A profile likelihood approach to measure the top quark mass in lepton+jets in CMS

LHC Top Working Group Open Meeting

GEFÖRDERT VOM
Bundesministerium
für Bildung
und Forschung

CLUSTER OF EXCELLENCE
QUANTUM UNIVERSE

Universität Hamburg
DER FORSCHUNG | DER LEHRE | DER BILDUNG

Christoph Garbers, Peter Schleper, Hartmut Stadie

University of Hamburg

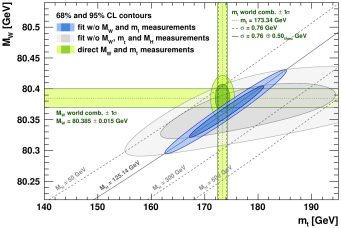
June 16, 2022

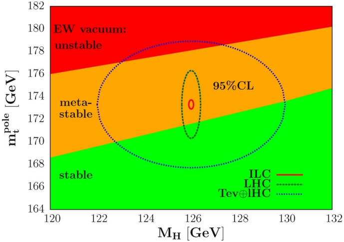
Top quark mass

The top quark mass (m_t) is a prominent input for SM consistency checks. Top quark loops contribute to the Higgs boson mass and top-bottom quark loops to the W boson mass.

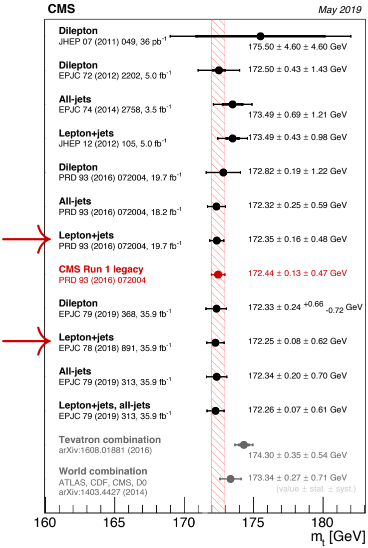
Direct measurements of m_t are among the most precise measurements of the LHC experiments.

- Categorized by the $t\bar{t}$ decay signature
- Templates derived from sim. are fit to the invariant mass distribution
- Invariant mass of the W boson is used to correct the jet energy (in addition to the centrally provided jet energy corrections (JEC))





The Top Quark $t\bar{t} \to$ l+jets channel



$t\bar{t} \to \ell +$jets channel was used in the most precise m_{t} measurement (PRD 93, 0720)

Useful channel for precision measurements due to

- branching ratio
- easy to trigger
- only one ν

The first analysis on data at $\sqrt{s} = 13\,\mathrm{TeV}$ (EPJC-78-891) could not surpass it with the same analysis approach.

This top quark mass measurement

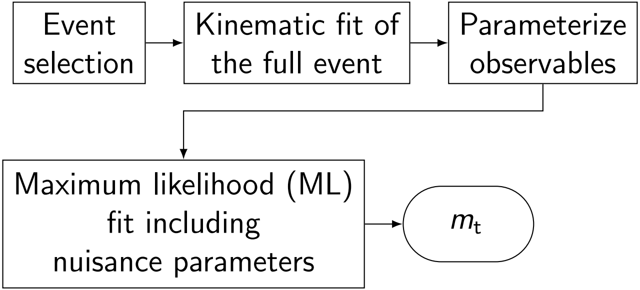
Analysis with reconstruction of data from 2016 also used in EPJC-78-891.

The biggest uncertainties were due to jet energy correction and, color reconnection modelling

Expect even more accurate uncertainty than the former analyses due to:

- Legacy data reconstruction
- CP5 UE tune
- More events in simulation variation samples
- More systematic variations via event based weights
- Use more observables
- Include all sources of uncertainty as nuisance parameters in the likelihood

Structure of the analysis



Event selection

Kinematic fit of the full event

Parameterize observables

Maximum likelihood (ML) fit including nuisance parameters

m_t

Event Selection

	EPJC-78-891	this analysis
JSON	13TeV Collision 16 36 fb^{-1}	
	23Sep2016ReReco	07Aug2017
signal MC	TT powheg-pythia8	
	MiniAODv2 80X CUETP8M2T4 tune	MiniAODv3 94X CP5 tune

