

# JRJC 2021

## b-Jet energy calibration with ATLAS Run 2 data using $t\bar{t}$ lepton+jets events

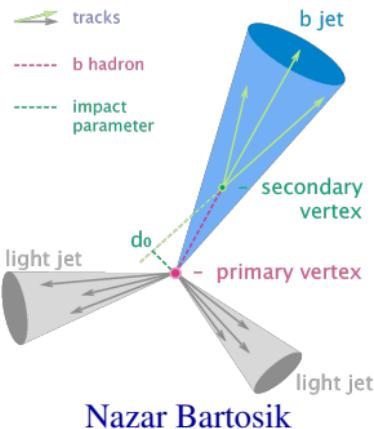
Romain Bouquet

Supervisor:

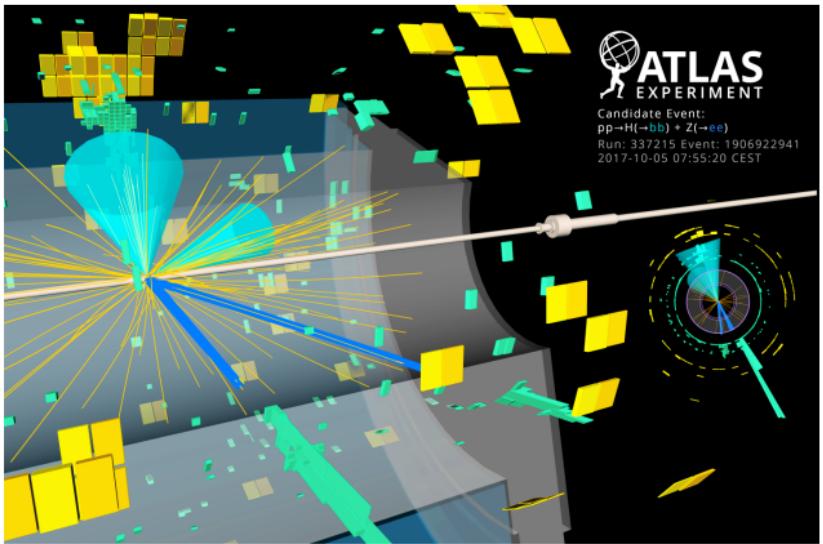
Giovanni Marchiori

Jet = Cone of hadrons and other particles produced by the hadronization of a quark or gluon

**bJES = b-Jet Energy Scale = Correction of b-jet energy applied after data acquisition**  
to compensate the imperfect calibration of the ATLAS detector and jet calibration procedure



Nazar Bartosik



Event display:  $H(\rightarrow b\bar{b}) + Z(\rightarrow e^+e^-)$ , ATLAS Collaboration

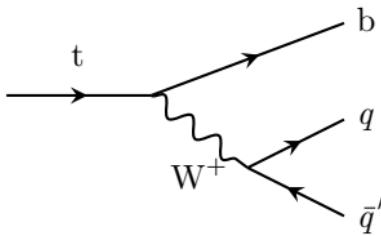
b-jets play a leading role in many analyses

e.g. main decay of the Higgs is  $H \rightarrow b\bar{b}$

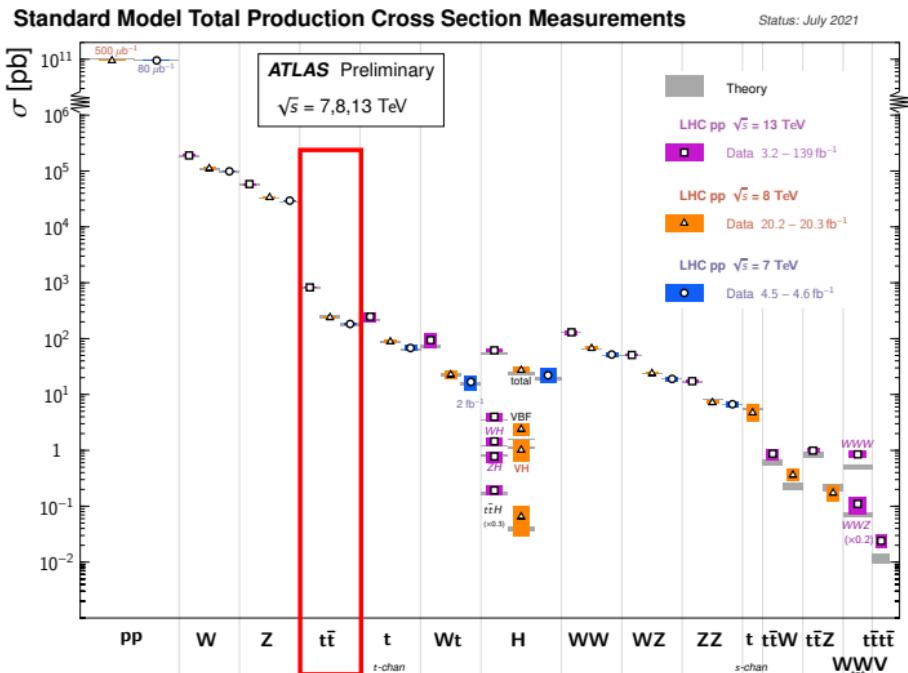
⇒ Need for a dedicated calibration of b-jet energy

# Looking for events with b jets...

LHC = "top factory"  $\rightarrow t\bar{t}$  events  
 (LHC Run 2,  $\sqrt{s} = 13 \text{ TeV}$ ,  $\mathcal{L} = 139 \text{ fb}^{-1}$ )



- $\text{BR}(t \rightarrow Wb) \approx 100\%$
- $\sigma(t\bar{t} \rightarrow X) = 832 \text{ pb}$   
 [Particle Data Group]
- Use  $m_{qb}$  invariant mass to calibrate  $b$ -jet energy

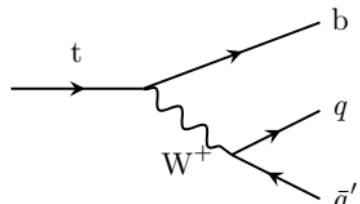


# Energy scale factor $\alpha$ & Template method

➤ Energy scale factor ( $\alpha$ ):  $E'_b = (1 + \alpha)E_b$ ,  $p'_{Tb} = (1 + \alpha)p_{Tb}$

Compare reconstructed  $m_{qqb}$  invariant mass distributions

$t \rightarrow Wb \rightarrow qq'b$  in Data and MC simulation to determine  $\alpha$



➤ Compute  $m_{qqb}$  distribution for Data

➤ Create MC templates  $m_{qqb}$  distribution shifting b-jet energy by:

$(1 + \alpha_i) \rightarrow$  one  $m_{qqb}$  distribution per value  $\alpha_i$

$$E'^{\text{MC}}_b(\alpha_i) = (1 + \alpha_i)E^{\text{MC}}_b, \quad p'^{\text{MC}}_{Tb}(\alpha_i) = (1 + \alpha_i)p^{\text{MC}}_{Tb}$$

➤ Compute  $\chi^2(\alpha_i)$  between MC template and (Pseudo-)Data using  $m_{qqb}$  distributions

→ find minimum  $\alpha_{\min}$  of  $\chi^2(\alpha)$  curve

➤ Determine statistical and systematic uncertainties:  $\delta\alpha^{\text{stat}}$  &  $\delta\alpha^{\text{syst}}$

**NB: MC samples are taken as reference to correct data:  
the final correction will be applied in Data**

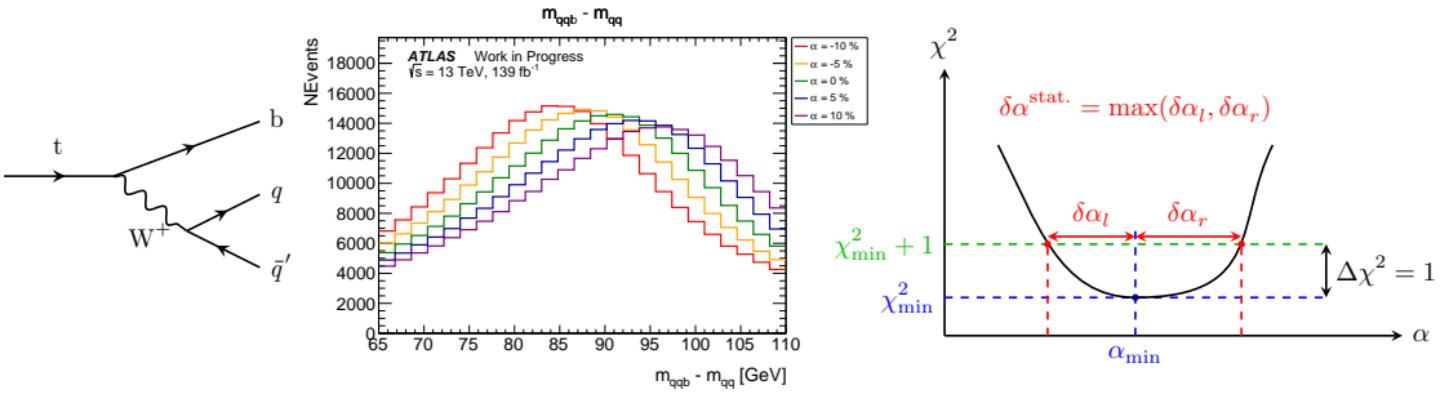
$$E_b^{\text{Data corrected}} = E_b^{\text{Data}} / (1 + \alpha_{\min}), \quad p_{Tb}^{\text{Data corrected}} = p_{Tb}^{\text{Data}} / (1 + \alpha_{\min})$$

**NEW MEASUREMENT:** bJES was previously assumed to be equal to the light jet calibration  
+ uncertainties extracted from MC simulations

bJES measurement using  $t\bar{t}$  lepton+jets was never performed in ATLAS  
⇒ Test procedure using Pseudo-Data then use Data

# MC templates for $m_{qqb} - m_{qq}$ distribution & Statistical error determination $\delta\alpha^{\text{stat}}$

Use  $m_{qqb} - m_{qq}$  instead of  $m_{qqb}$  to reduce sensitivity to light jet calibration  
 Shift of  $b$ -jet energy  $\Rightarrow$  Shift of the  $m_{qqb} - m_{qq}$  distribution

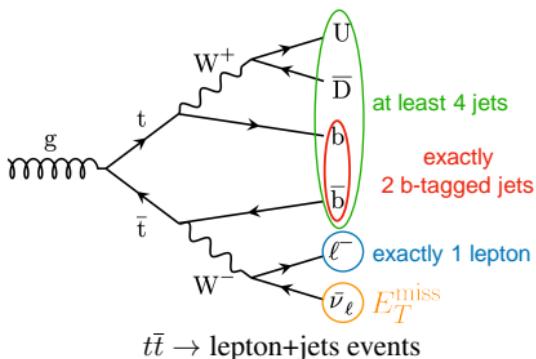


$\chi^2 \rightarrow$  compare shifted MC template and (Pseudo-)Data

$$\chi^2(\alpha_i) = \sum_k \frac{(O_k - E_k(\alpha_i))^2}{O_k} = \sum_k \frac{(N_k^{\text{Data}} - N_k^{\text{MC}}(\alpha_i))^2}{N_k^{\text{Data}}}$$

$$\chi^2_{\min} := \chi^2(\alpha_{\min})$$

To increase precision on  $\alpha_{\min} \pm \delta\alpha^{\text{stat}}$  use penalized spline ( $\approx$  mix between fit and interpolation)  
 with smoothing parameter  $\lambda$

**Selection criteria**

$$\gg p_T^{\text{lepton}} \geq 40 \text{ GeV} \quad \gg E_T^{\text{miss}} \geq 30 \text{ GeV} \quad \gg m_T^W + E_T^{\text{miss}} \geq 60 \text{ GeV}$$

**Exactly 4 jets:**

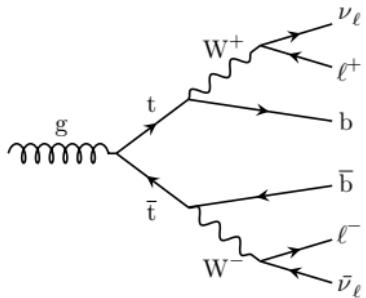
- **Exactly two light jets:**  $q_1$  and  $q_2$
- **Exactly 2 b-tagged jets:** among the 2 b-tagged jets, select the b-jet such that the invariant mass  $m_{q_1 q_2 b, \min}$  is the closest to the top quark mass:

$$\left| m_t - m_{q_1 q_2 b, \min} \right| = \min_{b_k, k=1,2} \left| m_t - m_{q_1 q_2 b_k} \right|$$

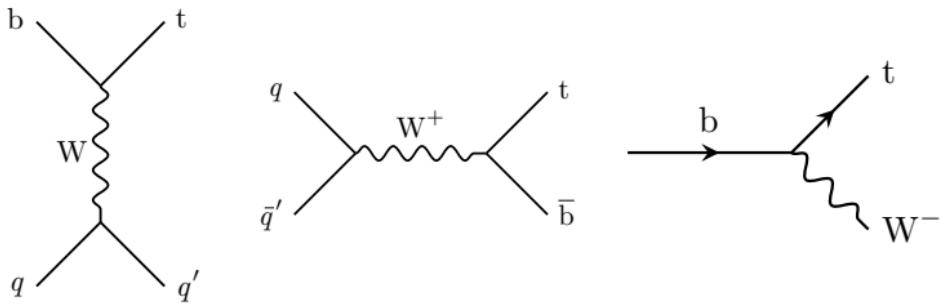
**Additionally require:**

$$\left| m_W - m_{q_1 q_2} \right| \leq 50 \text{ GeV} \quad \text{and} \quad 130 \text{ GeV} \leq m_{q_1 q_2 b, \min} \leq 210 \text{ GeV}$$

- $t\bar{t}$  dilepton events:



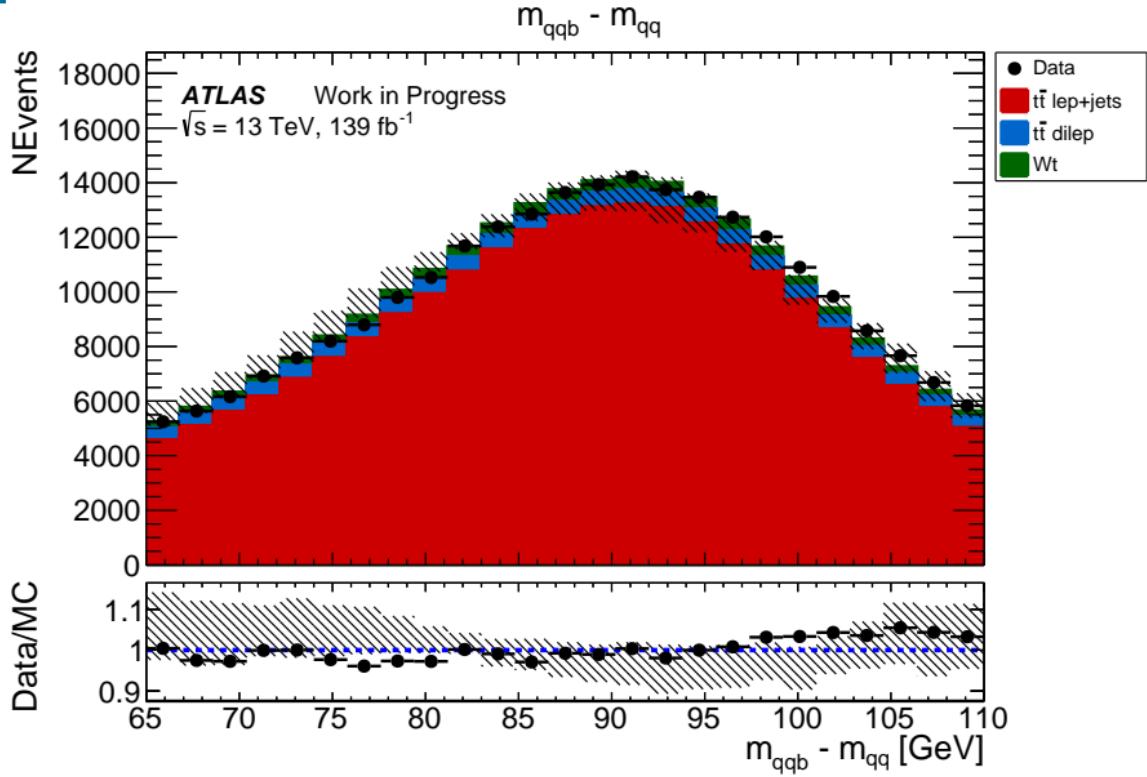
- Single top events (t-channel, s-channel, Wt)



Single top t-channel and s-channel contributions are negligible  
⇒ only consider  $t\bar{t}$  dilepton and single top Wt contamination

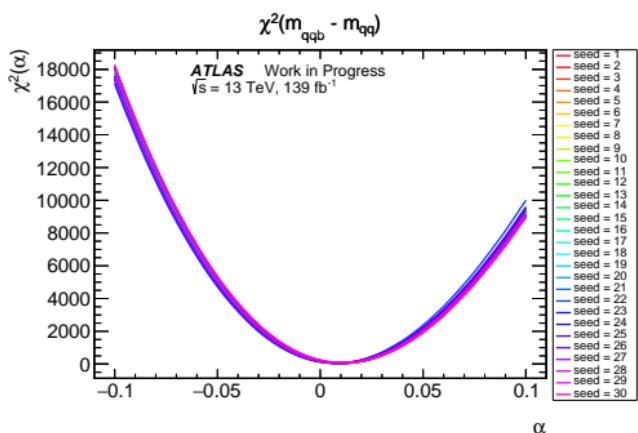
# $m_{qqb} - m_{qq}$ distribution: Data vs Nominal MC sample with no bJES corrections (Run 2, $\sqrt{s} = 13$ TeV, $\mathcal{L} = 139 \text{ fb}^{-1}$ )

8 / 19

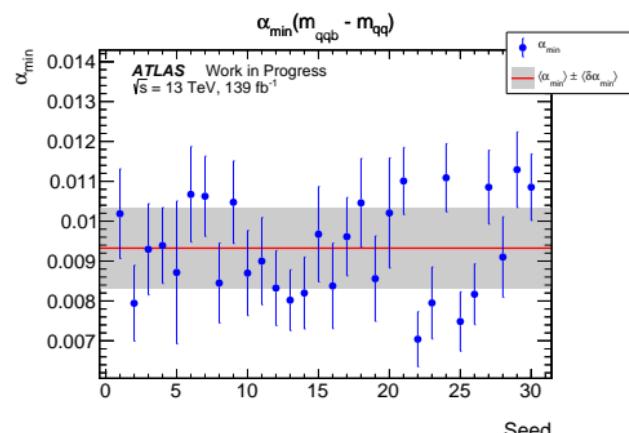


# Inclusive measurement in Data: $\alpha_{\min}^{\text{nominal}} \pm \delta\alpha^{\text{stat}}$ for Nominal FS $m_{qqb} - m_{qq}$ (OP spline)

$\chi^2(\text{FS})$



$\alpha_{\min}(\text{FS}) = 0.93\% \pm 0.10\%$



# Systematic uncertainties

$$\delta\alpha^{\text{syst}} = \alpha_{\min}^{\text{systematic}} - \alpha_{\min}^{\text{nominal}}$$

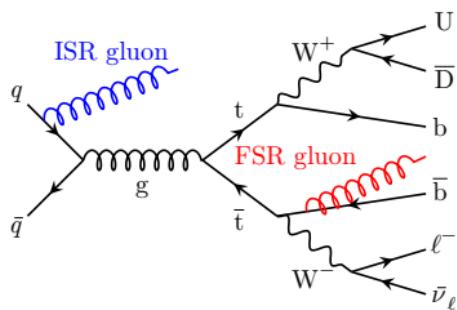
Compute  $\delta\alpha^{\text{syst}} = \alpha_{\min}^{\text{systematic}} - \alpha_{\min}^{\text{nominal}}$   
 using  $\alpha_{\min}^{\text{nominal}}$  (FS) or  $\alpha_{\min}^{\text{nominal}}$  (AFII) whether the systematic is FS or AFII samples  
**Total of 296 uncertainties**

➤ **219 Alternative weights** including :

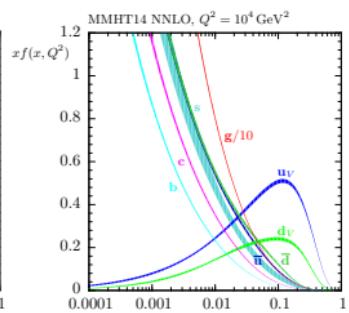
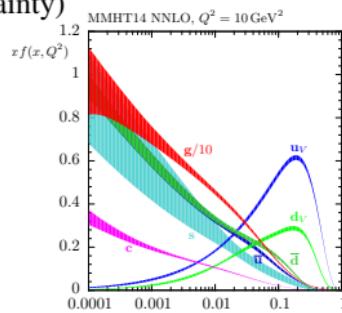
FS: b-tagging , scale  $\mu_R$  &  $\mu_F$ , ISR, FSR...

FS: PDF sets (PDF4LHC) compare results of alternative PDF sets (90901→90930)  
 to the nominal PDF set (90900)

AFII: RadHigh (ISR uncertainty)



Initial and final state radiation

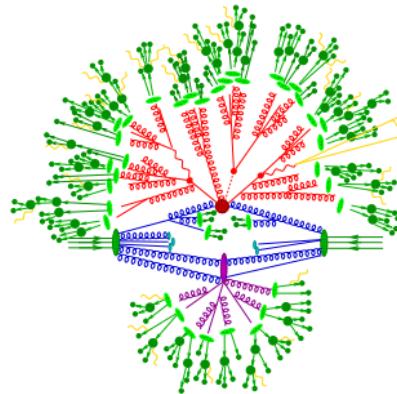


Parton distribution function uncertainties  
[Harland-Lang, L. A. and Martin, A. D.](#) and  
[Motylinski, P. and Thorne, R. S.](#)

# Systematic uncertainties

$$\delta\alpha^{\text{syst}} = \alpha_{\min}^{\text{systematic}} - \alpha_{\min}^{\text{nominal}}$$

- **72 Detector systematics** including (FS) :  
JER, NP Detector, NP Statistical, NP Modelling, Eta calibration, PileUp ...
- **2 MC Event generators uncertainties** (AFII) symmetrization of error as it is one sided systematics  
 $\delta\alpha^{\text{syst}}(\text{symmetrized}) = \pm\delta\alpha^{\text{syst}}(\text{one sided})$ :  
 parton shower modelling (use alternative shower, *Powheg+Herwig7*) &  
 matrix elements computation (use alternative generator, *aMCatNLO*)

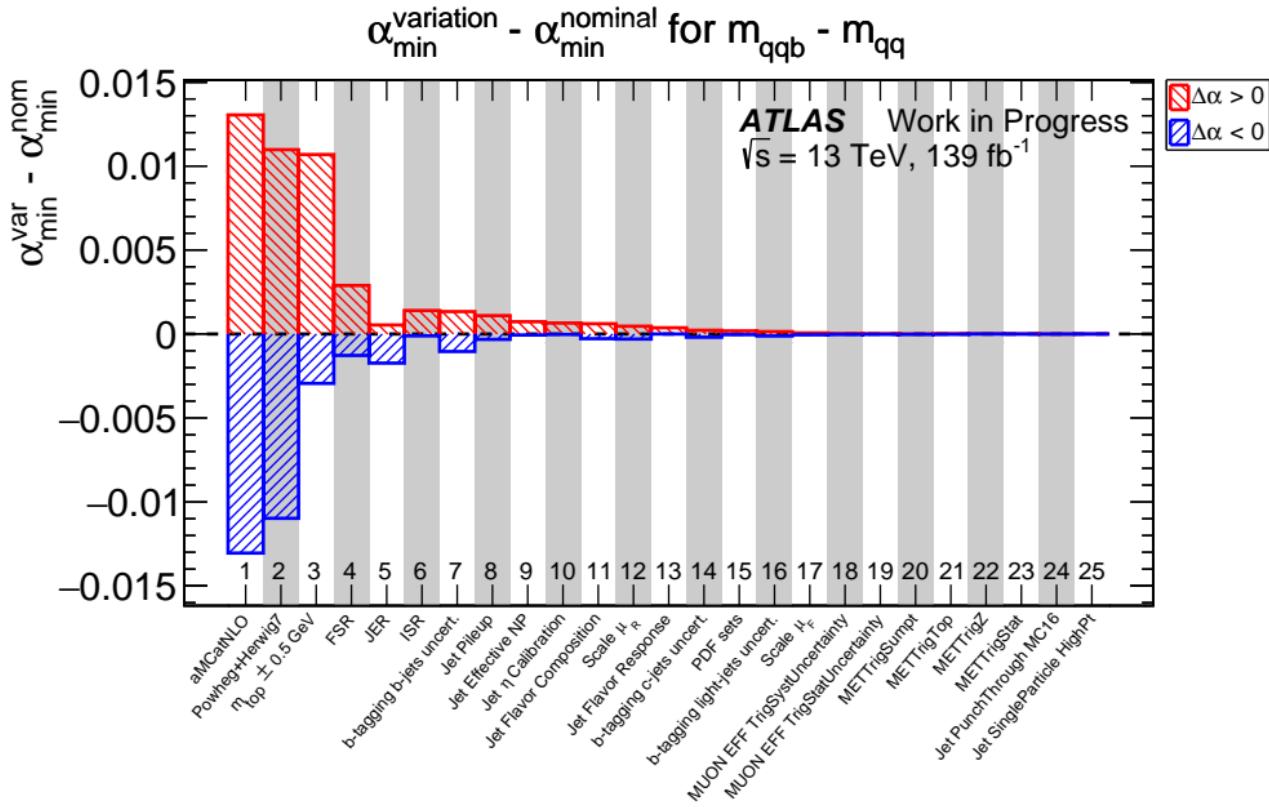


Sketch MC event generation, Stefan Höche

- **2 top mass variated samples** (AFII,  $m_t^{\text{Nominal}} \pm 0.5 \text{ GeV}$ ,  $m_t^{\text{Nominal}} = 172.5 \text{ GeV}$ ):  
 $m_{\text{top}} = 172.0 \text{ GeV}$  and  $173.0 \text{ GeV}$

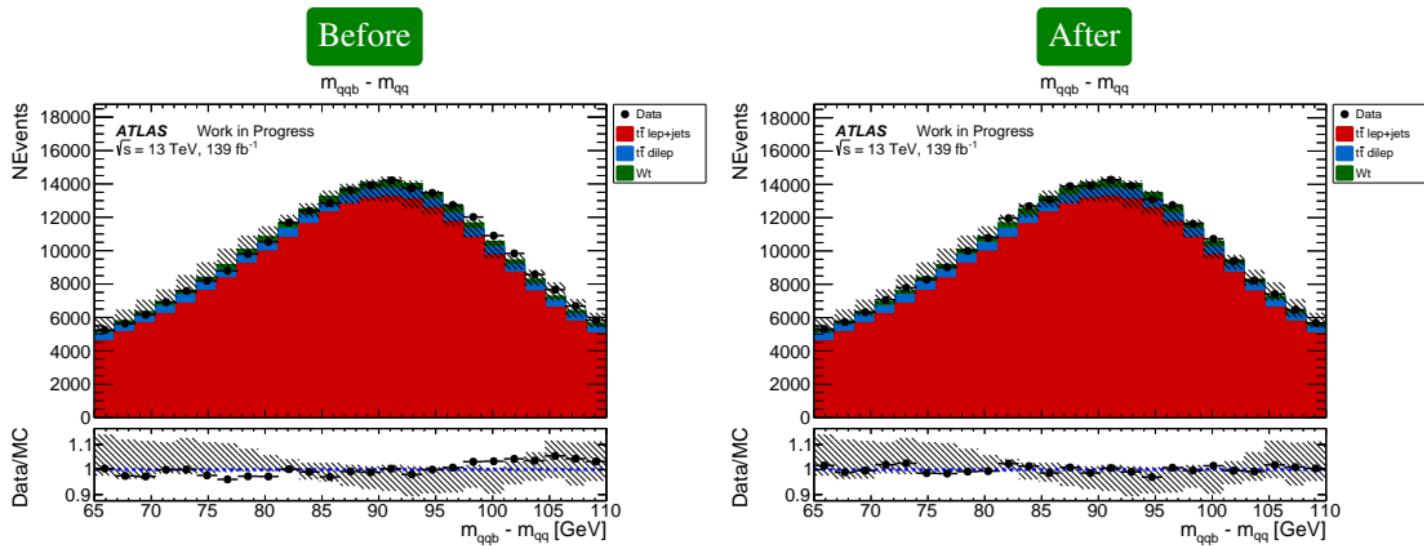
# Inclusive measurement in Data:

Systematic uncertainties  $\delta\alpha^{\text{syst}} = \pm^{+2.05\%}_{-1.75\%}$



# $m_{qqb} - m_{qq}$ distribution Data vs Nominal MC sample before and after $b$ -JES corrections

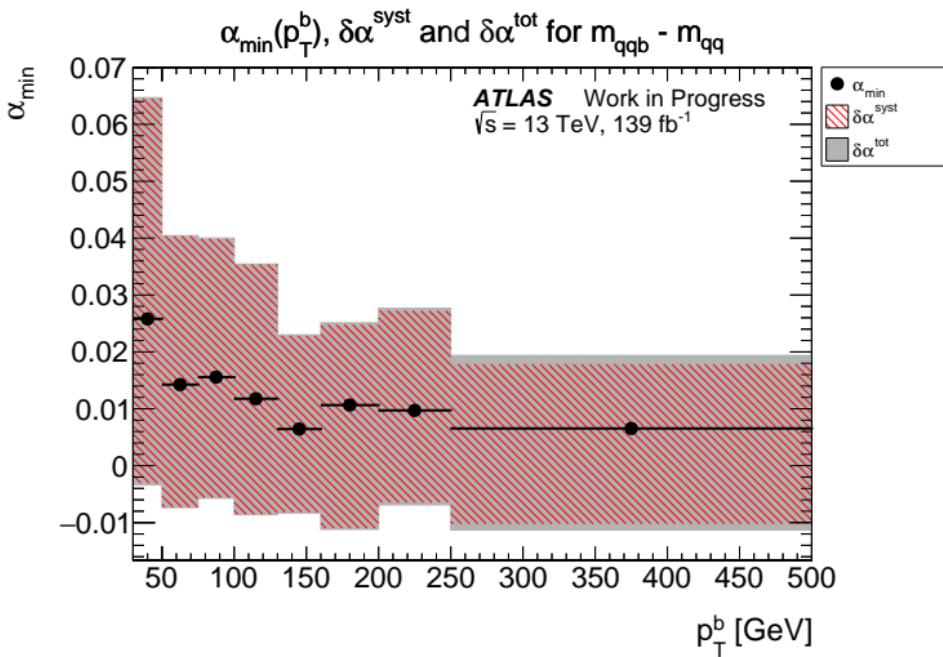
13 / 19



$$E_b^{\text{Data corrected}} = E_b^{\text{Data}} / (1 + \alpha_{\min}), \quad p_{Tb}^{\text{Data corrected}} = p_{Tb}^{\text{Data}} / (1 + \alpha_{\min})$$

$$\alpha_{\min} \pm \delta\alpha^{\text{stat}} \pm \delta\alpha^{\text{syst}} = 0.93\% \pm 0.10\% \pm {}^{+2.05\%}_{-1.75\%}$$

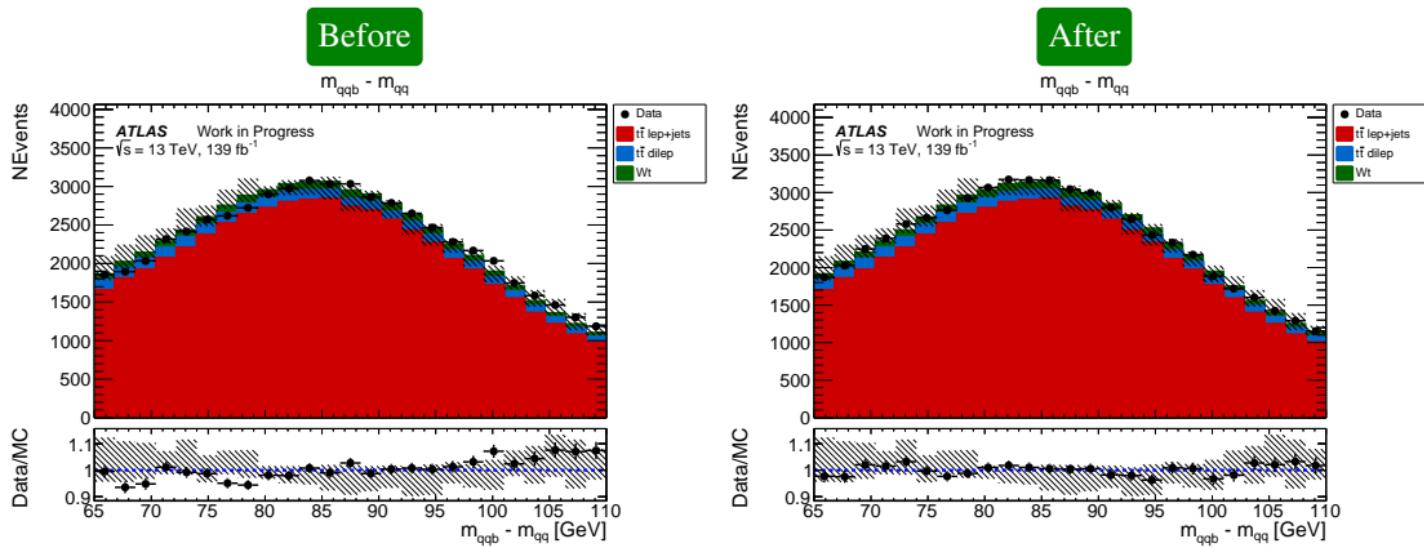
# Differential measurement w.r.t $p_T^b$



- Main systematic uncertainties are related to modelling and top mass uncertainties
- Inclusive and differential measurement are in agreement
- Good precision on the energy scale factor

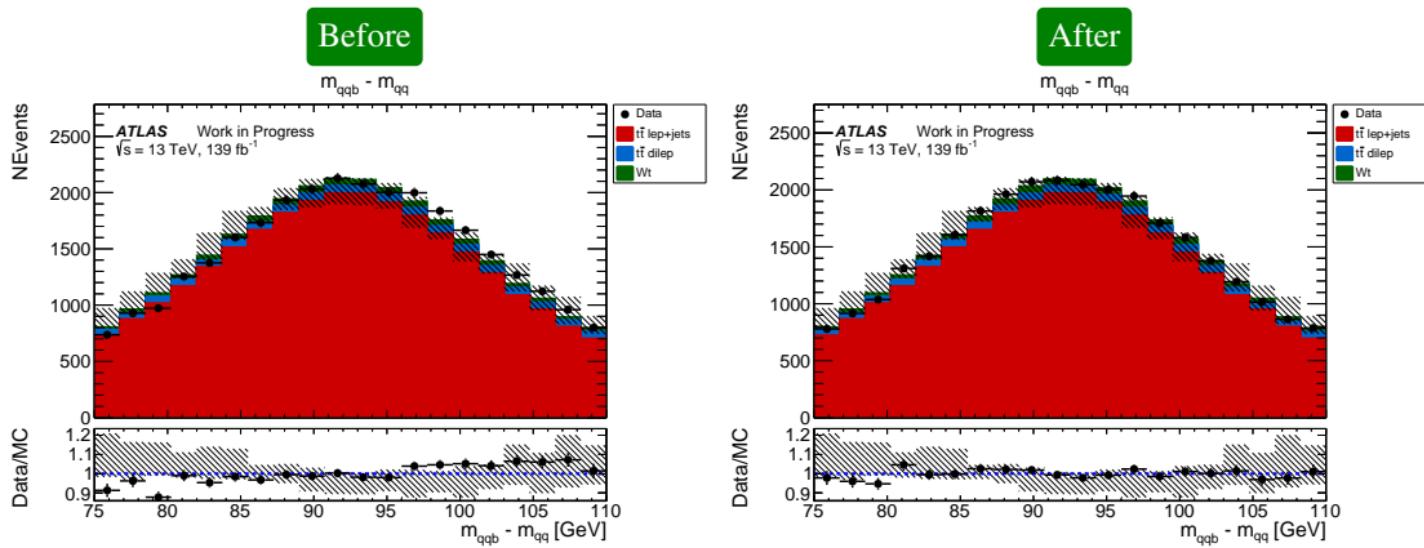
# $m_{qqb} - m_{qq}$ distribution Data vs Nominal MC sample before and after $b$ -JES corrections for $p_T^b \in [30, 50)$ GeV

15 / 19



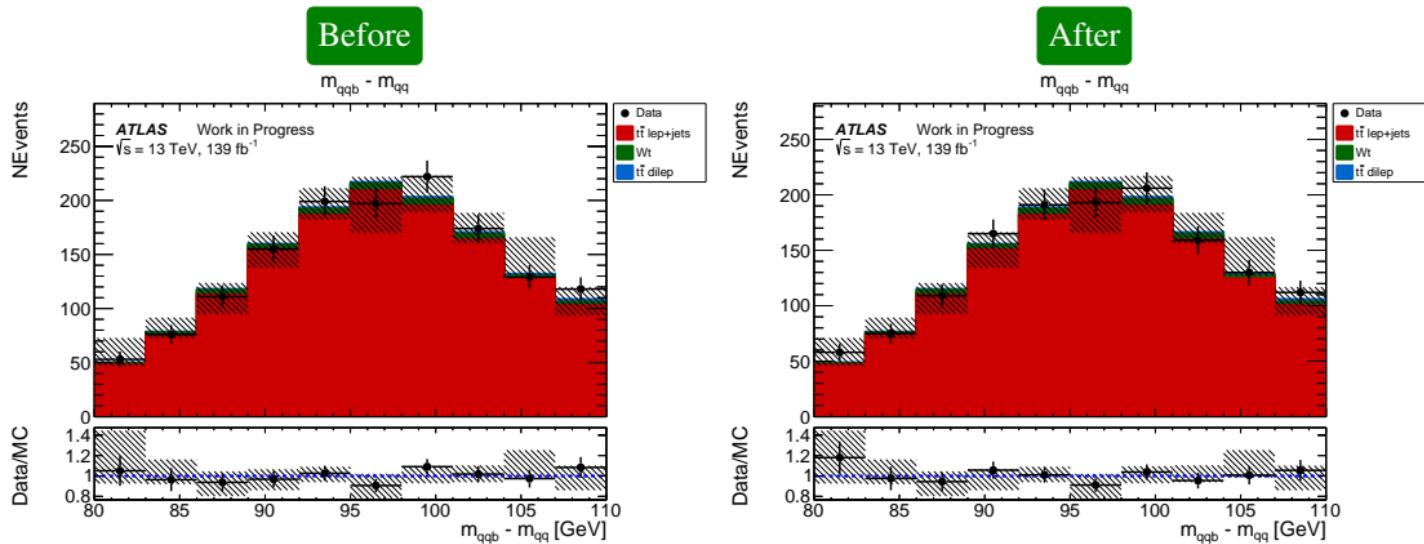
# $m_{qqb} - m_{qq}$ distribution Data vs Nominal MC sample before and after $b$ -JES corrections for $p_T^b \in [100, 130)$ GeV

16 / 19



# $m_{qqb} - m_{qq}$ distribution Data vs Nominal MC sample before and after $b$ -JES corrections for $p_T^b \in [250, 500] \text{ GeV}$

17 / 19



Correction less visible in high  $p_T^b$ -regions  
because the bJES corrections are closer to 0 when  $p_T^b$  increases

**Wanted to share those results with my oysters colleagues  
from La Rochelle...**



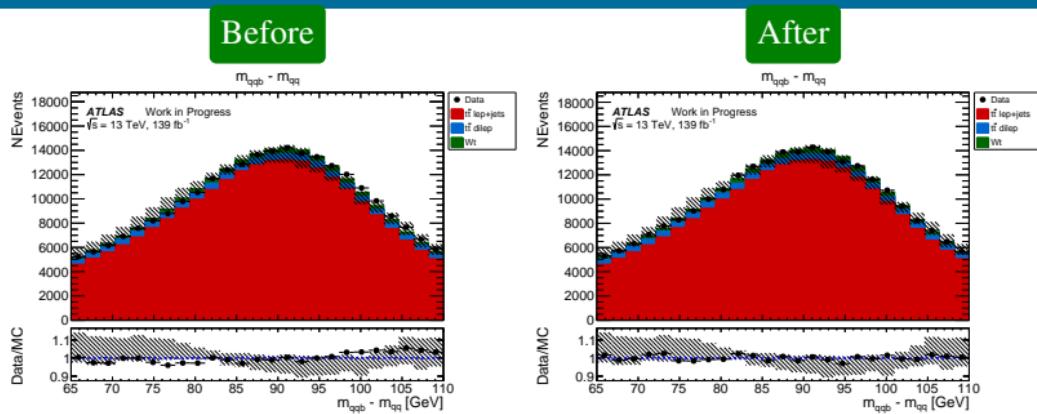
**... Could not find them** 

- First time b-jet energy scale factor is measured in Data
- Energy correction of the order of 1% with a total uncertainty of the order of 2%
- Differential measurement in 1 dimension w.r.t  $p_T^b$  as more variations are expected compared to  $\eta$
- Improvement of the MC vs Data agreement after applying bJES corrections in Data especially in low  $p_T^b$ -regions

**Thanks for your attention!**

# Backup

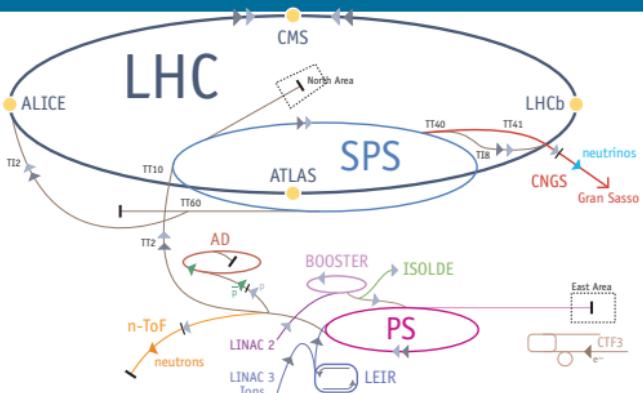
# $b$ -JES corrections... link with oyster



Lot of work...



