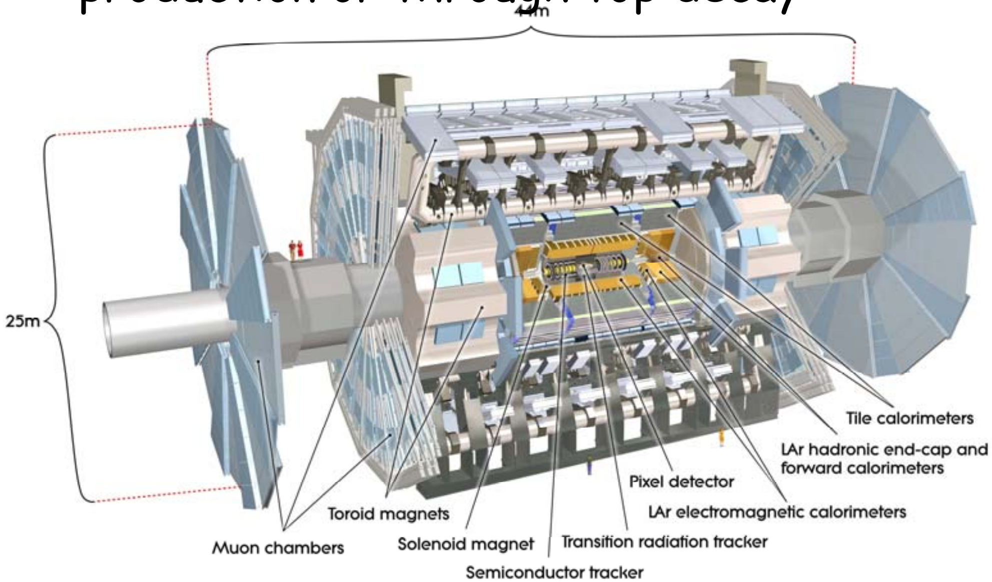
Daniel Bloch (IPHC Strasbourg) on behalf of ATLAS and CMS

Main interests for top studies:

large event rate at LHC ⇒

- detailed tests of Standard Model
- important calibration tool
- a main background for new physics searches
- the high mass of the top (173 GeV) may favour its couplings to new physics and motivates searches with associated top production or through top decay





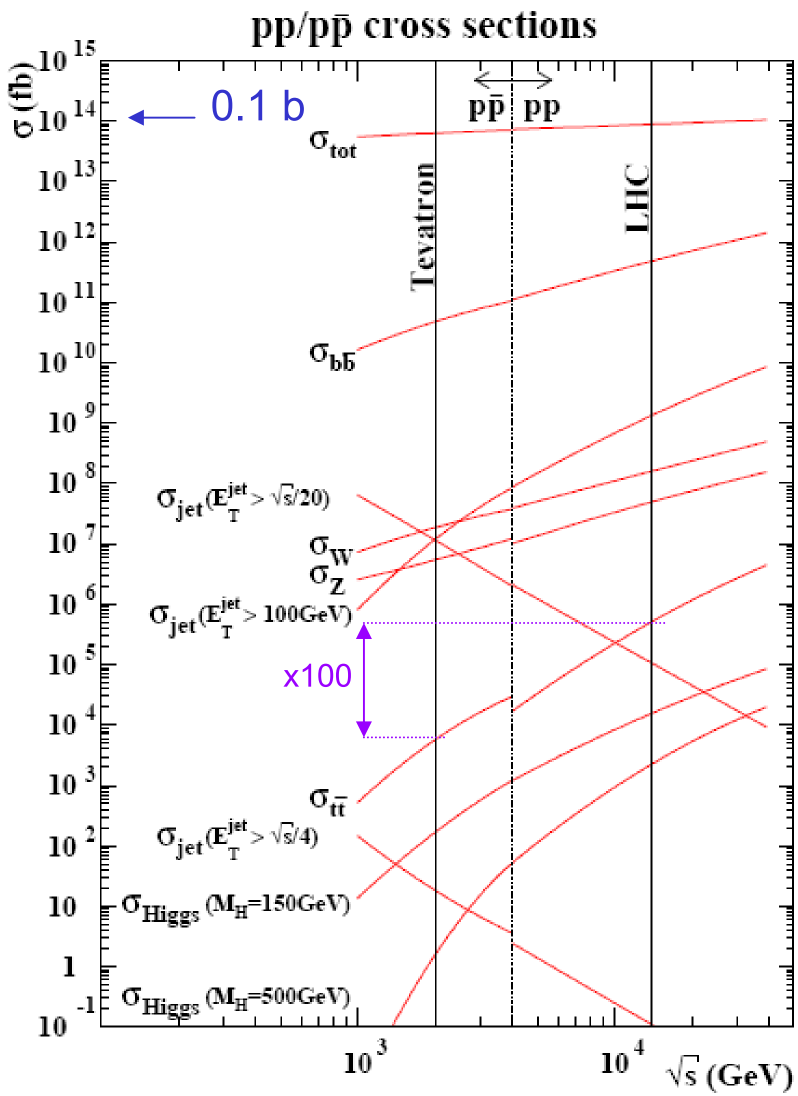
Here will focus on prospects for first measurements,

corresponding to a few 10 pb^{-1} up to a few 100 pb^{-1} at $\sqrt{s}$ = 10 TeV (although it will be at smaller $\sqrt{s}$ at the beginning of LHC).

LHC versus Tevatron

10 TeV LHC versus Tevatron:
- 50 times more top pairs and single top
- 5 times more W and Z

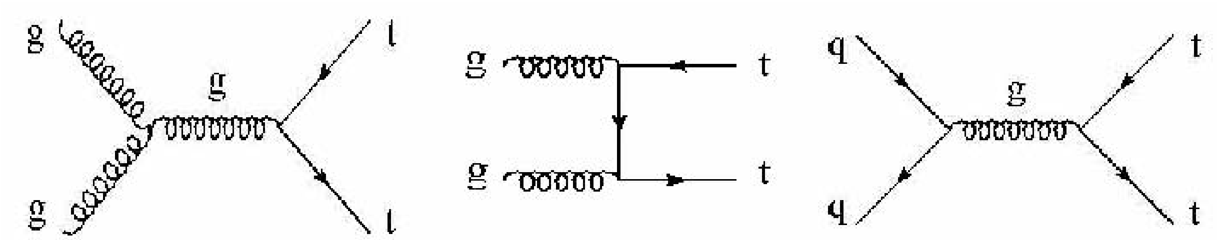
$\sqrt{s}$ (TeV)	pp cross section (pb) 7	10	14
QCD (p_t>80GeV)	~0.92 10^6 LO	~1.9 10^6 LO	~3.7 10^6 LO
$W \rightarrow l\nu$ +jets	24 000 (LO) * 1.14	40 000 (LO) * 1.14	60 000 (LO) * 1.14
$Z/\gamma \rightarrow ll$ +jets m_{ll}>50 GeV	2300 (LO) * 1.14	3700 (LO) * 1.14	7000 (LO) * 1.14
$W/Z \rightarrow l\nu/ll$ + cc/bb	~170	290 (LO)	450 (LO)
top pairs	187 NLO	414 NLO+NLL	908 NLO+NLL
single top (t channel)	65 NLO	130 NLO	246 NLO
WW/WZ/ZZ +jets (*$W/Z \rightarrow l\nu/ll$*)	3	9	14



top production at LHC

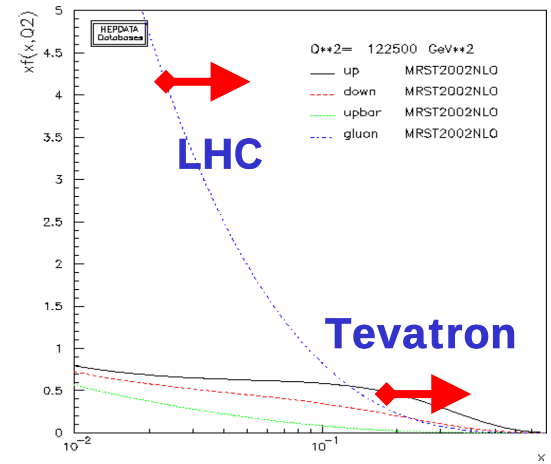
- top pair: strong production

~90% $gg \to t\bar{t}$ at 10-14 TeV LHC (~85% $q\bar{q} \to t\bar{t}$ at Tevatron)



