TECHNISCHE
UNIVERSITÄT
DRESDEN

**Fakultät Physik** Institut für Kern- und Teilchenphysik

Emmy
Noether-
Programm

Deutsche
Forschungsgemeinschaft
DFG

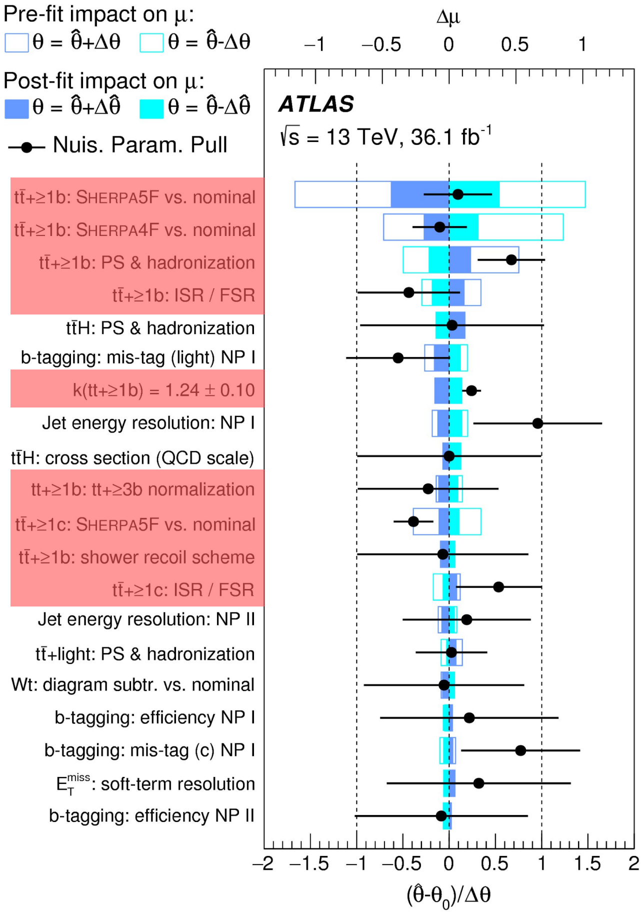
# Modelling and uncertainties for ttbb production as background to ttH(bb)

Frank Siegert

TOP 2019, Beijing, September 2019

# Why do we care so much about ttbb?

- ATLAS and CMS ttH(bb) analyses rely on MC modelling for irreducible ttbb background
  - included as template in profile likelihood fit
- Largest sources of uncertainty on extracted signal strength related to tt+HF MC modelling!
- What can we improve?
  - **ATLAS & CMS:** relied on NLO+PS ttbar so far! More accurate theory with NLO ttbb used only to reweight HF fractions (ATLAS) or cross-checks (CMS)
  - **Theory:** Large perturbative ttbb uncertainties even enlarged by NLO+PS algorithms
  - **Both**: More rigorous combination of inclusive tt+jets and ttbb predictions.

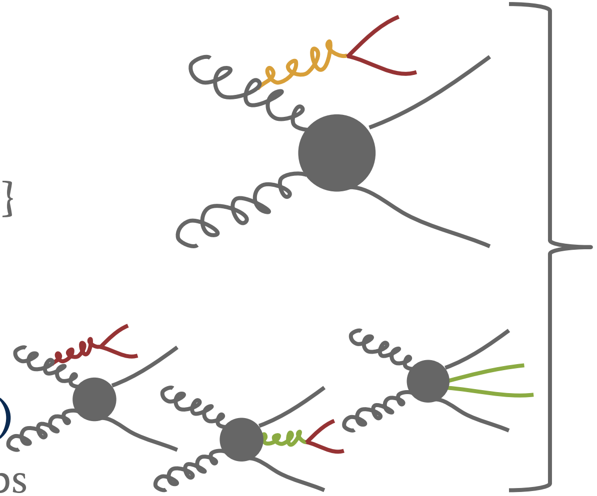


TECHNISCHE
UNIVERSITÄT
DRESDEN

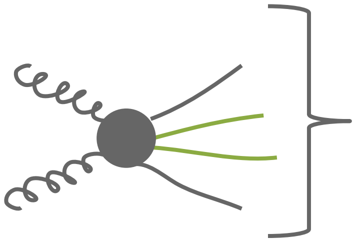
# Event generation for tt + heavy flavour

Traditional approaches for tt+HF MC predictions:

- “Inclusive” NLO+PS tt sample with HF production from **parton shower g→bb**
  - e.g. {Powheg,aMC@NLO}+{Pythia,Herwig}
- Multi-leg merged tt+jets sample with HF from **higher-order MEs** (hard b’s) or **parton shower g→bb** (soft/collinear b’s)
  - e.g. MG5_aMC+Pythia, Sherpa+OpenLoops
- NLO+PS ttbb using **matrix elements** with **massive** b-quarks
  - e.g. Powheg+OpenLoops+Pythia8, Sherpa+OpenLoops



“5-flavour” schemes

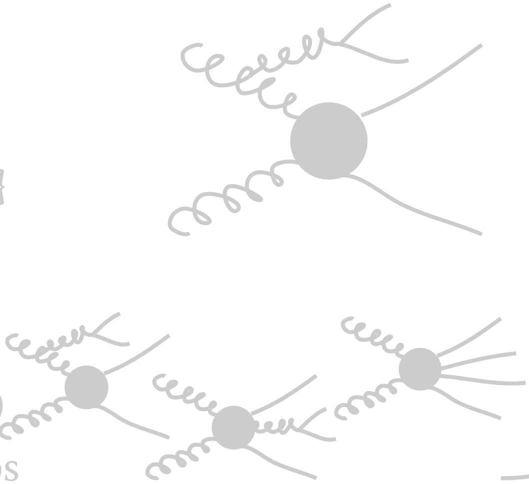


“4-flavour” schemes

# Event generation for tt + heavy flavour

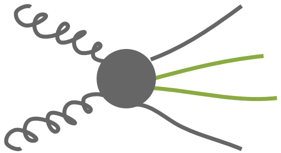
Traditional approaches for tt+HF MC predictions:

- “Inclusive” NLO+PS tt sample with HF production from **parton shower g→bb**
  - e.g. {Powheg,aMC@NLO}+{Pythia,Herwig}
- Multi-leg merged tt+jets sample with HF from **higher-order MEs** (hard b’s) or **parton shower g→bb** (soft/collinear b’s)
  - e.g. MG5_aMC+Pythia, Sherpa+OpenLoops



“5-flavour” schemes

- NLO+PS ttbb using **matrix elements** with **massive** b-quarks
  - e.g. Powheg+OpenLoops+Pythia8, Sherpa+OpenLoops