1 lepton + 4 jets

- HLT: isolated muon (electron) with $p_T > 24(27)$ GeV
- Muon (electron) selection: $p_T > 26(29)$ GeV and $|\eta| < 2.4$
- Veto on events with additional leptons
- Four anti-$\mathrm{k}_\mathrm{t}^{R=0.4}$ jets with $p_T > 30$ GeV, $|\eta| < 2.4$, $\Delta R(\text{muon,jet}) > 0.3$
- b-tagging: *DeepJet* (1% mis-tag, 78% efficiency)
  - At least two b-tags in selected jets

  difference to EPJC-78-891: DeepJet instead of CSVv2 ($\epsilon_{\text{bTag WP medium}}$: 70%→78%)
- Fit event kinematics to $t\bar{t}$-hypothesis, cut on $P_{\text{gof}} \geq 0.2$ (for most observables)

events in $P_{\text{gof}} \geq 0.2$ selection:

# μ+jets events:	140 362
# e+jets events:	87 265

That are about 40% more events than selected in EPJC-78-891, in addition a observable for 511 833 events with $P_{\text{gof}} < 0.2$ is included.

Kinematic reconstruction

Fit the event kinematics to a $\mathrm{t\bar{t}}$ hypothesis

Input: p_T and angles of the jets and lepton and missing p_T

Constraints:

- $m^{\text{fit}}_{t_{\text{hadr}}} = m^{\text{fit}}_{t_{\text{lept}}}$
- $m^{\text{fit}}_{W} = 80.4\,\text{GeV}$
- p_T balance

use goodness-of-fit

$P_{\text{gof}} = \exp(-\frac{1}{2}\chi^2)$

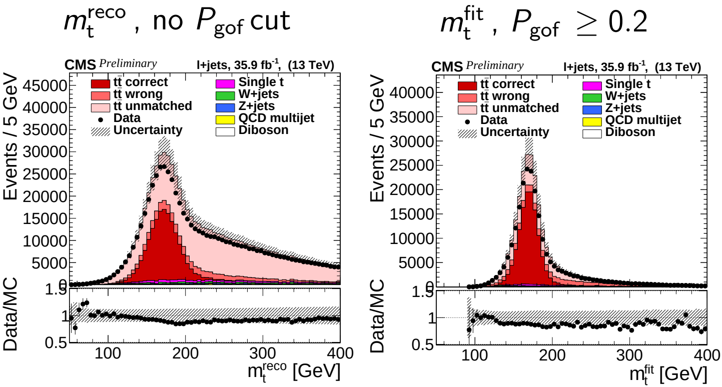
cut on $P_{\text{gof}} \geq 0.2$

Jet-parton assignments:

- Correct assignment
- Wrong: wrong jets-parton assignment, e.g. bottom quark jets switched
- Unmatched: matching to other jets / matching did not work

	no P_{gof} cut	$P_{\text{gof}} \geq 0.2$
$\mathrm{t\bar{t}}$ correct	20 %	47 %
$\mathrm{t\bar{t}}$ wrong	8 %	16 %
$\mathrm{t\bar{t}}$ unmatched	72 %	37 %
signal fraction	91 %	95 %
#events	739 460	227 627

Effect of the kinematic fit and P_{gof} sel.



m_t^{reco}, no P_{gof} cut

m_t^{fit}, $P_{\mathrm{gof}} \geq 0.2$

CMS-PAS-TOP-20-008 Fig.1

The kinematic fit and $P_{\mathrm{gof}} \geq 0.2$ selection improves:

- Jet-parton assignment
- Signal fraction
- Resolution of the invariant top quark mass distribution

Mass extraction method

A negative log-likelihood is minimized to extract m_{t} .

- Likelihood depends on m_{t} and nuisance parameters $\vec{\theta}$
- $\vec{\theta}$ incorporates the sysematic. uncertainty, ± 1 correspond to $\pm 1\sigma$ variations

The likelihood is $\lambda(m_{\mathrm{t}}, \vec{\theta}|\mathrm{data}) = P(\mathrm{data}|m_{\mathrm{t}}, \vec{\theta}) \cdot P(\vec{\theta})$

$P(\mathrm{data}|m_{\mathrm{t}}, \vec{\theta})$ is the density function for the probability to observe the data for given m_{t} and $\vec{\theta}$ values derived from simulation.
$P(\vec{\theta})$ encodes the prior knowledge on the values of the nuisance parameters.

