Determination of the Top-Quark Mass from the m_{lb} Distribution

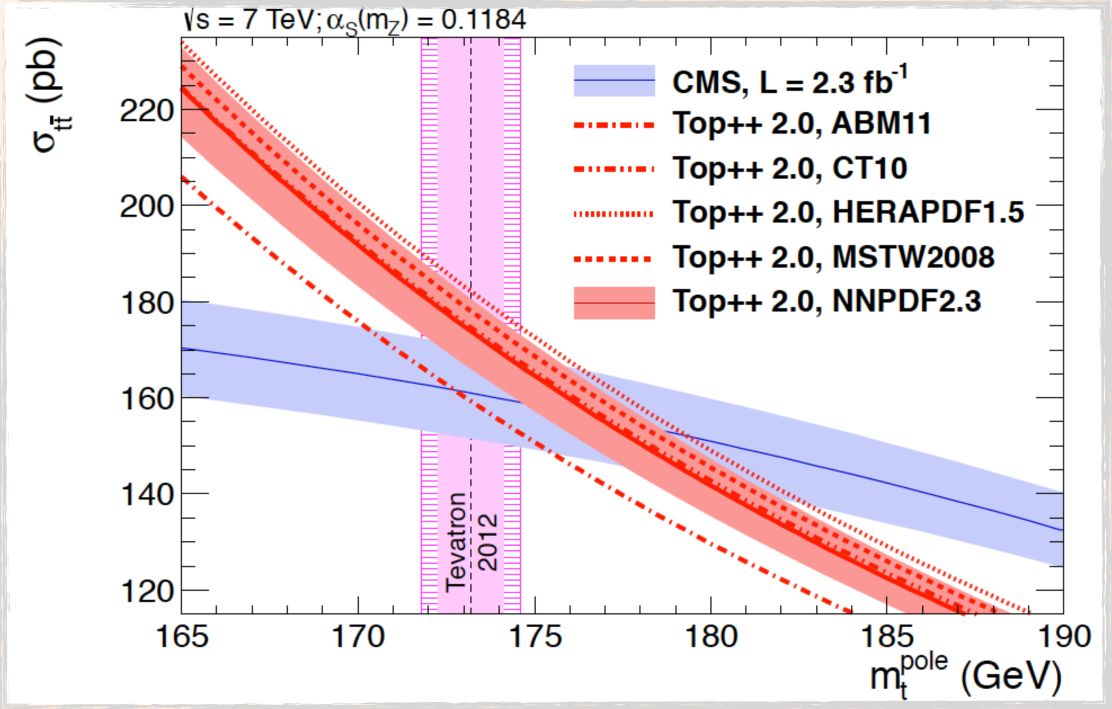
in Dileptonic Top-Quark Pair Events at $\sqrt{s}$= 8 TeV

Jan Kieseler

05.03.2015

- Motivation -
- Event Selection & m_{lb} Definition-
- Extraction Method & Results-

- Inclusive tt cross sections used to extract $m_{t,pole}$ at NNLO
- Taking advantage of different slope of $\sigma_{pred}(m_t)$ and $\sigma_{meas}(m_t)$



(PLB 728 (2014) 496)

Next step:

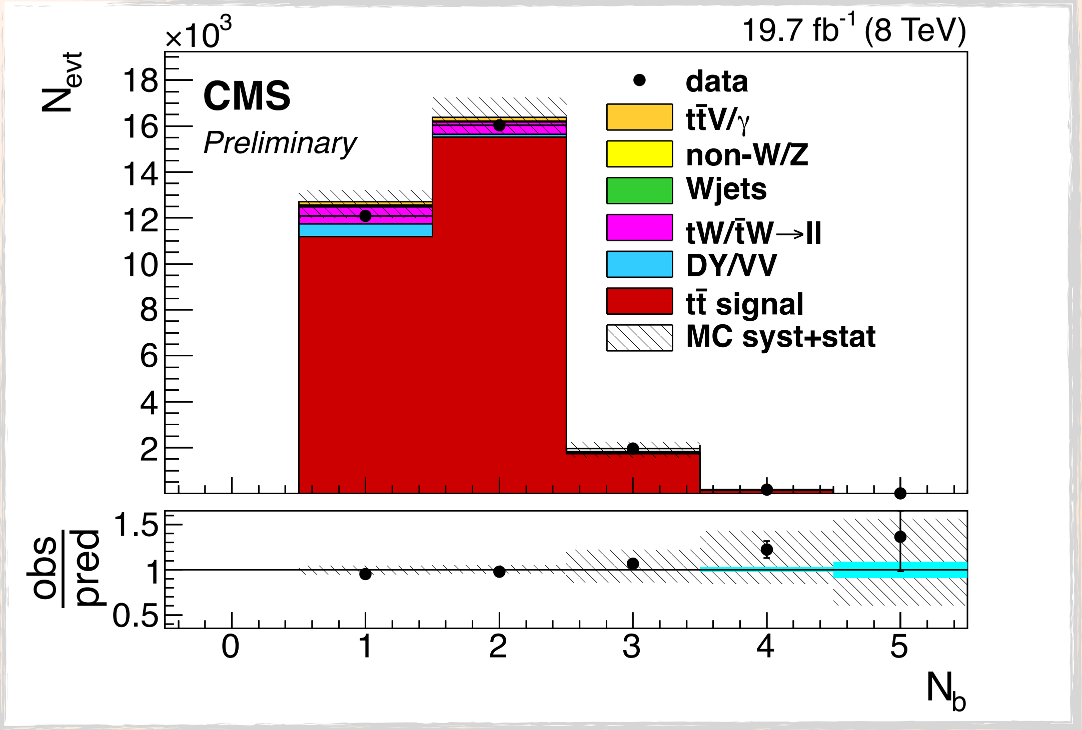
- Extract m_t from differential cross sections
- Possibly more sensitive to m_t, more clear theoretical interpretation
- Step to simultaneous m_t, pdf, α_s extraction

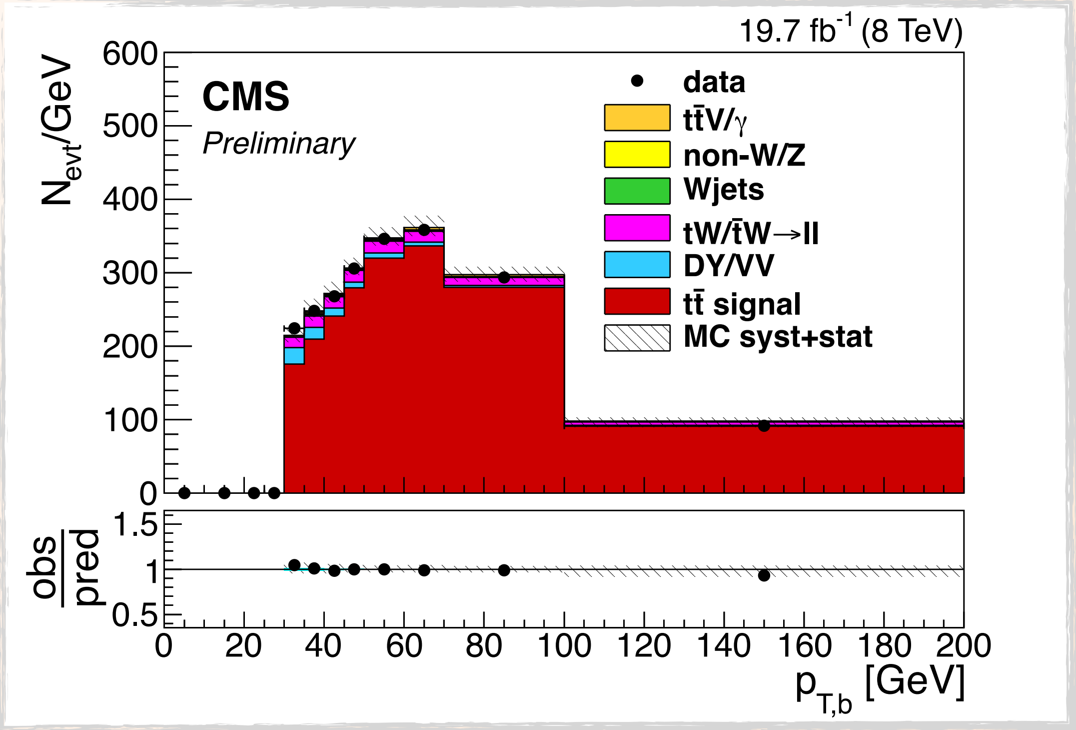
➡ Observable sensitive to m_t, measurement + prediction