... for a small result

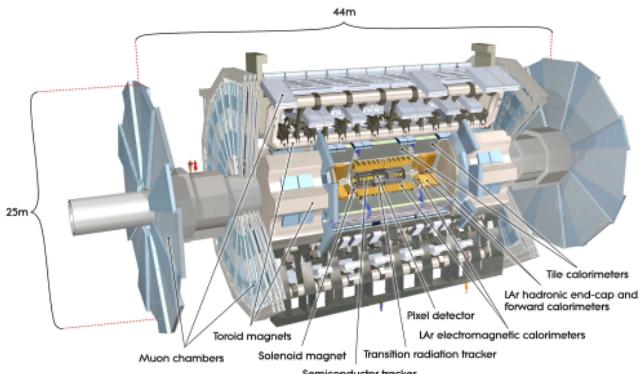


Julie Haffner

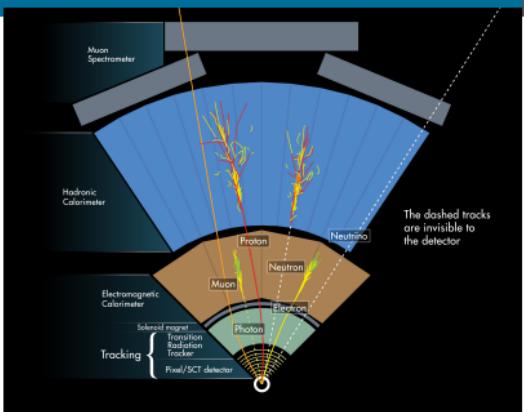
- 27 km of circumference
- $pp$  collisions at  $\sqrt{s} = 7, 8, 13$  TeV (but not only... sometimes  $Pb - Pb$  collisions)
- Bunch crossing every 25 ns
- $\langle \mu \rangle = 30$  inelastic collisions per bunch crossing
- 4 main detectors:
  - **ATLAS & CMS**: general purpose detectors to precisely study the SM
  - **LHCb**: studies differences between matter and antimatter with bottom and charm hadrons
  - **ALICE**: heavy-ion collisions and properties of the quark-gluon plasma produced in those collisions

# The ATLAS detector and its coordinates system

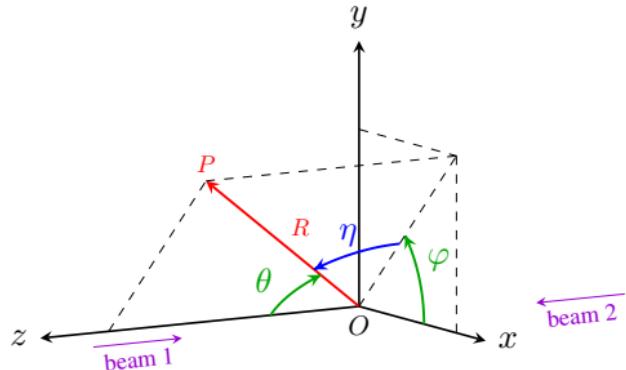
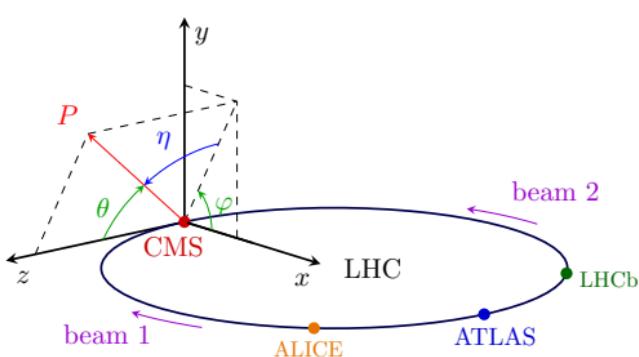
④ 4 / 35



Joao Pequenao

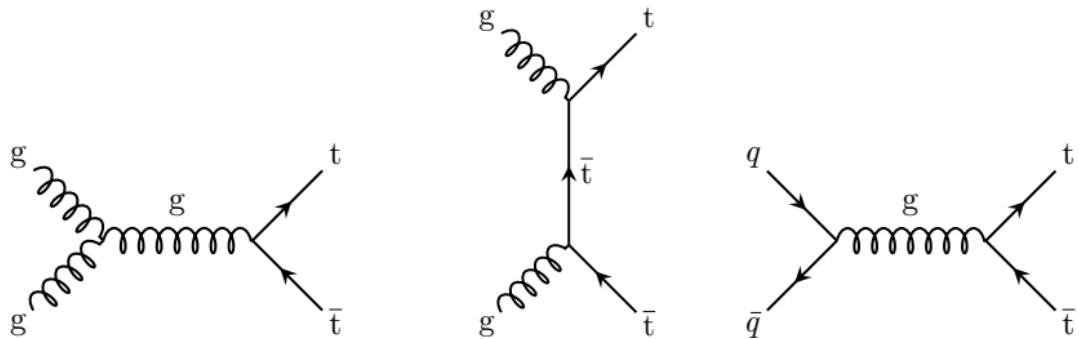


Joao Pequenao



➤ Transverse momentum  $p_T = \sqrt{p_x^2 + p_y^2}$

➤ Pseudorapidity  $\eta = -\ln(\tan(\theta/2))$

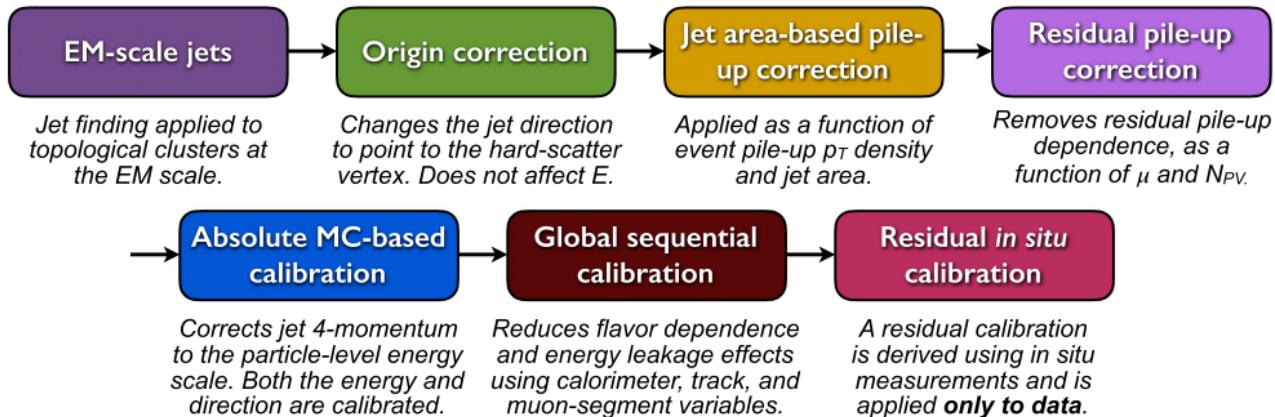


(a) gluon-gluon fusion  
( $s$ -channel)

(b) gluon-gluon fusion  
( $t$ -channel)

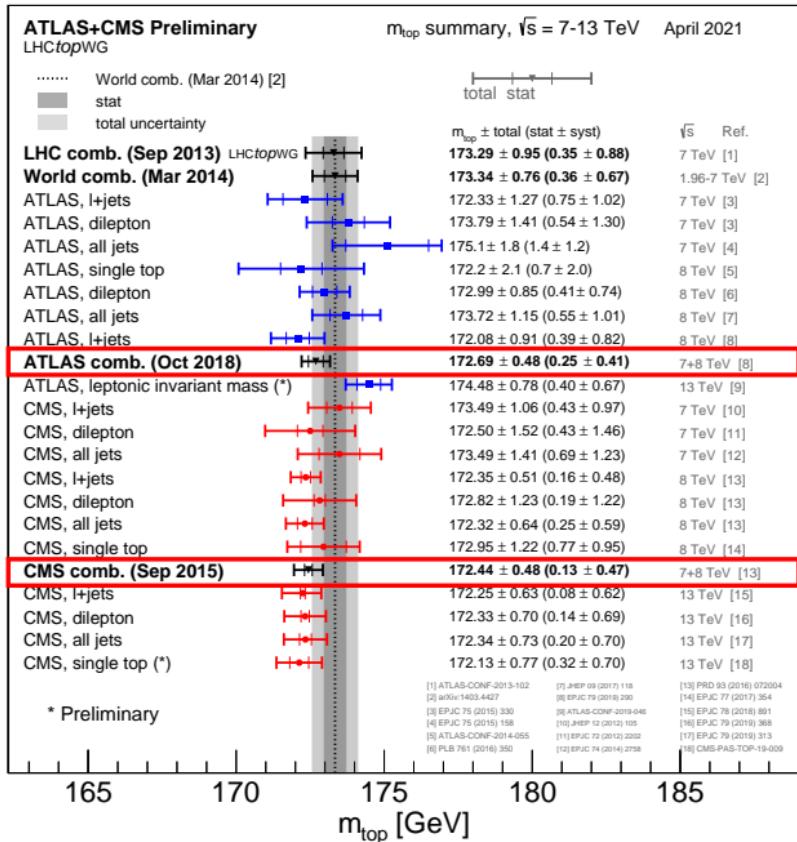
(c) quark-antiquark  
annihilation

# ATLAS Jet calibration



ATLAS Collaboration

# Top quark mass measurements by ATLAS and CMS



LHC Top Physics Working Group

- [1] ATLAS-CDF-2013-102
- [2] ATLAS-CDF-2014-027
- [3] EPJC 75 (2015) 330
- [4] EPJC 75 (2015) 158
- [5] ATLAS-CDF-2014-055
- [6] PLB 761 (2016) 350
- [7] JHEP 09 (2017) 118
- [8] ATLAS-CDF-2014-048
- [9] ATLAS-CDF-2019-048
- [10] JHEP 12 (2012) 105
- [11] EPJC 72 (2012) 2032
- [12] EPJC 74 (2014) 2758
- [13] PRD 93 (2016) 072004
- [14] EPJC 77 (2017) 364
- [15] EPJC 78 (2018) 891
- [16] EPJC 79 (2018) 368
- [17] EPJC 79 (2018) 313
- [18] CMS-PAS-TOP-19-009

# $p_T^b$ -regions: details

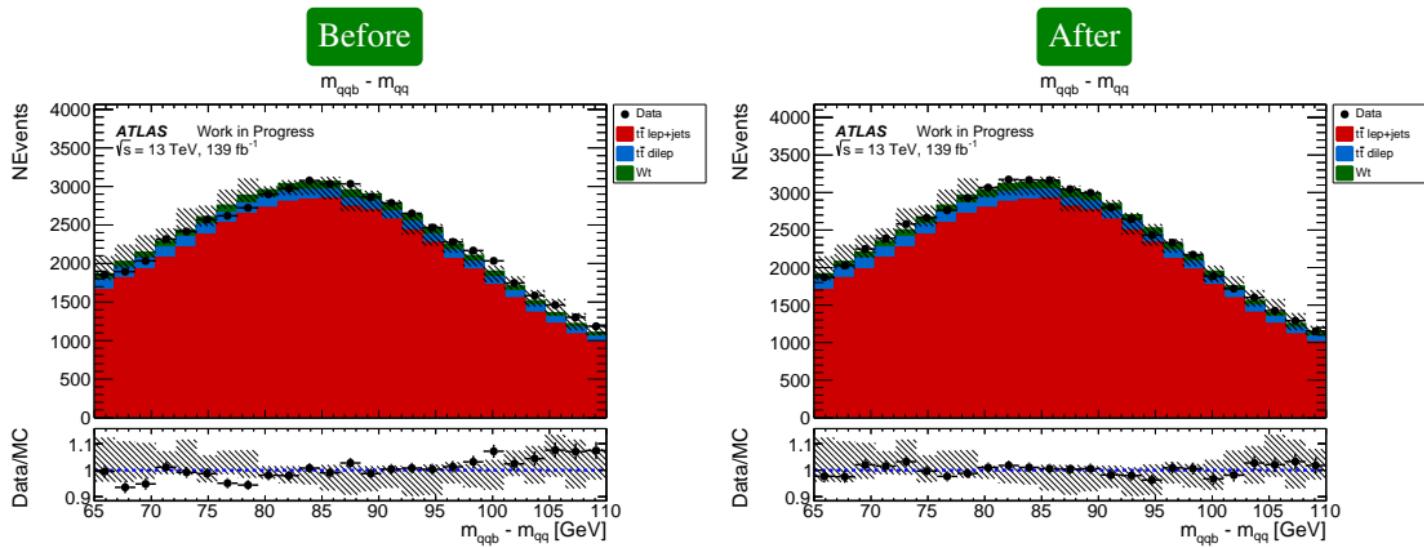
$p_T^b$ regions	$m_{qqb} - m_{qq}$	number of bins	$m_{qqb} - m_{qq}$ bounds
[30, 50) GeV		25	[65, 110] GeV
[50, 75) GeV		25	[65, 110] GeV
[75, 100) GeV		20	[70, 110] GeV
[100, 130) GeV		20	[75, 110] GeV
[130, 160) GeV		20	[75, 110] GeV
[160, 200) GeV		20	[80, 110] GeV
[200, 250) GeV		15	[80, 110] GeV
[250, 500) GeV		10	[80, 110] GeV

Table:  $m_{qqb} - m_{qq}$  distribution: number of bins and bounds depending on the  $p_T^b$ -region

- For the inclusive measurement 25 bins with bounds = [65, 110] GeV for the  $m_{qqb} - m_{qq}$  distribution
- Reasonably large enough  $p_T^b$ -regions
- Low number of bins to avoid bin fluctuations
- The bounds of the  $m_{qqb} - m_{qq}$  distribution are chosen so that the peak position is centered and both side on the left and right of the peak have roughly the same importance

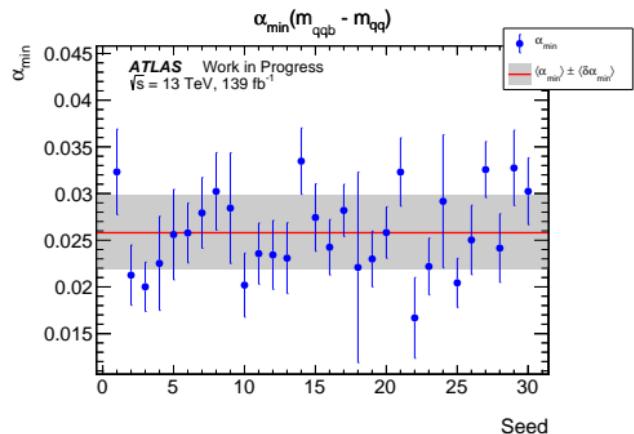
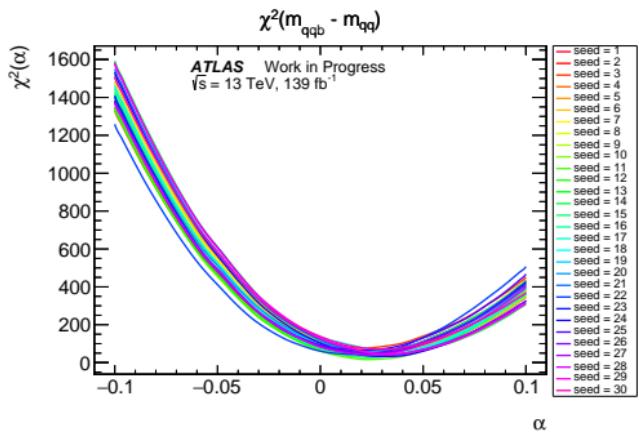
# $m_{qqb} - m_{qq}$ distribution Data vs Nominal MC sample before and after $b$ -JES corrections for $p_T^b \in [30, 50)$ GeV