The used templates P are presented on the following slides.

These templates are derived on distributions that are highly dependent on m_{t} or promise to hone in on some of the uncertainties.

Five independent observables are chosen.

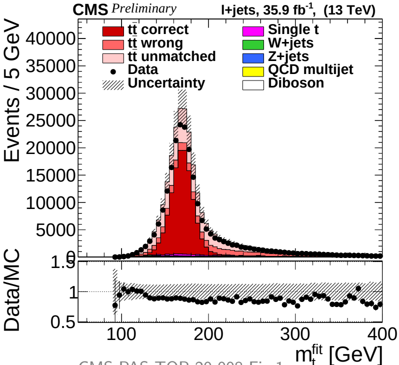
Observable $m_{\mathrm{t}}^{\mathrm{fit}}$

$m_{\mathrm{t}}^{\mathrm{fit}}$ is the main observable, used for events with $P_{\mathrm{gof}} \geq 0.2$

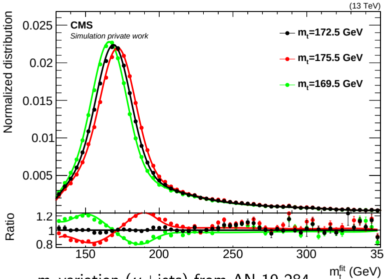
Template function:

$$P\left(m_{\mathrm{t}}^{\mathrm{fit}}\right) = f_{\mathrm{sig}} V\left(m_{\mathrm{t}}^{\mathrm{fit}} \,|\mu, \sigma\right) + \sum_{n=0}^{4} p_n T_n\left(m_{\mathrm{t}}^{\mathrm{fit}}\right)$$

with $V\left(x|\mu,\sigma\right) = \int_{-\infty}^{\infty} G\left(x',\sigma\right) L\left(x-x',\mu\right) dx$ and Chebyshev polynomials T_n, up to the order $n = 4$, defined as $T_0(x) = 1$, $T_1(x) = x$, $T_{n+1}(x) = 2xT_n(x) - T_{n-1}(x)$.



CMS-PAS-TOP-20-008 Fig.1



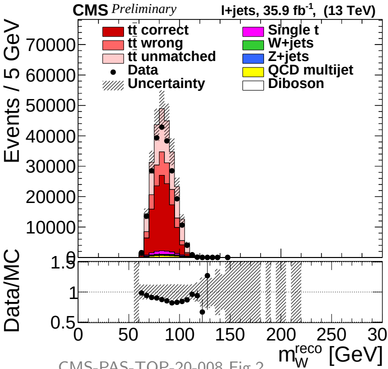
m_{t} variation (μ+jets) from AN-19-284

Observable $m_\mathrm{W}^\mathrm{reco}$

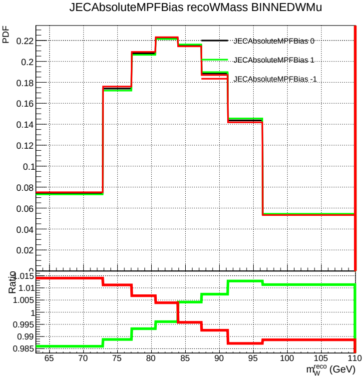
$m_\mathrm{W}^\mathrm{reco}$ gives a good handle on how much the jet energy and resolution needs to be corrected (in addition to standard CMS JEC), its inclusion reduces the JEC uncertainties

As template 8 bins of equal integral are used. This distribution is used for all following observables. It has 7 d.o.f. due to normalisation.

$m_\mathrm{W}^\mathrm{reco}$ is used for events with $P_\mathrm{gof} \geq 0.2$



CMS-PAS-TOP-20-008 Fig.2



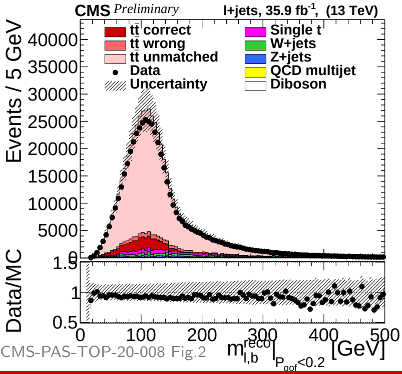
abs. MPF bias JEC variation (μ+jets) from AN-19-284 appendix

