# $t\bar{t}$+HF theory predictions

Stefano Pozzorini

based on

T. Jezo, J. Lindert and S.P. [arXiv:1802.00426]

and HXSWG studies in collabration with

F. Siegert, M. V. Garzelli, T. Jezo, J. Krause, A. Kardos, J. Lindert,
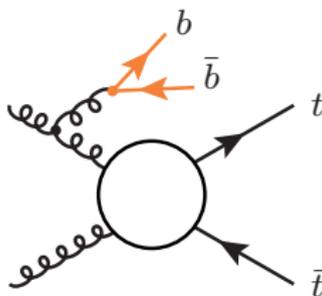R. Podskubka, C. Reuschle, M. Zaro

Higgs Toppings Workshop, Benasque, 29 May 2018

Universität
Zürich UZH

# Foreword



$t\bar{t}H(b\bar{b})$ **searches dominated by theory systematics of** $t\bar{t} + b$**-jet background**

- $t\bar{t} + b$-jet data help, but precise "extrapolation" to signal region calls for $t\bar{t} + b$-jet shape uncertainties at 10% level . . .

- variety of NLO MC tools available for $pp \to t\bar{t}b\bar{b}$ . . .

- . . . but $pp \to t\bar{t}b\bar{b}$ remains a nontrivial multi-particle multi-scale QCD process

- better understanding of its QCD dynamics and NLOPS theory systematics crucial for assessment of TH uncertainties

# Outline

# Option 1: NLOPS $t\bar{t}$ 5F (e.g. Powheg)

**$t\bar{t}b\bar{b}$ described through $t\bar{t}j$ tree MEs plus $g \to b\bar{b}$ shower splittings**



$gg \to t\bar{t}g$ + FS splittings

$\bar{b}g \to t\bar{t}\bar{b}$ + IS splittings
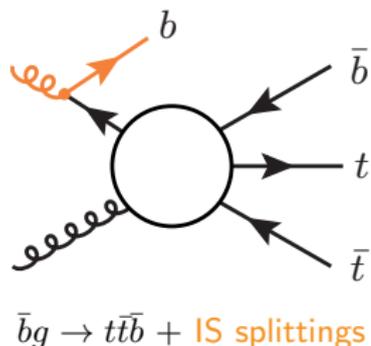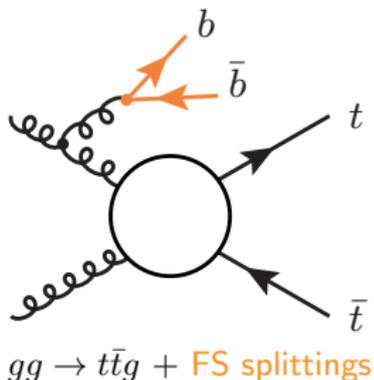
## Precision vs accuracy

- precision lower than LO but parton shower allows for accurate tuning to data
- residual uncertainties difficult to quantify

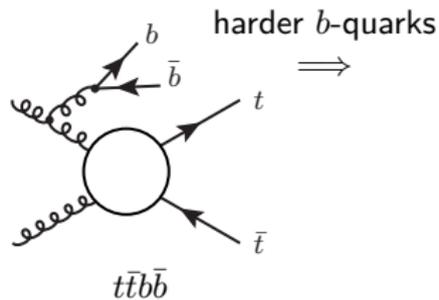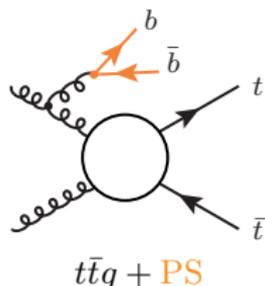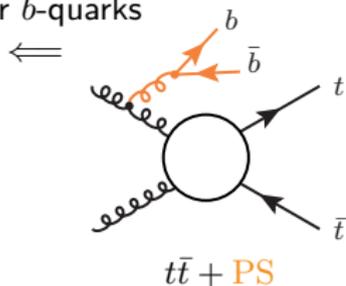## Calls for improved description based on $t\bar{t}b\bar{b}$ MEs

⇒ testable prediction with higher precision and more realistic uncertainties

⇒ possible tensions with data more instructive than tuning a non predictive MC!

# Option 2: (N)LO merging $t\bar{t} + 0, 1, 2$ jets 5F

$t\bar{t}b\bar{b}$ **described through** $t\bar{t} + 0, 1, 2$ **jet MEs and** $g \to b\bar{b}$ **shower splittings**



softer $b$-quarks $\Longleftarrow$

$t\bar{t} + \text{PS}$

$t\bar{t}g + \text{PS}$

$t\bar{t}b\bar{b}$

harder $b$-quarks $\Longrightarrow$
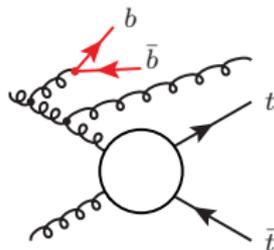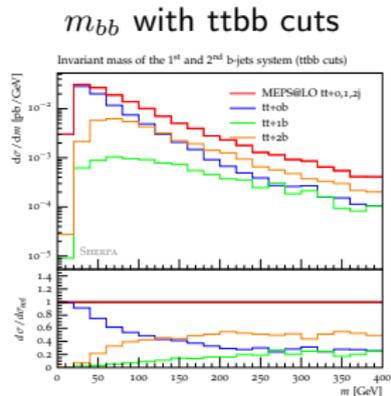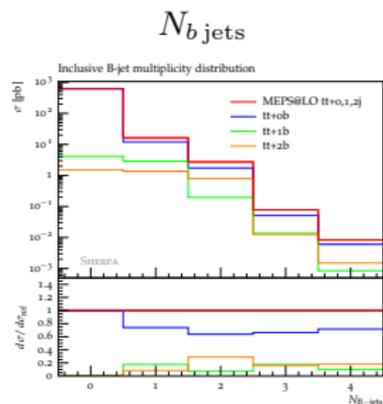
## Merging cut $Q_{\text{cut}}$

- separates ME regions ($k_T > Q_{\text{cut}}$) from shower regions ($k_T < Q_{\text{cut}}$)
- low $Q_{\text{cut}}$ maximises precision but can lead to prohibitive CPU cost at NLO
- in 5F scheme finite $Q_{\text{cut}}$ mandatory to avoid $g \to b\bar{b}$ singularity of MEs with $m_b = 0$

# $t\bar{t}b\bar{b}$ mostly from shower in $t\bar{t}$+multi-jet merging [1802.00426]

$t\bar{t} + 0, 1, 2$ **jet LO merging with** $Q_{\text{cut}} = 20\,\text{GeV}$



$N_{b\,\text{jets}}$

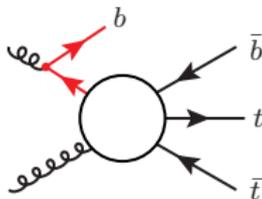

$m_{bb}$ with ttbb cuts



**Observables with $\geq 1$ additional b-jets**

- dominated by MEs with 2 light jets and no $b$-jets (up to $Q \sim 100\,\text{GeV}$)!
- due to the fact that $g \to b\bar{b}$ typically softer wrt 1st and 2nd splitting

$\Rightarrow$ **direct description in terms of $t\bar{t}b\bar{b}$ MEs seems preferable**

# Option 3: (N)LOPS $t\bar{t}b\bar{b}$ in 4F scheme
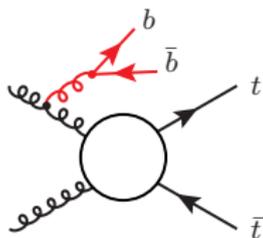


**4F $pp \to t\bar{t}b\bar{b}$ MEs with $m_b > 0$ at NLO+PS**

- describe full $b$-quark phase space

$\Rightarrow$ NLOPS accuracy for $t\bar{t} + 2\,b$-jet and $t\bar{t} + 1\,b$-jet observables! [Cascioli et al '13]

- 80% LO uncertainty reduced to 20–30% at NLO [Bredenstein et al. '09–'10; Bevilacqua et al. '10]

- include $b$-jet production from IS and FS $g \to b\bar{b}$ collinear splittings

# $t\bar{t}b\bar{b}$ dominated by FS $g \to b\bar{b}$ splittings [1802.00426]

### $t\bar{t}b\bar{b}$ **topologies with FS $g \to b\bar{b}$ splittings**

- dominant in full ttbb and ttb phase space
- notion of $g \to b\bar{b}$ splittings and IS/FS separation seems ill defined at large $\Delta R_{bb}$, $m_{bb}$, $p_{T,b}$ due to sizable interferences
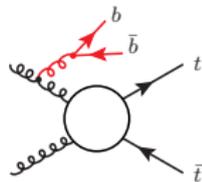
### $t\bar{t}b\bar{b}$ **topologies with IS $g \to b\bar{b}$ splittings**

- mostly clearly subdominant (no need for 5F scheme resummation)
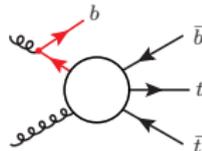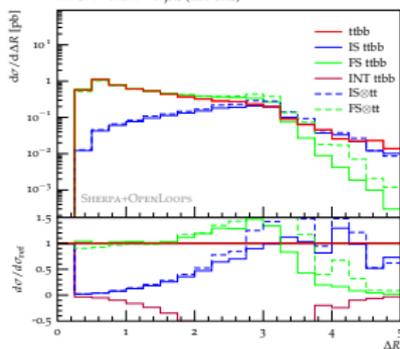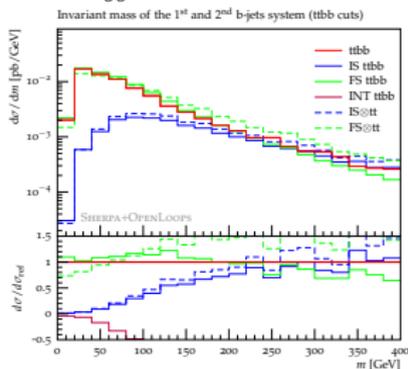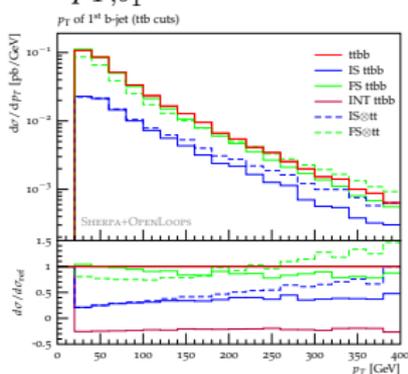


supports choice of 4F scheme with $m_b > 0$ and no $b$-quark PDF
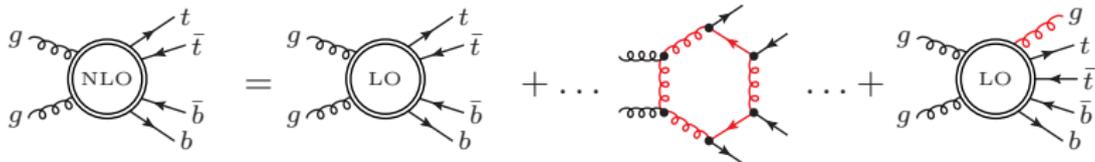
## Nontrivial features of $pp \to t\bar{t}b\bar{b}$ at NLO

- 34 LO diagrams and $> 1000$ NLO diagrams
- 6 external coloured partons
- large uncertainty from $\sigma_{t\bar{t}b\bar{b}} \propto \alpha_S^4(\mu_R)$
- multiple scales from 5 to 500 GeV (gap between $b\bar{b}$ and $t\bar{t}$ systems)

## Nontrivial NLOPS issues

- matching/shower effects up to 30% in Higgs region
- due to double $g \to b\bar{b}$ splittings [Cascioli et al '13]

  $\Rightarrow$     crucial to understand $g \to b\bar{b}$ splittings and matching+shower uncertainties

Invariant mass of the 1st and 2nd b-jets system (ttbb cuts)



$p_T$ of 1st light-jet (ttbb cuts)
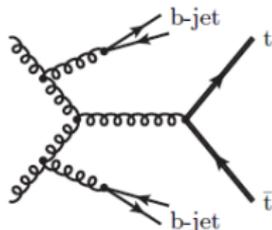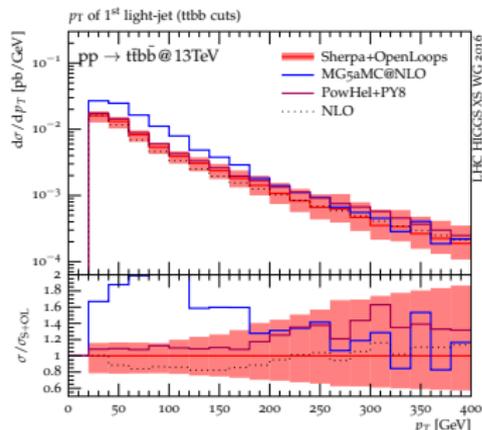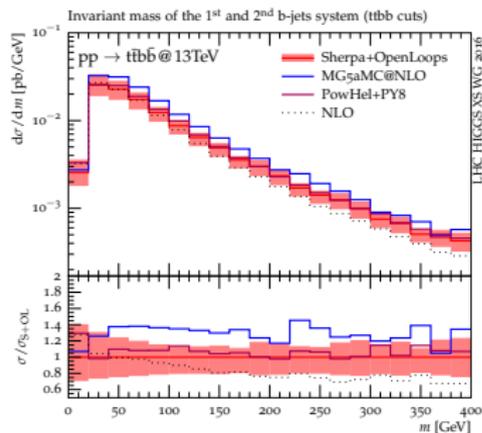
## MG5aMC@NLO+PY8 (4F) vs Sherpa (4F)

- 40% NLOPS/NLO enhancement of $t\bar{t} + 2$b XS in MG5

- related to sizeable enhancement of NLO radiation at $p_T \sim 100$ GeV

- sensitive to resummation scale (scalup) in MG5

  **Question (still open):** large uncertainty or not?!

## PowHel+PY8 (5F) vs Sherpa (4F)

- much better agreement

- but 5F scheme in Powhel not appropriate for collinear $g \to b\bar{b}$ splittings (ad-hoc cuts)

  **Question:** small theory uncertainty or accidental?

**Recent news**

- now MG5 supports $H_T$-based resummation scale $\mu_Q = f(\xi) H_T / 2$

- two new Powheg 4F $t\bar{t}b\bar{b}$ generators [arXiv:1709.06915, arXiv:1802.00426]

- new MatchBox+Herwig 4F $t\bar{t}b\bar{b}$ generator

$\Rightarrow$ ongoing campaign of $t\bar{t}b\bar{b}$ MC studies within HXSWG

# Outline

**Original 5F $t\bar{t}b\bar{b}$ POWHEL upgraded to 4F scheme with $m_b > 0$**

$\Rightarrow$ now applicable to entire $b$-jet phase space!