⑥ 9 / 35

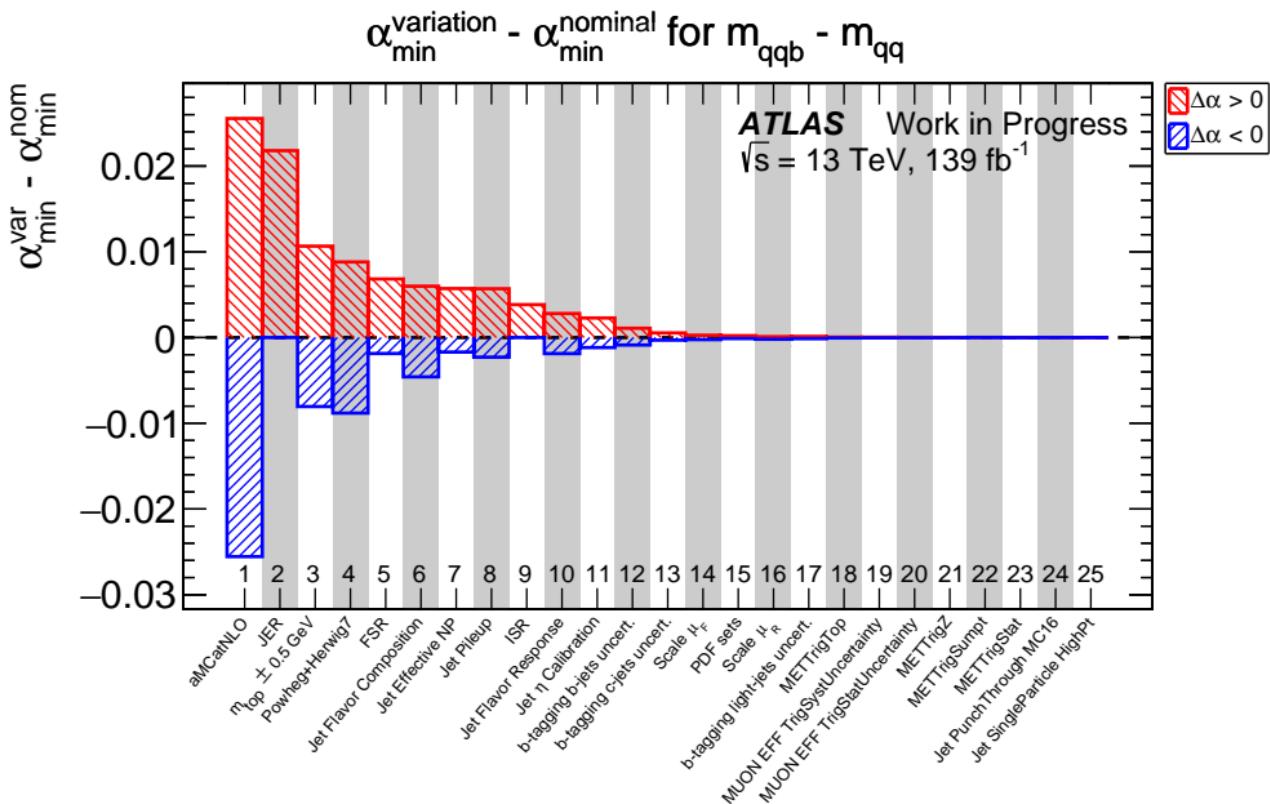


$\alpha_{\min}^{\text{nominal}} \pm \delta\alpha^{\text{stat}}$  for Nominal FS  $m_{qqb} - m_{qq}$   
 for  $p_T^b \in [30, 50] \text{ GeV}$   
 (OP spline)

⑥ 10 / 35

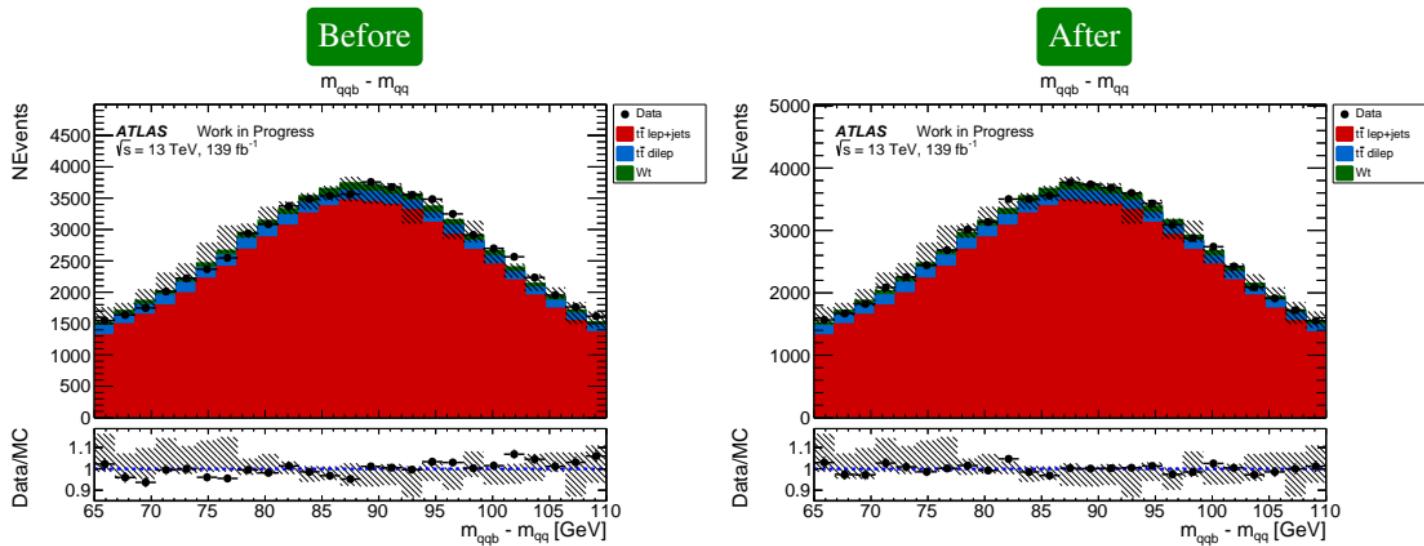


# Systematic uncertainties $\delta\alpha^{\text{syst}}$ for $p_T^b \in [30, 50] \text{ GeV}$



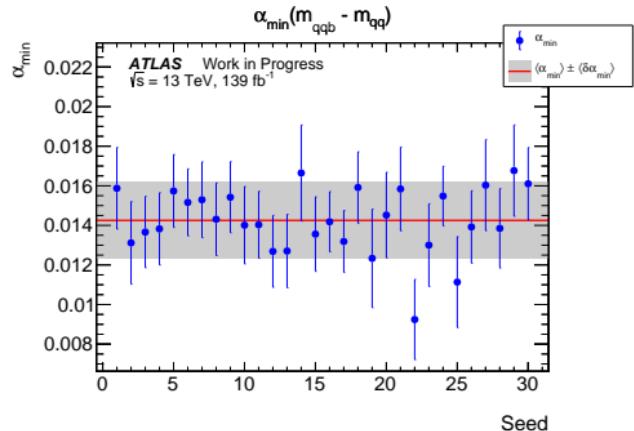
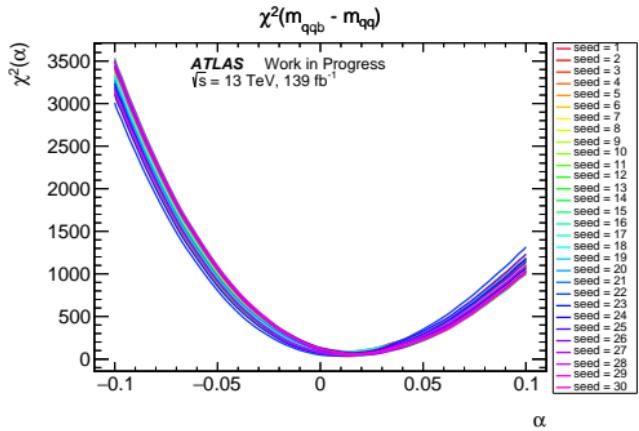
# $m_{qqb} - m_{qq}$ distribution Data vs Nominal MC sample before and after $b$ -JES corrections for $p_T^b \in [50, 75)$ GeV

⑥ 12 / 35

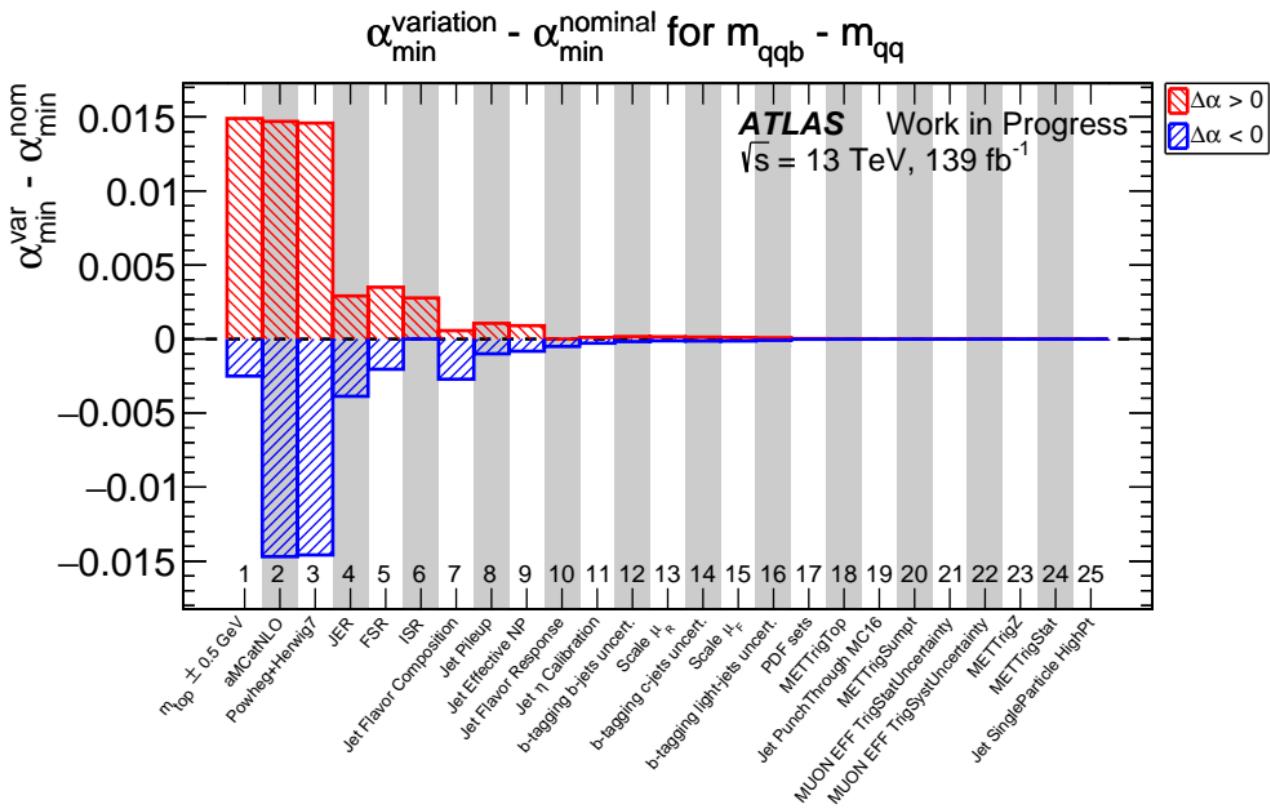


$\alpha_{\min}^{\text{nominal}} \pm \delta\alpha^{\text{stat}}$  for Nominal FS  $m_{qqb} - m_{qq}$   
 for  $p_T^b \in [50, 75] \text{ GeV}$   
 (OP spline)

⑥ 13 / 35

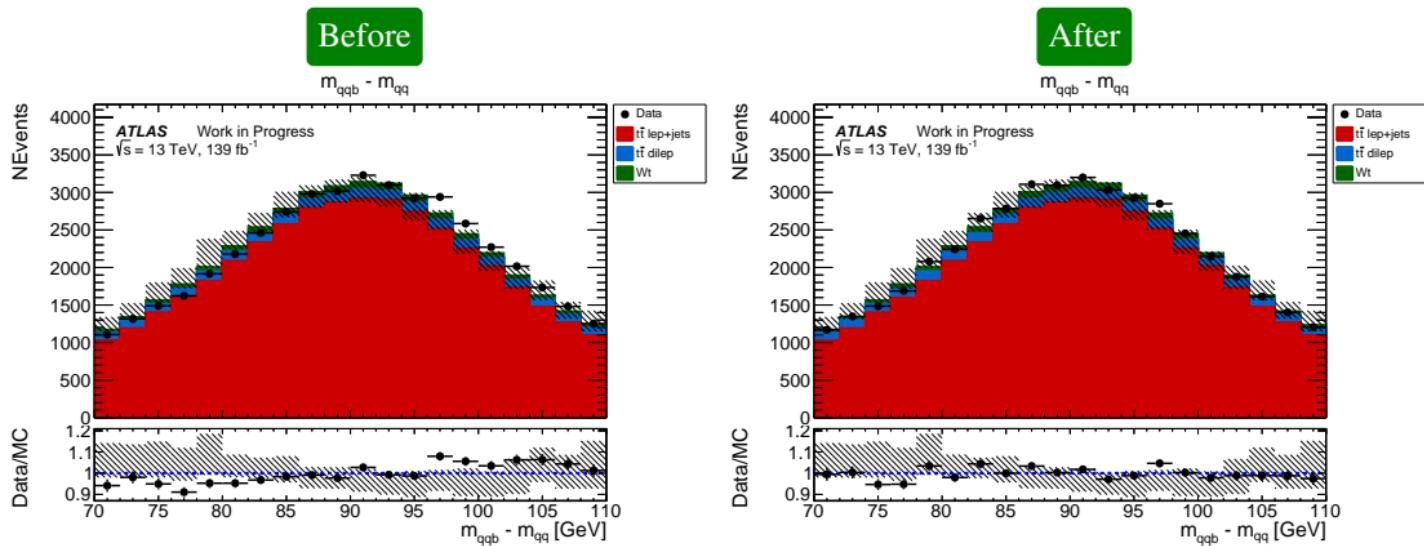


# Systematic uncertainties $\delta\alpha^{\text{syst}}$ for $p_T^b \in [50, 75] \text{ GeV}$



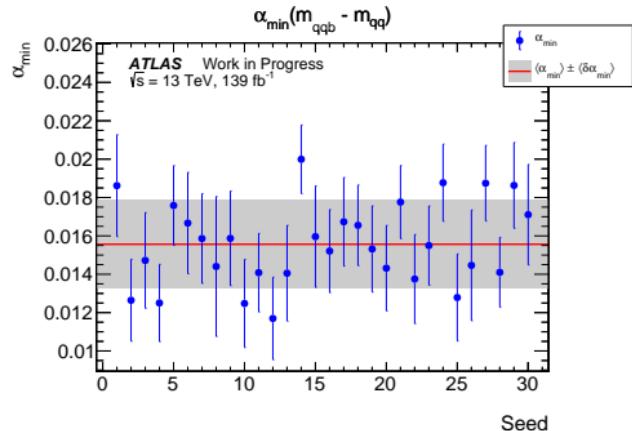
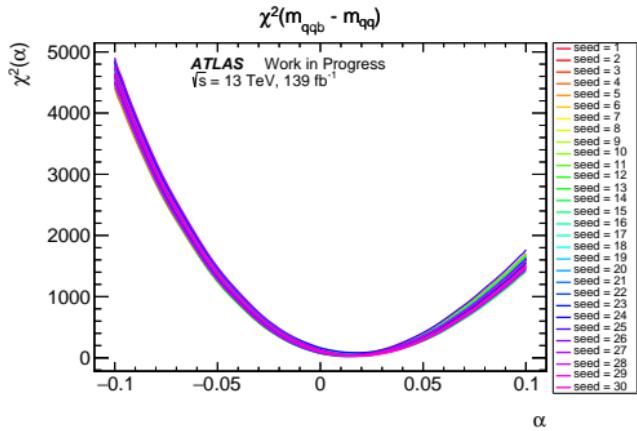
# $m_{qqb} - m_{qq}$ distribution Data vs Nominal MC sample before and after $b$ -JES corrections for $p_T^b \in [75, 100)$ GeV

