Five-flavour scheme predictions for $t\bar{t}b\bar{b}$ at next-to-leading order accuracy in MG5_aMC@NLO

Rikkert Frederix (Lund University),
Tetiana Moskalets (Southern Methodist University)

based on: EPJC 84, 763 (2024)

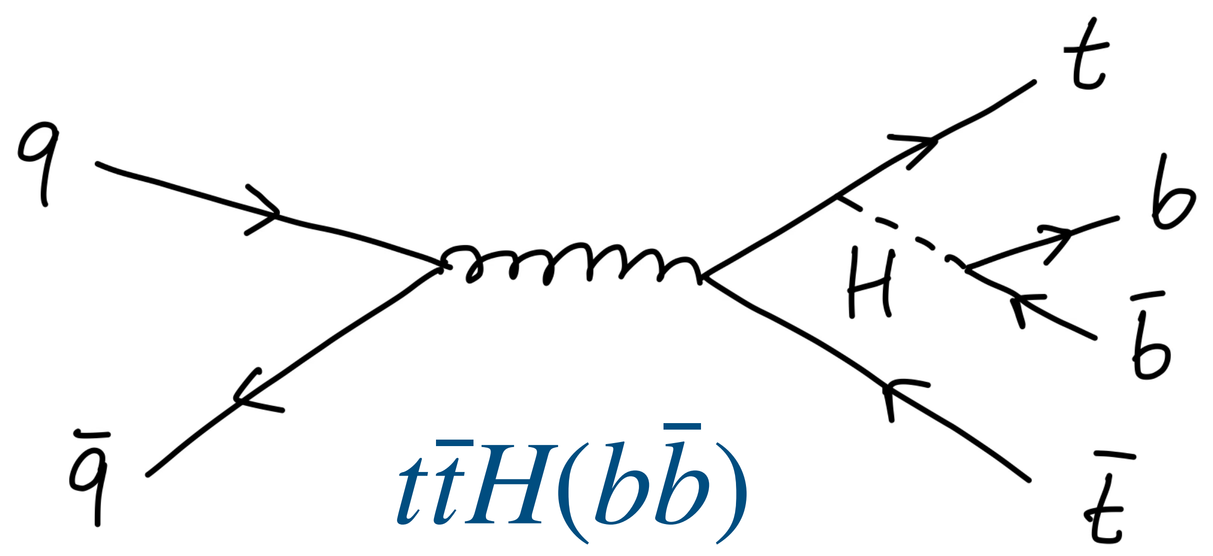
LHC Top WG meeting
11–13 November 2024

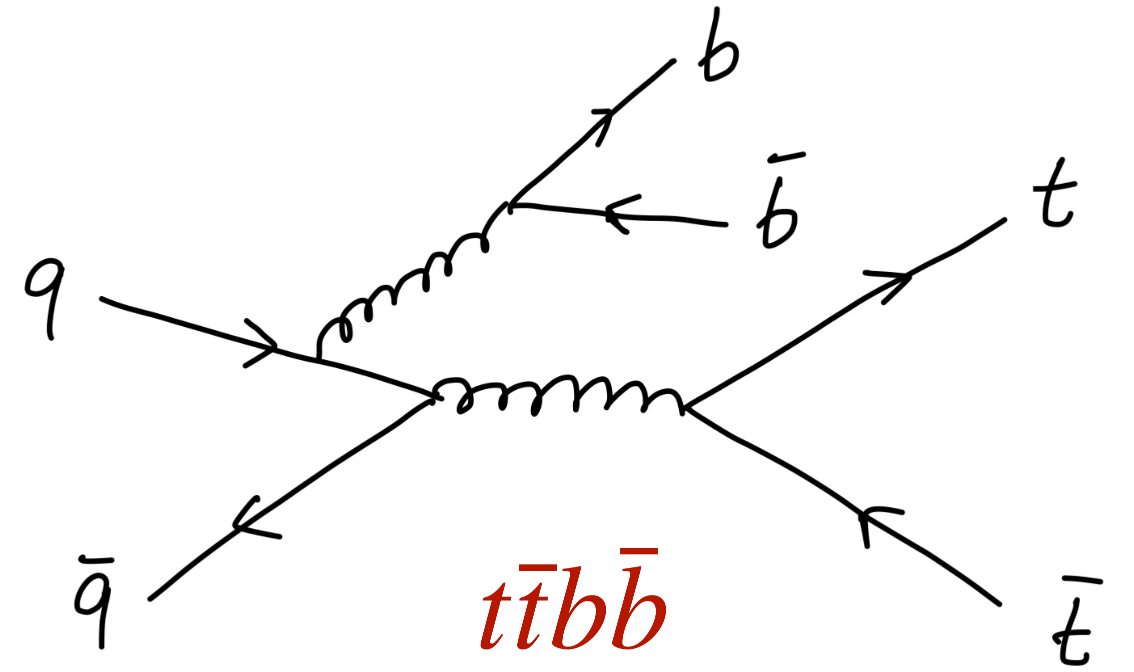
Role of the $t\bar{t}b\bar{b}$ process in the physics analyses

$t\bar{t}b\bar{b}$ represents a significant background in measurements probing the top Yukawa coupling

- $t\bar{t}H(\rightarrow b\bar{b})$ analyses
  - $t\bar{t}b\bar{b}$ is the dominant background
  - modelling uncertainty is currently a limitation

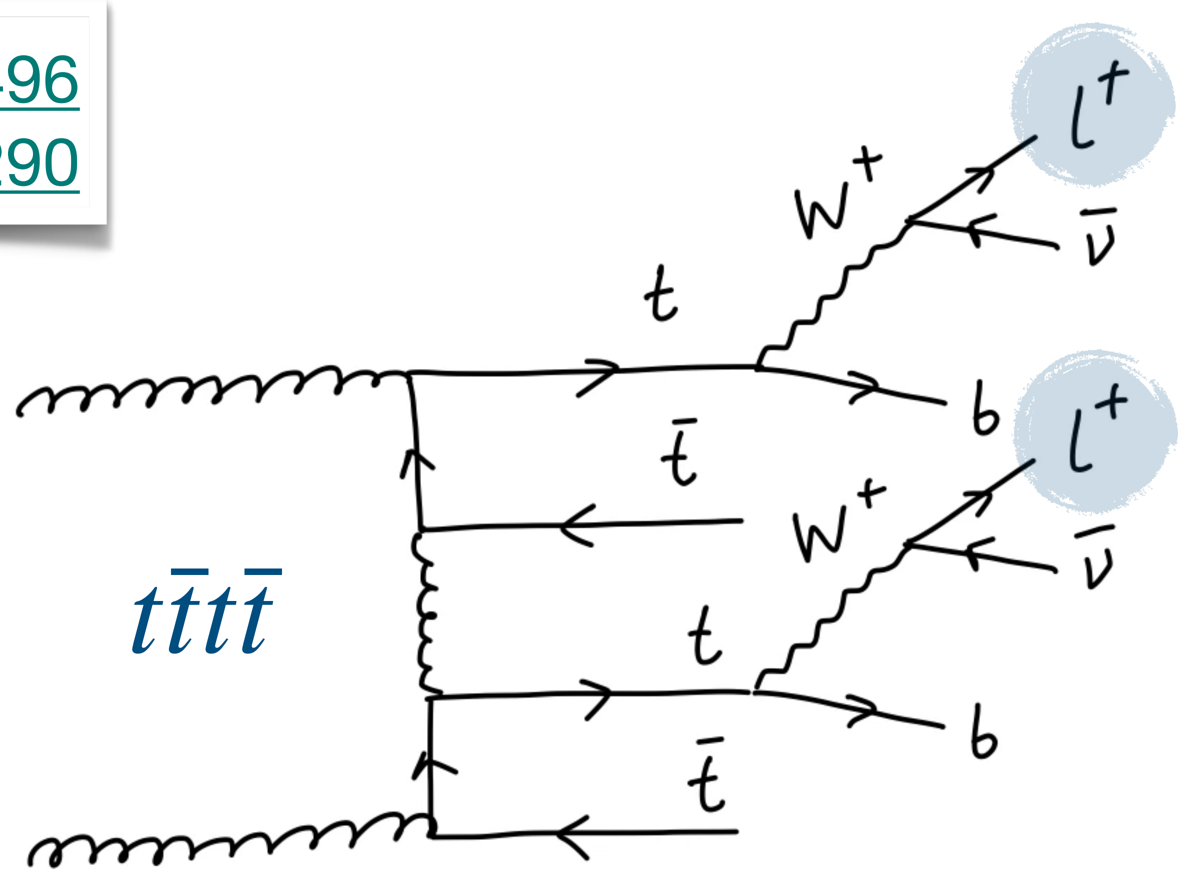
latest $t\bar{t}H(\rightarrow b\bar{b})$ from ATLAS PLB 849 (2024) and CMS arXiv:2407.10896

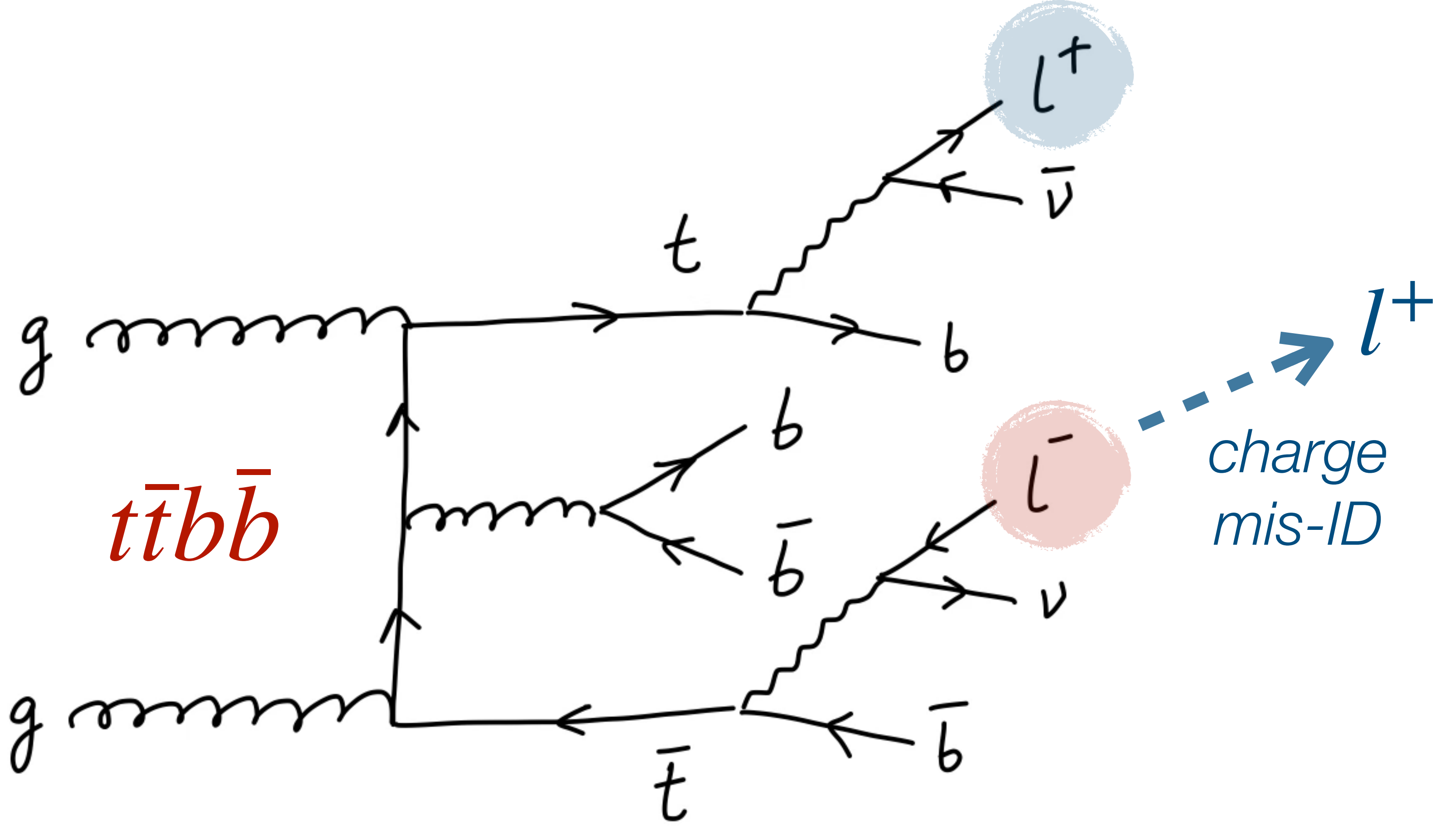




- 4-top analyses
  - $t\bar{t}$+jets (with additional b-jets) is the main source of fake/non-prompt and charge-mis-identification backgrounds