$\Rightarrow$ consistent comparison against other 4F $t\bar{t}b\bar{b}$ generators possible

**Comparison to 8 TeV CMS data (20/fb)** (agreement but statistics still low)

# PowhegBox+OpenLoops [Jezo, Lindert, Moretti, S.P 1802.00426]
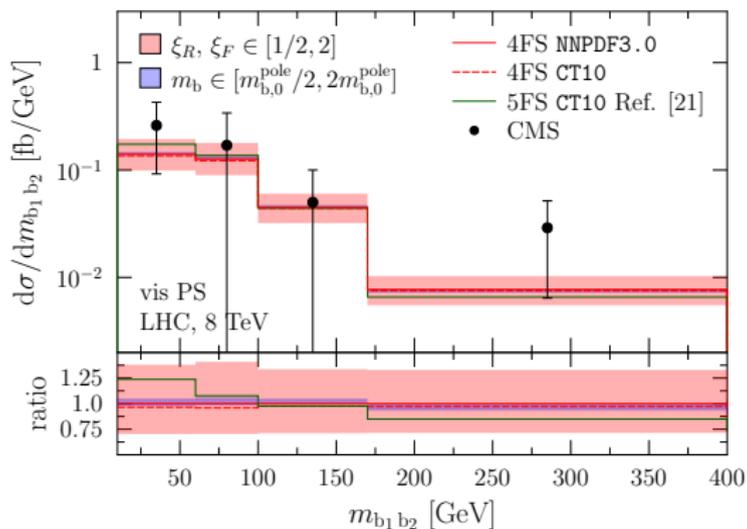
## Differences wrt Powhel

- PowhegBox–RES framework and OpenLoops MEs ($\Rightarrow$ fast)
- spin-correlated top decays
- separation of soft/hard radiation extended to FS radiation

## Restriction of soft (resummation) region

$$k_T \lesssim h_{\mathrm{damp}} = H_T/2 \qquad \text{and} \qquad \frac{R_{\mathrm{soft}}(\Phi_R)}{B(\Phi_B) \otimes K_{\mathrm{soft/coll}}(\Phi_{\mathrm{rad}})} < h_{\mathrm{bzd}} = 2$$

$\Rightarrow$ avoids resummation in regions $p_{T,b} < k_T < h_{\mathrm{damp}}$ where soft/coll factorisation not fulfilled

$\Rightarrow$ guarantees high stability wrt $h_{\mathrm{damp}}$ variations
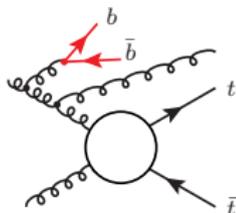


## Very large NLO $K$-factor

- Typically in $t\bar{t}b\bar{b}$ literature: $\sigma_{\mathrm{NLO}}/\sigma_{\mathrm{LO}} \sim 1.2$ based on LO inputs for $\sigma_{\mathrm{LO}}$
- Using NLO inputs throughout (like in NLOPS local $K$-factor): $\sigma_{\mathrm{NLO}}/\sigma_{\mathrm{LO}} \sim 1.9$

  $\Rightarrow$ origin of large correction and perturbative convergence to be understood!

# NLOPS vs NLO Powheg $t\bar{t}b\bar{b}$ predictions [1802.00426]
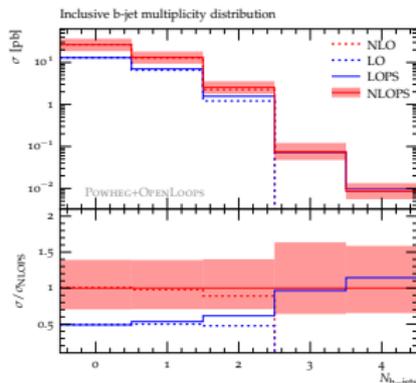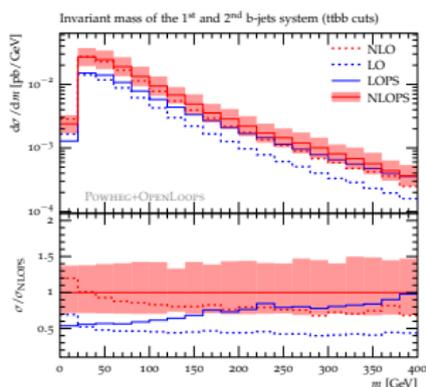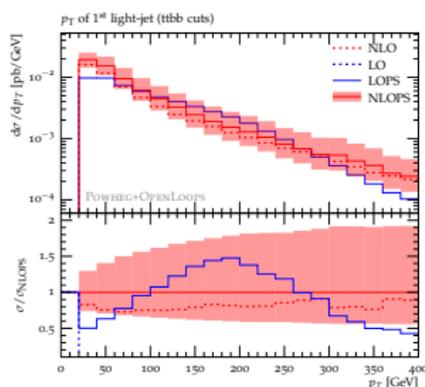


**Moderate NLOPS/NLO corrections within scale-variation bands**

- $10\%$ for $\sigma_{tt+2b}$
- 20–30% at $m_{bb} \sim 100\,\mathrm{GeV}$ (double splittings)

**Shape of light-jet $p_T$**

- NLOPS quite similar to fixed-order NLO
- LOPS/NLOPS suggests that PY8 overestimates radiation with $p_T \sim 200\,\mathrm{GeV}$
  $\Rightarrow$ related to effects in MG5+PY8?

$N_b$  $M_{b_1 b_2}$  $p_{T,j_1}$
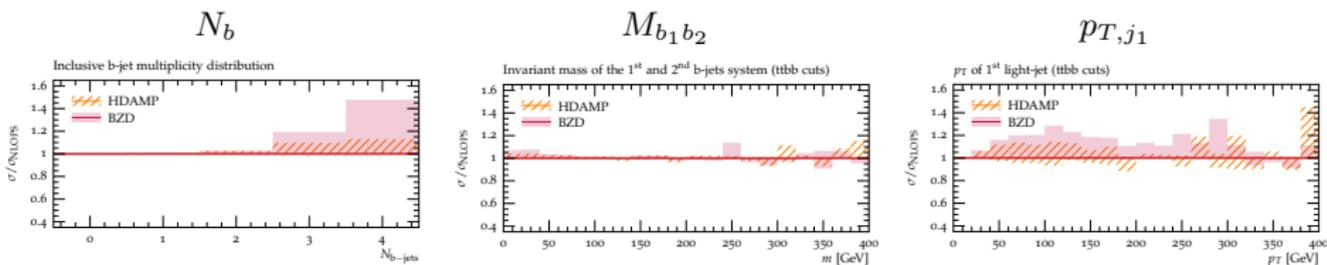
**Dependence on matching scales** ($h_{\mathrm{damp}} = H_T/4, H_T/2, H_T, 1.5m_t$ and $h_{\mathrm{bzd}} = 2, 5, 10$)

- inclusive $t\bar{t} + b$-jet observables remarkably stable at <span style="color:red">percent level</span>

- jet-$p_T$ spectrum stable at 10–20% level

<span style="color:blue">High stability guaranteed by $h_{\mathrm{bzd}}$ restriction of Powheg resummation</span>

$N_b$ — Inclusive b-jet multiplicity distribution

$M_{b_1 b_2}$ — Invariant mass of the 1st and 2nd b-jets system (ttbb cuts)

$p_{T,j_1}$ — $p_T$ of 1st light-jet (ttbb cuts)
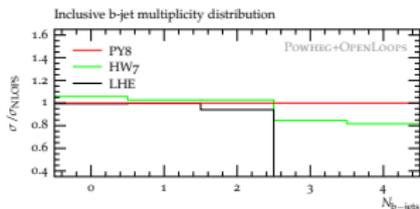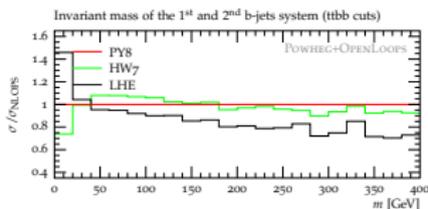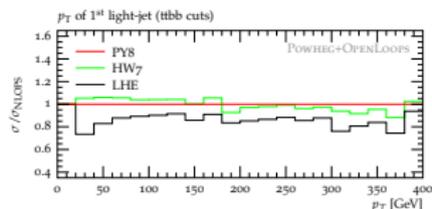
## LHEs vs NLOPS

- Significant NLOPS effects in $m_{bb}$ (double splittings)

## Powheg+Pythia8 vs Powheg+Herwig

- dependence on choice of parton shower $<<$ QCD scale dependence

High stability thanks to independence of $1^{st}$ Powheg emission wrt parton shower

**Variations of scalup, $g \to b\bar{b}$ splittings and choice of $\alpha_S$ in PY8**
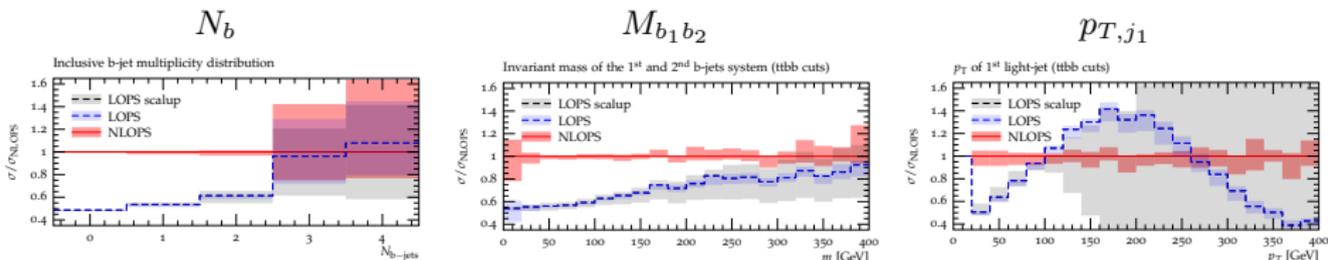


**LOPS variations** dominated by factor-2 variations of scalup=$H_T/2$

- strongly reduced at NLO, especially for jet-$p_T$

**NLOPS variations** dominated by $g \to b\bar{b}$ modeling and $\alpha_S$ variations in PY8

- remarkably small NLOPS uncertainties

- again because $1^{st}$ Powheg radiation independent of shower

- double-splitting effects stable wrt variations of $g \to b\bar{b}$ in PY8

# Powheg $t\bar{t}b\bar{b}$ vs other (N)LOPS tools  I

## $t\bar{t}b\bar{b}$ **Powheg+PY8 vs Sherpa**

- Powheg with all matching+shower (no scale) uncertainties vs nominal Sherpa
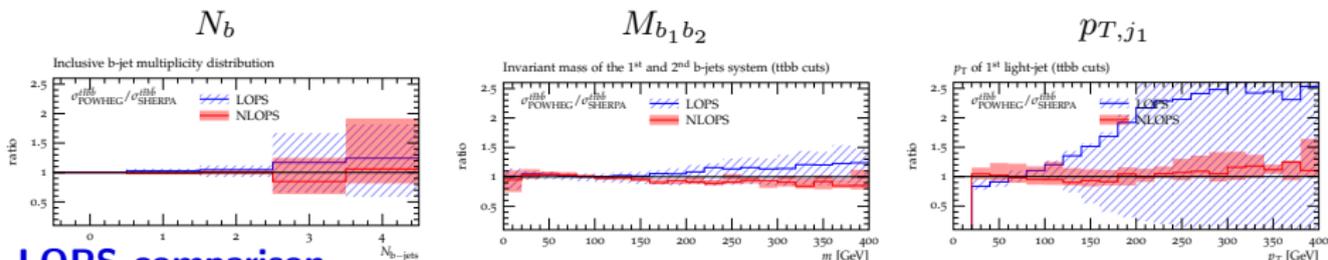


**LOPS comparison**

- radiation of Sherpa shower less hard than PY8
- ⇒ likely to have beneficial effects on MC@NLO matching in Sherpa

  (while Powheg matching does not suffer from PY8 excess)

**NLOPS comparison**

- NLOPS uncertainties and Powheg/Sherpa differences clearly reduced at NLO
  ⇒ different showers and matching methods agree at better than 10% level!
- using Sherpa 2.2 recoil scheme yields more significant (but still moderate) differences (see later)