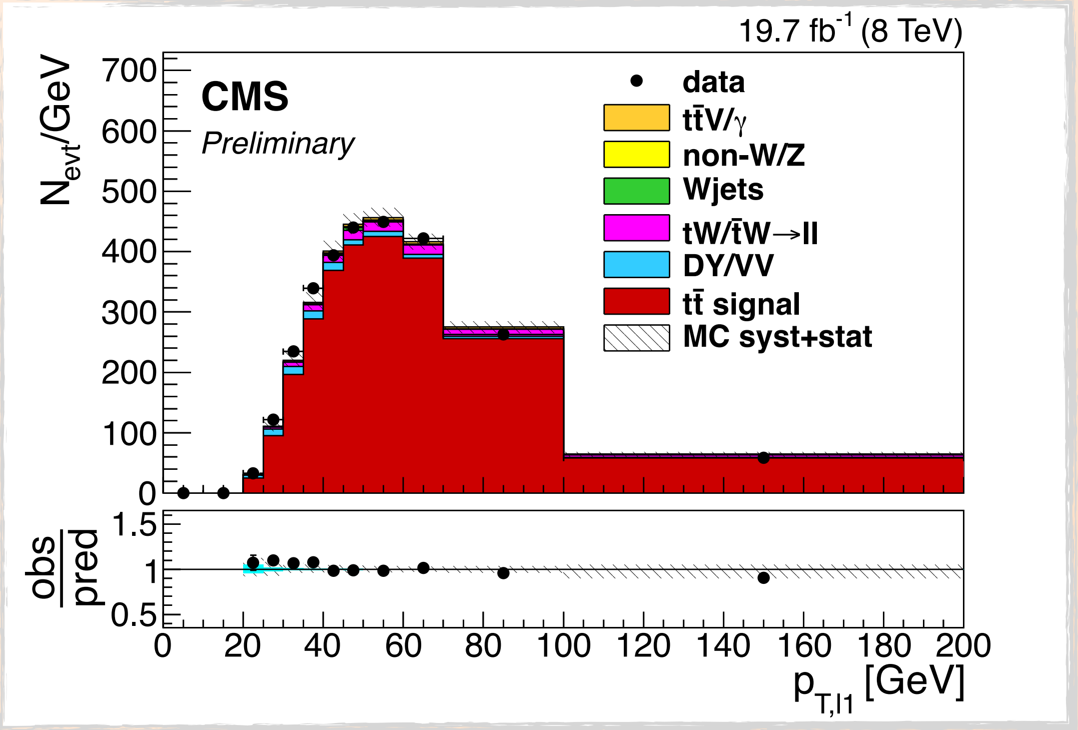
Reference: CMS PAS-TOP-14-014

Event Selection

- Full 2012, 8 TeV dataset (19.7 fb^{-1})
- Signal: MadGraph+Pythia (7 mass-points)
- Selection
  - ≧2 opposite charged isolated leptons **eμ**
  - Remove low m_{ll} < 20 GeV
  - ≧2 high p_T jets
  - ≧1 b-tagged jet