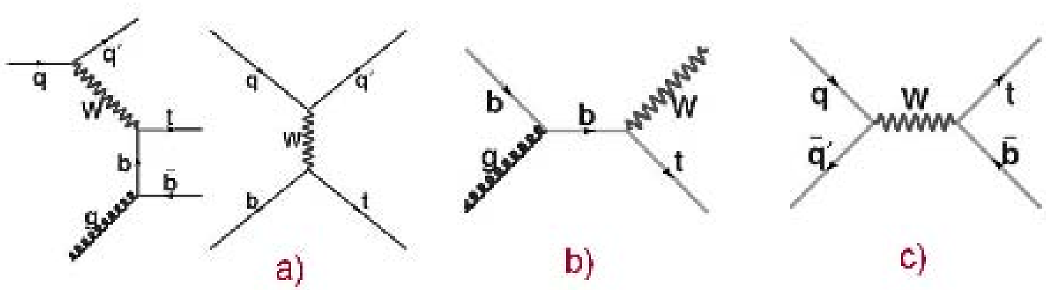
10 TeV: 414 ± 42 pb

(±10% renormalization and factorization scales, ±2% PDF)



Most studies presented here are performed at $\sqrt{s}$ = 10 TeV

- single top: electroweak production



10 TeV: 130 pb 29 pb 5 pb

allow a direct measurement of Wtb coupling

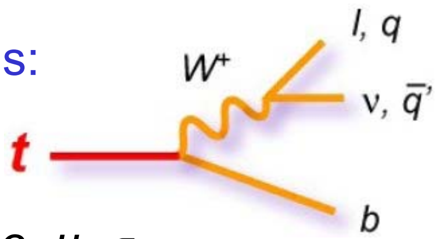
top pair final states

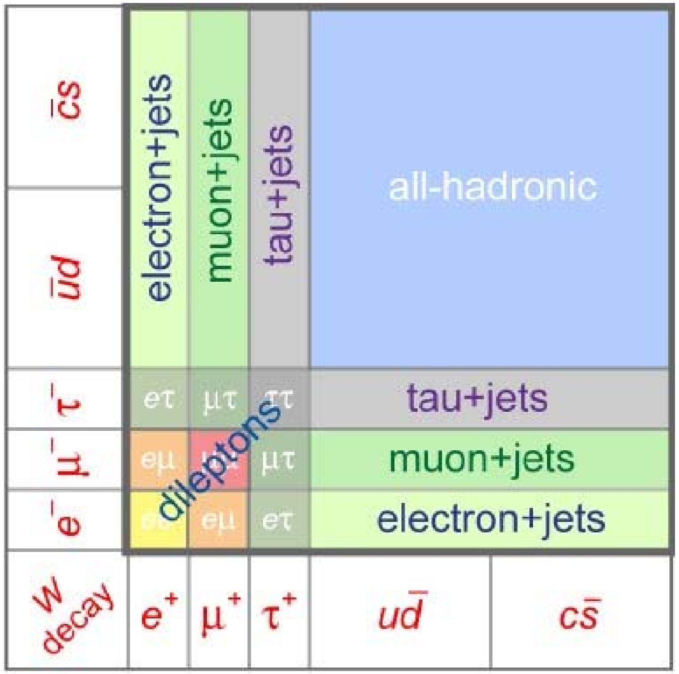
In SM with 3 generations:

BR($t \rightarrow Wb$) ≈ 1

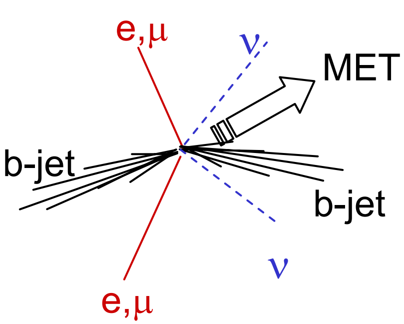
BR($W \rightarrow qq$) ~ 67%

BR($W \rightarrow l\nu$) ~ 11%, $l = e, \mu, \tau$

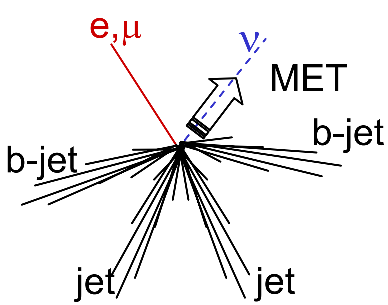




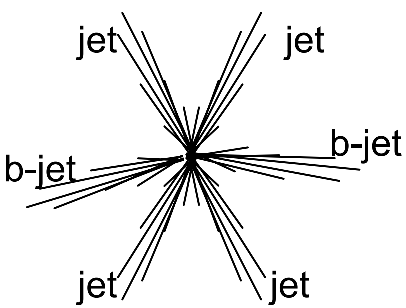
Most studies performed with dileptons or l+jets (with l = e, μ or $\tau \rightarrow e/\mu\ \nu\nu$)



dileptons
(BR~6%)



lepton + jets
(BR~34%)



all hadronic
(BR~46%)

Early top physics program

With a few 10 pb^{-1}:

- top re-discovery and first top pair cross-section measurement

With 100-300 pb^{-1}:

- BR($t \rightarrow Wb$) / BR($t \rightarrow Wq$)
- re-evidence for single top (t channel)
- first polarization studies in top production and decay
- first search for high mass $t\bar{t}$ resonances

With > 1 fb^{-1}:

- single top cross-section
- top mass
- top properties: top charge, *W* helicity in top decay, top spin correlation,
- search for rare top decays and for top in SUSY/BSM

Also top pair events will be used for **calibration with a few 100 pb^{-1}:**

- overall jet energy scale (using the *W* mass constraint)
- b-tagging validation
- MET validation

e/μ trigger, id & isolation

Trigger for top:

- *e*: based on em and had calo clusters at L1, matched to tracks at HLT
  here request $p_T(e) > 15$ GeV in CMS and ATLAS
- *μ*: based on segments in muon system at L1, matched to tracks at HLT
  here request $p_T(\mu) > 9$ (15) GeV in CMS (ATLAS)

Lepton identification:

- *e*: use calo shower shapes in η and φ and had/em energy fraction
- *μ*: matching quality between muon segments and inner tracking
  may also require MIP deposit in calorimeters

Efficiency measurements: will use $Z \rightarrow ee / \mu\mu$ tag-and-probe (with one well identified lepton and the other lepton as a probe)

Isolation: tracks and calo deposits within a cone $\Delta R = \sqrt{(\Delta\eta^2+\Delta\varphi^2)}$ around the lepton:

- CMS: use scalar sum of track p_T and calo E_T relative to *e/μ* p_T
  within $\Delta R < 0.2 – 0.3$
- ATLAS: cut on calorimetric $E_{Tiso} < 6$ GeV within $\Delta R < 0.2$

early $t\bar{t}$ cross-section

Considered scenarios (just an exercise in fact)

- CMS: with 20 pb^{-1} μ+jets or *e*+jets, 10 pb^{-1} *ee*+$\mu\mu$+*e*μ
- ATLAS: with 200 pb^{-1} μ+jets or *e*+jets , 200 pb^{-1} *ee*+$\mu\mu$+*e*μ

(including tau decays into *e* or μ)

- Rely on simple selections: no or limited MET cut, no b-tagging

- QCD and W/Z+jets have large uncertainties at LHC,
  also MC is not trustable for the amount of fake leptons

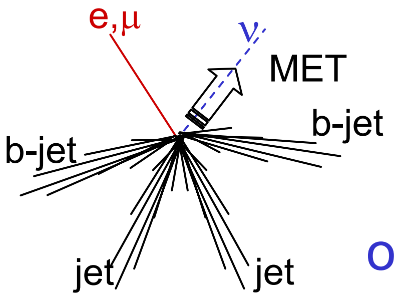
  (fake ~10^{-4} to 10^{-3} of QCD multijet events)

  $\Rightarrow$ estimate them with data driven tools !
- Rely on MC for other physics backgrounds:
  hadronic *tt*, single top, *WW* / *WZ* / *ZZ*

- Measurement:

$$\sigma_{t\bar{t}} = \frac{N_{obs} - N_{bgd}}{A \times \epsilon \times \mathcal{L}}$$

with weighting (CMS) or max. likelihood (ATLAS) technique
to combine the dilepton channels



lepton (e/μ) +jets: selections

only 1 muon or 1 electron with $p_T > 20$ GeV
and $|\eta| < 2.4 - 2.5$ (but no electron within $1.4<|\eta|<1.5$)