# Powheg $t\bar{t}b\bar{b}$ vs other (N)LOPS tools II

**Powheg+PY8:** $t\bar{t}b\bar{b}$ **vs inclusive** $t\bar{t}$

- all matching+shower (no QCD scale) uncertainties only for $t\bar{t}$ generator
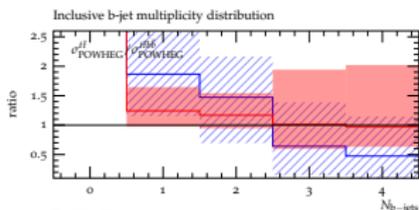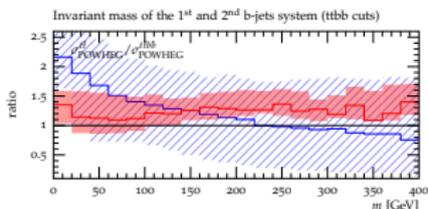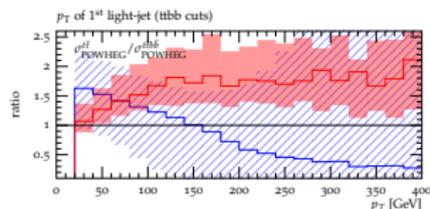


**LOPS**

- uncertainties beyond factor 2
- large differences in $N_b$, $m_{bb}$ and jet-$p_T$

**NLOPS**

- differences strongly reduced at NLOPS ("Powheg miracle")
- $t\bar{t}$ exceeds $t\bar{t}b\bar{b}$ by only $\sim 20\%$ in $N_b$ and $m_{bb}$ shape is OK
  (100% excess in the jet-$p_T$ tail)

Motivation for $t\bar{t}b\bar{b}$ NLOPS lies in smaller (see previous plots) and better defined theory uncertainties

# Outline

# Idea and goals

**Main goal:** $t\bar{t}b\bar{b}$ **theory uncertainty estimates for** $t\bar{t}H(b\bar{b})$

- comparing MC against data not sufficient since MC needed for extrapolations

- comparing different NLOPS MC tools is (only) the starting point . . .

- . . . we need intrinsic uncertainty of individual MC $\Rightarrow$ should explain MC differences

**Roadmap**

- optimal choice of settings for coherent (apple-to-apple) comparison

- variations to isolate/rank uncertainties of fixed-order, matching and shower origin

- identify+understand leading sources of MC differences/uncertainties

- TH uncertainties recommendations (for $t\bar{t}H$ searches and $t\bar{t}b\bar{b}$ measurements)

$\Rightarrow$ Uniform framework for $t\bar{t} + b$-jets TH systematics in ATLAS+CMS

**Status** https://twiki.cern.ch/twiki/bin/view/LHCPhysics/ProposalWwbbbb

- preliminary results (also limited by MC statistics)

- no conclusion but good progress and interesting open questions/hypotheses

# Rivet analysis and tools

**New $t\bar{t}b\bar{b}$ Rivet analyses for HXSWG comparisons** (J. Lindert, T. Jezo)

- with stable tops (extended wrt YR4): 60 observables (ttb and ttbb cuts)

- with dileptonic top decays: 84 observables (WW4b and WW3b cuts)

- public results ⇒ benchmarks for MC simulation in ATLAS/CMS and theory

**5 MC tools, 2 NLOPS methods, 3 showers, 10 contributing authors**

| Tool | MC@NLO | Powheg | Pythia 8.2 | Herwig 7.1.2 | Sherpa2.2.4 | MC contacts |
|------|:------:|:------:|:----------:|:------------:|:-----------:|-------------|
| SHERPA2.2+OPENLOOPS | x | | | | x | F. Siegert, J. Krause |
| MG5_AMC@NLO | x | | x | x | | M. Zaro |
| MATCHBOX+OPENLOOPS | x | | | x | | C. Reuschle, R. Posdkubka |
| POWHEG+HELAC | | x | x | x | | M.V. Garzelli, A. Kardos |
| POWHEGBOX+OPENLOOPS | | x | x | x | | T. Jezo, J. Lindert |
| | 3 | 2 | 3 | 4 | 1 | |

# Comparison 6 MC with top decays (WW4b cuts)

**Inputs** (here and in the following)

- same inputs as in HXSWG YR4 (but default shower tunes here)
- limited statistics
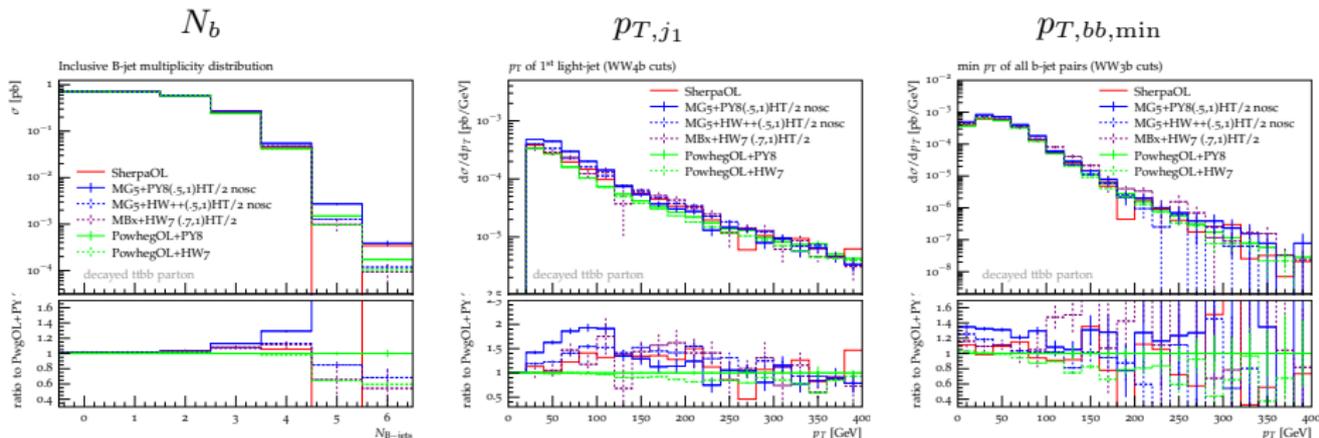


**Features observed with stable tops confirmed**

- now 20% spread of $WW + 4b$ XS and factor-2 in jet spectrum

(present studies focussed back on stable $t\bar{t}b\bar{b}$)

# How to interpret MC comparison

**Different types of dependencies**

- parton shower choice and tune
- choice of matching method and related parameters
- perturbative apsects (QCD scales, PDFs)

**Idea**

- disentangle perturbative/matching/shower dependence
- in particular: isolate irreducible differences due to matching method

Ongoing discussion on matching settings tricky but instructive. . .

# MC@NLO vs Powheg matching (how to compare?)

**Splitting of radiation**: $S$-events (soft/singular) and $H$-events (hard/remnant)

$$\mathrm{d}\sigma_S = \mathrm{d}\Phi_B \bar{B}(\Phi_B) \left[ \Delta(t_{\mathrm{IR}}) + \Delta(k_T) \frac{R_{\mathrm{soft}}(\Phi_R)}{B(\Phi_B)} \Phi_{\mathrm{rad}} \right] \quad \mathrm{d}\sigma_H = \mathrm{d}\Phi_R \left[ R(\Phi_R) - R_{\mathrm{soft}}(\Phi_R) \right]$$

**Soft radiation integrated out in** $\bar{B}$ $\Rightarrow$ $\bar{B}/B =$ local $K$-factor

$$\bar{B}(\Phi_B) = B(\Phi_B) + V(\Phi_B) + \int \mathrm{d}\Phi_{\mathrm{rad}} R_{\mathrm{soft}}(\Phi_B, \Phi_{\mathrm{rad}})$$

**Powheg vs MC@NLO difference only in** $R_{\mathrm{soft}}$

Powheg: $R_{\mathrm{soft}}(\Phi_R) = R(\Phi_R) g_{\mathrm{soft}}(\Phi_{\mathrm{rad}}, h_{\mathrm{damp}})$      matrix element

MC@NLO: $R_{\mathrm{soft}}(\Phi_R) = B(\Phi_B) \otimes K_{\mathrm{shower}}(\Phi_{\mathrm{rad}}) g_{\mathrm{soft}}(\Phi_{\mathrm{rad}}, \mu_Q)$    parton shower

**Soft profile** $g_{\mathrm{soft}}(\Phi_{\mathrm{rad}}, \mu_Q)$
- restricts $R_{\mathrm{soft}}$ below $\mu_Q$ (resummation scale), e.g. $\theta(\mu_Q^2 - k_T^2)$

$\Rightarrow$ **ideal choice for consistent comparison**: $h_{\mathrm{damp}} = \mu_Q$ and same $g_{\mathrm{soft}}$    ...?

# Choice of shower starting scale (separation of $S/H$ events)

**Implementation of** scalup **based on recommendation** $\mu_Q = h_{\mathrm{damp}} = H_T/2$

- different choices of $S$- and $H$-events ($\mathrm{scalup}_S$, $\mathrm{scalup}_H$)
- distributed with different profiles: see ($\mathrm{scalup}_{\mathrm{mean}}, \mathrm{scalup}_{\mathrm{max}}$) in the table

| MC | method | $\mathrm{scalup}_S$ | $\mathrm{scalup}_H$ | comments |
|---|---|---|---|---|
| MG5 | MC@NLO | $(.55, 1)\mu_Q$ | $(1, 1)\mu_Q$ | |
| MatchBox | MC@NLO | $(.7, 1)\mu_Q$ | $(.7, 1)\mu_Q$ | |
| Sherpa | MC@NLO | $(1, 1)\mu_Q$ | 1st $k_T$ or $\mu_Q$ | |
| PowhegOL | Powheg | ME | 1st $k_T$ | no scalup dependence |
| Powhel | Powheg | ME | 1st $k_T$ | no scalup dependence |

## Comments

- MG5 authors "synchronise" $\mathrm{scalup}_{\mathrm{max}} \Rightarrow \mathrm{scalup}_{\mathrm{mean}} \sim H_T/4$ (no consensus so far)

- important: study $\mathrm{scalup}_S$ and $\mathrm{scalup}_H$ dependence separately

# Choice of PDFs+$\alpha_S$ (studies by M.Zaro see next breakout session)

## Employed PDFs+$\alpha_S$ values

|  | label | scheme and PDFs | $\alpha_S(M_Z)$ | $\alpha_S/\alpha_S^{4F}$ |
|---|---|---|---|---|
| Fixed-order NLO | 4F NLO | NNPDF30_nlo_af_0118 | 0.112 | 1 |
| HW7 showering | 5F LO PS | MMHT14 LO (HW tune) | 0.1262 | 1.125 |
| PY8 showering | 5F LO PS | NNPDF2.3 QCD+QED LO (Monash) | 0.1365 | 1.219 |

## MC implementation

|  |  |  | 1st emission | | higher emissions |
|---|---|---|---|---|---|
| MC | method | $\bar{B}$ | $H$-event | $S$-event |  |
| PowhegOL | Powheg | 4F NLO | 4F NLO | 4F NLO | 5F LOPS |
| Powhel | Powheg | 4F NLO | 4F NLO | 4F NLO | 5F LOPS |
| Sherpa | MC@NLO | 4F NLO | 4F NLO | 4F NLO | 4F NLO |
| MG5 | MC@NLO | 4F NLO | 4F NLO | 5F LOPS | 5F LOPS |
| MatchBox | MC@NLO | 4F NLO | 4F NLO | 5F LOPS | 5F LOPS |

## Open questions

- $S$-radiation with 5F LO PDF $\Rightarrow$ jet-$p_T$ spectrum with too large $\alpha_S$
- use PS tunes with reasonable $\alpha_S$?
- consistency issues also from $b \to bg$ ISR in 1st emission?

# Comparison with stable tops (ttbb cuts)



$N_b$      $M_{b_1 b_2}$      $\Delta R_{b_1 b_2}$
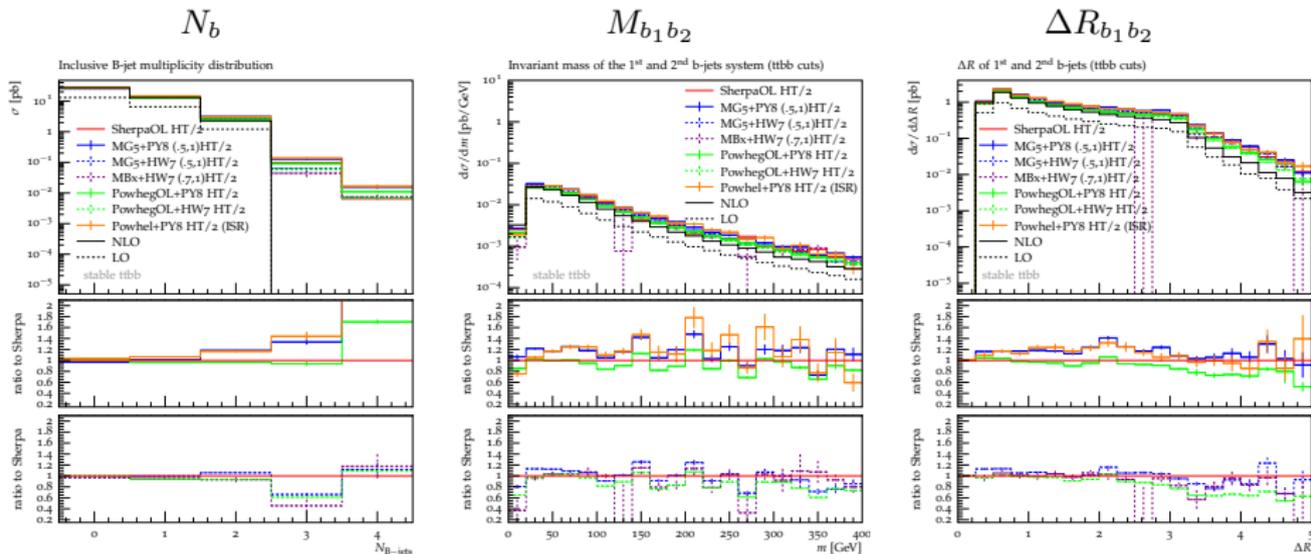
- NLOPS+PY8 and Sherpa predictions (1st ratio):
  PowhegOL $\simeq$ Sherpa while MG5+PY8 $\simeq$ Powhel+PY8 (lack of FS $h_{\mathrm{damp}}$?)