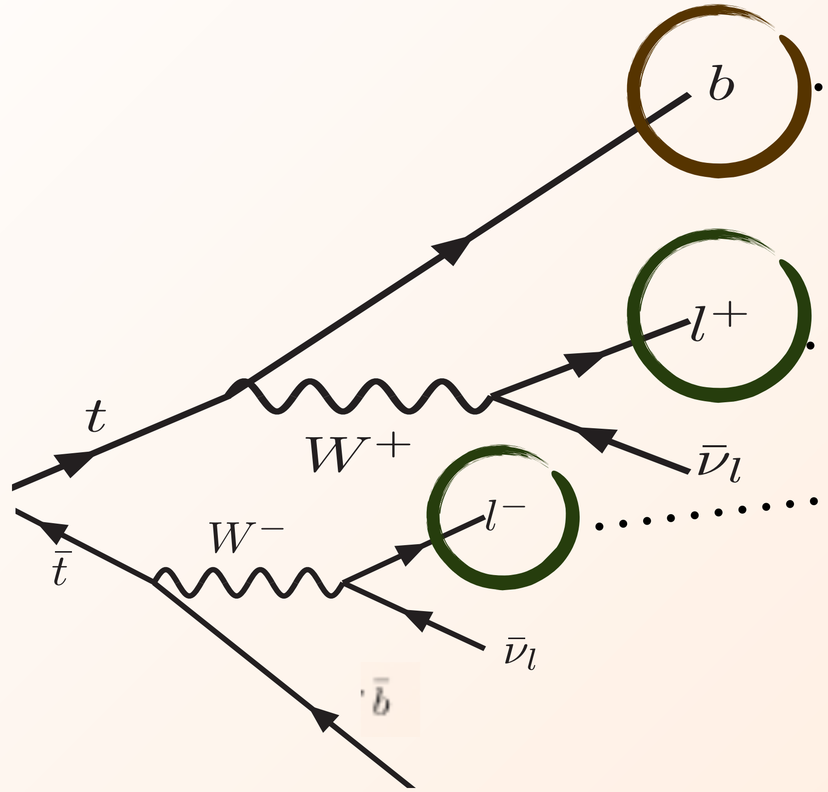






Observable $m_{lb,min}$

Definition on **generator** and **reconstruction** level



- leading b-quark from top / leading b-jet
- $|\eta| < 2.4$, $p_T > 30$ GeV

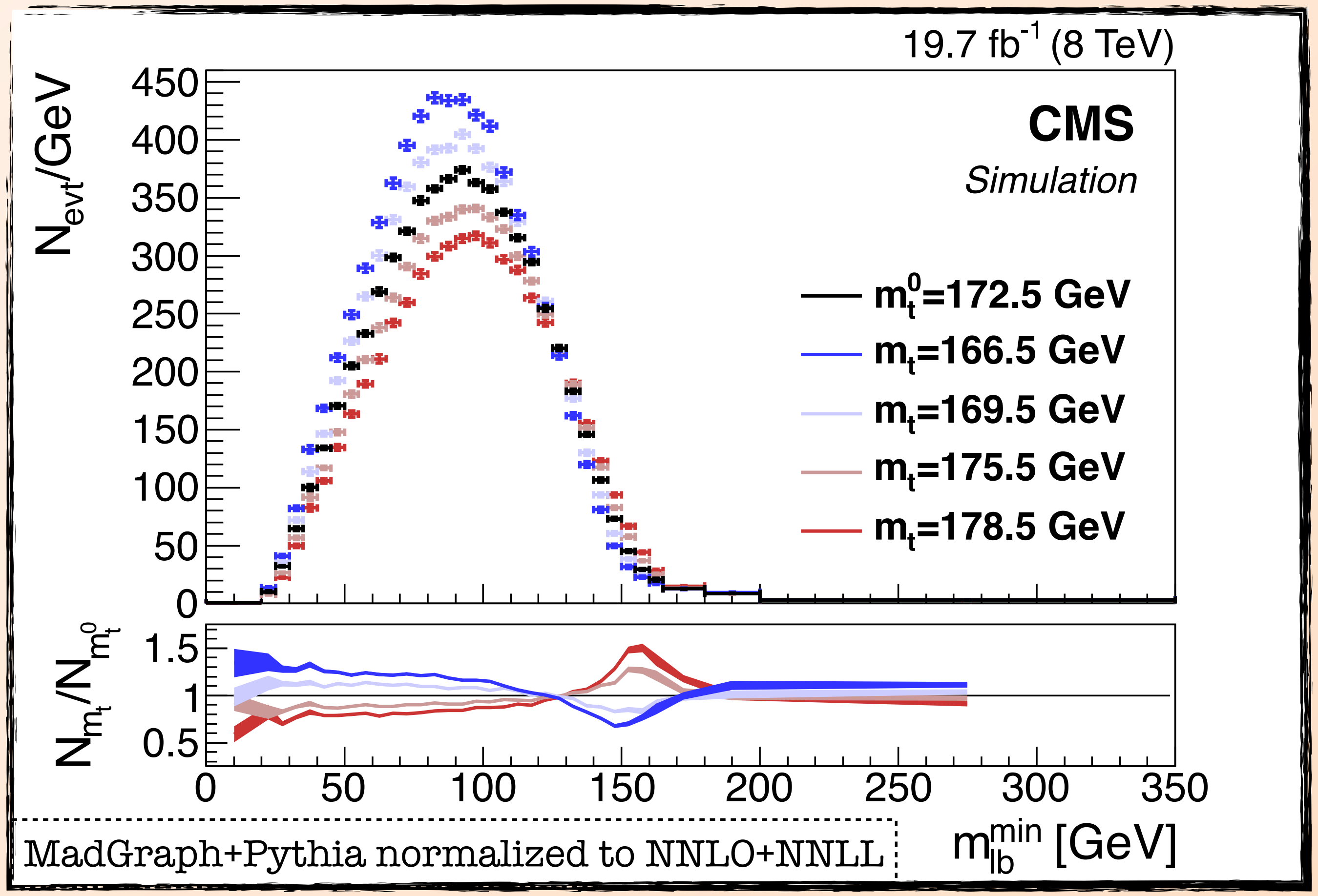
- leptons from top (e/μ/tau(→e,μ)) / selected oppositely charged lepton pair
- $|\eta| < 2.4$, $p_T > 20$ GeV

$m_{lb,min}$: smaller m_{lb} of 2 combinations

➡ Leading b-jet: smaller JES uncertainties

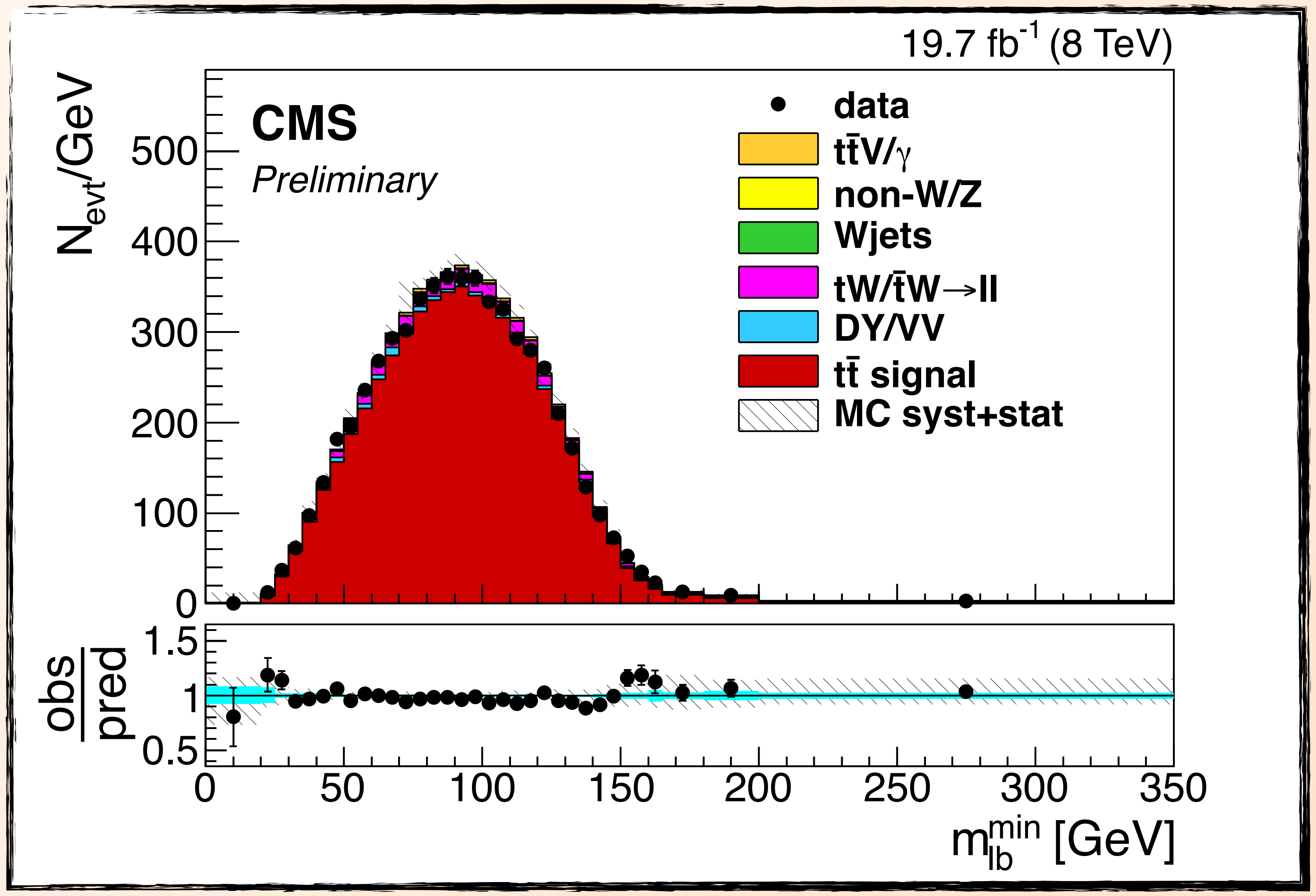
➡ Lepton kinematics known to high precision