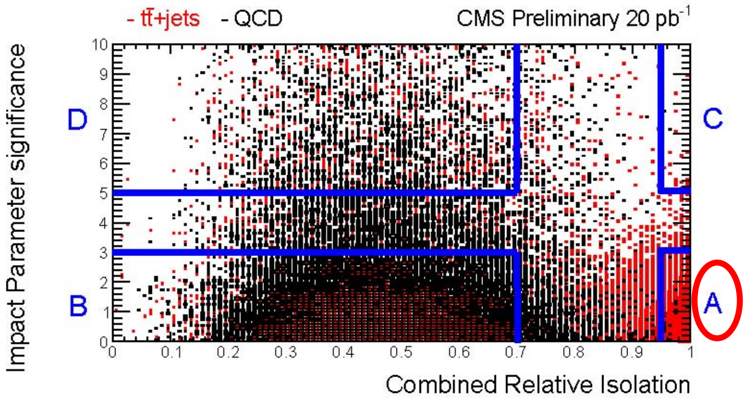
- μ transverse impact parameter < 200 μm in CMS (reduce *b/c* decays), reject the event if other non-isolated muon in ATLAS
- isolation cut based on track+calo in CMS, only on calo in ATLAS
- require ≥4 jets with $p_T>30$ (40) GeV and $|\eta|<2.4$ (2.5) in CMS (ATLAS)
  (in ATLAS the 4th jet can have $p_T>20$ GeV)
- in ATLAS, 2 kinds of selection: using MET with MET > 20 GeV
  or not using MET but with tighter kinematic cuts:
  $p_T^{e(\mu)} > 40$ (30) GeV, $|\eta^{e(\mu)}| < 1.4$ (1.5), $\Sigma_{e/\mu,j2,j3,j4}\, p_T > 160$ GeV,
  (4th jet $p_T > 30$ GeV for μ+jets)
- no b-tagging

e/μ+jets: QCD bgd estimate

μ **CMS:**

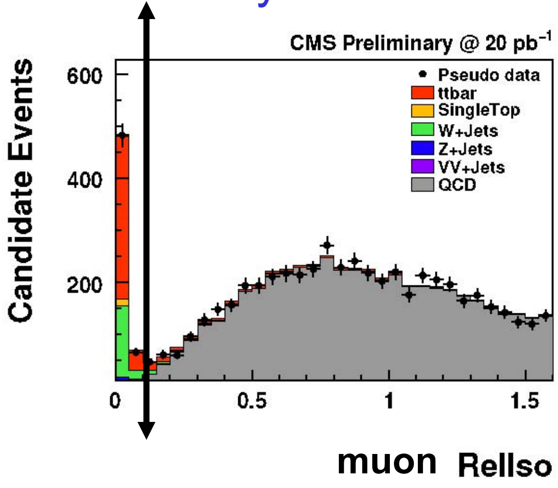
- quadrant method:
  using 2 weakly correlated variables, muon impact parameter significance and muon isolation:

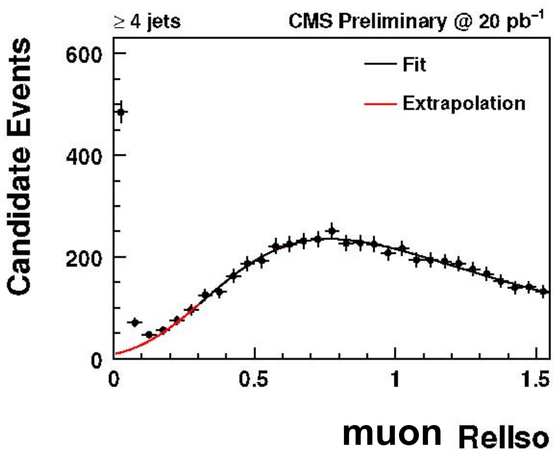
  $$N_A/N_B = N_C/N_D \Rightarrow N_A = N_B N_C/N_D$$



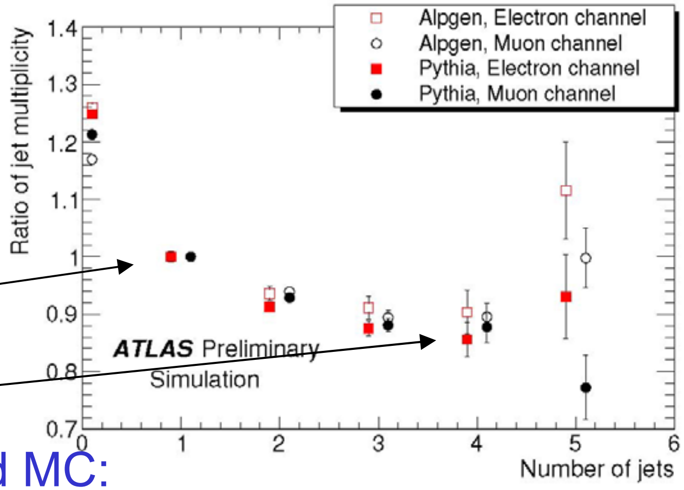
e/μ **CMS:**

- use real data sample to estimate the QCD background (with low MET and tighter *Z→ee/μμ* veto) on the lepton isolation
- extrapolation of lepton isolation below the cut
- check consistency for events with 1, 2 jets and apply to 3, ≥ 4 jets





e/μ+jets: W+jets bgd estimate



e/μ +jets ATLAS:

- ratio of *W* and *Z* amounts has much smaller uncertainty than *W* amount alone
- this *W/Z* ratio can be measured in a control region (CR) with 1 jet (p_T>20 GeV) then extrapolated to a signal region (SR) with ≥ 4 jets)
- using only ratios SR/CR for *W* and *Z* in data and MC:

$$C_{MC} = \frac{(W^{SR}/W^{CR})_{MC}}{(Z^{SR}/Z^{CR})_{MC}} \quad (W^{SR}/W^{CR})_{data} = (Z^{SR}/Z^{CR})_{data} \cdot C_{MC}$$

- Limitations due to statistics: with 200 pb^{-1}, expect

in CR: ~10k (16k) $Z \to ee$ ($\mu\mu$), 150k (190k) $W \to e\nu$ ($\mu\nu$); in SR: ~80 (150) $Z \to ee$ ($\mu\mu$)

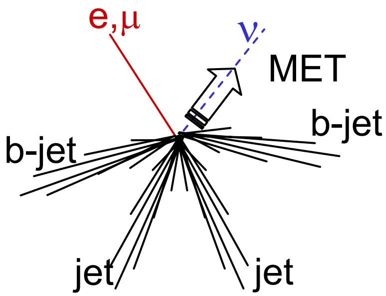
and to the residual QCD bgd (estimated also from data with ±50% uncertainty)

⇒ overal uncertainty of about ±20% on *W*+jets normalization

μ+jets CMS:

- With ~200 pb^{-1}: can estimate the *W*+jets background from the *W* charge asymetry $(N_{W+} - N_{W-}) / N_W = (N_{\mu+} - N_{\mu-}) / (N_{\mu+} + N_{\mu-})$

  ⇒ statistical uncertainty of about ±20%, but PDF systematics to be evaluated



e/μ+jets: hadronic top

The amount of top pair events can be extracted from the reconstructed hadronic top mass

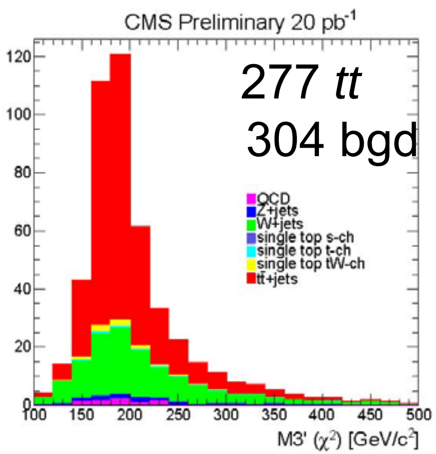
⇒ use the 3 jets combination with highest scalar sum of p_T (CMS and ATLAS)

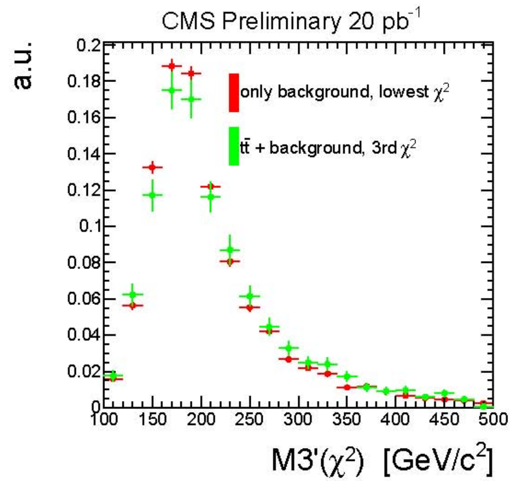
- or in **CMS *μ*+jets:**

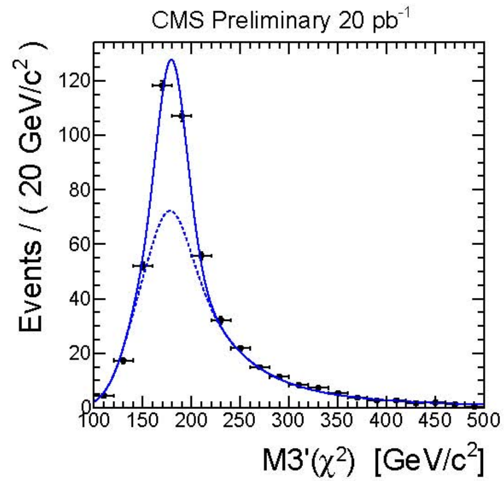
keep the 4 jets combination with smallest χ^2 of W_{jj} / t_{jjj} / $t_{jl\nu}$ masses

$$\chi^2 = \frac{(m_{j_1 j_2} - m_W)^2}{\sigma_{jj}^2} + \frac{(m_{j_1 j_2 j_3} - m_t)^2}{\sigma_{jjj}^2} + \frac{(m_{\mu\nu j_4} - m_t)^2}{\sigma_{\mu\nu j}^2},$$