Observables $m_{\ell\mathrm{b}}^{\mathrm{reco}}\big|_{P_{\mathrm{gof}}<0.2}$

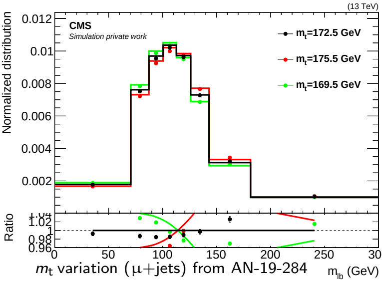
$m_{\ell\mathrm{b}}^{\mathrm{reco}} = \sqrt{\left(P_{lepton}^{reco} + P_{b}^{reco}\right)^2}$ is used in the $t\bar{t} \rightarrow$ 2l+jets mass measurement

The b$_{\mathrm{lep}}$ assignment in events cut by the P_{gof} selection is in most cases still correct and $m_{\ell\mathrm{b}}^{\mathrm{reco}}\big|_{P_{\mathrm{gof}}<0.2}$ includes information and dependencies not included in former analyses.

As template 8 bins of equal integral are used.



CMS-PAS-TOP-20-008 Fig.2



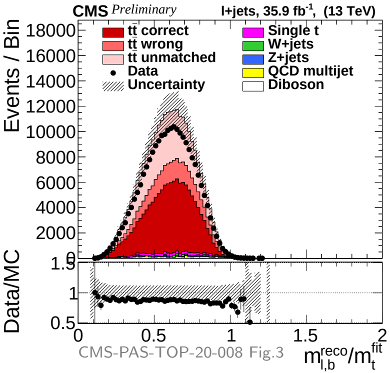
m_t variation (μ+jets) from AN-19-284

Observables $m_{\ell b}^{\mathrm{reco}} / m_{\mathrm{t}}^{\mathrm{fit}}$

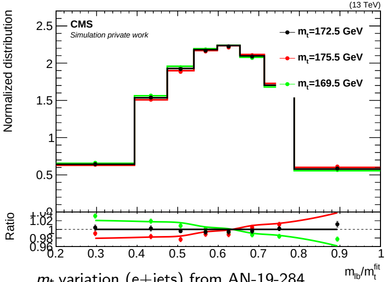
$m_{\ell b}^{\mathrm{reco}} / m_{\mathrm{t}}^{\mathrm{fit}}$ is used as a way to include $m_{\ell b}^{\mathrm{reco}}$ in $P_{\mathrm{gof}} \geq 0.2$ uncorrelated from $m_{\mathrm{t}}^{\mathrm{fit}}$.

This observable is less jet energy sensitive than the other observables

As template 8 bins of equal integral are used. $m_{\ell b}^{\mathrm{reco}} / m_{\mathrm{t}}^{\mathrm{fit}}$ is used for events with $P_{\mathrm{gof}} \geq 0.2$



CMS-PAS-TOP-20-008 Fig.3



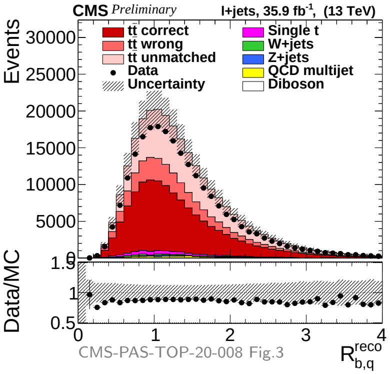
m_{t} variation (e+jets) from AN-19-284

Observables $R_{\mathrm{bq}}^{\mathrm{reco}}$

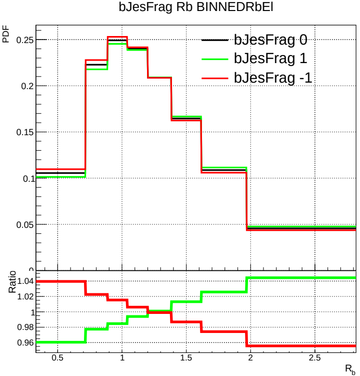
$R_{\mathrm{bq}}^{\mathrm{reco}} = \frac{p_{\mathrm{T_{b1}}}^{\mathrm{reco}} + p_{\mathrm{T_{b2}}}^{\mathrm{reco}}}{p_{\mathrm{T_{q1}}}^{\mathrm{reco}} + p_{\mathrm{T_{q2}}}^{\mathrm{reco}}}$ has been used in the ATLAS m_{t} measurement (EPJC-79-290).

