TECHNISCHE UNIVERSITÄT DRESDEN
Fakultät Physik Institut für Kern- und Teilchenphysik

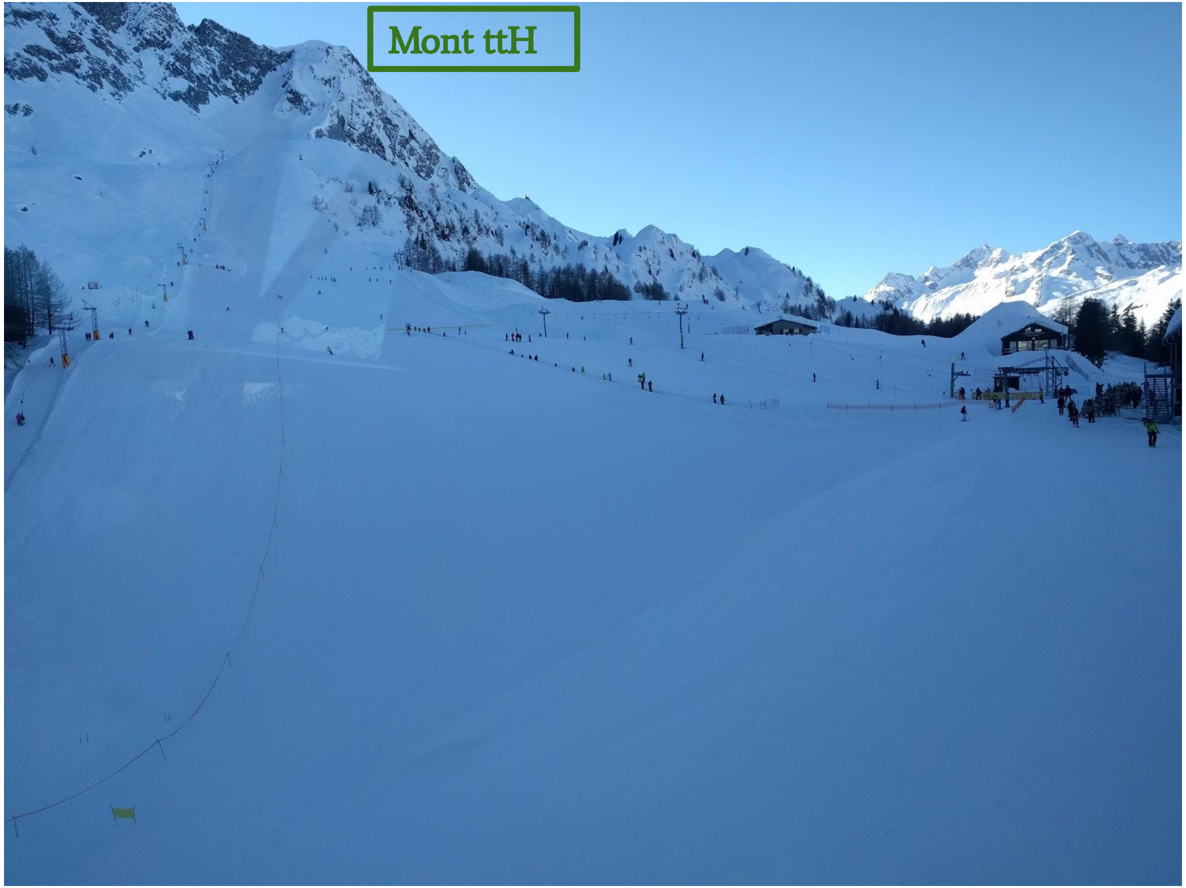
Emmy Noether-Programm
Deutsche Forschungsgemeinschaft
DFG

Modelling and uncertainties for backgrounds to ttH

Frank Siegert

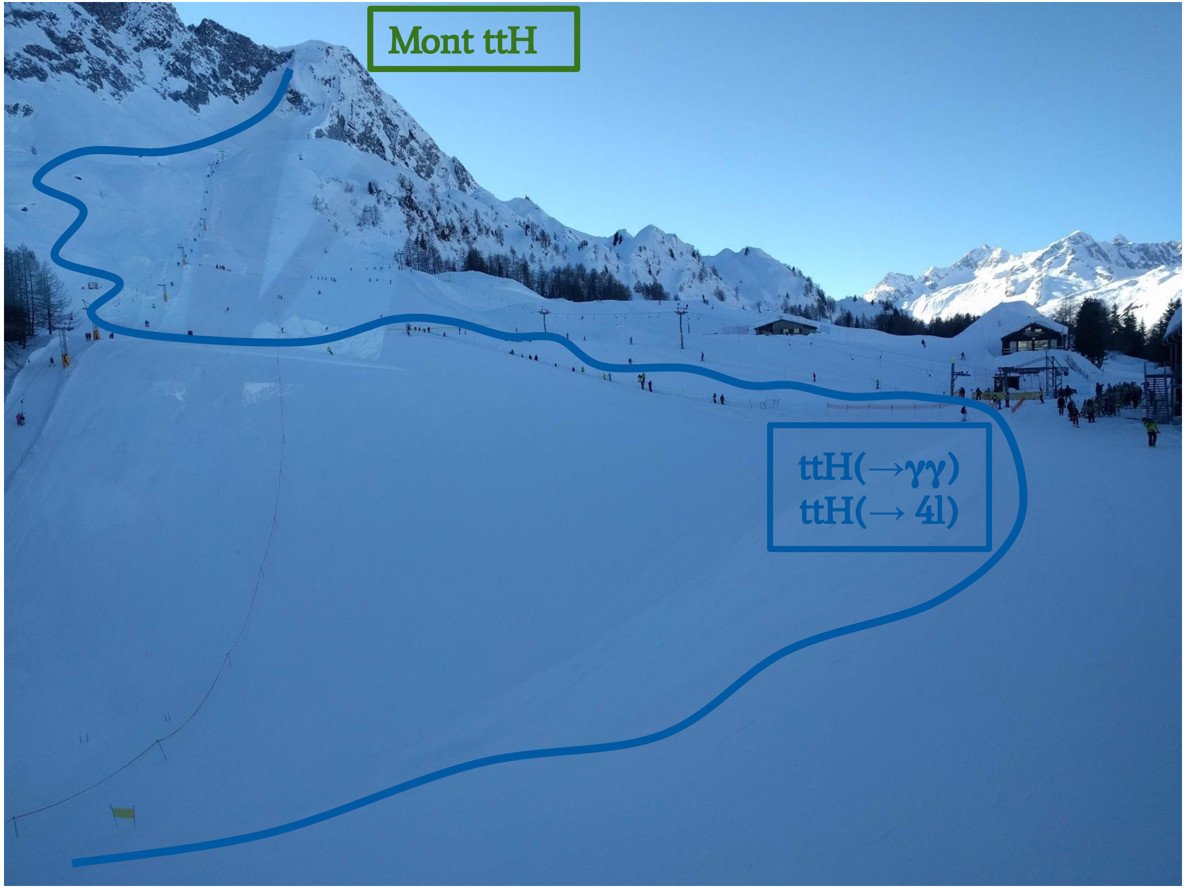
Zürich Phenomenology Workshop, January 2020

The ttH background modelling challenge



There are different ways to master Mont ttH!