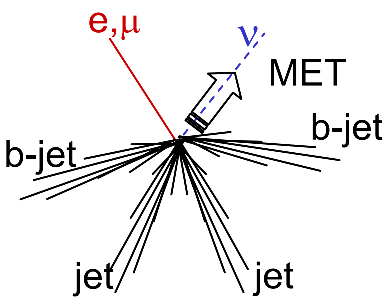
the combinatorial bgd can be inferred from data with a less good χ^2:







e/μ+jets: cross-section



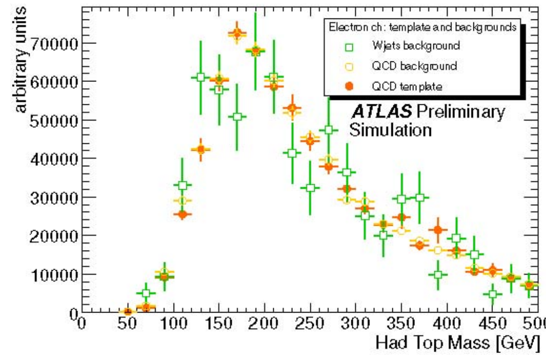
or in **ATLAS *e/μ*+jets:**

request 2 jets within ±10 GeV of M_W

⇒ good bgd rejection and allow to

estimate the shape of the combinatorial bgd

by inverting M_W cut or by using QCD multijet data



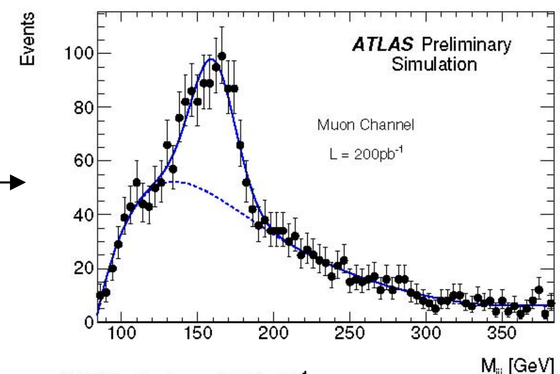
- Then fit of hadronic mass (or η(muon) in CMS) → amount of $t\bar{t}$ → cross-section:

μ or e ATLAS (200 pb⁻¹):

± 3 % (stat) ±13 % (syst) (count)

±15 % (stat) ±10 % (syst) (fit) ⟶

± 6 % (stat) ±12 % (syst) (fit no MET)



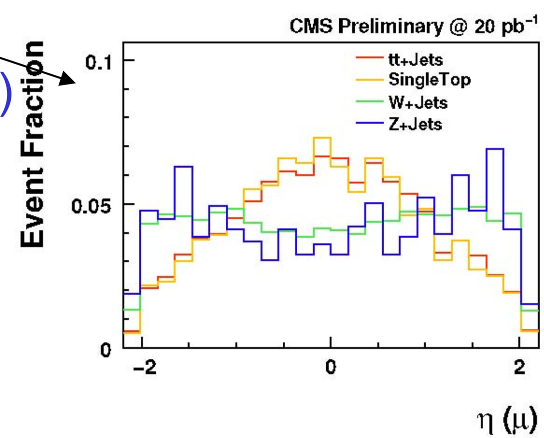
μ CMS (20 pb⁻¹):

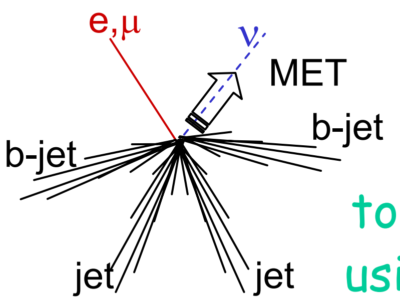
±18% (stat) ±19% (syst) (fit η(muon))

± 9% (stat) ±22% (syst) (boost dec. tree)

e **CMS (20 pb⁻¹):**

±23% (stat) ±20% (syst) (fit mass)



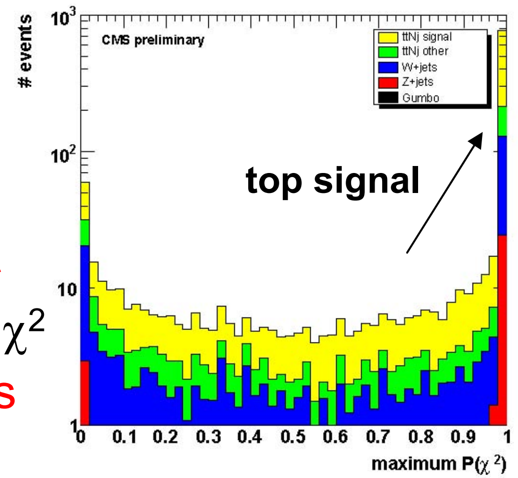


Jet energy scale with top

top pair events in e/μ + jets can be used to calibrate the jet energy, using the well known W mass (and eventually top mass) constraint(s).

Measuring top mass at ±1 GeV requires jet energy at ±1% !

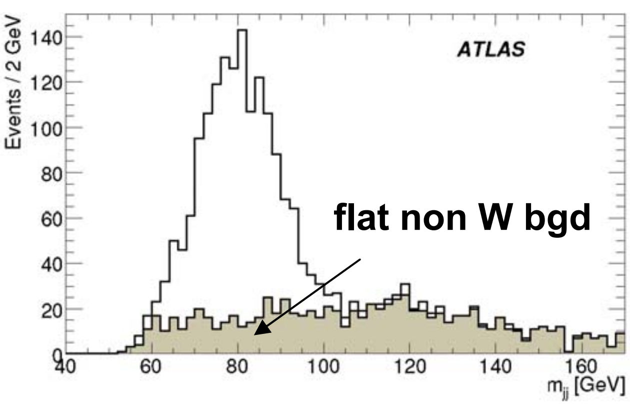
Studies have been performed in CMS and ATLAS at 14 TeV.



- **CMS μ+jets (100 pb^{-1}):** cut on likelihood ratio based on jets (p_T, η, φ) and b-tag discriminant to get the correct jet combination for each top. Parametrize resolutions on jets (p_T, η, φ) and perform a kinematic fit *with W and top mass constraints*. Retain the jet energy correction with smallest χ^2

  ⇒ ±1% precision on jet energy, both for b and light quarks

- **ATLAS e/μ+jets 800 pb^{-1}:**



exactly 4 jets of p_T>40 GeV, with 2 b-tagged jets

⇒ W→jj mass (*no top mass constraint*) with 80% purity

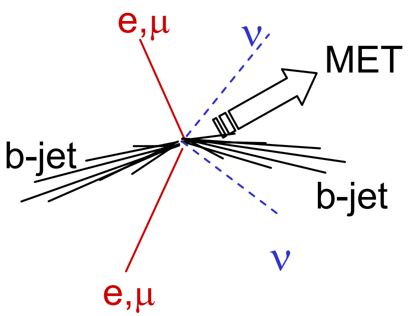
2 methods: iterative rescaling or templates.

Systematics due to bgd estimate, bias due to p_T cut, … should be less than ±0.5%.