It gives an additional handle on flavor-dependent jet energy scales.

As template 8 bins of equal integral are used. $R_{\mathrm{bq}}^{\mathrm{reco}}$ is used for events with $P_{\mathrm{gof}} \geq 0.2$



CMS-PAS-TOP-20-008 Fig.3



bJES frag. variation (e+jets) from AN-19-284

Nuisance Fit: Parameterisation

Each α_k is a free parameter of a observable template

Parameterize free parameters linear in $\mathbf{m_t}$

$$\alpha_k(m_t) = \alpha_k^0 + s_k^0 (m_t - 172.5\,\mathrm{GeV})$$

The parameters α_k are $\mu, \sigma, f_{sig}, p_n|_{n=0}^4$ for each m_t^{fit} observable and 7 bin heights for each other.

Include all uncertainty sources l as nuisance parameters θ_l in this parameterisation. ($l = 0$ corresponds not to a nuisance parameter but the m_t dependence)

Use factorized approach

$$\alpha_k(m_t, \vec{\theta}) = (\alpha_k^0 + s_k^0 (m_t - 172.5\,\mathrm{GeV})) \prod_l (1 + s_k^l \theta_l)$$

The default simulation corresponds to $\vec{\theta} = \vec{0}$. The θ_l are constrained by $\mathrm{Gaus}(\theta_l|0,1)$ corresponding to a variation of $\pm 1\sigma$.

Nuisance Fit: Slope Uncertainties

To account for the uncertainty on the slopes, corresponding to the statistic limitation of the simulation, additional parameters β_k and $\vec{\omega_k}$ are added.

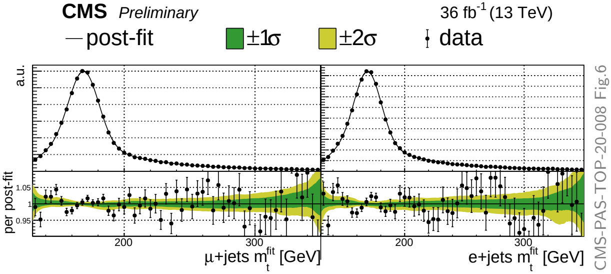
Uncertainty of the slopes

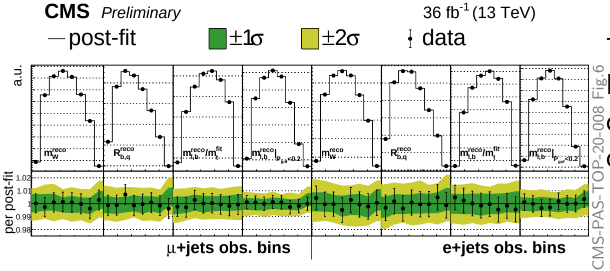
$$\alpha_k(m_{\mathrm{t}}\,,\vec{\theta},\beta_k,\vec{\omega}_k) =$$
$$(\alpha_k^0+\beta_k+s_k^0\,(m_{\mathrm{t}}\,-172.5\mathrm{GeV})+\omega_k^0\cdot 1\mathrm{GeV})\prod_l(1+s_k^l\theta_l+\omega_k^l)$$

β_k's and ω_k's are constrained by multi-dimensional Gaussian functions around zero with their variance equal to the covariance matrices from the slopes

These slope uncertainties are only added for systematic variation from dedicated samples. That are 5 of the 72 nuisance parameters. The other variation are done via event based weights or jet energy variation.

Post-fit distributions





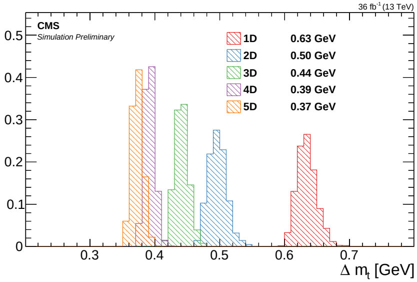
Post-fit probability density functions with error bands, compared to data

The distributions can describe the data well.

