Evidence for four-top-quarks production with the ATLAS detector at the Large Hadron Collider

Paolo Sabatini [a], on behalf of the ATLAS Collaboration

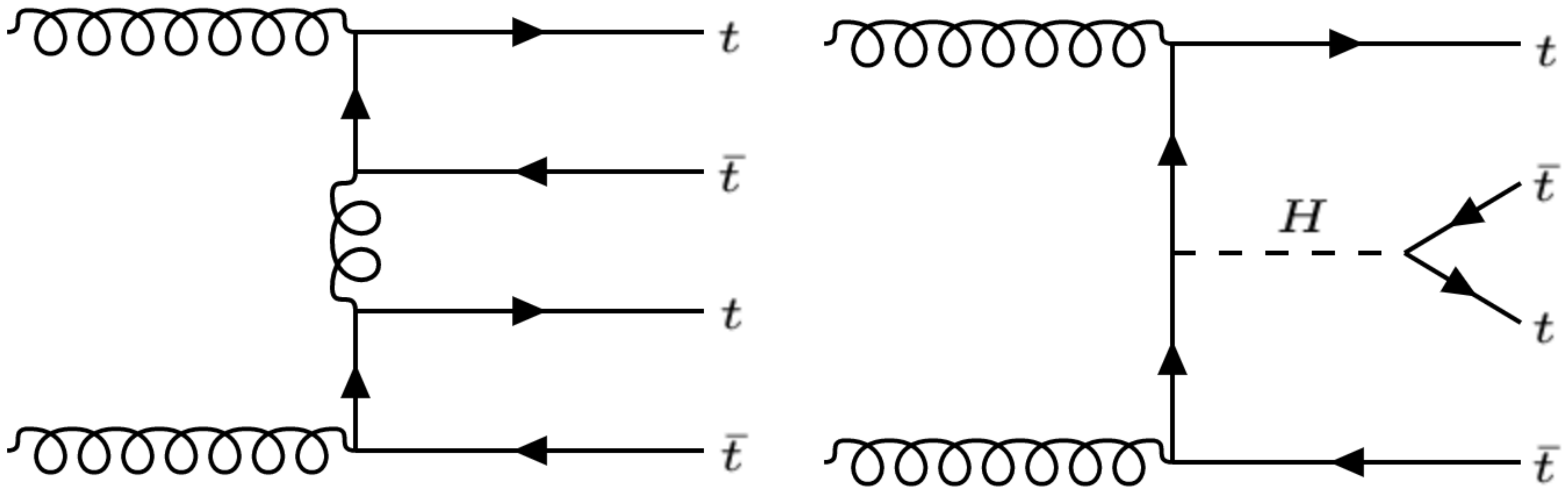
[a] Instituto de Física Corpuscolar, Valéncia (CSIC/UV)

EPS-HEP 2021 **T07: Top and Electroweak Physics** **30/07/2021**

Four top-quark production

The production of four-top quark is one of the rarest events involving top-quarks in the final state.

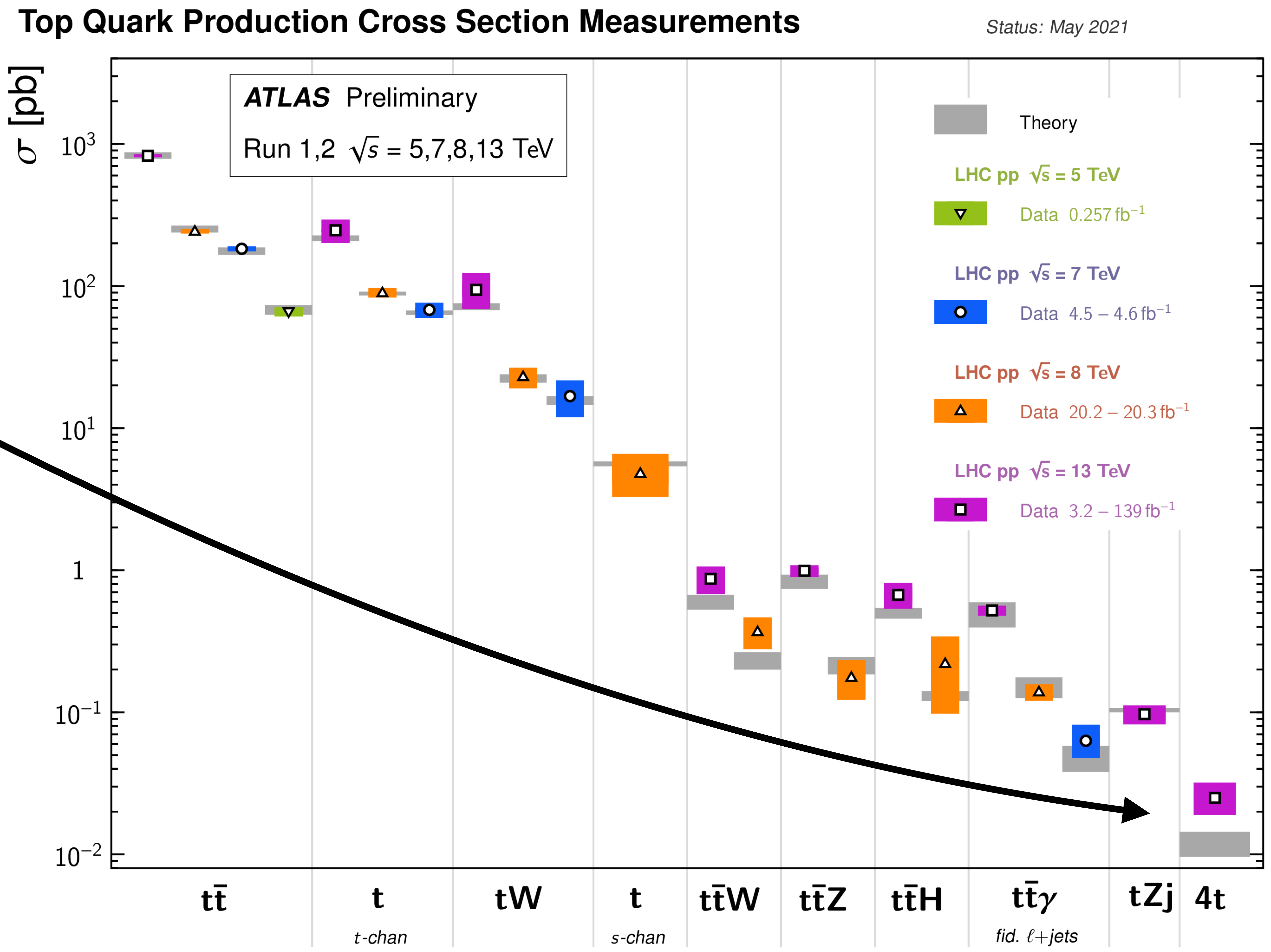
Very interesting test of SM validity which describes the top-quark sector over **5** orders of magnitude!