Extrapolating to a lower luminosity (50 pb^{-1})

⇒ ±2% precision on jet energy for light quarks

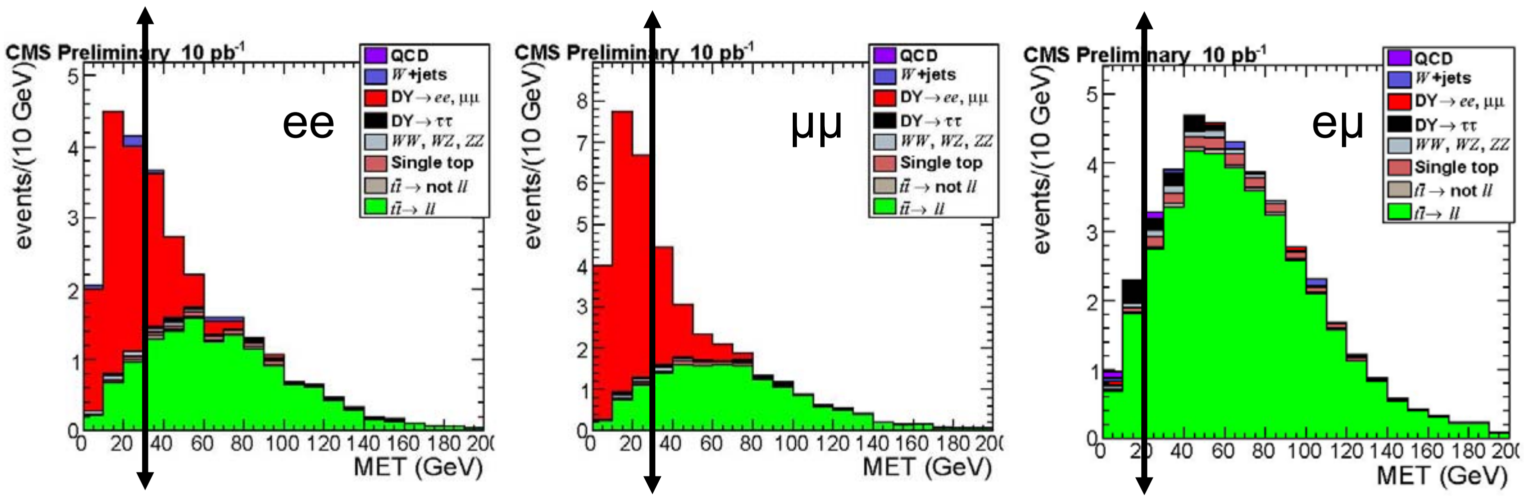
Note that these are overall resolutions, much more statistics is needed for detailed p_T, η studies



dilepton+jets: selections

consider *ee, μμ, eμ* of opposite charge
p_T > 20 GeV for each lepton
(|η| < 1.37 or 1.52 < |η| < 2.47 for e in ATLAS)

- reject *ee* and *μμ* pairs within $M_Z \pm$ 15 (5) GeV in CMS (ATLAS) → remove DY
- request isolated leptons, keep the 2 highest p_T (or more isolated) leptons
- request ≥ 2 jets with p_T > 30 (20) GeV and |η| < 2.4 (2.5) in CMS (ATLAS)
- MET > 30-35 (20) GeV in *ee/μμ* (*eμ*) events → remove DY, QCD



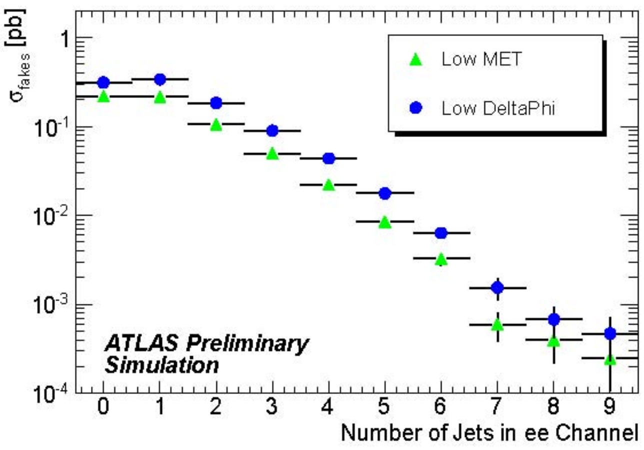
dilepton+jets: data driven bgd estimate

DY+jets:

- **CMS 10 pb^{-1}:** get nb. of events inside *Z* window, use MC to extrapolate outside (residual bgd corrected from $e\mu$ events) $\rightarrow \pm 30\%$ uncertainty
- **CMS 100 pb^{-1}:** infer the DY shape from low MET events
- **ATLAS 200 pb^{-1} :** use quadrant method in MET / *Z* mass plane, advantage: only ratio of events are used, for Data and MC separately.

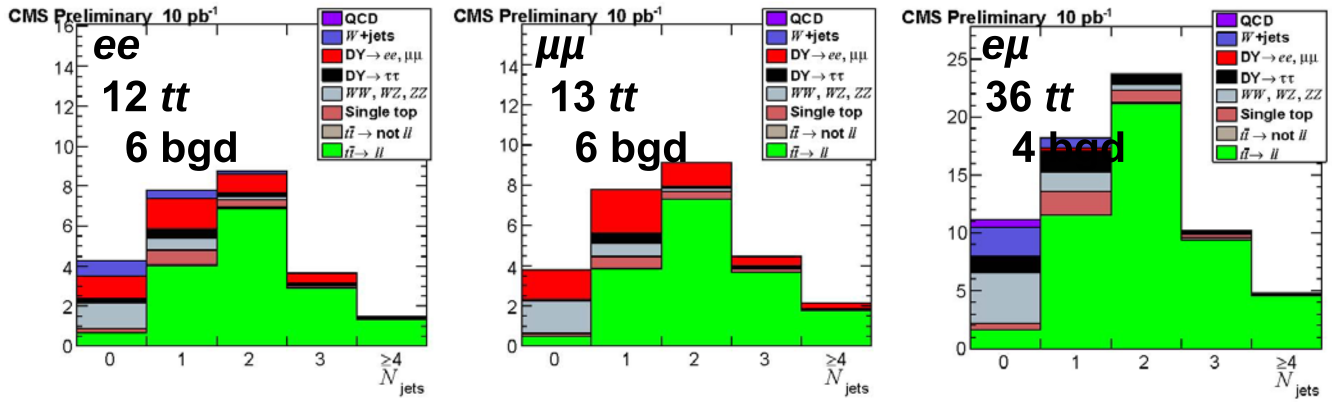
Fake leptons (mainly e) in *W*+jets and QCD:

- **CMS 10 pb^{-1}:** relaxed lepton id and isolation, compared to full selection in multijet events $\rightarrow \pm 50\%$ uncertainty
- **ATLAS 200 pb^{-1}:** define loose (L) and tight (T) lepton id, compute loose-to-tight efficiencies for real (r) and fake (f) leptons; efficiencies can be measured:
  in r from tag-and-probe $Z \rightarrow ll$;
  in f from 2 orthogonal low/high MET regions;
  then estimate the nb. of r and f leptons by solving 3 equations in TT, TL, LL;
  to be parameterised w.r.t. N_{jets}.



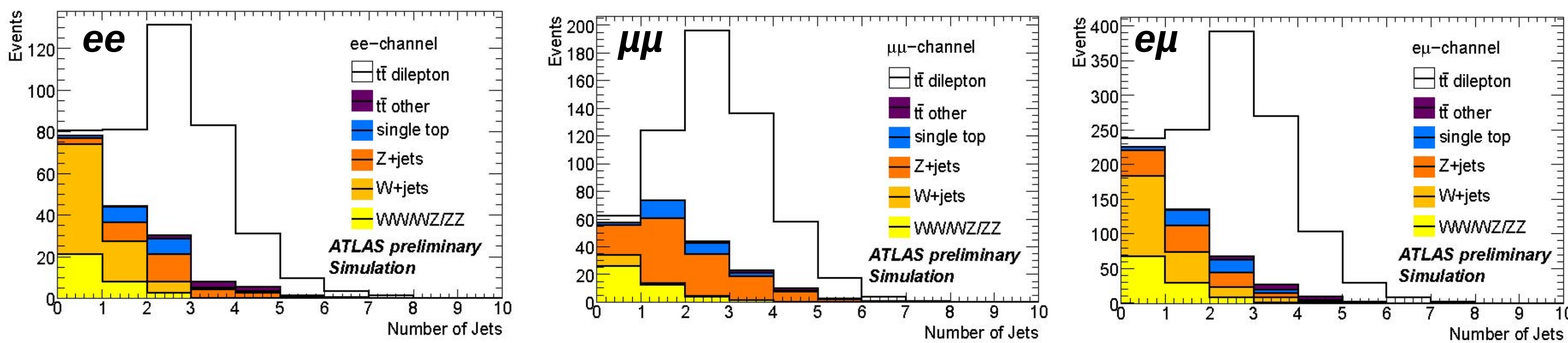
dilepton+jets: $t\bar{t}$ cross-section

- **CMS 10 pb^{-1}:** combined (weighting) σ_{tt} ±15% (stat) ±10% (syst)