latest $t\bar{t}t\bar{t}$ from ATLAS EPJC 83 (2023) 6, 496 and CMS PLB 847 (2023) 138290

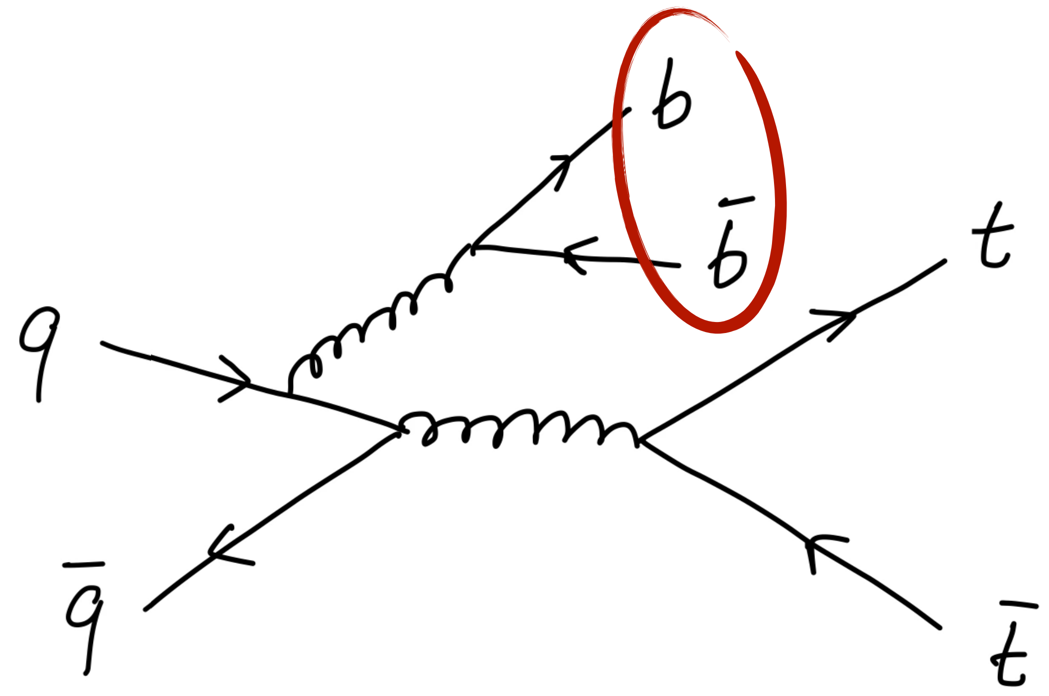


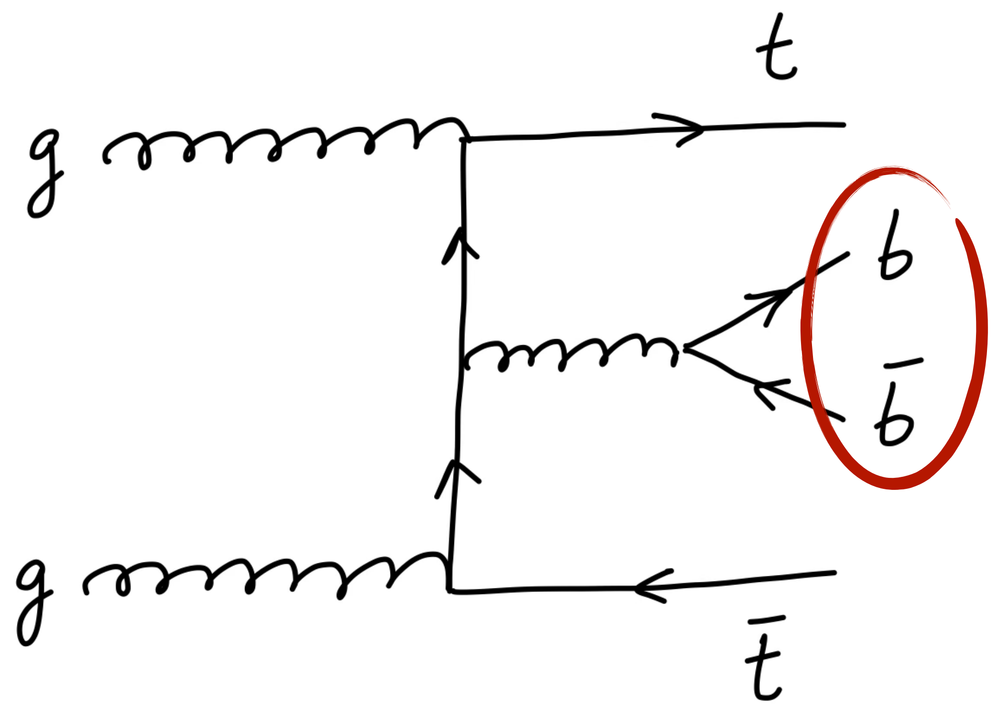


Simulation of the $t\bar{t}b\bar{b}$ process

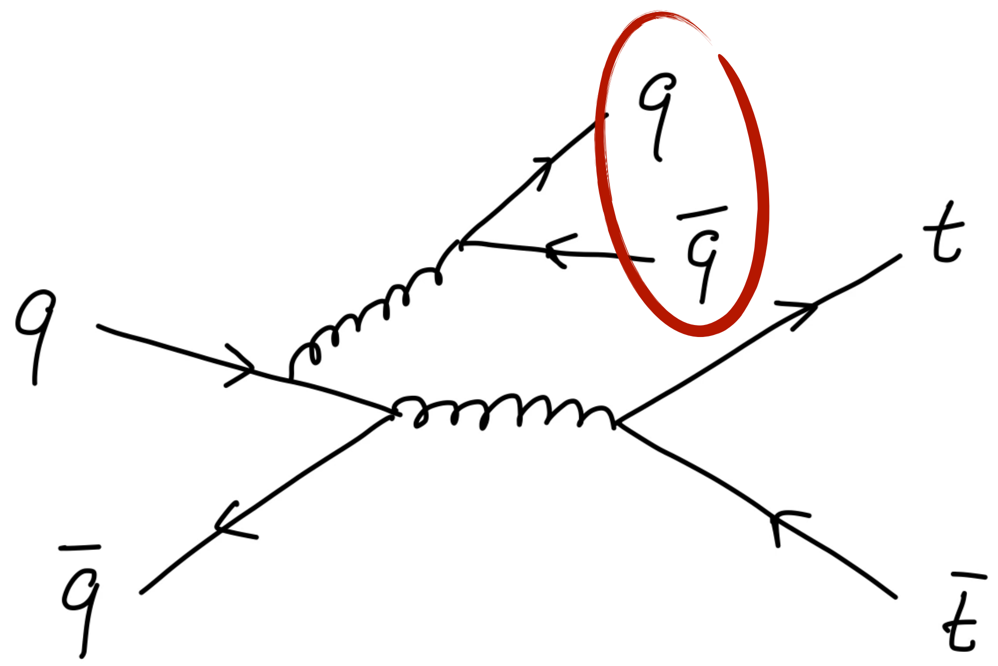
- **Two primary theoretical frameworks:** four-flavour scheme (4FS) and five-flavour scheme (5FS)
- **Alternative:** "fusion" method (or variable flavor number scheme)
  - Merges aspects of both the 4FS and 5FS calculations
  - Currently, the additional jets in the 5FS component are only computed at LO

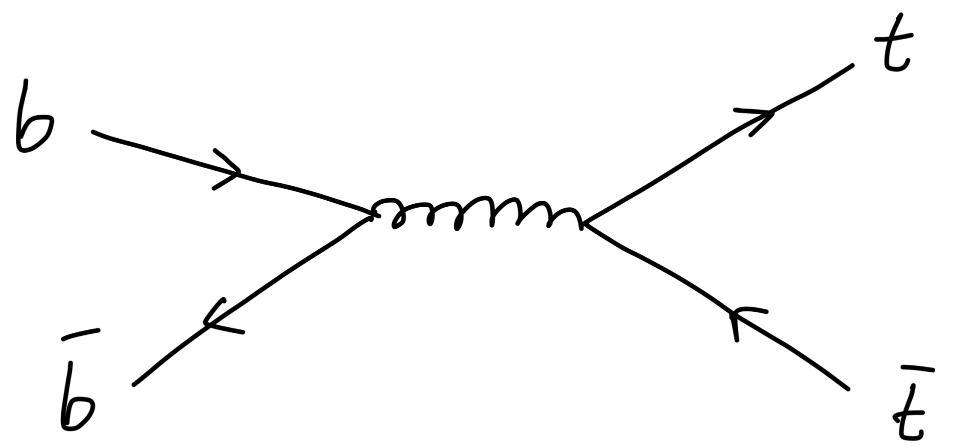
Höche, Krause, Siegert (2019)
Ferencz, Höche, Katzy, Siegert (2024)

