Top quark pair cross-section measurement using CMS data at a 7 TeV centre-of-mass energy

jérémie lellouch
on behalf of the CMS collaboration

CRIMEA2011
September 2011

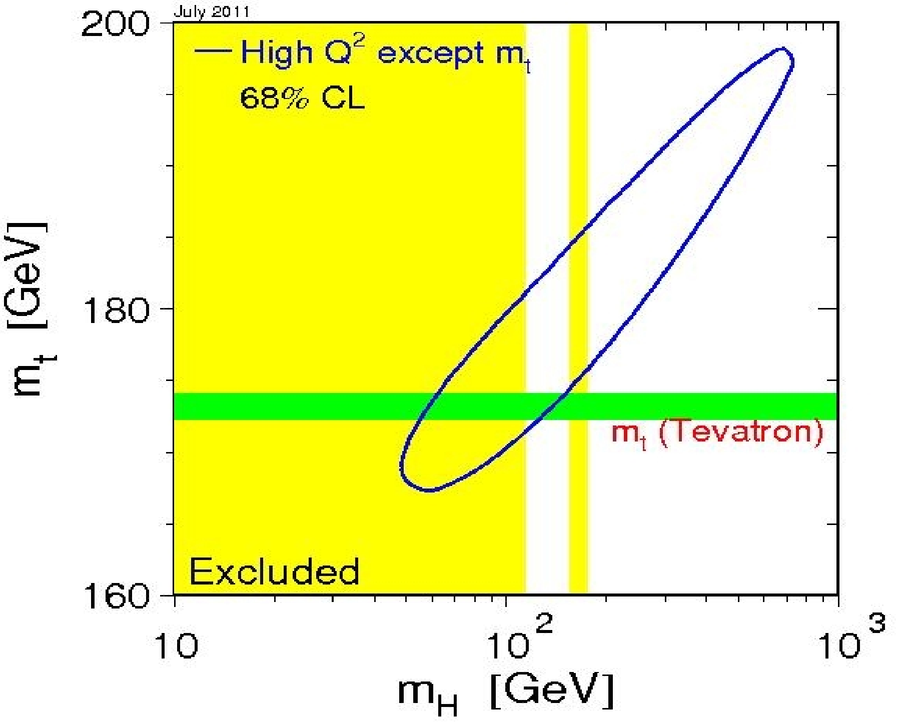
Table of contents

- Introduction: motivation for top quark studies
- LHC and the CMS detector
- Production and decay of the top quark at the LHC
- Cross-section measurement strategy
- Selected review of the CMS results
- Combination of the CMS results and comparison with theory
- Conclusion and outlook

Motivation for top quark studies

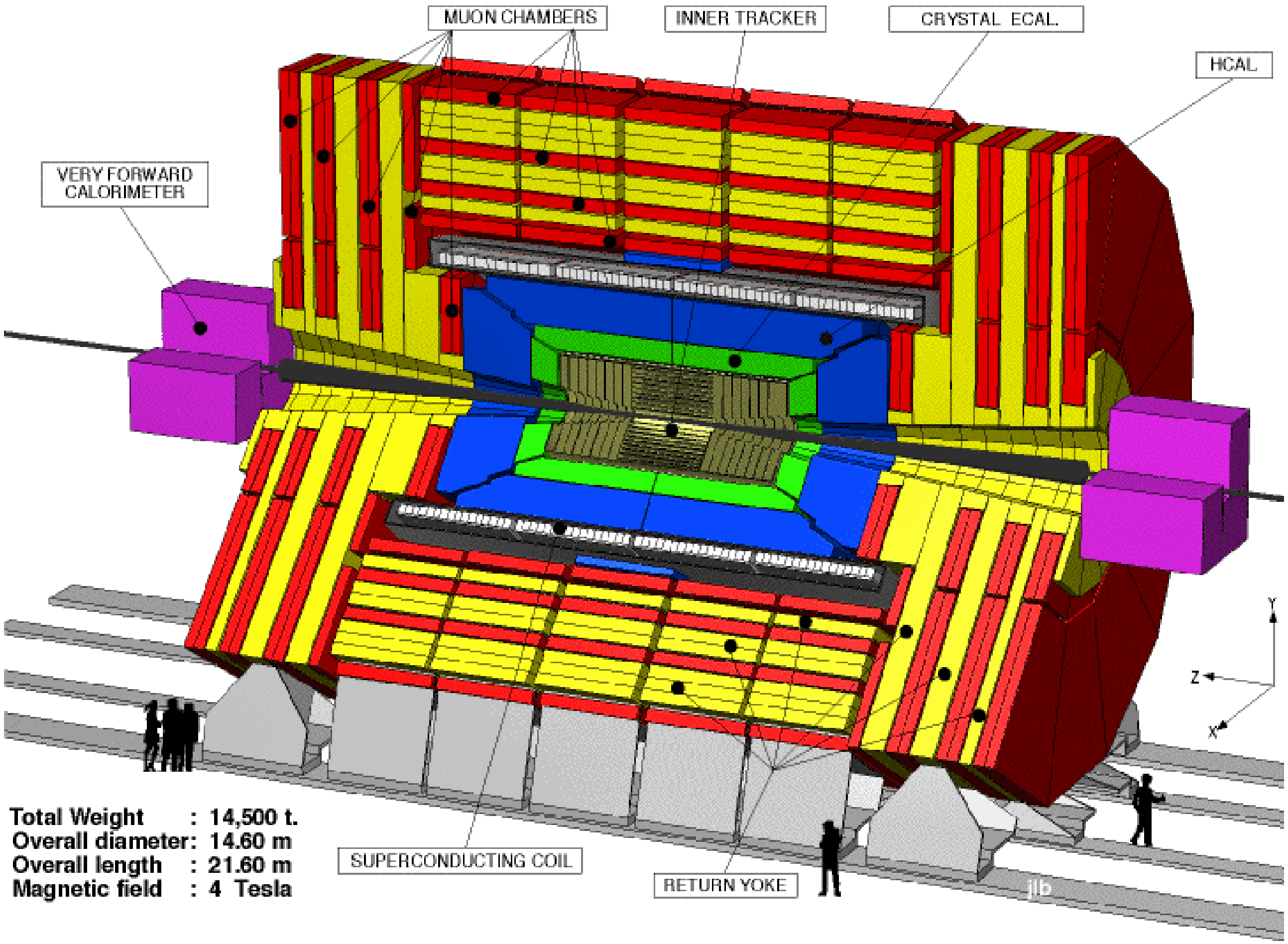
- Discovered in 1995 (Tevatron) but properties not very well known
- Decays before hadronisation → access to bare quark level
- A key to the secrets of the Standard Model and beyond

  - Maximum coupling to the Higgs field → low Higgs mass preferred
  - Allows precision measurements of SM parameters: total and diferential cross-sections, mass, charge, asymmetry, also V_{tb} element of the CKM matrix,...
  - Good window to new physics



  - Also a good benchmark topology for a wide array of analysis ingredients: lepton identification, jet energy scale, b-jet identification,...