The ttH background modelling challenge



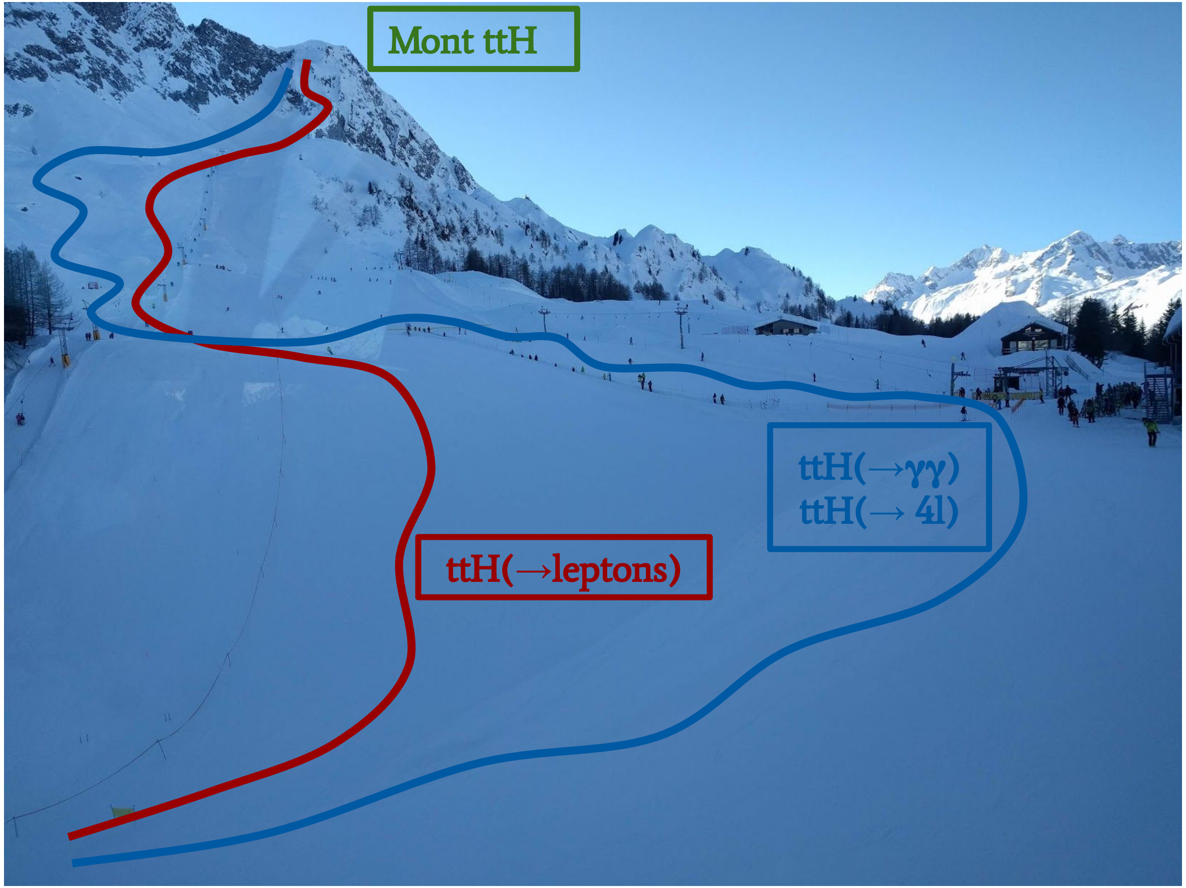
The Slow and Easy:

ttH(→γγ)

ttH(→ZZ→4l)

- ~no background challenge due to clean peak
- slow (but steady!) progress with data taking

The ttH background modelling challenge



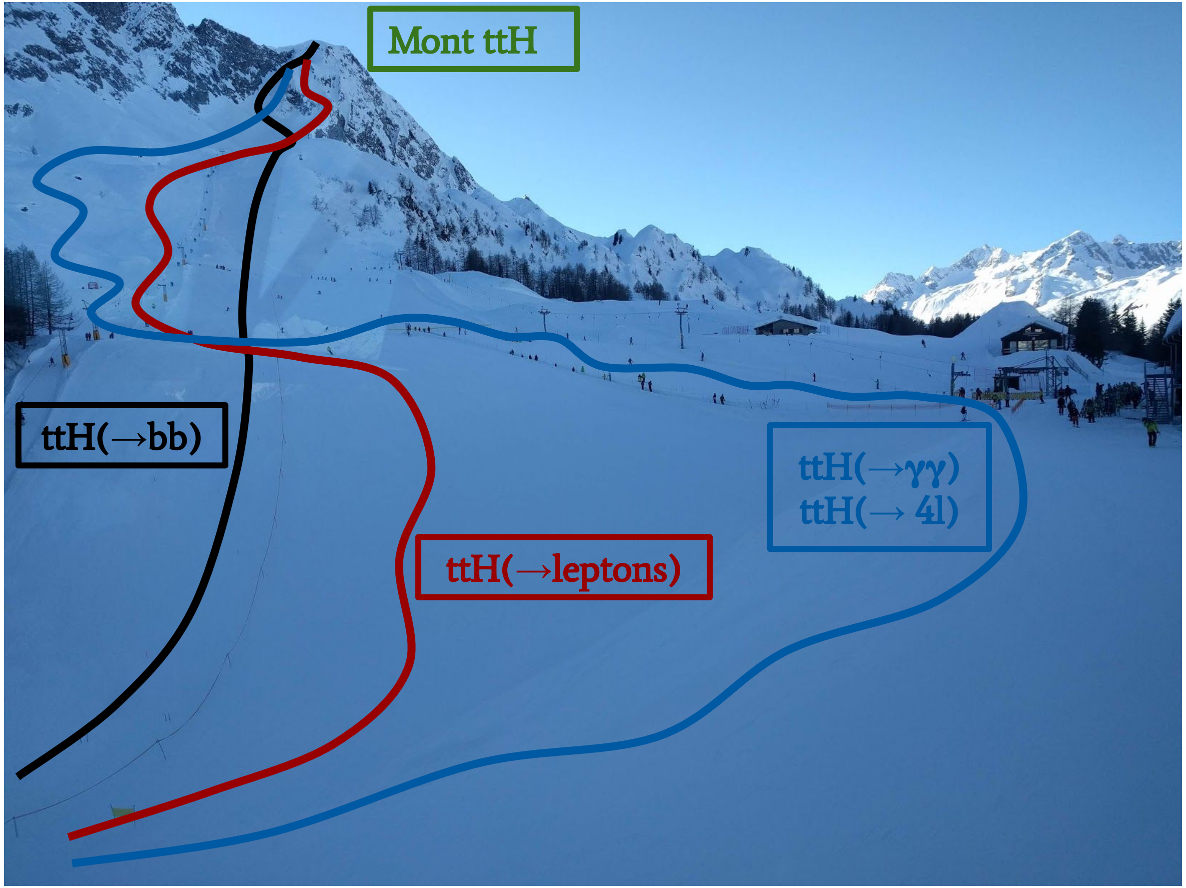
Middle Ground:

ttH(→WW/ττ)

- Faster data taking due to higher BRs
- Background challenges mainly from ttV
  → QCD/EW?
  → Fakes/Exp?

[see Lisa's and Davide's talks]

The ttH background modelling challenge



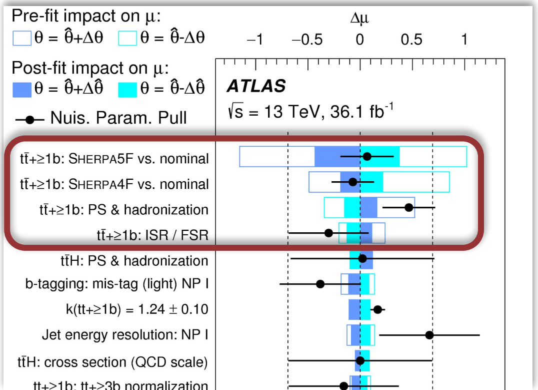
Steep and Dangerous:

ttH(→bb)

- Highest rate
- Many pitfalls!