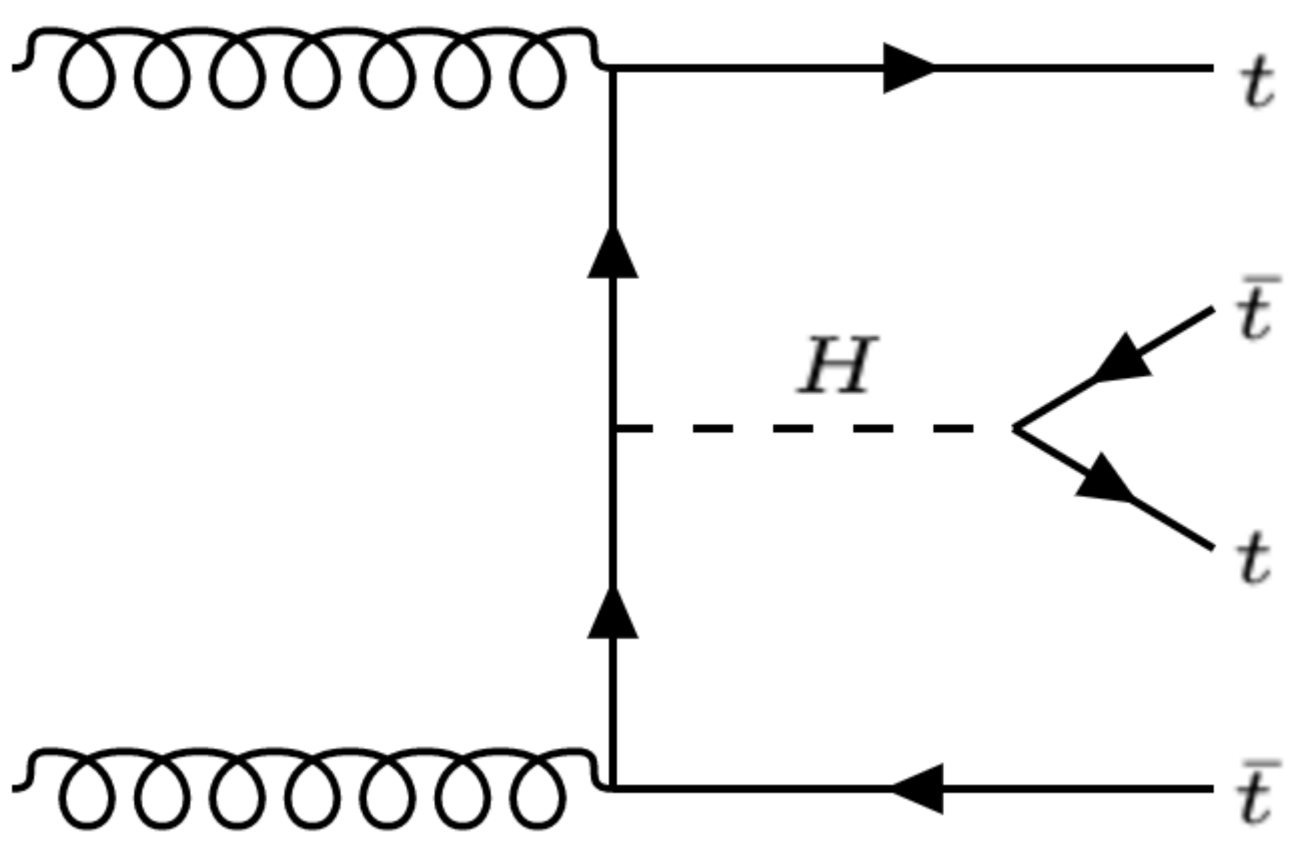
Cross-section predicted by SM @ NLO (EW+QCD)

JHEP 02 (2018) 031

$$\sigma_{t\bar{t}t\bar{t}} = 12 \pm 20\,\%\ \text{fb}$$



Interesting for many reasons



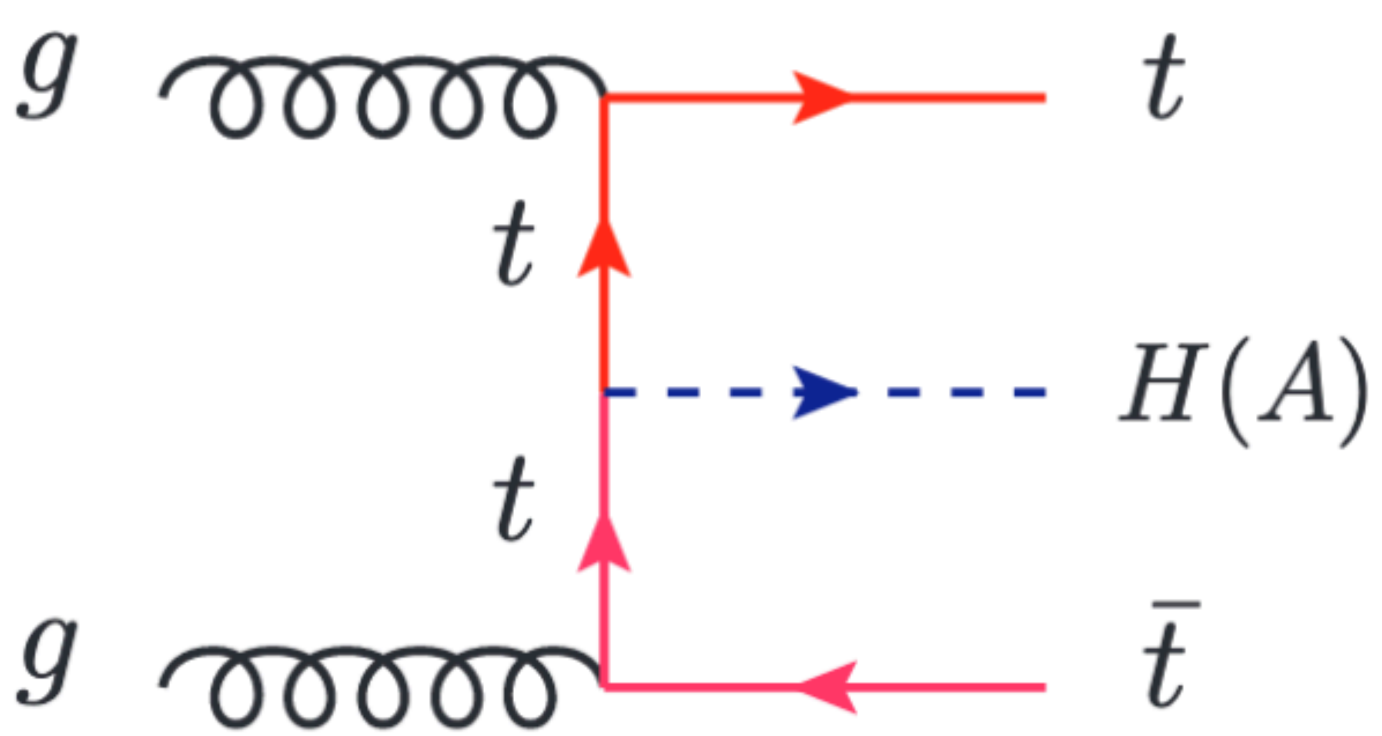
Four top-quark production is sensitive to top-Yukawa coupling.