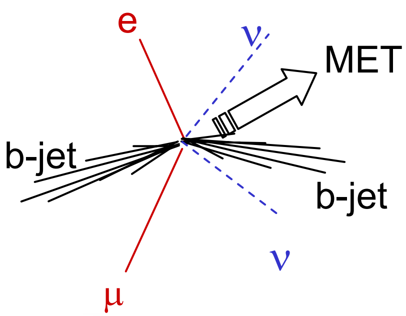


S/B~2 in *ee*, *μμ* (main bgd=DY), ~9 in *eμ*

- **ATLAS 200 pb^{-1}:** combined (max. likelihood) σ_{tt} ± 3% (stat) ± 9% (syst)



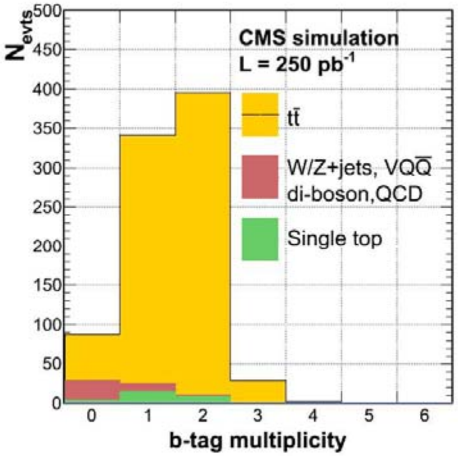
- Note that a quite pure signal could be easily obtained with b-tagging:
  S/B~10 (16) in all channels if ≥ 1(2) b-tagged jet (CMS)

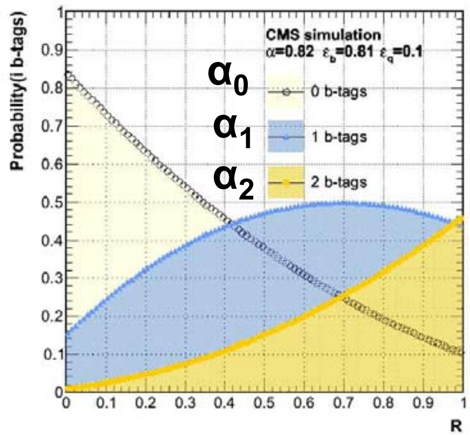


R=BR(t→Wb)/BR(t→Wq)

A consistency test of the 3 families SM (expect R=1).

CMS: dileptons (eμ) + jets, 250 pb^{-1}: with loose b-tagging





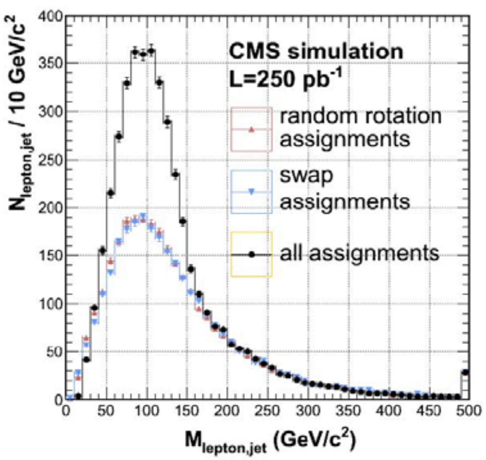
nb. of observed b-tagged jets depends on : R and b-tag efficiency ε_b

⇒ fix R = 1 and fit ε_b

or take ε_b from data measurement and fit R+another free parameter

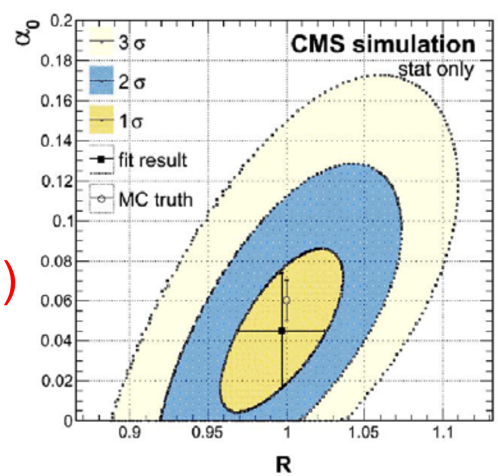
(α_0 = fraction of events with 0 b-jets)

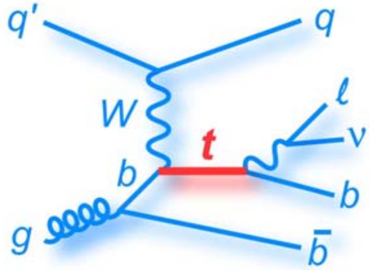
The nb. of right b-jet assignments can be directly measured on data:



from fit to $M_{lepton\text{-}jet}$ distribution where the wrong assignment bgd can be estimated from data

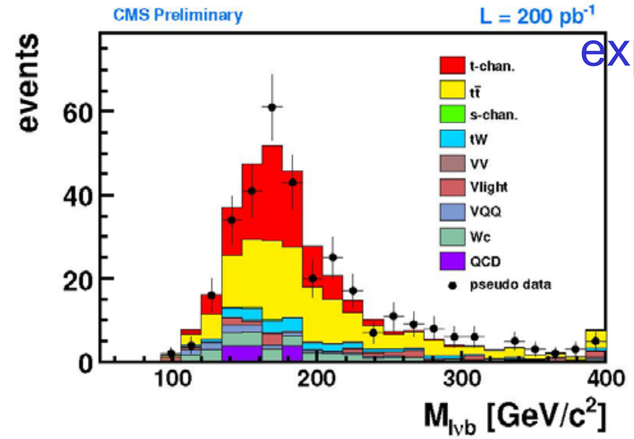
⇒ gives R within ± 2%(stat) ± 9%(syst ε_b) ± 3%(syst bgd)





Single top in μ+jets

CMS 200 pb^{-1}:
similar to top pairs, but
- exactly 2 jets of p_T > 30 GeV with only 1 tight b-tag (no other loose),
- reduce QCD with transverse W mass $M_T(W)$ > 50 GeV,
- then top mass reconstruction:



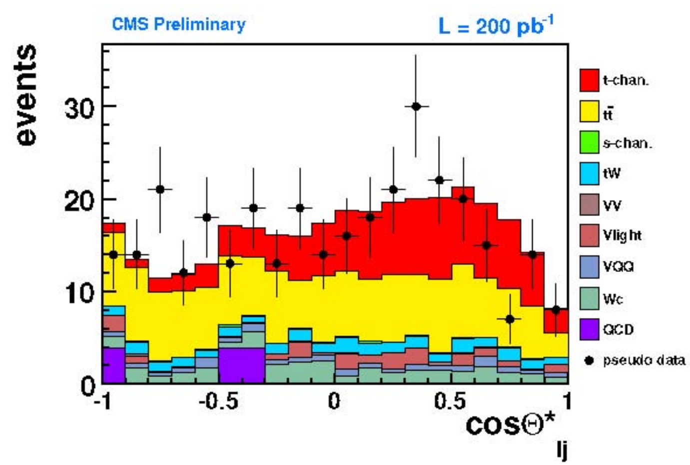
expect S/B~0.4
~100 t chan,
~ 25 *tW* + s
~140 *tt*
~ 70 *W/Z*+X

Data driven QCD and W+jets bgd :
from $M_T(W)$ fit:
QCD shape from revert μ isolation and no b-tag
W shape from $Z\to\mu\mu$ sample with reweighting to $W\to\mu\nu_\mu$ $\Rightarrow$ bgd yield ± 50%

Single top cross section:
from fit to polarization angle: with $\cos\theta_{lj}^* < 0.75$
with flat bgd and MC signal

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

$\Rightarrow$ uncertainty ±35% (stat) ± 15% (syst)
sensitivity $\approx 2.7\sigma$ for 200 pb^{-1}
(would be >5σ above 700 pb^{-1})

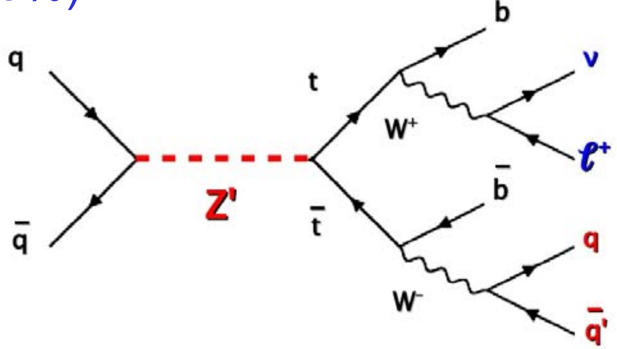


Search for high mass $t\bar{t}$ resonances

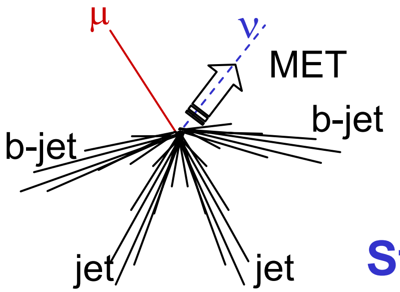
Several Beyond Standard Model extensions predict enhanced coupling to the 3rd family top quark.
Here search for a narrow Z' resonance decaying into top pairs.
Tevatron limit on leptophobic Z': $m_{Z'} > 820$ GeV