Most seriously:
Background modelling
for pp → ttbb

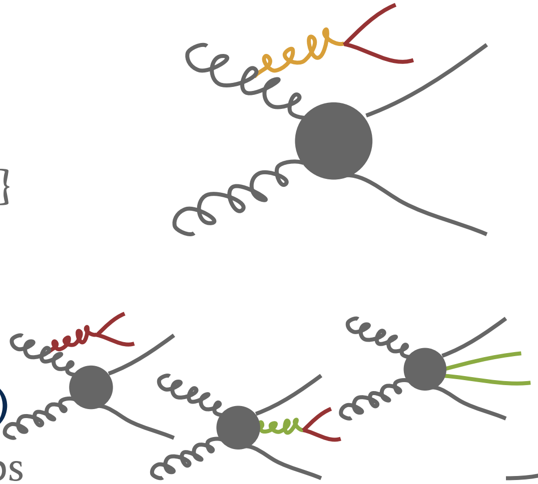
[focus of this talk]



Event generation for ttbb

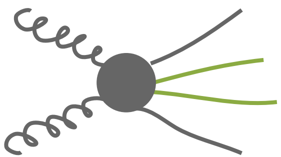
Traditional approaches for ttbb MC predictions:

- "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



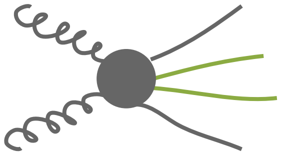
"4-flavour" schemes

Traditional approaches for ttbb 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

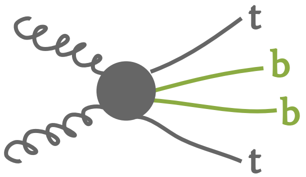
Final state g→bb **dominant**

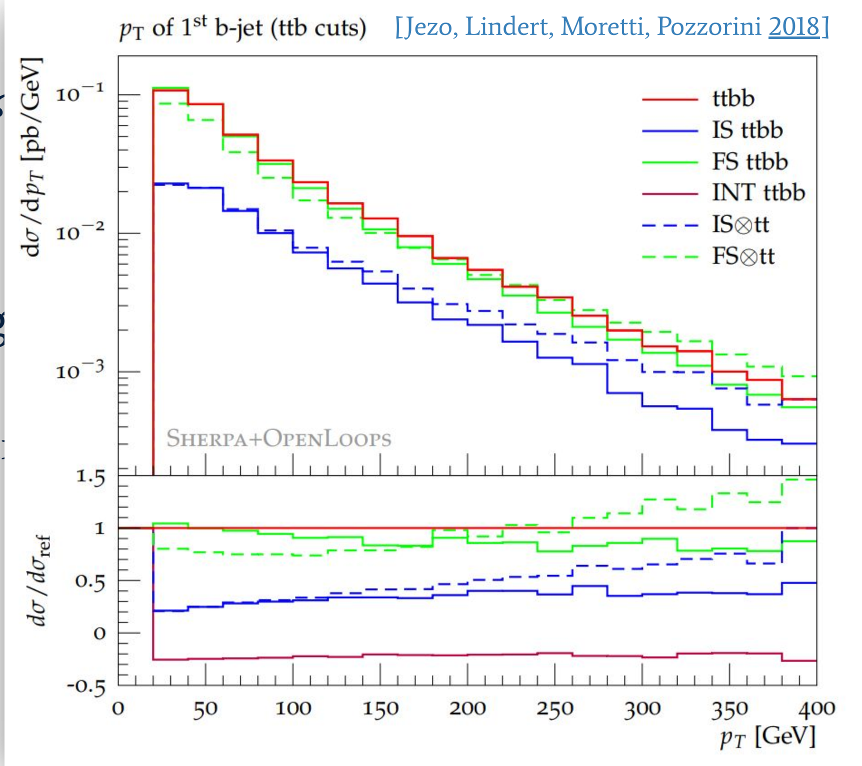
- massive b's → no (jet) cuts!
- collinear g→bb produced in ME

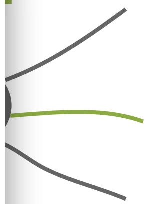
No initial state b in 4FS PDFs

- IS g→bb in ME

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







Final state g … in 4FS PDFs

- massive … ME
- collinea…

Anatomy of 4FS NLO+PS for ttbb

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

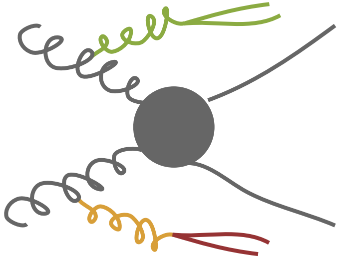
Final state g→bb **dominant**

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

No initial state b in 4FS PDFs

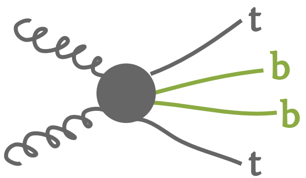
- IS g→bb in ME

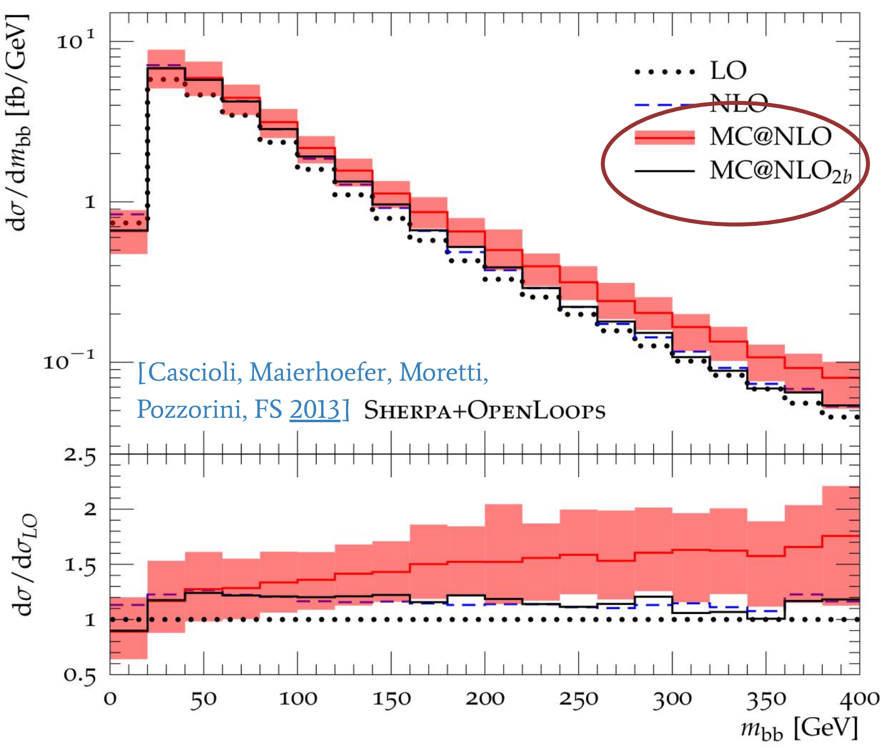
- Matched to parton shower for additional emissions
  - **“double-splitting”** contribution becomes relevant!



Anatomy of 4FS NLO+PS for ttbb

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





Final state … in 4FS PDFs

- massive … ME
- collinea…

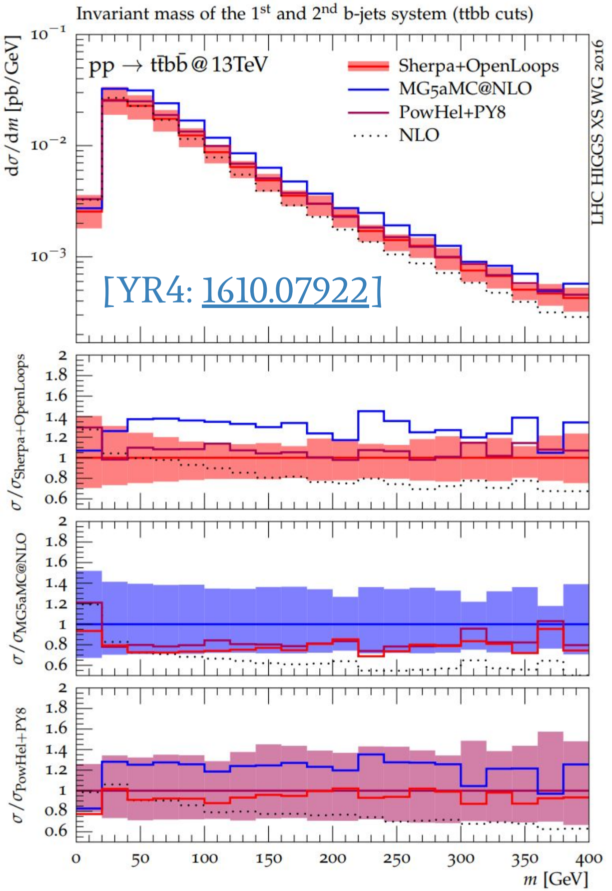
- Matched to par…
additional emis…
  - "double-spli…
  becomes re…

The past

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 ttH(bb)?)?