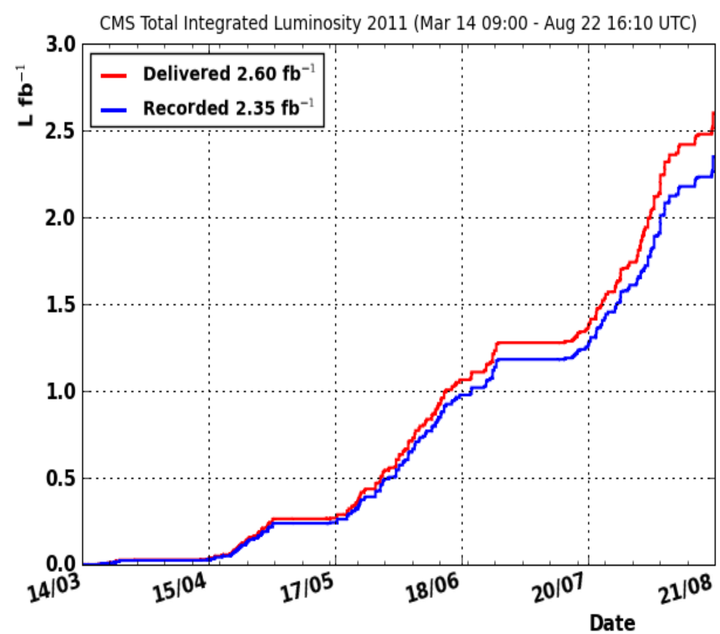
The CMS detector



36 pb^{-1} recorded in 2010

2.3 fb^{-1} already on tape in 2011

Machine and data acquisition system perform very well



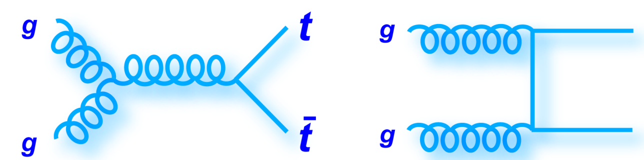
Good lepton identification

Hermetic calorimetry, fine lateral granularity

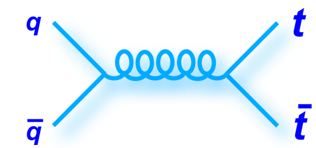
Excellent all-silicon tracking → allows for improved jet and MET reconstruction ("particle-flow" algorithms) and efficient b-tagging

Production and decay at the LHC

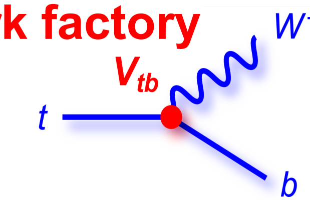
- Produced in pairs via gluon fusion or quark annihilation
- Gluon fusion dominates at the LHC (7 TeV) while quark annihilation dominates at the Tevatron (~ 2 TeV)



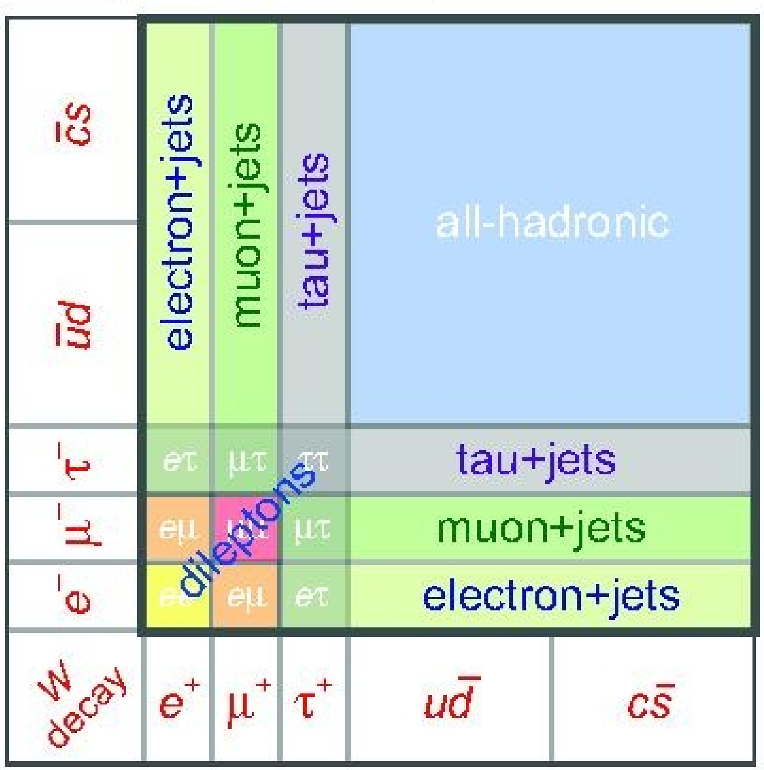
Kidonakis [PRD 82 (2010) 114030]: σ_{tt} (7 TeV) = 163 + 11 – 10 pb



- Production cross-section ~ 7pb at the Tevatron → the LHC is a top quark factory
- The top decays to Wb → decay channels defined by the W decay



Top Pair Decay Channels



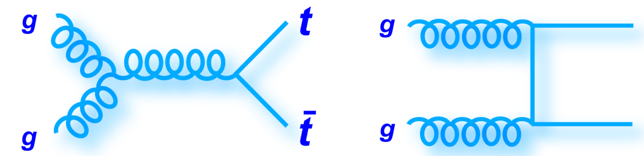
All-hadronic channel (BR ~ 45%)
High branching fraction
Very challenging: high multijet background, difficult to trigger

Lepton+jets channel (BR ~ 30% + τ+jets ~ 15%)
Cleaner signature, better resolution
Presence of missing transverse energy (neutrino from W decay)
τ channel tricky

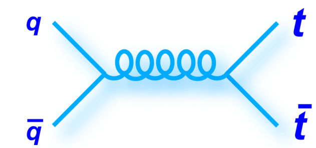
Dilepton channel (BR ~ 5% + τ+l ~ 5%)
Very clean
High missing transverse energy (2 neutrinos)
τ channels more difficult

Production and decay at the LHC

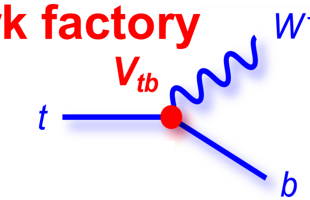
- Produced in pairs via gluon fusion or quark annihilation
- Gluon fusion dominates at the LHC (7 TeV) while quark annihilation dominates at the Tevatron (~ 2 TeV)



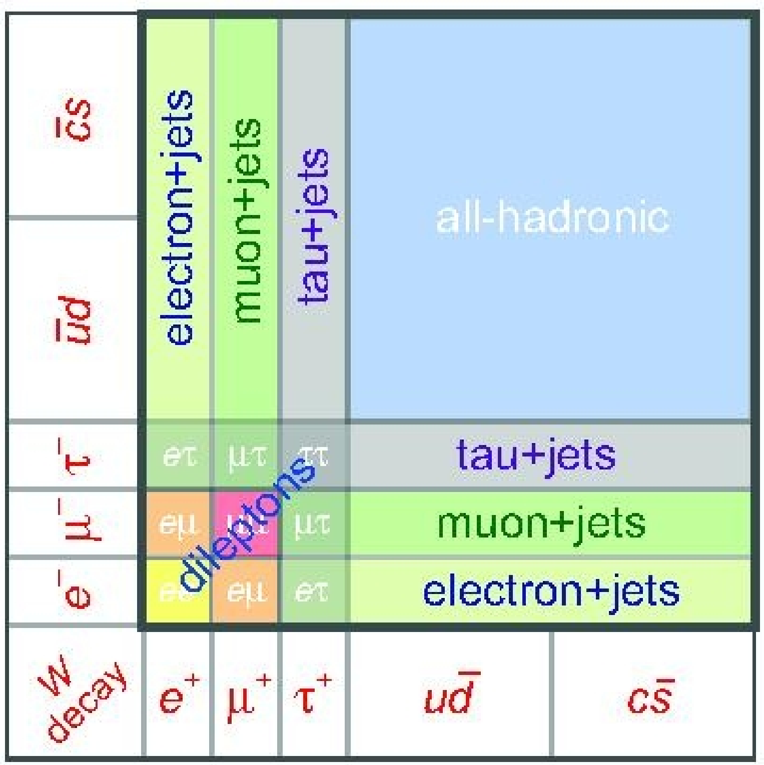
Kidonakis [PRD 82 (2010) 114030]: σ_{tt}(7 TeV) = 163 + 11 – 10 pb