- NLOPS+HW7 and Sherpa predictions closer to each other
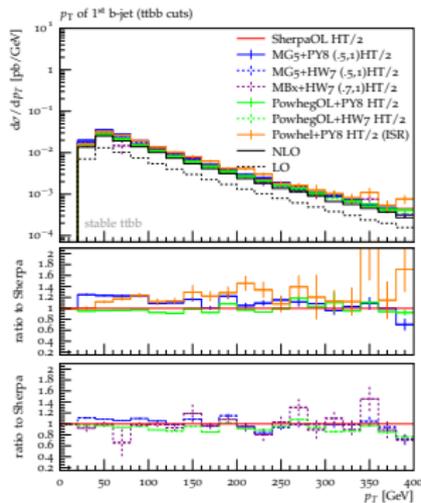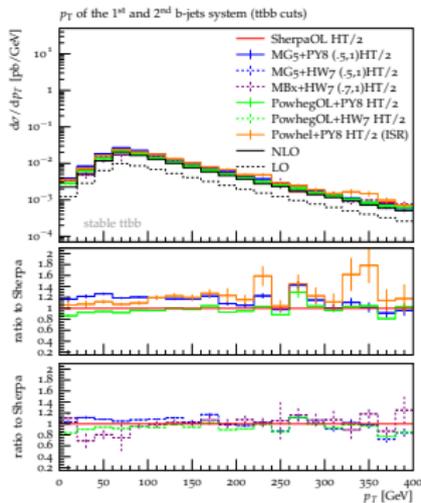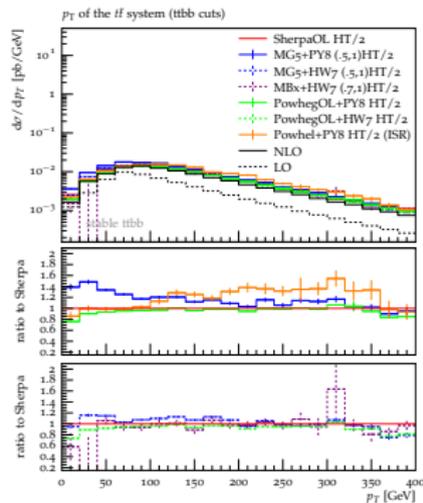
- in general good agreement in shape of inclusive observables
- different shapes of MG5+PY8 and Powhel+PY8 in $p_T$ of $t\bar{t}$

# Spectrum of ligh-jet radiation
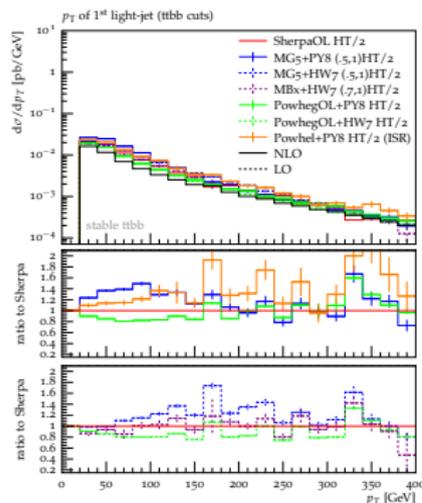


**Familiar picture of MC differences**

- Normalisation changed a bit, but YR4-like shape differences persist

**Current interpretation** (hypothesis)

- large local $K$-factor applied to soft events distorts jet-$p_T$ spectrum
- no effect on total XS, but related recoil can shift $b$-jet $p_T$

$\Rightarrow$ migrations between different $N_b$ bins

$\Rightarrow$ enhancement of $t\bar{t}+2b$ cross section

**Effect depends on relative importance of $S/H$ contributions**

- can be ehnanced in MG5+PY8 due to
  - scalup implementation = soft profile (theory uncertainty)
  - overestimate of soft radiation by PY8 (in part unphysical)

    To be discussed!

# PY8/HW7 (in MG5 and Powheg+OL)



$N_b$
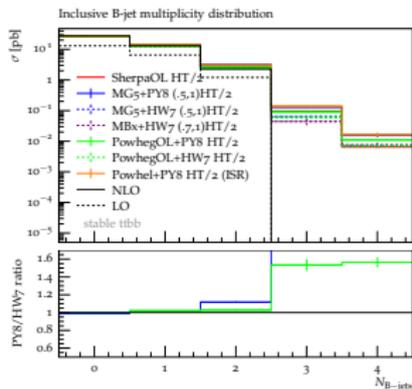
$M_{b_1 b_2}$
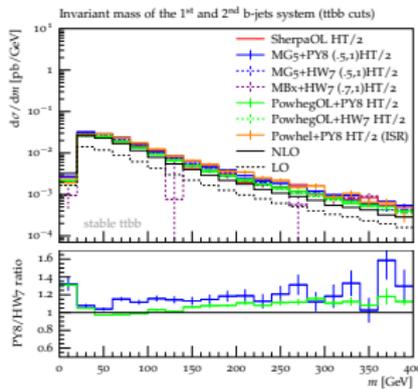
$p_{T,j_1}$
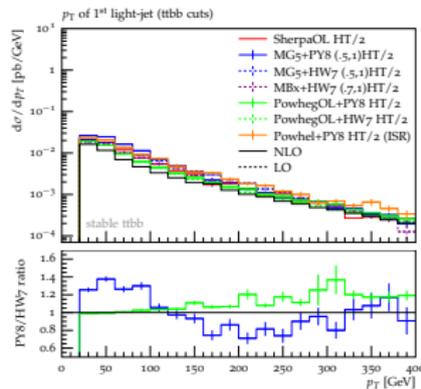
- PowhegOpenLoops more stable
- aMC_MG5+Herwig closer to the other MC tools

# Relative effect of double $g \to b\bar{b}$ splittings (for 5 MCs)
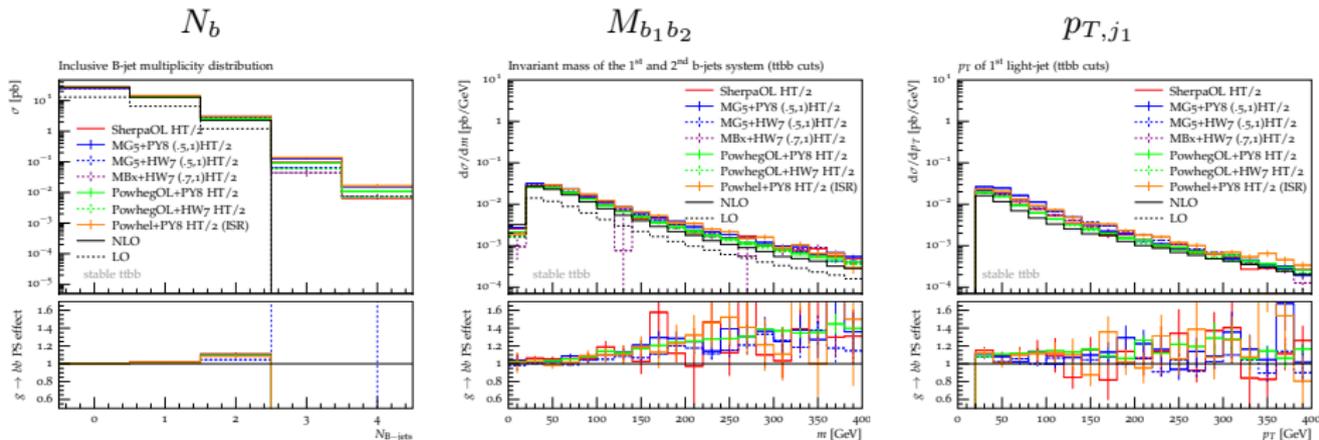


- does not explain sizable shape differences

# NLOPS/NLO and $\mu_Q$,hdamp dependence



$N_b$ — Inclusive B-jet multiplicity distribution

$M_{b_1 b_2}$ — Invariant mass of the 1st and 2nd b-jets system (ttbb cuts)

$p_{T,j_1}$ — $p_T$ of 1st light-jet (ttbb cuts)
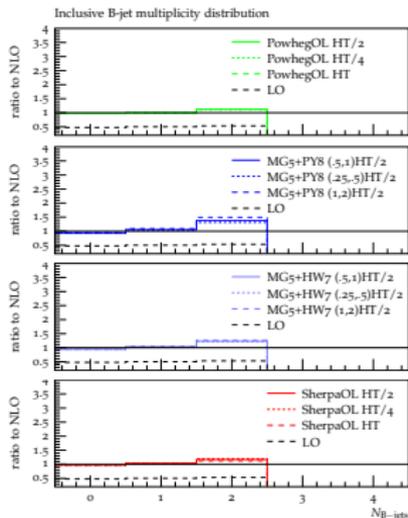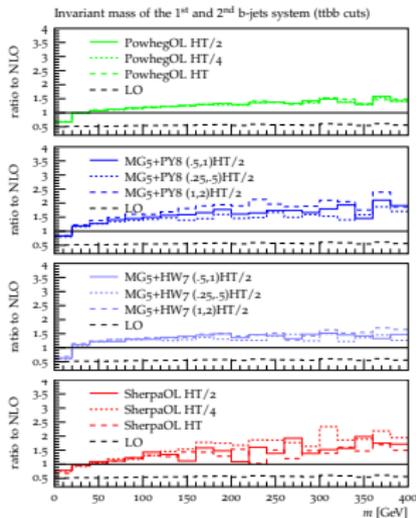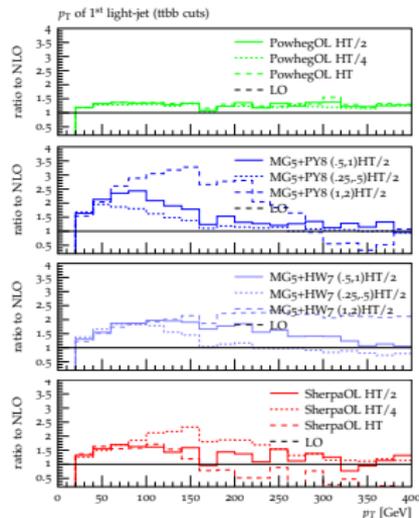
- Powheg very stable
- similar trend but different $\mu_Q$ dependence in MG5+PY8, MG5+HW and Sherpa (new recoil scheme)

# Conclusions and Outlook

**8 different 4F $t\bar{t}b\bar{b}$ NLOPS tools**

- opportunity to disentangle and understand sources of TH uncertainty

- contribution of MC authors crucial: many thanks!

**Important goals still ahead of us**

- achieve consensus on interpretation of tool differences/uncertainties

- address perturbative uncertainties, in particular for shapes (see backup slides)

$\Rightarrow$ thorough recommendations for $t\bar{t} + b$-jet theory uncertainty estimates

# Discussion session

# Missing large logarithms from $g \to b\bar{b}$ fragmentation? I

**Probability of $g \to b\bar{b}$ in a hard gluon jet** [Mangano and Nason, PLB 285 (1992)]

$$\rho(Q^2, K^2) = \int_{2m_b}^{Q} \mathrm{d}K \, P_{g \to b\bar{b}}(K) \times n_g(Q^2, K^2)$$

$g \to b\bar{b}$ splitting probability at virtuality $K^2 = m_{b\bar{b}}^2$

$$P_{g \to b\bar{b}}(K) = \frac{\alpha_S(K^2)}{3\pi K} \left(1 + \frac{2m_b^2}{K^2}\right) \sqrt{1 - \frac{4m_b^2}{K^2}}$$

Multiplicity of gluons with virtuality $K^2$ in hard-gluon jet with $p_T = Q$

$$n_g(Q^2, K^2) = \left[\frac{\ln(Q^2/\Lambda^2)}{\ln(K^2/\Lambda^2)}\right]^a \cosh\left[\sqrt{\frac{2C_A}{\pi b}} \left(\sqrt{\ln(Q^2/\Lambda^2)} - \sqrt{\ln(K^2/\Lambda^2)}\right)\right]$$