“4-flavour” schemes

# Anatomy of 4FS NLO+PS for ttbb

- 2→4 NLO QCD matrix elements with massive b-quarks

Final state g→bb **dominant**

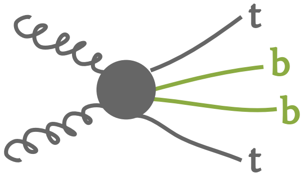
- massive b’s → no (jet) cuts!
- collinear g→bb produced in ME

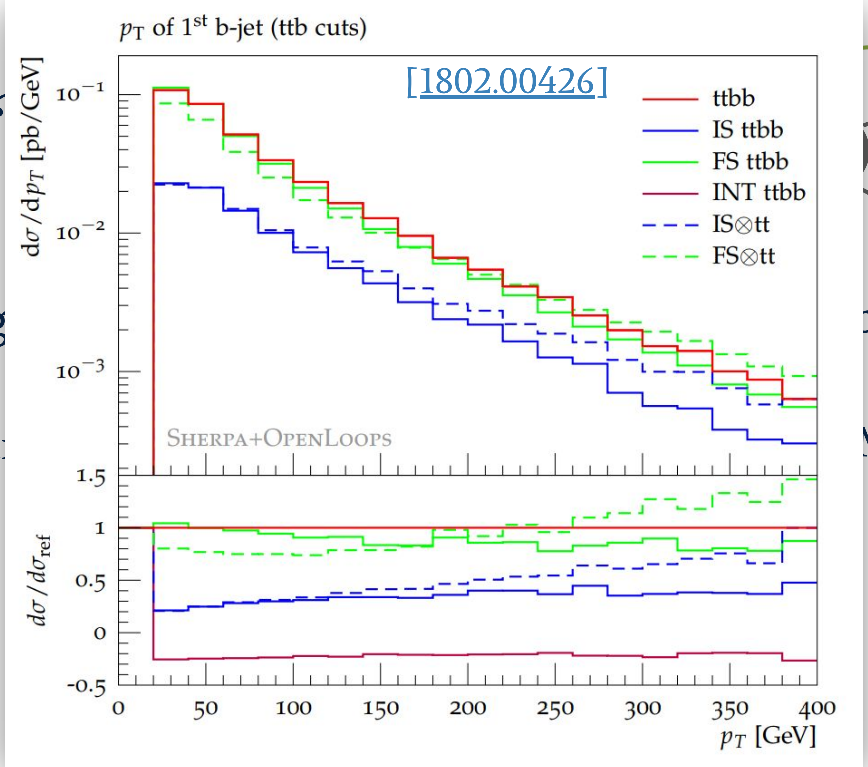
No initial state b in MEs

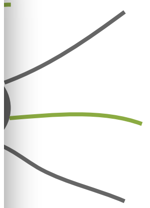
- 4FS PDFs
- IS g→bb in ME

# Anatomy of 4FS NLO+PS for ttbb

- 2→4 NLO QCD matrix elements with massive b-quarks







Final state g … in MEs

  - massive … 
  - collinea … ME

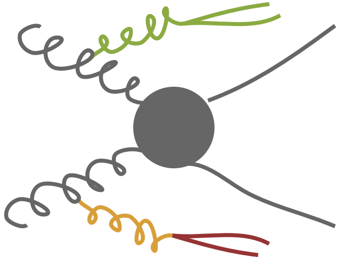
# Anatomy of 4FS NLO+PS for ttbb

- 2→4 NLO QCD matrix elements with massive b-quarks

Final state g→bb **dominant**
  - massive b’s → no (jet) cuts!
  - collinear g→bb produced in ME

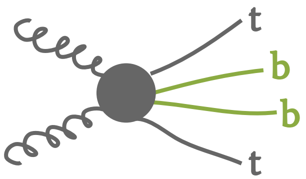
No initial state b in MEs
  - 4FS PDFs
  - IS g→bb in ME

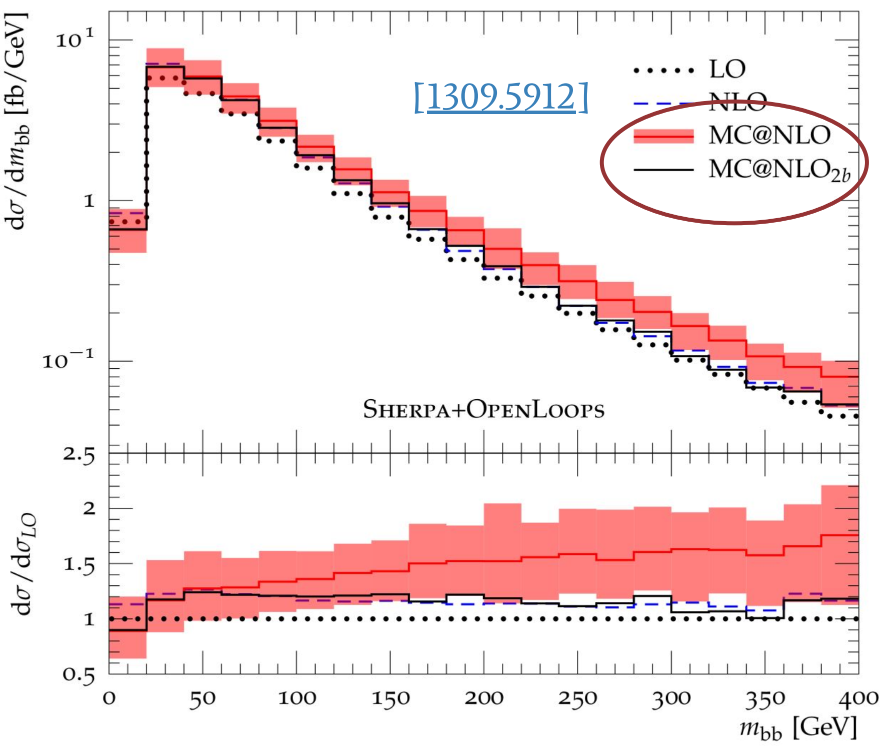
- Matched to parton shower for additional emissions
  - **“double-splitting”** contribution becomes relevant!