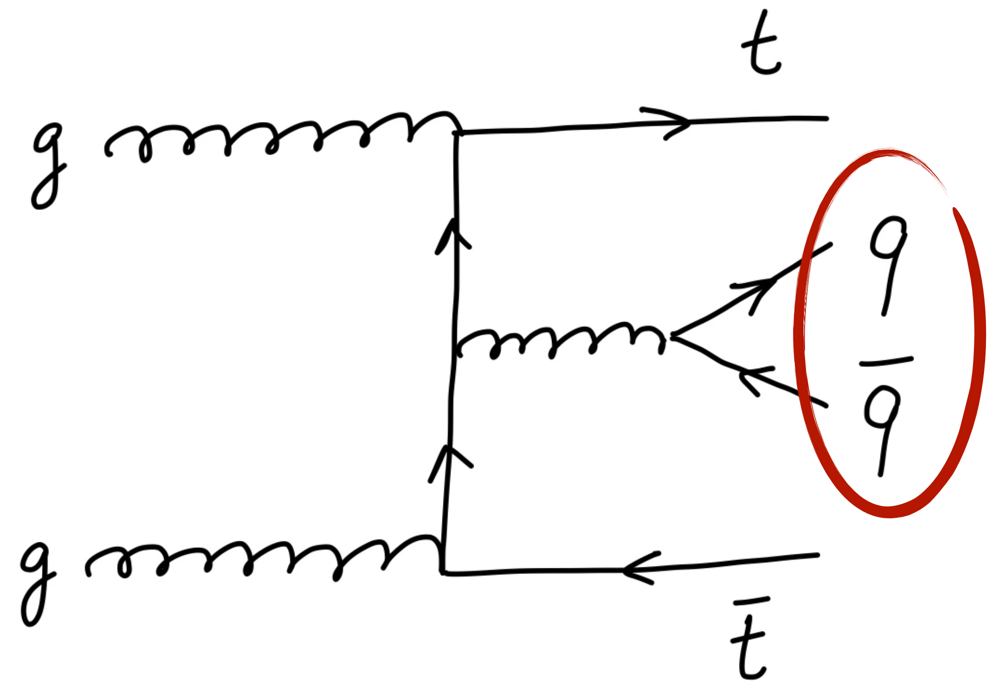




	4FS	5FS
***b*-quarks in the matrix element**	massive	massless
***b*-quarks included in the PDF?**	no	yes
renormalisation scheme	on-shell	$\overline{\text{MS}}$
final state	exclusively $t\bar{t}b\bar{b}$	inclusive $t\bar{t}$ + jets







Simulation of the $t\bar{t}b\bar{b}$ process in the 4FS

Buccioni, Kallweit, Pozzorini, Zoller (2019)

- 4FS calculations are usually the most precise at fixed order
  - b-quark mass effects taken into account
  - The processs can be generated down to any energies

- Calculation with a certain number of jets at fixed order is reliable only if there are no scale hierarchies
  - $t\bar{t}b\bar{b}$ production is a multi-scale process
  - Large mass difference between the top and bottom quarks → **large logarithms** $\left(\log(m_b/p_{\mathrm{T},b}) \text{ or } \log(p_{\mathrm{T},b}/\sqrt{\hat{s}})\right)$
  - Difficult to choose optimal renormalisation and factorisation scales

see the discussion in the LHC Higgs Xsec WG report arXiv:1610.07922

Cascioli, Maierhöfer, Moretti, Pozzorini, Siegert (2014)
Ježo, Lindert, Moretti, Pozzorini (2018)

- Challenges arise when matching to a parton shower:
  - Parton shower radiation can produce additional b-quarks
  - Jets generated by the shower can be harder than the matrix-element-level bottom quarks
  - We need only the subleading b-quarks to come from the parton shower, but not the leading ones
    - Not fully understood how the parton shower radiation should be constrained

Simulation of the $t\bar{t}b\bar{b}$ process in the 5FS

- Generate an inclusive $t\bar{t}$ + jets sample, select b-jets after parton showering

Frixione, Nason, Webber (2003)
Frixione, Nason, Ridolfi (2007)
Hoeche, Krauss, Maierhoefer, Pozzorini, Schonherr, Siegert (2015)
Mazzitelli, Monni, Nason, Re, Wiesemann, Zanderighi (2022)

- Massless b-quarks → **large logarithms do not arise in the matrix element**

- Large scale hierarchies between the top quarks and the jets can be resummed by a multi-jet merging procedure
  - For example, FxFx merging in MadGraph5_aMC@NLO

Frederix, Frixione (2012)

- **Accurate parton-shower approximation for all softer jets**
  - hardest parton shower jets are always softer than the softest matrix element jets, which is not always the case in the 4FS

merging scale

$$p_{\mathrm{T}}(\mathrm{PS\ jets}) \;<\; \mu_Q \;<\; p_{\mathrm{T}}(\mathrm{ME\ jets})$$

except for jets coming from the higher-multiplicity sample

- b-quark mass effects:
  - Important in the collinear/IR region ← incorporated into parton shower splitting fuctions
  - Missing in the matrix element, but they are less relevant for the hard b-quarks

Simulation of the $t\bar{t}b\bar{b}$ process in the 5FS

- But generating $t\bar{t}$ + 0,1,2 jets @ NLO accuracy requires substantional computing resources

Hoeche, Krauss, Schonherr, Siegert (2013) | Frederix, Frixione (2012) | Plätzer (2013)

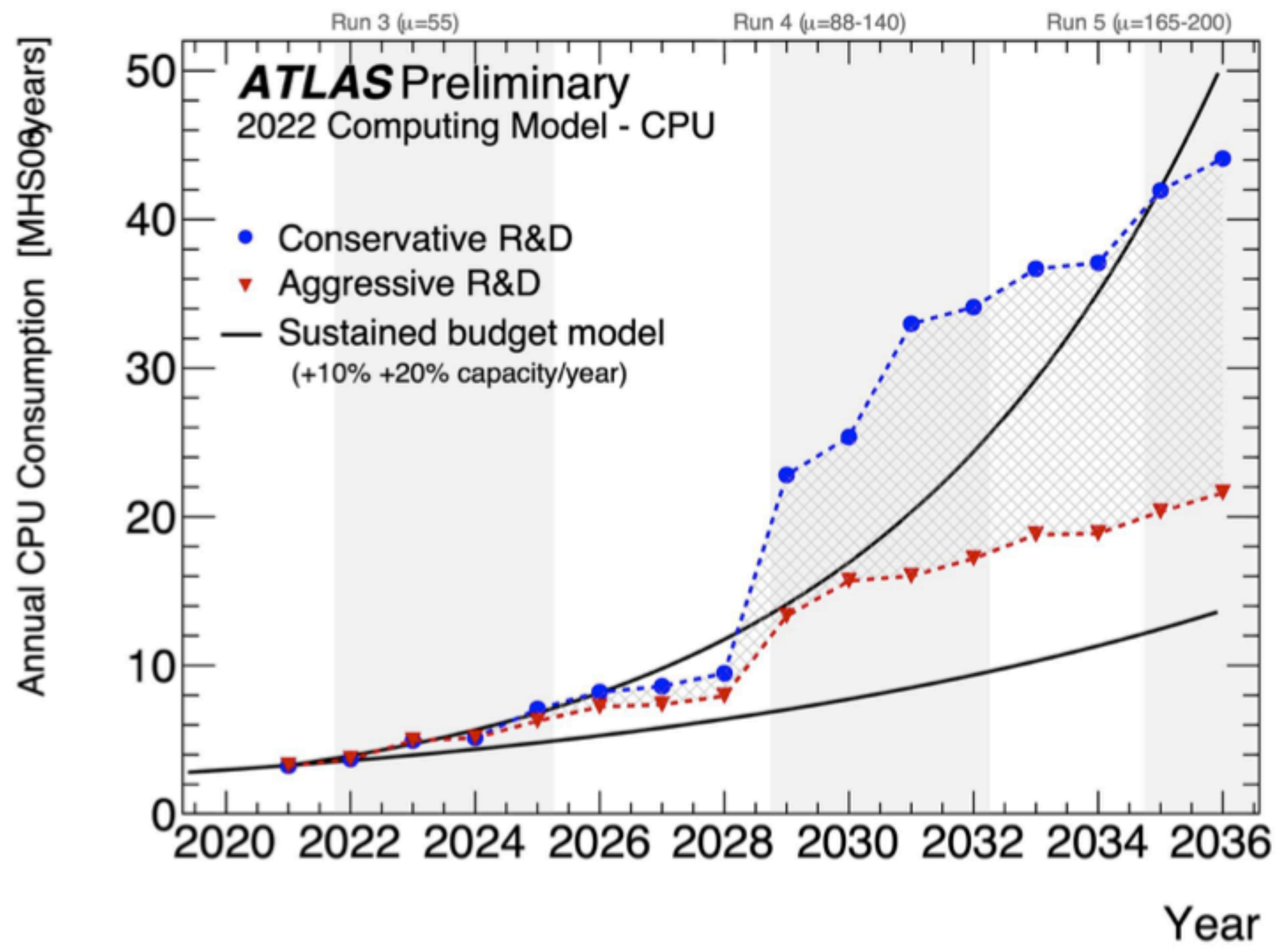
- Selection efficiency of $t\bar{t}b\bar{b}$ is low (percent level)
  - $gg \rightarrow t\bar{t}gg$ dominates

number of instructions to calculate a process in MadGraph5_aMC@NLO

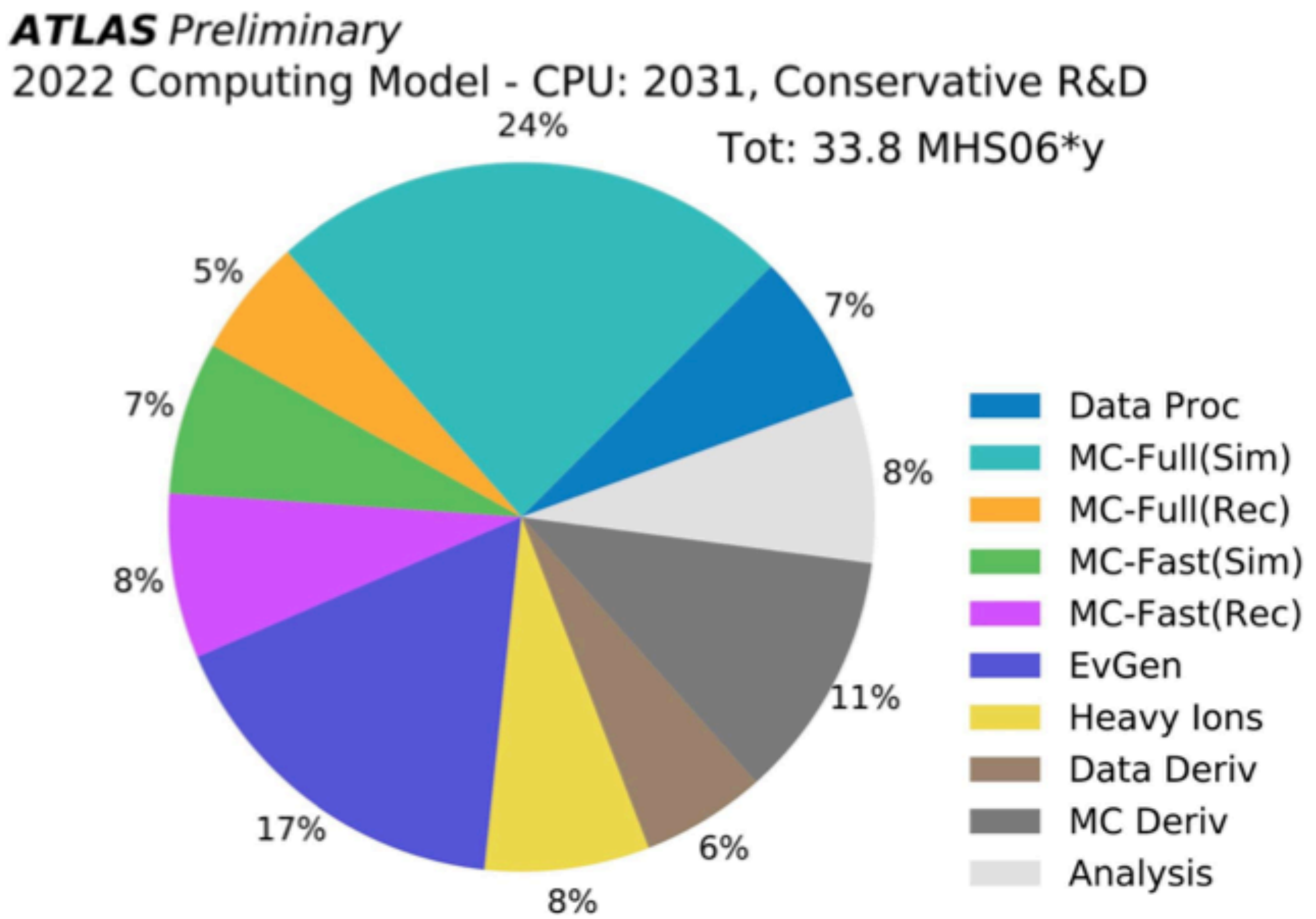
	$gg \to t\bar{t}$	$gg \to t\bar{t}g$	$gg \to t\bar{t}gg$
madevent	13G	470G	11T
matrix1	3.1G (23%)	450G (96%)	11T (>99%)
↳ ext	450M (3.4%)	3.3G (<1%)	7.3G (<1%)
↳ int	1.9G (14%)	160G (35%)	2T (19%)
↳ amp	530M (4.0%)	210G (44%)	5.5T (51%)