⑥ 15 / 35

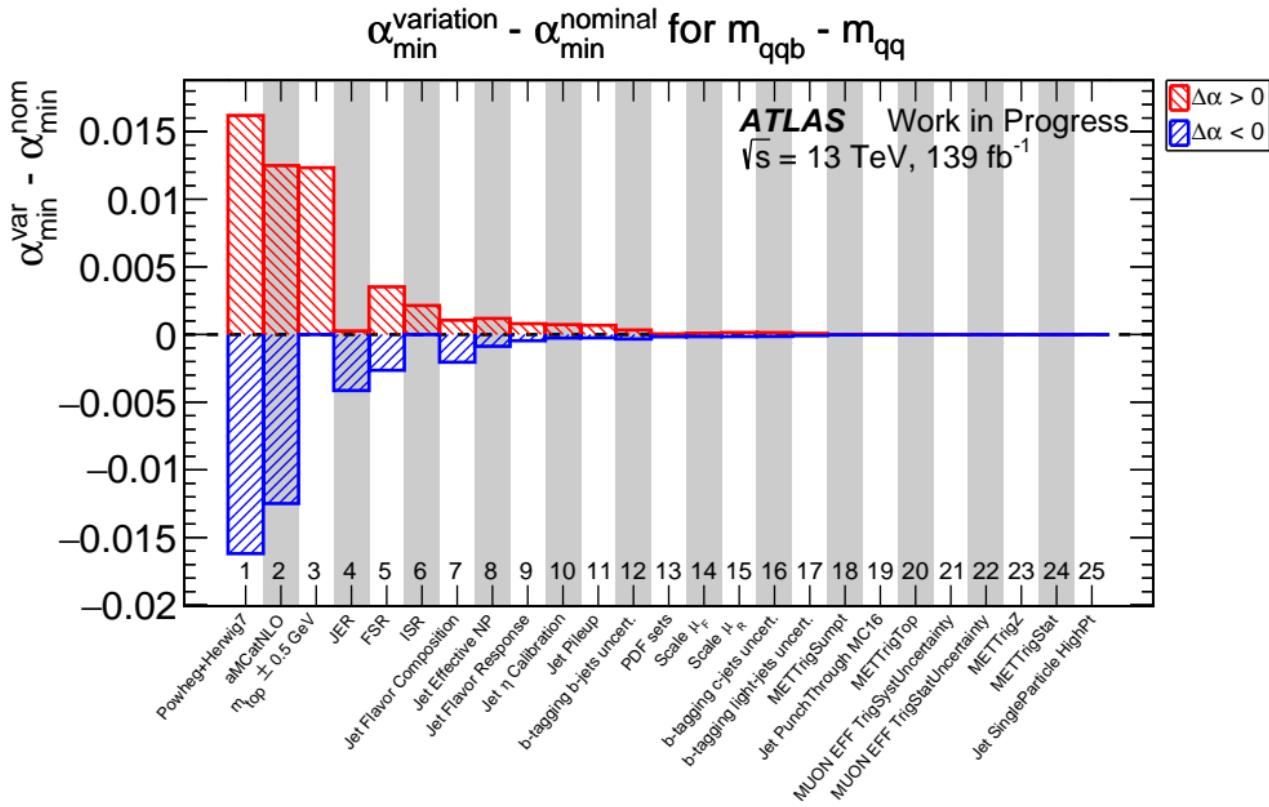


$\alpha_{\min}^{\text{nominal}} \pm \delta\alpha^{\text{stat}}$  for Nominal FS  $m_{qqb} - m_{qq}$   
 for  $p_T^b \in [75, 100] \text{ GeV}$   
 (OP spline)

⑥ 16 / 35

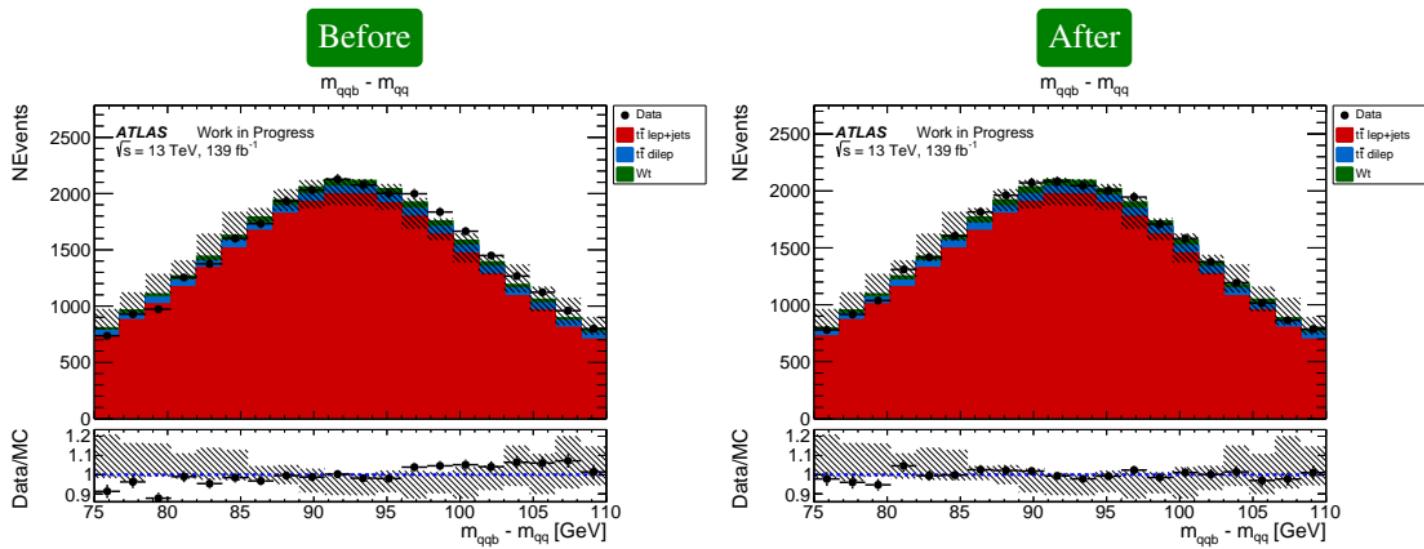


# Systematic uncertainties $\delta\alpha^{\text{syst}}$ for $p_T^b \in [75, 100) \text{ GeV}$



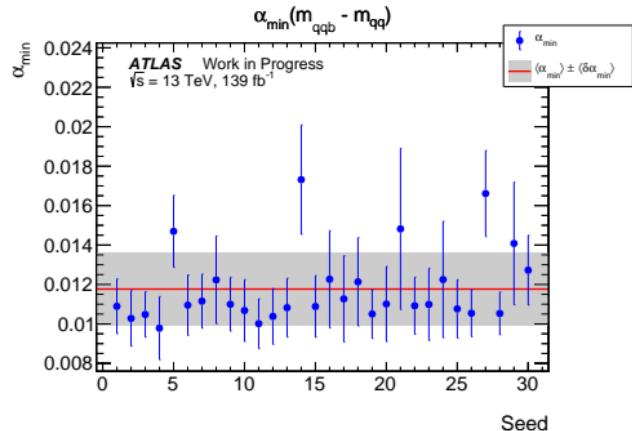
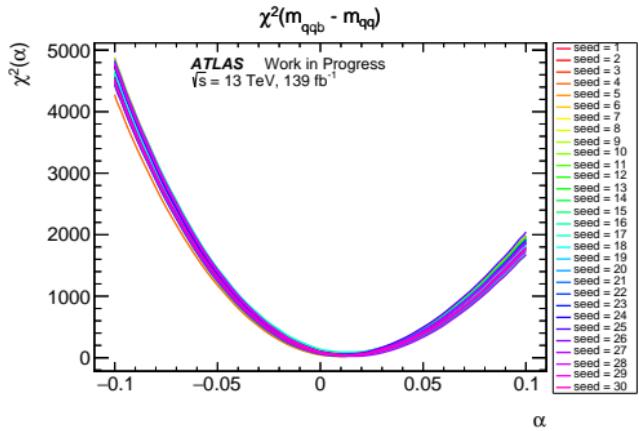
# $m_{qqb} - m_{qq}$ distribution Data vs Nominal MC sample before and after $b$ -JES corrections for $p_T^b \in [100, 130)$ GeV

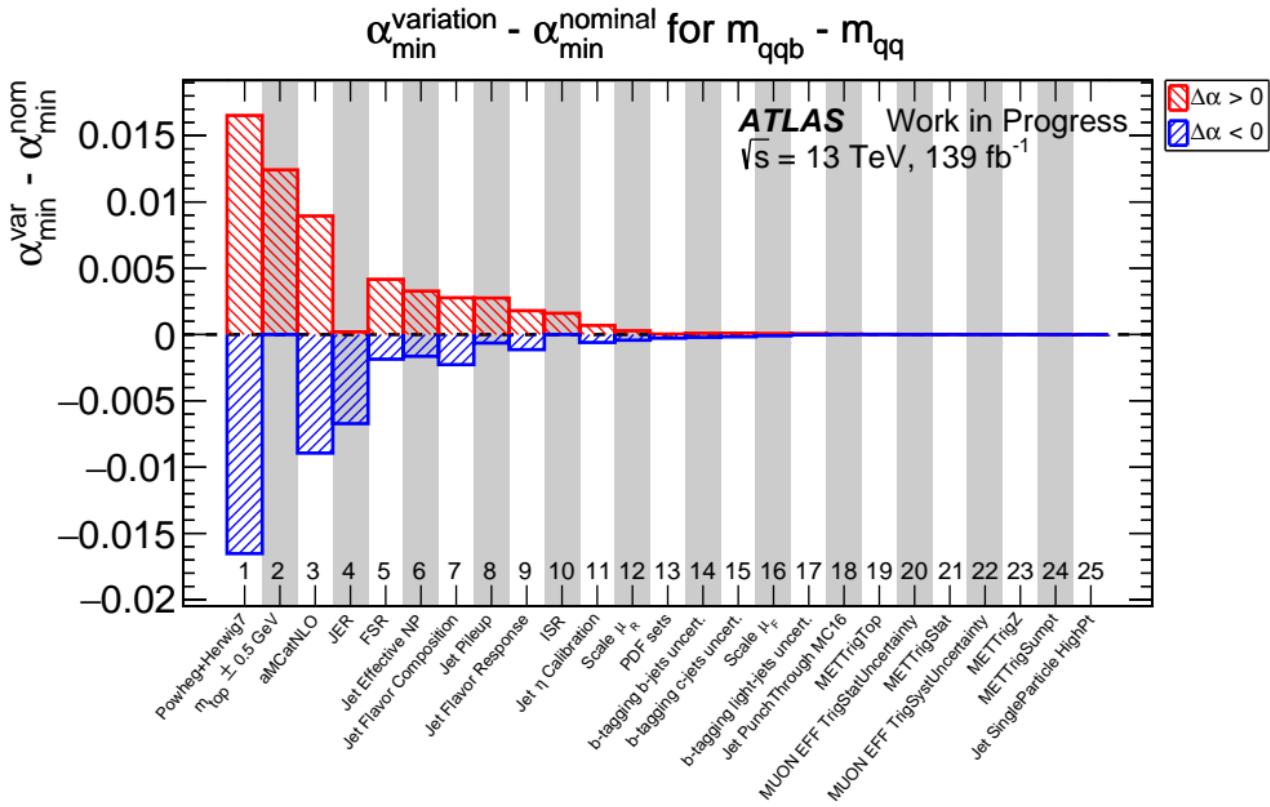
© 18 / 35



$\alpha_{\min}^{\text{nominal}} \pm \delta\alpha^{\text{stat}}$  for Nominal FS  $m_{qqb} - m_{qq}$   
 for  $p_T^b \in [100, 130] \text{ GeV}$   
 (OP spline)

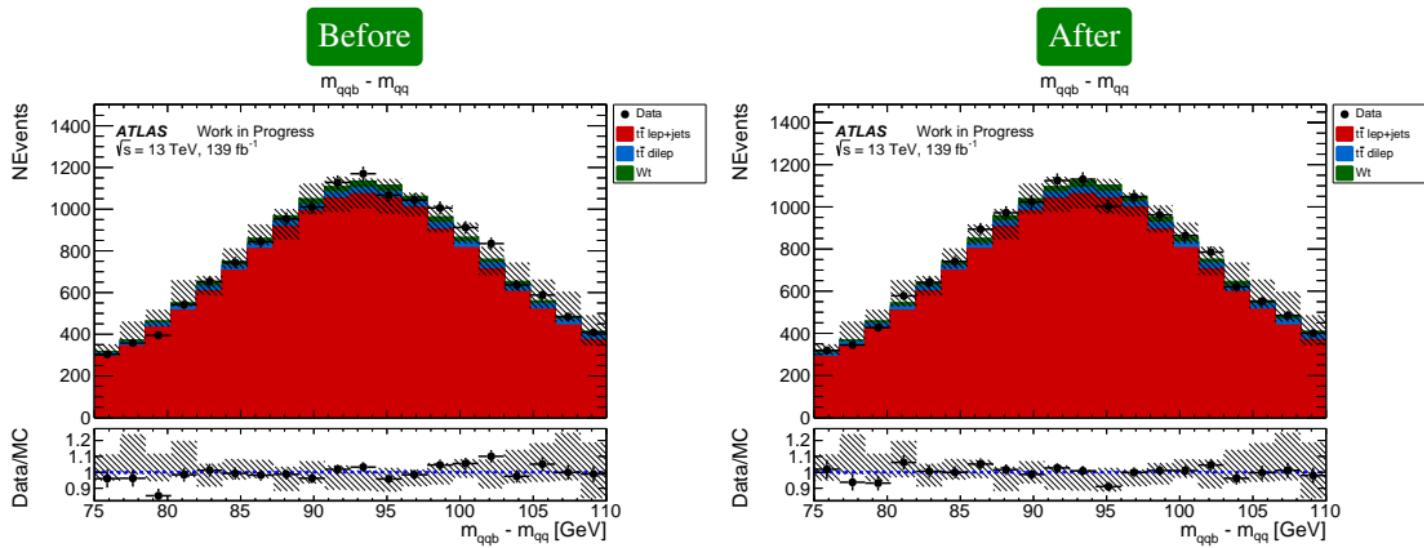
⑥ 19 / 35





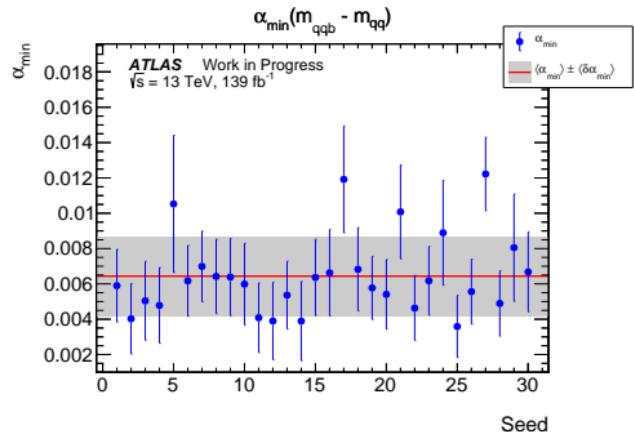
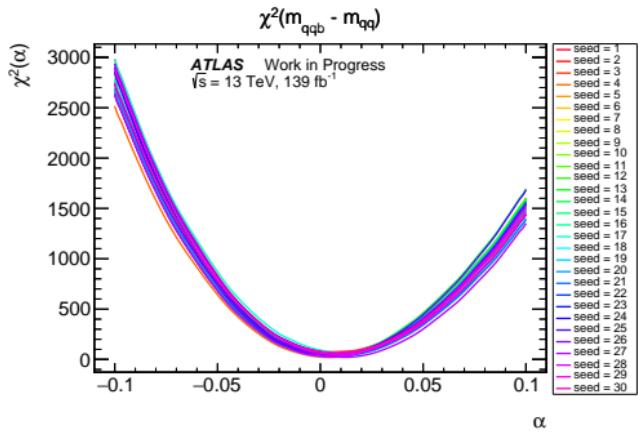
# $m_{qqb} - m_{qq}$ distribution Data vs Nominal MC sample before and after $b$ -JES corrections for $p_T^b \in [130, 160)$ GeV