- Here assumes Z' with SM couplings and relative width $\Gamma/m \sim 1\%$ (less than experimental resolution $\sigma_m/m \sim 5\text{-}10\%$)

- CMS 100 pb^{-1} at 10 TeV:
  μ + jets or fully hadronic,
  ATLAS 1 fb^{-1} at 14 TeV : μ + jets



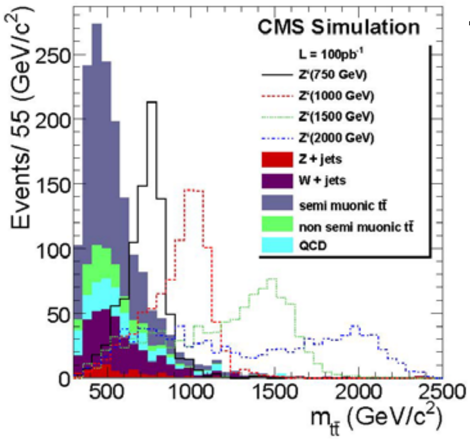
- Challenge for m(tt) > 2 TeV:
  identify highly boosted $t \rightarrow bqq$ decays as a monojet,
  with under control backgrounds: QCD, W+jets and standard top pair.



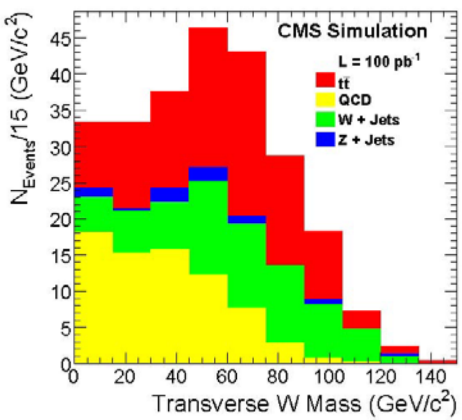
high m_{tt}: μ+jets (standard)

Standard selection (no boost top)

isolated μ, but reduce $b \to \mu$ with
. $\Delta R(\mu,\text{jet}) > 0.4$, $p_T^{rel}(\mu,\text{jet}) > 35$ GeV,
reduce QCD with ≥ 4 jets of $p_T > 35$ GeV,
reduce *Z* with 2nd lepton veto,
choose mass combination with smallest χ^2
$\Rightarrow \varepsilon_{tt} \approx$ 13-17%, $\varepsilon_{Z'} \approx$ 24% up to 2 TeV



100 pb^{-1}
expect
~900 SM tt
~700 bgd

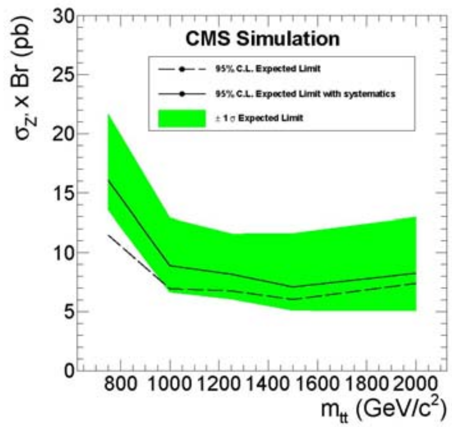


Data driven QCD and W+jets bgd estimate:

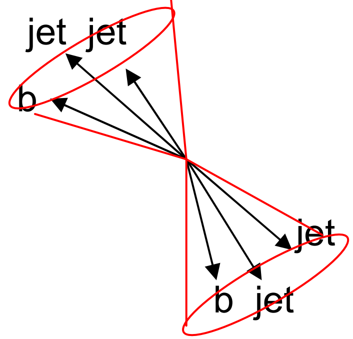
reverse $\Delta R(\mu,\text{jet})$ and $p_T^{rel}(\mu,\text{jet})$ cuts,
or require $H_T = \Sigma\, p_T < 350$ GeV,
and fit the transverse *W* mass
(fixing *W*+jets and *Z*+jets shape from MC)

CMS 100 pb^{-1} 10 TeV:

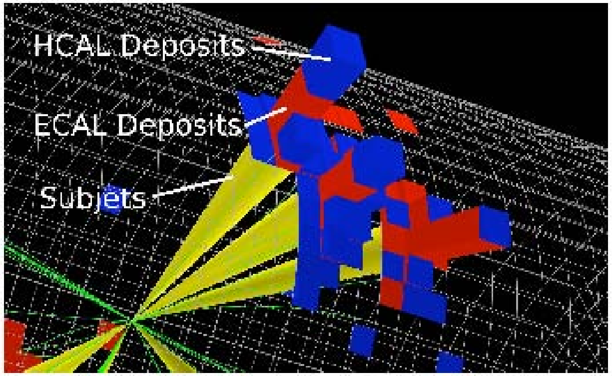
95% CL σ·BR exclusion



8.3 pb for m_{tt} = 2 TeV

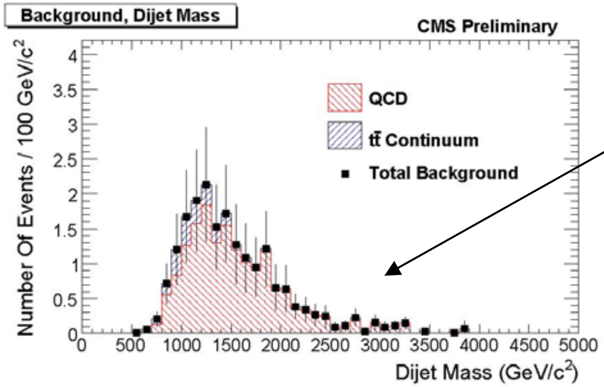


high m_{tt}: all hadronic (boost)



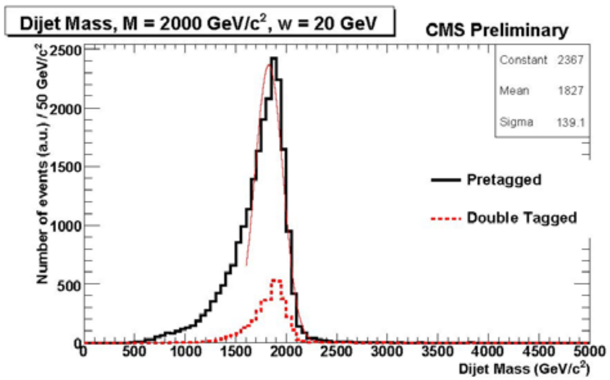
Strategy for boosted top jets:
k_T like algo for monojet substructure kinematic cuts (using *W* and top mass constraints) to reject about 98% of light *q/g* jets at p_T > 600 GeV, with 46% efficiency for top $\rightarrow$ tagged monojets

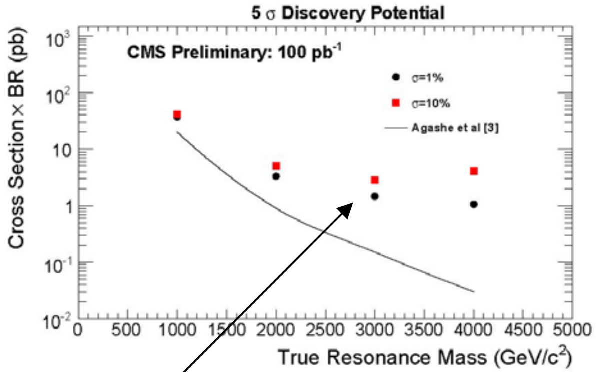
- **Selections:** 2 tagged monojets with p_T > 250 GeV, |y| < 2.5



Data driven bgd estimate:
W/Z+jets, QCD, single top using anti-tagged jets

top pair continuum taken from MC with ±factor 2 systematic

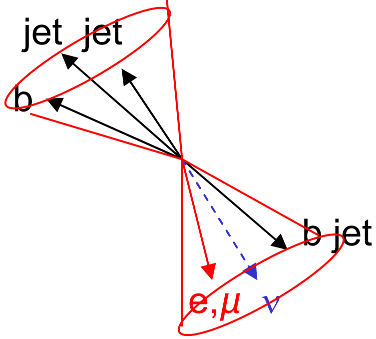




CMS 100 pb^{-1} 10 TeV:

for m_{tt} = 2 (3) TeV:
discovery up to to σ·BR = 4.0 (1.6) pb
95% CL exclusion up to 1.5 (0.7) pb

high m_{tt}: e/μ+jets (boost)



Selections for boosted top jets:

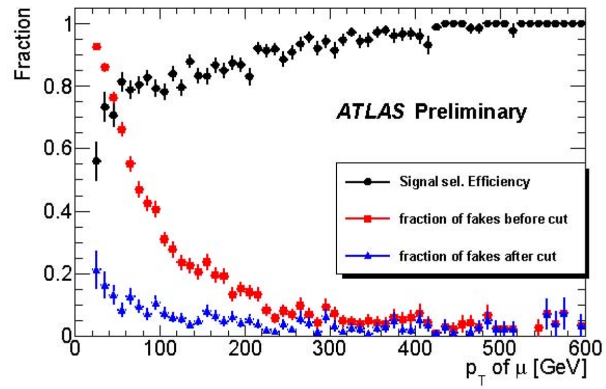
leptonic top: (no isolation !)