Total Predicted Uncertainty

histogram		set label				
observable	category	1D	2D	3D	4D	5D
$m_{\mathrm{t}}^{\mathrm{fit}}$	$P_{\mathrm{gof}} \geq 0.2$	x	x	x	x	x
$m_{\mathrm{W}}^{\mathrm{reco}}$	$P_{\mathrm{gof}} \geq 0.2$		x	x	x	x
$m_{\ell\mathrm{b}}^{\mathrm{reco}}$	$P_{\mathrm{gof}} < 0.2$			x	x	x
$m_{\ell\mathrm{b}}^{\mathrm{reco}} / m_{\mathrm{t}}^{\mathrm{fit}}$	$P_{\mathrm{gof}} \geq 0.2$				x	x
$R_{\mathrm{bq}}^{\mathrm{reco}}$	$P_{\mathrm{gof}} \geq 0.2$					x

CMS-PAS-TOP-20-008 tab.1



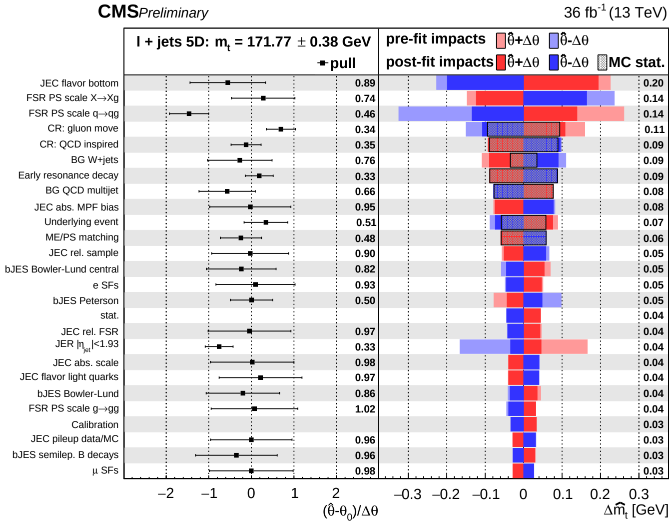
CMS-PAS-TOP-20-008 Fig.4

The total predicted uncertainty is reduced by the inclusion of each additional observables.

The biggest improvement comes from including $m_{\mathrm{W}}^{\mathrm{reco}}$.

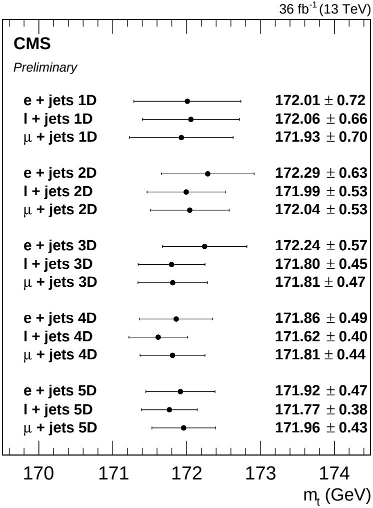
Result and impacts

$$m_{\mathrm{t}}^{\mathrm{MC}} = 171.77 \pm 0.38\,\mathrm{GeV},\ \text{this includes}\ \sigma_{\mathrm{stat}} = 0.04\,\mathrm{GeV}$$



CMS-PAS-TOP-20-008 Fig.7

Results in all categories

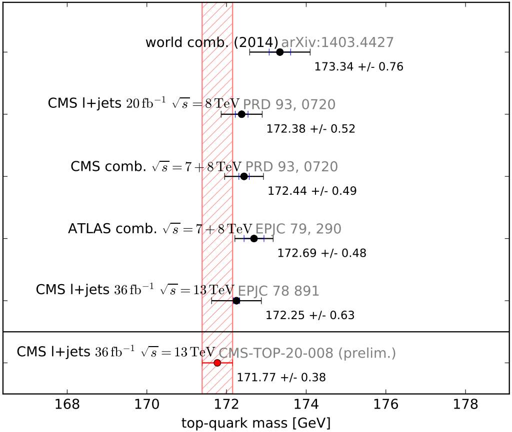


CMS-PAS-TOP-20-008 Fig.5

- *2D e+jets* matches the uncertainty on the former m_t measurement on this data
- *3D μ+jets* surpasses the former most precise m_t measurement.
- The final result is $m_t^{\text{(l+jets 5D)}} = 171.77 \pm 0.38\,\text{GeV}$
- Measured m_t values are consistent with the final result
- *l+jets* results are all outside the *μ+jets* and *e+jets* results (with the same observable setting)
  - Nuisance parameters get measured at higher absolute values when more data is included
  - Constrains on the nuisance parameters alter the reference $t\bar{t}$ simulation, changing the measured value

Compared to former m_{t} measurements

Context among former top quark mass measurement. The red band is this result.