It provides an orthogonal investigation with respect to $t\bar{t}H$ measurement

Phys. Rev. D 95 053004   Phys. Rev. D 99 113003

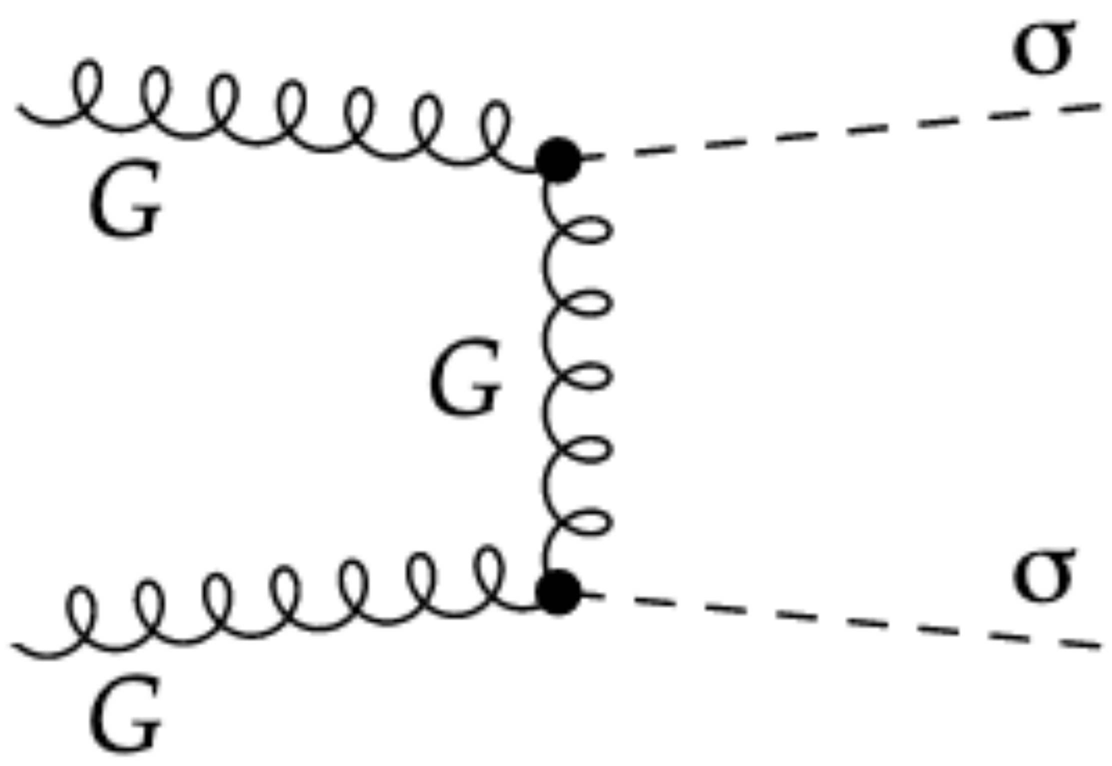
Many Beyond Standard Model (BSM) scenarios provide large enhancement of the cross-section