from O. Mattelaer's talk

➡ **5FS approach is computationally demanding!**
- This will become even more relevant when producing Monte-Carlo for the HL-LHC era



CERN-LHCC-2022-005

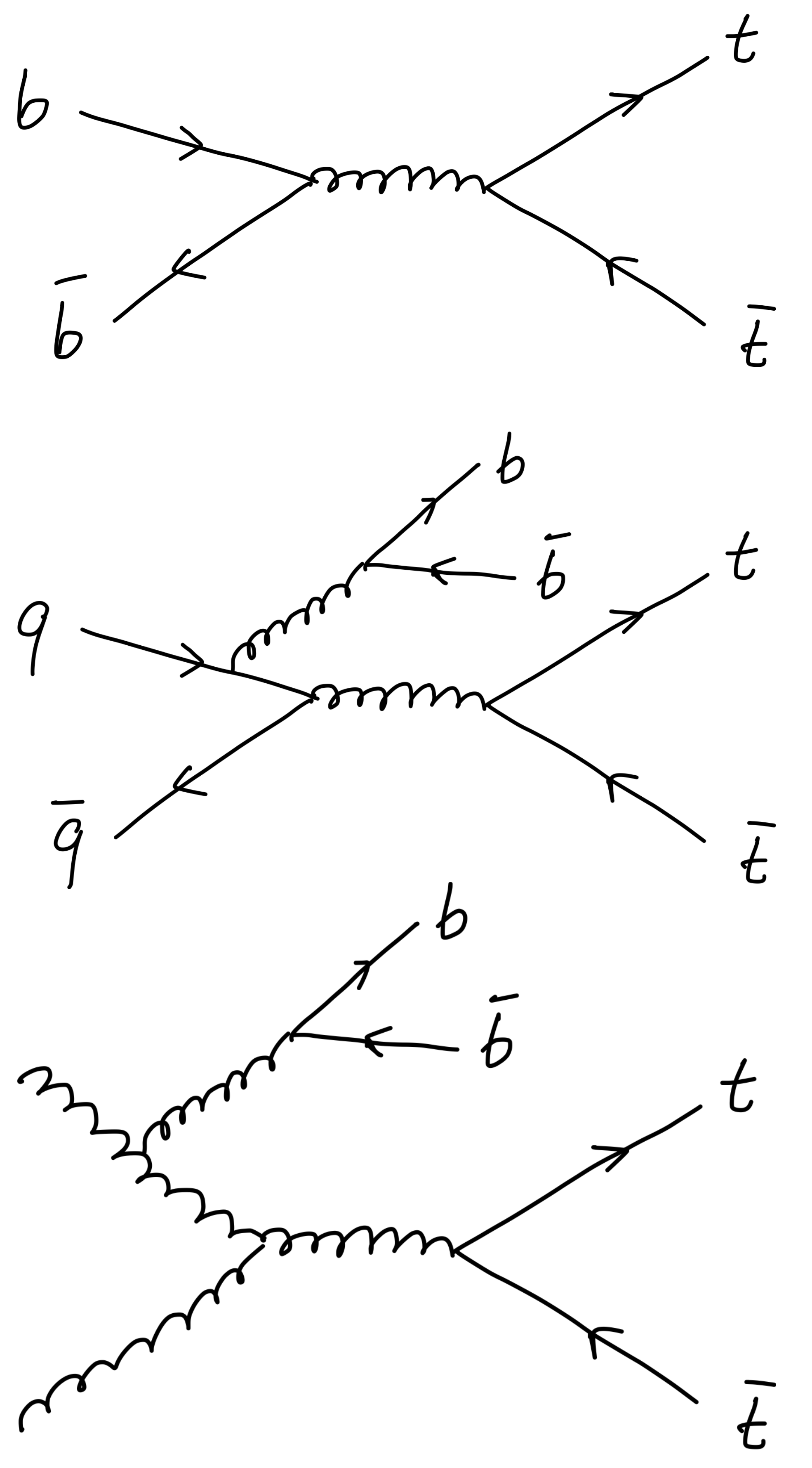


b-flavour enhancement in the matrix element

✓ We proposed a novel method to enhance the b-jet selection efficiency in the 5FS approach

R. Frederix, TM
EPJC 84, 763 (2024)

- Augment the generation probability of bottom quark flavour in the short-distance event generation
  - During phase-space integration and unweighting, multiply the weight of each contribution containing external b-quarks by w_{enh}
  - For bottom quarks can be generated in the initial or final state
    - $gg \to t\bar{t}b\bar{b}(g)$
    - $gb \to t\bar{t}bg(\to b\bar{b})$
    - $bb \to t\bar{t}q\bar{q}(g)$
    - …
- To compensate for this and to preserve the cross-section, multiply the weight of events with external b-quarks by $1/w_{\text{enh}}$



examples of the enhanced subprocesses

b-flavour enhancement in the matrix element

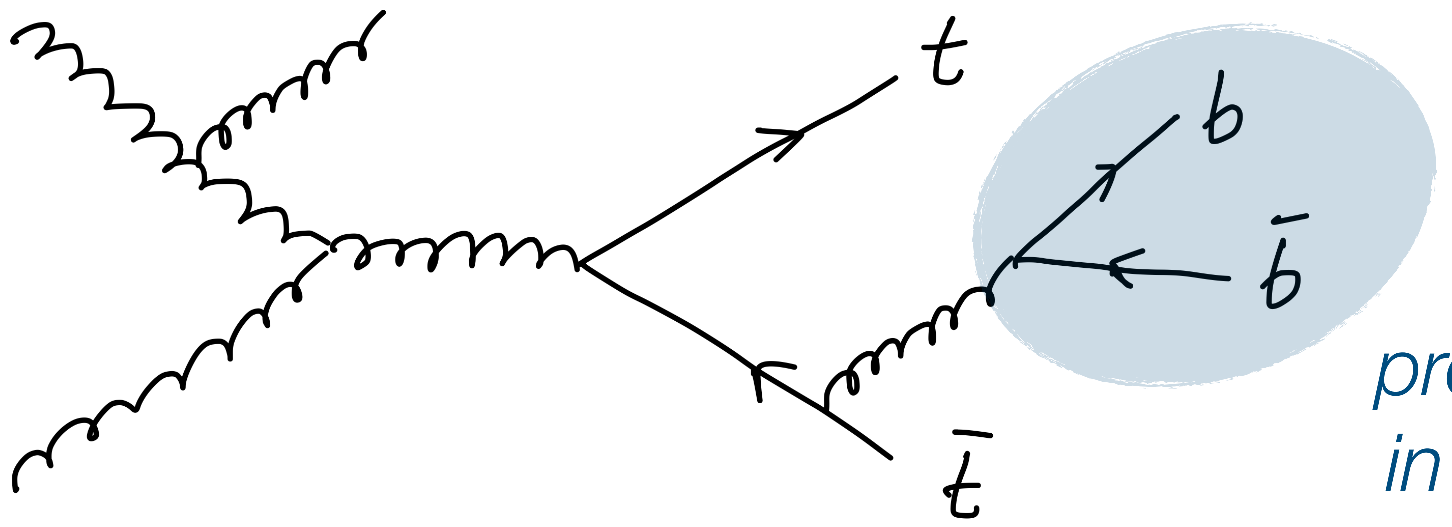
✓ We proposed a novel method to enhance the b-jet selection efficiency in the 5FS approach

R. Frederix, TM
EPJC 84, 763 (2024)

- This procedure is implemented in the MadGraph5_aMC@NLO
  - enhancement factor w_{enh} can be set by a new parameter, `bflav_enhancement`, in the runcard file
  - **The new feature will become part of an upcoming release**

✳ ***NB:*** *hard processes like $gg \to t\bar{t}gg$ which can yield a $t\bar{t}b\bar{b}$ event after a $g \to b\bar{b}$ splitting in the parton shower will not get enhanced $\Rightarrow$ the fraction of $t\bar{t}b\bar{b}$ events is increased by a factor smaller than w_{enh}*