The inclusion of additional observables and nuisance parameters result in a lower value of the measured top quark mass.

The new result is compatible with the prior l+jets measurements.

Summary

- Inclusion of nuisances parameters in the fit helps to hone in on systematic uncertainties on the top quark mass.
- Including $m_{\ell\mathrm{b}}^{\mathrm{reco}}$ for events formerly excluded by the P_{gof} cut, $m_{\ell\mathrm{b}}^{\mathrm{reco}}/m_{\mathrm{t}}^{\mathrm{fit}}$ and $R_{\mathrm{bq}}^{\mathrm{reco}}$ decreases the uncertainty in the direct measurement by additional 150 MeV

The final result is:

$$m_{\mathrm{t}}^{\mathrm{MC}} = 171.77 \pm 0.38\,\mathrm{GeV}\ \left(\frac{\sigma_{m_{\mathrm{t}}}}{m_{\mathrm{t}}} = \pm 0.22\%\right)$$

This includes $\sigma_{\mathrm{stat}} = 0.04\,\mathrm{GeV}$ and $\sigma_{\mathrm{calibration}} = 0.03\,\mathrm{GeV}$

- Its biggest uncertainty source is JEC flavor bottom as in prior analyses.
- Followed by the FSR PS scale uncertainties that is bigger than in former analyses, as their anti-correlated effects are split
- The limit from simulation statistic of variation samples still considerable.

This result surpasses the prior measurement on the same data by 0.25 GeV and is the most precise top quark mass measurement by 0.12 GeV.

Preliminary publication: CMS-PAS-TOP-20-008

Thank you for your attention!

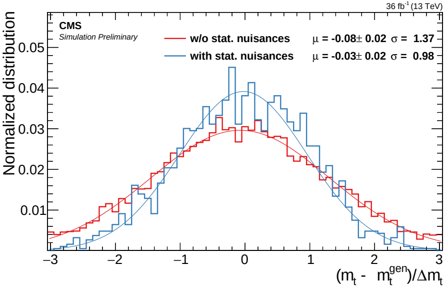
Backup

Predicted uncertainty pull

The chosen parameterization was validated with multitude of checks.

For example the pull distributions from toy studies with and without the additional stat. nuisance parameters were compared.
For each toy, mass and nuisance parameter values are pulled from their prior distributions, the slopes are varied within their uncertainty. From the resulting templates histograms are generated and their bins again varied within their predicted statistical uncertainty.

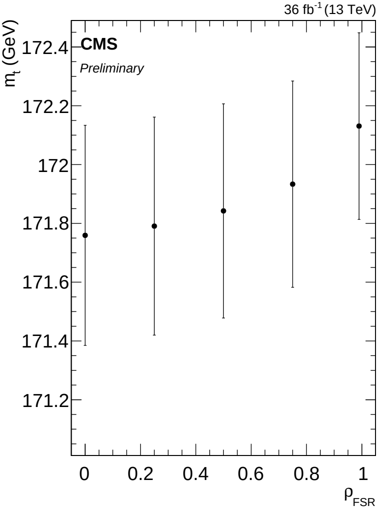
The handling of the simulation statistic works.



CMS-PAS-TOP-20-008 Fig.4

Split FSR correlation dependence

- FSR PS scales are varied independently for different particle types
- In the former analysis this scales were fully correlated from dedicated samples, the split scales from event based weights were not available.
- m_{t} measurement depends on the assumed correlation between the FSR PS scales
- Full correlation is not physical as the splitting happens at different scales for different particle types
- Small m_{t} shift for $\rho_{\mathrm{FSR}} < 0.5$ and the measured scales are not compatible
  → FSR PS scales are used uncorrelated



CMS-PAS-TOP-20-008 Fig.5