- Production cross-section ~ 7pb at the Tevatron → the LHC is a top quark factory
- The top decays to Wb → decay channels defined by the W decay



Top Pair Decay Channels



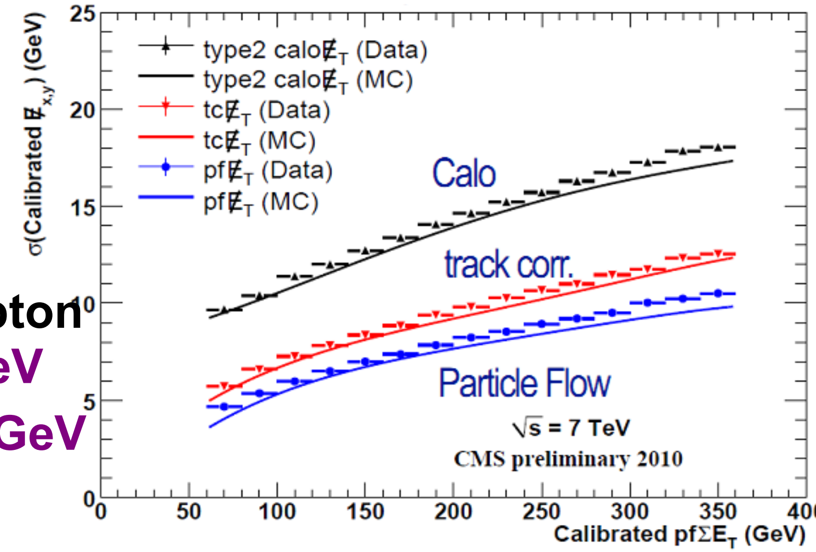
Today: present a selection from all the exploited channels from 2010 (36 pb^{-1}) and 2011 (1 fb^{-1})

Strategy for cross-section extraction

- Need to build a model to compare with data
- Signal shape is given by simulated samples (MADGRAPH)
- Background shapes
    - are given by simulated samples
    - can be extracted from data: multijet background, but also Z/W + jets
- Data-driven techniques rely on isolating "side-band" data enriched in background but depleted in signal; contamination from other backgrounds can be removed by using simulation
- Since we want to perform a cross-section measurement, perform a binned likelihood fit of the different background / signal shapes to data or do a counting experiment

The lepton + jets channel (1)

- **Final state includes one leptonically-decaying W boson: presence of missing transverse energy**
  need good MET resolution, use particle flow

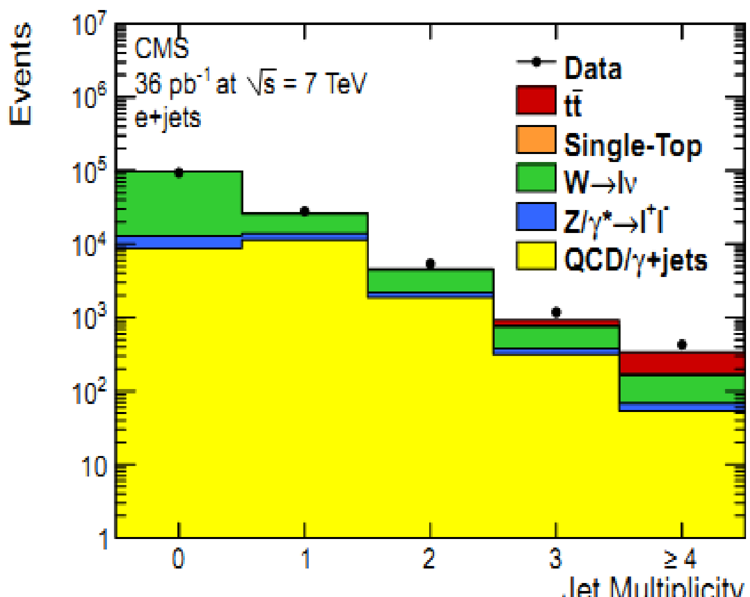


- **Event selection**
  - **Single-lepton triggers, exactly one isolated lepton**

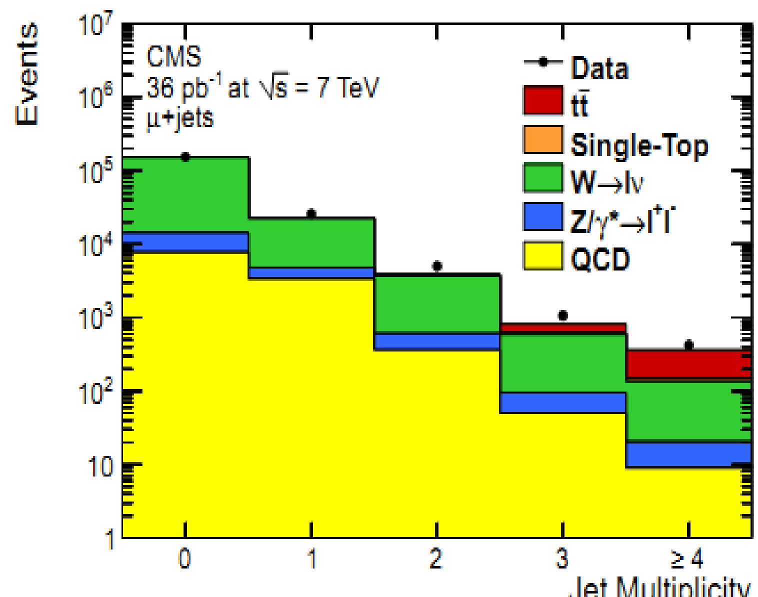
electron: η < 2.5, relative isolation < 0.1, p_T > 30 GeV

muon: η < 2.1, ΔR(m, jet)<0.3, rel. iso.<0.05, p_T > 20 GeV

  - **Veto di-lepton events**
  - **High-p_T central jets (η < 2.4, p_T > 30 GeV)**