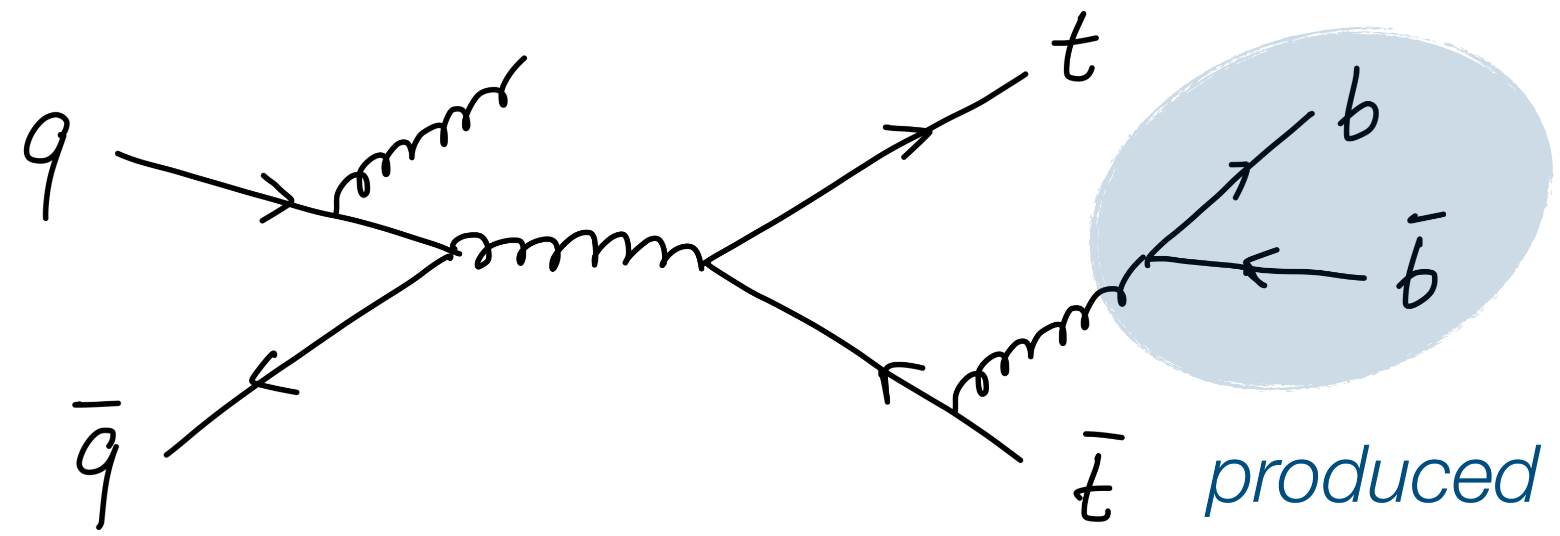
- *also, too high enhancement factors (>100) cause instabilities which result in large statistical fluctuations*



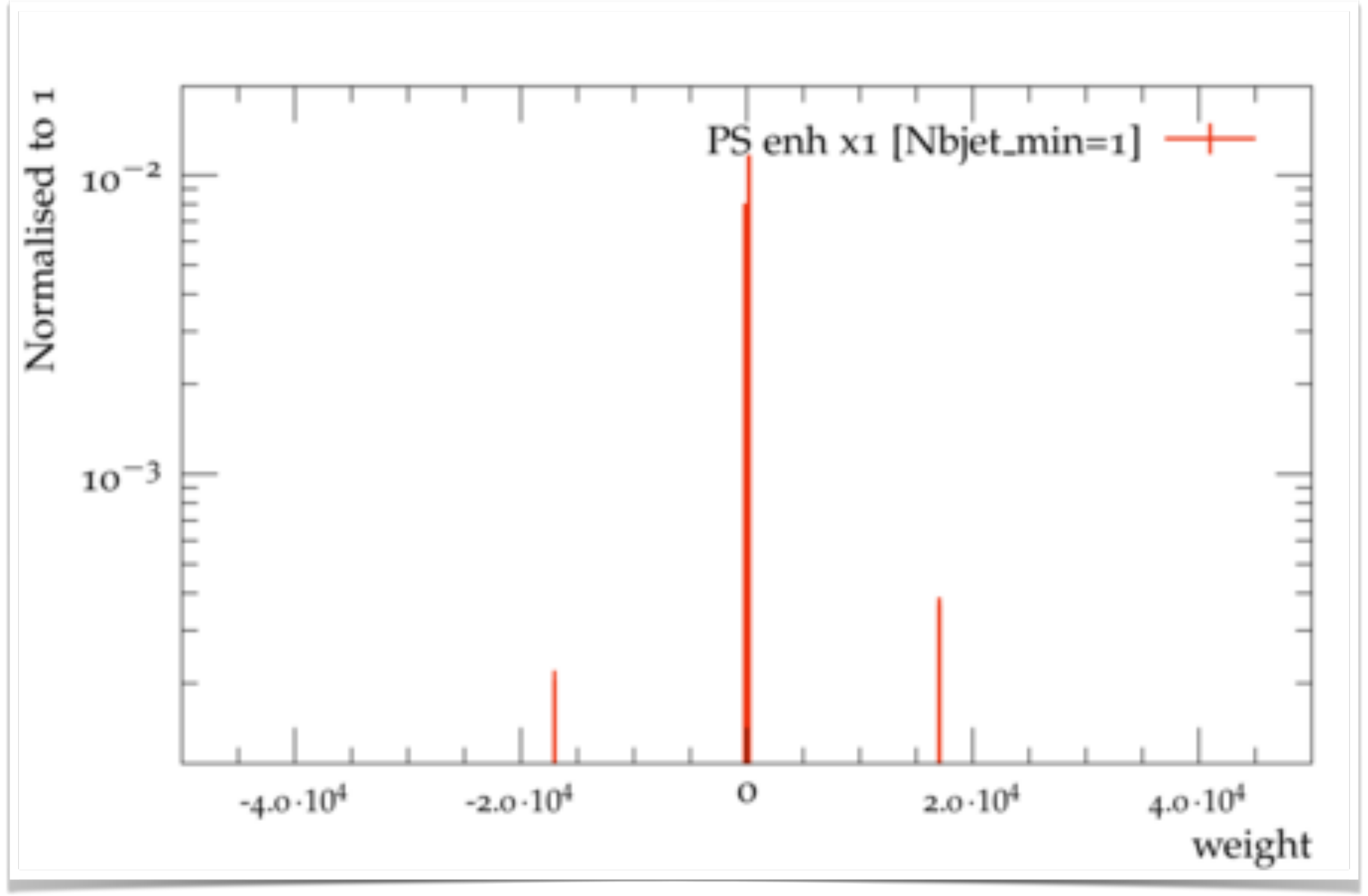
this diagram is not enhanced — can we enhance it in the PS?

b-flavour enhancement in the parton shower?

- A similar biasing strategy can be potentially applied in the parton shower
- Pythia8 has a built-in mechanism for enhancing splitting probabilities, in particular $g \to b\bar{b}$ ones
  - In versions ≥ 8.311
- In practice:
  - Even moderate enhancement in the PS causes significant widening of the event weight distribution
  - **Large weights deteriorate the statistics** → cancels the improvement from the b-enhancement completely

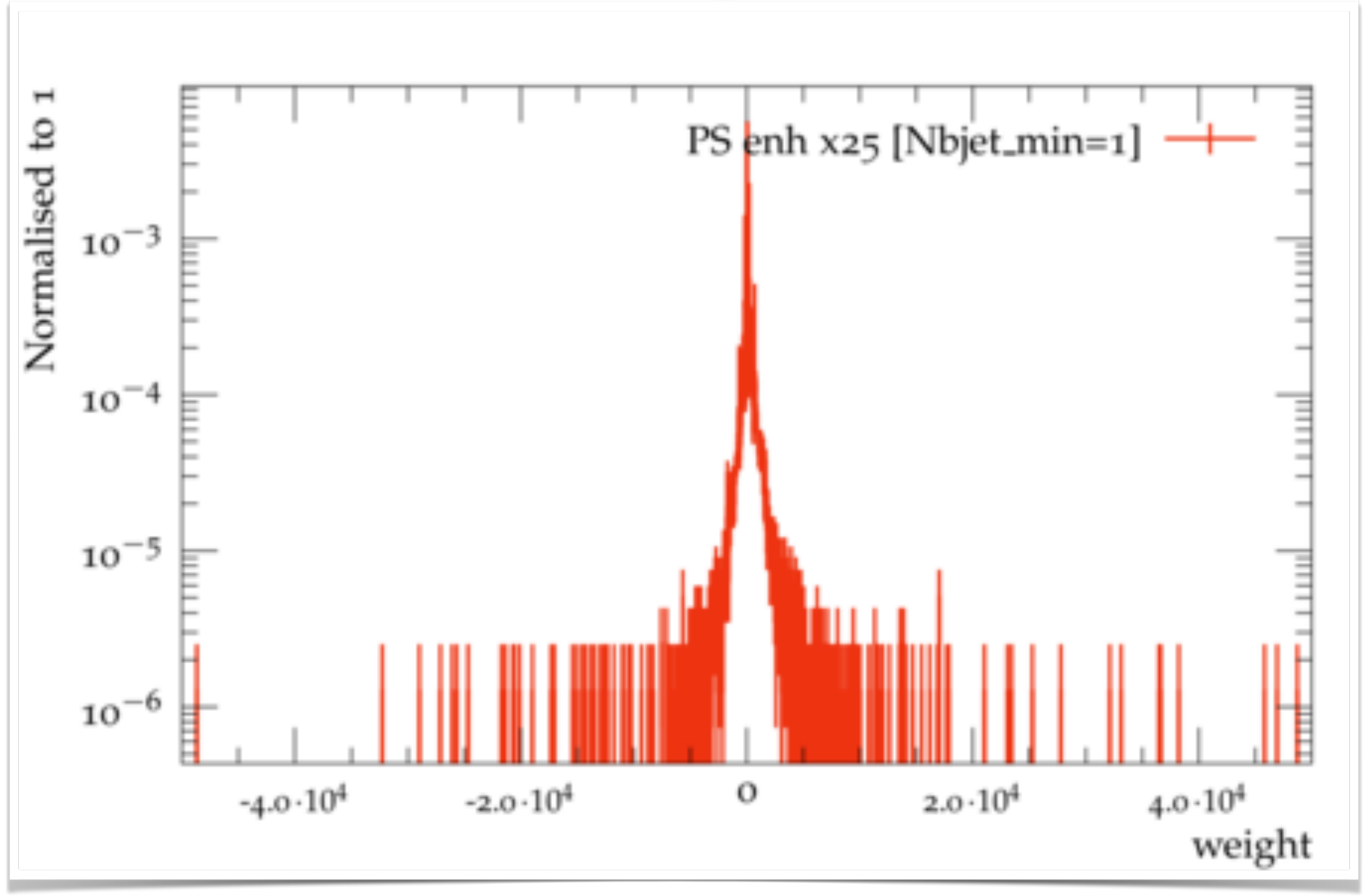


event weights w/o enhancement in the PS



$g \to b\bar{b}$ enhancement →

event weights with enhancement in the PS



Generation setup for the 5FS sample and comparison to the 4FS

5FS $t\bar{t}$ + jets sample with b-enhancement in the ME

R. Frederix, TM
EPJC 84, 763 (2024)

- **MadGraph5_aMC@NLO $t\bar{t} + 0,1,2$ jets @NLO sample, FxFx merged**
- **Enhancement factor $w_{\text{enh}} = 100$**
- **Renormalisation/factorisation scales:** central values for are taken from the FxFx merging
  - 7-point variations
- **Merging scale:** 40 GeV
  - variations: 70 and 100 GeV
- **Shower starting scale:** $H_T/2$
  - variation: $H_T/4$
- Generation-level cut of 20 GeV on jet p_T
- **Matched to the Pythia8 parton shower**
- **Not including**:
  - hadronisation
  - underlying events
  - top quark decay

taking an envelope as a total uncertainty

to reduce the generation time and to simplify the analysis, and because we focus on the differences in the ME

Truth-level analysis in Rivet:

- anti-k_T jets ($R > 0.4$)
  - $p_T > 25$ GeV
  - $|\eta| < 2.5$
- jets containing at least one bottom quark are identified as b-jets
- consider two scenarios:
  - at least 1 b-jet
  - at least 2 b-jets

