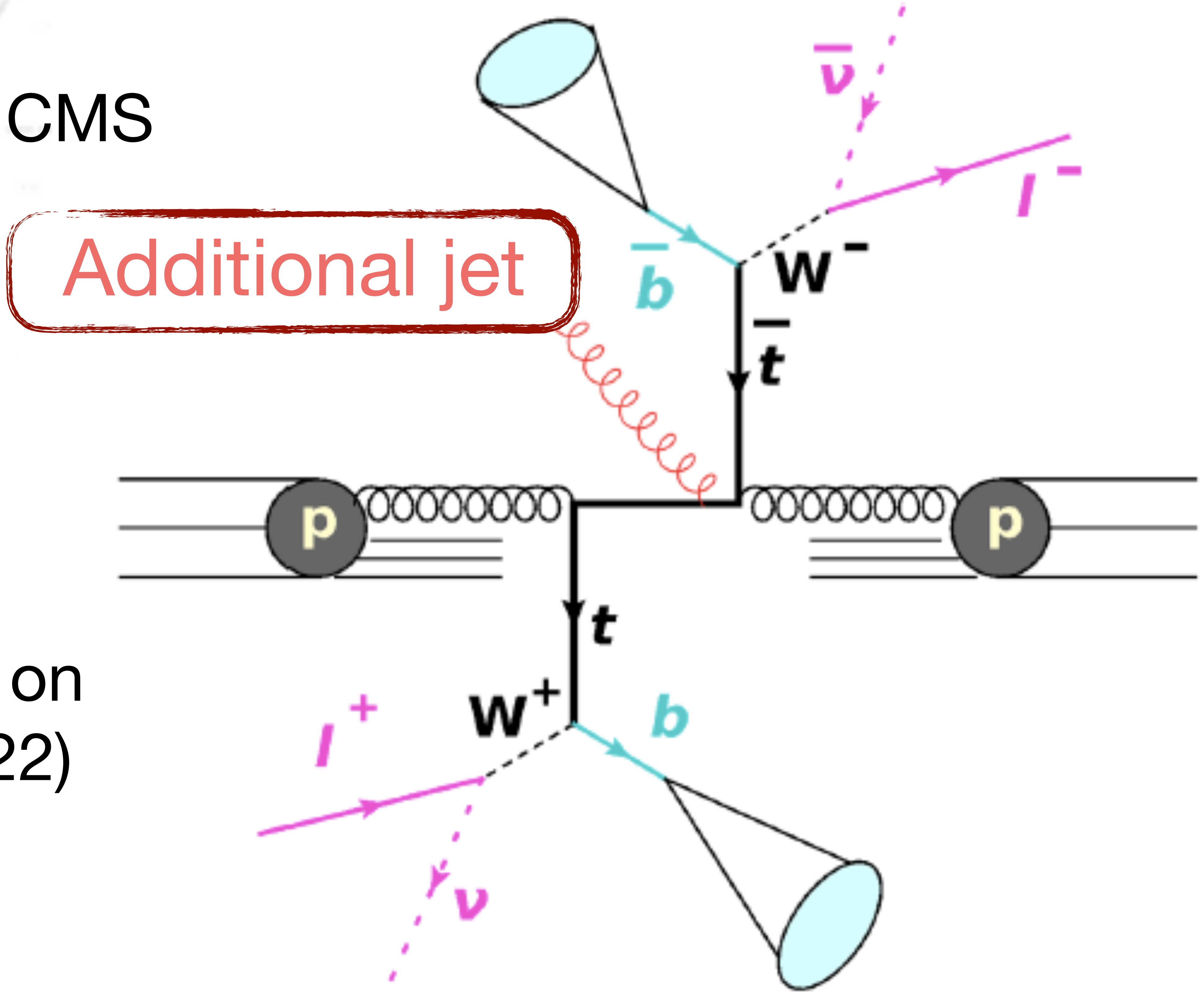
Measurement of the top quark pole mass from $t\bar{t}$+jet events at 13 TeV

[arXiv:2207.02270]

Sebastian Wuchterl (DESY) for CMS



15th International Workshop on
Top Quark Physics (TOP 2022)

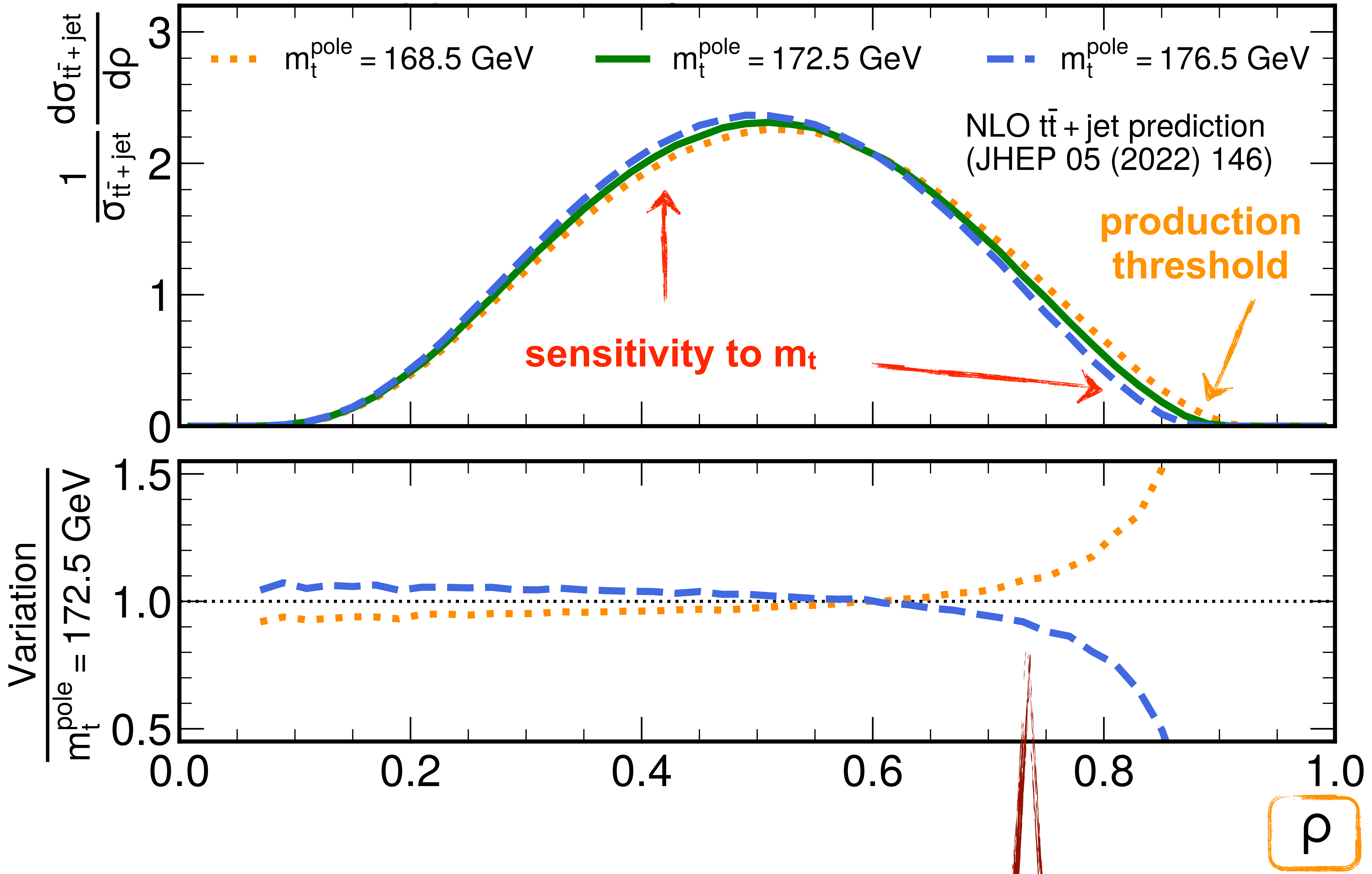
ρ observable for $t\bar{t}$+jet production

- To extract m_t^{pole}, measurements compared to **well-defined theoretical observables**
  - See M. Defranchis' talk yesterday
  - Increase sensitivity to m_t^{pole} by requiring additional jet → $t\bar{t}$+jet

→ Observable ρ **particularly suitable [1]**

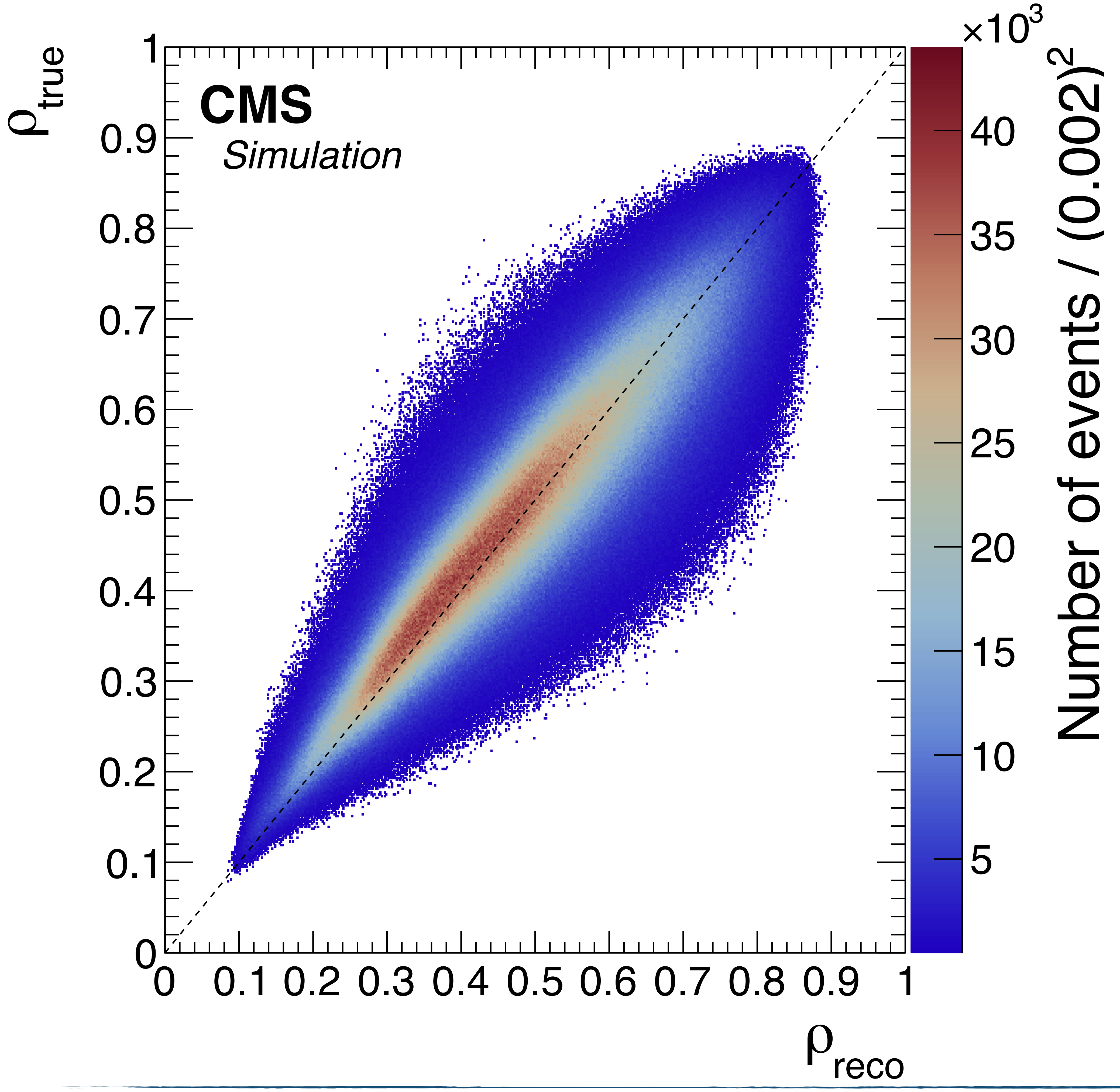
$$\mathcal{R}(m_t, \rho) = \frac{1}{\sigma_{t\bar{t}+\text{jet}}} \frac{d\sigma_{t\bar{t}+\text{jet}}}{d\rho}$$

$$\text{with } \boxed{\rho = \frac{2\,m_0}{m_{t\bar{t}+\text{jet}}}},\ m_0 = 170\,\text{GeV}$$



[1] arXiv:1303.6415

MVA based kinematic reconstruction



- Regression neural network
  - Basic 4-momenta & high level as input variables
  - Including also solutions of analytical methods
  - 100% reconstruction efficiency
  - **Factor two improvement** in ρ resolution wrt. to two common CMS methods (4-8%)

- $t\bar{t}$+jet split in 4 bins of ρ, treated as independent processes **[0., 0.3, 0.45, 0.7, 1.]**

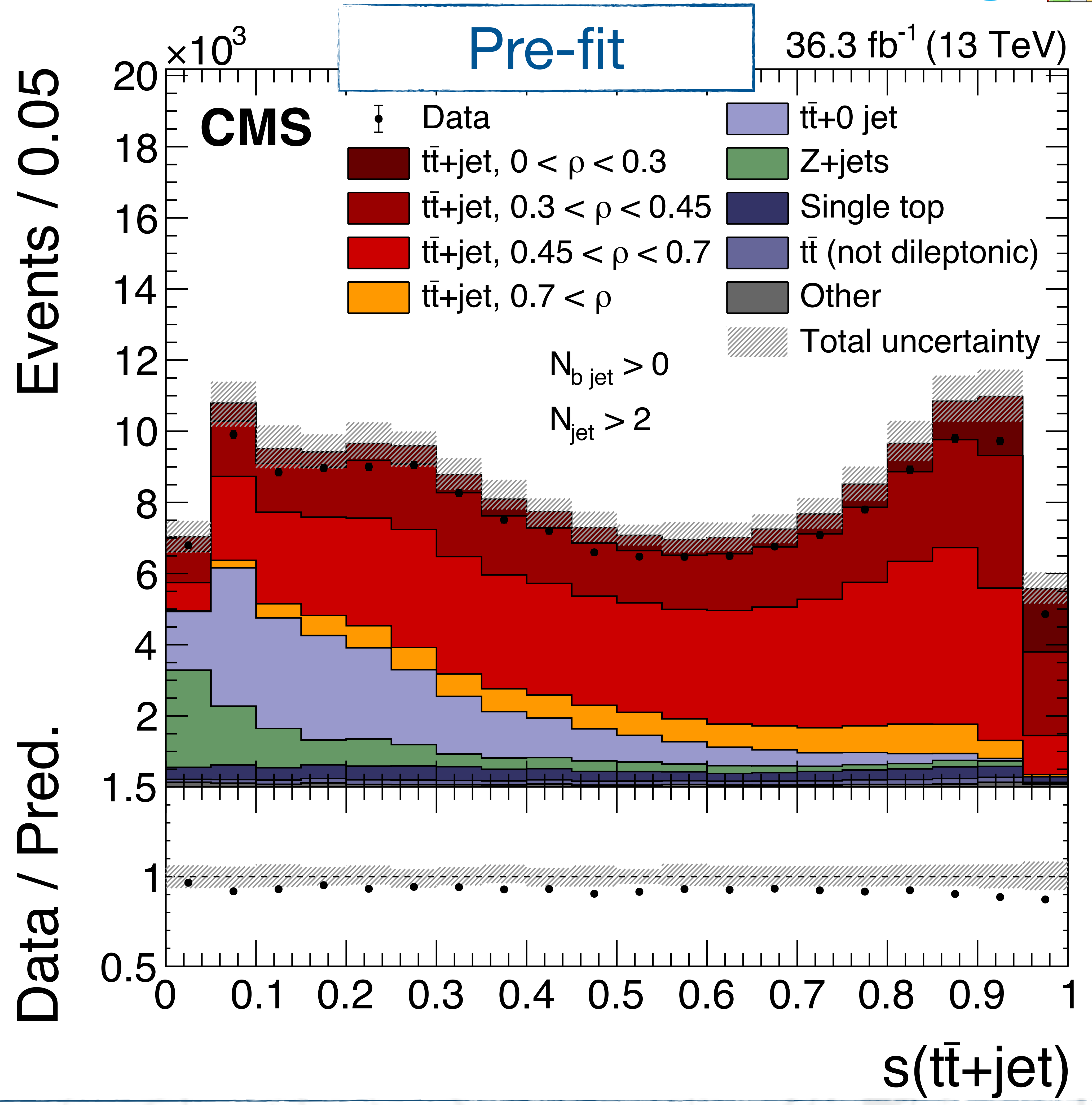
Event classification

- Best discrimination to maximize sensitivity
- Multi-classifier neural network

  $t\bar{t}$+jet signal / $t\bar{t}$+0 jet background / Z+jets
- **Same performance** for all ρ values and **bias removed**

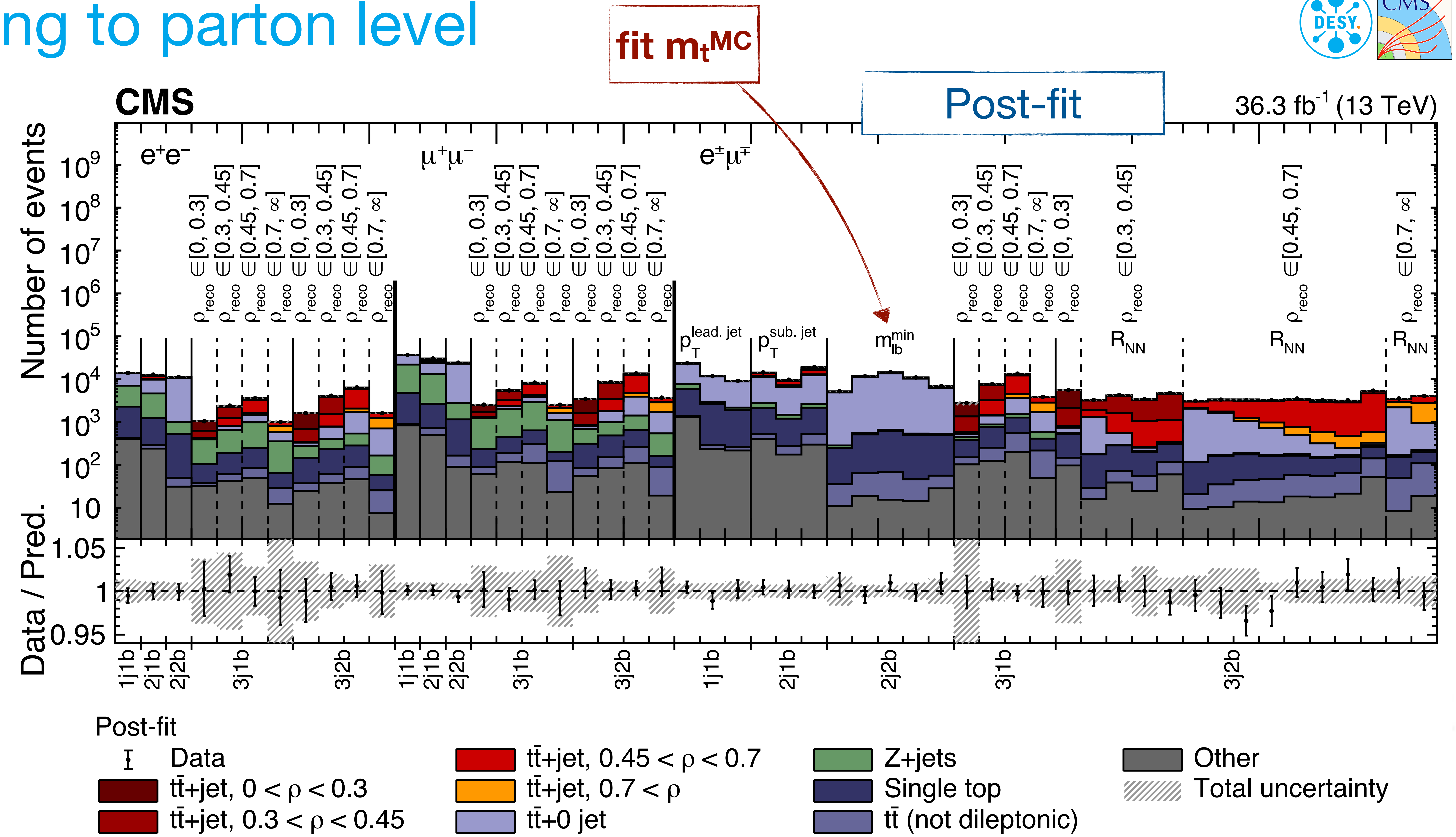
➔ Use relative response as auxiliary variable in fit

$$R_{\mathrm{NN}} = \frac{s(t\bar{t} + \mathrm{jet})}{s(t\bar{t} + \mathrm{jet}) + s(t\bar{t}\ \mathrm{bkg})}$$