# Anatomy of 4FS NLO+PS for ttbb

- 2→4 NLO QCD matrix elements with massive b-quarks





Final state … in MEs

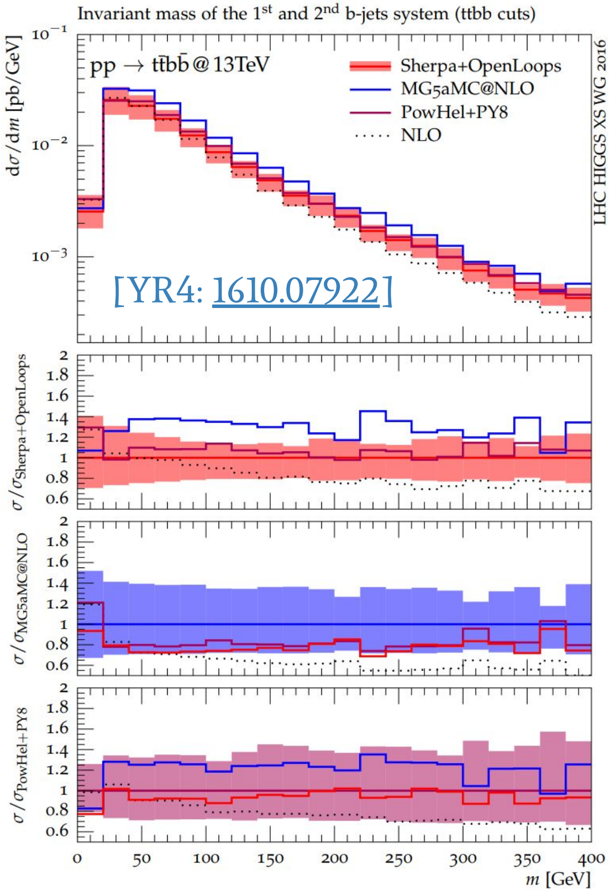
- massive …
- collinea … ME

- Matched to par… additional emis…
  - “double-spli… becomes re…

# MC programs for 4FS ttbb at NLO+PS

- Several tools on the market
  - Sherpa + OpenLoops [1309.5912]
  - PowHel + Pythia/Herwig [1709.06915]
  - PowhegBox + OpenLoops + Pythia/Herwig [1802.00426]
  - MG5_aMC + Pythia/Herwig
  - Herwig7 + OpenLoops
- History of out-of-the-box comparisons:
  - Large discrepancies
  - Partially due to large perturbative uncertainties
  - But also beyond!
    - Parton Shower?
    - NLO+PS matching algorithm?

**Improve** or accept as **uncertainties** (and kill ttHbb?)?



[YR4: 1610.07922]

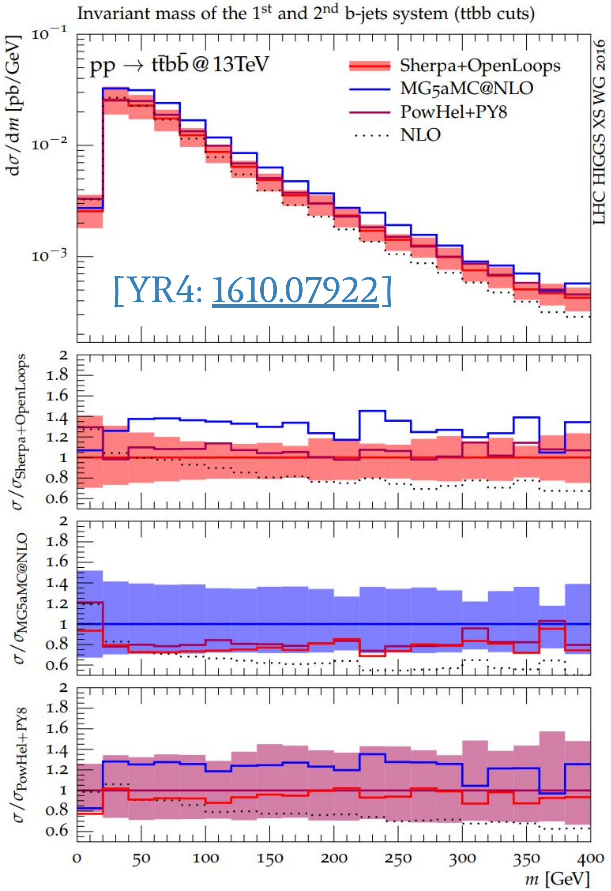
# MC programs for 4FS ttbb at NLO+PS

- Several tools on the market
  - Sherpa + OpenLoops [1309.5912]
- F
    - » NL matching algorithm?

Arguably one of the most complex processes for NLO+PS matching

→ Strong challenge to understand unc's as prototype for other processes!

**Improve** or accept as **uncertainties** (and kill ttHbb?)?

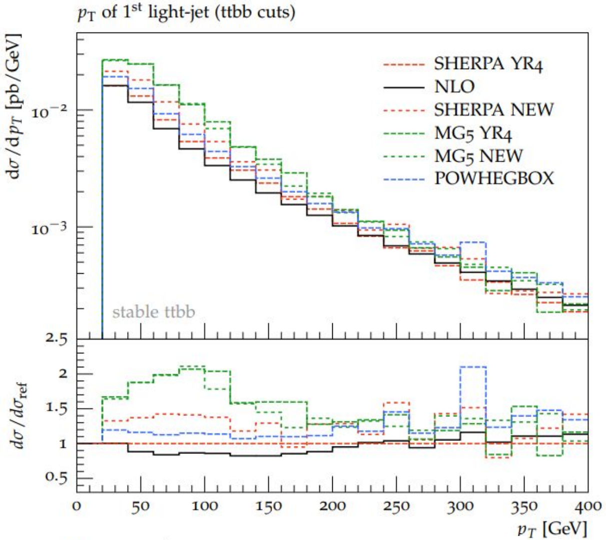


[YR4: 1610.07922]

"

# Diagnosis: Tuned comparisons

- Tuned comparison effort to compare matching and parton shower between various tools
  → Isolate algorithmic unc's in:
  - NLO+PS matching
  - Parton shower (e.g. recoil scheme effects)
- New input from PowhegBox implementation and ttbbj NLO calculation helps pin down discrepancies
- Common [Rivet routine] for tt+1b and tt+2b final states in context of ttH subgroup in HXSWG



[from ttH-HXSWG]

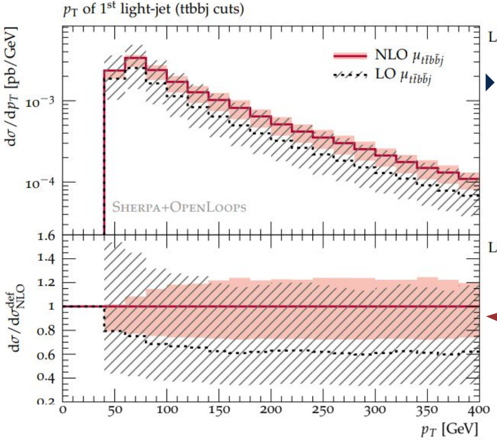
| Tool | MC@NLO | Powheg | Pythia 8.2 | Herwig 7.1.2 | Sherpa2.2.4 | MC contacts |
|---|---|---|---|---|---|---|
| Sherpa2.2+OpenLoops | x | | | | x | F. Siegert, J. Krause |
| MG5_aMC@NLO | x | | x | x | | M. Zaro |
| MatchBox+OpenLoops | x | | | x | | C. Reuschle, R. Posdkubka |
| Powheg+Helac | | x | x | x | | M.V. Garzelli, A. Kardos |
| PowhegBox+OpenLoops | | x | x | x | | T. Jezo, J. Lindert |
| | 3 | 2 | 3 | 4 | 1 | + S. Pozzorini |