Perturbative expansion in $\alpha_S = \alpha_S(Q^2) = \left[b \ln(Q^2/\Lambda^2)\right]^{-1}$

$$\frac{\mathrm{d}\rho(Q^2, K^2)}{\mathrm{d}K} = \frac{\mathrm{d}\rho(Q^2, K^2)}{\mathrm{d}K}\bigg|_{\mathrm{LO}} \times \left[1 + \alpha_S \left(C_1 L^2 + \dots\right) + \alpha_S^2 \left(C_2 L^4 + \dots\right) + \dots\right]$$

with double logarithms $L = \ln(K^2/Q^2)$

**Distribution $\mathrm{d}\rho(Q^2, K^2)/\mathrm{d}K$ at LO, NLO and NNLO for $Q = 50\,\text{GeV}$**
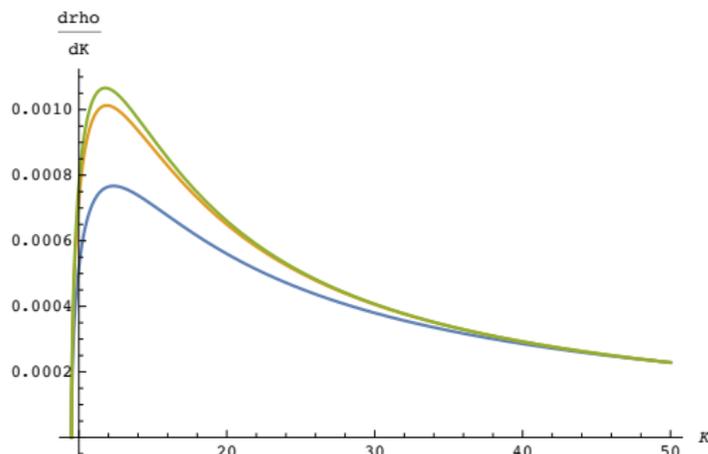


- higher-order effects well approximated by NLO
- peak close to threshold ($K \gtrsim 2m_b$) but long tail

**Total $g \to b\bar{b}$ probability**

| $Q[\text{GeV}]$ | LO | NLO | NNLO | NLO/LO | NNLO/NLO |
|---|---|---|---|---|---|
| 50 | 2.08% | 2.44% | 2.51% | 1.17 | 1.03 |
| 100 | 2.73% | 3.50% | 3.71% | 1.28 | 1.06 |
| 500 | 3.84% | 6.06% | 7.05% | 1.59 | 1.16 |

# Changes wrt YR4 HXSWG studies (errata corrige)
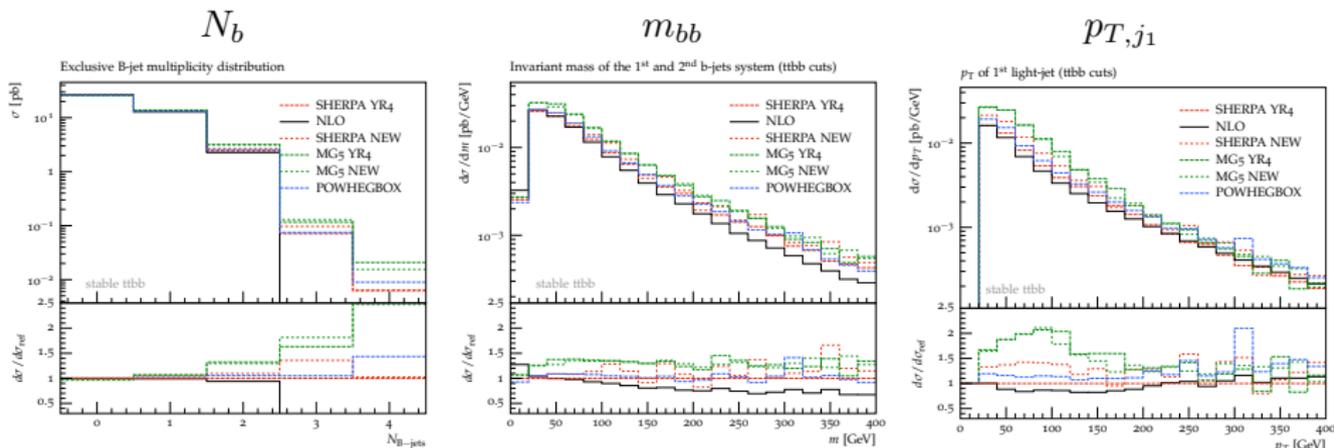


## Bottom line

- MG5+PY8 did not change significantly (in spite of $\hat{s} \to H_T$ based scalup)
- Sherpa moved in the direction of MG5+PY8
  - $+35\%$ in the jet-$p_T$ spectrum (but little impact on inclusive shapes)
  - due to new default recoil scheme (for 2nd and higher emissions)
  - and other changes (to be clarified in detail)

# Comparison of different showers and recoil schemes



$p_{T,j_1}$
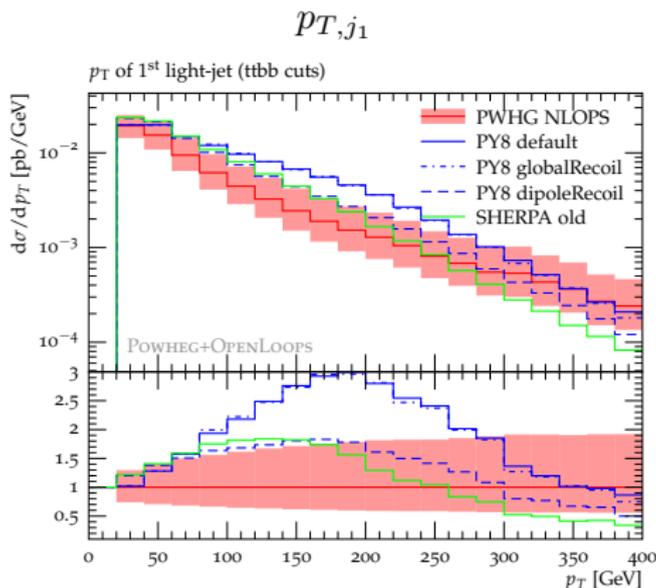
**LOPS with different showers and recoil schemes** (overall NLO normalisation)

- large MC effects may be due to the recoil effects of QCD radiation on $b$-jets
- PY8 dipole recoil scheme more consistent with NLOPS radiation spectrum, hower not supported in MC@NLO matching
- also Sherpa (YR4 recoil scheme) more consistent with NLOPS

# Distortion of ligh-jet radiation spectrum (reloaded)



$p_{T,j_1}$
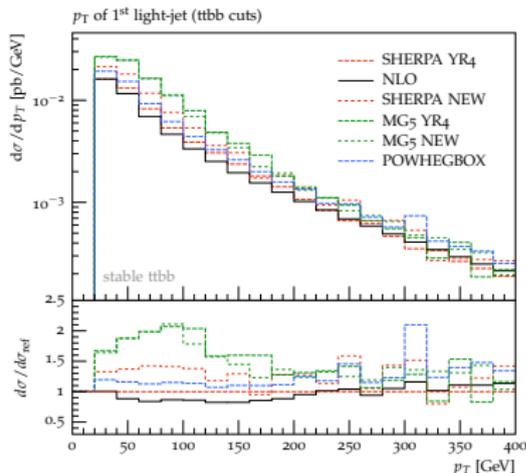
**Current interpretation** (hypothesis)

- large local $K$-factor applied to S-events distorts jet-$p_T$ spectrum
- no effect on total XS, but related recoil can shift $b$-jet $p_T$
- $\Rightarrow$ migrations between different $N_b$ bins
- $\Rightarrow$ enhancement of $t\bar{t}+2b$ cross section
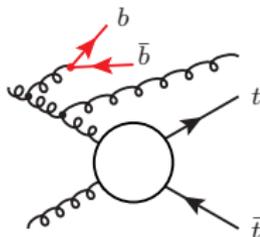
**Effect depends on relative importance of $S/H$ contributions**

# Natural separation approach

**Compare hardness of $g \to b\bar{b}$ splitting to $p_T$ of NLO radiation**

- $p_T(\text{jet}) < k_T(g \to b\bar{b}) \quad \Rightarrow \quad$ soft
- $p_T(\text{jet}) > k_T(g \to b\bar{b}) \quad \Rightarrow \quad$ hard
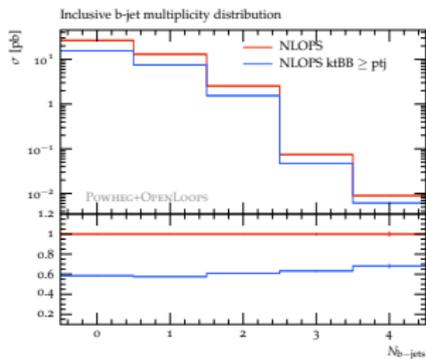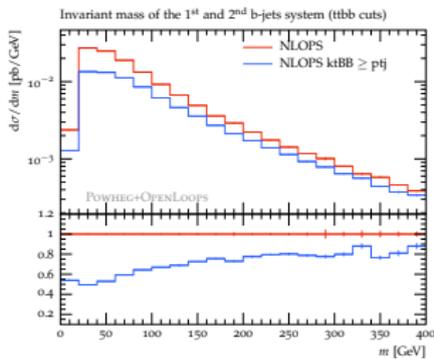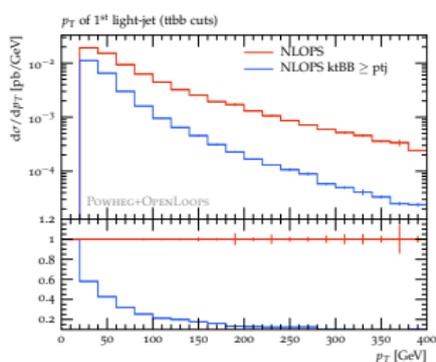


- roughly $1/2$ of $t\bar{t}b\bar{b}$ cross section involves a jet harder than $b$-jet system
- it is natural to treat it as H-contribution in NLOPS framework

# Comparison of S/H separation in various tools
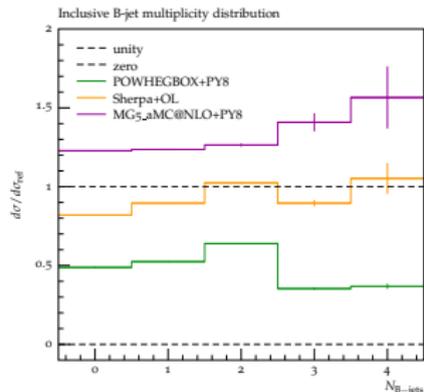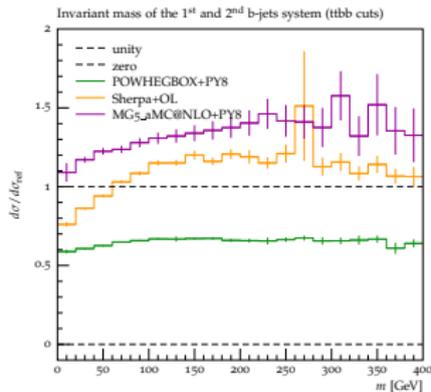


$N_b$         $m_{bb}$         $p_{T,j_1}$
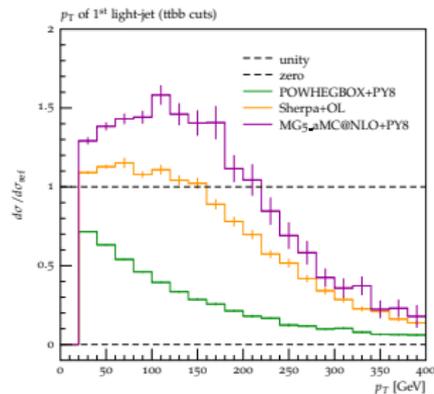
- Powhe: S-contribution $\sim 50\%$, i.e. comparable to $k_T(b\bar{b}) < p_T(\text{jet})$ (as a result of $h_{\text{bzd}}$)

- MC@NLO tools: in Sherpa and especially in MG5+PY8, $S$-contribution overestimates full XS and must be compensated by negative H-contribution

# Interplay of $K \gg 1$ and negative $\sigma_H$ in MC@NLO

$$\frac{\mathrm{d}\sigma}{\mathrm{d}\Phi_B} = \underbrace{\bar{B}_{\mathrm{soft}}(\Phi_B)\left[\Delta(t_{\mathrm{IR}}) + \Delta(k_T)\mathcal{K}_{\mathrm{soft}}(\Phi_1)\,\mathrm{d}\Phi_1\right]}_{\mathrm{S-events\ (LHE\times shower)}} + \underbrace{\left[R(\Phi_R) - B(\Phi_B)\mathcal{K}_{\mathrm{soft}}(\Phi_1)\right]\mathrm{d}\Phi_1}_{\mathrm{H-events}}$$

**Soft radiation approximated by paron shower in the soft region** $k_T \lesssim \mu_Q$

$$R(\Phi_R) \longrightarrow B(\Phi)\mathcal{K}_{\mathrm{soft}}(\Phi_1) = B(\Phi)K_{\mathrm{shower}}(\Phi_1)\,g_{\mathrm{soft}}(\Phi_1, \mu_Q)$$

**and integrated out in**

$$\bar{B}_{\mathrm{soft}}(\Phi_B) = B(\Phi_B) + V(\Phi_B) + B(\Phi_B)\int \mathrm{d}\Phi_1 \mathcal{K}_{\mathrm{soft}}(\Phi_1)$$

**Matching distorted by** $K$**-factor** $\bar{B}_{\mathrm{soft}}/B \gtrsim 2$ **and** $\underbrace{(R - B\mathcal{K}_{\mathrm{soft}})}_{\mathrm{H-weight}} < 0$

$$\frac{\mathrm{d}\sigma}{\mathrm{d}\Phi_B \mathrm{d}\Phi_1} = R + \underbrace{\left[\frac{\bar{B}_{\mathrm{soft}}}{B}\Delta - 1\right]}_{\gtrsim\,100\%\mathrm{distortion}} B\mathcal{K}_{\mathrm{soft}} = \underbrace{\left(\frac{\bar{B}_{\mathrm{soft}}}{B}\Delta\right)}_{\mathrm{max\ resummation}} R + \underbrace{\left[\frac{\bar{B}_{\mathrm{soft}}}{B}\Delta - 1\right]}_{\gtrsim\,100\%}\underbrace{(B\mathcal{K}_{\mathrm{soft}} - R)}_{>0}$$

⇒ strongly enhanced positive correction beyond "max resummation": unphysical?