MC programs for 4FS ttbb at NLO+PS

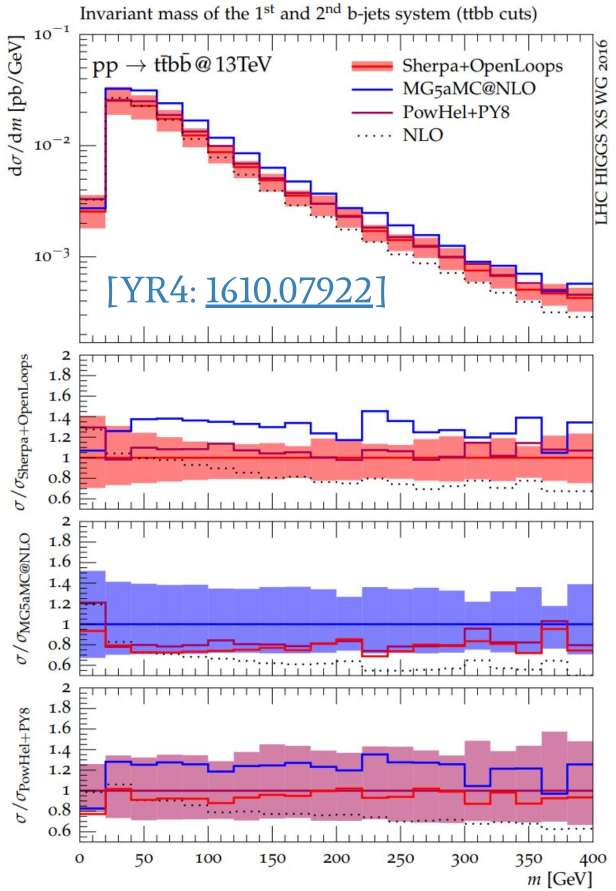
- Several tools on the market
  - Sherpa + OpenLoops [1309.5912]
  - PowHel + Pythia/Herwig [1709.06915]
  - …[1802.00426]

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

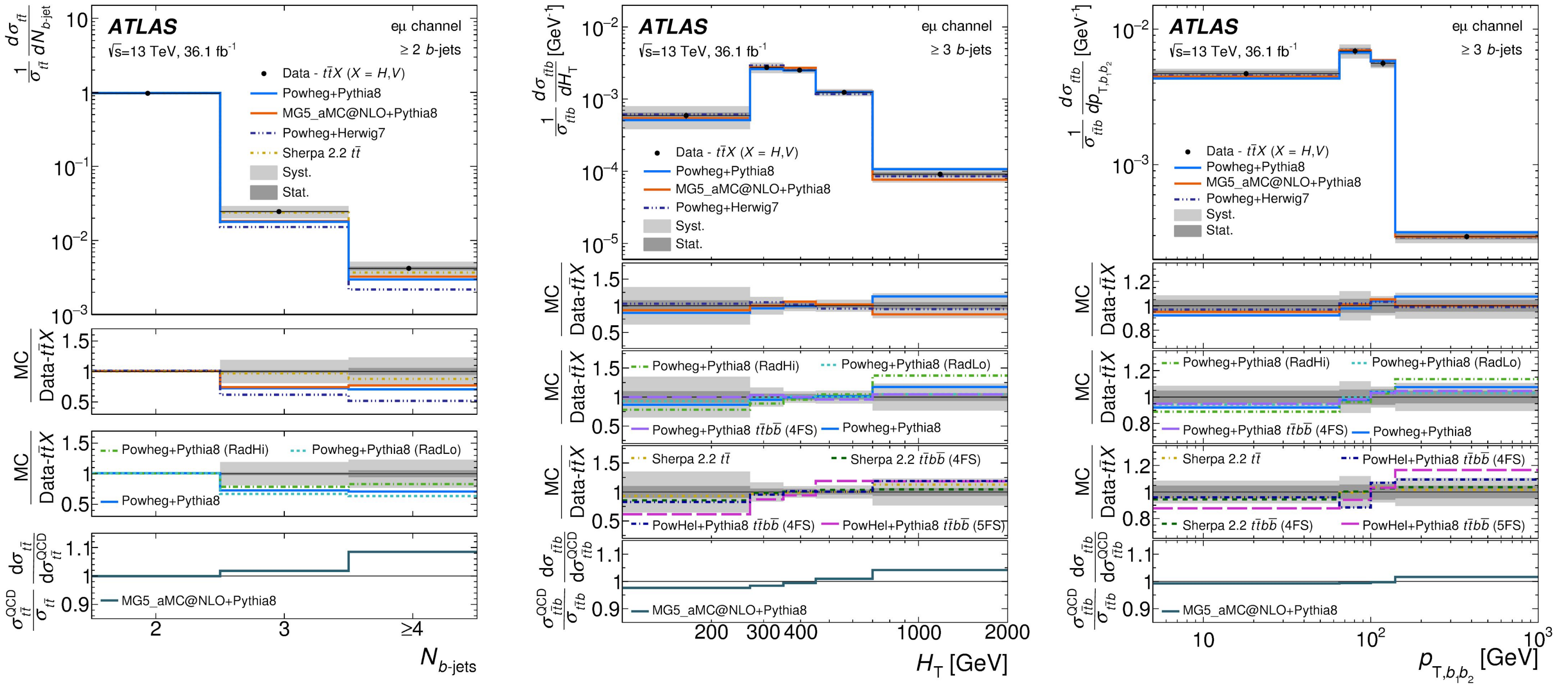
→ Strong challenge to understand unc's as prototype for other processes!

- …
  - » …ching algorithm?

Improve or accept as uncertainties (and kill ttH(bb)?)?



- ATLAS measurement of differential ttbb XS at 13 TeV [1811.12113]