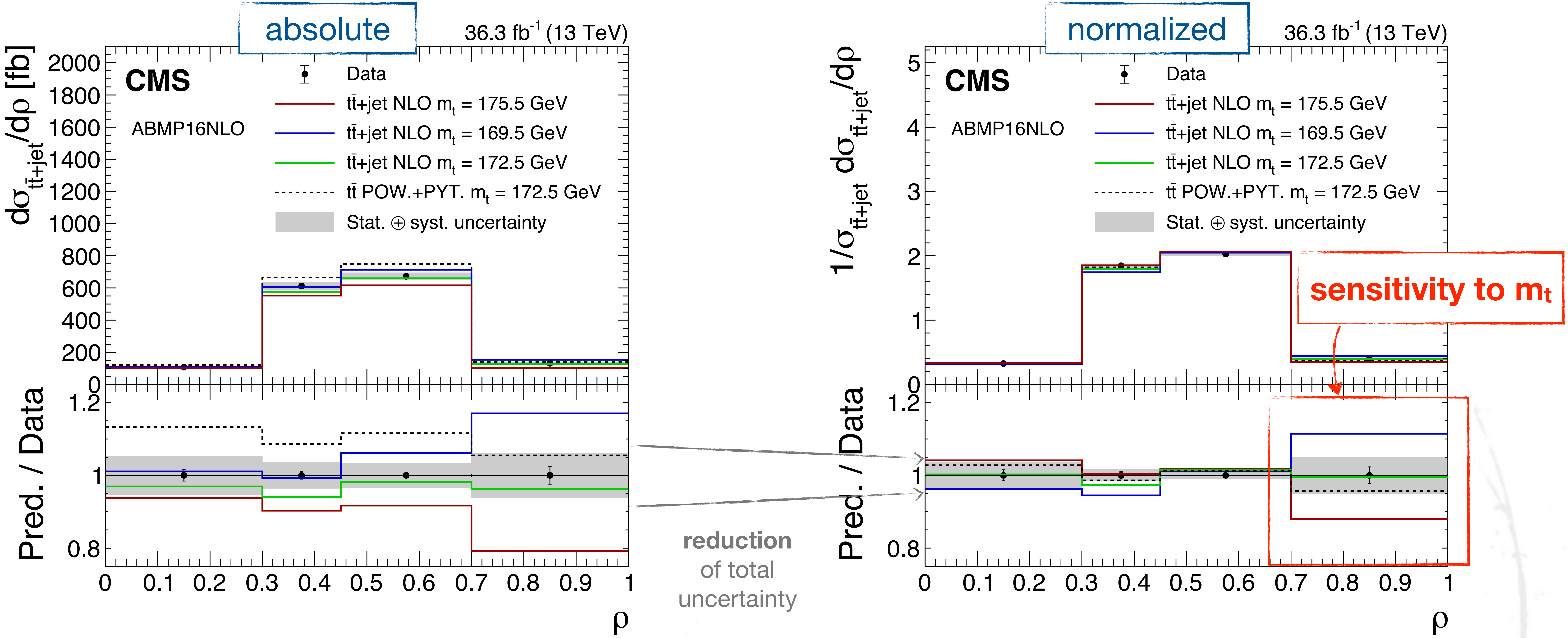


Likelihood unfolding to parton level

- Constrain backgrounds & systematics using **event categories**
- Maximize acceptance
- Very **good post-fit agreement** between data/prediction



Results: Absolute and normalised differential cross section



Event selection

kin. reconstruction

classification

categorisation

unfolding

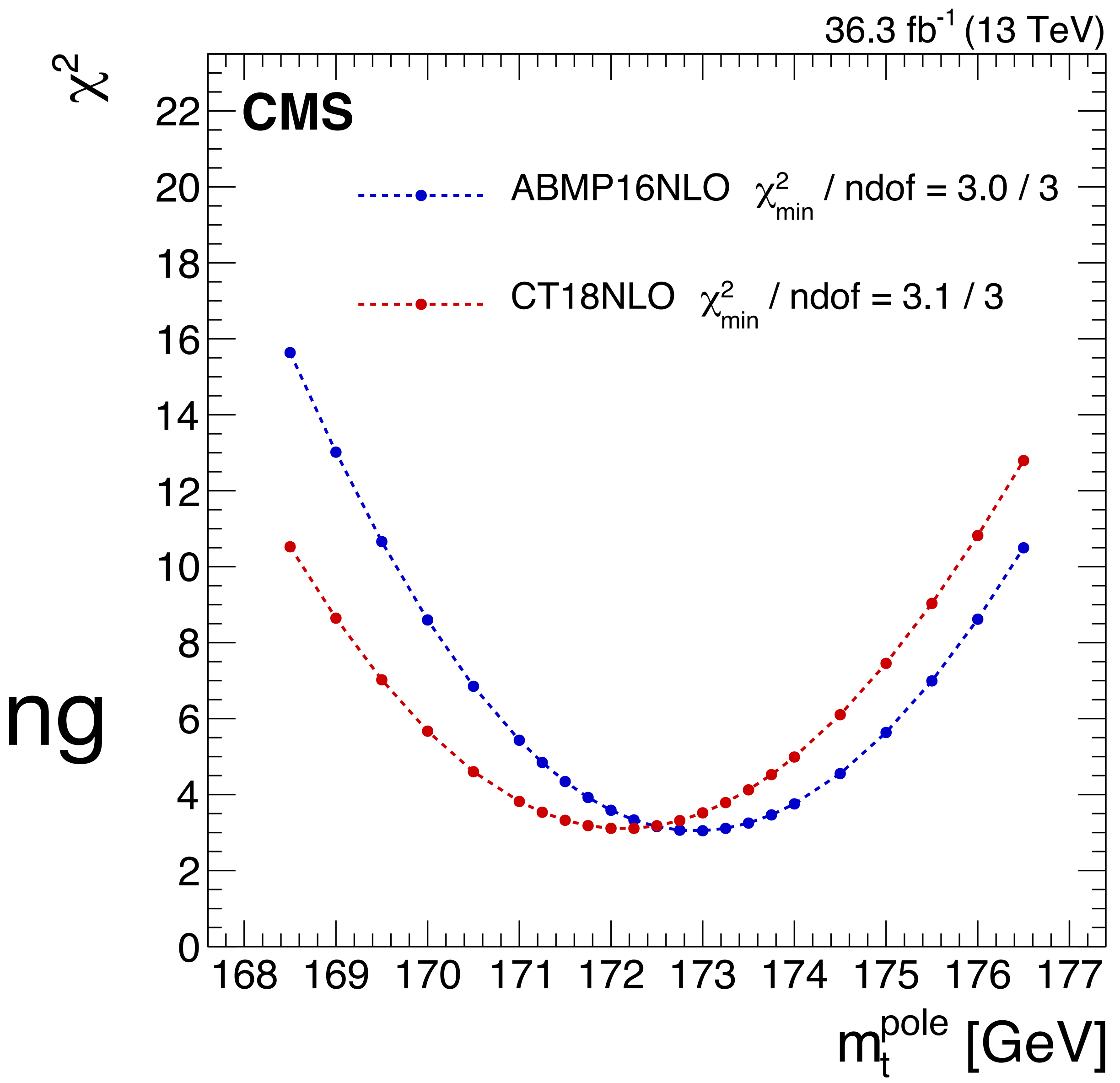
normalised differential cross section

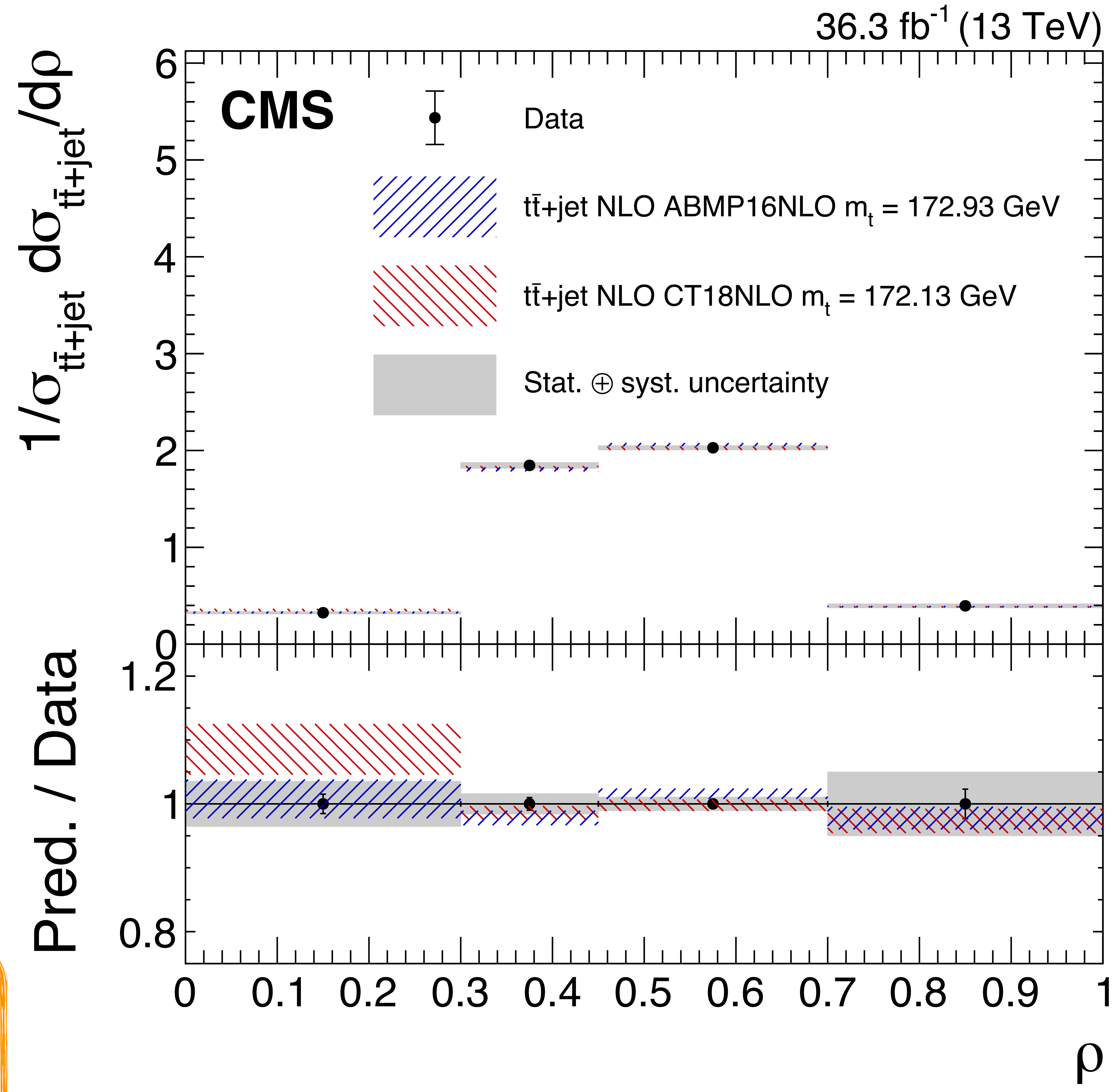
NLO theory prediction

fit of top mass

Results: Top quark mass extraction

- χ^2 scan and fit: $\chi^2 = \Delta^T V^{-1} \Delta$
  - Full covariance matrix
  - PDF uncertainties in Δ
  - w/ extrapolation uncertainties
    - Ignore constraints on modeling uncs. outside acceptance
- Scale uncertainty from envelope





$m_t^{pole} = 172.16 \pm 1.34\,\text{(fit+PDF+extr)}\,^{+0.50}_{-0.40}\,\text{(scale)}$ (CT18NLO)

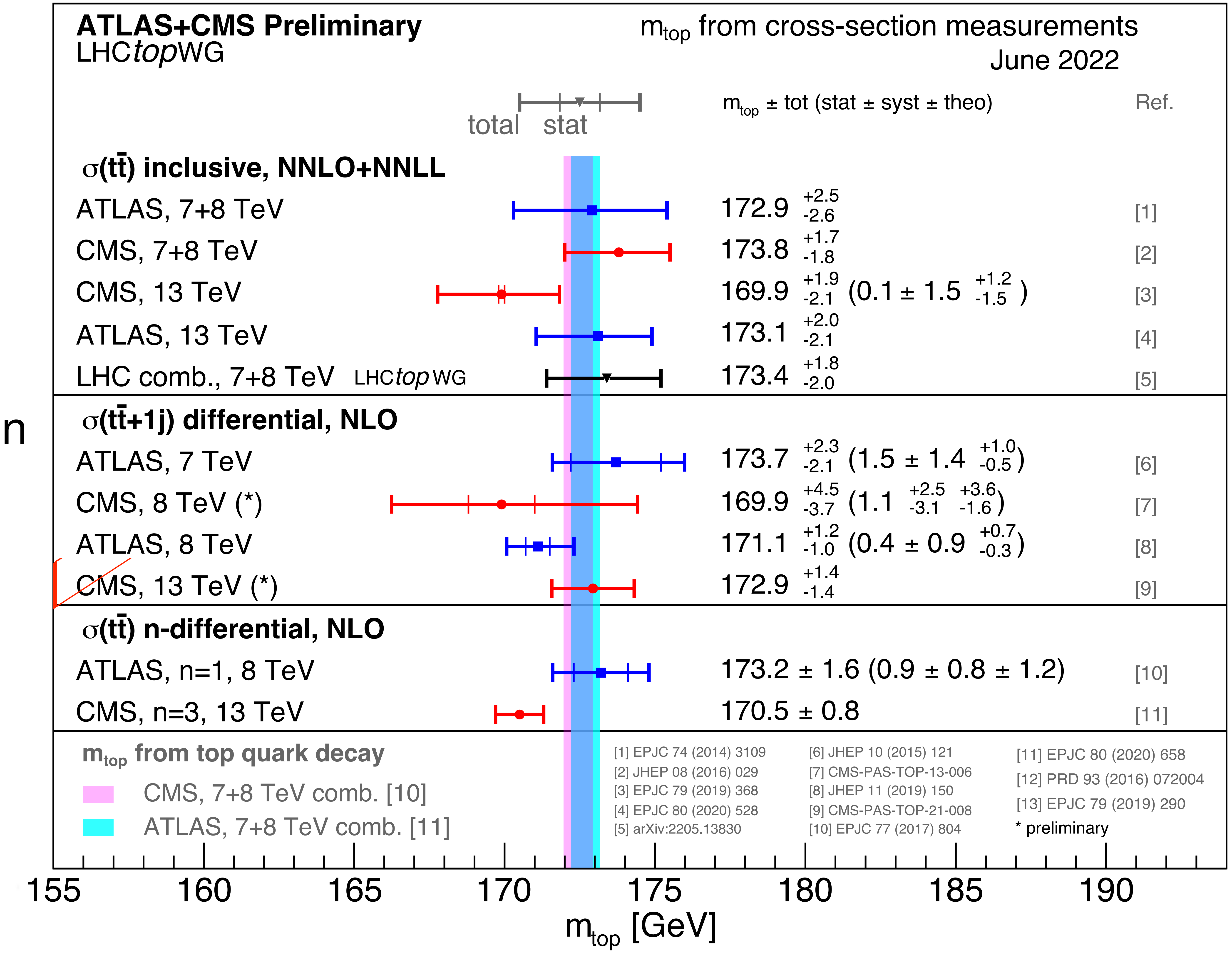
$m_t^{pole} = 172.94 \pm 1.26\,\text{(fit+PDF+extr)}\,^{+0.51}_{-0.43}\,\text{(scale)}$ (ABMP16NLO)

- Good agreement with unfolded data

Conclusions

Thank you for your attention!