# Therapy: Tuned matching [Preliminary]

- Differences suspected as combination of 2 effects in MC@NLO matching:

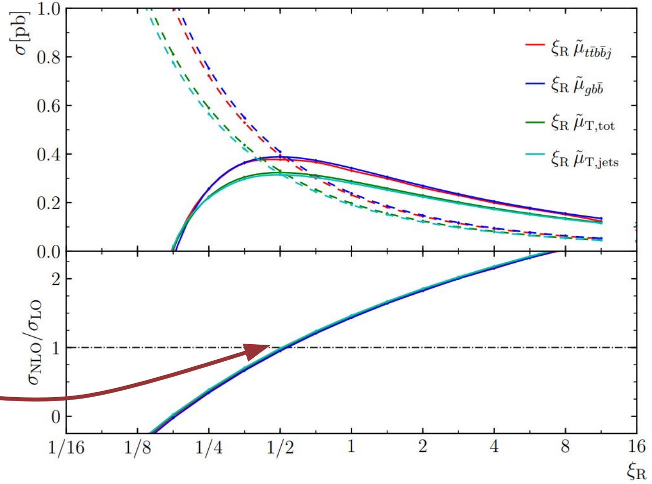
$$\mathrm{d}\sigma^{(\mathrm{NLO+PS})} = \mathrm{d}\Phi_B \bar{\mathcal{B}} \left[ \underbrace{\Delta(t_0, \mu_Q^2)}_{\text{unresolved}} + \underbrace{\int_{t_0}^{\mu_Q^2} \mathrm{d}t \frac{\mathcal{R}_{\mathrm{PS}}}{\mathcal{B}} \Delta(t, \mu_Q^2)}_{\text{resolved, singular}\equiv\mathbb{S}} \right] + \mathrm{d}\Phi_R \underbrace{[\mathcal{R} - \mathcal{R}_{\mathrm{PS}}]}_{\text{resolved, non-singular}\equiv\mathbb{H}}$$

  - large K-factor~1.9
  - spuriously large $R_{PS}$ in MC@NLO matching with MadGraph5_aMC@NLO + Pythia/Herwig



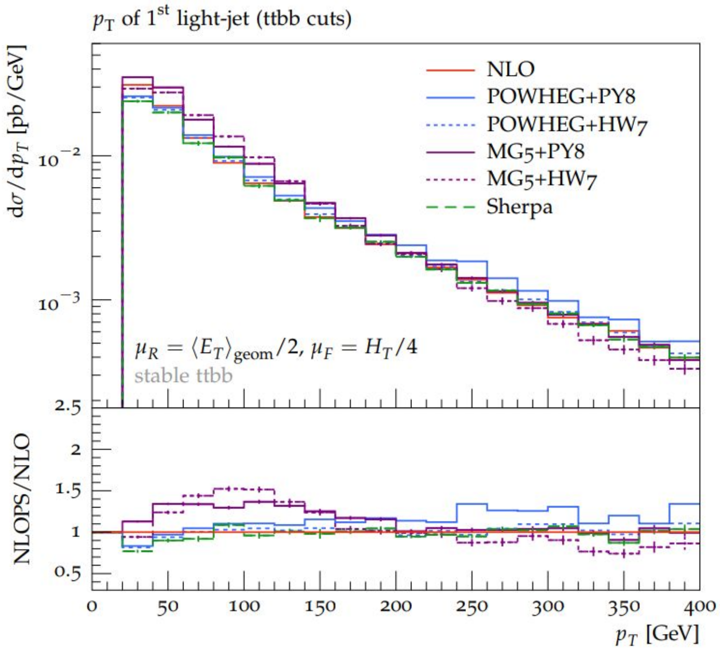
- Fixed-order studies of **ttbbj@NLO** with OpenLoops2+Sherpa
  [Buccioni, Kallweit, Pozzorini, Zoller 2019]
  - Reduced $\mu_R$ stabilises K-factor
  - No significant shape distortions

New benchmark for NLO+PS progs!



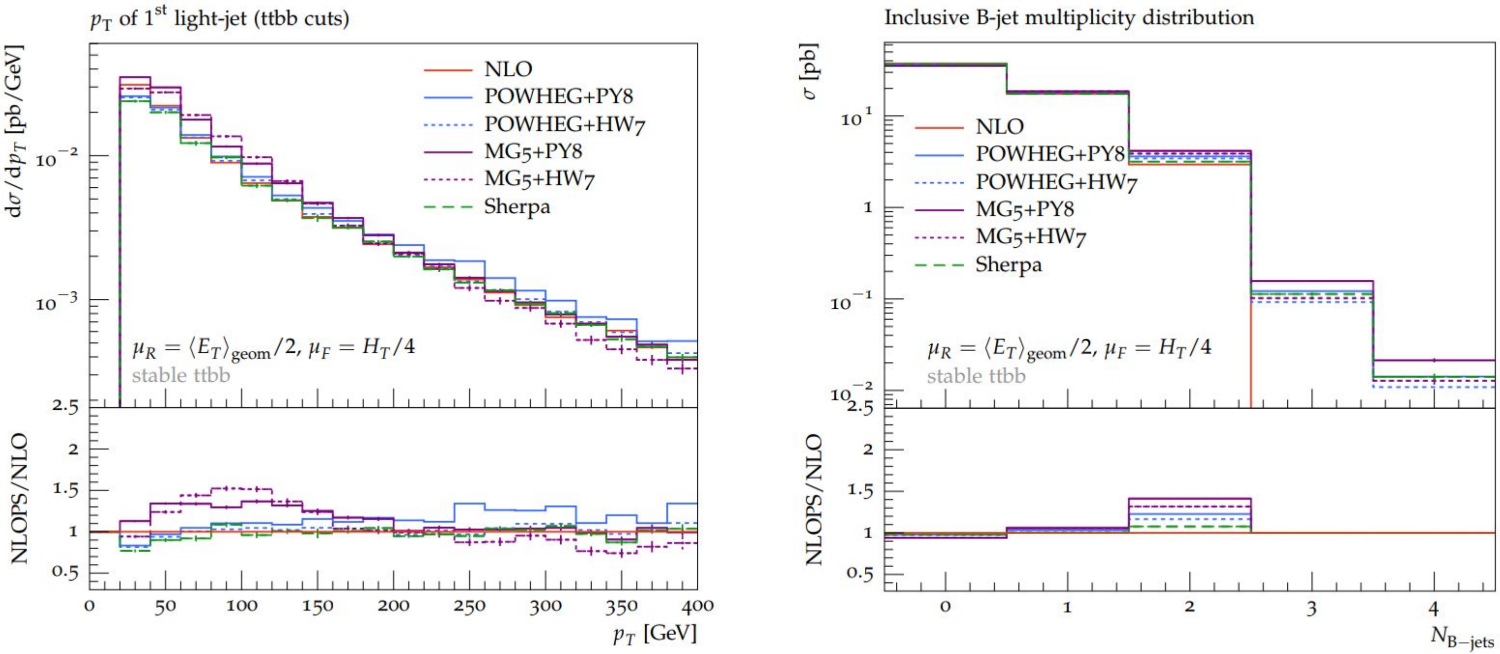
# Therapy: Tuned matching [Preliminary]