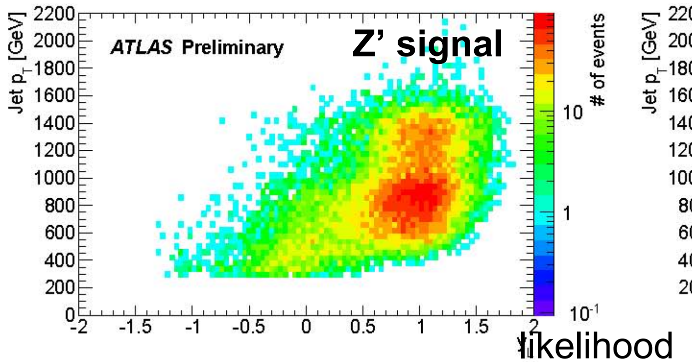
highest p_T *e* or *μ* within $\Delta R(l,\text{jet}) < 0.6$

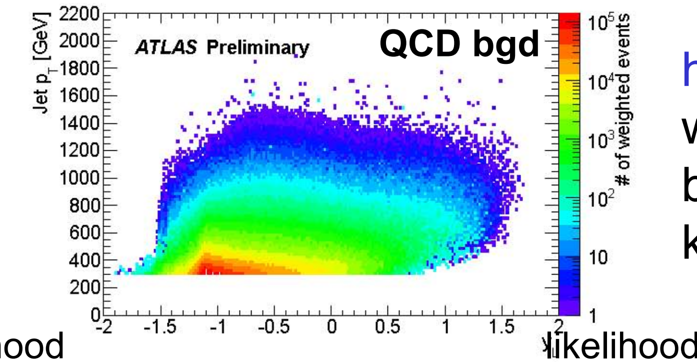
reduce fakes, b→e/μ and combinatorial background

with 2D cut in $p_T^{rel}(l,\text{calo})\cdot\Delta R$ vs m_{calo}/m_{l+calo}

request MET>20 GeV and estimate ν momentum







hadronic top monojet:

with k_T algo

build likelihood based on

k_T splittings and jet mass

Z' mass resolution: can reach $\sigma_m/m \sim 5\%$, but needs proper out-of-cone corrections

Z' and bgd efficiencies: after likelihood selection, *W*+jets are removed,

main bgd is top pairs + QCD: $\varepsilon_{Z'} \sim 0.1\text{-}0.4$, $\varepsilon_{QCD} \sim 10^{-5}\text{-}10^{-4}$ for jet $p_T > 600$ GeV

ATLAS 1 fb^{-1} 14 TeV, for m_{tt} = 2 (3) TeV: 95% CL exclusion up to 0.6 (0.2) pb

the method is quite promising, even with reduced luminosity and cm. energy

Conclusion

- The top physics program at LHC is rich and opens a new window for detailed standard model tests and for the search of new physics

- Observing top events in the early running is challenging because it requests the simultaneous reconstruction and identification of most physics objects: jets, muons, electrons

- but also missing transverse energy and b-tagging for calibration purpose and for more detailed studies

- A new energy domain will be explored at LHC,
  some early surprise may happen, even in top physics.

BACK-UP

Monte Carlo

- Use full simulation (generator+GEANT4) for signal studies and background rejection.
- However with ideal (1 fb^{-1}) calibration and alignment, and no pile-up.
- Normalize processes to NLO cross-sections (when known).
- Use fast simulation for some systematic studies.

CMS:

- *tt +jets*, *W/Z +jets, W +bb/cc, WW/WZ/ZZ +jets*, *single top*: **MadGraph** matrix element interfaced to **PYTHIA** for parton shower,
- *QCD multijets*: **PYTHIA+GEANT4** filtered with muon or electron

ATLAS:

- *tt +jets*: **MC@NLO** generator + **HERWIG** for fragm. + hadronization
- *single top*: **AcerMC** generator + **PYTHIA**
- *W/Z +jets, W +bb/cc*: **ALPGEN** generator matched to **HERWIG**
- *WW/WZ/ZZ +jets*: **HERWIG** or **MC@NLO** + **HERWIG**
- *QCD multijets*: **PYTHIA** dijets, or **HERWIG** fast simulation

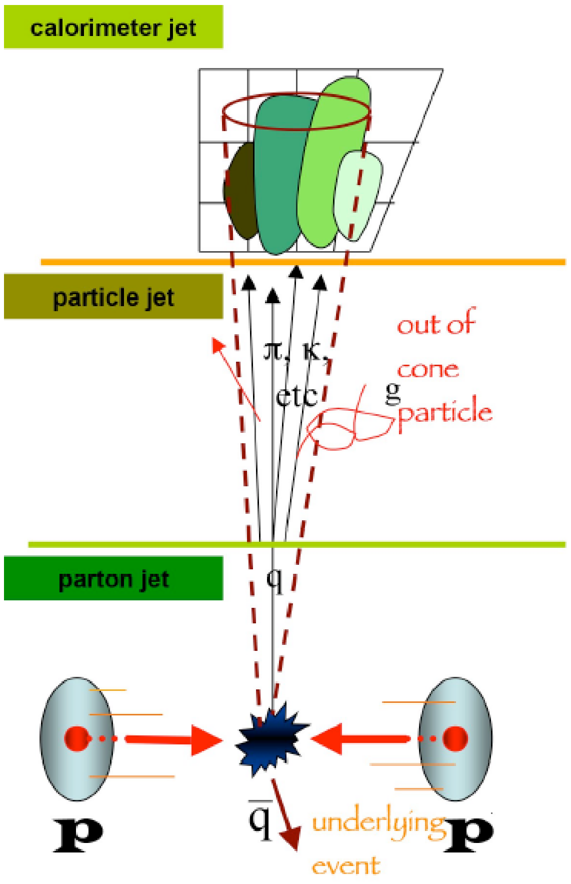
Jet and Missing E_T (MET)

Jet id: use cone algorithms with calo clusters within

$\Delta R=\sqrt{(\Delta\eta^2+\Delta\varphi^2)}$

- CMS: seedless infrared-safe cone algo with $\Delta R < 0.5$
- ATLAS: requires $\Delta R < 0.4$

don't consider jets too close from an electron with $\Delta R(\mathrm{jet},e) < 0.3$ (0.2) in CMS (ATLAS)



Correct for many effects (in E_T, η bins), as

- non-linearities in calo response,
- loss of charged particles in B field bending
- poorly instrumented regions ($|\eta| \sim 1.5$)

Systematics: assume ±10% in CMS,
±5%($|\eta|<3.2$) to ±10%($|\eta|>3.2$) in ATLAS

MET: computed as the vector sum of calo towers within $|\eta| < 5$, corrected for identified muons, electrons and for calibrated jets

$t\bar{t}$ cross-section: systematics

- **Luminosity:** assumed to ±10% (±10-20%) in CMS (ATLAS)

- **Jet energy scale:** ±5-8% (±3-4%) for dileptons, ± 15-17% (± 4-9%) for l+jets

- **Trigger and Lepton id efficiency :** ±6% (±3-5%) in CMS (ATLAS)

- **Theory:** typically ±6-7% for dileptons, ± 9-11% for l+jets
  - **ttbar model:** compare MadGraph, Pythia, MC@NLO in CMS (± 2-10%), MC@NLO, AcerMC and Alpgen in ATLAS (± 3-7%)
  - **ISR/FSR:** vary Λ_{QCD} and p_T cut off ±3% (± 4-13%) in CMS (ATLAS)
  - **single top, WW/WZ/ZZ:** vary theoretical uncertainties ⇒ ±1%
  - **PDF:** apply CTEQ6.6 reweighting ⇒ ±5% (±2%) in CMS (ATLAS), also compare with MRST2006nnlo in ATLAS

- **Data driven bgd estimates:** ±6% (±6%) for *eμ*, ±12% (±2-10%) for *ee/μμ*
  ±6-7% (±3-8%) for l+jets in CMS (ATLAS)