- **First** measurement of $\mathbf{m_t^{pole}}$ **using** $\mathbf{t\bar{t}}$**+jet** in CMS / at 13 TeV
  - MVA techniques for ρ reconstruction and classification
  - Unfolding via profiled likelihood fit
  - Extraction of m_t^{pole} at NLO
- Triggered improvements in theory prediction
  - Implementation of dynamic scale
- **< 0.8 % precision,** compatible with previous measurements



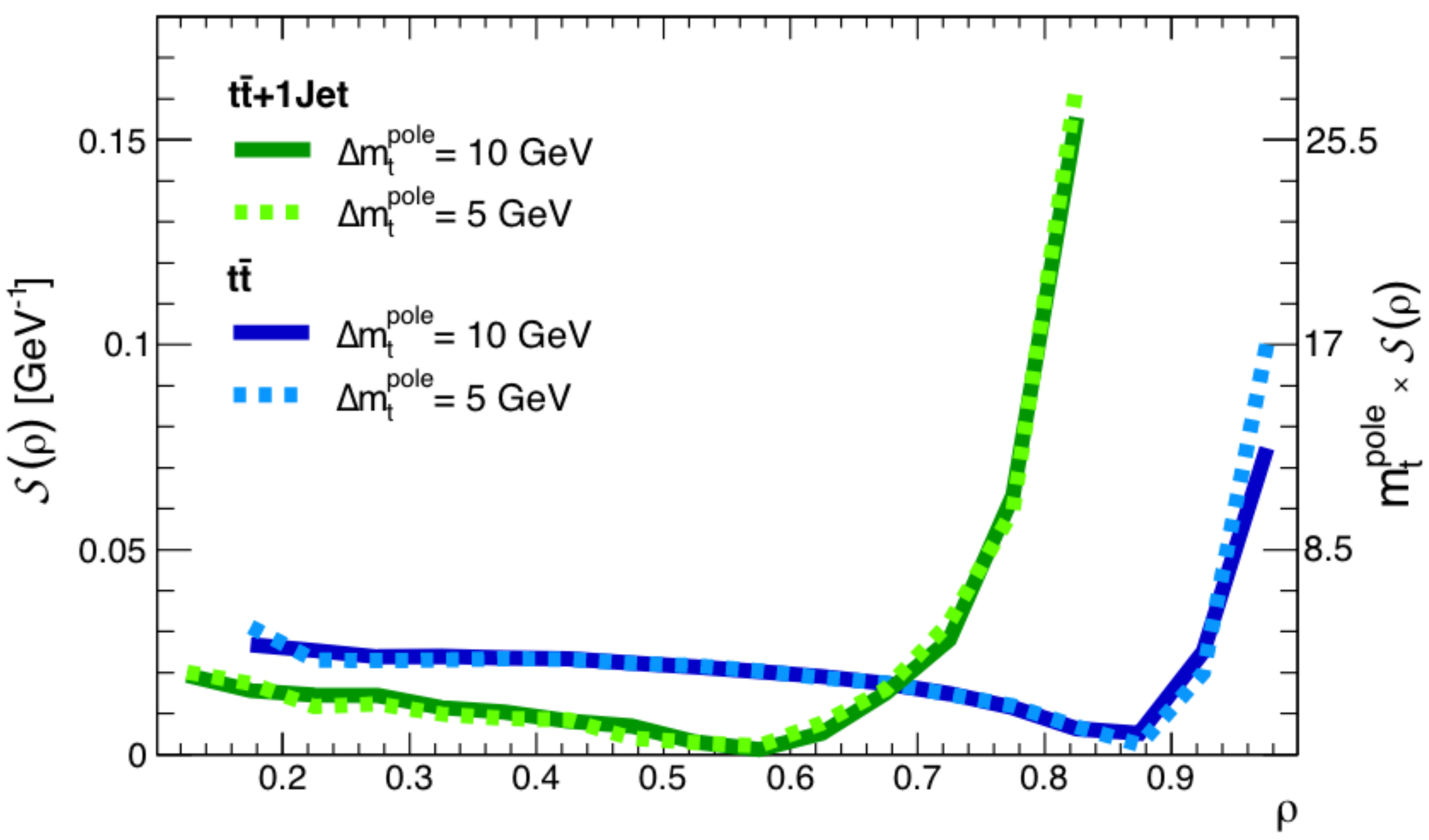
Backup

Motivation

Advantages of ρ measurement

$$\frac{1}{\sigma_{t\bar{t}+\text{jet}}}\frac{d\sigma_{t\bar{t}+\text{jet}}}{d\rho}\ldots$$

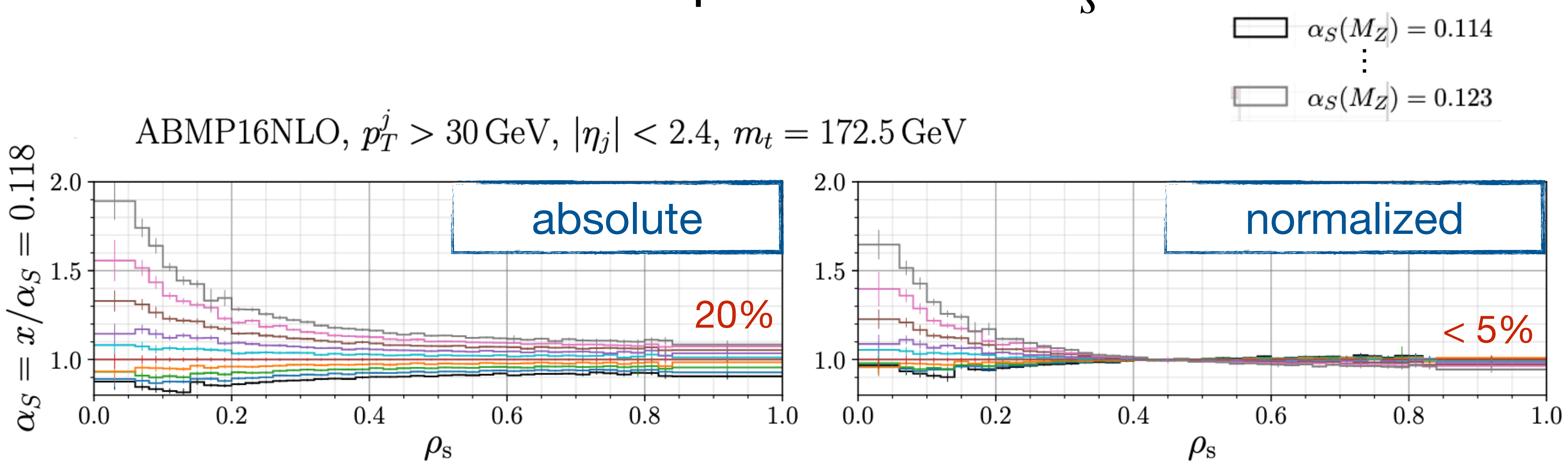
…has higher sensitivity to m_t^{pole} than for $t\bar{t}$ production



$$\mathcal{S}(\rho_s) = \sum_{\Delta=\pm 5-10\,\text{GeV}} \frac{|\mathcal{R}(170\text{ GeV},\rho_s)-\mathcal{R}(170\text{ GeV}+\Delta,\rho_s)|}{2|\Delta|\mathcal{R}(170\text{ GeV},\rho_s)}$$

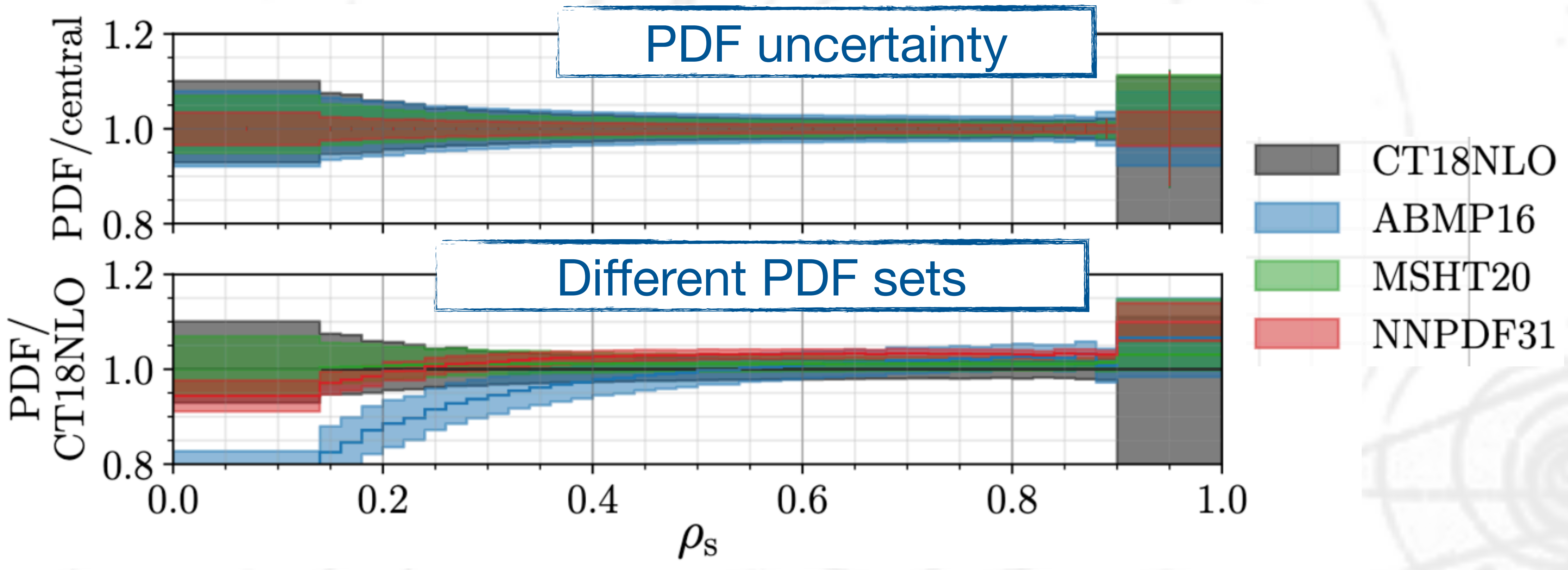
[1] arXiv:1303.6415

…has a small dependence on α_s



JHEP 05 (2022) 146

…has a residual PDF dependence



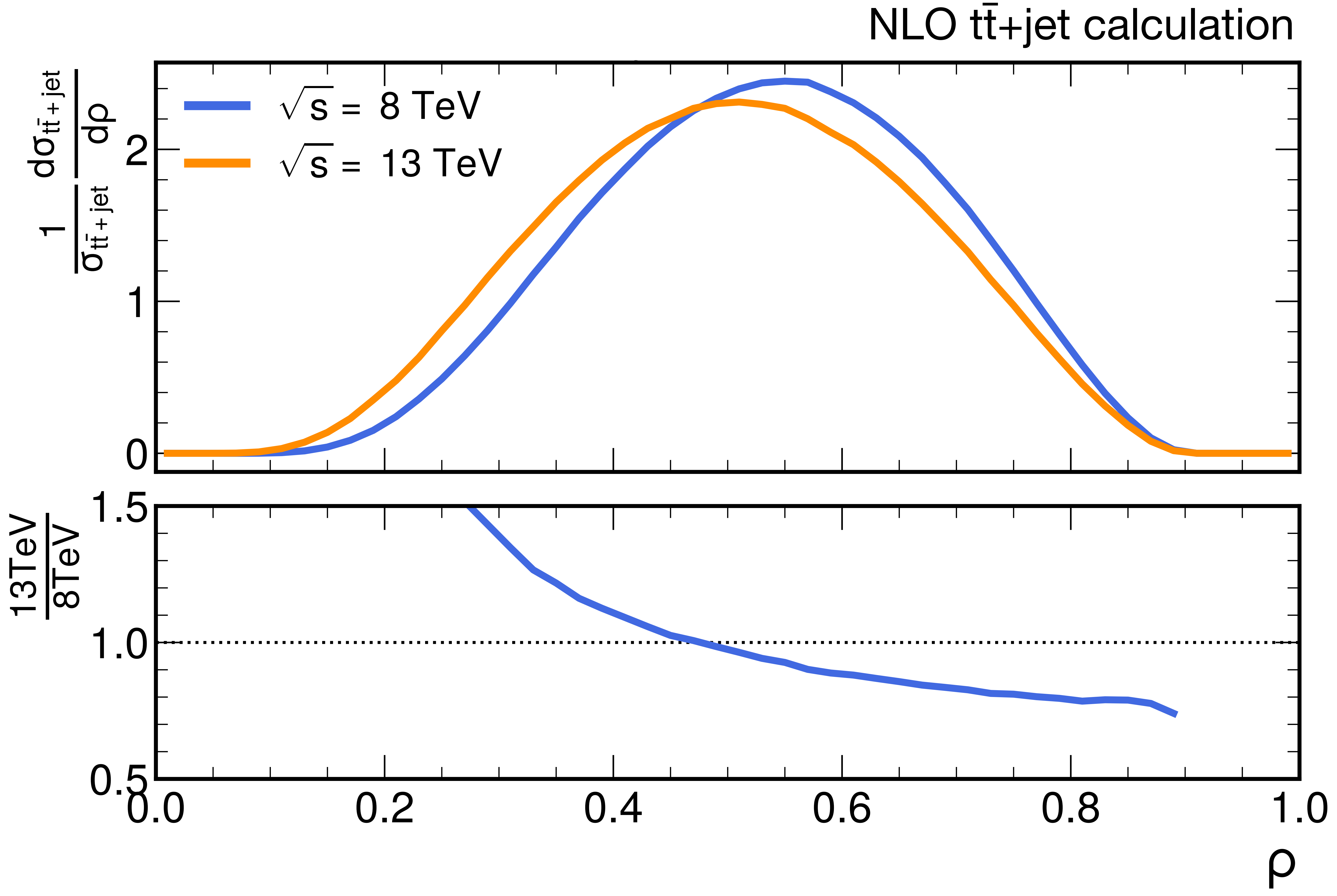
From $\sqrt{s}$ = 8 TeV to 13 TeV

Improvements & drawbacks

- Change in $\sqrt{s}$ leads to change of shape of $\mathscr{R}$ distribution
  → Less sensitivity at 13 TeV

- Progress in theory prediction:
  - Dynamic scale implemented
    → Flat LO/NLO scale factors
    → Symmetric and reduced scale uncertainties

JHEP 05 (2022) 146



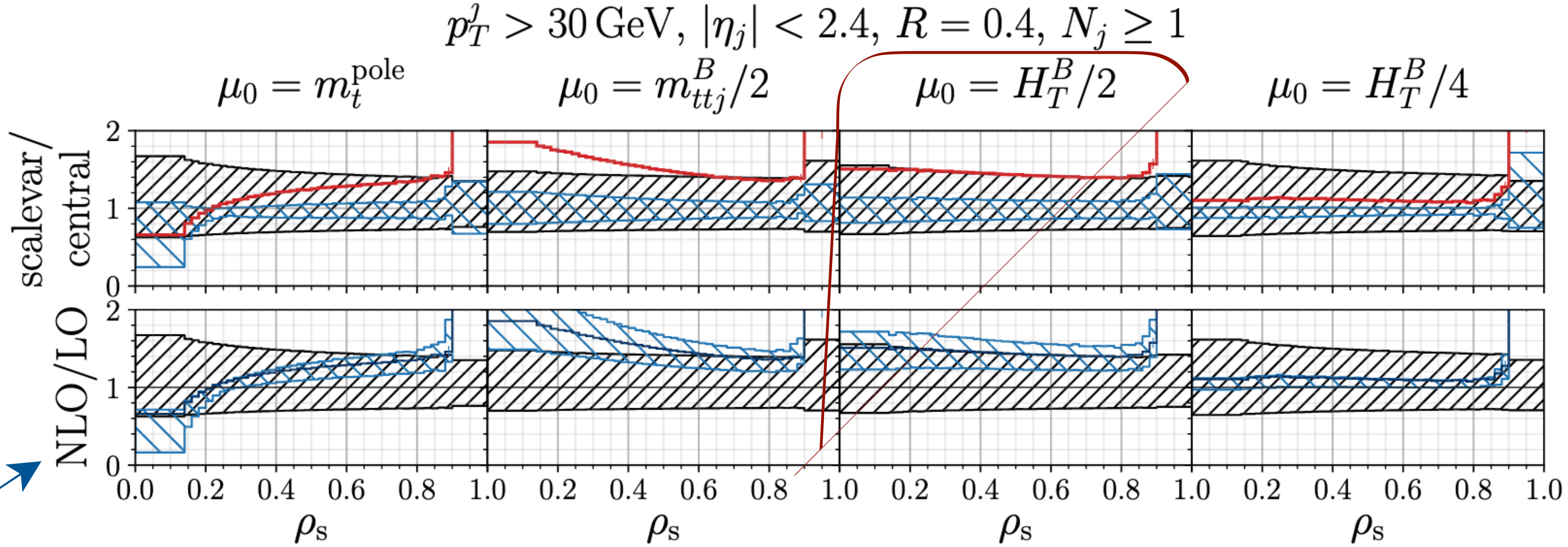
From $\sqrt{s}$ = 8 TeV to 13 TeV

Improvements & drawbacks

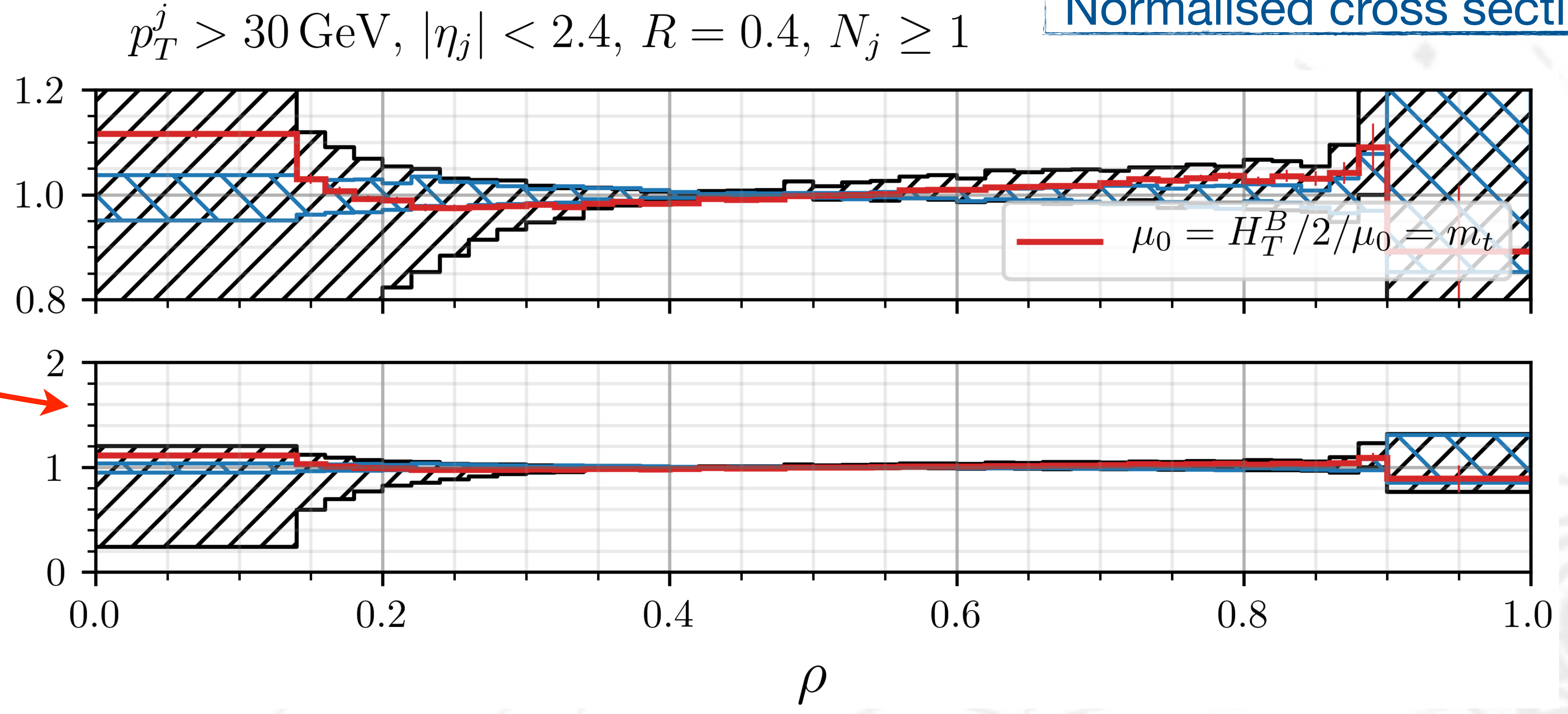
- Change in $\sqrt{s}$ leads to change of shape of $\mathscr{R}$ distribution
  - → Less sensitivity at 13 TeV
- Progress in theory prediction:
  - Dynamic scale implemented
    - → Flat LO/NLO scale factors
    - → Symmetric and reduced scale uncertainties