4FS $t\bar{t}b\bar{b}$ sample

R. Frederix, TM
EPJC 84, 763 (2024)

- **MadGraph5_aMC@NLO+Pythia8** NLO+PS $t\bar{t}b\bar{b}$ sample
- **Renormalisation/factorisation scales**:
  - central values:

$$\mu_R = (E_{T,t} E_{T,\bar{t}} E_{T,b} E_{T,\bar{b}})^{1/4}$$
$$\mu_F = \frac{1}{2}(E_{T,t} + E_{T,\bar{t}} + E_{T,b} + E_{T,\bar{b}})$$

  - 7-point variations

following the recommendations in the LHC Higgs Xsec WG report arXiv:1610.07922

- **Shower starting scale:** $H_T/2$
- Generation-level cut of 20 GeV on jet p_T
- **Matched to the Pythia8 parton shower**
- **Not including**:
  - shower starting scale uncertainty
  - matching scheme uncertainty
  - hadronisation
  - underlying events
  - top quark decay

← *expected to be sizeable, (see the LHC HXS WG report) but is non-trivial to assess exactly*

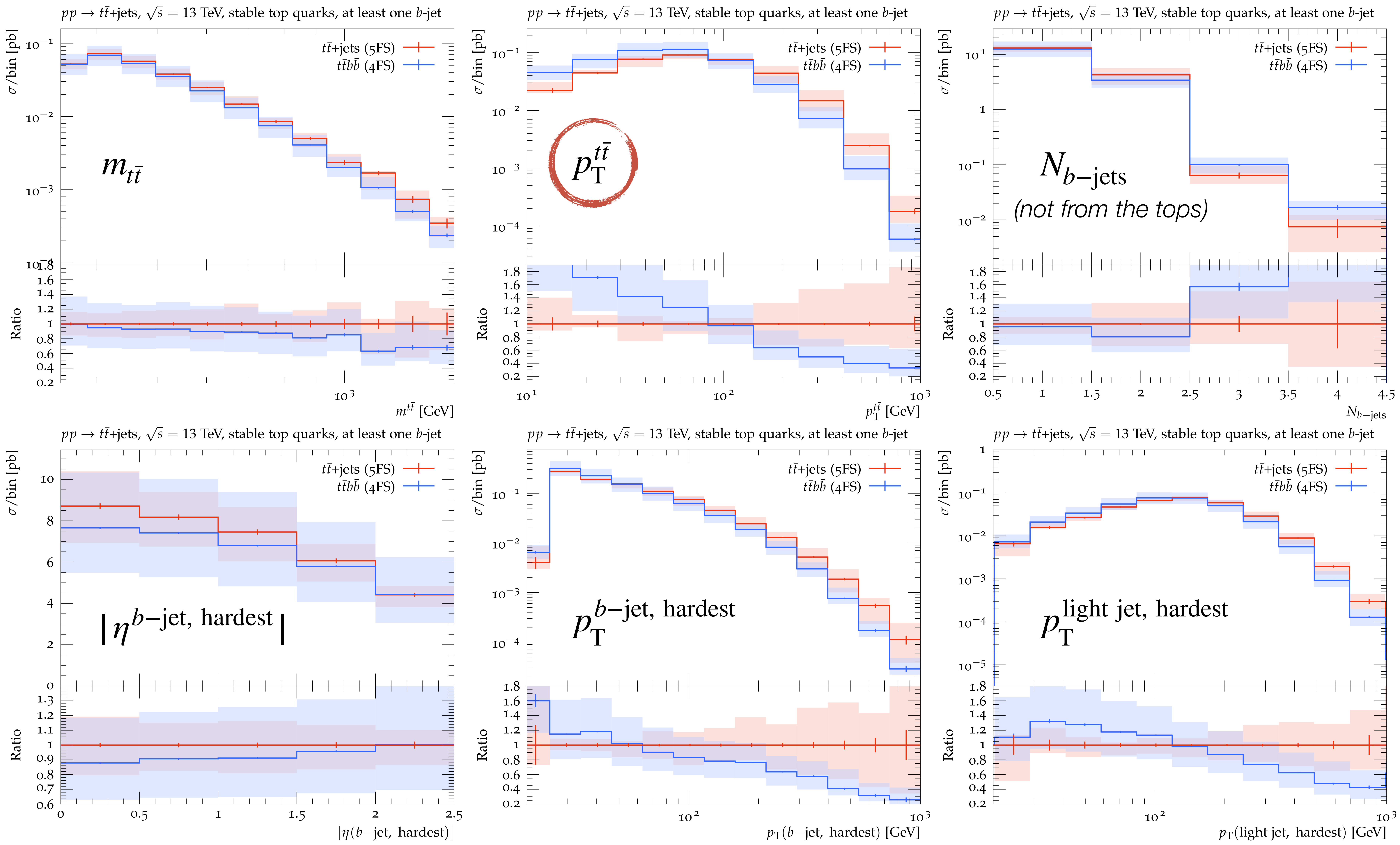
Truth-level analysis in Rivet:

- anti-k_T jets ($R > 0.4$)
  - $p_T > 25$ GeV
  - $|\eta| < 2.5$
- jets containing at least one bottom quark are identified as b-jets
- consider two scenarios:
  - at least 1 b-jet
  - at least 2 b-jets

5FS vs 4FS: at least 1 b-jet scenario

R. Frederix, TM
EPJC 84, 763 (2024)

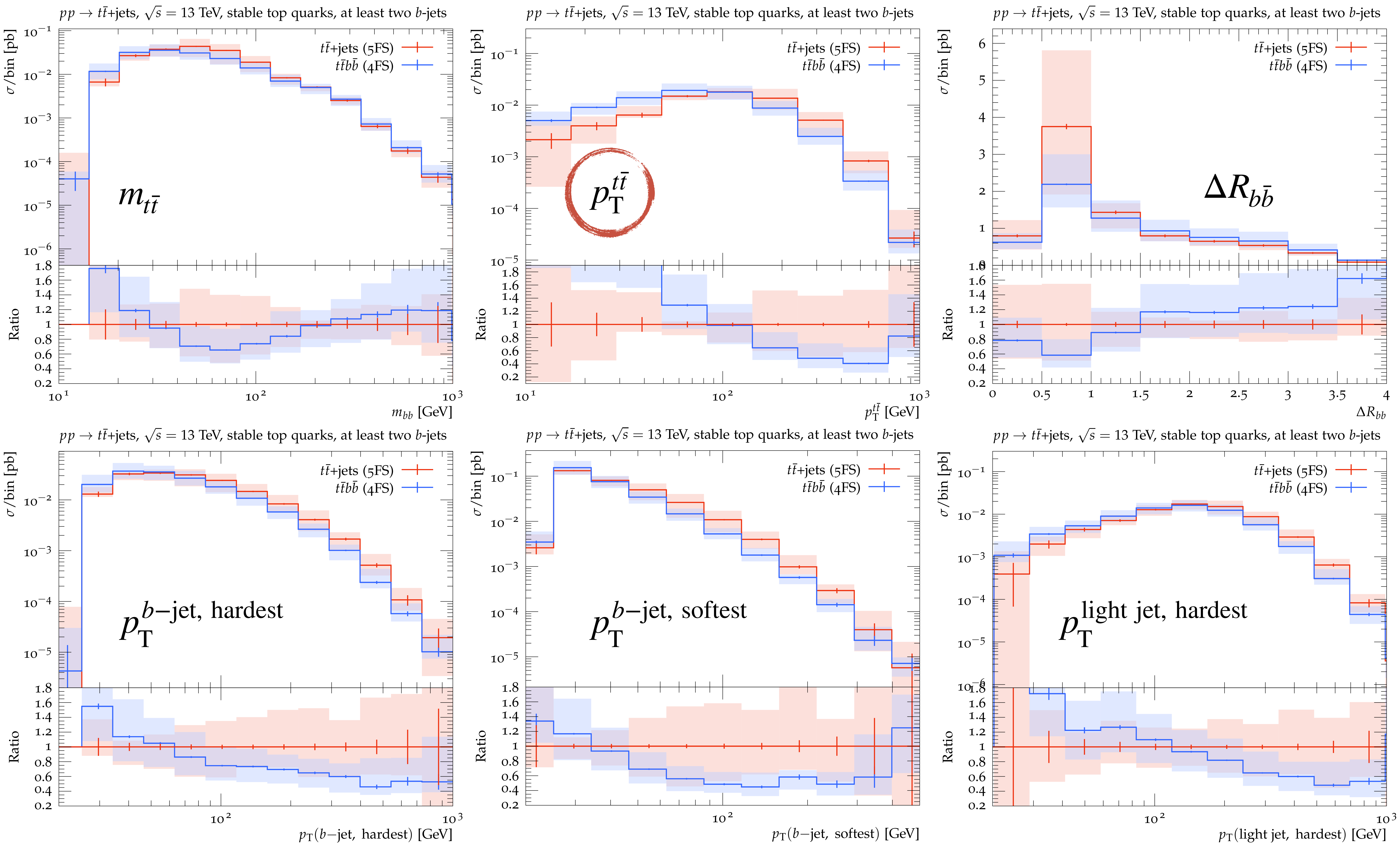
- For most of the variables, 4FS and 5FS predictions are compatible within the uncertainty bands
- 5FS uncertainty is more reliable than the 4FS one, since the 4FS matching uncertainty is expected to be significant but is not included
- $p_T^{t\bar{t}}$: **5FS predicts a much harder spectrum than 4FS** (but this difference is expected to be covered by the full 4FS uncertainty)
  - ➡ We investigated it further, see next slides



5FS vs 4FS: at least 2 b-jets scenario

R. Frederix, TM
EPJC 84, 763 (2024)

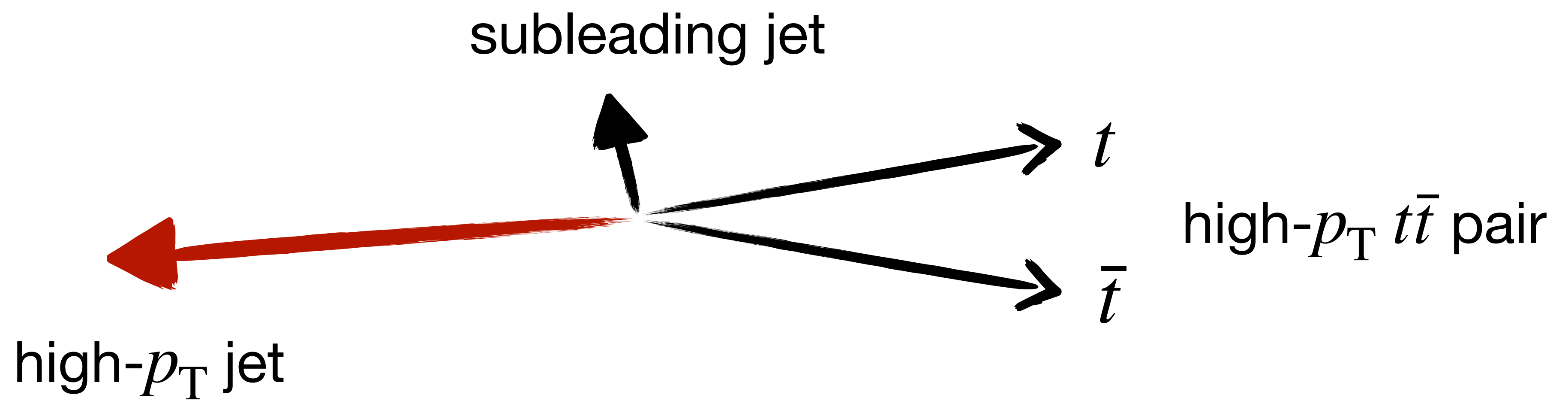
- Similar picture as for the ≥ 1 b-jet selection
- Difference in the $p_T^{t\bar{t}}$ spectrum (expected to be covered by the full 4FS uncertainty)
- The rest of the variables are in agreement