- Application of reduced scale to tuned NLO+PS comparisons
  - improved agreement between NLO+PS tools for light-jet spectrum



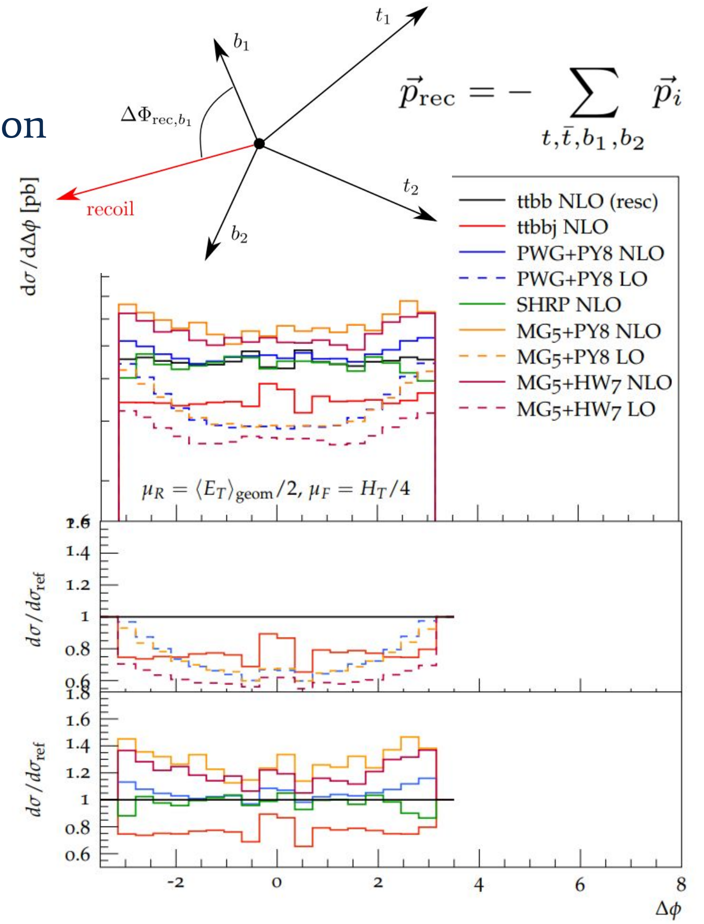
# Therapy: Tuned matching [Preliminary]

- Application of reduced scale to tuned NLO+PS comparisons
  - improved agreement between NLO+PS tools for light-jet spectrum
  - still sizable O(40%) differences in $N_{2b}$ region → origin?

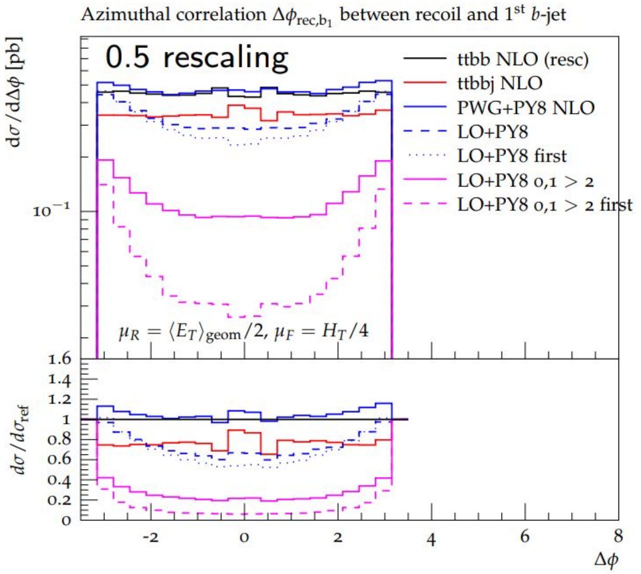


# Recoil observables [Preliminary]

- Large shower recoil effect on b-jets
  - strong recoil to b-jet in Pythia8 already in 1st emission → ruled out by ttbbj NLO
  - survives in MC@NLO matching procedures



$$\vec{p}_{\mathrm{rec}} = -\sum_{t,\bar{t},b_1,b_2} \vec{p}_i$$



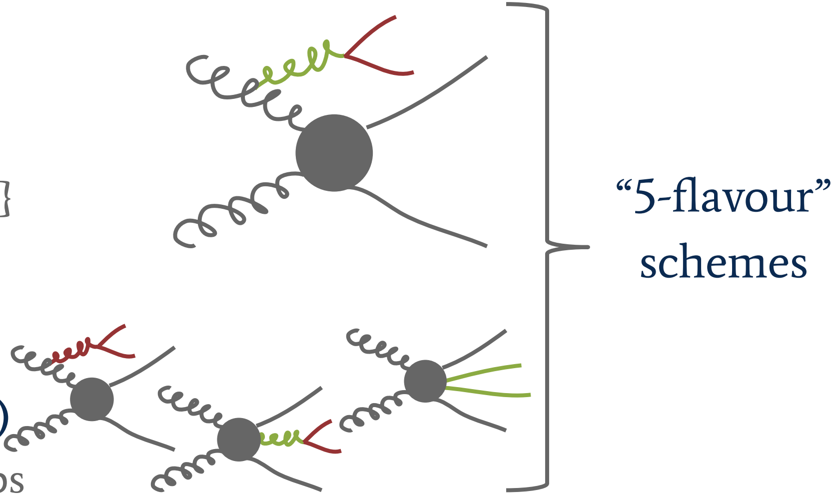
- Have to accept these differences as matching uncertainties?
- Not surprising, since $\langle p_T^{\mathrm{jet}}\rangle$ ~ 10 x $\langle p_T^{\mathrm{bjet2}}\rangle$!