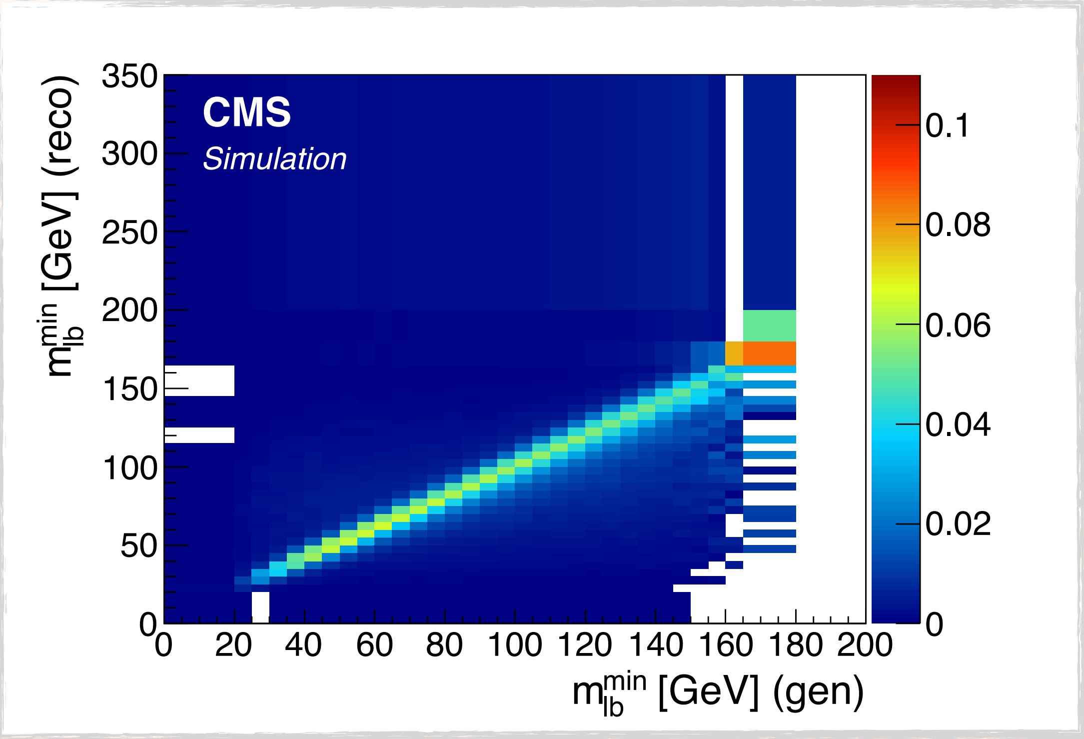
➡ Visible phase space (vis. PS)



Observable $m_{lb,min}$ (2)

- Good data/MC agreement
- Low background
- Only small contribution of migrations from outside visible PS





- τ decays, efficiencies included in response matrix

$$\vec{x}_{reco} = \mathcal{L} \cdot M^{resp} \; \vec{x}_{pred}$$

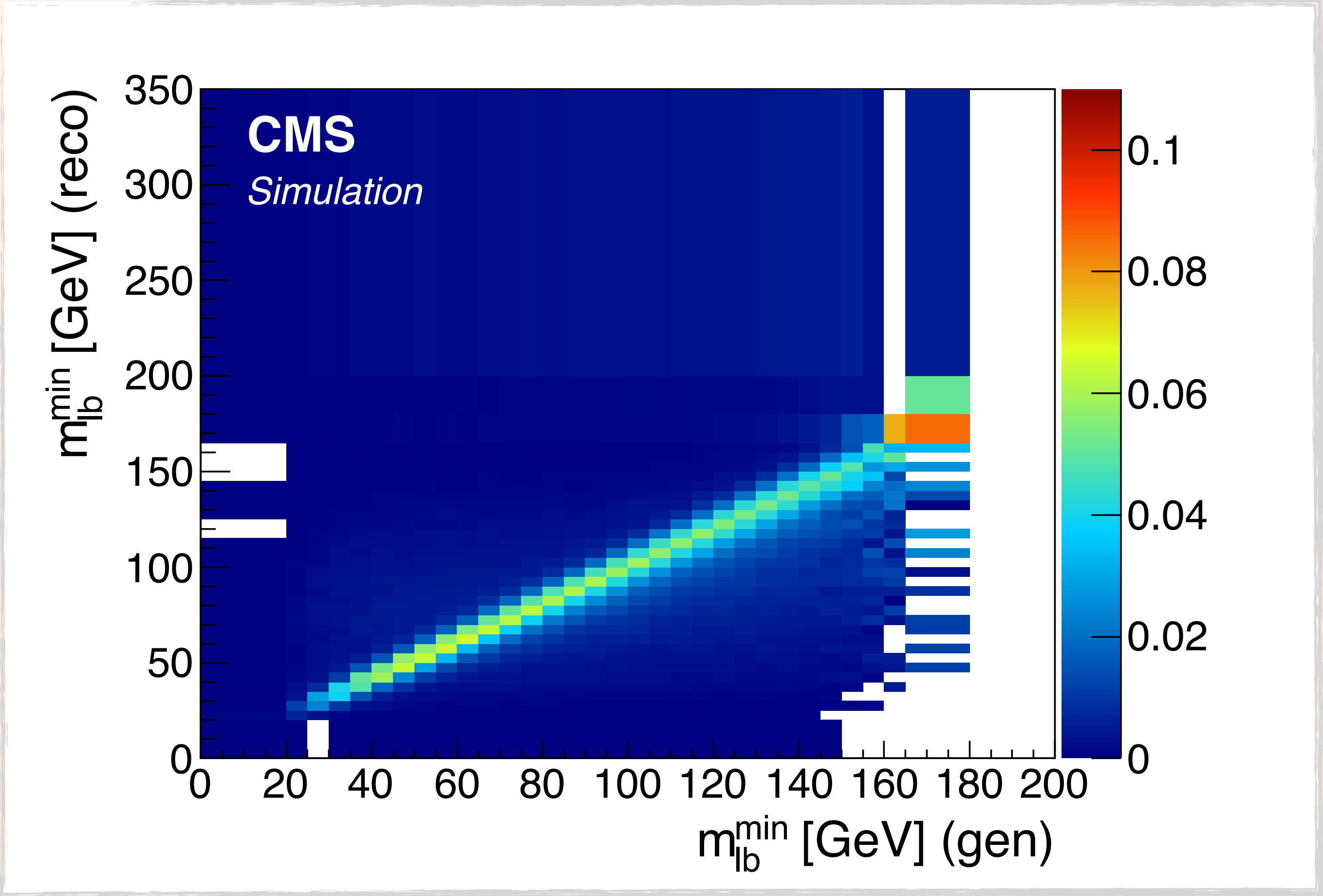
- Strong correlation between reconstructed and generated $m_{lb,min}$

Folding

- Compare reconstruction level distributions (MadGraph + detector sim.)
- Compare other prediction to data on reconstruction level?
- Fold theory:

$$M_{ij}^{resp} = N_{ij} / \sum_{j=0}^{n} N_{ij}$$

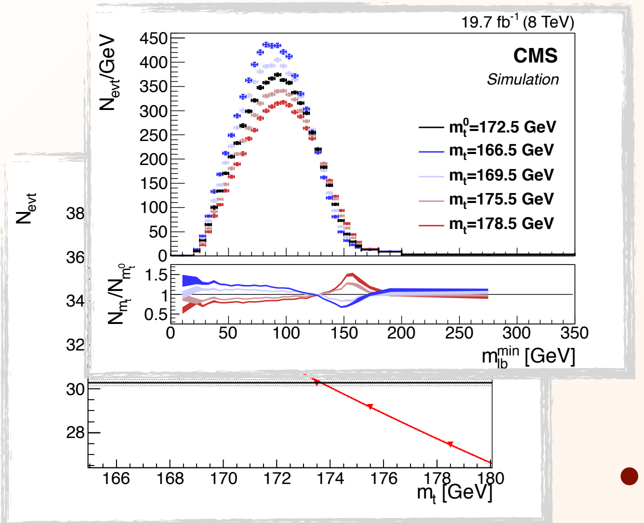
$$\vec{x}_{reco} = \mathcal{L} \cdot M^{resp} \quad \vec{x}_{pred}$$



- Separate matrix for each systematic variation and m_t
  - Detector modeling uncertainties
  - Signal modeling uncertainties (reconstruction efficiencies)
- No additional unfolding/regularization uncertainties
- No stat. correlations between bins of reconstructed distribution
- Here: Fold prediction within visible phase space