2HDM



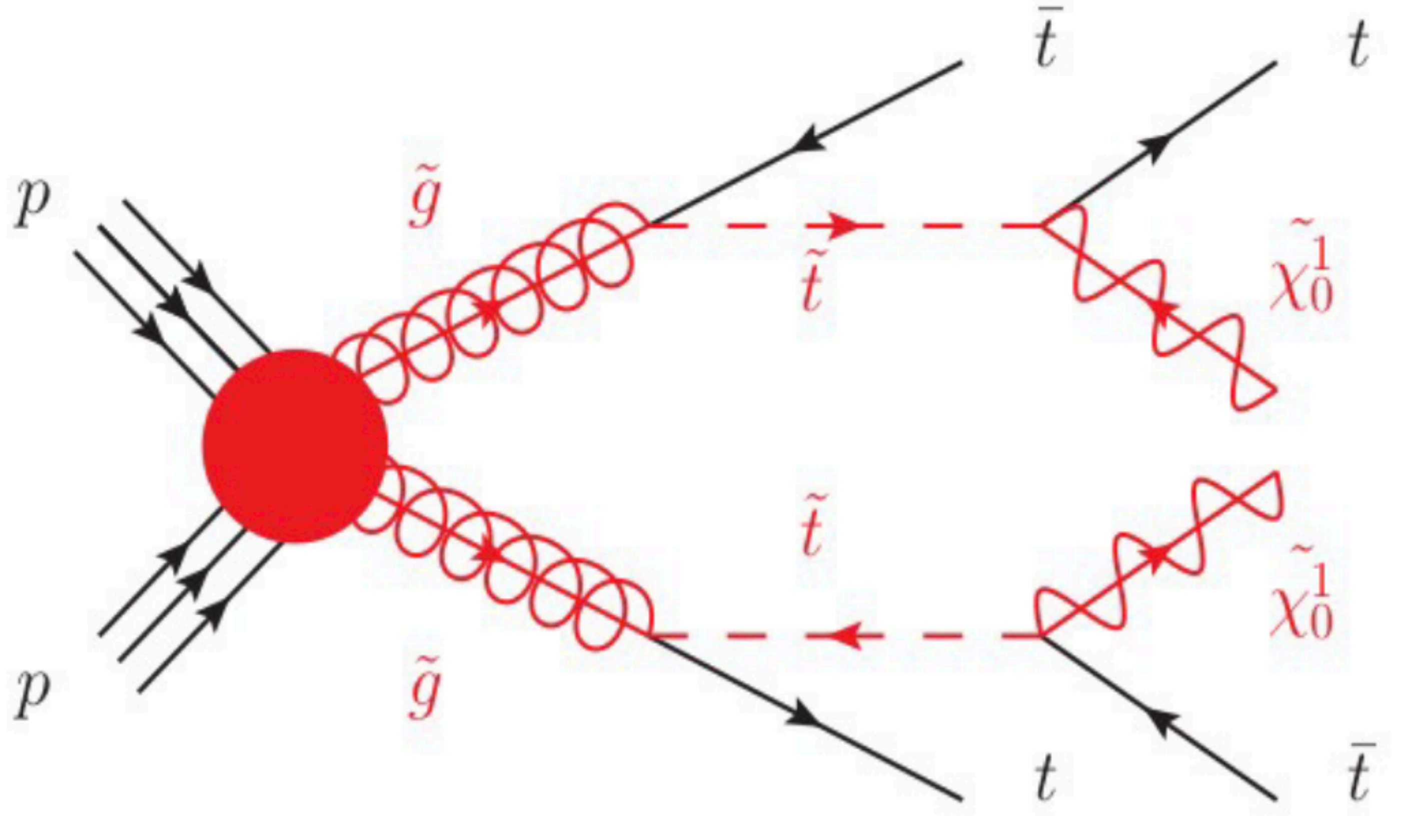
JHEP01 (2017) 018
JHEP 06 (2015) 137

Scalar gluons



J. Phys. G: Nucl. Part. Phys. 36 075001
JHEP 04 (2013) 043

SUSY

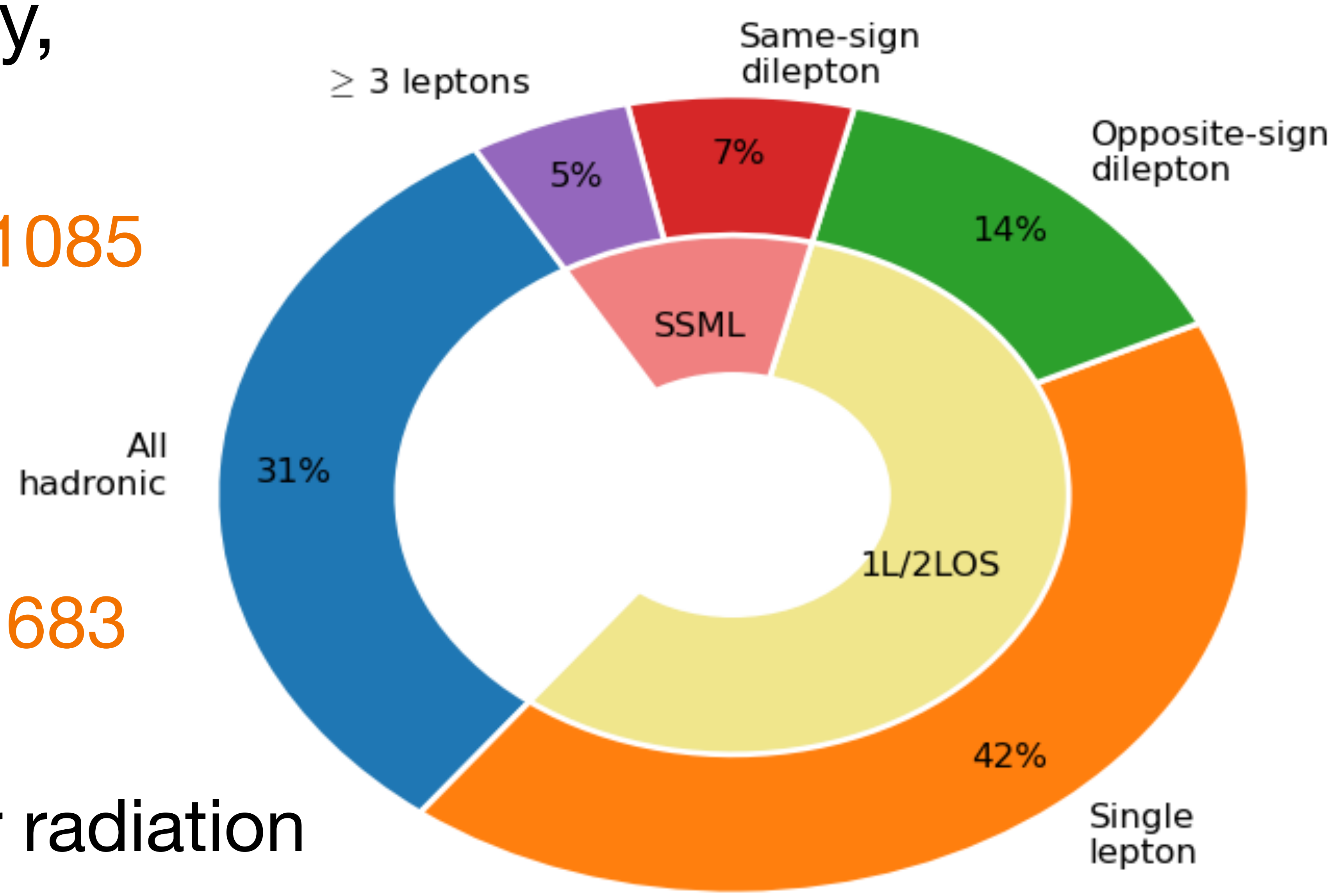


Phys. Report 110 (1984) 1-2
Phys. Lett. B 76 (1978) 5

In the context of EFT, this production is parametrised by a $t\bar{t}t\bar{t}$ contact-interaction term   1802.07237

Detection channels and process signature

Each top-quark in the final state can decay leptonically or hadronically, leading to many possible detection channels.



Same-sign dilepton + **multi-lepton channel** (SSML) EPJC 80 (2020) 1085

Corresponding to ~12% of total events.

Facing with modelling of $t\bar{t}+V$ at large b-jet multiplicity

Dealing with many sources of instrumental backgrounds

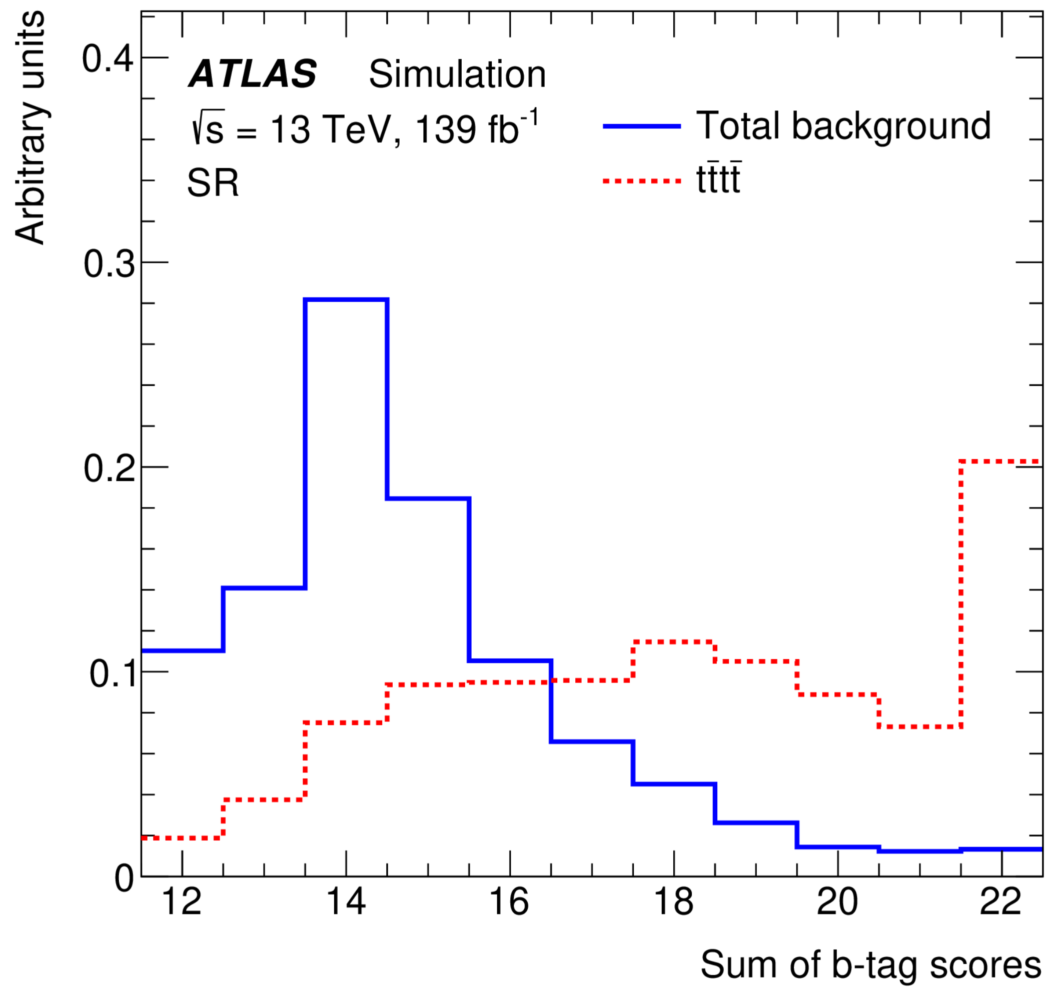
Single lepton + **opposite-sign dilepton channel** (1L2LOS) 2106.11683