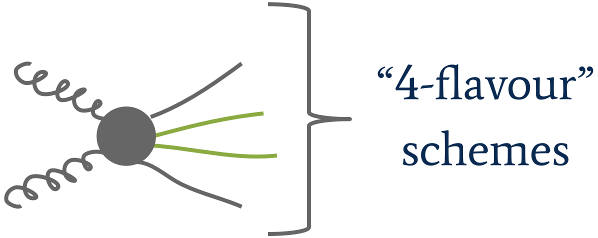
How to reduce uncertainties in hard jet configurations?

# Recap: Event generation for tt + heavy flavour

Traditional approaches for tt+HF MC predictions:

- “Inclusive” NLO+PS tt sample with HF production from **parton shower g→bb**
  - e.g. {Powheg,aMC@NLO}+{Pythia,Herwig}
- Multi-leg merged tt+jets sample with HF from **higher-order MEs** (hard b’s) or **parton shower g→bb** (soft/collinear b’s)
  - e.g. MG5_aMC+Pythia, Sherpa+OpenLoops
- NLO+PS ttbb using **matrix elements** with **massive** b-quarks
  - e.g. Powheg+OpenLoops+Pythia8, Sherpa+OpenLoops

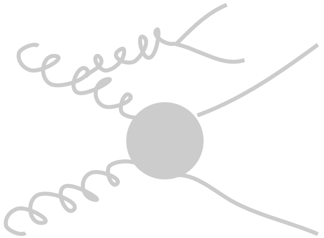




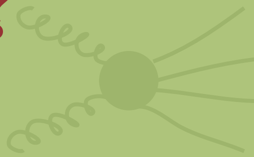
# Recap: Event generation for tt + heavy flavour

Traditional approaches for tt+HF MC predictions:

- "Inclusive" NLO+PS tt sample with HF production from **parton shower g→bb**
  - e.g. {Powheg,aMC@NLO}+{Pythia,Herwig}
- Multi-leg merged tt+jets sample with HF from **higher-order MEs** (hard b's) or **parton shower g→bb** (soft/collinear b's)
  - e.g. MG5_aMC+Pythia, Sherpa+OpenLoops
- NLO+PS ttbb using **matrix elements** with **massive** b-quarks
  - e.g. Powheg+OpenLoops+Pythia8, Sherpa+OpenLoops



"5-flavour" schemes



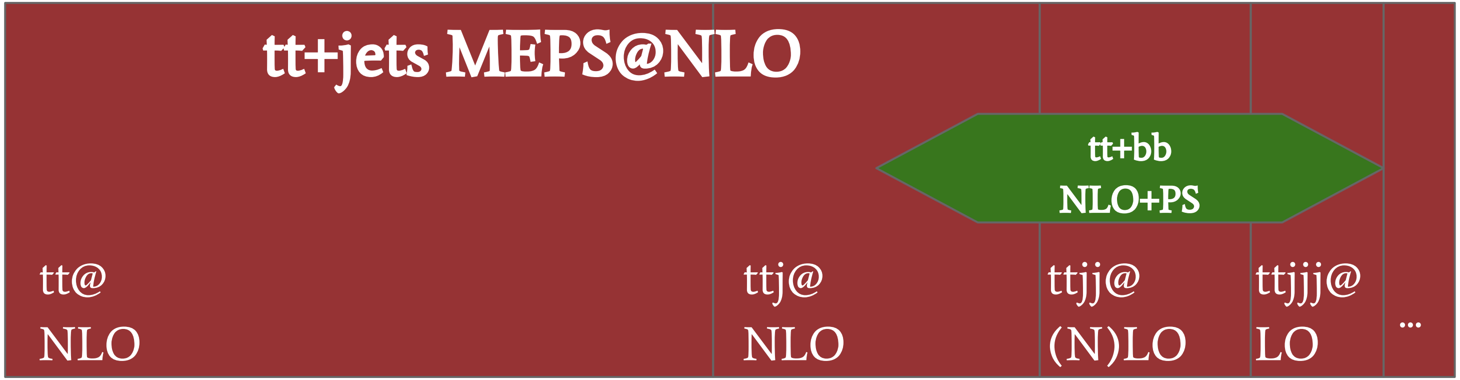
"4-flavour" schemes

**Can we combine 4-flavour and 5-flavour multileg?**

# Fusing X+bb and X+jets in the Sherpa MC

aka “Multi-jet merging in a variable flavour number scheme”

[1904.09382]



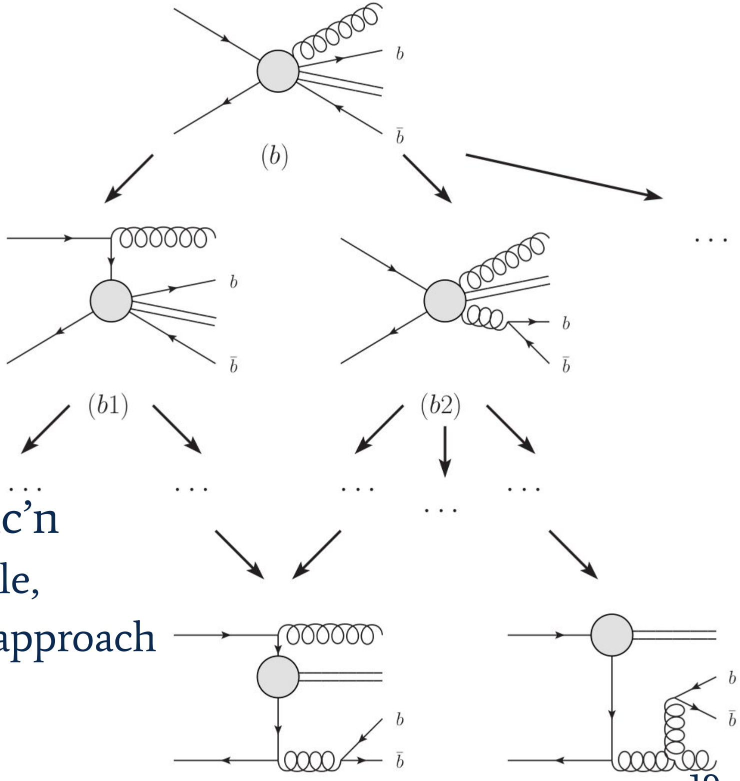
**Three main ingredients:**

1. Interpreting ttbb as merged contribution
2. Overlap removal
3. Matching 4F/5F in PDFs and $\alpha_S$

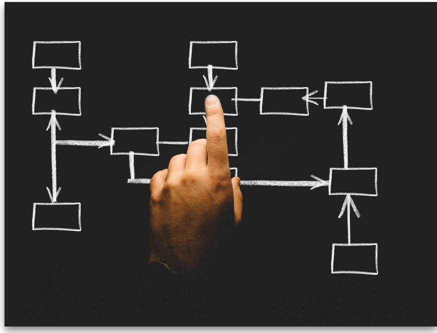
Can be applied for LO and NLO merging!