# Backup slides

## Generic jet content of $t\bar{t} + 0, 1, 2\,$jet LO merging
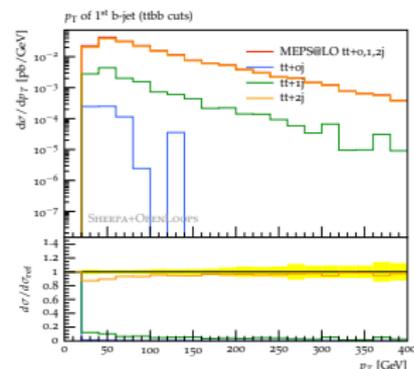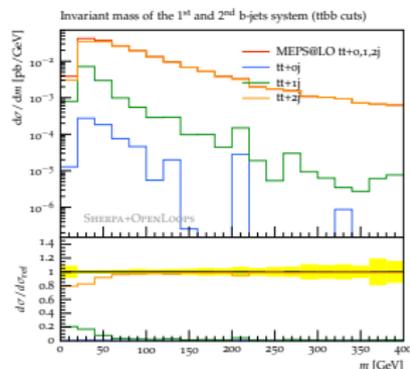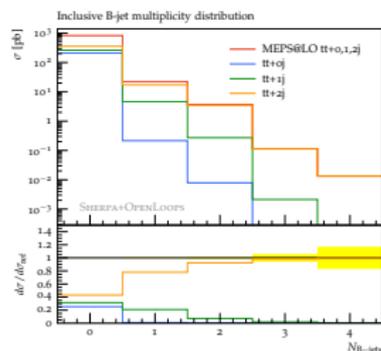


## Cross sections and distributions with $\geq 1$ additional b-jets

- dominated by $t\bar{t} + 2\,$jet MEs . . .

# Dependence on resummation scale $\mu_Q$ (shortly after YR4)



Invariant mass of the 1st and 2nd b-jets system (ttbb cuts)

$p_T$ of 1st light-jet (ttbb cuts)

- Sherpa+OpenLoops
- MG5aMC@NLO
- S+Ol $_{\sqrt{\hat{s}}/2}$
- MG5 $_{\sqrt{\hat{s}}/2}$
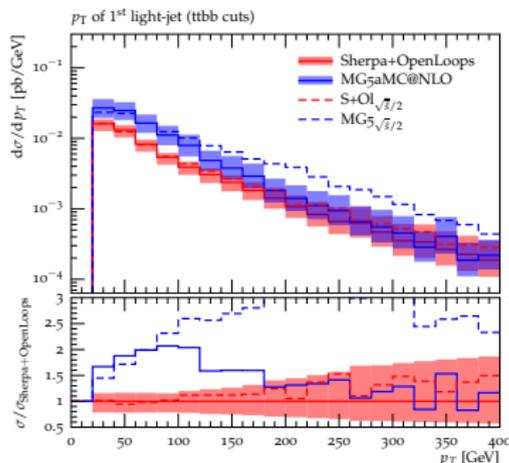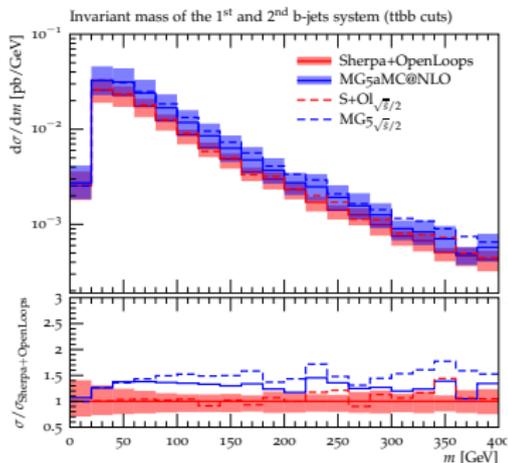
**Nominal MG5_aMC and Sherpa+OpenLoops predictions in YR4**

- MG5_aMC supports only* $\mu_Q = f(\xi)\sqrt{\hat{s}} \Rightarrow$ smearing function restricted to $0.1 < f(\xi) < 0.25$ to mimic recommended $\mu_Q = H_T/2$ implemented in Sherpa

**$\mu_Q$ variations enhance the discrepancy**

- $\mu_Q = \sqrt{\hat{s}}/2$ in Sherpa to mimic MG5_aMC default choice $0.1 < f(\xi) < 1$
- strong $\mu_Q$-sensitivity of MG5_aMC $\Rightarrow$ much more pronounced deviations

* Ongoing studies with new MG5 version supporting $H_T/2$. See talks by Zaro & Neu.

# NLOPS $t\bar{t}b\bar{b}$ 4F with SHERPA+OPENLOOPS [Cascioli et al '13]

## Convergence of 4F scheme but unexpected MC@NLO enhancement

|  | $ttb$ | $ttbb$ | $ttbb\,(m_{bb} > 100)$ |
|---|---|---|---|
| $\sigma_{\mathrm{LO}}[\mathrm{fb}]$ | $2644^{+71\%+14\%}_{-38\%-11\%}$ | $463.3^{+66\%+15\%}_{-36\%-12\%}$ | $123.4^{+63\%+17\%}_{-35\%-13\%}$ |
| $\sigma_{\mathrm{NLO}}[\mathrm{fb}]$ | $3296^{+34\%+5.6\%}_{-25\%-4.2\%}$ | $560^{+29\%+5.4\%}_{-24\%-4.8\%}$ | $141.8^{+26\%+6.5\%}_{-22\%-4.6\%}$ |
| $\sigma_{\mathrm{NLO}}/\sigma_{\mathrm{LO}}$ | 1.25 | 1.21 | 1.15 |
| $\sigma_{\mathrm{MC@NLO}}[\mathrm{fb}]$ | $3313^{+32\%+3.9\%}_{-25\%-2.9\%}$ | $600^{+24\%+2.0\%}_{-22\%-2.1\%}$ | $181^{+20\%+8.1\%}_{-20\%-6.0\%}$ |
| $\sigma_{\mathrm{MC@NLO}}/\sigma_{\mathrm{NLO}}$ | 1.01 | 1.07 | 1.28 |

## Large enhancement ($\sim$30%) in Higgs region from double $g \to b\bar{b}$ splittings



Mass of first two b-jets (ttbb cuts)
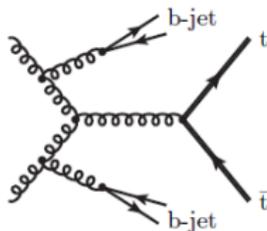
## One $g \to b\bar{b}$ splitting from PS

$\Rightarrow$ TH uncertainties related to matching, shower and 4F/5F schemes crucial!
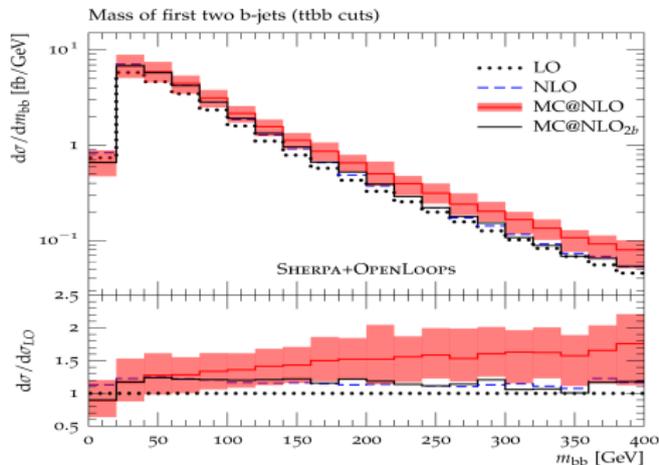
# Setup for $t\bar{t}b\bar{b}$ 4F Powheg+OpenLoops predictions [arXiv:1802.00426]

**Aspects identical to HXSWG YR4**

- NNPDF30_NLO_as_0118_nf_4
- $\mu_R = (E_{T,t}E_{T,\bar{t}}E_{T,b}E_{T,\bar{b}})^{1/4}$
- $\mu_F = H_T/2$,
- $h_{damp} = H_T/2$,

**Matching scale variations**

- $h_{damp} = H_T/4, H_T/2, H_T, 1.5m_t$
- $h_{bzd} = 2, 5, 10$

**Shower and PDFs for showering**

- A14 Pythia tune with $\alpha_S(M_Z) = 0.127$
- NNPDF2.3 LO 5F PDFs

# NLOPS subtleties for multi-scale problems [1802.00426]

**Matching based on factorisation of $S$-radiation wrt hard $t\bar{t}b\bar{b}$ process**

$$R_{\text{soft}}(\Phi_R) \simeq B(\Phi_B) \otimes K_{\text{soft/coll}}(\Phi_{\text{rad}}) \qquad \text{for} \quad k_T < h_{\text{damp}} \sim m_t$$

**What about radiation with $p_{T,b} < k_T < h_{\text{damp}}$? Soft or hard?**

- $t\bar{t}b\bar{b}$ factorisation can fail and factorising hard $t\bar{t}+$jet subprocess can be more appropriate

- example: hard jet radiation in the direction of $b\bar{b}$ system

    - $\Phi_B \to \Phi_R$ FKS mappings $\Rightarrow b\bar{b}$ system absorbs jet recoil and becomes much softer

    - $R(\Phi_R)$ enhancement that violates $ttbb$ factorisation

- similar issues expected also in MC@NLO matching
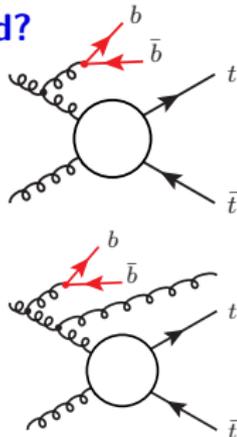
**Powheg "safety" system:** resummation only if $R_{\text{soft}} < h_{\text{bzd}} \times B \otimes K_{\text{soft/coll}}$

$$g_{\text{soft}}(\Phi_{\text{rad}}, h_{\text{damp}}, h_{\text{bzd}}) = \frac{h_{\text{damp}}^2}{h_{\text{damp}}^2 + k_T^2} \; \theta\Big(h_{\text{bzd}}B(\Phi_B) \otimes K_{\text{soft/coll}}(\Phi_{\text{rad}}) - R(\Phi_R)\Big)$$

$\Rightarrow$ high stability wrt $h_{\text{damp}}$ variations

# More observables with top decays (limited statistics)



$\Delta R_{bb,\min}$
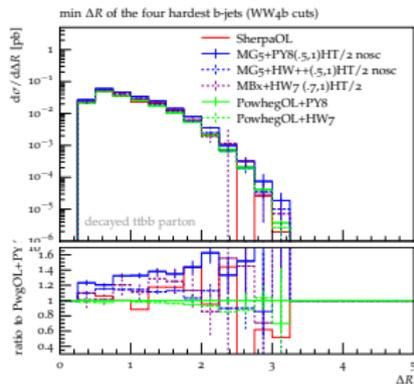
$M_{bb,\min}$
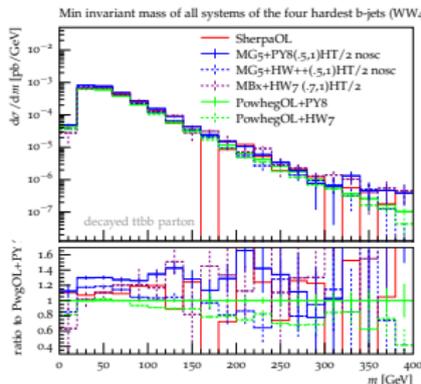
$M_{bb,\max}$
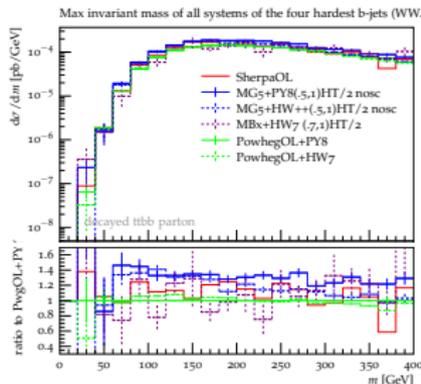
# Spectrum of ligh-jet radiation