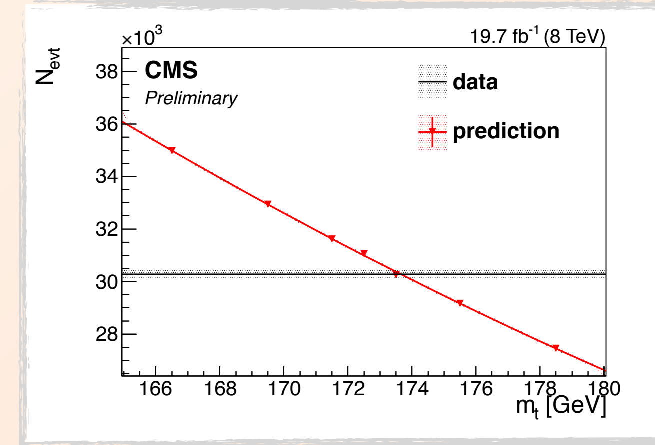
➡ Reconstruction level prediction can be directly compared to measured points

Method Overview



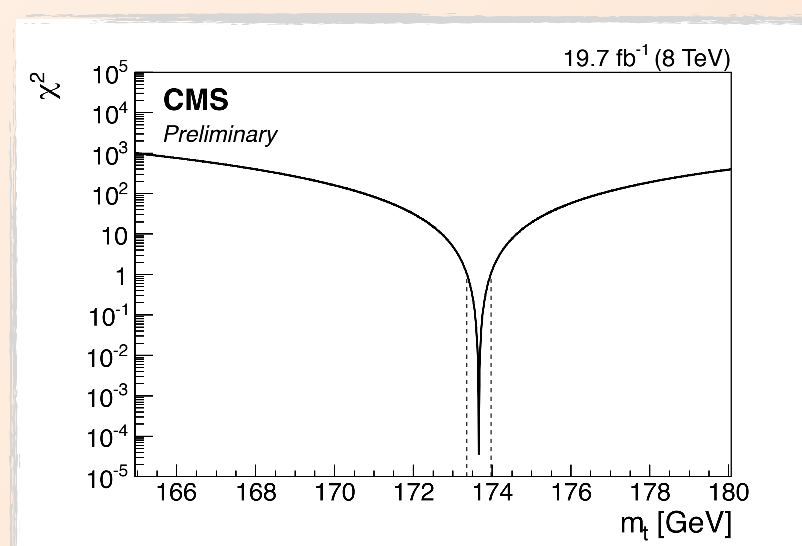
prediction:
either
or
+ data

$$N_{\text{pred},i}(m_t)$$



- For each m_{lb} bin:
  - Derive cont. m_t dependence of pred (fit)
  - Construct χ^2 (m_t) per bin
- Combine bins
  - No stat correlations → sum χ^2 over bins
- Determine m_t
  - Repeat for each syst. variation

Likelihood

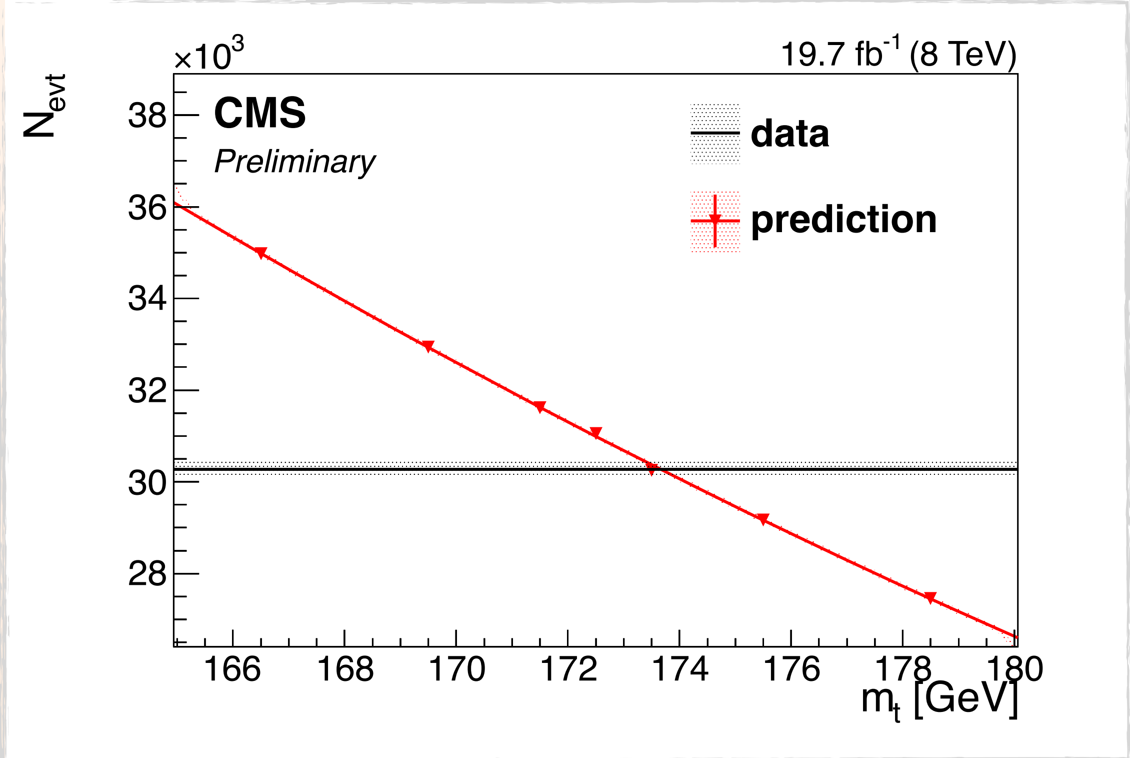


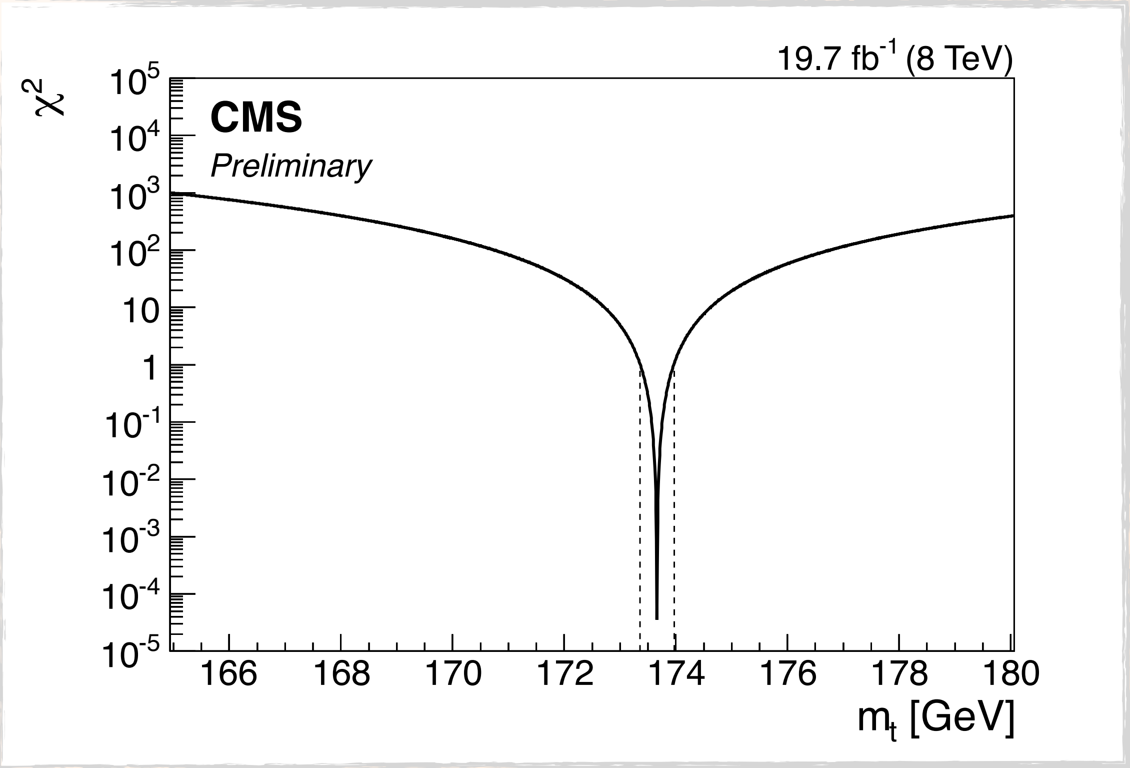
m_t

rate
shape & rate
shape

Modeling Likelihood (m_t)