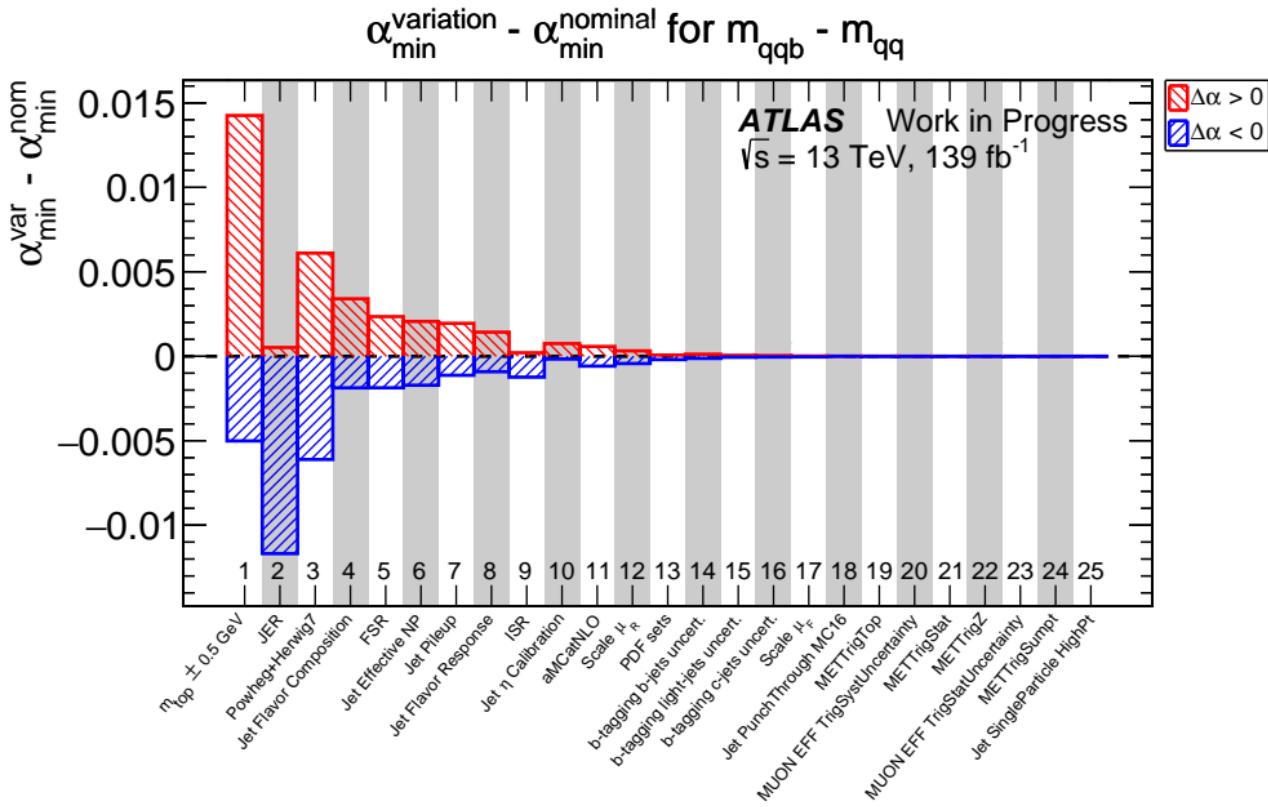
© 21 / 35



$\alpha_{\min}^{\text{nominal}} \pm \delta\alpha^{\text{stat}}$  for Nominal FS  $m_{qqb} - m_{qq}$   
 for  $p_T^b \in [130, 160] \text{ GeV}$   
 (OP spline)

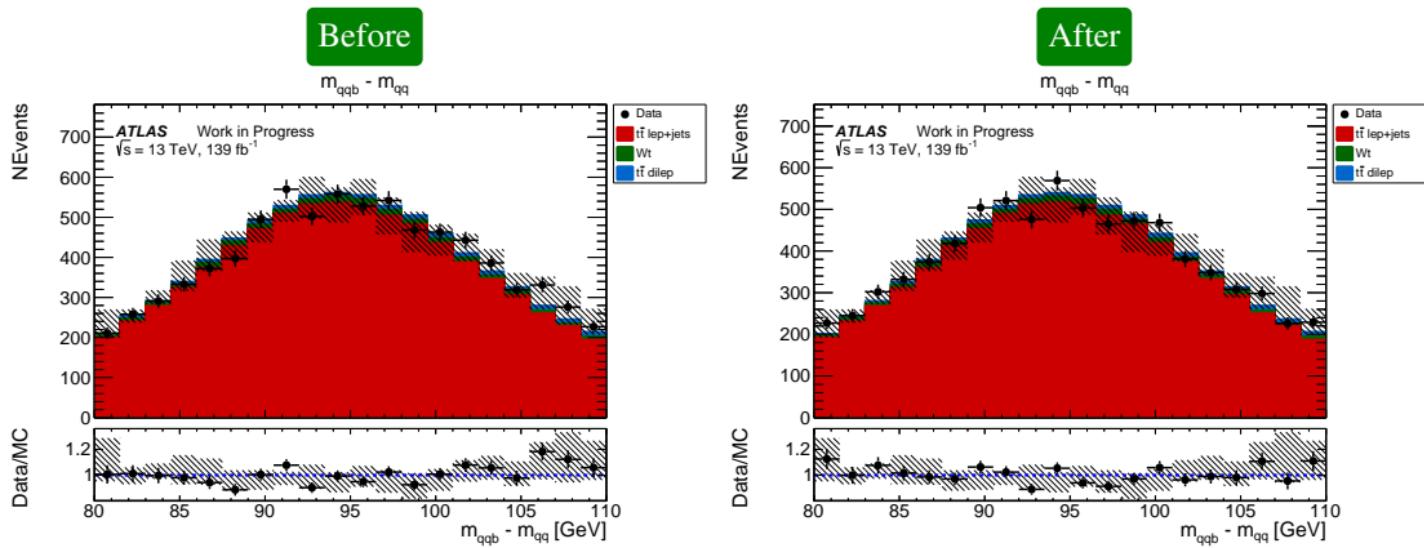
⑥ 22 / 35





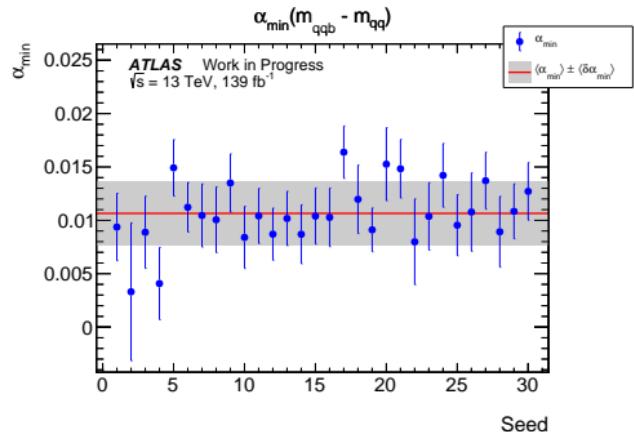
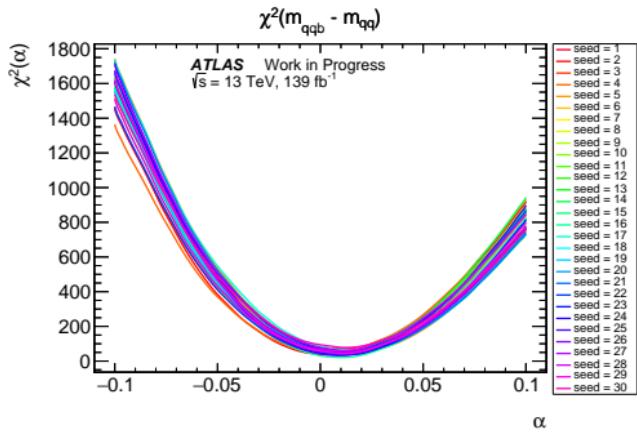
# $m_{qqb} - m_{qq}$ distribution Data vs Nominal MC sample before and after $b$ -JES corrections for $p_T^b \in [160, 200)$ GeV

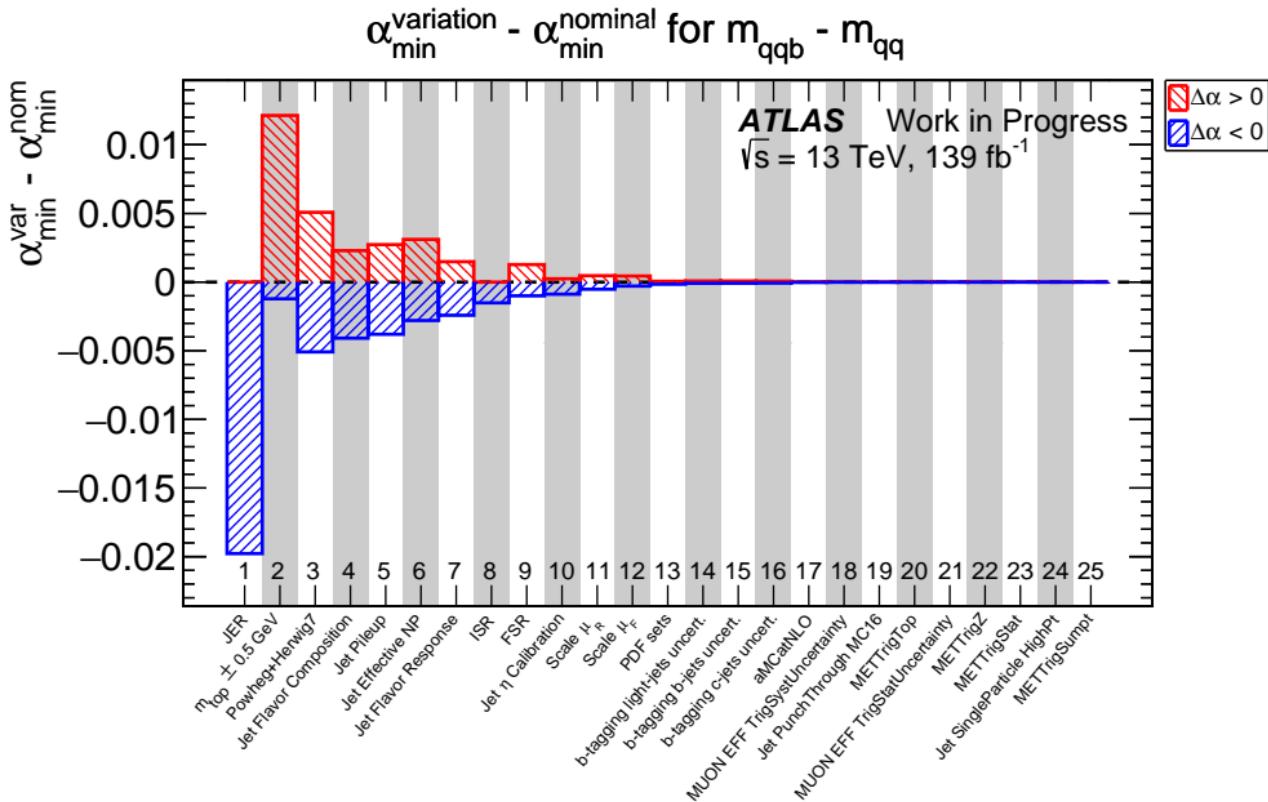
⑥ 24 / 35



$\alpha_{\min}^{\text{nominal}} \pm \delta\alpha^{\text{stat}}$  for Nominal FS  $m_{qqb} - m_{qq}$   
 for  $p_T^b \in [160, 200] \text{ GeV}$   
 (OP spline)