$p_{T,j_1}$
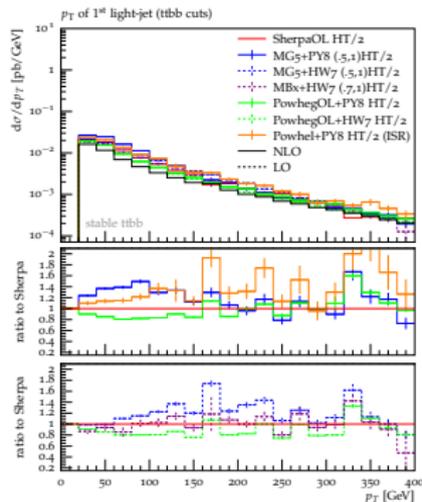
$p_{T,t\bar{t}b\bar{b}}$ (ttbb cuts)
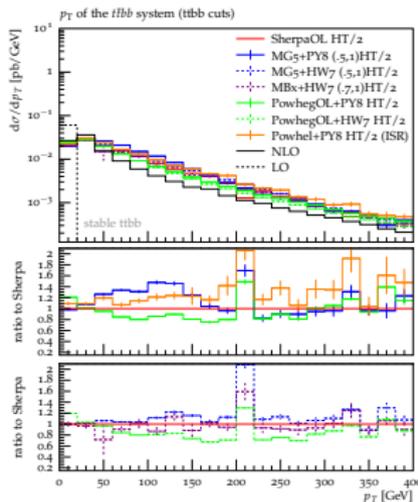
$p_{T,t\bar{t}b\bar{b}}$ (no cuts)
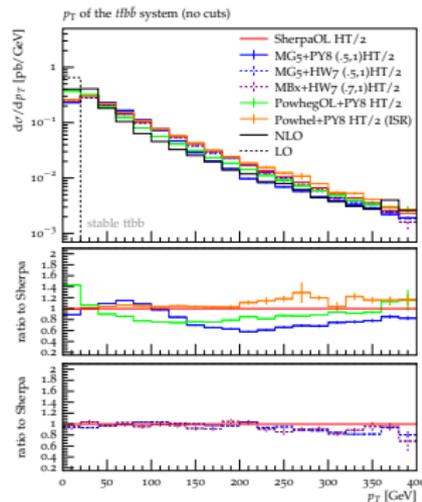
- familiar picture in spectrum of radiation/recoil spectrum
- normalisation changes but shape different persists if $b$-jet cuts removed
- hypotheis: distortion of jet-spectrum due to large local $K$-factor and different $S/H$ separation
- ⇒ jet recoil transferred to $b$-jets ⇒ $b$-jet bin migrations

# Hadronisation effects in $t\bar{t}b\bar{b}$ MC comparisons

**Motivation of theory studies w.o. top decays and hadornisation**

- top decays are trivial (well understood EW interactions) but render the analysis of $b$-quark production in $WWb\bar{b}b\bar{b}$ final states quite cumbersome

- switching off top decays is very useful in order to investigate the QCD dynamics of $b$-production in $pp \to t\bar{t}b\bar{b}$ (which dominates TH uncertainties!)

- since top quarks carry SU(3) charge, also hadronisation needs to be switched off

**Possible bias of MC comparisions?**

- switching off hadronisation could bias comparisons of different showers (Pythia, Sherpa, Herwig) due to dependencies on unphysical dependences (e.g. IR cutoff)

- irrelevant for Powheg+PY8 vs MG5+PY8 comparison (same shower)

- for Sherpa vs MG5+PY8 we have assessed this effect comparing LOPS simulations of $H + b$-jet production (as proxy of $t\bar{t}b\bar{b}$ production) finding non-negligible but rather small hadronisation effects wrt the observed differences in $t\bar{t}b\bar{b}$ production

see https://twiki.cern.ch/twiki/bin/view/LHCPhysics/LOpphHadronisation

## Why NLO Matching for $t\bar{t}b\bar{b}$ Production in 4F (and not 5F) Scheme
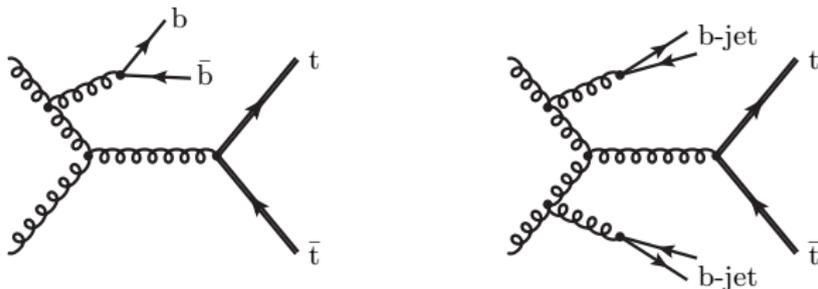


**5F scheme ($m_b = 0$):** $t\bar{t}b\bar{b}$ MEs cannot describe collinear $g \to b\bar{b}$ splittings

$\Rightarrow$ *inclusive* $t\bar{t}$+b-jets simulation (quite important for exp. analyses!) requires $t\bar{t}$g+PS, i.e. $t\bar{t}+ \leq 2$ jets NLO merging [Höche, Krauss, Maierhöfer, S. P. , Schönherr, Siegert '14 ]

see talk by F. Krauss

**4F scheme ($m_b > 0$):** $t\bar{t}b\bar{b}$ MEs cover full b-quark phase space

$\Rightarrow$ MC@NLO $t\bar{t}b\bar{b}$ sufficient for inclusive $t\bar{t}$+b-jets simulation

- access to **new $t\bar{t} + 2$b-jets production mechanism** wrt 5F scheme: double collinear $g \to b\bar{b}$ splittings (surprisingly important impact on $t\bar{t}H(b\bar{b})$ analysis!)

**Sherpa's MC@NLO master formula** [Frixione, Webber '02; Höche, Krauss, Schönherr, Siegert '11]

$$\sigma_n^{\mathrm{MC@NLO}} = \int d\Phi_n \Bigg[ \mathcal{B}(\Phi_n) + \mathcal{V}(\Phi_n) + \mathcal{B}(\Phi_n) \otimes \mathcal{I} \Bigg] \left\{ \Delta(\mu_Q^2, t_{\mathrm{IR}}) + \int_{t_0}^{\mu_Q^2} d\Phi_1 \mathcal{S}(\Phi_1) \, \Delta(\mu_Q^2, t) \right\}$$

$$+ \int d\Phi_{n+1} \Bigg[ \mathcal{R}(\Phi_{n+1}) - \mathcal{B}(\Phi_n) \otimes \mathcal{S}(\Phi_1) \Bigg]$$

- shower resummation effectively acts starting from $\mathcal{O}(\alpha_s^2)$, and iterated emissions yield fully realistic events
- inclusive observables with $n$ $(n+1)$ particles preserve NLO (LO) accuracy

**Factorisation and Resummation scales** (available phase space for QCD emission)

$$\mu_{\mathrm{F}} = \mu_Q = \frac{1}{2}(E_{\mathrm{T,t}} + E_{\mathrm{T,\bar{t}}})$$

**Scale choice crucial due to $\alpha_S^4(\mu^2)$ dependence** (80% LO variation)

- widely separated scales $m_b \leq Q_{ij} \lesssim m_{t\bar{t}b\bar{b}}$ can generate huge logs
- CKKW inspired scale adapts to b-jet $p_{\mathrm{T}}$ and guarantees good pert. convergence
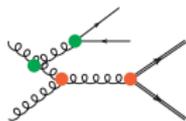

$\mu_{\mathrm{R}}^4 = E_{\mathrm{T,t}} E_{\mathrm{T,\bar{t}}} E_{\mathrm{T,b}} E_{\mathrm{T,\bar{b}}} \quad \Rightarrow \quad \alpha_S^4(\mu_{\mathrm{R}}^2) = \alpha_S(E_{\mathrm{T,t}}^2)\alpha_S(E_{\mathrm{T,\bar{t}}}^2)\alpha_S(E_{\mathrm{T,b}}^2)\alpha_S(E_{\mathrm{T,\bar{b}}}^2)$

## Approach proposed in YR4

**NLOPS 4F $t\bar{t}b\bar{b}$ sample**

- can be applied in its full phase space (no generation cuts)
- $\Rightarrow$ inclusive description of $t\bar{t}+ \geq 1b$-quarks
- includes also contributions corresponding to $gb \rightarrow t\bar{t}b$ in the 5F scheme

**Inclusive $t\bar{t} + X$ sample**

- needs to be restricted to $t\bar{t} + 0\, b$-quarks to avoid double counting
- $\Rightarrow$ veto events containing $b$-quarks not arising from showered top decays or MPI or UE

**Possible implementations**

- $t\bar{t} + X$ and $t\bar{t}b\bar{b}$ samples independent samples
- reweighting of $t\bar{t} + X$ sample through $t\bar{t}b\bar{b}$ in the $t\bar{t}+ \geq 1b$-quarks region

## Refinement for region of small $p_{T,b}$

**Caveat**

- $t\bar{t}b\bar{b}$ sample yields (small) contribution to $t\bar{t} + 0\,b$-jet categories of EXP analysis
- $t\bar{t} + 0\,b$-jet categories (dominated by $t\bar{t}$+gluons/light-quarks) can bias $t\bar{t}b\bar{b}$ fit
- ⇒ preferable to restrict $t\bar{t}b\bar{b}$ to $t\bar{t} + b$-jet categories

**Proposal: smooth matching of $t\bar{t} + X$ and $t\bar{t}b\bar{b}$ samples**

- using smearing function of leading b-jet $p_T$, such as

$$\xi(p_{T,b}) = \begin{cases} 0 & \equiv \text{ pure } t\bar{t} + 0b & \text{for} \quad p_{T,b} < p_{T,\min} \\ \frac{1}{2}\left[1 - \cos\left(\pi \frac{p_{T,b} - p_{T,\min}}{p_{T,\max} - p_{T,\min}}\right)\right] & & \text{for} \quad p_{T,\min} < p_{T,b} < p_{T,\max} \\ 1 & \equiv \text{ pure } t\bar{t} + \geq 1b & \text{for} \quad p_{T,b} > p_{T,\max} \end{cases}$$

- with transition region in the vicinity of experimental $b$-jet threshold,
  e.g. $[p_{T,\min}, p_{T,\max}] = [15, 25]\,\text{GeV}$
- same matching procedure should be used in ATLAS and CMS for a transparent
  comparison and combination of EXP results

## Scale choices (YR4) and uncertainties (no proposal yet)

**Factorisation ($\mu_Q$) and resummation ($\mu_Q$) scales**
$$E_{T_i} = \sqrt{m_i^2 + p_{T,i}^2}$$

$$\mu_F = \mu_Q = \frac{H_T}{2} = \frac{1}{2} \sum_{i=t,\bar{t},b,\bar{b}} E_{T,i}$$

$\mu_Q \equiv$ shower starting scale is a free paramater in MC@NLO (not in Powheg)

**CKKW-like (softer) renormalisation scale**

$$\mu_R = \mu_{\mathrm{CKKW}} = \prod_{i=t,\bar{t},b,\bar{b}} E_{T,i}^{1/4}$$

**Scale variations (leading uncertainty)** $\sim$20-30%

- factor-2 variations of $\mu_R$ and $\mu_F$ $\Leftrightarrow$ normalisation
- "kinematic" variations of $\mu_R, \mu_F, \mu_Q$ $\Leftrightarrow$ shape
- variations of $\mu_Q$ in MC@NLO and $h_{\mathrm{damp}}$ in Powheg $\Leftrightarrow$ NLOPS matching

**Other variations**

- PDF variations (only few percent)
- shower variations: tune variations, shower recoil scheme, ...

# Correlation of TH uncertainties between categories

## Categories

- $t\bar{t}h(b\bar{b})$ analyses based on simultaneous fit of MC to data in various categories with different # of light- and $b$-jets
- correlations crucial to constrain background in signal region (with multiple $b$-jets)

## Between $t\bar{t}$+light-jet and $t\bar{t} + b$-jet categories

- uncertainties should be uncorrelated

## Between sub-categories (e.g. $ttb$, $ttbb$, $ttB$)

- uncertainties should be correlated

  **Motivation**: independent shower, matching and ME variations account for different types of uncertainties (e.g. related to collinear $g \to b\bar{b}$ splittings or hard b-production) $\Rightarrow$ no need of separate categories with uncorrelated uncertainties

**Scale variations for shape (not for normalisation) uncertainties**

**Consider (aggressive but not fully unreasonable)** *kinematic distortions* **of** $\mu_R, \mu_F, \mu_Q$ using various combinations of the variables

$$\mu_{\text{CMMPS}} = \prod_{i=t,\bar{t},b\bar{b}} E_{T,i}^{1/4}, \qquad m_{b\bar{b}}, \qquad H_{T,b(t)} = E_{T,b(t)} + E_{T,\bar{b}(\bar{t})}, \qquad H_T = H_{T,t} + H_{T,b}$$

| Scale | default | glo-HT | glo-Mt | glo-soft | R-Mbb | R-HTb | R-HTt | Q-CMMPS | Q-Mt |
|---|---|---|---|---|---|---|---|---|---|
| $\mu_R$ | $\mu_{\text{CMMPS}}$ | $H_T/2$ | $m_t$ | $\mu_{\text{CMMPS}}$ | $(m_t m_{b\bar{b}})^{1/2}$ | $\left(m_t H_{T,b/2}\right)^{1/2}$ | $\left(m_t H_{T,t}/2\right)^{1/2}$ | $\mu_{\text{CMMPS}}$ | $\mu_{\text{CMMPS}}$ |
| $\mu_F$ | $H_{T,t}/2$ | $H_T/2$ | $m_t$ | $\mu_{\text{CMMPS}}$ | $H_{T,t}/2$ | $H_{T,t}/2$ | $H_{T,t}/2$ | $H_{T,t}/2$ | $H_{T,t}/2$ |
| $\mu_Q$ | $H_{T,t}/2$ | $H_T/2$ | $m_t$ | $\mu_{\text{CMMPS}}$ | $H_{T,t}/2$ | $H_{T,t}/2$ | $H_{T,t}/2$ | $\mu_{\text{CMMPS}}$ | $m_t$ |
| Cuts | $\Delta\sigma/\sigma$ | $\Delta\sigma/\sigma$ | $\Delta\sigma/\sigma$ | $\Delta\sigma/\sigma$ | $\Delta\sigma/\sigma$ | $\Delta\sigma/\sigma$ | $\Delta\sigma/\sigma$ | $\Delta\sigma/\sigma$ | $\Delta\sigma/\sigma$ |
| $ttb$ | 0% | −41% | −27% | +4.7% | +2.3% | 1.1% | −32% | −3.5% | −0.3% |
| $ttbb$ | 0% | −33% | −17% | −0.7% | +0.2% | 3.4% | −22% | −6.4% | −1.1% |
| $ttbb_{100}$ | 0% | −29% | −13% | −9.2% | −5.6% | +2.5% | −17% | −14% | −2.9% |

glo single global scale: hard, fixed and softer

R renormalisation scale (dominant!): modify or avoid b-jet dependence

Q resummation-scale (PS uncertainties): softer and fixed

26

**Additional $m_b$ and PDF variations** with potential impact on shape (and normalisation)

|  | $M_b = 5.0$ | $M_b = 4.5$ | CTEQ 4F | MSTW$_{37}$ | MSTW$_{38}$ |
|---|---|---|---|---|---|
| Cuts | $\Delta\sigma/\sigma$ | $\Delta\sigma/\sigma$ | $\Delta\sigma/\sigma$ | $\Delta\sigma/\sigma$ | $\Delta\sigma/\sigma$ |
| $ttb$ | $-3.5\%$ | $+4.4\%$ | $-10\%$ | $-0.1\%$ | $+2.6\%$ |
| $ttbb$ | $-0.7\%$ | $+2.7\%$ | $-9.3\%$ | $+0.2\%$ | $+4.2\%$ |
| $ttbb_{100}$ | $-0.1\%$ | $+4.4\%$ | $-7.8\%$ | $-0.7\%$ | $+6.9\%$ |