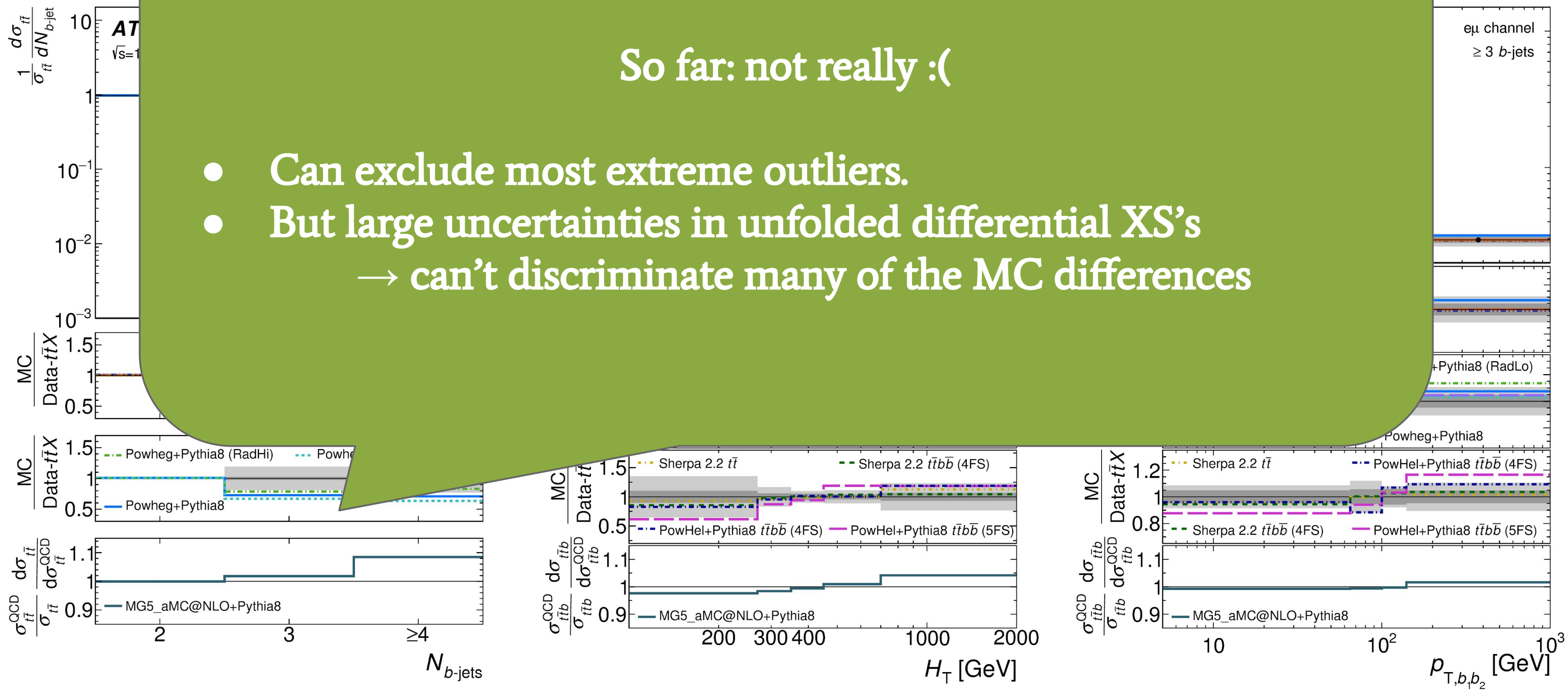


Can we learn from experimental data?

- AT

So far: not really :(

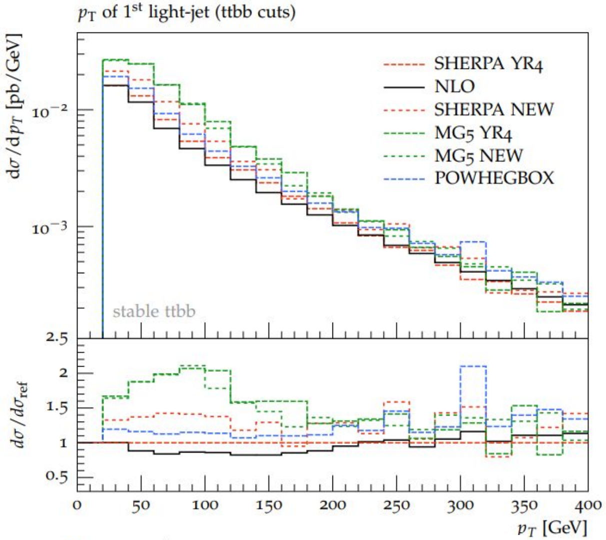
- Can exclude most extreme outliers.
- But large uncertainties in unfolded differential XS's
  → can't discriminate many of the MC differences



The present

Starting point: 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



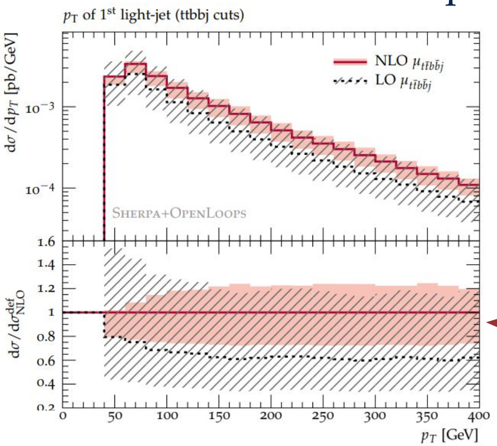
Tool	MC@NLO	Powheg	Pythia 8.2	Herwig 7.1.2	Sherpa2.2.4	MC contacts
SHERPA2.2+OPENLOOPS	x				x	F. Siegert, J. Krause
MG5_AMC@NLO	x		x	x		M. Zaro
MATCHBOX+OPENLOOPS	x			x		C. Reuschle, R. Posdkubka
POWHEG+HELAC		x	x	x		M.V. Garzelli, A. Kardos
POWHEGBOX+OPENLOOPS		x	x	x		T. Jezo, J. Lindert
	3	2	3	4	1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

New input: ttbbj@NLO calculation

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

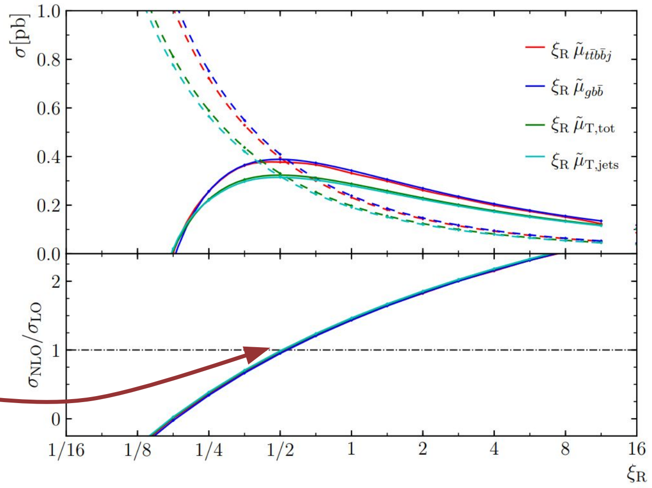
$$d\sigma^{(\mathrm{NLO+PS})} = d\Phi_B \bar{\mathcal{B}} \Big[\underbrace{\Delta(t_0, \mu_Q^2)}_{\text{unresolved}} + \underbrace{\int_{t_0}^{\mu_Q^2} dt \frac{\mathcal{R}_{\mathrm{PS}}}{\mathcal{B}} \Delta(t, \mu_Q^2)}_{\text{resolved, singular} \equiv \mathbb{S}} \Big] + d\Phi_R \underbrace{[\mathcal{R}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

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



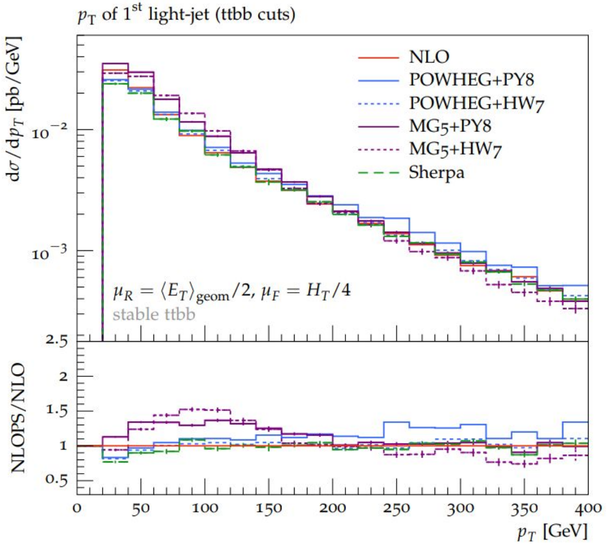
- Fixed-order studies of **ttbbj@NLO** with OpenLoops2+Sherpa [Buccioni, Kallweit, Pozzorini, Zoller 2019]
  - Reduced μ_R stabilises K-factor
  - No significant shape distortions

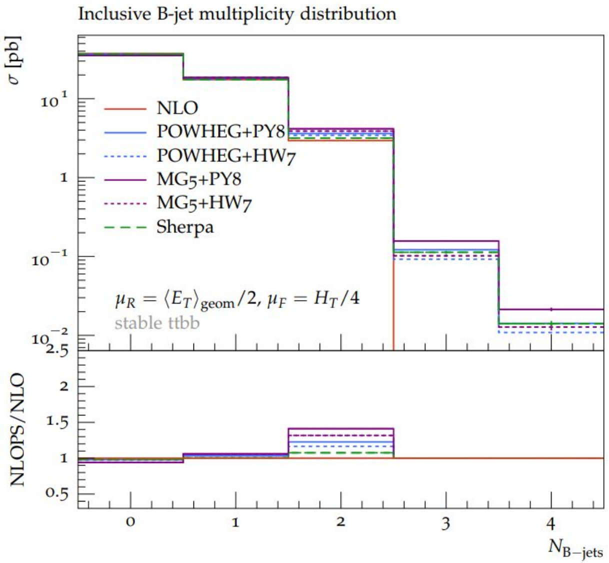
New benchmark for NLO+PS progs!