electron channel



muon channel

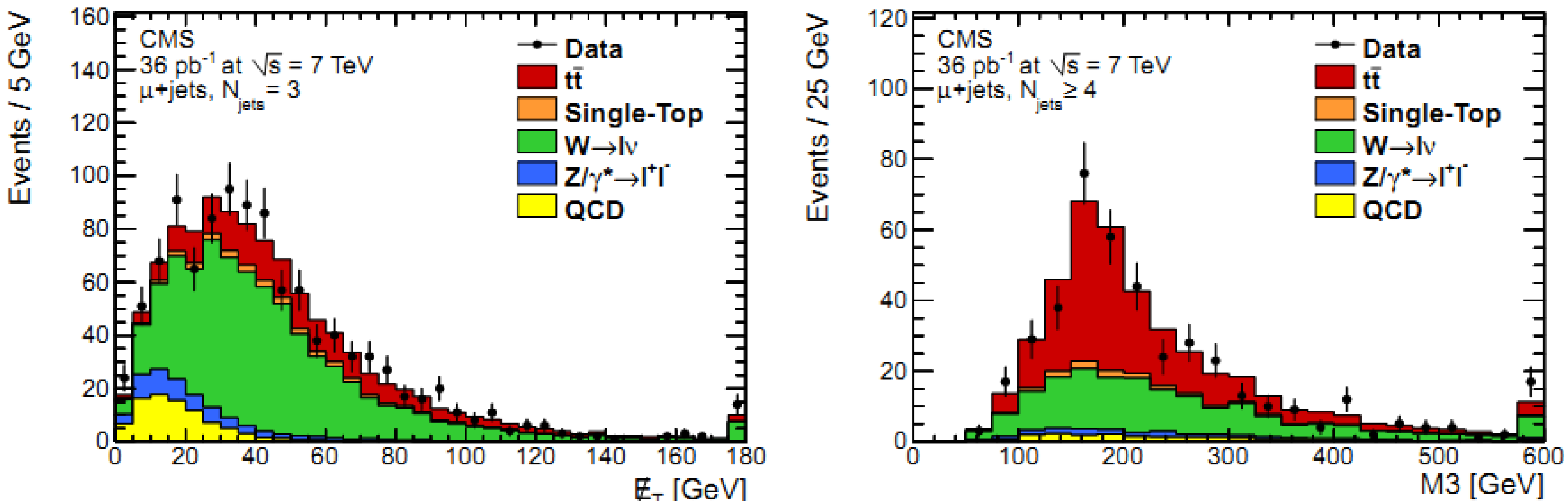
36 pb^{-1}

- **Main backgrounds: W+jets; QCD, estimated from data with loosely isolated leptons**

The lepton + jets channel (2)

- **Strategy: simultaneous kinematic fit of signal and background templates to data in the 3-jet and >= 4-jet sample**
  - **Simultaneous fit allows better shape constraint**
  - **Fit on MET in the 3-jet sample, M3 (mass of hadronically decaying top) in the >=4-jet sample – most discriminating variables**

3 jets - MET **36 pb^{-1}** **>= 4 jets - M3**



muon channel

Distributions after fit

- **Combined e /μ result: perform simultaneous MET and M3 fits on both the electron and the muon channel; one parameter per channel for QCD since sources are different (semi-leptonic decays in e/μ, also pion-rich jets faking electrons in e)**

$$\sigma_{tt} = 173 + 36 - 32 \text{ (stat +syst)} \pm 7 \text{ (lumi) pb}$$

Total systematic uncertainty ~ 20%

Dominated by jet energy scale ~ 18%, factorisation scale ~ 7%

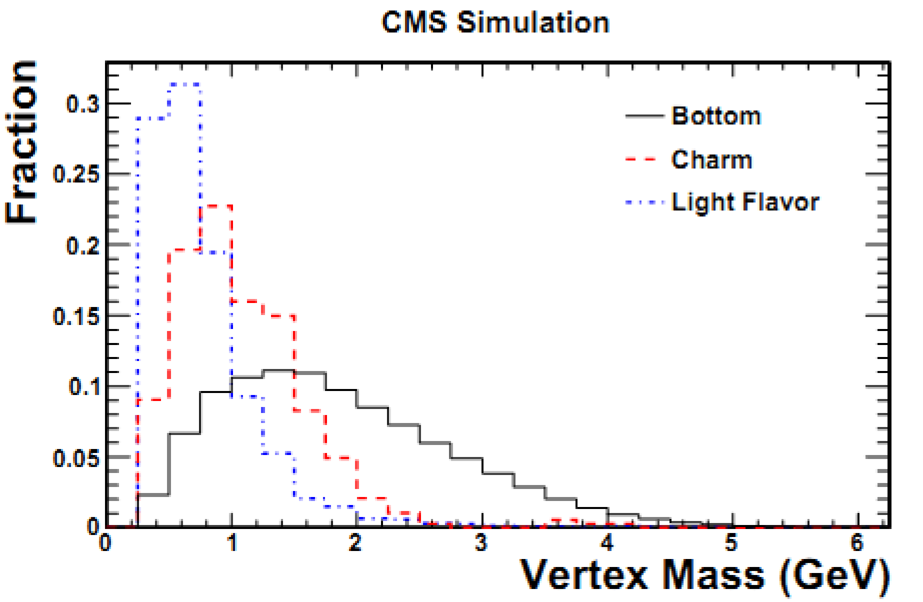
Result with 36 pb^{-1}, analysis with 1 fb^{-1} being reviewed

The lepton + jets channel (3)

- Use flavour information → b-tagging
  - B hadrons have long lifetime → presence of secondary vertex
  - Build discriminant from displaced charged tracks

- Choose working point with 55% b-tag efficiency and 1.5% mis-tag rate

- Use secondary vertex mass to discriminate against light-flavour background
  $M(SVX) = M(\Sigma \text{ 4-vectors}_{SVX})$ with $M(\text{particle}) = M_{\pi}$



- Perform simultaneous maximum likelihood fit on M(SVX) in bins of N_{jets} and $N_{b\text{-}tag}$

- Main systematics: jet energy scale, Q^2-scale, b-tagging efficiency → correlated parameters which can bring large variations to the yields in each bin

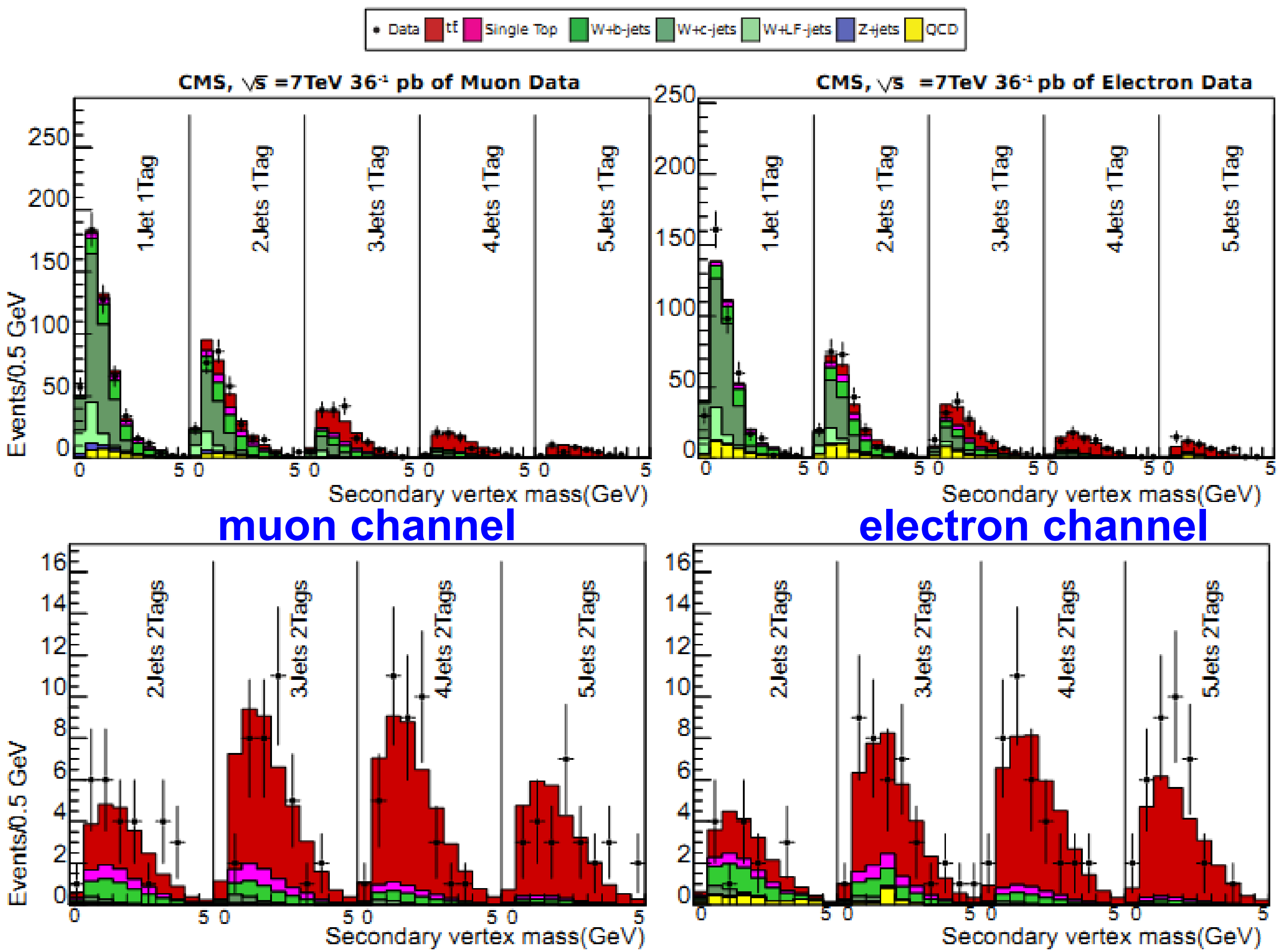
- Therefore include them in likelihood fit as nuisance parameters, in effect simultaneously measuring the signal and background contributions