- Fit mt-dependence of prediction, N_{pred}, with 2nd order polynomial
  - Includes backgrounds
  - Use statistical uncertainty (mostly background) from central point → σ_{pred}
  - Parametrization uncertainties not included here (→ see systematics)
- Data yield N_{data}: constant with σ_{data}

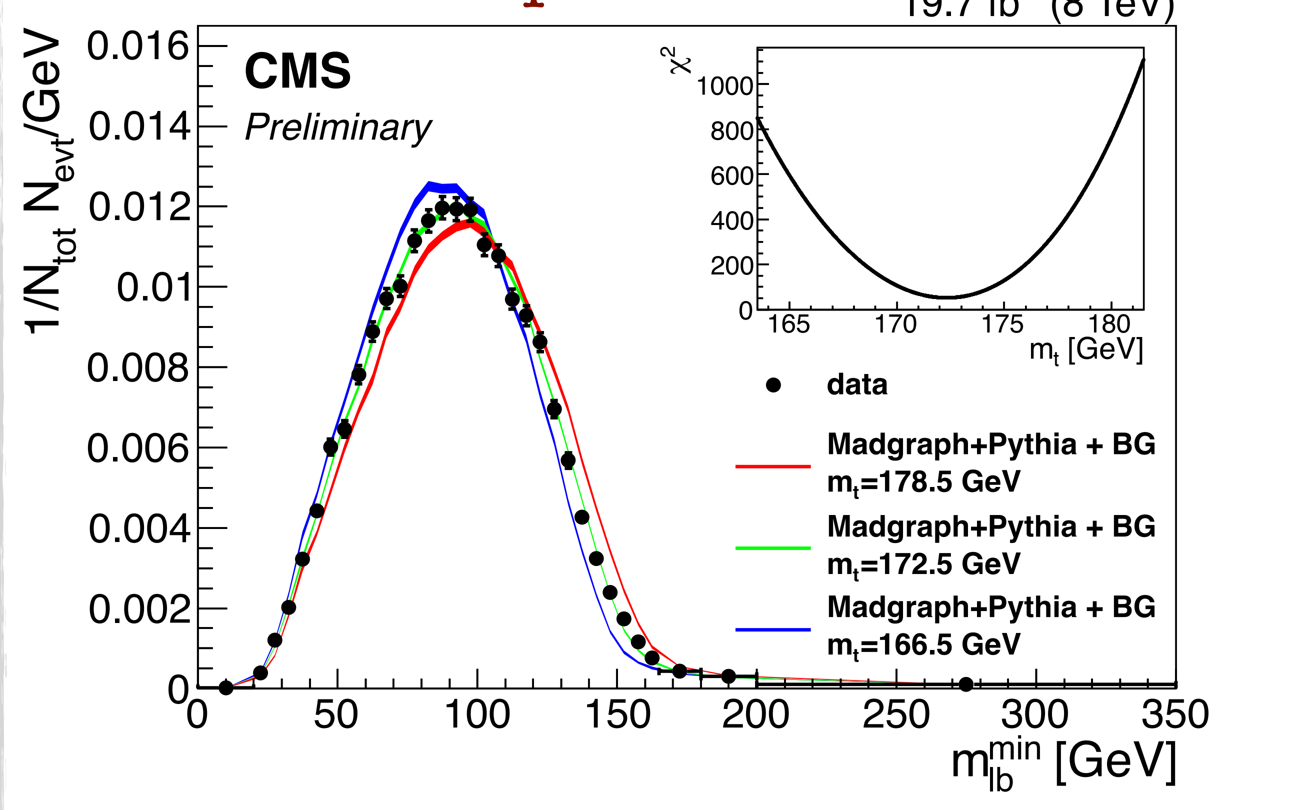




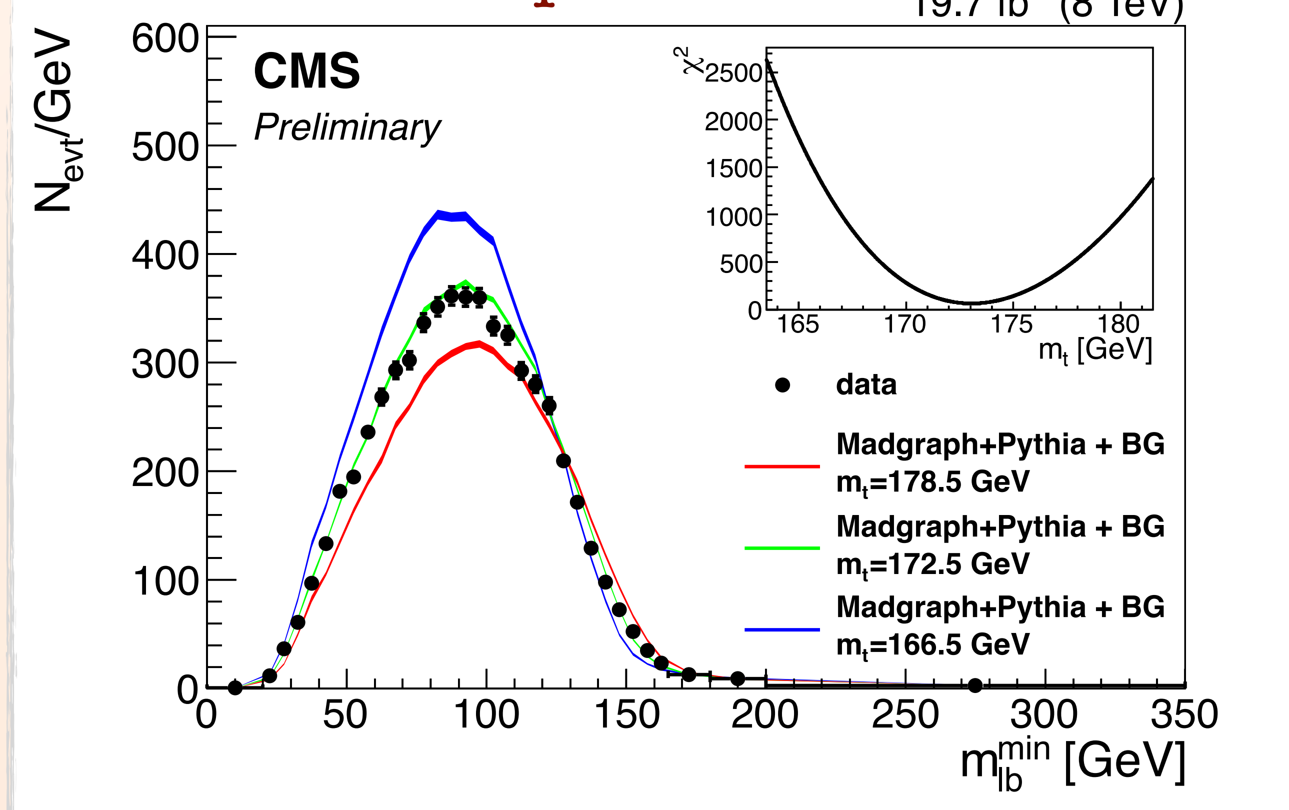
- Define estimator $$\chi^2 = \frac{(N_{\text{pred}} - N_{\text{data}})^2}{\sigma^2_{\text{pred}} + \sigma^2_{\text{data}}}$$
- Extract top mass from minimum of $\chi^2(m_t)$
- For total rate or in bins of any distribution (m_{lb})