Tuned matching [Preliminary]

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





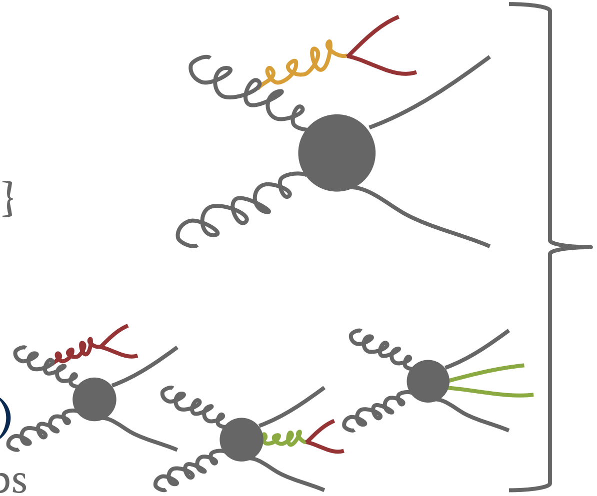
- Direct benchmark studies of **MC vs. ttbbj@NLO** starting now
  → Aim: Uncertainty prescription for experimental analyses

~~The~~ One future

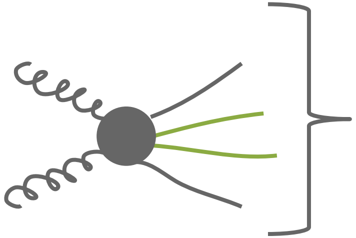
Recap: Event generation for ttbb

Traditional approaches for ttbb 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

Recap: Event generation for ttbb

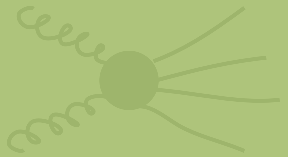
Traditional approaches for ttbb MC predictions:

- "Inclusive" NLO+PS tt sample with HF production from **parton shower g→bb**
  - e.g. {Powheg,aMC@NLO}+{Pythia,Herwig}

"5-flavour" schemes

- Multi-leg merged tt+jets sample with HF from **higher-order MEs** (hard b's) or **parton shower g→bb** (soft/collinear b's)

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

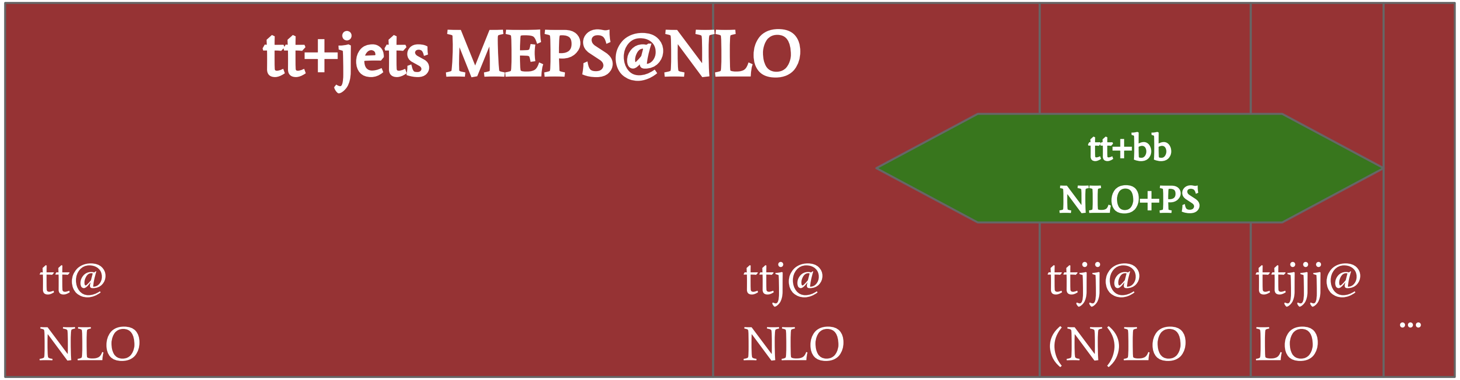


"4-flavour" schemes

Combining 4-flavour ttbb and 5-flavour tt+jets?

Fusing ttbb and tt+jets in the Sherpa MC

aka "Multi-jet merging in a variable flavour number scheme" [Höche, Krause, FS 2019]



Three main ingredients for Fusing:

1. Interpreting ttbb as merged contribution
2. **Overlap removal**
3. Matching 4F/5F in PDFs and α_S

Can be applied for LO and NLO merging!

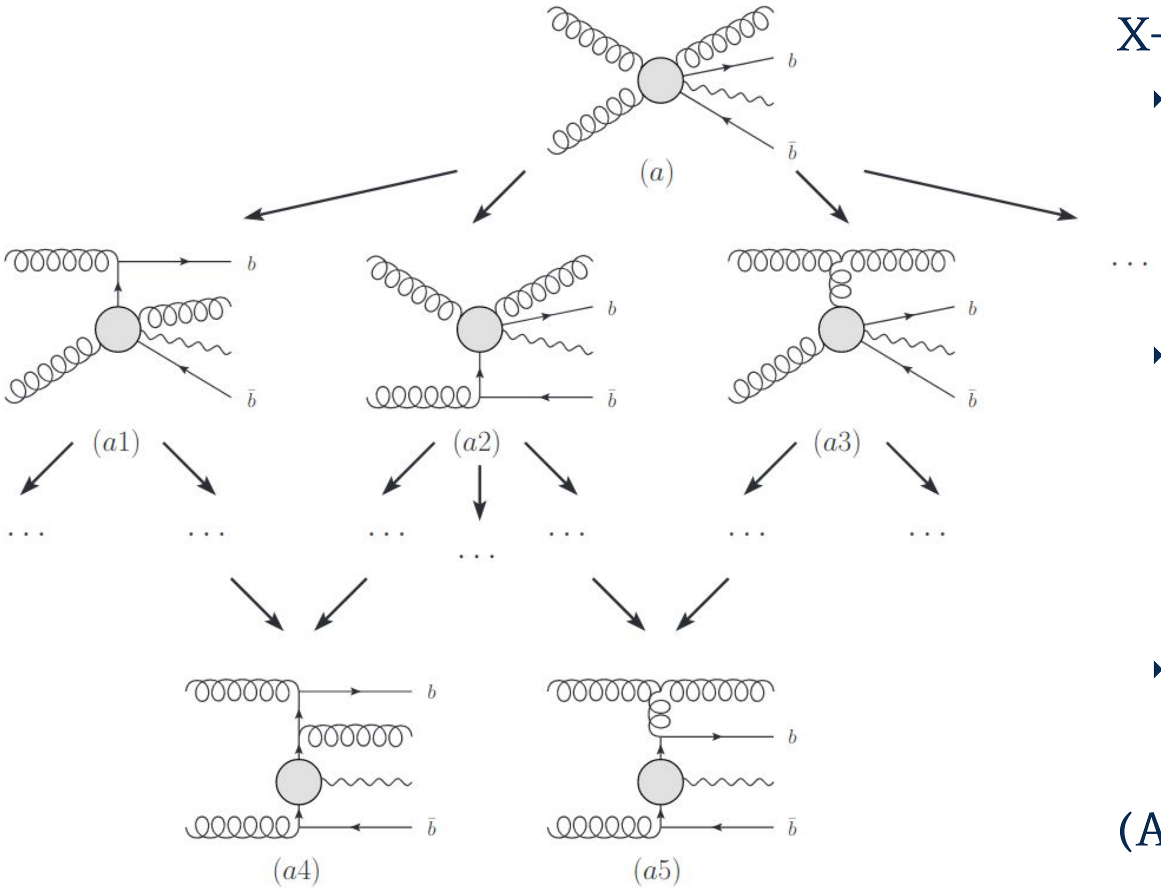
Heavy Flavour Overlap Removal

- 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!)