Corresponding to ~60% of total events.

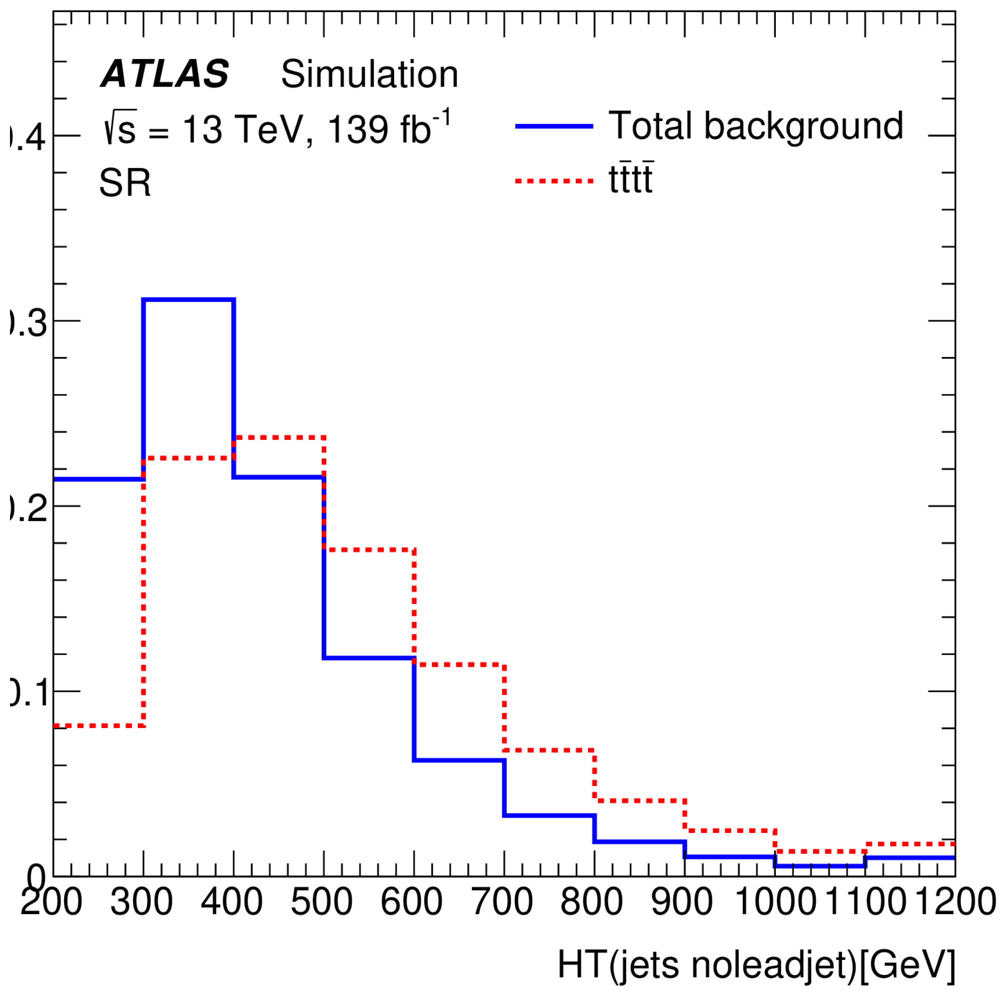
Background dominated by $t\bar{t}$ production associated with large jet heavy-flavour radiation

How can we separate $t\bar{t}t\bar{t}$ from the total background?