Extraction of m_t (MadGraph+Pythia)

shape



shape & rate

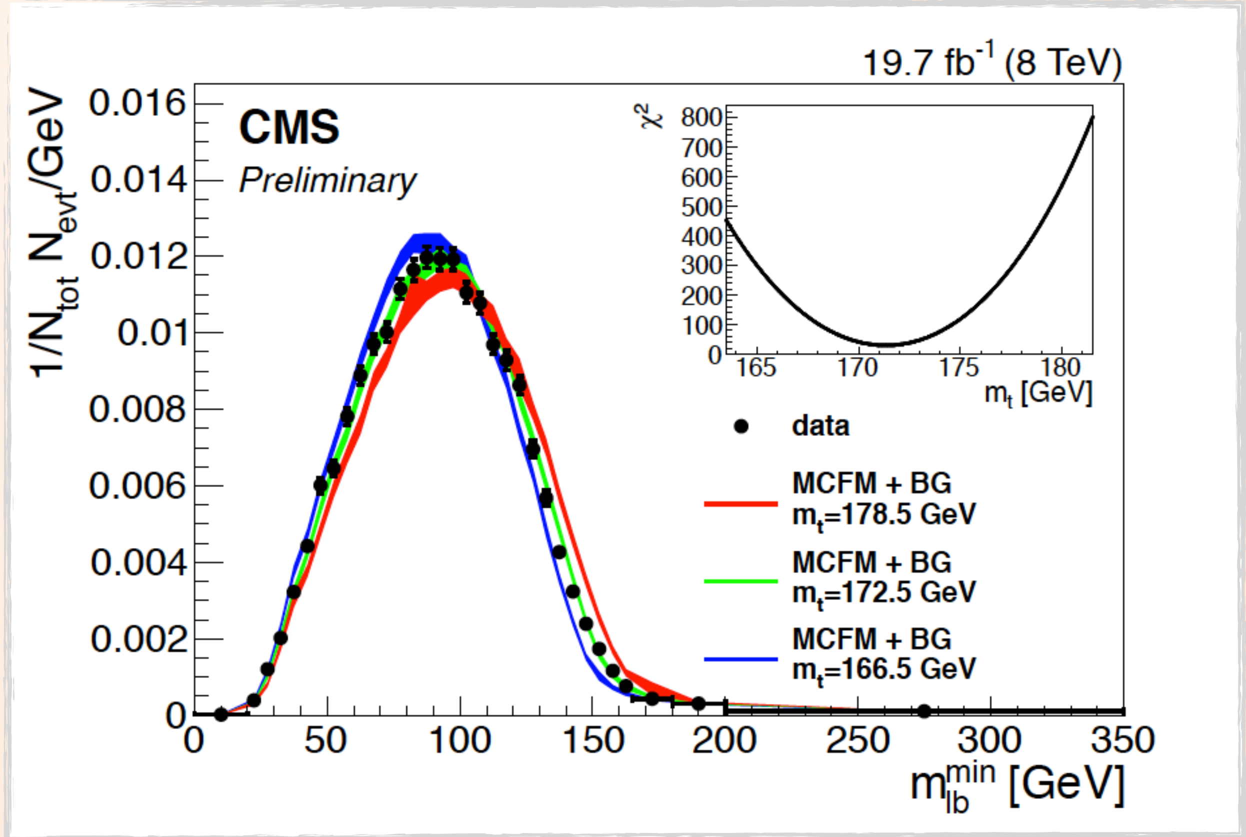


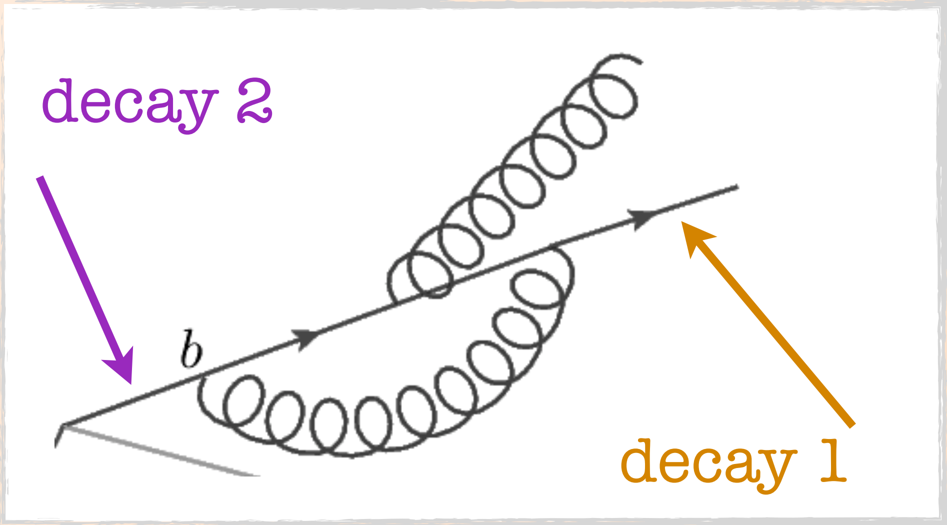
only stat uncertainties

- MadGraph+Pythia + background processes
- Compare data to different m_t hypothesis in prediction
  - for shape only, shape & rate and rate-only
- Repeated for each systematic uncertainty

Extraction using MCFM

- Use MCFM to simulate $m_{lb,min}$ distribution in. vis PS
- Fold to detector level and extract top mass
  - normalized (rate underestimated)
- Pure MCFM NLO / LO
- NLO production + decay 1
  - Real and virt. corrections in decay & production
- **NLO production + decay 2**
  - Real and virt. corrections in production only (Closest to MadGraph ME level)
- Plain LO (no FSR)