Example at LO



X+jets candidates:

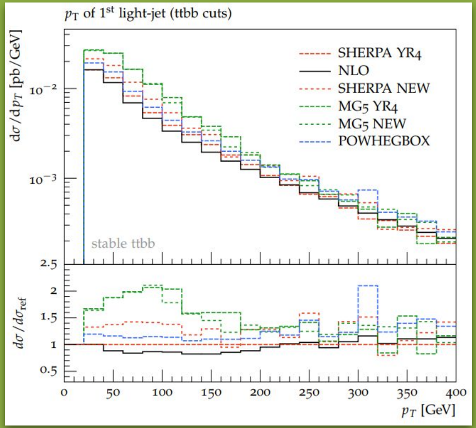
- (a) → (a3) → (a5)
  "soft light jet, hard b-jets"
  → **direct component**
  - vetoed from X+jets
- (a) → (a2) → (a5)
  "soft b-jets, hard light jet"
  → **fragmentation component**
  - kept in X+jets
  - not produced in X+bb (Sud's)
- (a) → (...) → (a4)
  → **fragmentation component**

(At NLO all these would be direct.)

Interesting features of fusing approach

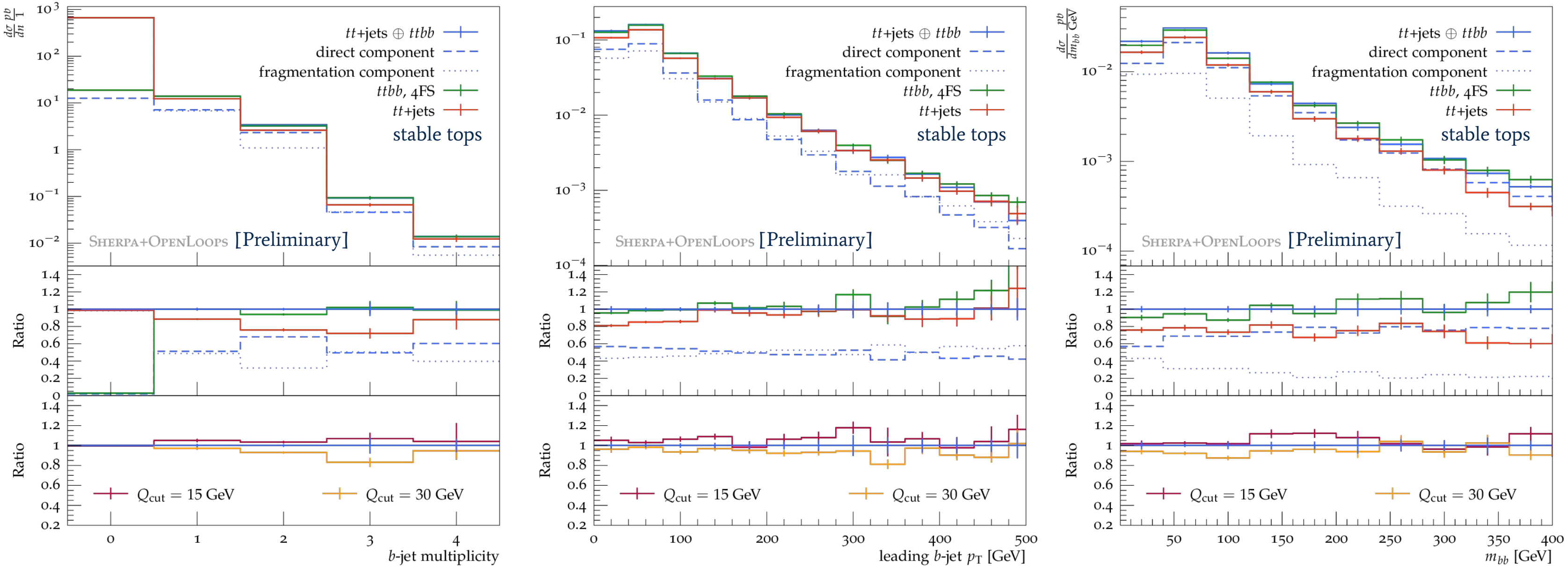
- **"Inverse" hierarchies** treated naturally
  - Emissions from gluon in g->bb relevant for ttbb [Ridolfi, Ubiali, Zaro 2019]
  - Particularly relevant for configurations with hard (additional) gluons and soft b's!
    - Not freaks: $\langle p_T^{jet} \rangle \sim 10 \times \langle p_T^{bjet2} \rangle$!
- Formally reproduces **FONLL** [Forte, Napoletano, Ubiali 2016] at accuracy of ME/PS used
  - "MC implementation" of $\sigma^{\text{FONLL}} = \sigma^{(5)} - \sigma^{(4),(0)} + \sigma^{(4)}$
  - Currently many developments in parton shower land → soon NLL showers based on NLO splitting functions
- Consistent treatment of tt+jets and ttbb: explore **correlation** strategies

Remember:



Sneak preview: Application to ttbb

- Implemented in Sherpa 2.2 and applied as fusion of MEPS@NLO tt + 0,1j@NLO + 2,3j@LO and massive ttbb@NLO



[Katzy, Krause, Pollard, FS in prep]

- 2-bjet production dominated by direct component, but 1-bjet observables with ~equal contributions from direct and fragmentation configurations!

Conclusions

Englisch	Deutsch ▲
–	ADJ schwer | schwerer | am schwersten ... ⊞ SYNO diffizil | heikel ... ⊞
heavy {adj}	**schwer** [auch fig.]
difficult {adj}	**schwer** [hart, anstrengend, schwierig]

- Which problems do we need to tackle?
  - **ATLAS & CMS:** Rely significantly on NLO+PS tt so far!
    More accurate theory with NLO ttbb used only to reweight HF fractions (ATLAS) or for cross-checks (CMS)
  - **Theory:** Large perturbative ttbb uncertainties even enlarged by NLO+PS algorithms
  - **Both**: More rigorous combination of tt+jets and ttbb predictions.