  JHEP 05 (2022) 146

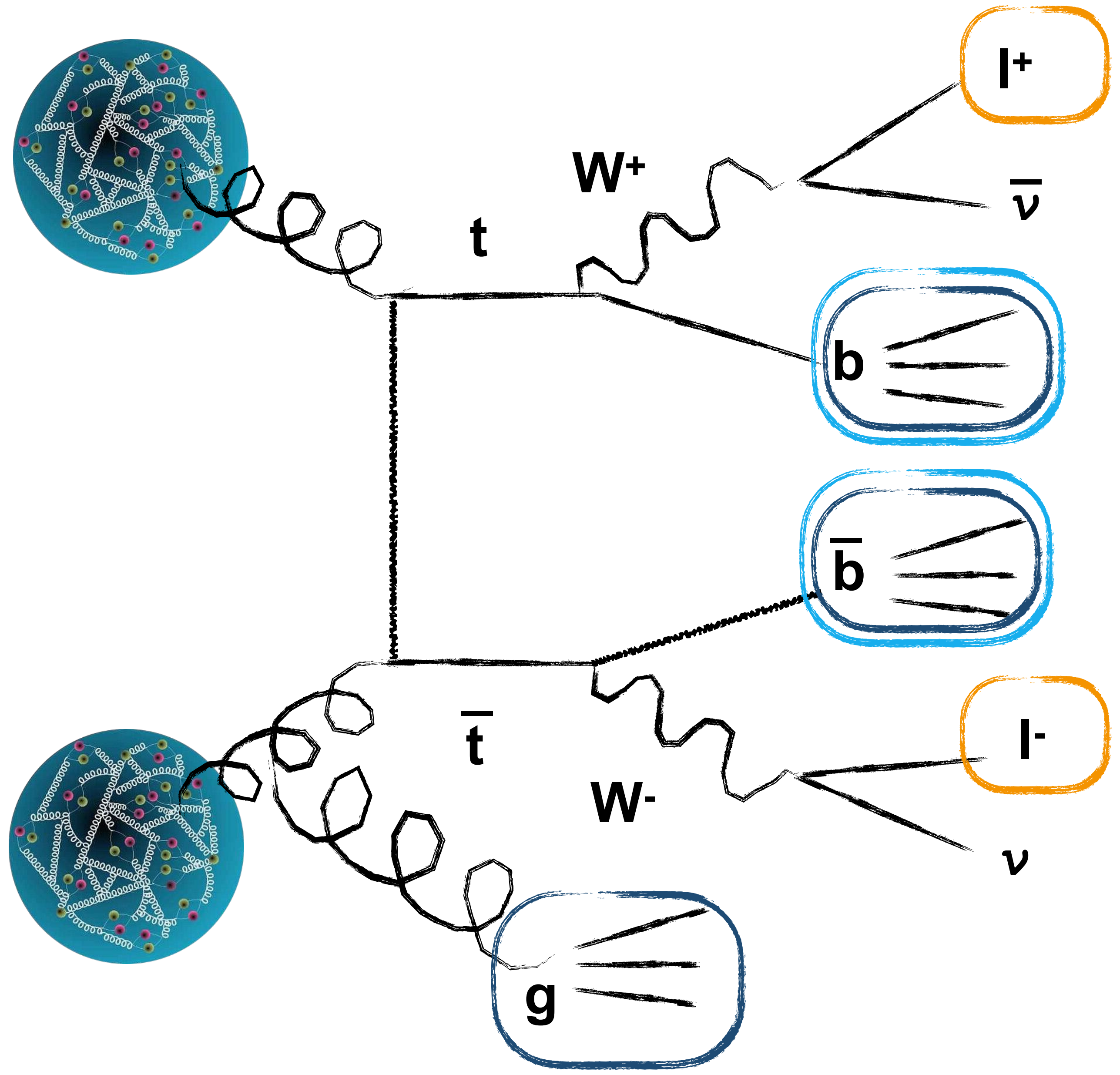
Absolute cross section



Normalised cross section

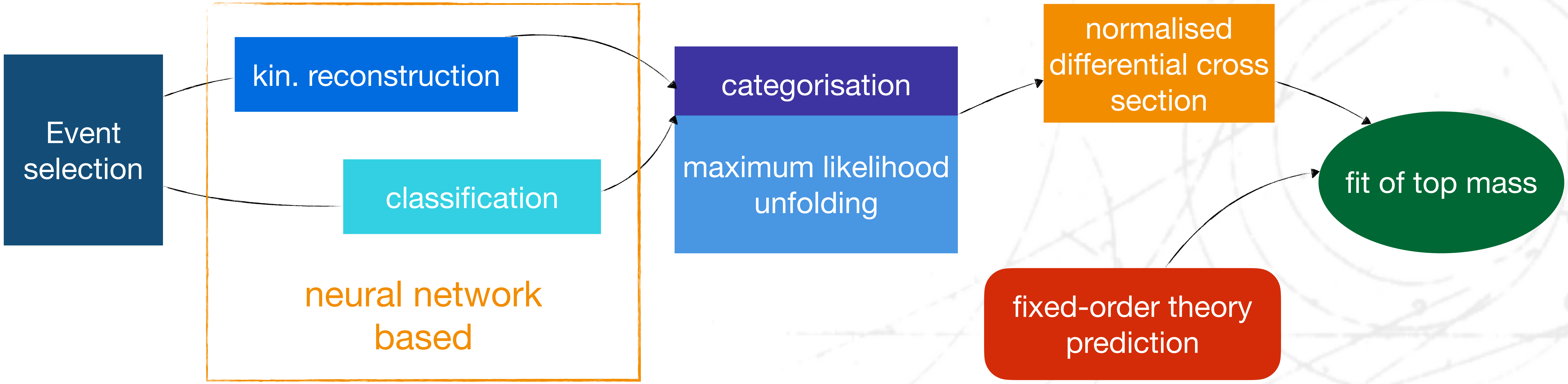


Selection & Analysis strategy



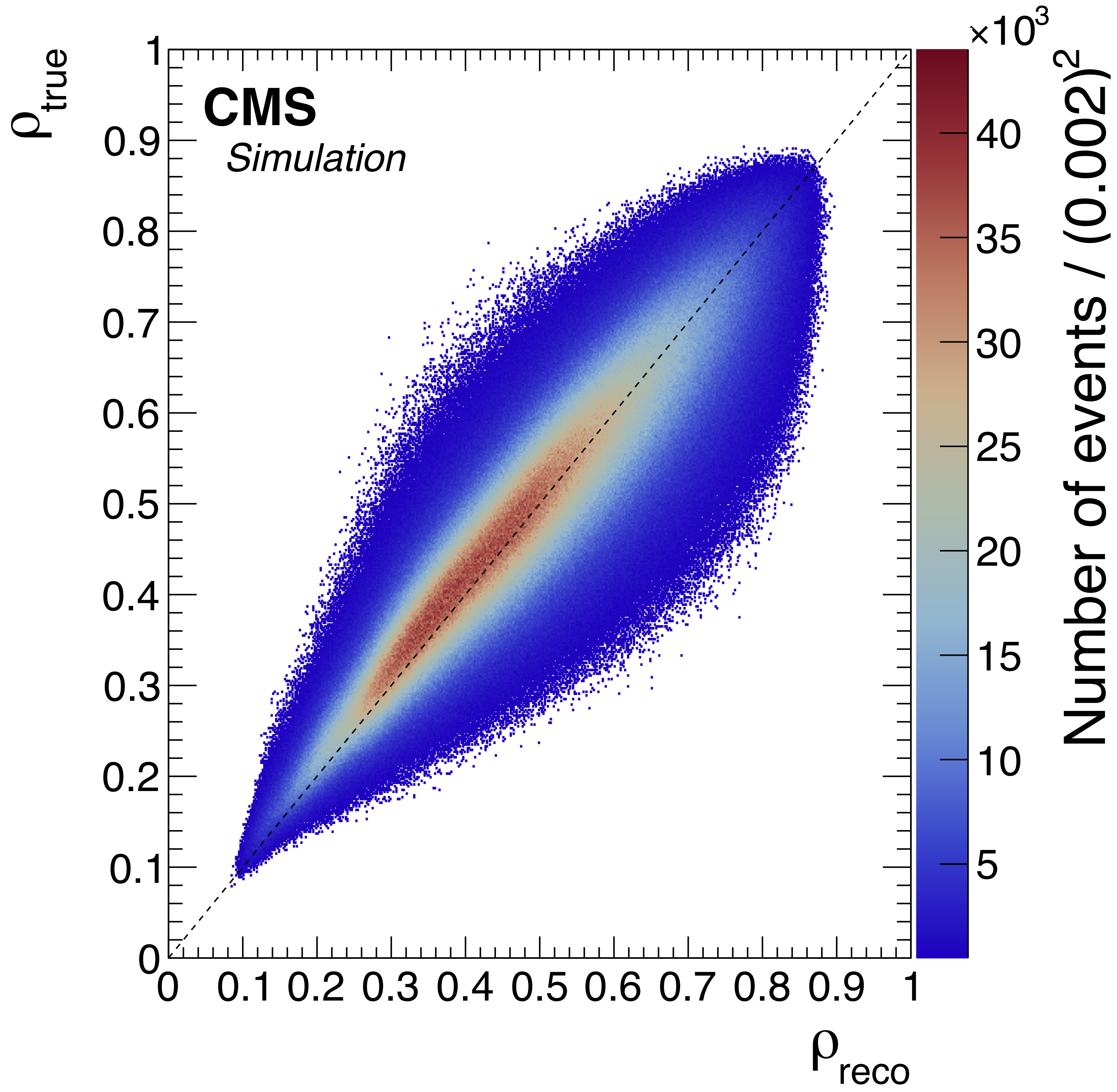
- 2 opposite-charged leptons:
  $\mu^+\mu^-$, $e^\pm\mu^\mp$, e^+e^-
  $p_T > 25$ (20) GeV & $|\eta| < 2.4$
- Jets with $p_T > 30$ GeV & $|\eta| < 2.4$
- NN based kinematic reconstruction of ρ

- e^+e^-, $\mu^+\mu^-$ channel:
  - Suppress Z+jets background
- At least 1 b jet:
  - Loose working point (10% mistagging rate)
  - b jet energy regression

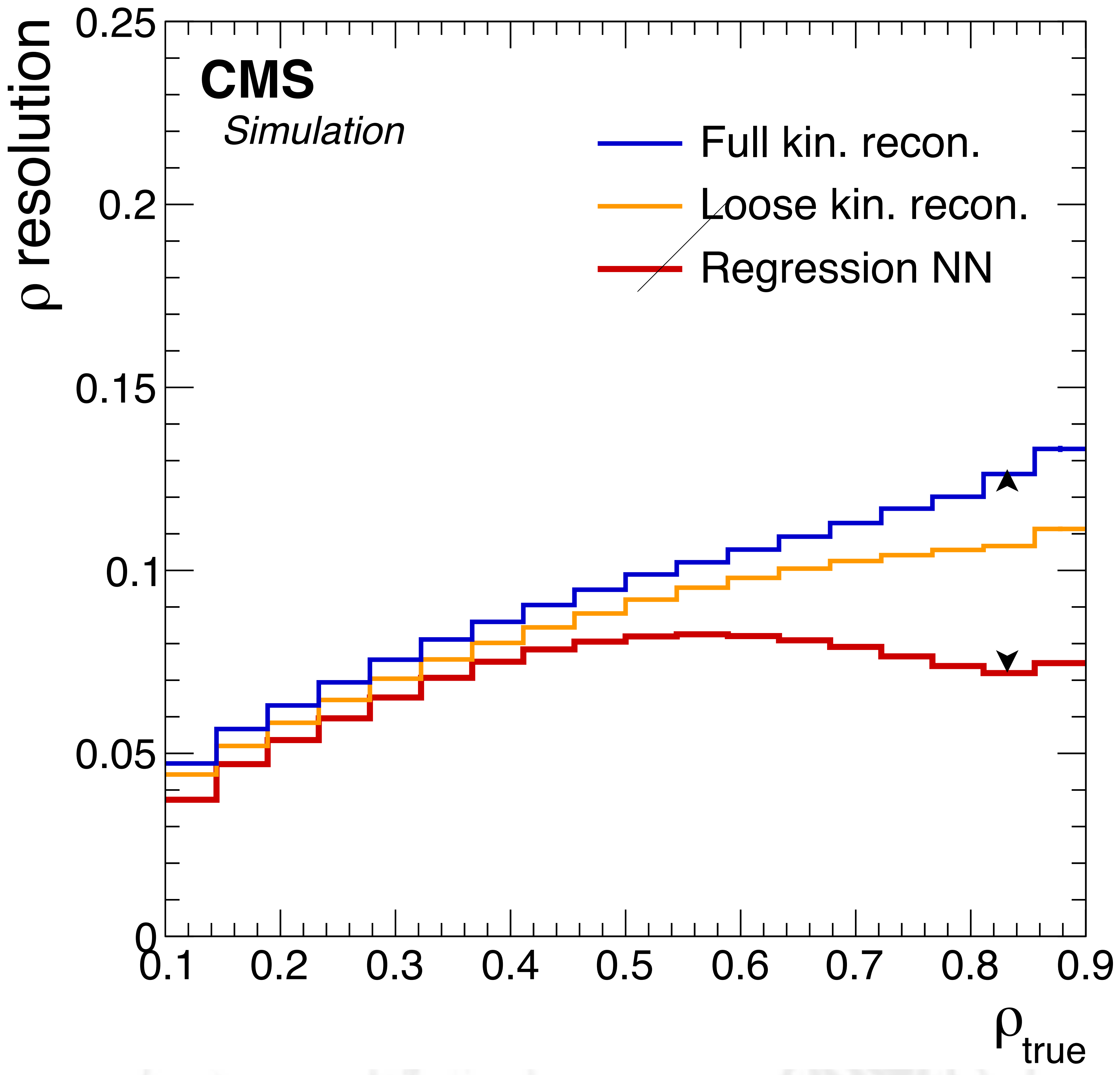


Kinematic reconstruction using neural networks

- Neural network regression for analysis
  - Basic 4-momenta & high level input variables
  - Uses also solutions of analytical methods
  - Has 100% reconstruction efficiency



Factor two improvement wrt. to two common CMS methods:



- Full kin. reconstruction using mass constraints
  - top/antitop four-momenta solved individually
- Loose kin. reconstruction without m_t constraint
  - Solve only for $t\bar{t}$ system

Binning in ρ

Data vs. prediction

- Use neural-network reconstruction for the measurement
  - Same response in data & simulation
- Determine binning in ρ at parton + detector level:
  - Use same symmetric binning
  - Based on studies of:
    - Purity & stability
    - Conditioning of unfolding problem

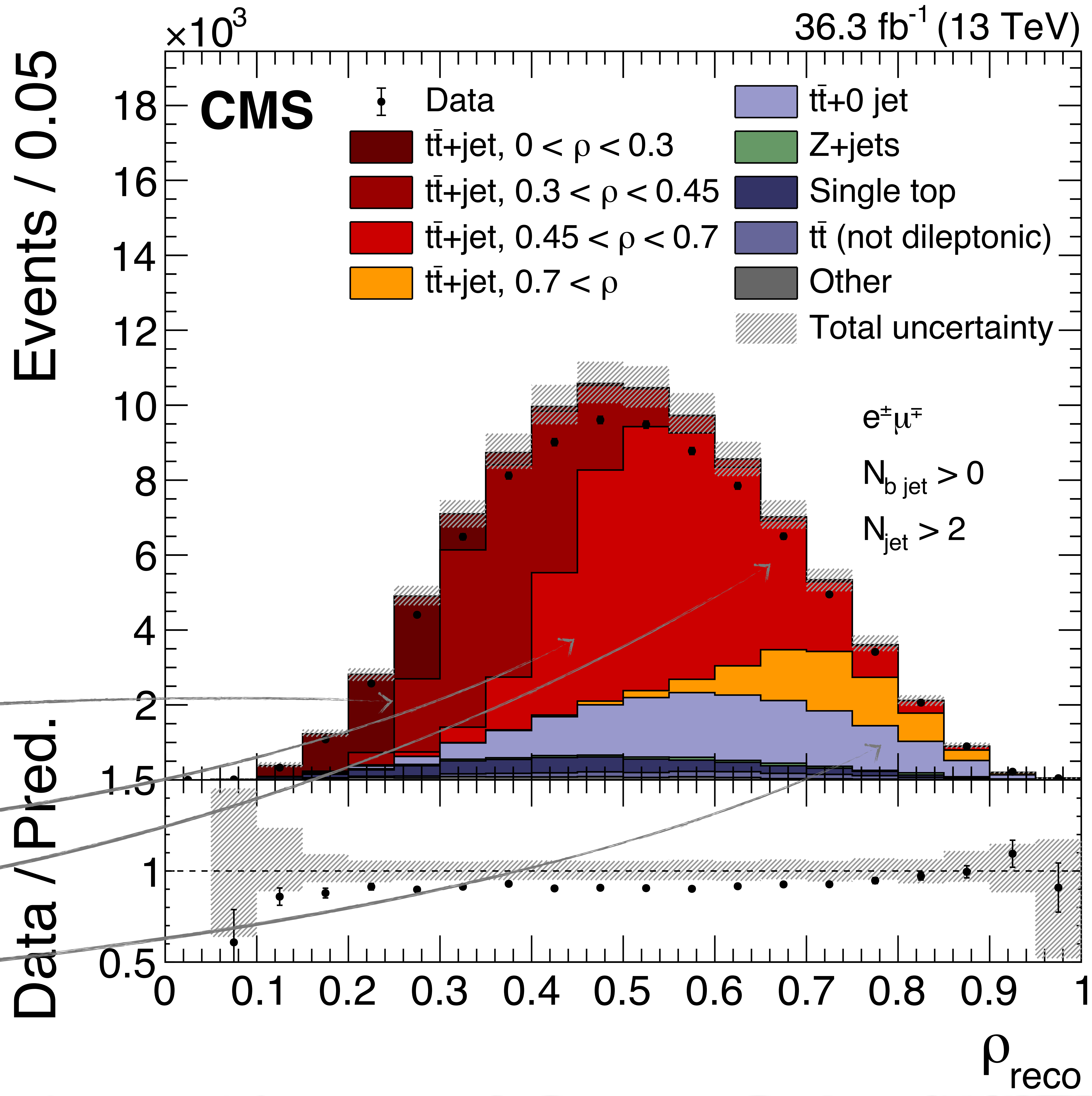
→ 4 bins:

[0, 0.3]

[0.35, 0.45]

[0.45, 0.7]

[0.7, 1]



Maximum likelihood unfolding

Overview & event categorization

- Multidimensional fit to directly measure at parton level
  - → Maximize acceptance
  - → Constrain syst. uncs. using nuisance parameters $\boxed{\vec{\lambda}}$
- Achieved via:
  - Event categories (e.g. $\mu^+\mu^-$, $e^\pm\mu^\mp$, e^+e^-)
  - Bin in $N_{b\,\text{jet}}$ and ρ_{reco} to increase signal sensitivity
  - Use N_{jet} = 1,2 as control regions to constrain $t\bar{t}$ (+0 jet) background and uncertainties
    - Fit normalisation of $t\bar{t}$+0 jet bkg simultaneously

- Introduce signal strength parameters r^k for each bin

$$\boxed{\sigma^k_{t\bar{t}+\text{jet}}} = \int_{\rho^k_{low}}^{\rho^k_{high}} \frac{d\sigma_{t\bar{t}+\text{jet}}}{d\rho} d\rho \qquad r^k = \frac{\sigma^k_{t\bar{t}+\text{jet}}}{\sigma^k_{t\bar{t}+\text{jet}}(MC)}$$

$$\mathcal{L} = \prod_i \frac{e^{-\nu_i}\nu_i^{n_i}}{n_i!} \prod_j \pi(w_j) \prod_m \pi(\lambda_m)$$

- $\boxed{m_t^{MC}}$ free floating parameter to mitigate dependence

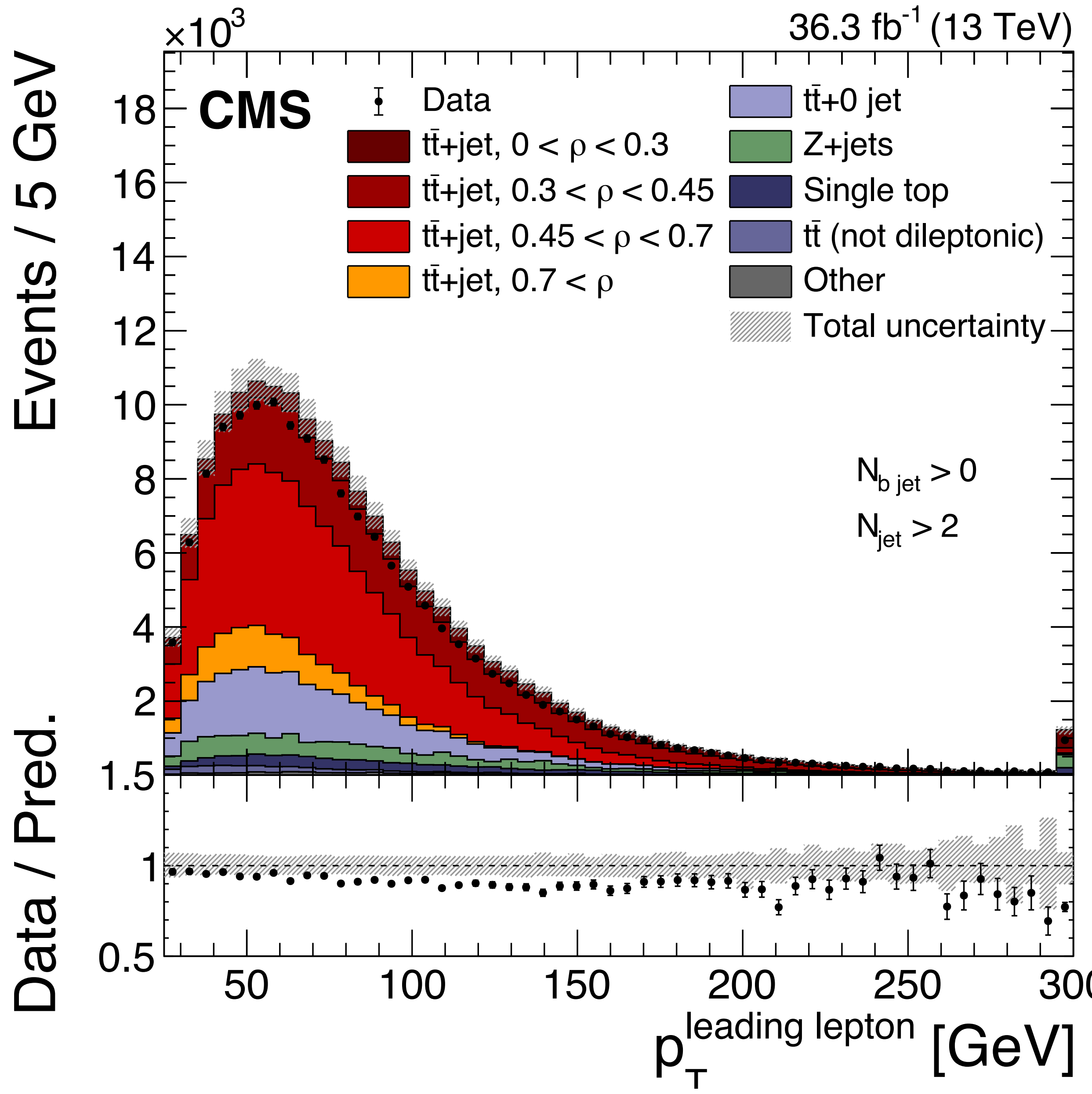
$$\nu_i = \sum_k s_i^k(\sigma^k_{t\bar{t}+\text{jet}}, m_t^{\text{MC}}, \vec{\lambda}) + \sum_j b_i^j(w_j, m_t^{\text{MC}}, \vec{\lambda})$$

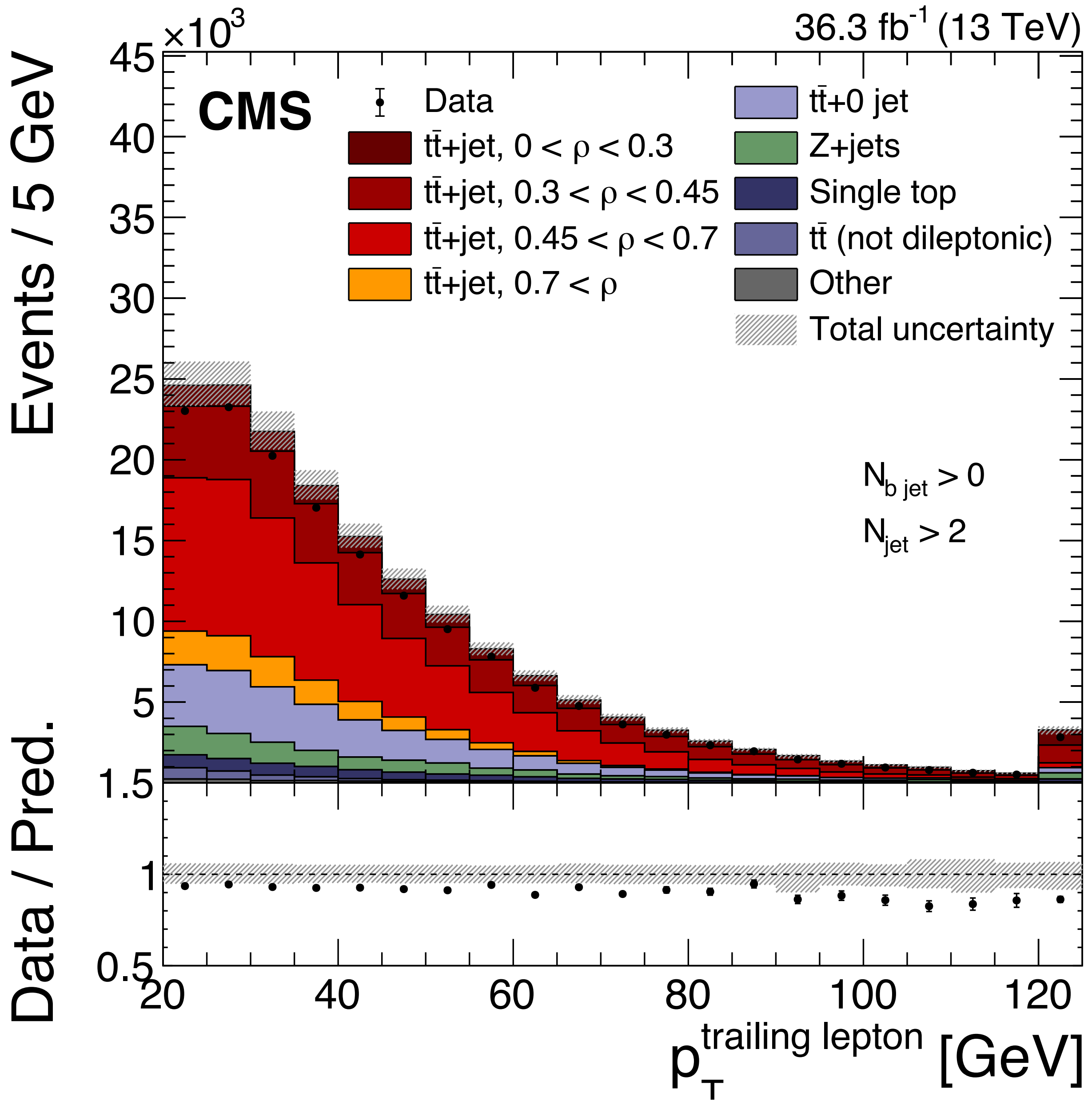
→ Consistent modeling of systematic uncertainties and correlations

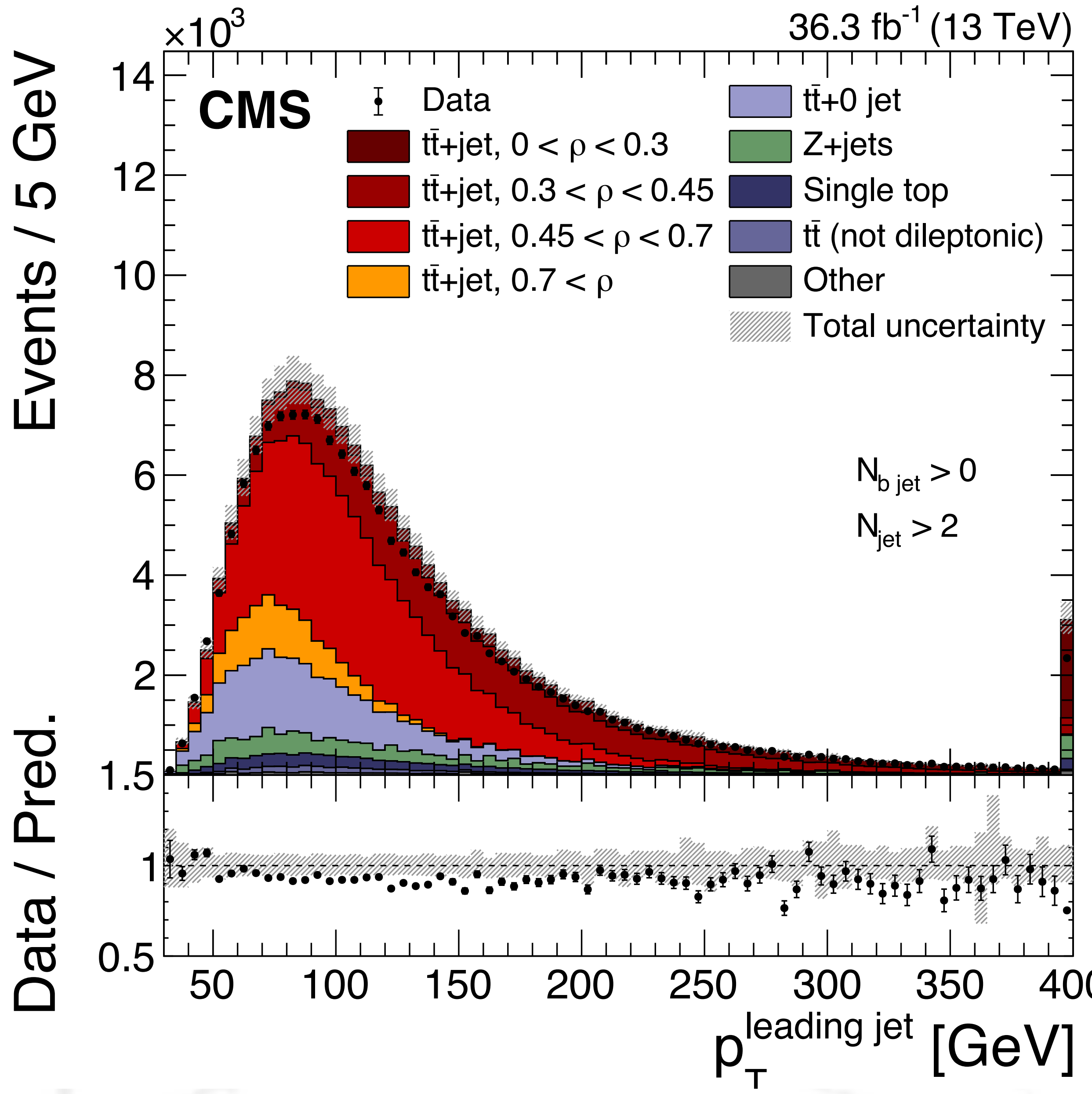
→ Background fitted and subtracted in fit

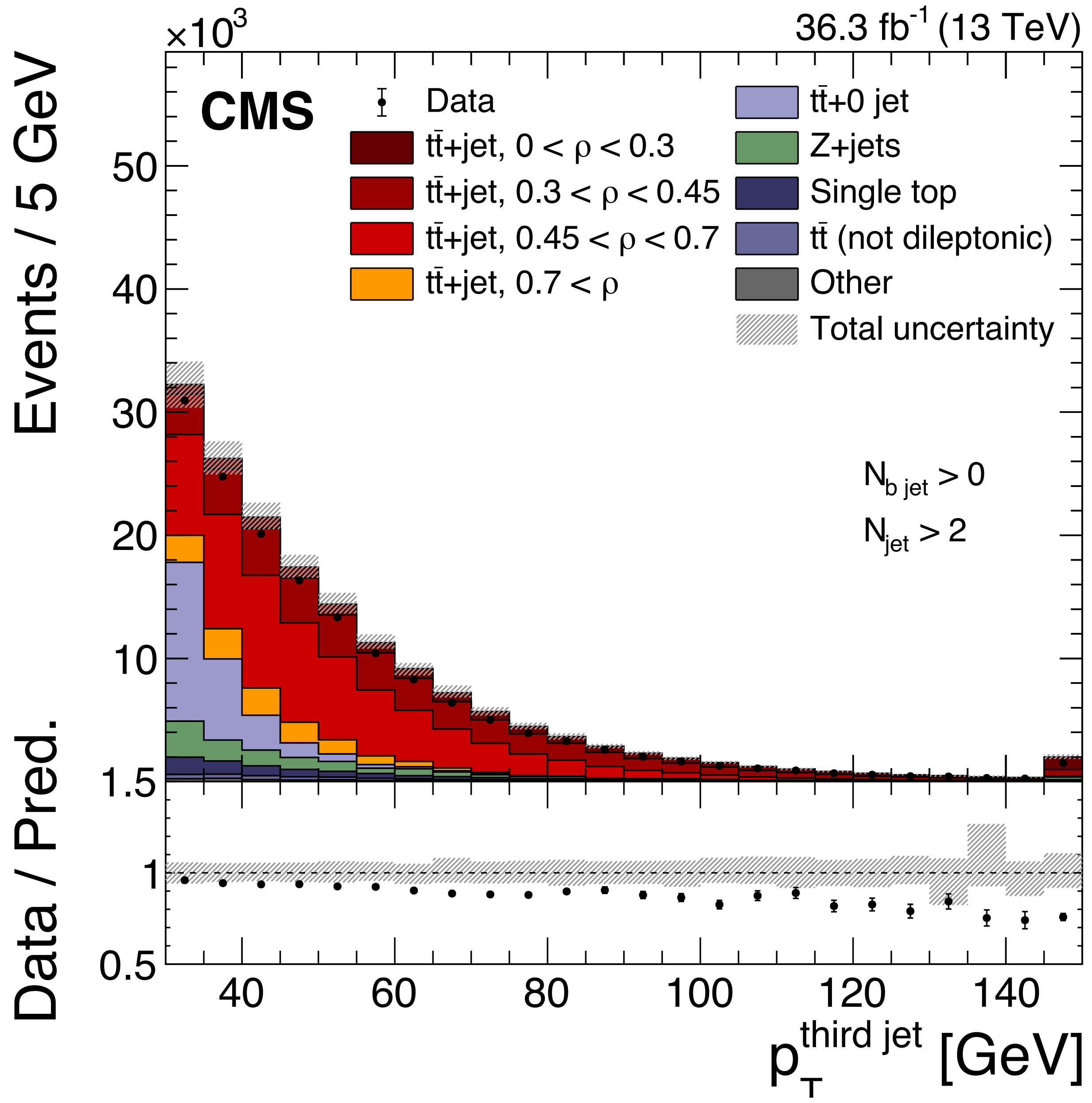
Distributions for $t\bar{t}$+jet - like selection