- Uncertainties similar to MadGraph+Pythia

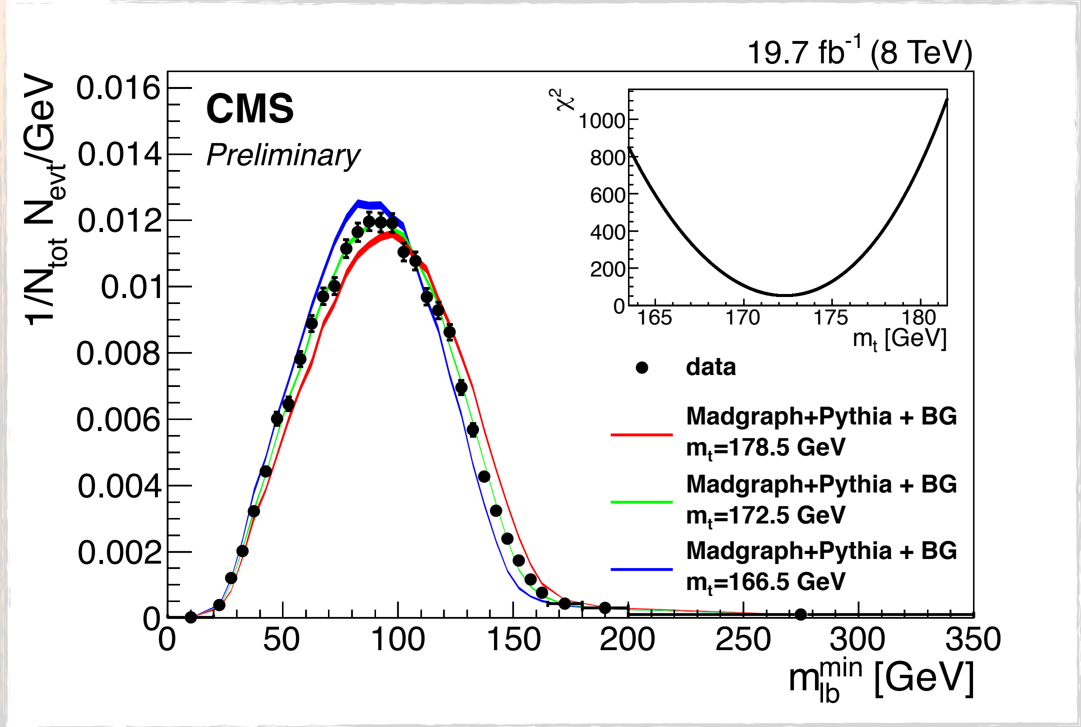
Results

- M_{lb} shows low m_t-sensitivity wrt. higher orders in production
- Overall consistent results from "shape", "shape & rate" and "rate" based extraction

Prediction	Fit method	Fitted m_t [GeV] from m_{lb}^{min}
MADGRAPH+PYTHIA	shape+rate	$173.1^{+1.9}_{-1.8}$
MADGRAPH+PYTHIA	rate	$173.7^{+3.5}_{-3.4}$
MADGRAPH+PYTHIA	shape	$172.3^{+1.3}_{-1.3}$
MCFM (LO)	shape	$171.5^{+1.1}_{-1.1}$
MCFM (NLO)	shape	$171.4^{+1.0}_{-1.1}$

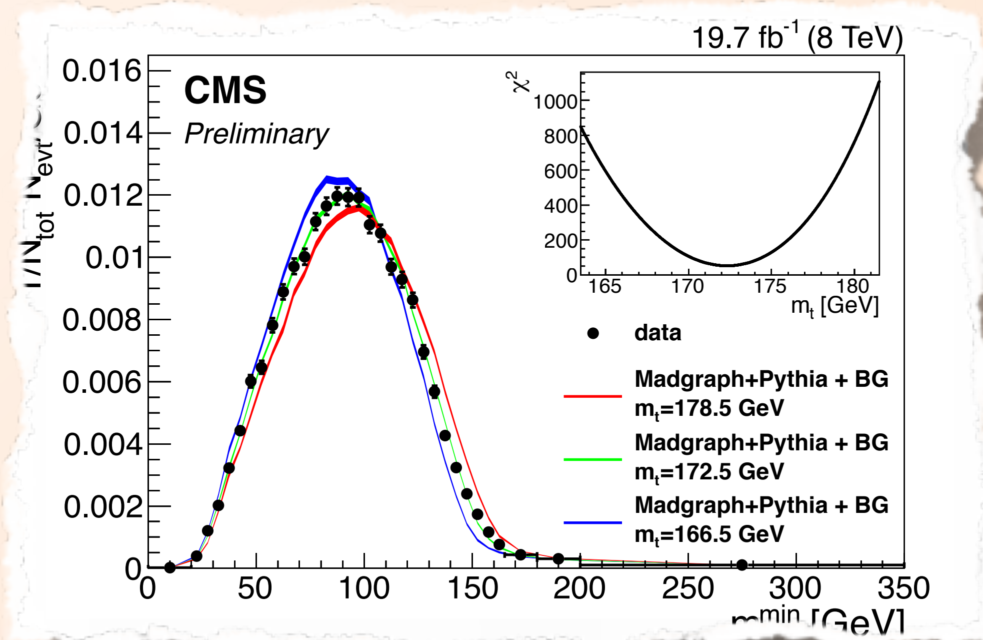
Dominant uncertainties (experimental/theory):

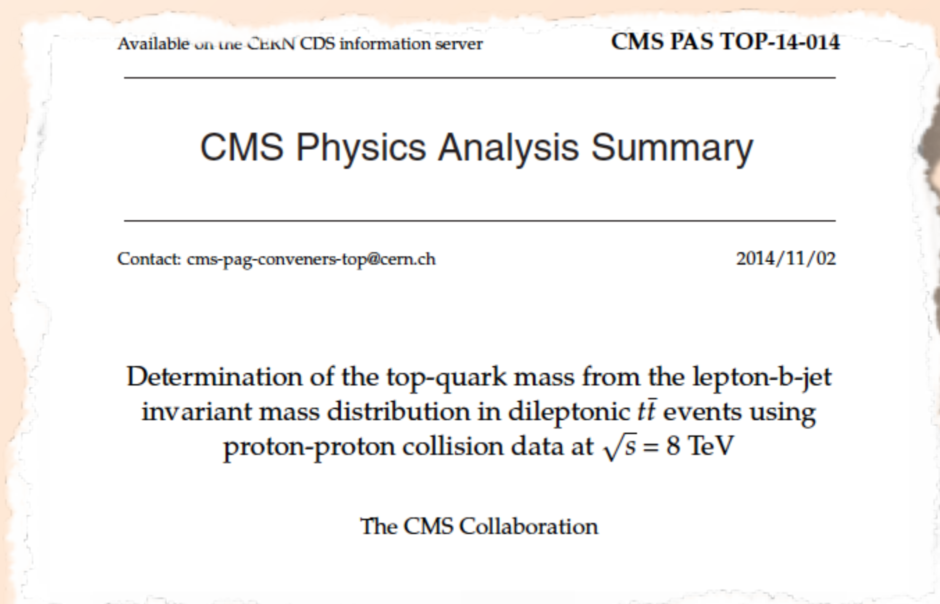
- Rate:
  - normalization uncertainties, JES
  - signal normalization
- Shape & rate:
  - partial cancellation of JES uncertainties
  - signal normalization
- Shape → smallest total uncertainty
  - JES
  - Q^2 scale



Summary

- Extracted top mass from reco-level m_{lb} distributions & total rate with MadGraph+Pythia
  - From shape: $m_t = 172.3 \pm 1.3$ GeV
  - All “MadGraph+Pythia mass”-values consistent
  - Partial cancellation of uncertainties in shape+rate extraction
- Extracted using folded MCFM LO/NLO
  - No unfolding/regularization effects
  - Low sensitivity of m_{lb} to production mechanism




Available on the CERN CDS information server — CMS PAS TOP-14-014

CMS Physics Analysis Summary

Contact: cms-pag-conveners-top@cern.ch — 2014/11/02

Determination of the top-quark mass from the lepton-b-jet invariant mass distribution in dileptonic $t\bar{t}$ events using proton-proton collision data at $\sqrt{s} = 8$ TeV

The CMS Collaboration

BACKUP

Closure Tests: Pseudo-Experiments

- Based on fitted $N_{pred}(m_t)$ (per bin)
- Data toys:
  - Random yield (Poisson) around N_{pred}
  - Same statistics as data
- Prediction toys:
  - Signal
    - Random yield (Poisson) with tt MC statistics
    - Uncorrelated for each mass point
  - Background
    - Random yield (Poisson) with background MC statistics
    - Correlated for each mass point
  - Scaled back to luminosity and added

- Extract m_t and evaluate $m_{t,in}$ vs. $m_{t,out}$ (x 10^5)

➡ No bias

➡ Stat uncertainties estimated correctly