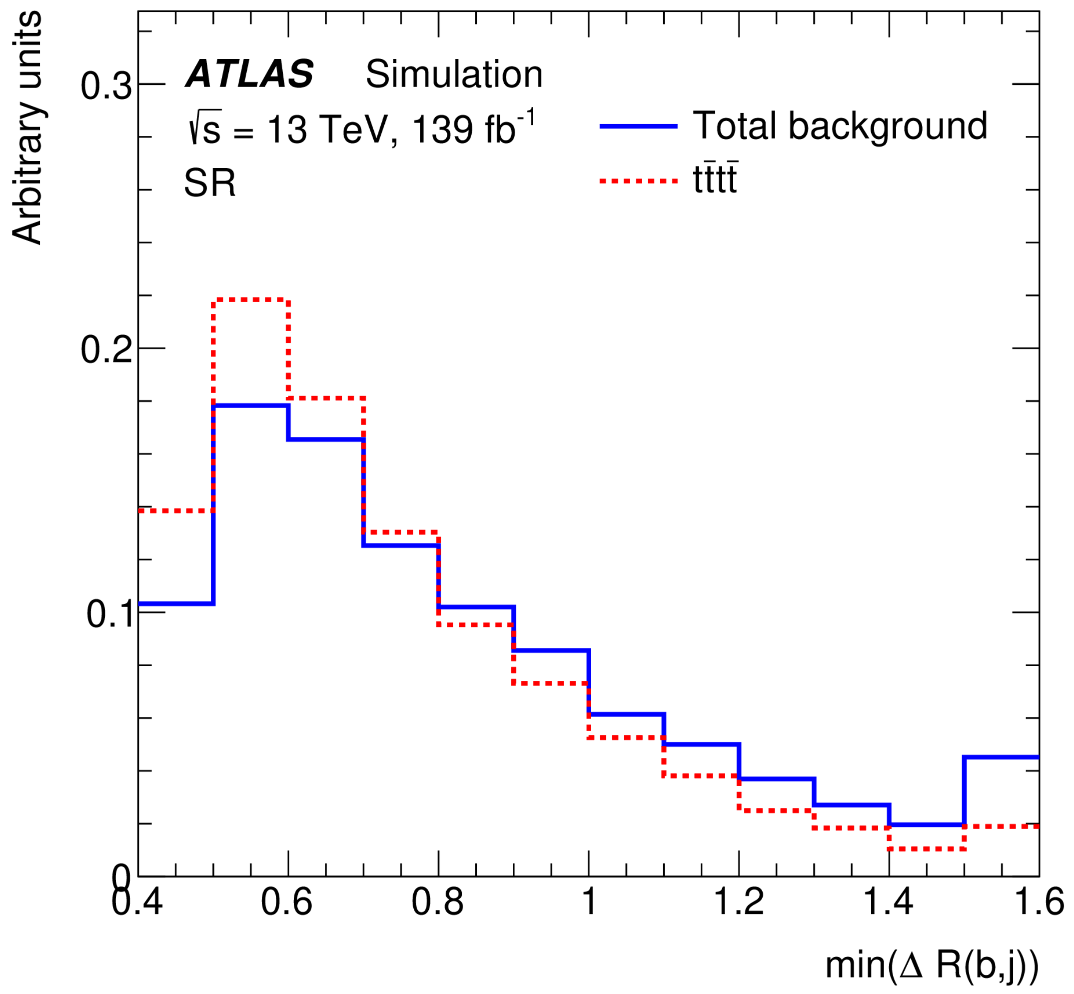
More & harder b-jets



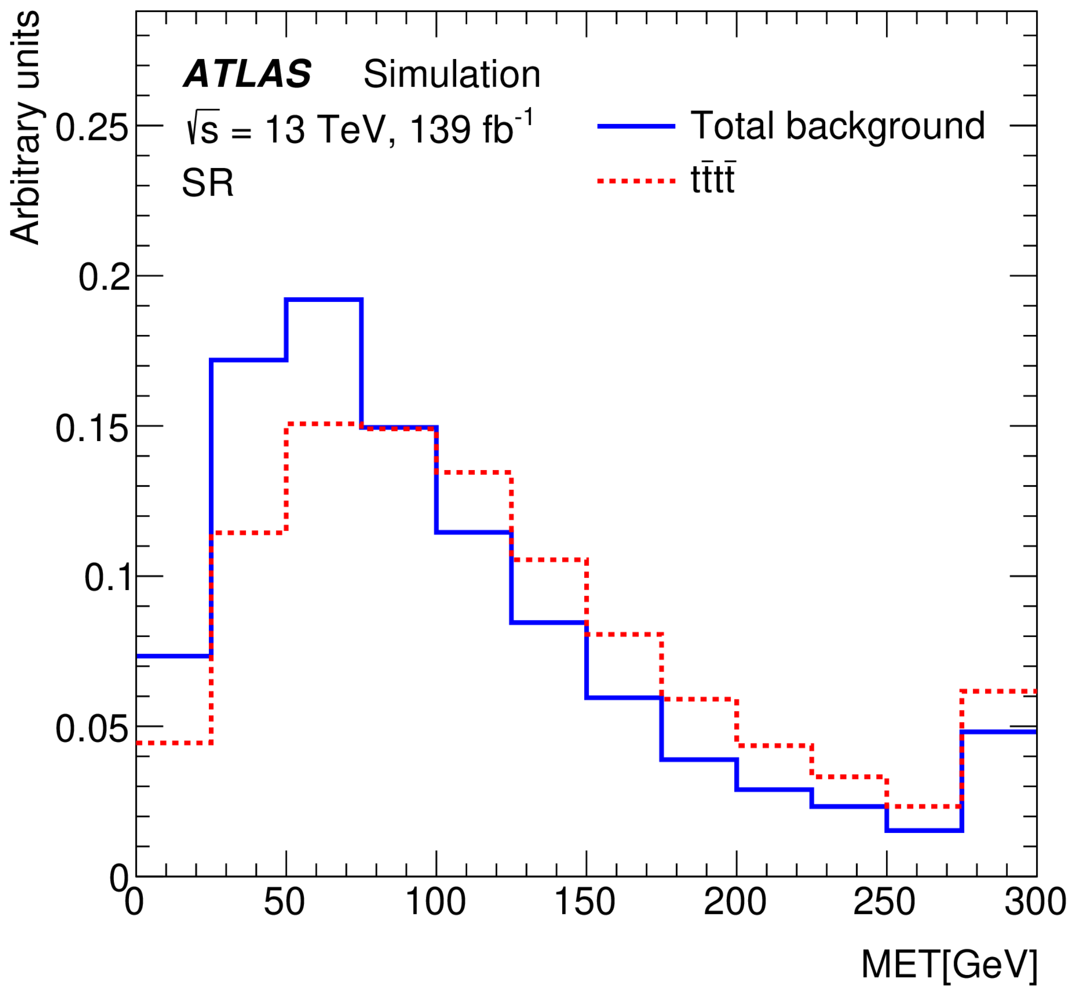
More & harder jets



Top decay chains



Large E_T^{miss}



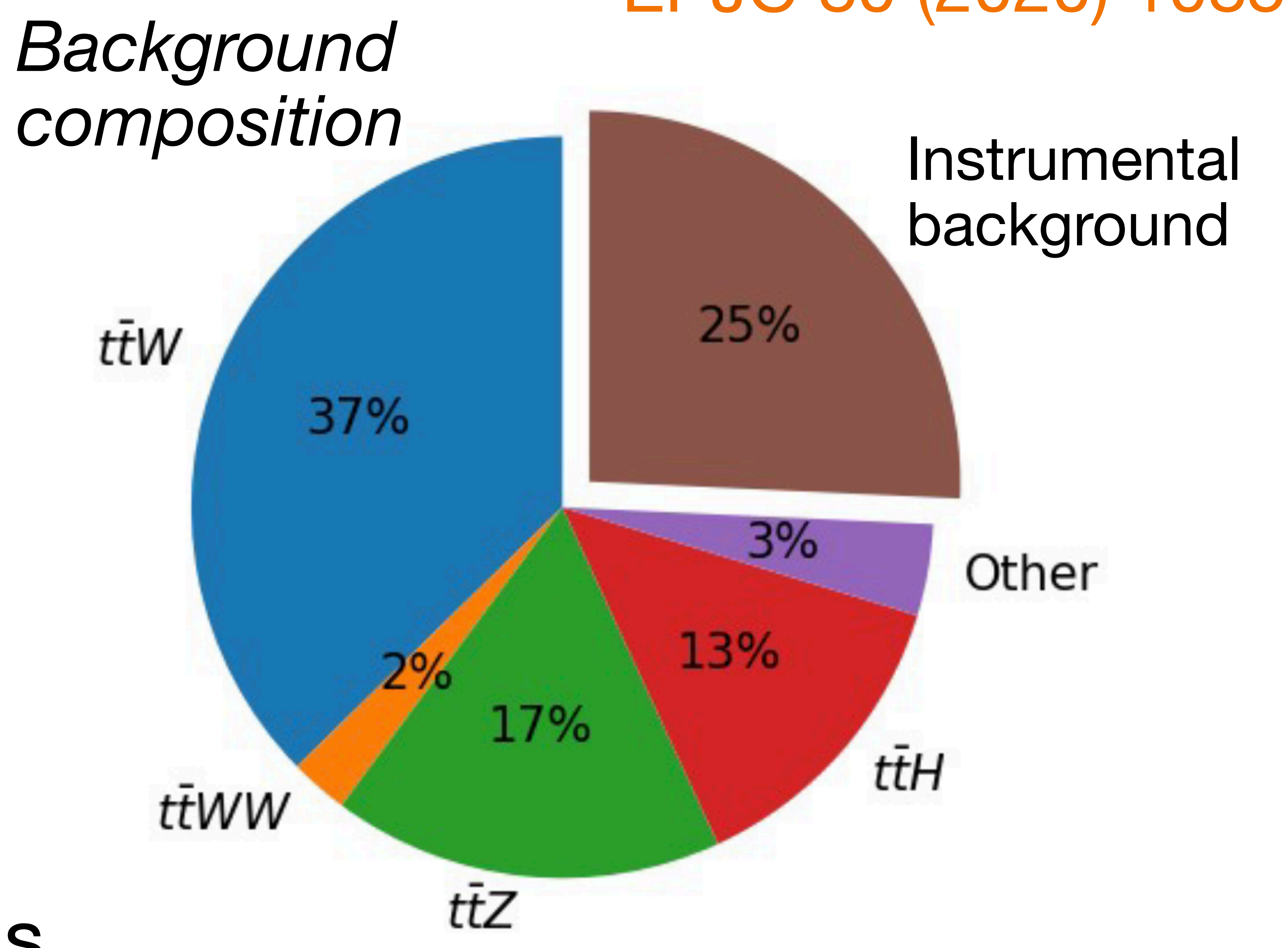
[SSML] Analysis strategy

EPJC 80 (2020) 1085