The lepton + jets channel (4)

- **Final distributions**

1 tag

2 tags



36 pb^{-1}

signal in red

secondary vertex mass

σ_{tt} = 150 ± 9 (stat.) ± 17 (syst.) ± 6 (lumi) pb

Total systematic uncertainty ~ 11%

Dominated by jet energy scale ~ 18%, factorisation scale ~ 7%

The di-lepton channel (1)

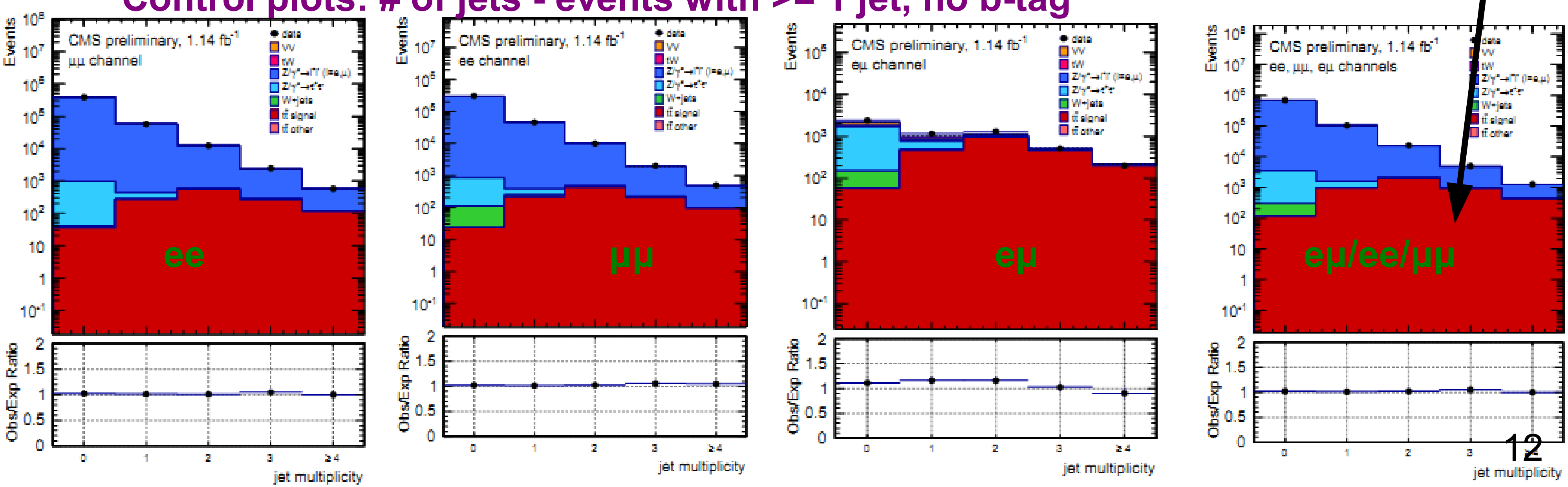
- Final state: 2 leptons, 2 b jets and missing transverse energy → **need good jet energy scale, b-tagging and lepton ID, but clean signature, low QCD background**

- Event selection
  - Di-lepton triggering
  - At least one pair of oppositely charged, isolated leptons with $p_T > 20$ GeV, $\eta < 2.4$ (2.5) for muons (electrons), one or two jets with $p_T > 30$ GeV, $\eta < 2.5$
  - MET > 30 GeV (2 jets) or 50 GeV (1 jet) in the ee and μμ channels only

- Main background: **Z/γ + jets**, estimated from data by counting number of Z + jets events having $76 < M_{ll} < 106$ GeV and comparing with this cut's efficiency in simulation

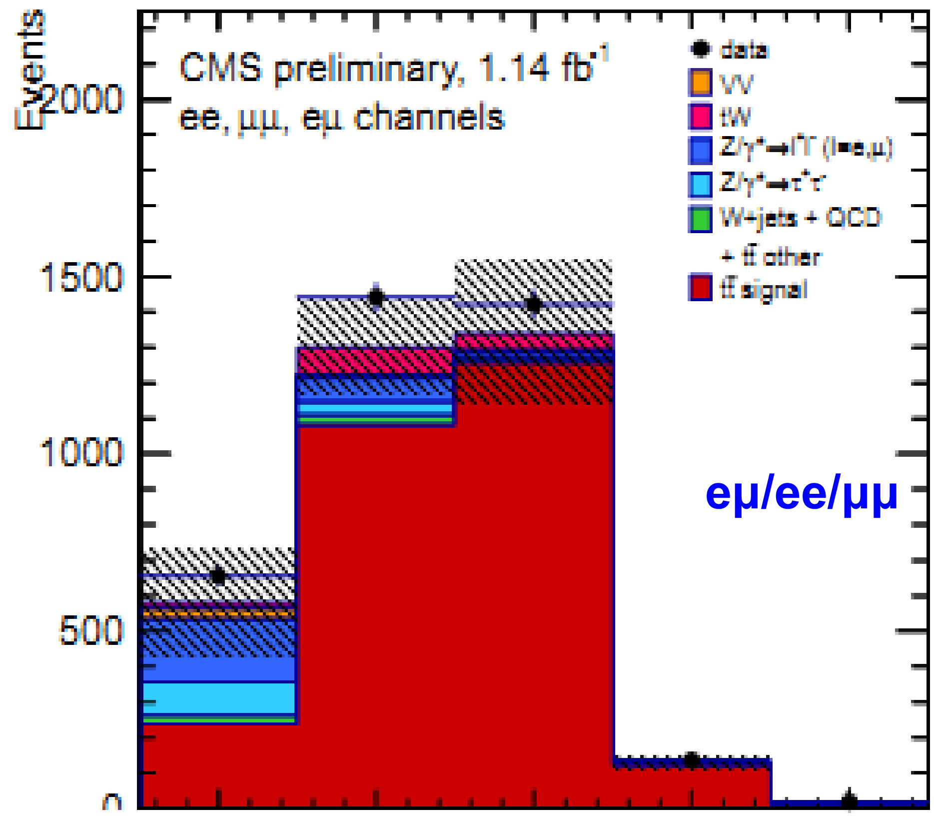
1.14 fb^{-1} Signal

Control plots: # of jets - events with >= 1 jet, no b-tag



The di-lepton channel (2)