5FS vs 4FS: differences in the $p_T^{t\bar{t}}$ distribution

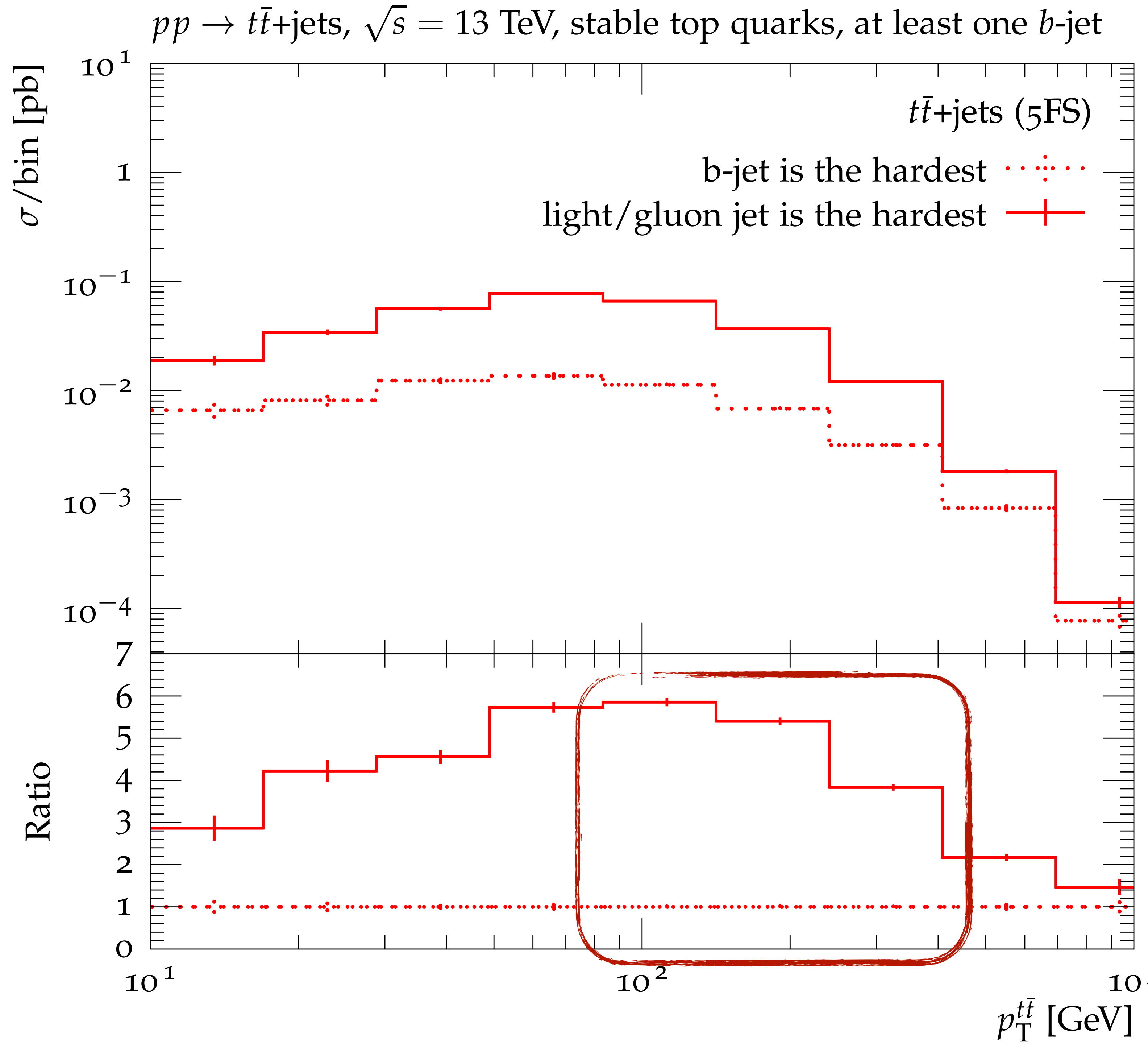
R. Frederix, TM
EPJC 84, 763 (2024)

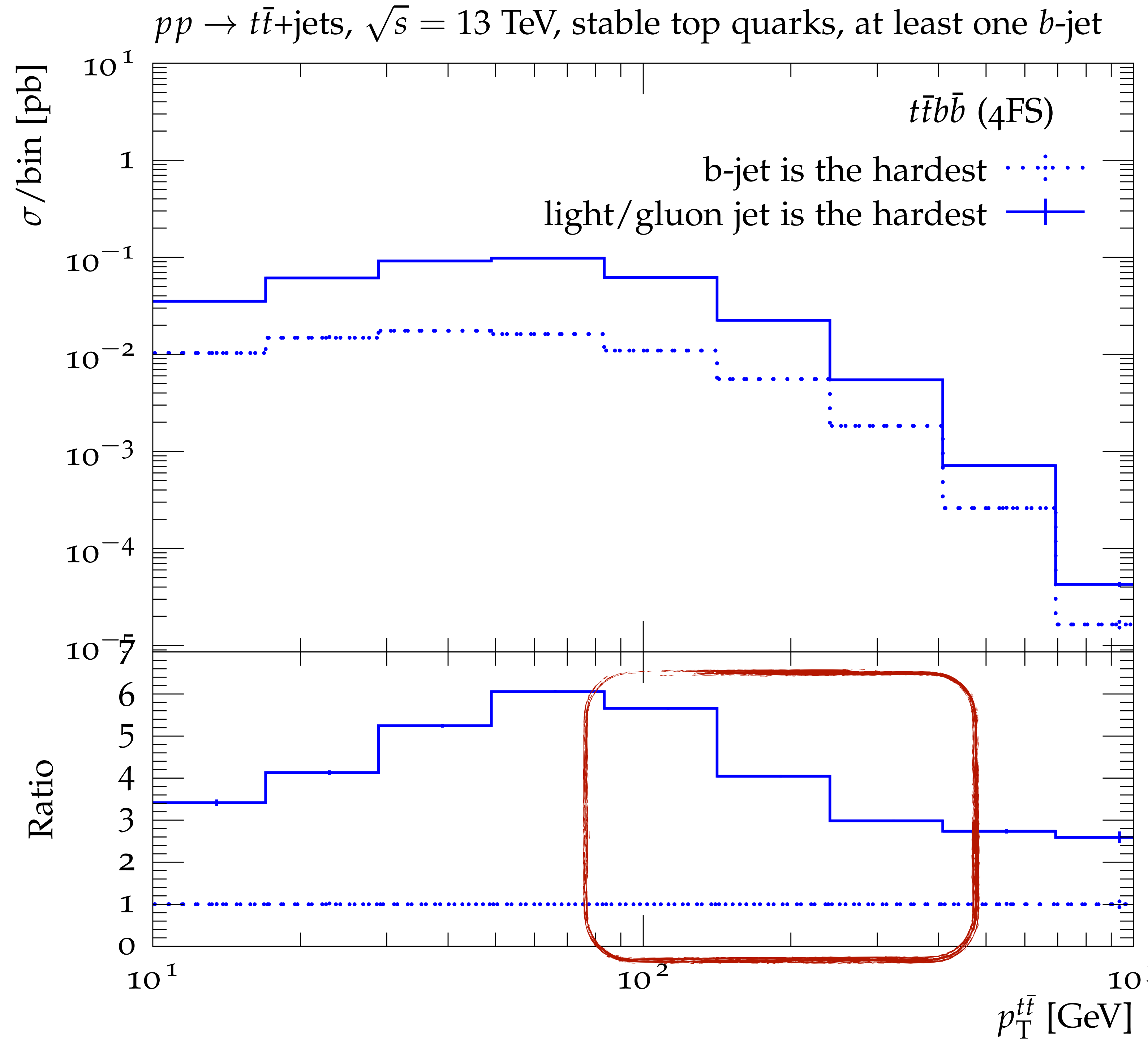
- At large $p_T^{t\bar{t}}$, it is kinematically most-likely that the $t\bar{t}$ pair recoils agains a single hard jet
- If the hardest jet is a light jet:
  - 5FS: described at NLO (most likely it is a gluon jet)
  - 4FS: described at LO or by the PS
    - No $t\bar{t}gg$ events from the ME
    - There is no hard gluon to recoil from



- For high $p_T^{t\bar{t}}$, the fraction of events with the hardest jet being light-flavoured is indeed larger in the 5FS
- But after $p_T^{t\bar{t}} \sim 500$ GeV the situation is opposite — why?
  - Let's look again at the jet p_T distributions…