# Step 1: Embedding ttbb as merged contribution

- ttj(j(...)) matrix elements treated in regular **tt+jets MEPS@NLO**:
  - clustering to get topology of ME emissions ("shower history")
  - core scale based on 2→2 process
  - application of $\alpha_S(\mu_R{}^2) \to \alpha_S(p_T{}^2)$ reweighting for each emission
  - application of Sudakov factors $\Delta(t_1, t_2)$ along internal lines (event vetoes) for correct resummation properties
- Now: Same applied to **ttbb NLO+PS** massive calc'n
  - remains separate standalone ttbb NLO+PS sample, but generated consistent with multi-leg merged approach



# Step 2: Heavy Flavour Overlap Removal



- HFOR used before in experiments in simplified form
  - dR(b,b)>0.4 → keep from ttbb ME
  - dR(b,b)<0.4 → keep from tt ME + bb from PS
- Here: from multi-leg merging prescription
  - Cluster **full event at PS level** using “reverse shower”
  - Look at **leading two emissions**
    - Heavy Flavour → keep from **ttbb NLO+PS** simulation (“**direct component**”)
    - Light Flavour → keep from **tt+jets MEPS@NLO** (“**fragmentation component**”)

⇒ Sub(sub)leading g→bb splittings not from ttbb ME, but from ttjjjj ME or from PS.

- (Extra: caution with b’s from “FSR” in top decay products!)

# Step 3: Matching 4F/5F in PDFs and $\alpha_S$

- For consistent combination with tt+jets we produce the massive **ttbb NLO+PS** with a 5F PDF

  → $m_b$ mismatch with massive NLO matrix elements

  - Looking at ideas from **FONLL** [Forte, Napoletano, Ubiali 2016] based on

    $$\sigma^{\text{FONLL}} = \sigma^{(5)} - \sigma^{(4),(0)} + \sigma^{(4)}$$

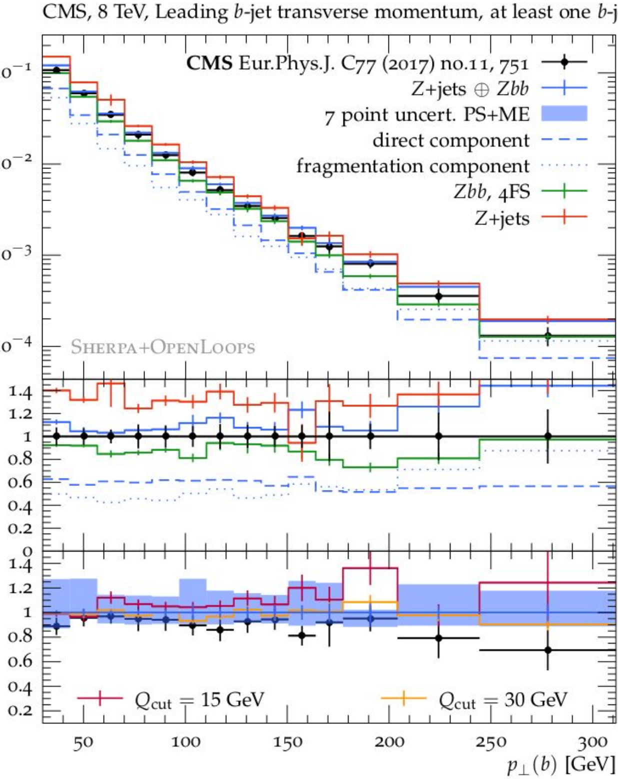
    we find that they are generated by prescription above!
  - NLO accuracy preserved from input matrix elements
  - LL/NLL accuracy according to shower used
    - Overlap removal and embedding of ttbb as merged contribution with LL shower automatically generates leading log matching term
    - Next-to-leading log would need explicit counterterms as event weights (complicated) or comes **automatically with NLL showers** in the future
- Additional event weights for mismatch between $\alpha_S$ evolution with $m_b = 0$ and virtuals with $m_b \neq 0$

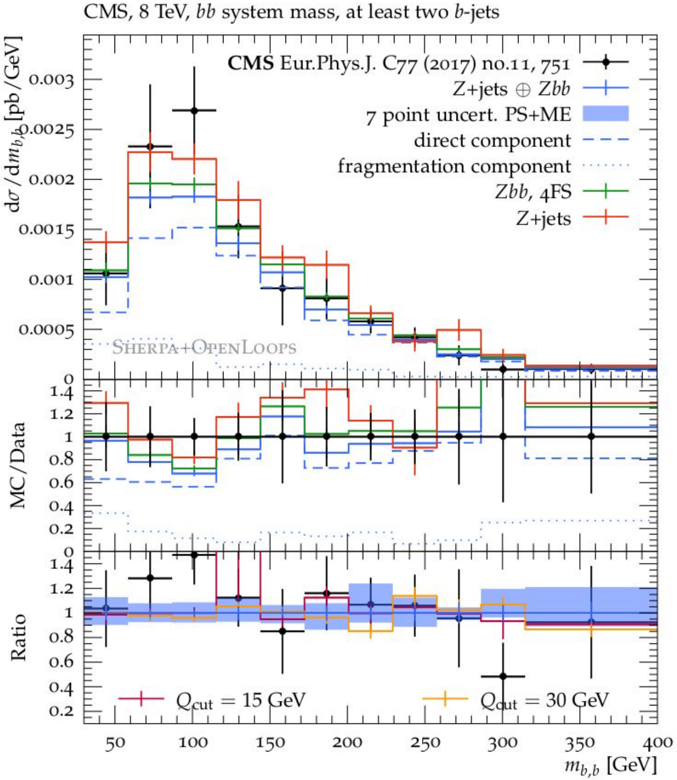
$$w^{\text{new}}_{q\bar{q}} = w_{q\bar{q}} \left(1 - \frac{4}{3} T_R \ln \frac{\mu_R^2}{Q^2} \frac{w^{\text{Born}}}{w^{\text{ME}}}\right)$$