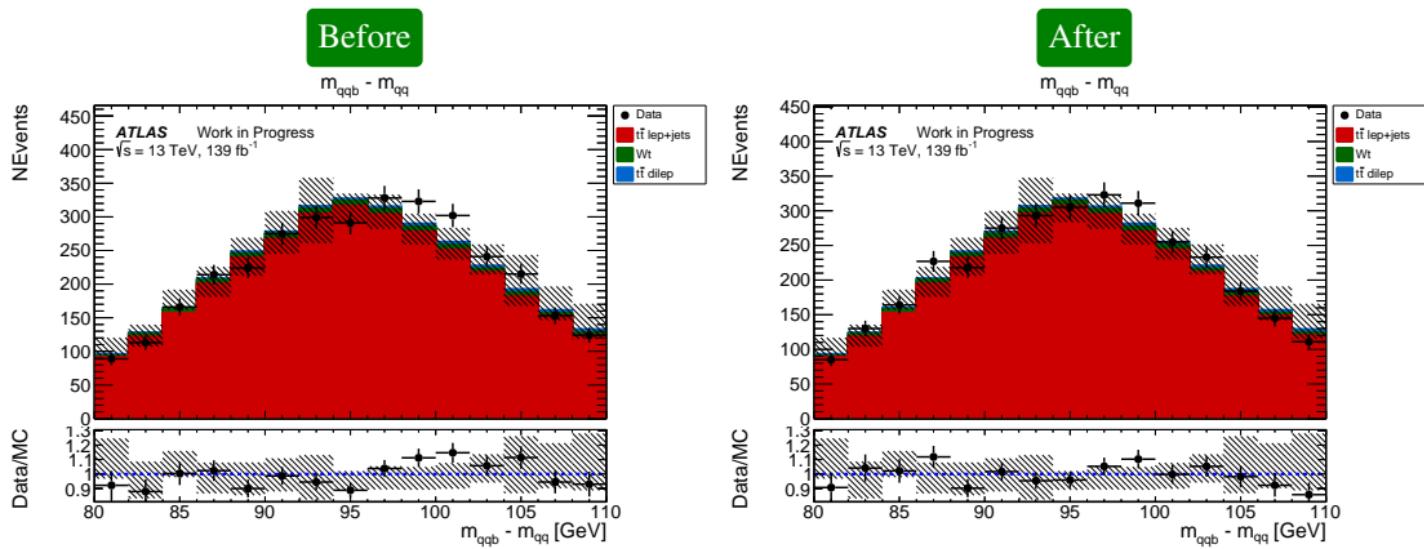
⑥ 25 / 35





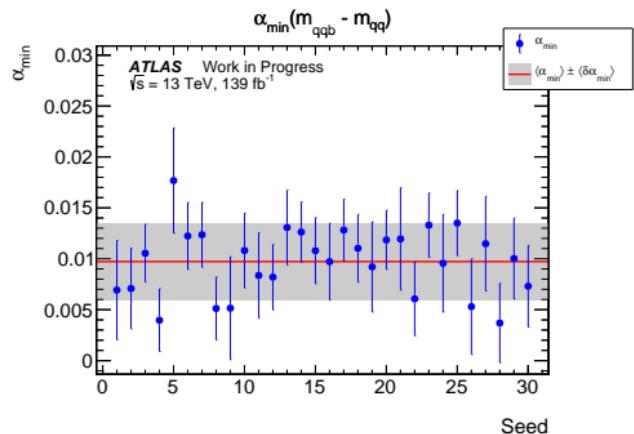
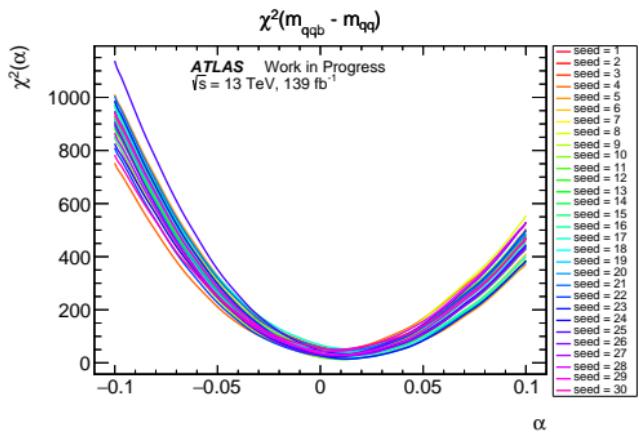
# $m_{qqb} - m_{qq}$ distribution Data vs Nominal MC sample before and after $b$ -JES corrections for $p_T^b \in [200, 250)$ GeV

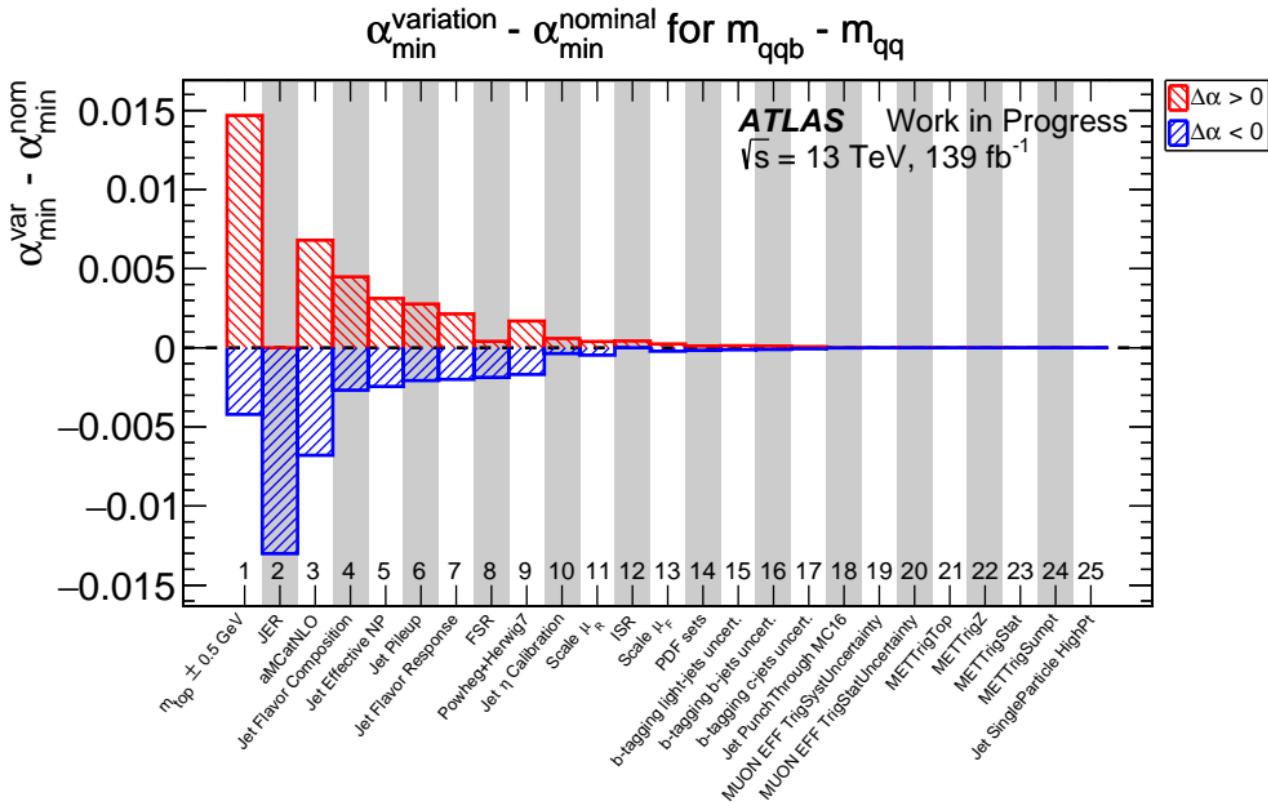
© 27 / 35



$\alpha_{\min}^{\text{nominal}} \pm \delta\alpha^{\text{stat}}$  for Nominal FS  $m_{qqb} - m_{qq}$   
 for  $p_T^b \in [200, 250] \text{ GeV}$   
 (OP spline)

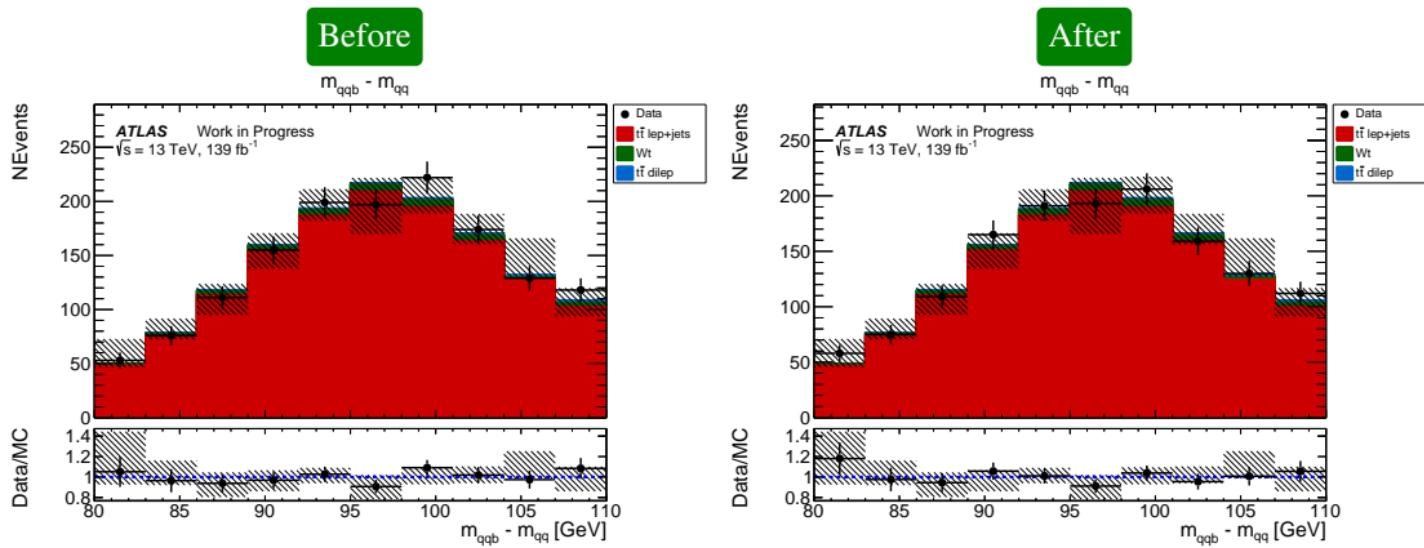
⑥ 28 / 35





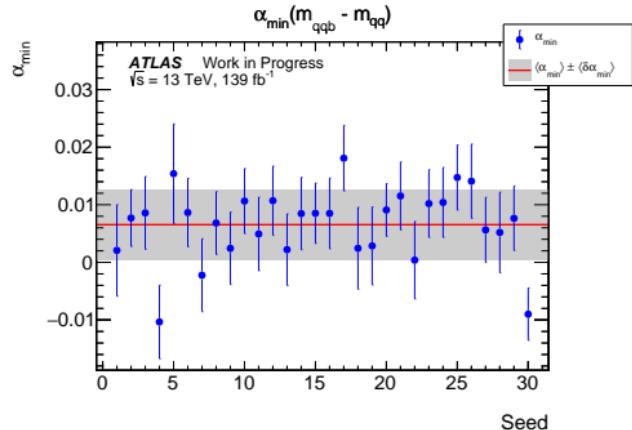
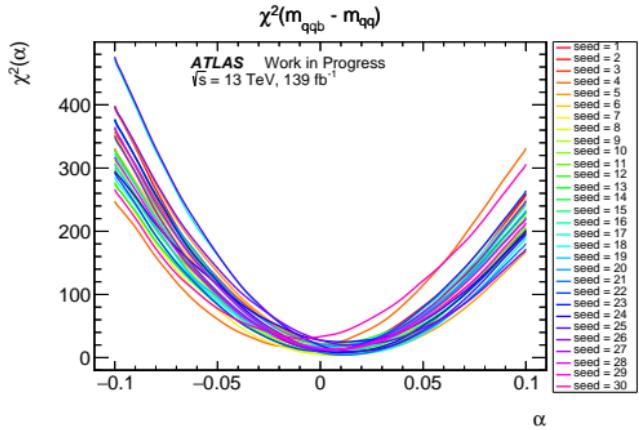
# $m_{qqb} - m_{qq}$ distribution Data vs Nominal MC sample before and after $b$ -JES corrections for $p_T^b \in [250, 500)$ GeV

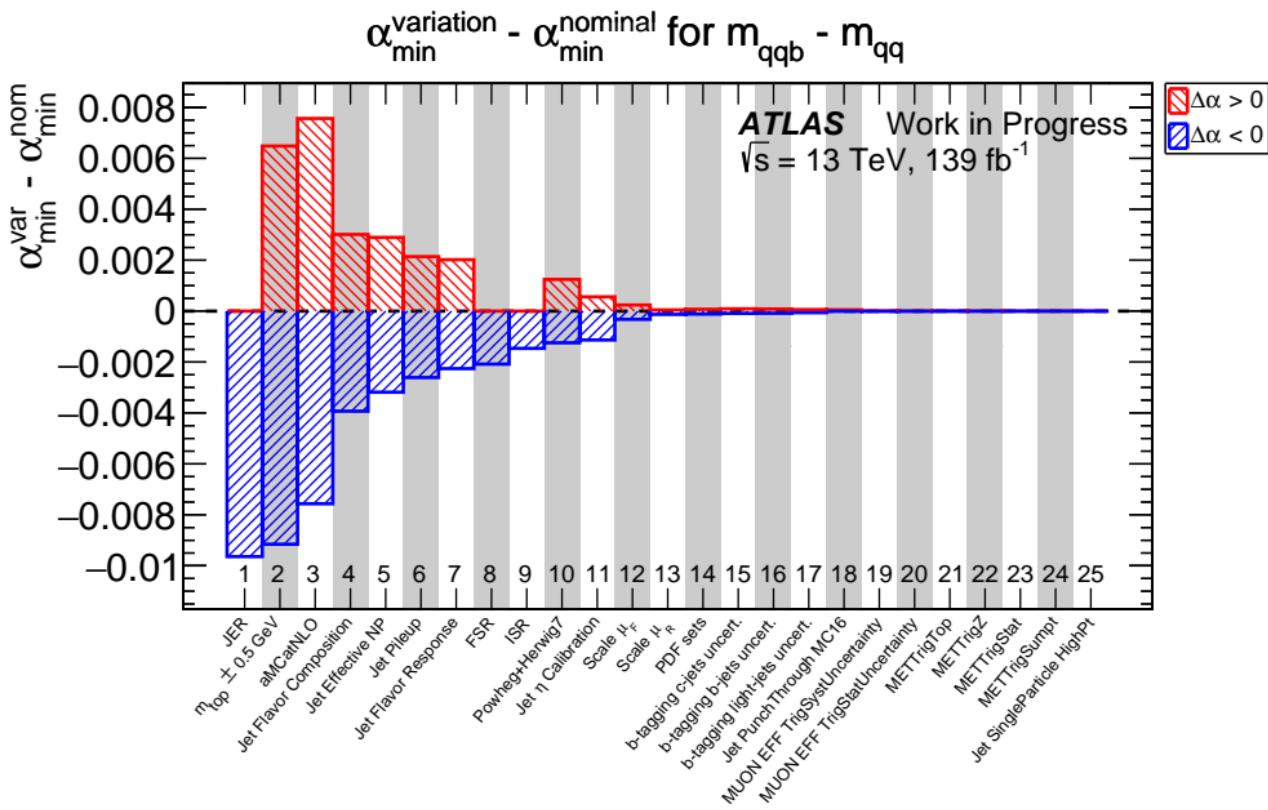
⑥ 30 / 35



$\alpha_{\min}^{\text{nominal}} \pm \delta\alpha^{\text{stat}}$  for Nominal FS  $m_{qqb} - m_{qq}$   
 for  $p_T^b \in [250, 500] \text{ GeV}$   
 (OP spline)

⑥ 31 / 35





# PseudoData test

PseudoData test:  $\alpha^{\text{true}} = 0.55\%$   
 $\alpha_{\text{min}}^{\text{nominal}} \pm \delta\alpha^{\text{stat}}$  for Nominal FS  $m_{qqb} - m_{qq}$   
(OP spline)

⑤ 34 / 35