- conservative b-mass variations $m_b = 4.75 \pm 0.25 \, \text{GeV}$ (impact on collinear regions)
- compare central MSTW to central CT10 PDF and MSTW variations with large gluon-shape distortion (MSTW eigenvector 19)

27

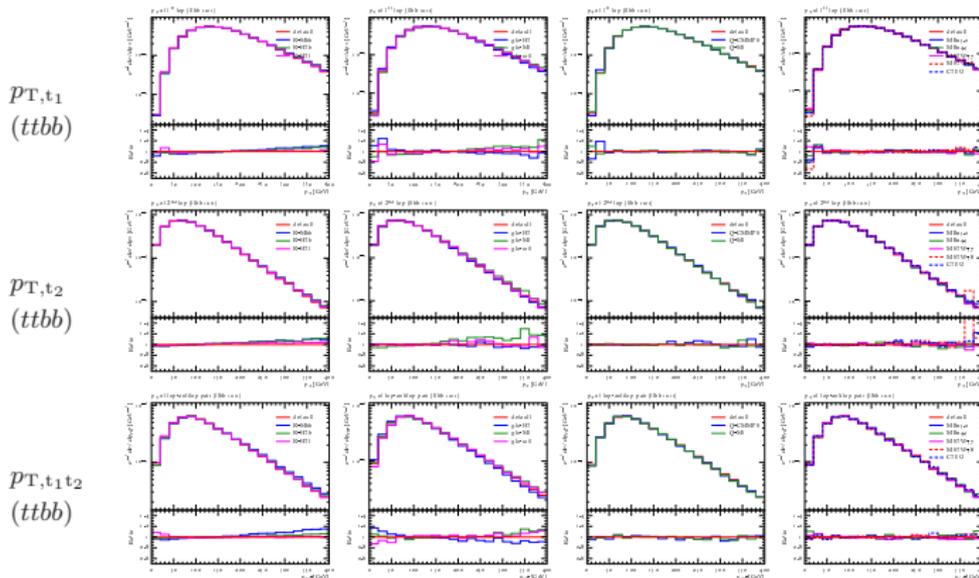### Shape variations of differential observables

The following plots show a representative selection of shape uncertainties

- normalisation uncertainties removed by normalising all distributions to one
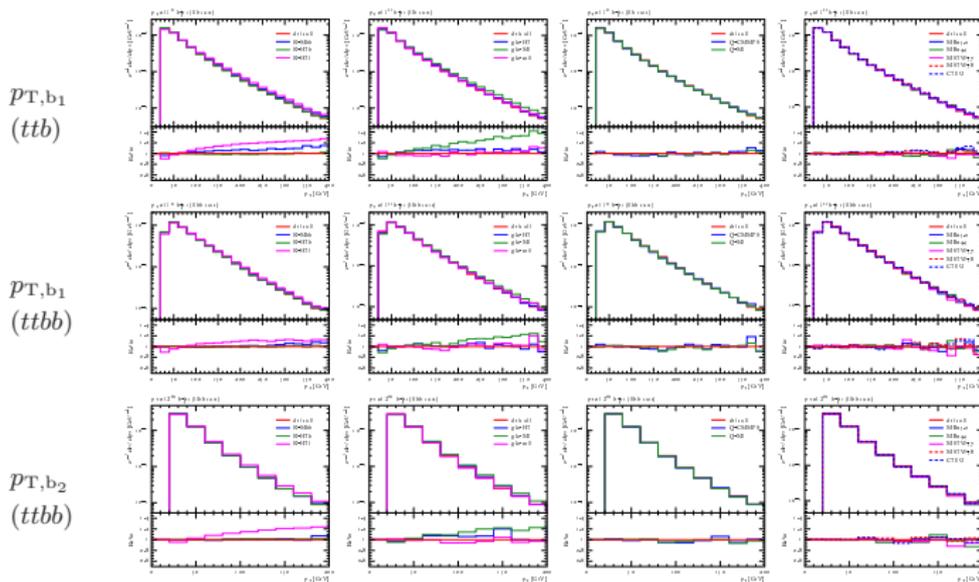- columns represent (1) R-type (2) glo-type (3) Q-type (4) $m_b$+PDFs variations

**Shape uncertainty of top-quark and b-jet rapidities**
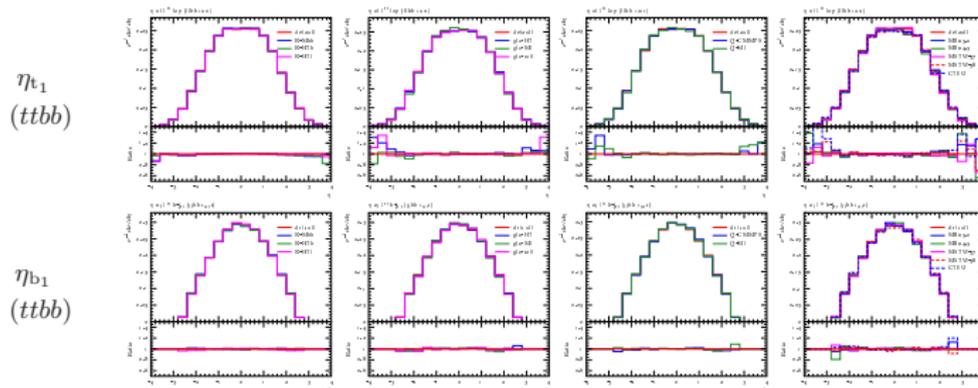


$\Rightarrow$ percent-level variations for $|\eta| < 2.5$; $\eta_b$ very stable

28

## Shape uncertainty of top-$p_{\mathrm{T}}$

$p_{\mathrm{T},\mathrm{t}_1}$
$(ttbb)$

$p_{\mathrm{T},\mathrm{t}_2}$
$(ttbb)$

$p_{\mathrm{T},\mathrm{t}_1\mathrm{t}_2}$
$(ttbb)$



$\Rightarrow \sim10\%$ variations (20% in the tails) driven by top-dependence of $\mu_{\mathrm{R}}$

29

**Shape uncertainty of b-jet $p_T$**

$p_{T,b_1}$
($ttb$)

$p_{T,b_1}$
($ttbb$)

$p_{T,b_2}$
($ttbb$)



$\Rightarrow$ ~10-20% variations (40% in the tails) driven by b-dependence of $\mu_R$

30

**Shape uncertainty of b-jet correlations**

$m_{b_1 b_2}$
($ttbb$)



$\Delta R_{b_1 b_2}$
($ttbb$)



$\Delta R_{b_1 b_2}$
($ttbb_{100}$)
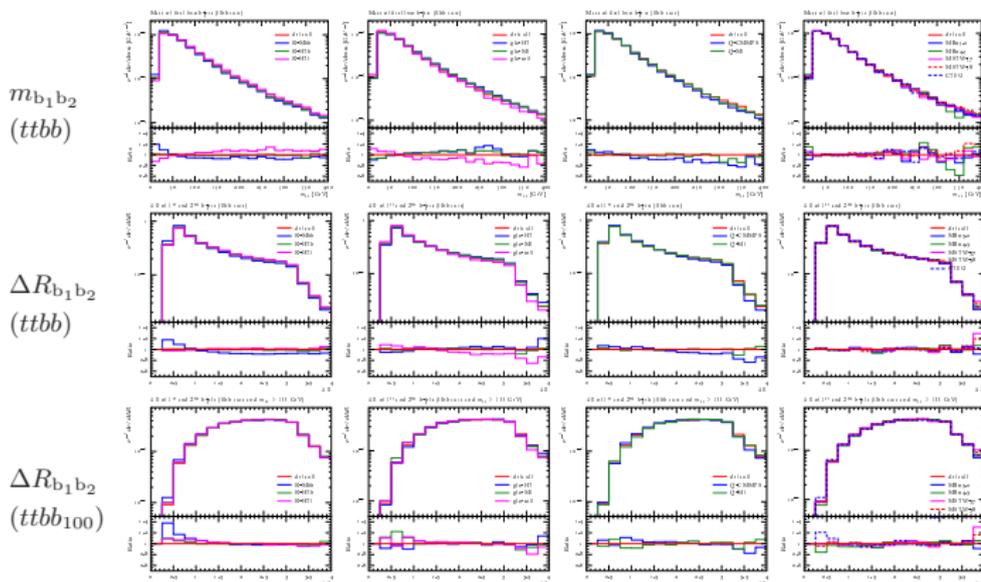


$\Rightarrow$ ~10-20% variations driven by b-dependence of $\mu_R$ (at small $m_{bb}$ and $\Delta R$) and (agressive) reduction of $\mu_Q$ in the tail

31

**Shape uncertainty of $1^{st}$ light-jet $p_T$**

$p_{T,j_1}$
$(ttb)$
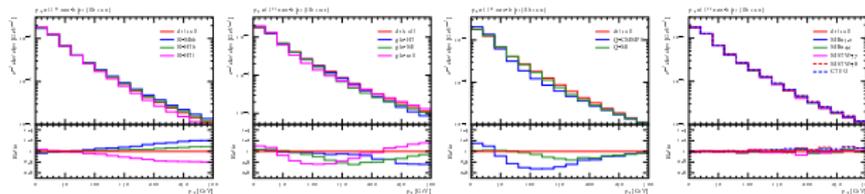
$p_{T,j_1}$
$(ttbb)$
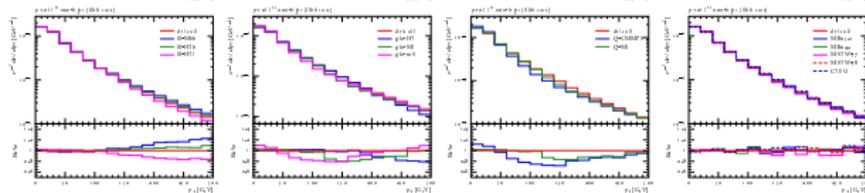


⇒ up to ∼30% variations at intermediate $p_T$ values. Indicates that the considered variations (dominated by choice of soft resummation scale) are (probably too) conservative

# Scale choices (YR4) and uncertainties (no proposal yet)

**Factorisation ($\mu_Q$) and resummation ($\mu_Q$) scales**
$$E_{T_i} = \sqrt{m_i^2 + p_{T,i}^2}$$

$$\mu_F = \mu_Q = \frac{H_T}{2} = \frac{1}{2} \sum_{i=t,\bar{t},b,\bar{b}} E_{T,i}$$

$\mu_Q \equiv$ shower starting scale is a free paramater in MC@NLO (not in Powheg)

**CKKW-like (softer) renormalisation scale**

$$\mu_R \quad = \quad \mu_{\mathrm{CKKW}} = \prod_{i=t,\bar{t},b,\bar{b}} E_{T,i}^{1/4}$$

**Scale variations (leading uncertainty)** $\sim$20-30%

- factor-2 variations of $\mu_R$ and $\mu_F$ ⇔ normalisation
- "kinematic" variations of $\mu_R, \mu_F, \mu_Q$ ⇔ shape
- variations of $\mu_Q$ in MC@NLO and $h_{\mathrm{damp}}$ in Powheg ⇔ NLOPS matching

**Other variations**

- PDF variations (only few percent)
- shower variations: tune variations, shower recoil scheme, . . .

# Correlation of TH uncertainties between categories

**Categories**

- $t\bar{t}h(b\bar{b})$ analyses based on simultaneous fit of MC to data in various categories with different # of light- and $b$-jets
- correlations crucial to constrain background in signal region (with multiple $b$-jets)

**Between $t\bar{t}$+light-jet and $t\bar{t} + b$-jet categories**

- uncertainties should be uncorrelated

**Between sub-categories** (e.g. $ttb$, $ttbb$, $ttB$)

- uncertainties should be correlated

  **Motivation**: independent shower, matching and ME variations account for different types of uncertainties (e.g. related to collinear $g \to b\bar{b}$ splittings or hard b-production) $\Rightarrow$ no need of separate categories with uncorrelated uncertainties