Conclusions

Englisch	Deutsch ▲
–	ADJ schwer | schwerer | am schwersten ... ⊞ SYNO diffizil | heikel ... ⊞
heavy {adj}	**schwer** [auch fig.]
difficult {adj}	**schwer** [hart, anstrengend, schwierig]

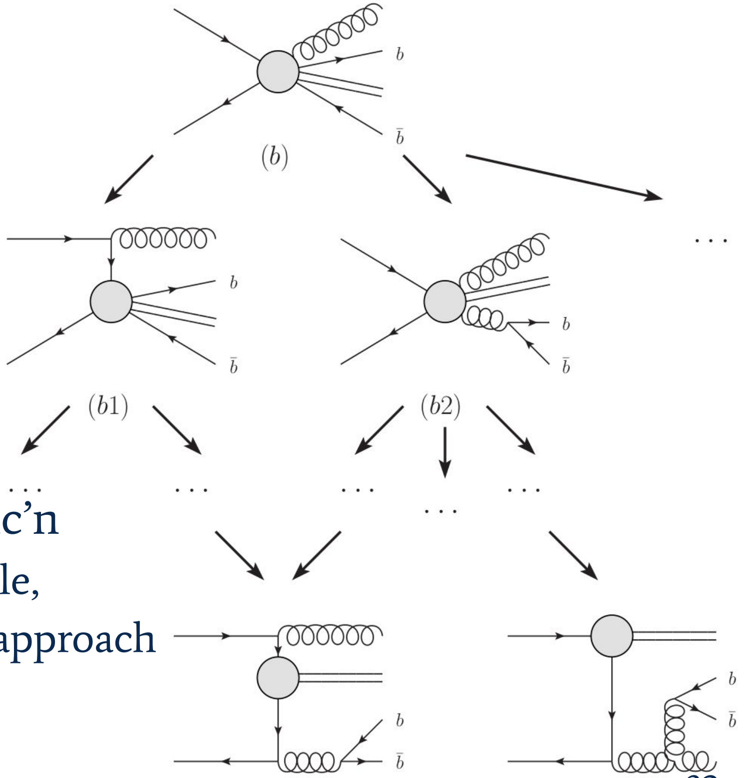
- Which problems do we need to tackle?
  - **ATLAS & CMS:** Rely significantly on NLO+PS tt so far!
    More accurate theory with NLO ttbb used only to reweight HF fractions (ATLAS) or for cross-checks (CMS)
  - **Theory:** Large perturbative ttbb uncertainties even enlarged by NLO+PS algorithms
  - **Both**: More rigorous combination of tt+jets and ttbb predictions.

Thanks for your attention!

Backup

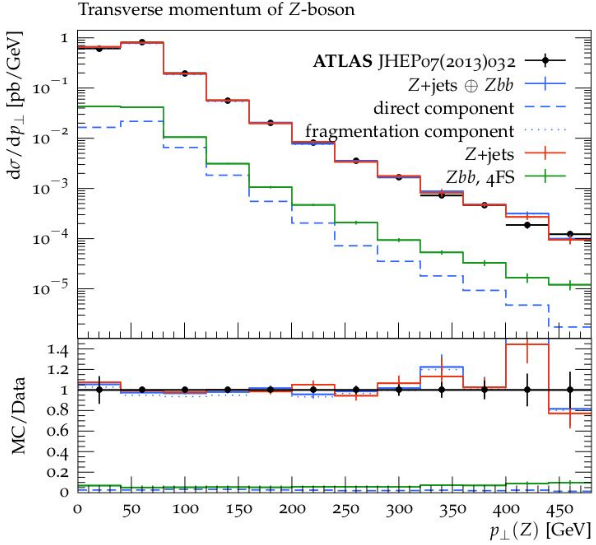
ttbb as merged contribution

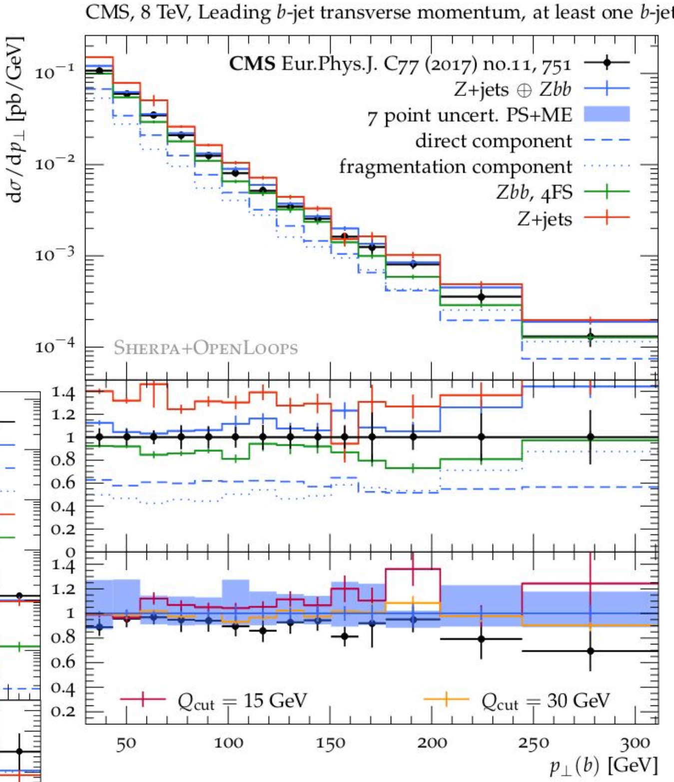
- ttj(j(...)) matrix elements treated in regular **tt+jets MEPS@NLO**:
  - clustering to get topology of ME emissions ("shower history")
  - core scale based on 2→2 process
  - application of $\alpha_S(\mu_R{}^2) \to \alpha_S(p_T{}^2)$ reweighting for each emission
  - application of Sudakov factors $\Delta(t_1, t_2)$ 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

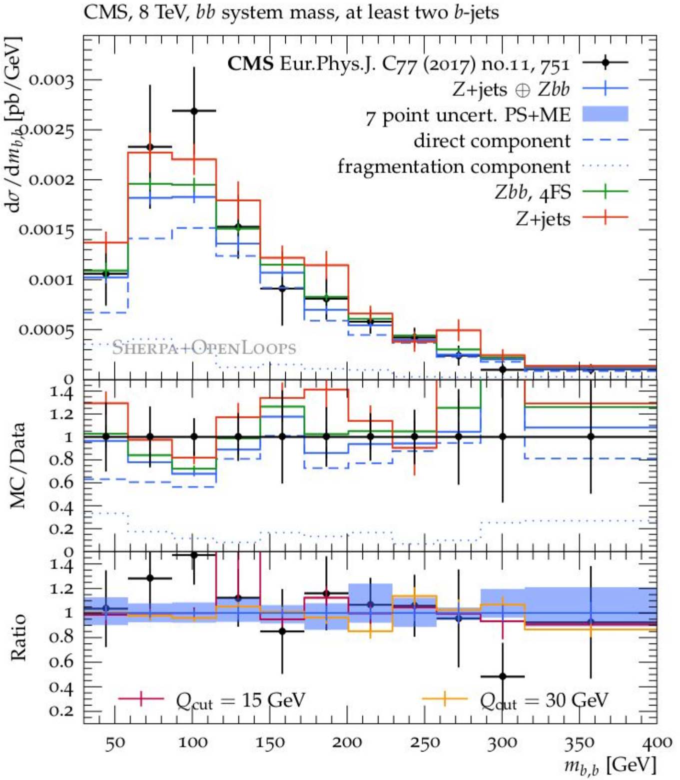


Fusing: Application to Z+HF production

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

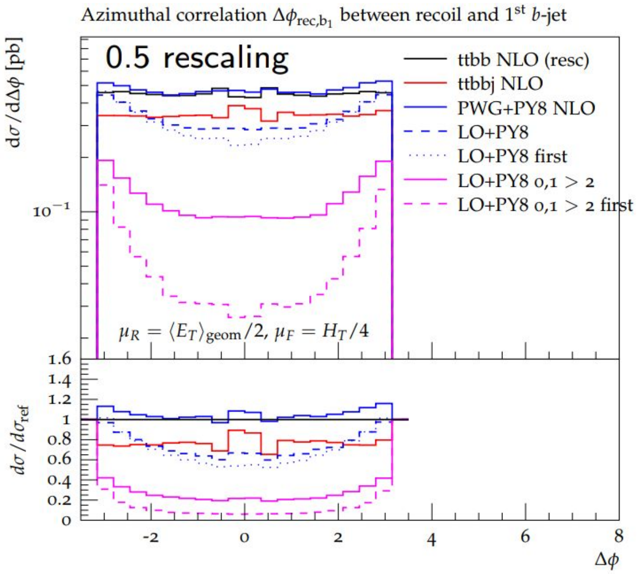






Recoil observables [Preliminary]

- 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



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

How to reduce uncertainties in hard jet configurations?