- Also use b-tagging to increase signal purity: working point chosen has ~80% efficiency and ~ 10% mis-tag rate
- Cross-section extraction: counting experiment



jets 0 1 2 3 >=4

1.14 fb^{-1}

b-tagging applied

$$\sigma_{tt} = 169.9 \pm 3.9\ (\text{stat}) \pm 16.3\ (\text{syst}) \pm 7.6\ (\text{lumi})\ \text{pb}$$

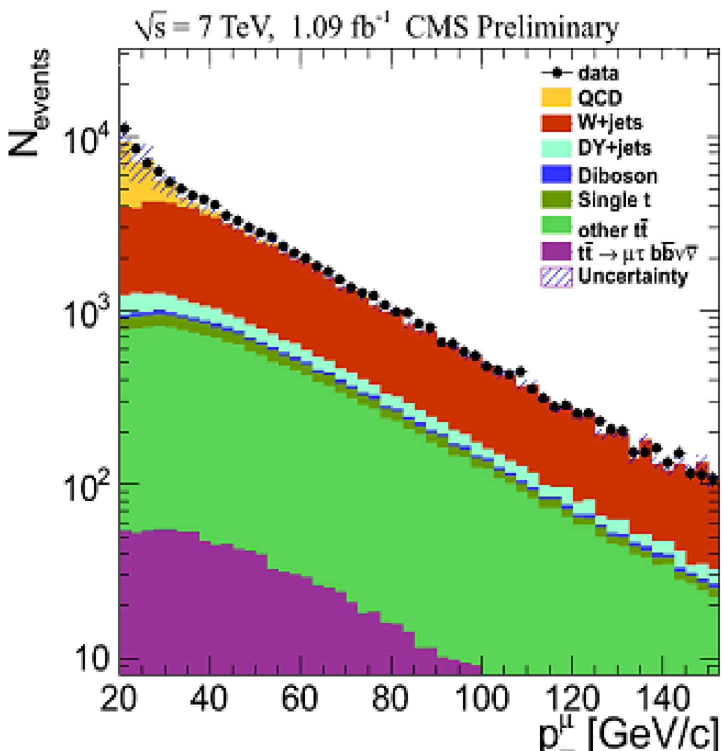
Total systematic uncertainty ~ 10%

Dominated by lepton selection ~ 4%, b-tagging ~5%

The μ/τ channel (1)

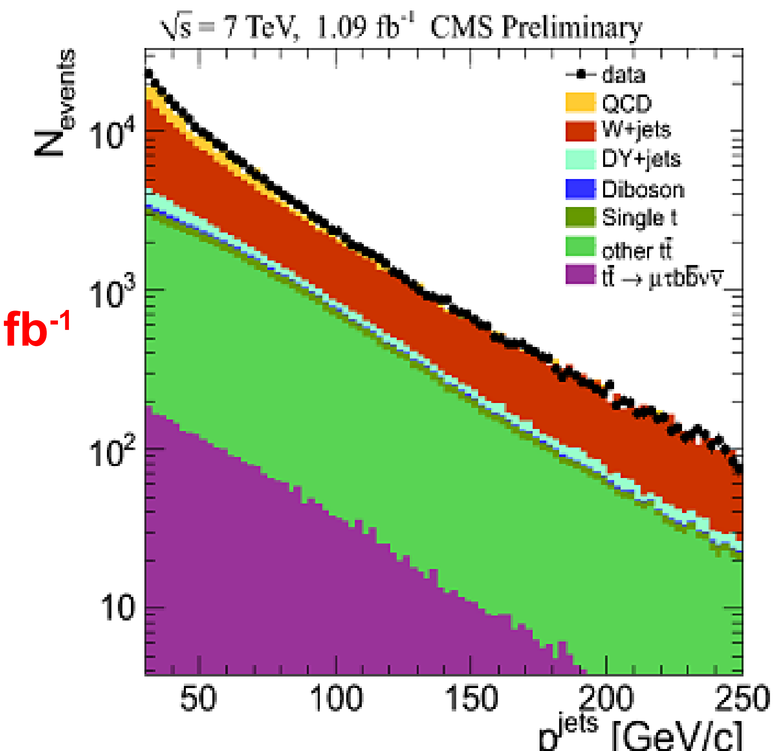
- Final state: 2 leptons of which one is a hadronically decaying τ, 2 b jets and missing transverse energy → tricky due to the fact that hadronic τ resemble jets

- Event selection
  - Single-muon trigger
  - One isolated muon with p_T > 20 GeV; one tau with p_T > 20 GeV
  - At least two jets with p_T > 20 GeV, one of which is b-tagged, and MET > 40 GeV



muon p_T

1.09 fb^{-1}

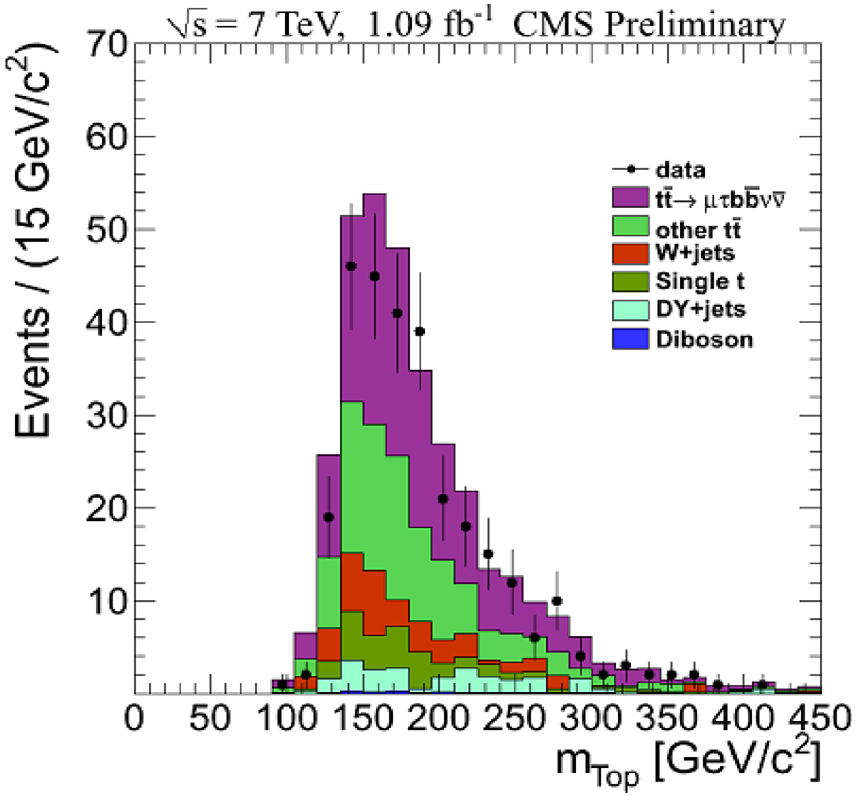


p_T of jets

>= 3 jets
no b-tag

The μ/τ channel (2)