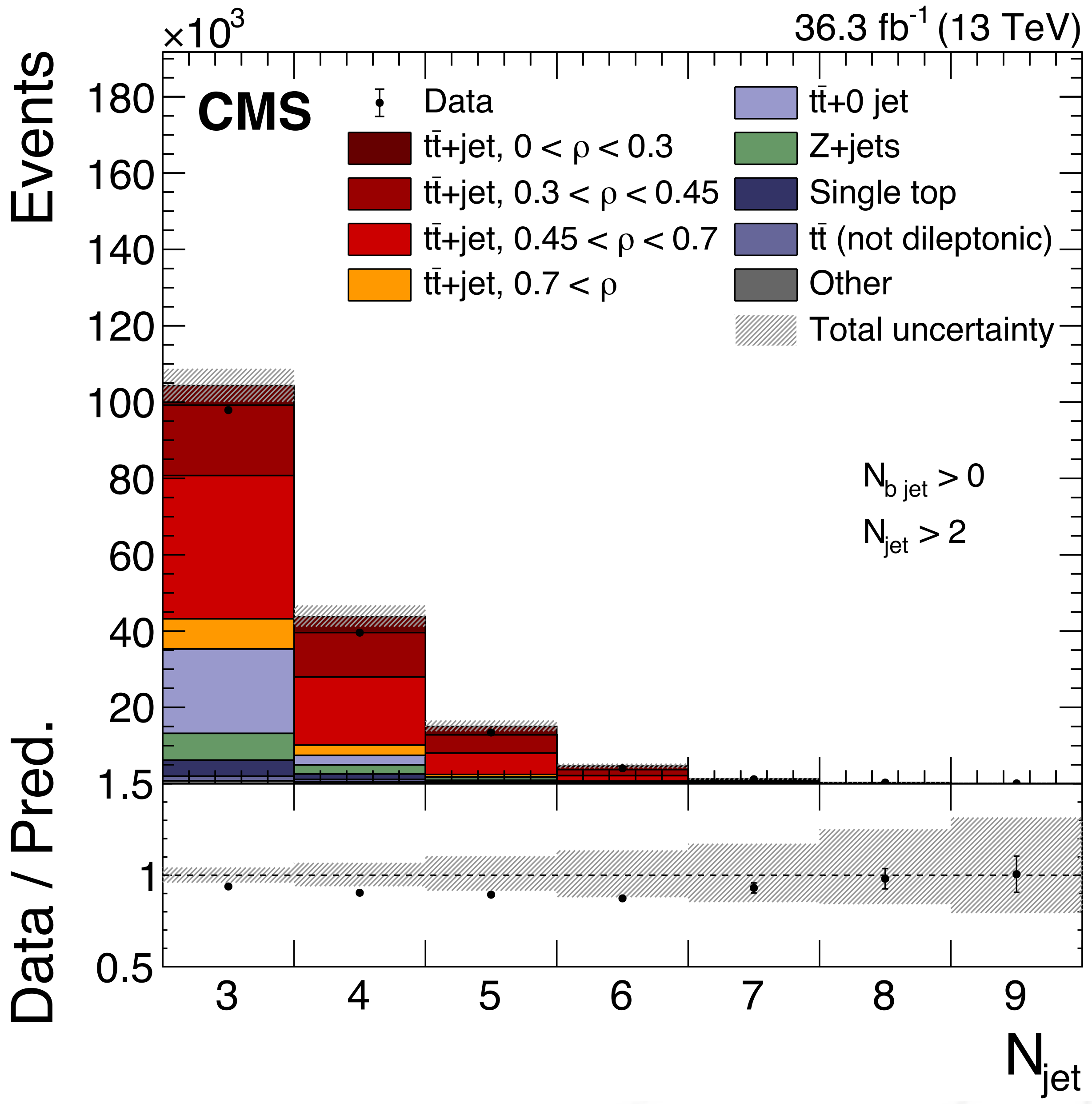
CMS

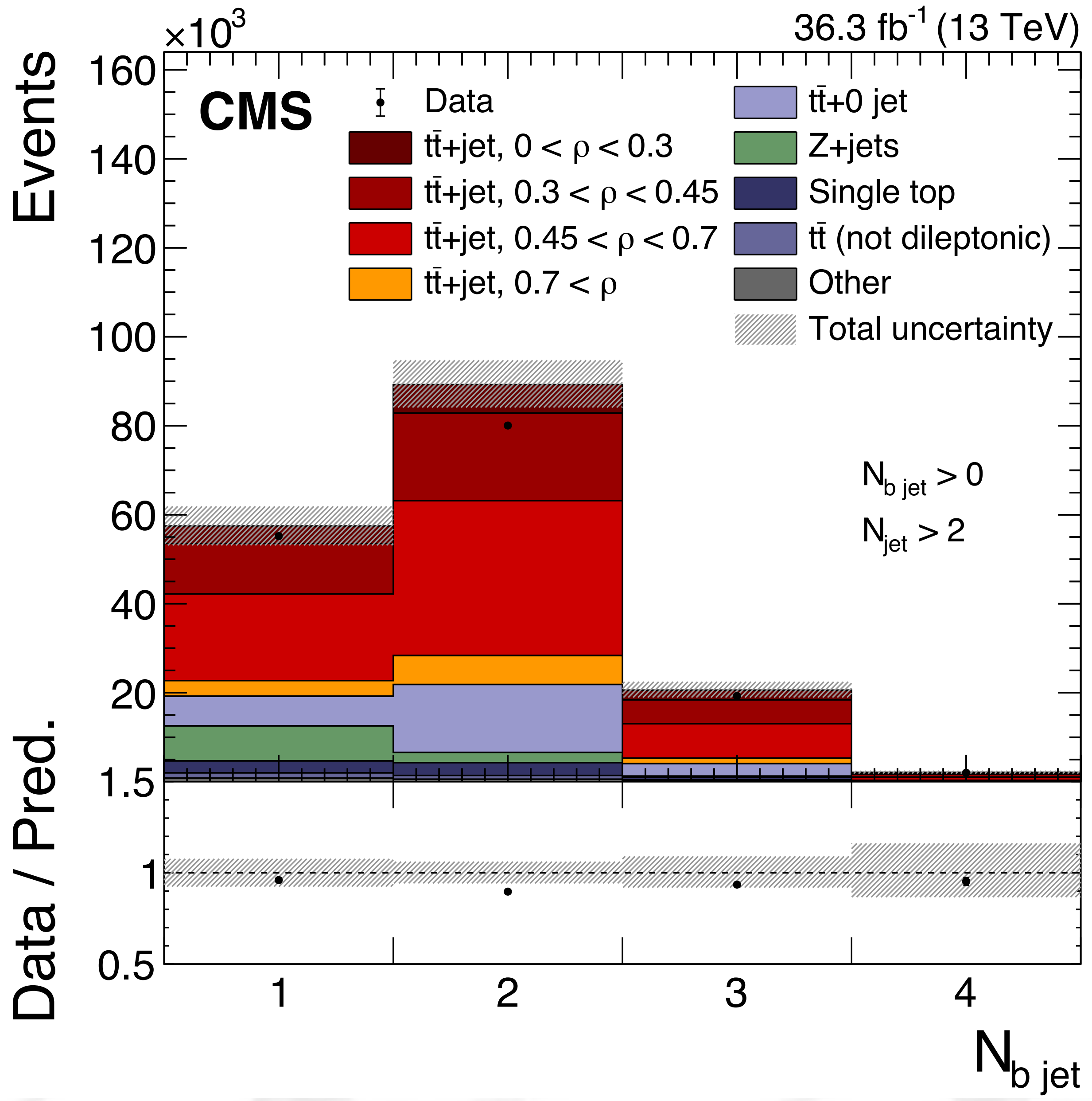






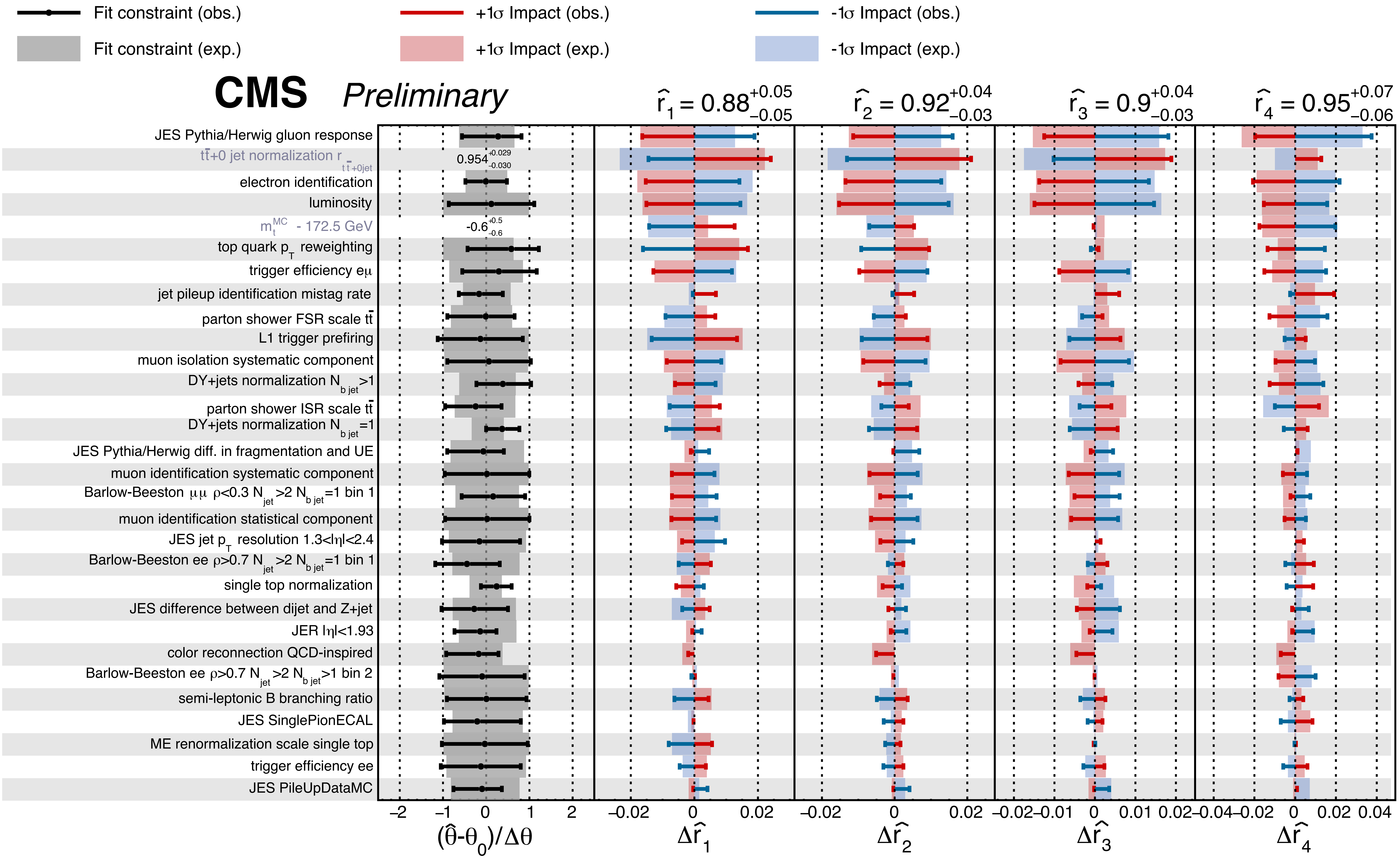




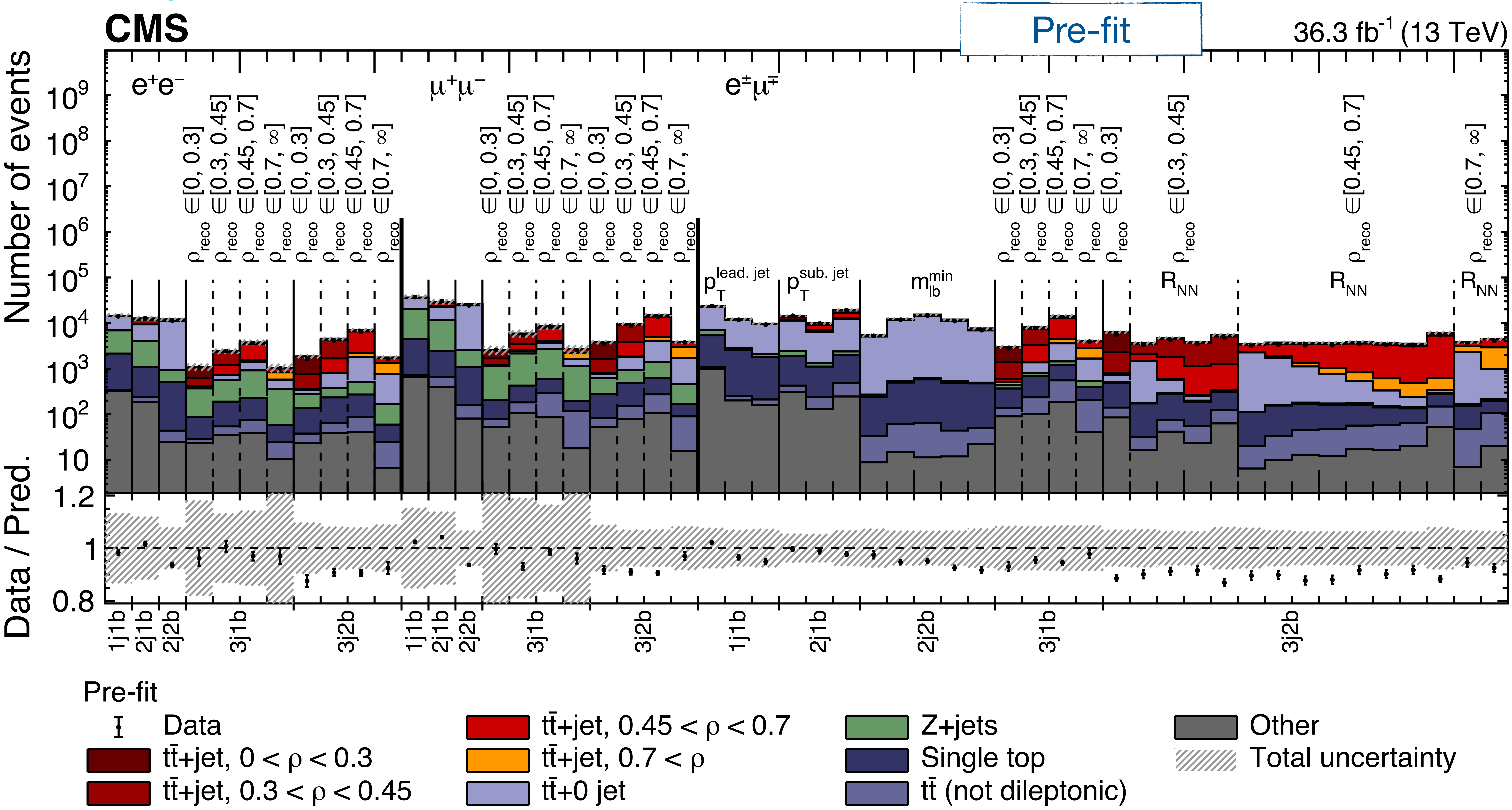


Nuisance paramater impacts

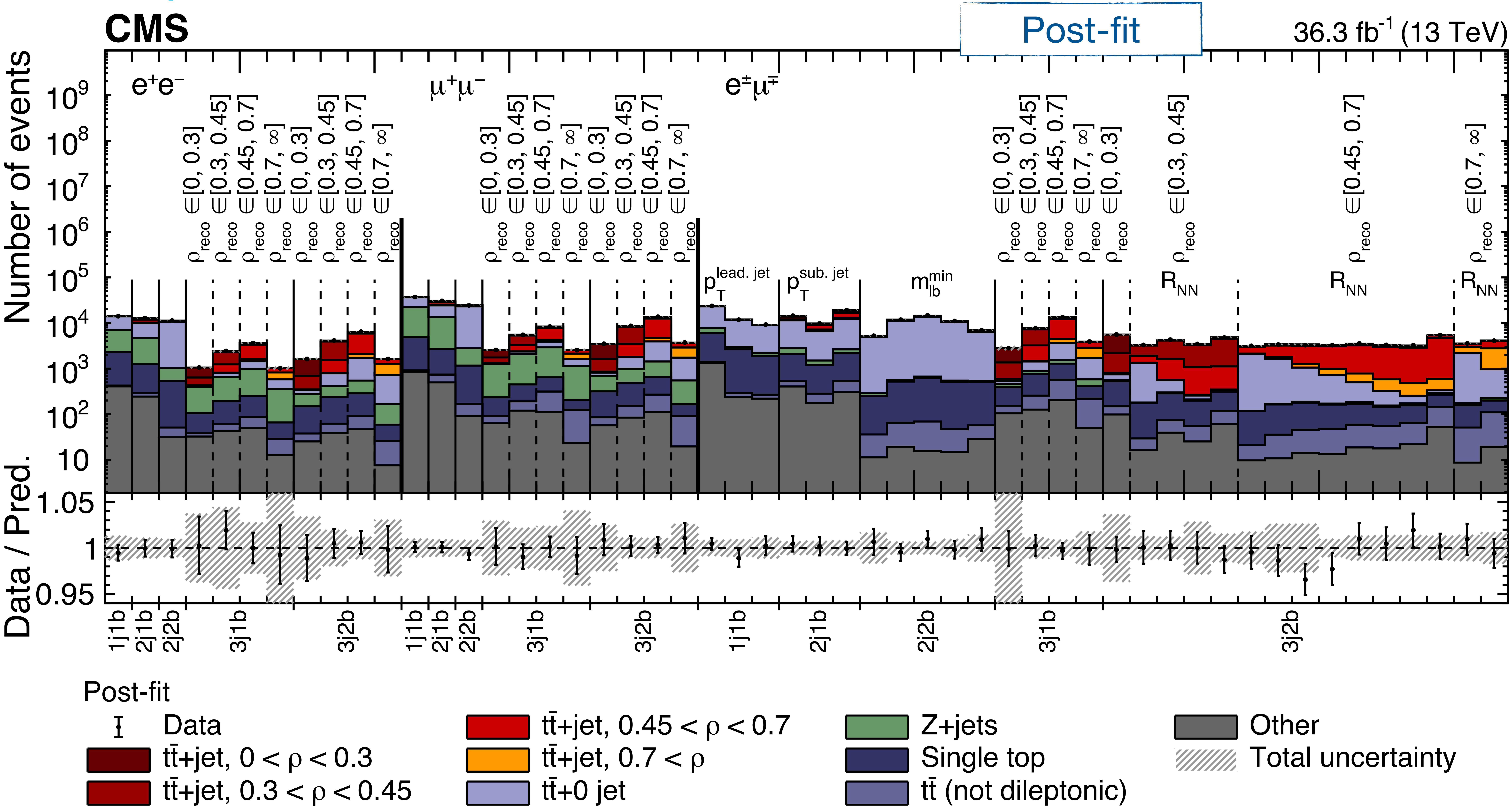
And constraints/pulls



Pre- & post-fit distributions



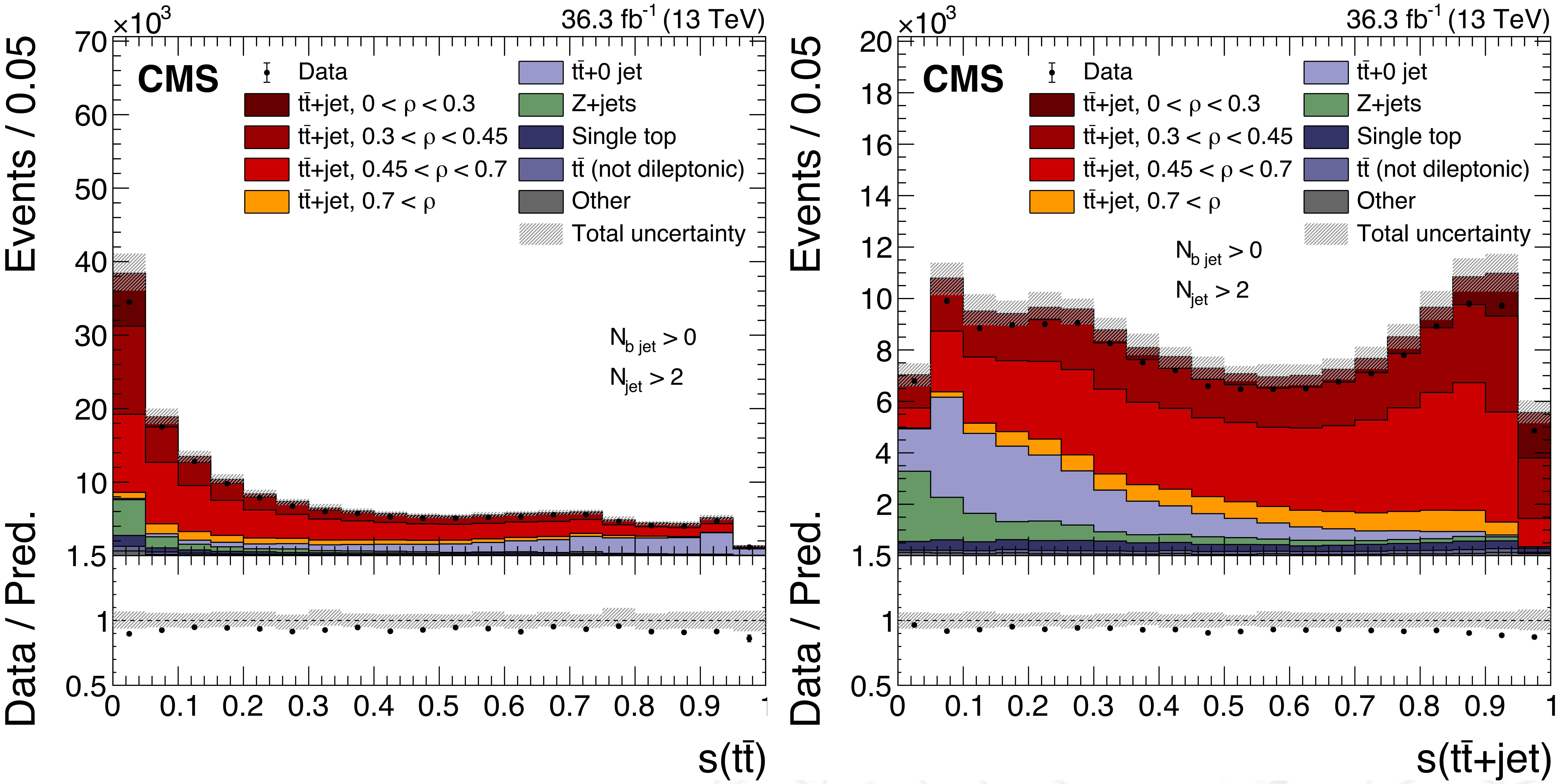
Pre- & post-fit distributions



MVA methods using neural networks

Event classification

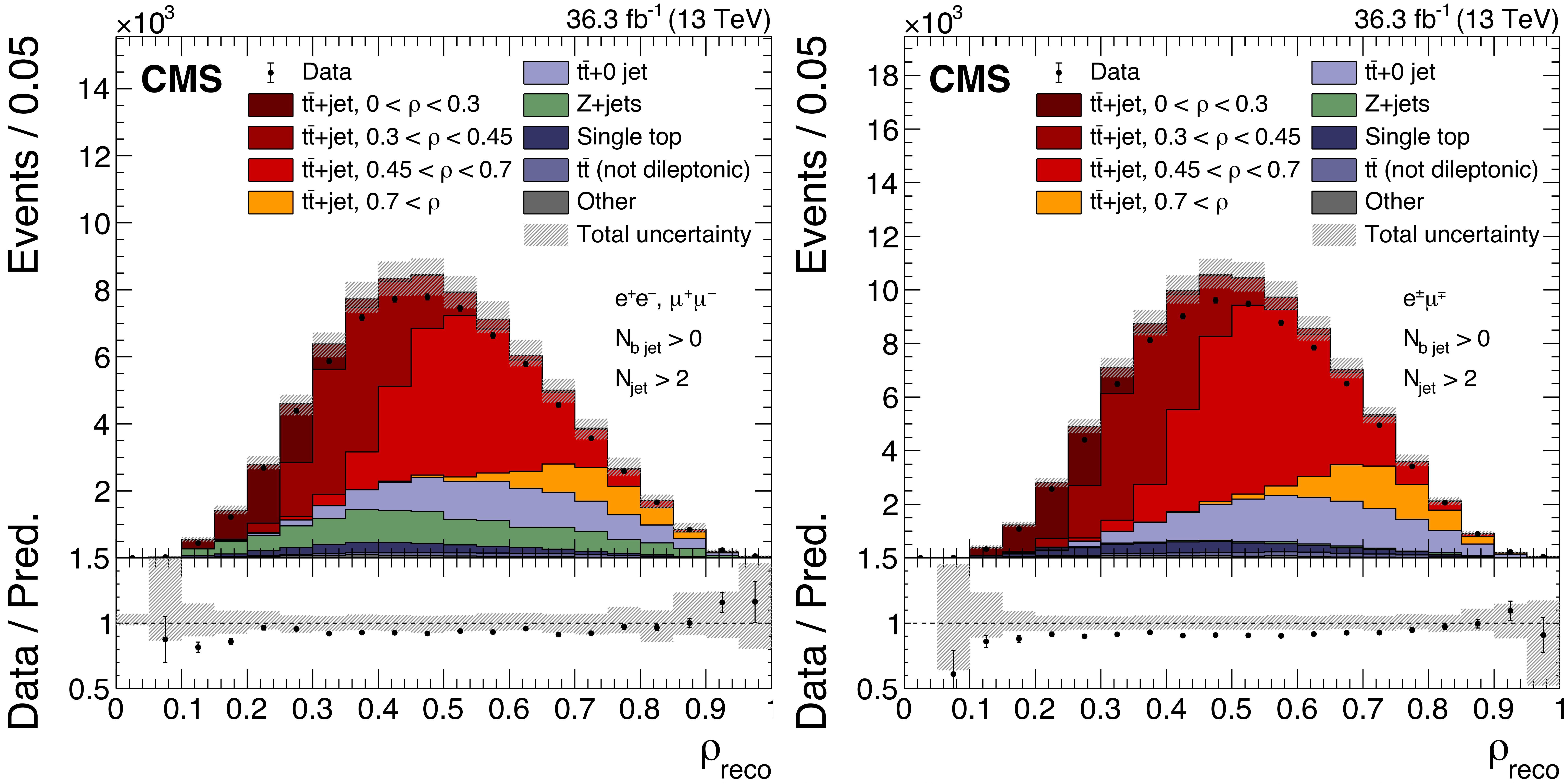
- trained on aMC@NLO simulation
- discriminate between three classes:
  - $t\bar{t}$+jet signal, $t\bar{t}$(+0 jet) background, Z+jets
- decorrelated output score from reconstructed ρ (using Unsupervised Domain Adaptation by Backpropagation)



MVA methods using neural networks

Kinematic reconstruction

- trained on aMC@NLO simulation



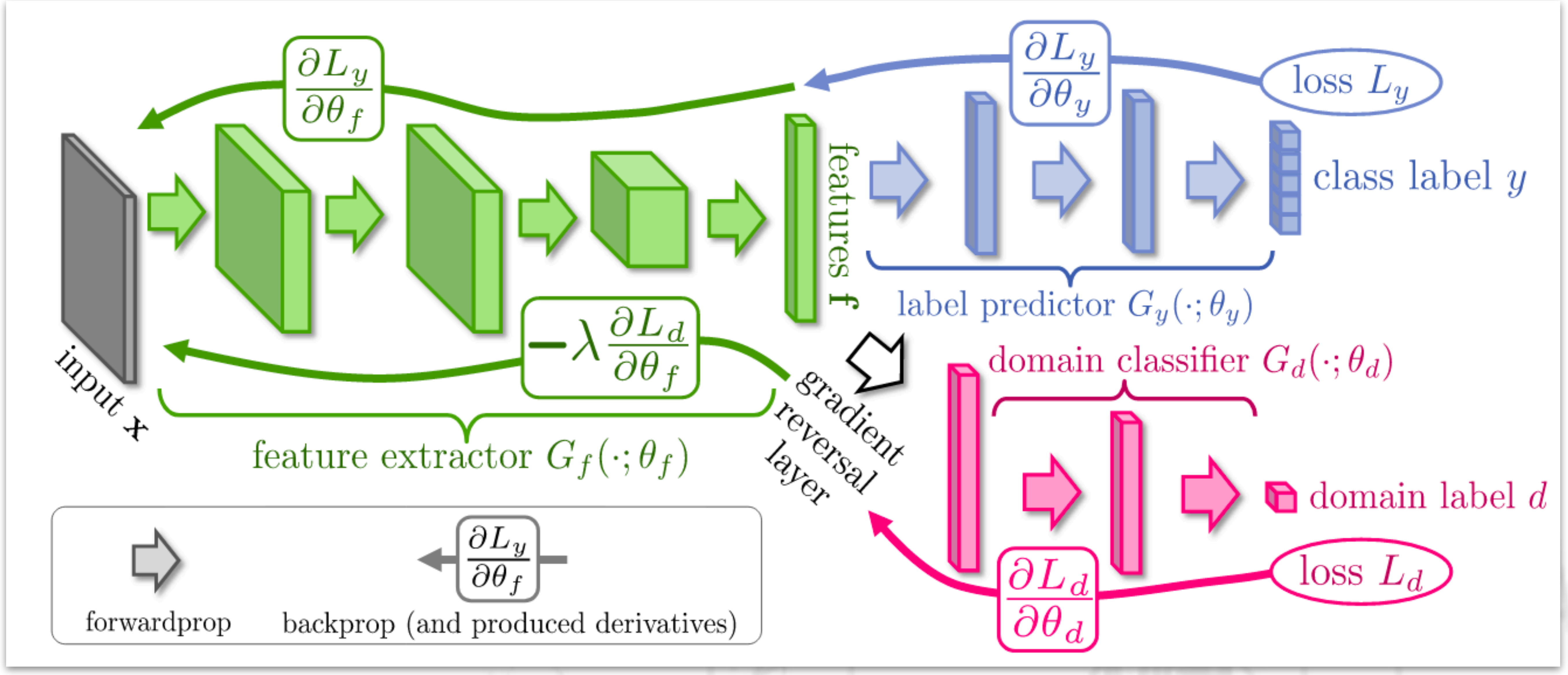
MVA methods using neural networks

Kinematic reconstruction

Unsupervised Domain Adaptation by

Backpropagation:

- decorrelation of the NN output from different variables, classes, training sets..
- inserting additional layers and output node targeting to regress ρ
    - add layer in front a layer that reverts the gradient
- classification performance not degraded



see arXiv:1409.7495

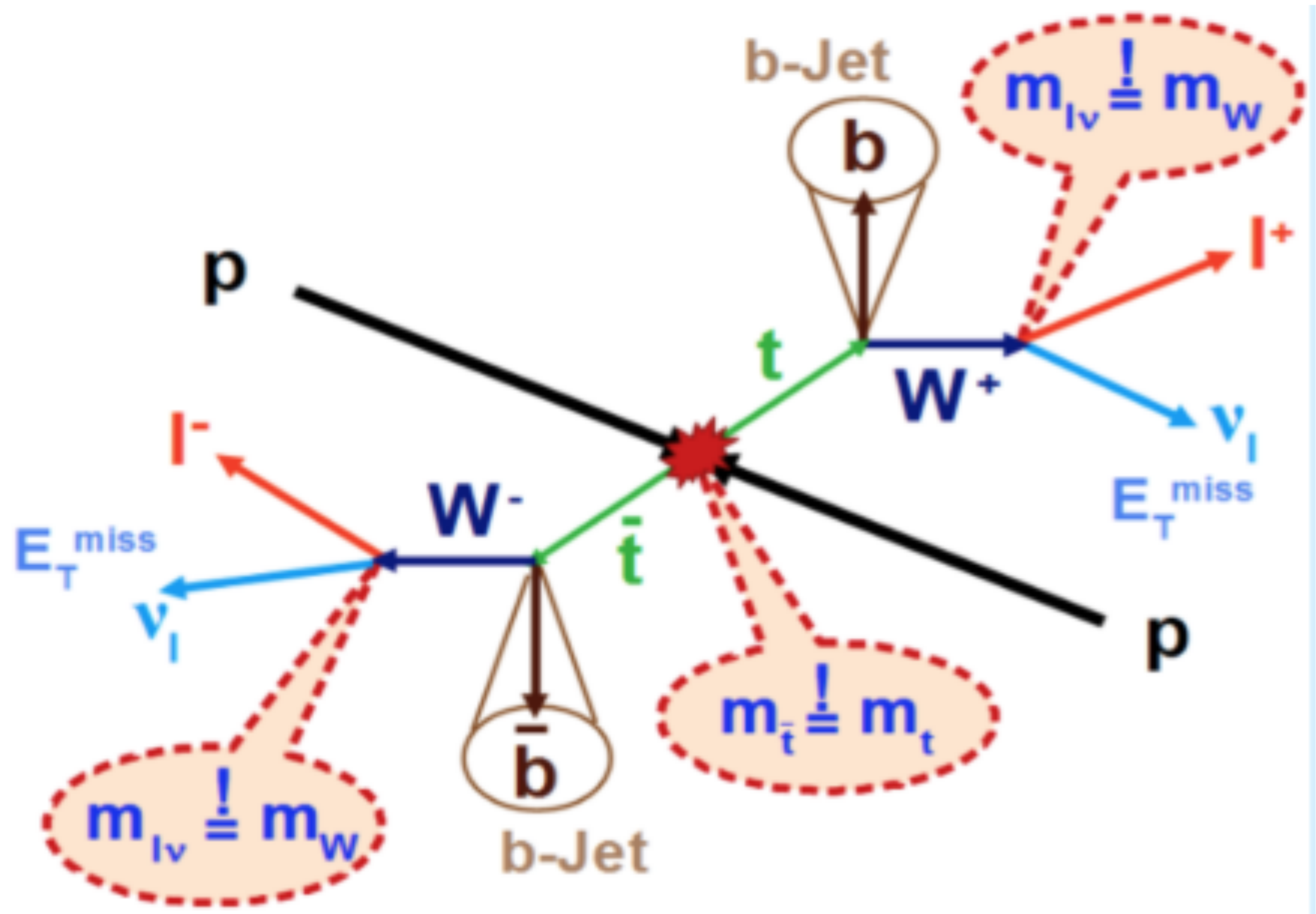
Kinematic Reconstruction of the $t\bar{t}$-system arXiv:1811.06625

- inputs:
  - 2 jets
  - 2 leptons
  - MET

- constraints:
  - m_t, $m_{\bar{t}}$ =172.5 GeV
  - m_{W+}, m_{W-} = 80.4 GeV
  - $p_T(\nu,\bar{\nu})$ = MET

- unknowns:
  - 3-momenta: ν, $\bar{\nu}$

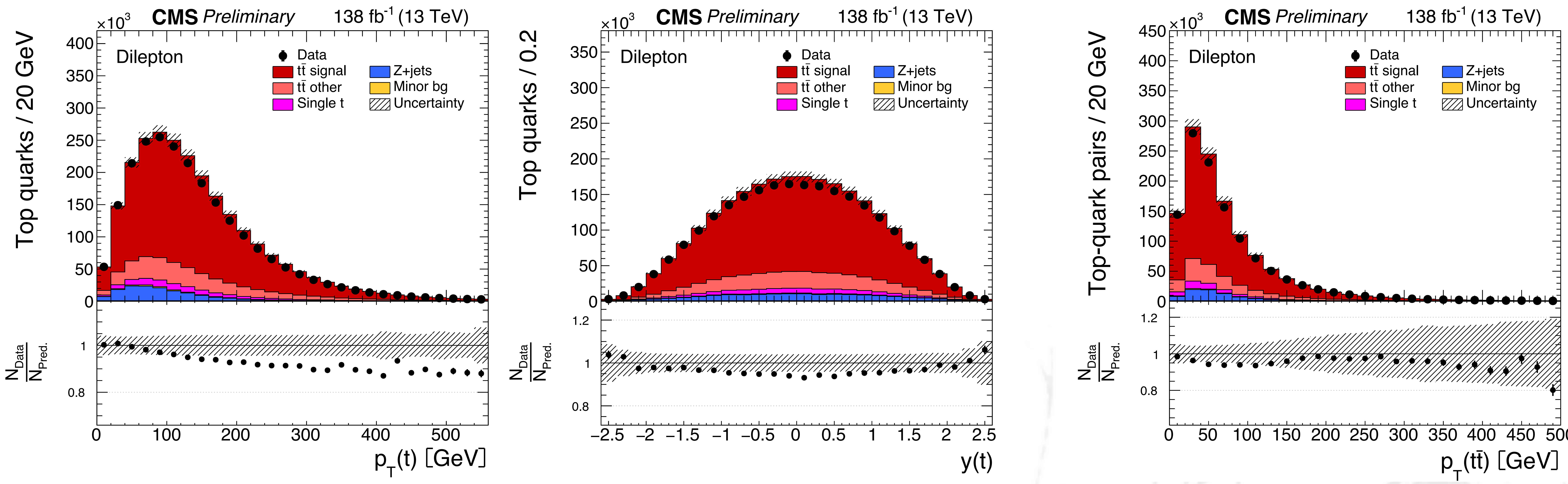
- solutions with b-tagged jets are preferred
- reconstruct event 100 times:
  - W mass smeared according to Breit-Wigner distribution
  - Lepton, b-Jet energies smeared according to detector resolution
  - Weights are calculated based on m_{lb} spectrum
  - Take weighted average as solution
- Efficiency > 90%



$$p^{top}_{x,y,z} = \frac{1}{w} \sum_{i=0}^{100} w_i \cdot \left(p^{top}_{x,y,z}\right)_i$$