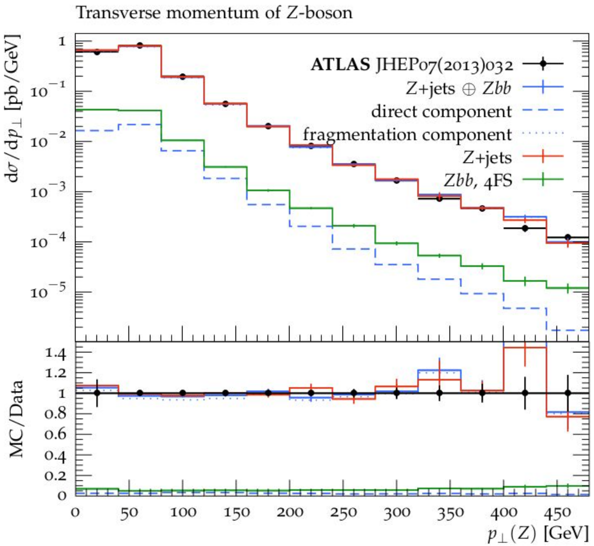
$$w^{\text{new}}_{gg} = w_{gg} \left(1 - \frac{4}{3} T_R \ln \frac{\mu_R^2}{m_b^2} \frac{w^{\text{Born}}}{w^{\text{ME}}}\right)$$

# Validation for Z+HF production

- Implementation in Sherpa 2.2
- First application to Z+HF, compared to CMS 8 TeV data

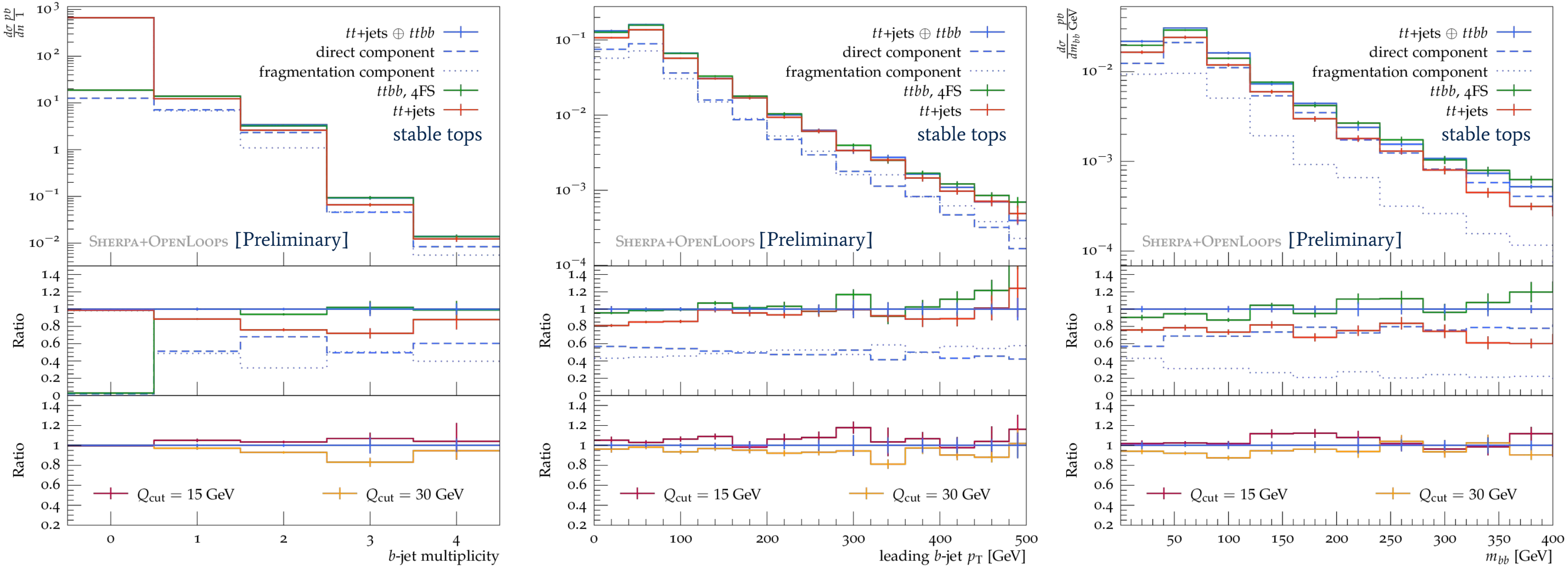

CMS, 8 TeV, Leading $b$-jet transverse momentum, at least one $b$-jet


CMS, 8 TeV, $bb$ system mass, at least two $b$-jets


Transverse momentum of $Z$-boson

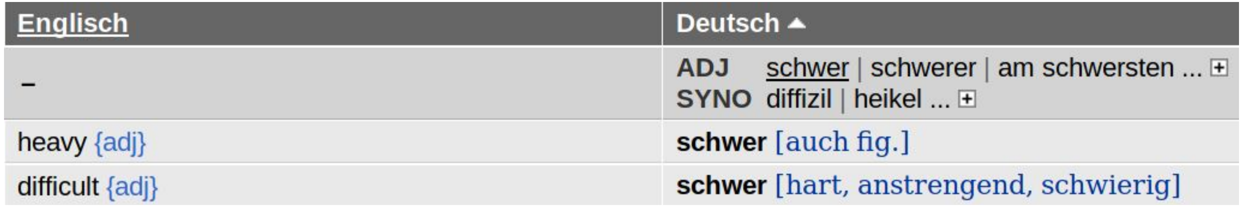
# Sneak preview: Application to ttbb

- Also applied as fusion of MEPS@NLO tt + 0,1j@NLO + 2,3j@LO and massive ttbb@NLO
- 2-bjet production dominated by direct component, but 1-bjet observables with equal contributions from direct and fragmentation configurations!



[Katzy, Krause, Pollard, FS in prep]

# Conclusions



<table>
<tr><th>Englisch</th><th>Deutsch ▲</th></tr>
<tr><td>–</td><td>ADJ schwer | schwerer | am schwersten ... ⊞<br>SYNO diffizil | heikel ... ⊞</td></tr>
<tr><td>heavy {adj}</td><td><b>schwer</b> [auch fig.]</td></tr>
<tr><td>difficult {adj}</td><td><b>schwer</b> [hart, anstrengend, schwierig]</td></tr>
</table>

- tt + heavy flavour predictions as background to ttH(→bb) are challenging
  - **NLO+PS matching** non-trivial and revealing spurious effects
- **ttbbj NLO** calculation valuable benchmark for NLO+PS ttbb simulations
- Large remaining uncertainties in $N_{2\text{ b-jet}}$, possibly due to recoil from hard jets?
- New **fusing algorithm** allows rigorous combination of tt+jets and ttbb MC simulations
  - More reliable simulation of configurations with large scale hierarchies (hard jets, soft b-jets)

Thank you for your attention!