- Main background: jets faking τ; calculate probability $w(\eta, p_T)$ that a jet fakes a τ on data selected with high-p_T jet trigger; don't consider the trigger-matched jet to avoid trigger bias, subtract real τ contribution with simulation



reconstructed top quark mass via kinematic fit after b-tagging and lepton opposite sign requirement

σ_{tt} = 148.7 ± 23.6 (stat) ± 26 (syst) ± 8.9 (lumi) pb

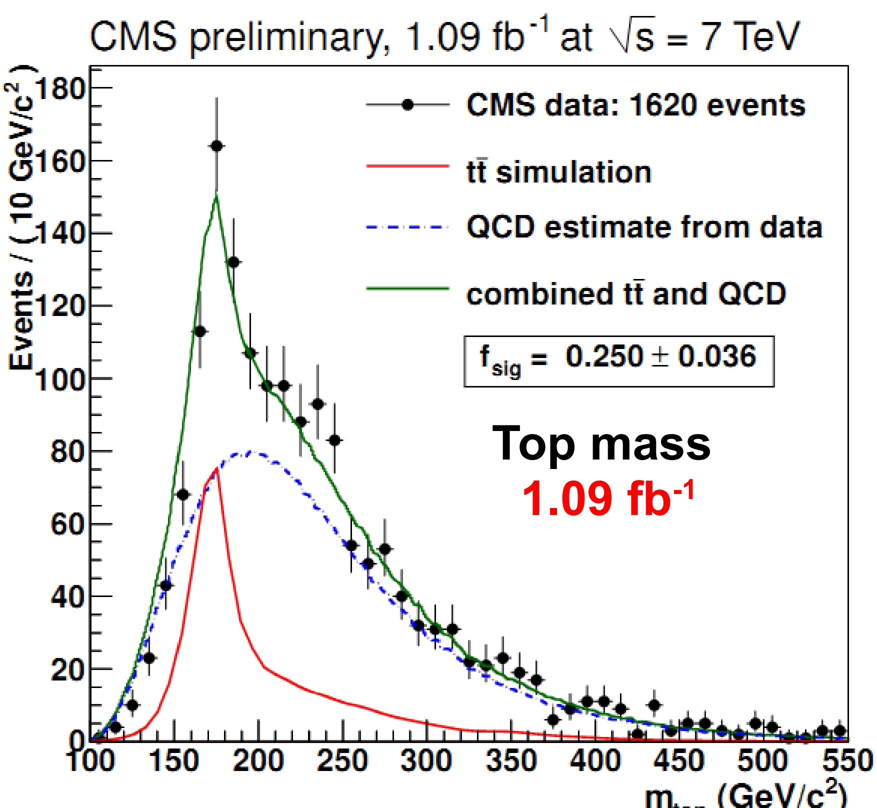
Total systematic uncertainty ~ 17%

Dominated by fake tau background estimation ~ 13%, τ-ID ~ 7%, b-tagging ~ 5%

The all-hadronic decay channel (1)

- **Very challenging**: the multijet background (from QCD interactions) dominates

- **Step 1: event selection**
  - 4 jets with p_T > 60 GeV, 1 jet with p_T > 50 GeV, 1 jet with p_T > 40 GeV
  - Keep events with more jets if p_T > 30 GeV
  - Require at least **2 b-tagged jets** with N_{tracks}(secondary vertex) >= 3 and decay length significance d_B > 2 → **38% efficiency with very low mis-tag rate (0.12%)**

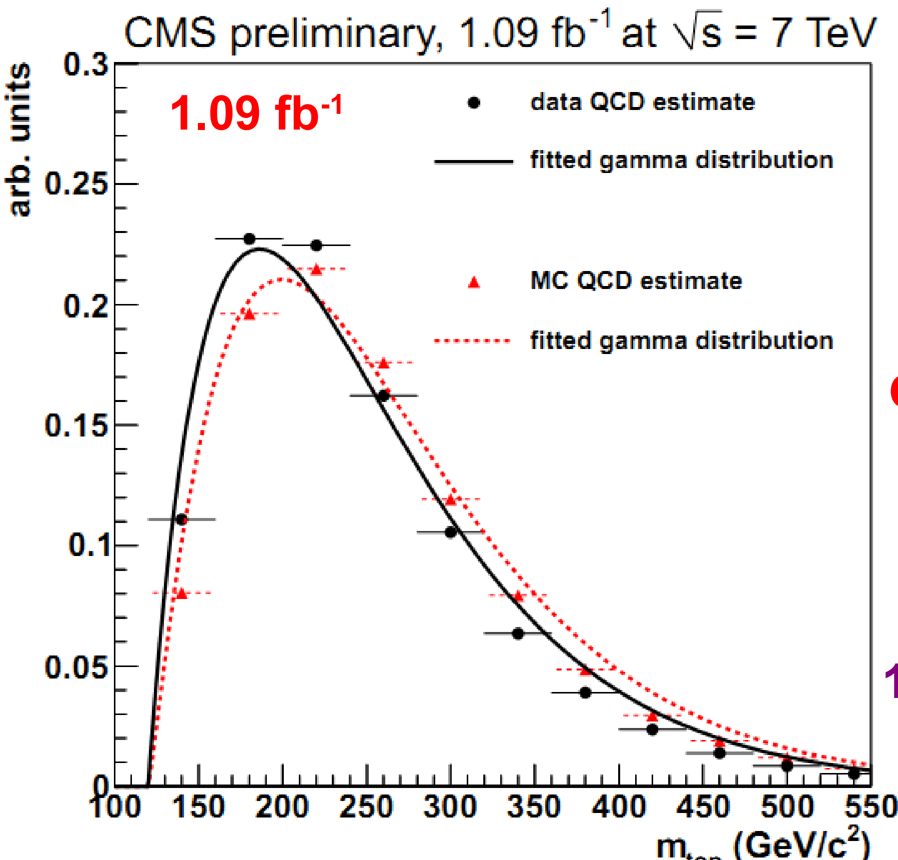


- **Step 2: kinematic fit**
  - Reconstruct two W bosons (m = 80.4 GeV) from the non-tagged jets
  - Two top quarks from the W and the tagged jets; assume $m_{top} = m_{antitop}$
  - Fit the combinations, keep events with $P(\chi^2)$ > 1%

	Signal fraction
Step 1 w/o b-tag	2%
Step 1 + b-tag	17%
Step 2	32%

The all-hadronic decay channel (2)

- **Model QCD background from data**
  - Select events with >= 6 no b-tagged jets (signal fraction < 1%)
  - Reweigh to match kinematics of b-tagged sample (p_T and η of jets)
  - Perform kinematic fit on 0-tag sample, assuming all jets to be b-jets, and adjust weight
  - The reweighed data events are used to estimate the shape of the QCD background
  - Cross-check with simulated events



- **Cross-section extraction via maximum likelihood fit of signal and background shapes to data**