Full kinematic reconstruction

From CMS-PAS-TOP-20-006



Kinematic Reconstruction of the $t\bar{t}$-system

Another approach

arXiv:1904.05237

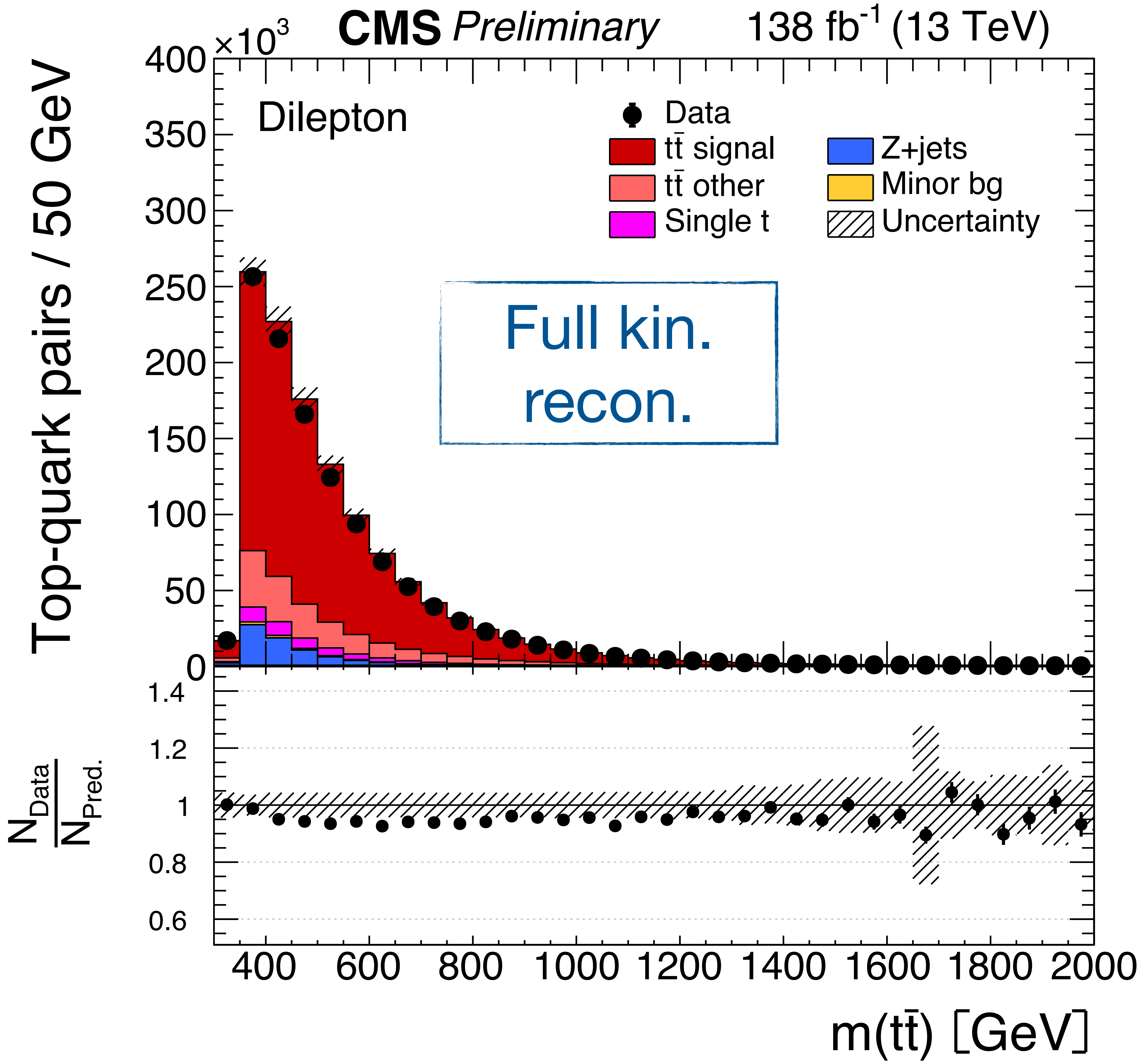
- "Loose kinematic reconstruction":
  - drop m_t requirement
    - no bias on top mass
  - reconstruct $\nu\bar{\nu}$ system as a whole
    - only total $t\bar{t}$-system reconstruction
  - $p_T(\nu\bar{\nu}) = p_T(MET)$
  - $p_z(\nu\bar{\nu}) = p_z(l\bar{l})$, $E(\nu\bar{\nu}) = E(l\bar{l})$
  - requirements: $M(\nu\bar{\nu}) \geq 0$, $M(W^+W^-) \geq 2m_W$

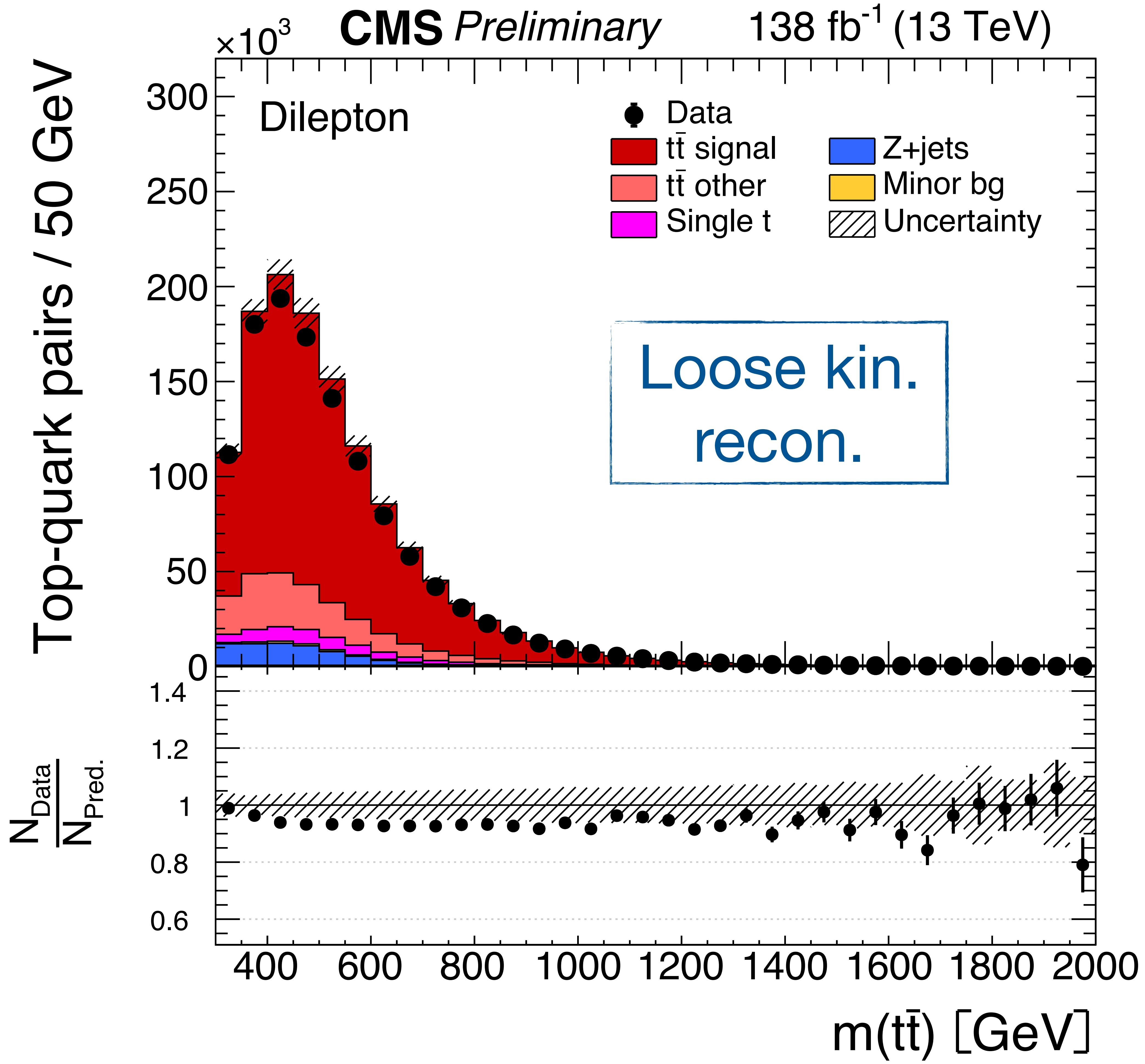
  - prefer solutions with higher b-tagged jet multiplicity over others

  - Obtained kinematics similar to full kinematic reconstruction
  - Efficiency similar

Full vs loose kinematic reconstruction

From CMS-PAS-TOP-20-006





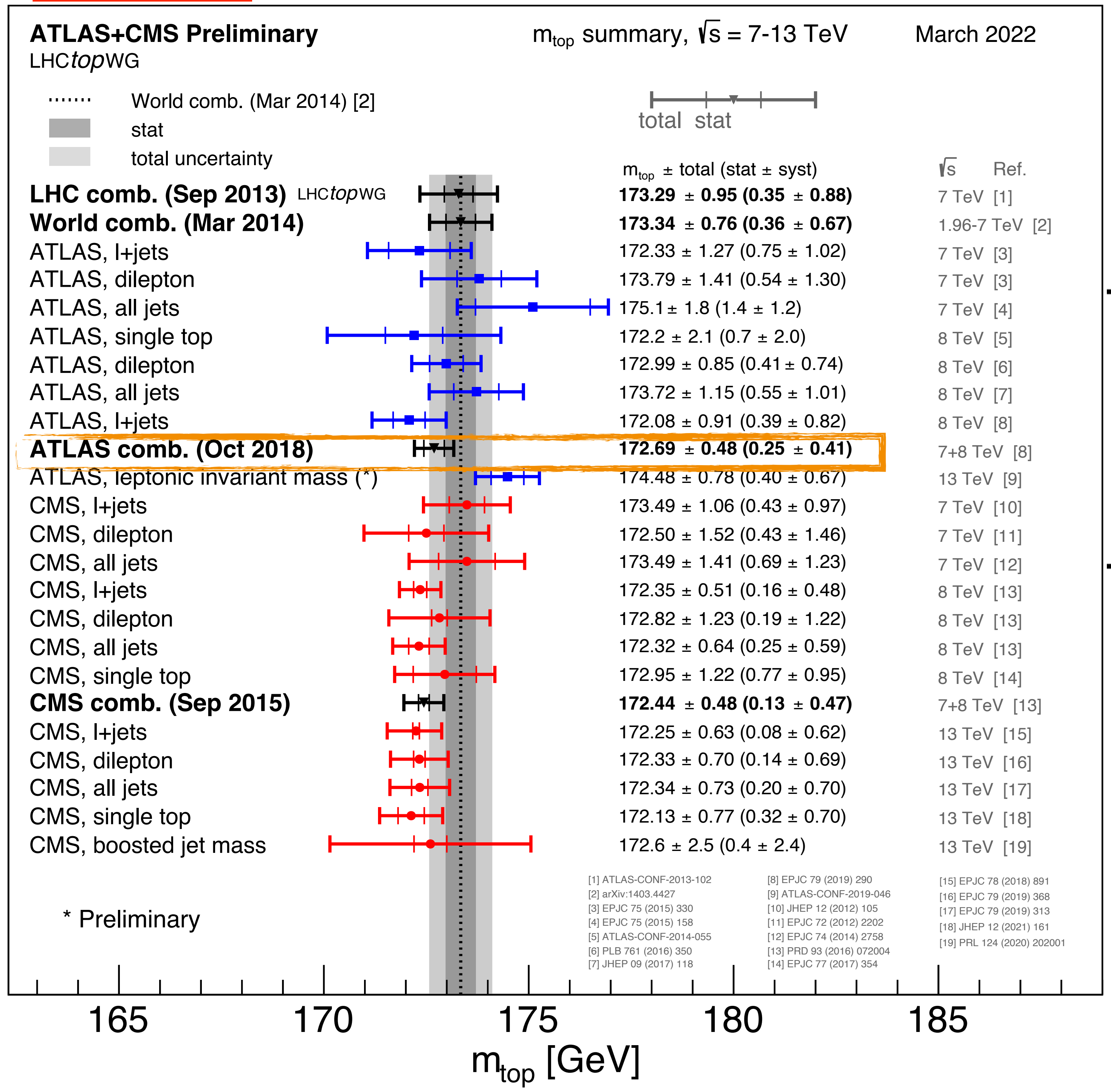
Top quark mass

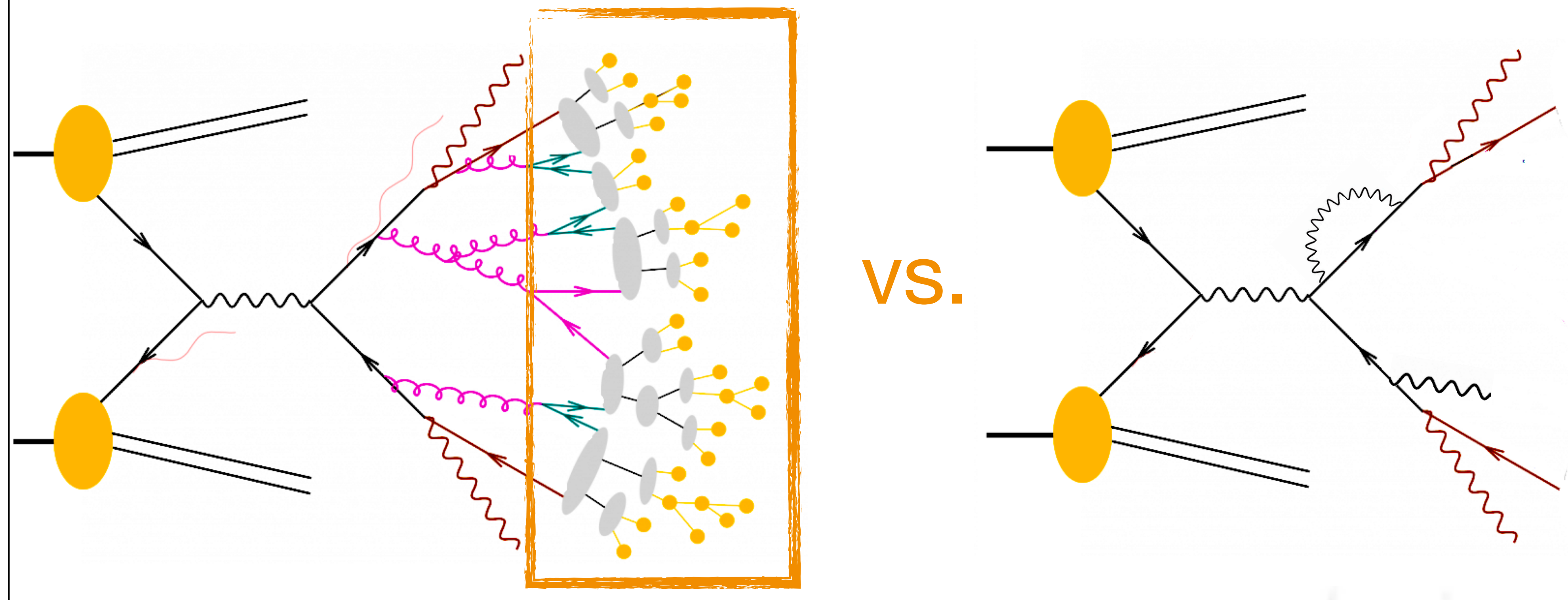
direct measurements m_t^{MC} ⟷ m_t indirect measurements

- measuring m_t^{MC} using reconstructed decay products
  - very high experimental precision
    - ~0.5 GeV
  - relies on details of MC simulation

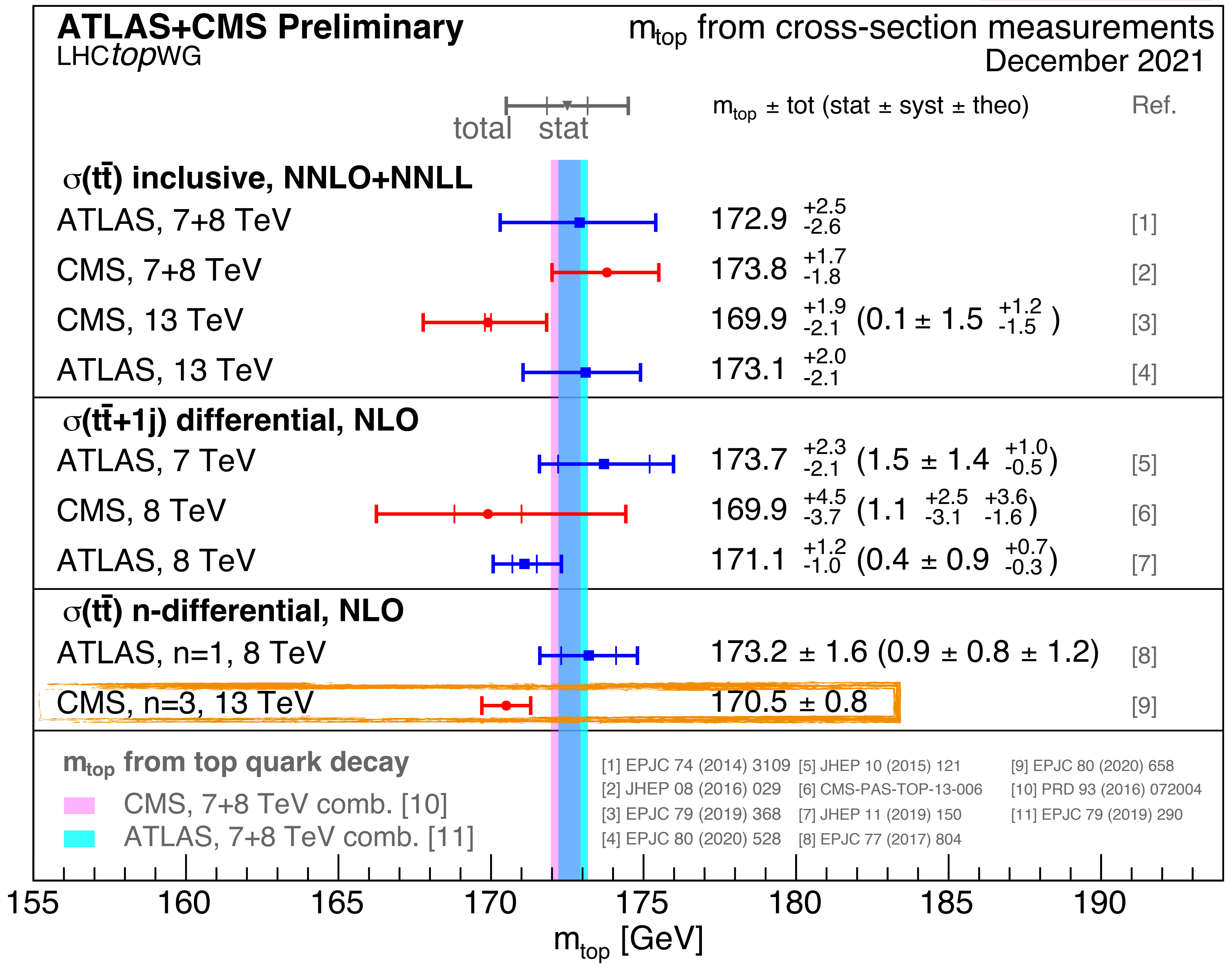
- extract m_t in well defined renormalisation scheme (pole, $\overline{MS}$, …)
- measuring cross section with direct sensitivity to m_t
  - either inclusive or differential

Reference





Reference



$$m_t^{MC} = m_t^{pole} \pm \Delta_{MC}\ O(1GeV)$$