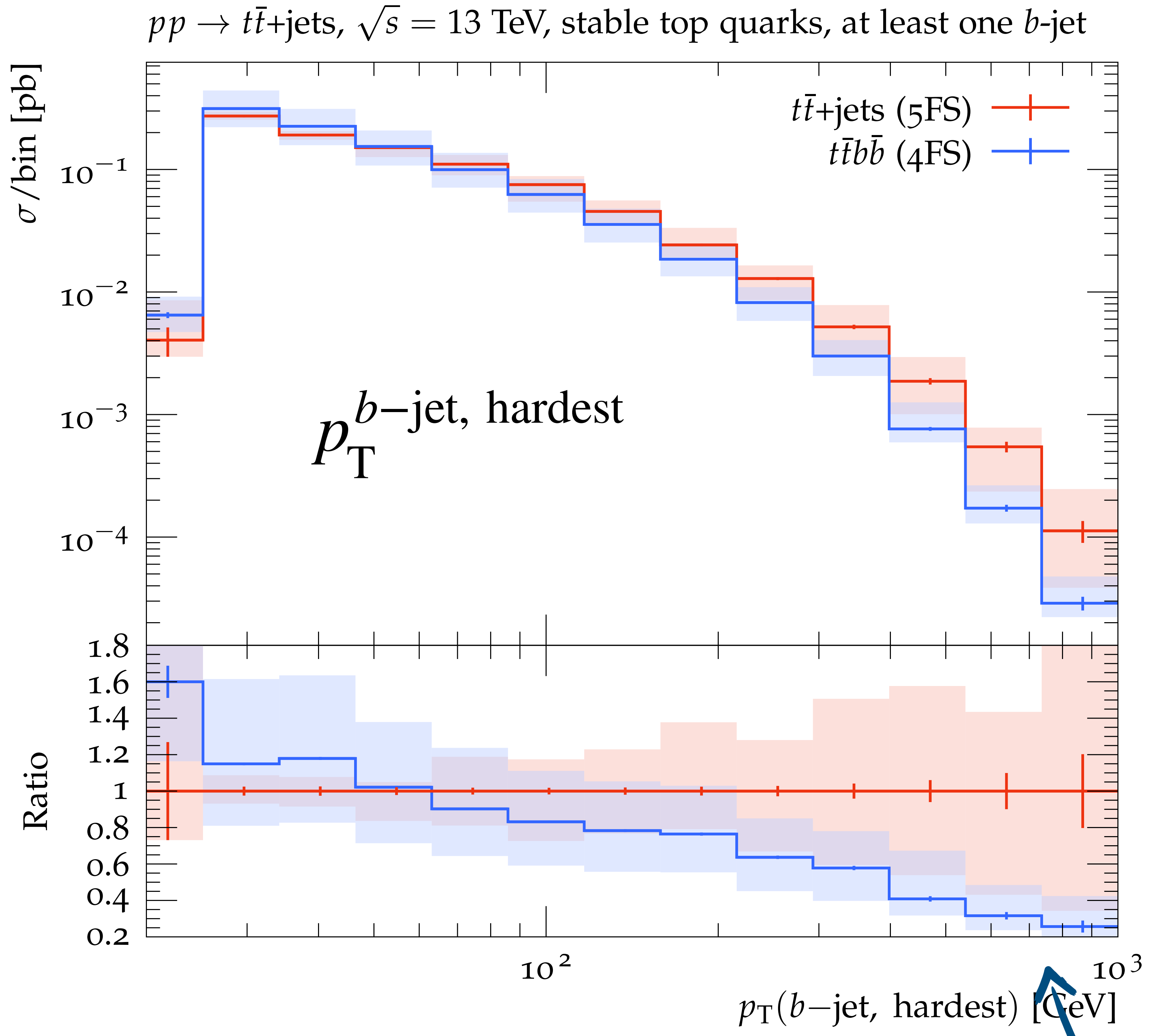
at least 1 b-jet selection

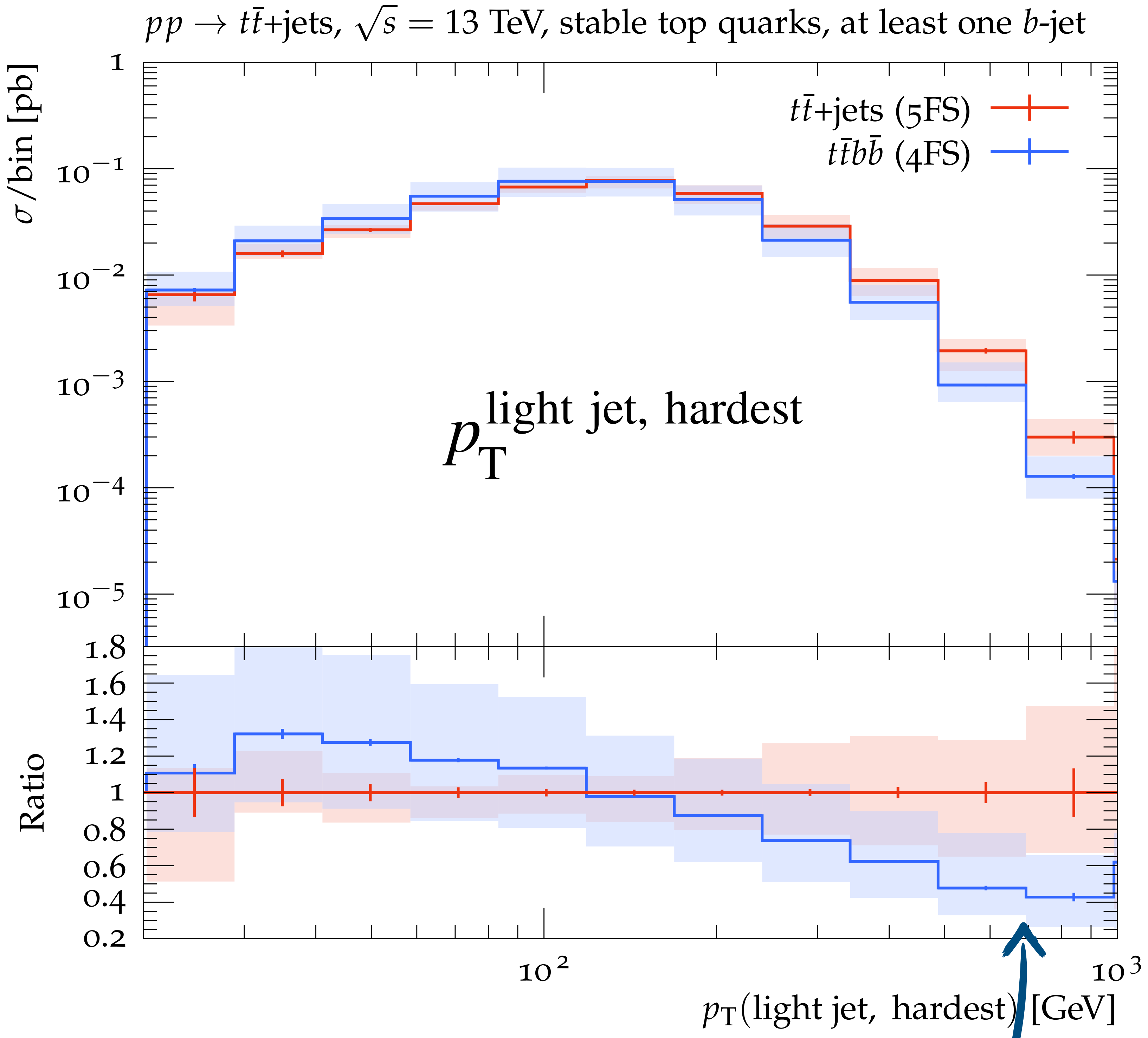




5FS vs 4FS: differences in the $p_T^{t\bar{t}}$ distribution

R. Frederix, TM
EPJC 84, 763 (2024)





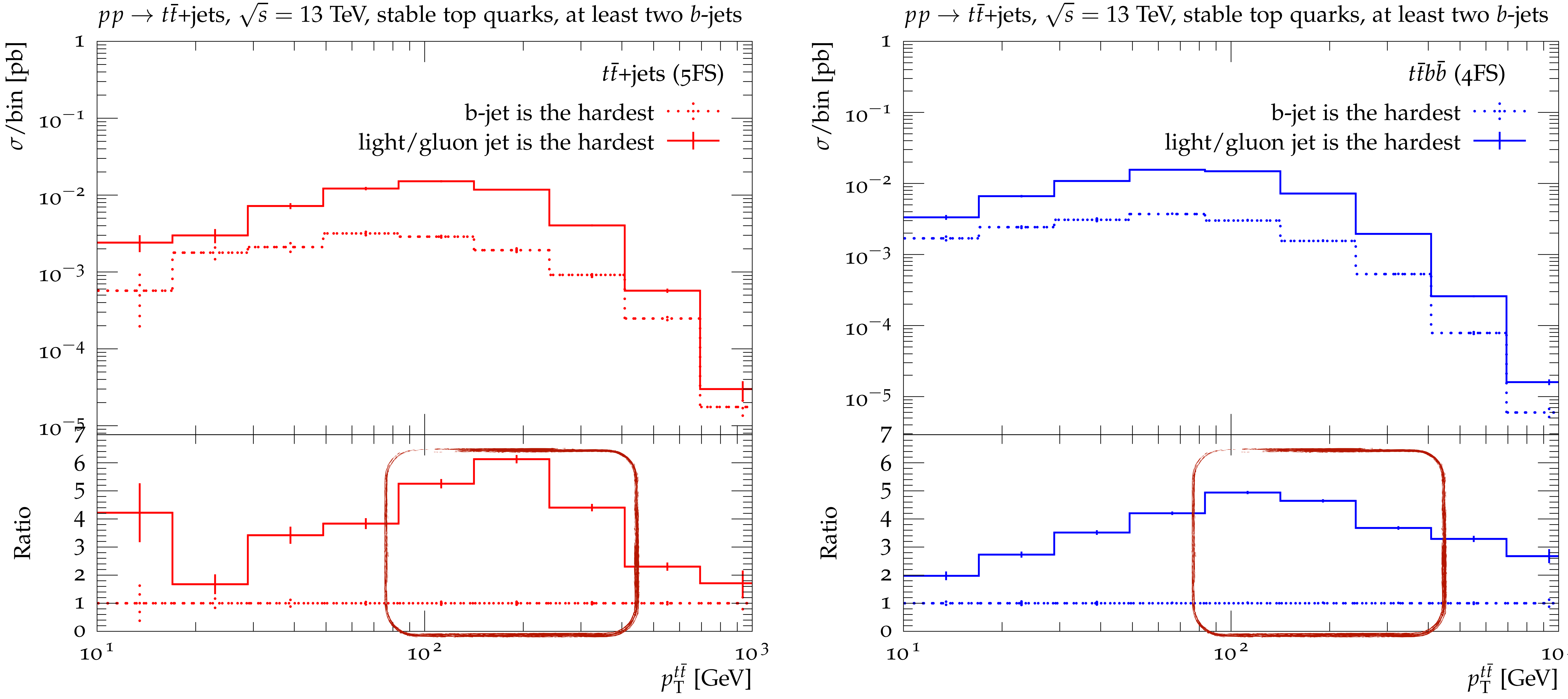
at very high p_T, the ratio of the 5FS over the 4FS predictions is larger for b-jets than for light jets

5FS vs 4FS: differences in the $p_T^{t\bar{t}}$ distribution

R. Frederix, TM
EPJC 84, 763 (2024)

- The difference in the fraction of the hardest light jets in even more pronouced in the ≥ 2 b-jet selection

at least 2 b-jets selection



➡ The reason for the large 5FS–4FS difference in the $p_T^{t\bar{t}}$ spectrum at large momenta is

- The correlation between $p_T^{t\bar{t}}$ and $p_T^{\text{light jet, hardest}}$
- Expected 5FS–4FS difference between the fraction of events with the hardest jet being light-flavoured

What about a comparison to data?

- Not done yet
- A study is planned within the ATLAS Physics Modelling Group
- Once the new MadGraph5_aMC@NLO release is out, the idea is to produce an "official" sample (5FS, FxFx, with enhancement) and check:
  - data/MC comparison using the ATLAS $t\bar{t}b\bar{b}$ Rivet routine
  - performance of the enhancement feature

To summarise:

- $t\bar{t}b\bar{b}$ production serves as a significant background process across various high-energy physics phenomena
- 5FS calculation of $t\bar{t}b\bar{b}$ at NLO yields the most accurate prediction for this process to date
  - no large logarithms appearing in the matrix element calculation
  - no complications when matching to a parton shower
- We compute the $t\bar{t}$ + jets process with up to 2 jets at NLO using the FxFx merging prescription and match it to the Pythia8 shower
- To improve the efficiency of selecting events with additional b-jets we enhance the probability of producing short-distance events with additional b-quarks using a newly implemented feature in the MadGraph5_aMC@NLO generator
  - This makes producing the $t\bar{t}b\bar{b}$ in the 5FS at NLO more viable, given the computational demands of the 5FS approach

✳ *Similar heavy-flavour enhancement could also be applied to the "fusion" method in Sherpa*