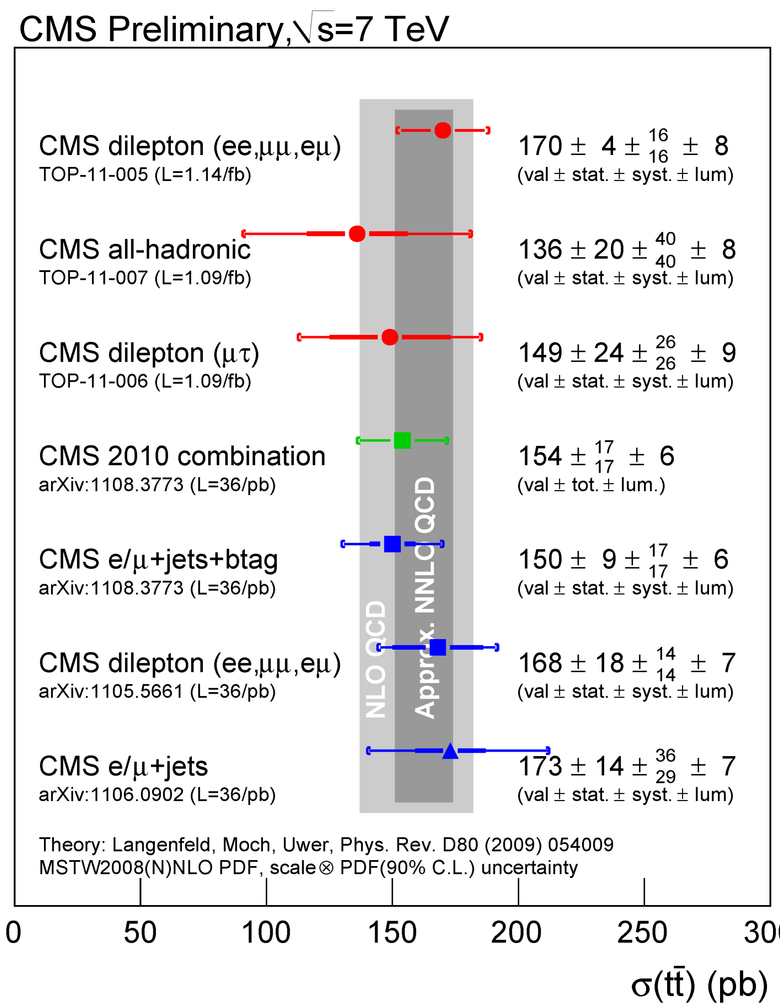
$$\sigma_{tt} = 136 \pm 20\ (\text{stat}) \pm 40\ (\text{syst}) \pm 8\ (\text{lumi})\ \text{pb}$$

Total systematic uncertainty ~ 29 %

Dominated by b-tagging (16%) and jet energy scale (14%)

12% assigned to QCD modelling (from ±5% variation in shape of gamma distribution fit to predicted points)

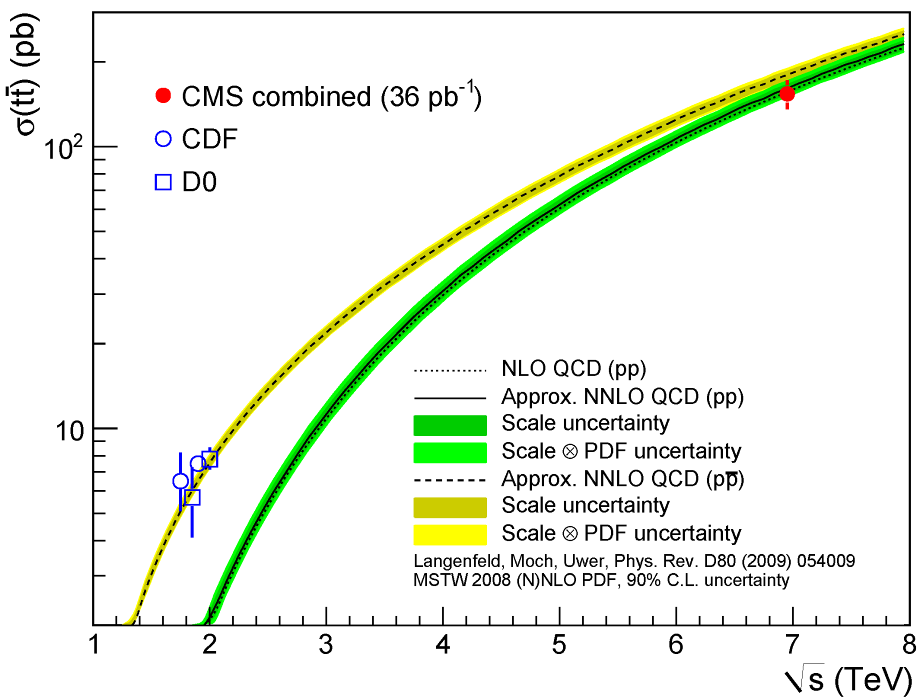
Combined results and comparison with theory



- Combine measurements using the best linear unbiased estimator technique
  - Assume uncorrelated systematics within each channel

σ_{tt} = 154 ± 17 ± 6 (lumi) pb

Total uncertainty ~ 12%



σ_{tt}^{NLO} (MCFM) = 158 + 23 – 24 pb

$\sigma_{tt}^{appNNLO}$ (HATHOR) = 164 + 11 – 13 pb

$\sigma_{tt}^{appNNLO}$ (Kidonakis) = 163 + 11 – 10 pb

CMS already starting to constrain NLO calculations

Conclusion and outlook

- The LHC and the CMS detector have been performing very well since physics data taking started
  - 36 pb^{-1} on tape in 2010, already > 2 fb^{-1} in 2011
  - Expect 3-5 fb^{-1} by end 2011

- Top cross-section measurements have been done in various channels
  - Good agreement with theory
  - Most analyses already out with > 1 fb^{-1}
  - Other analyses released with 36 pb^{-1} but pushing to update

- Top pair topology is an excellent benchmark for testing many analysis ingredients: reconstruction algorithms (particle-flow), lepton ID, b-tagging, data-driven background prediction techniques,.. → excellent understanding of detector

- CMS measurements now systematics-limited
  - CMS working hard to decrease dominant systematics: jet energy scale / MET, b-tagging, tau-ID,...

- CMS data already starting to constrain NLO calculations
- Differential measurements to come along with greater precision

Back-up

The BLUE method

- See L. Lyons, D. Gibaut; NIMA A270 1998 110-117 for details

- The Best Linear Unbiased Estimate method is a statistical technique used to combine several measurements of the same quantities obtained from eg. different channels

- Let's say we want to combine n measurements...

- Generate pseudo-experiments, using simulation, for each of those measurements, then perform measurement (e.g. likelihood fit) on each of them → n measures

- Check each measures for biases → check e.g. pull distributions for the n fit results

- If OK, we need to weight out n different results – α_p will be the total contribution or result p. To find out the weights, calculate a variance (shown here for n = 3):

$$\sigma^2_{combined} = \begin{pmatrix} \alpha_1 & \alpha_2 & \alpha_3 \end{pmatrix} \begin{pmatrix} \sigma_1^2 & \sigma_{12} & \sigma_{13} \\ \sigma_{12} & \sigma_2^2 & \sigma_{23} \\ \sigma_{13} & \sigma_{23} & \sigma_3^2 \end{pmatrix} \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{pmatrix}$$

- Compute the α factors by minimising the variance using the contrain $\Sigma\alpha_i = 1$

- For more details see also CDF and DØ papers on eg. top mass measurements or single-top evidence and discovery, where this technique was used