Analysis exploiting the full ATLAS Run 2 dataset (139 fb^{-1})

Selections:

- 2 same-sign leptons or ≥3 leptons (e/μ)
- ≥6 jets (of which ≥2 b-tagged)
- $H_T = \sum p_T > 500$ GeV



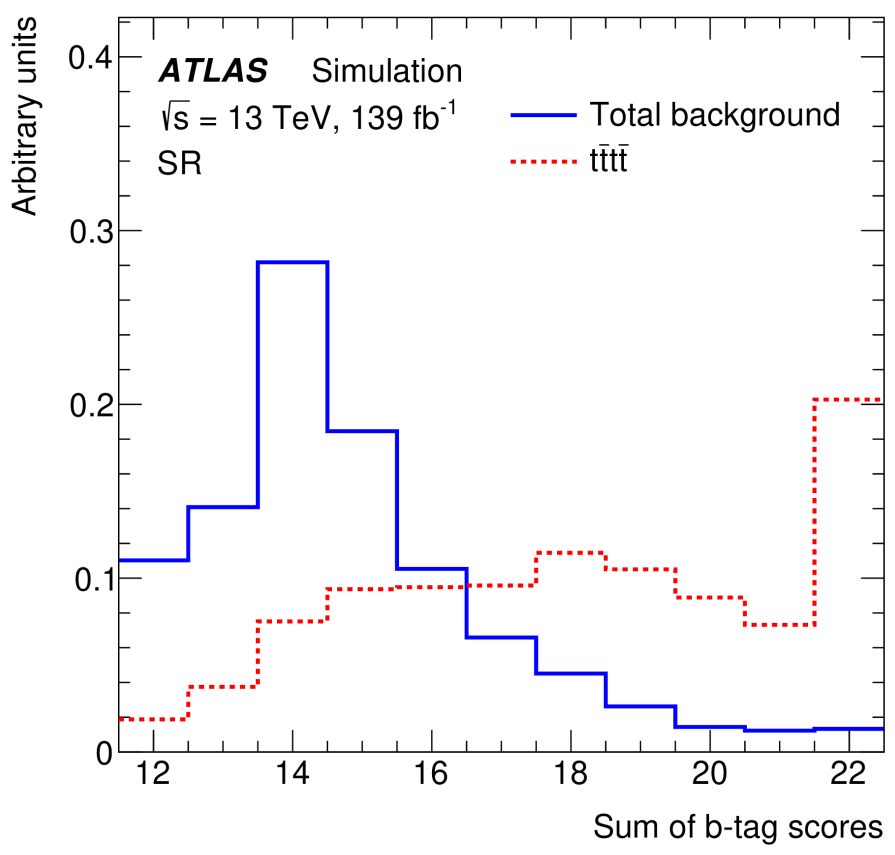
Irreducible backgrounds: top pairs production associated with bosons

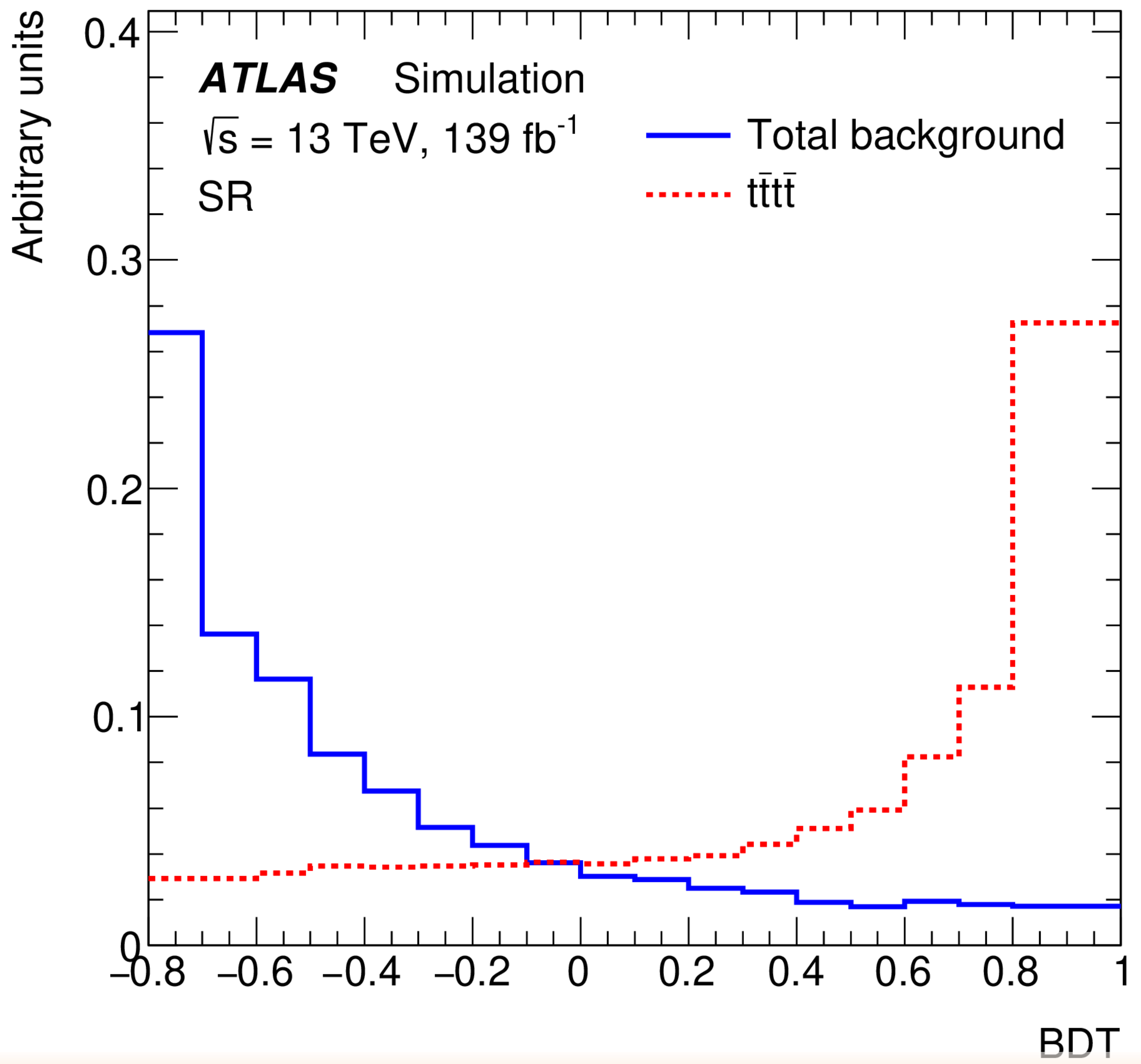
Significant contribution of fake/non-prompt leptons

Two main goals: signal separation and robust background estimation (next slide)

A Boosted-Decision-Tree classifier is trained to optimise signal-vs-background separation

- BDT score used as discriminating variable in SR for profile likelihood fit
- BDT score outperforming with respect to any other variable
- B-tagging information most important feature





[SSML] Instrumental background estimation

EPJC 80 (2020) 1085

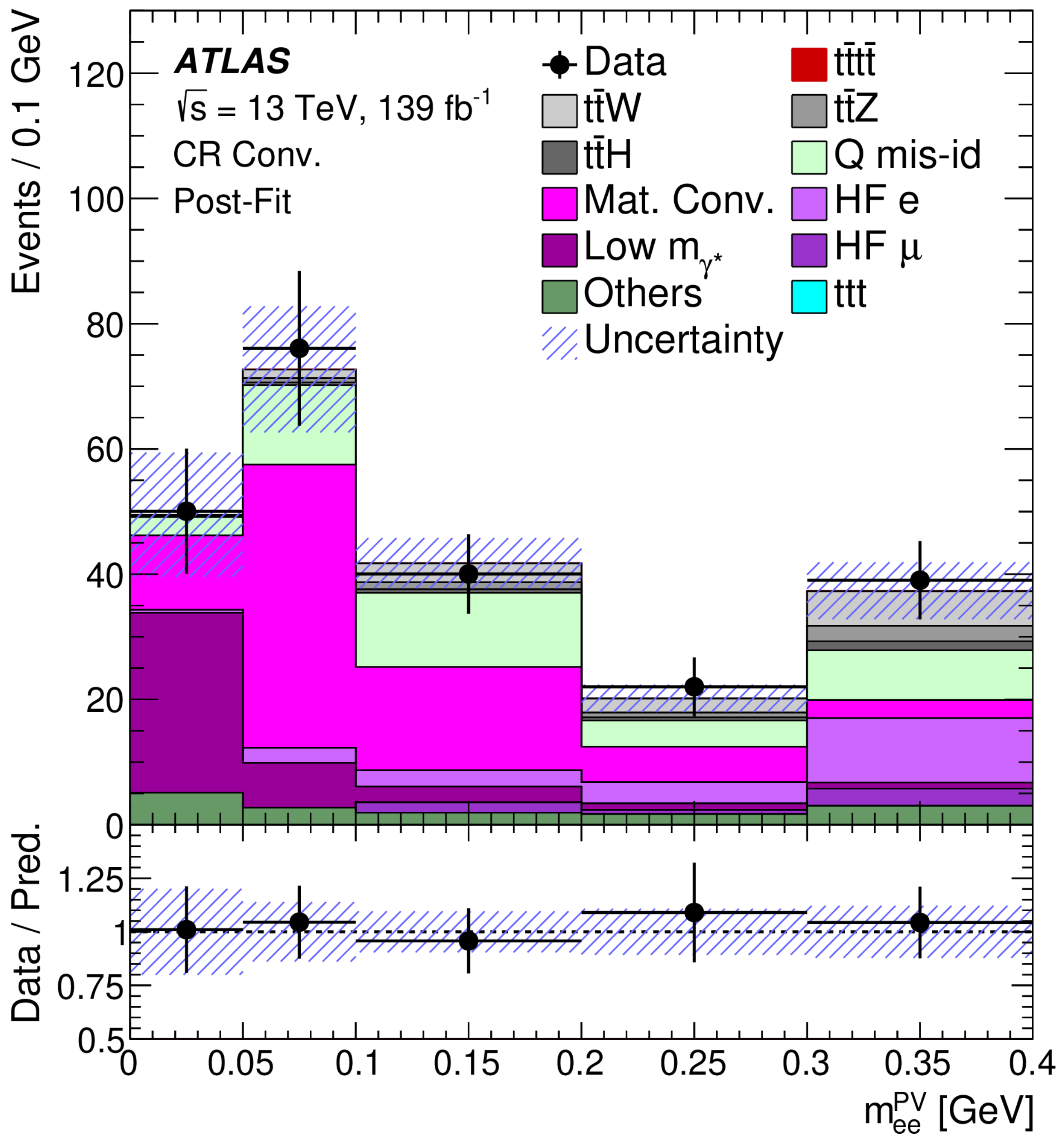
Charge mis-identification estimated with $Z(ee)$ events by measuring the charge mis-id. efficiency

Non-prompt leptons background & conversions estimated through **template method**

Dedicated control region at low E_T^{miss} to control these source of backgrounds

Material & γ^* conversions

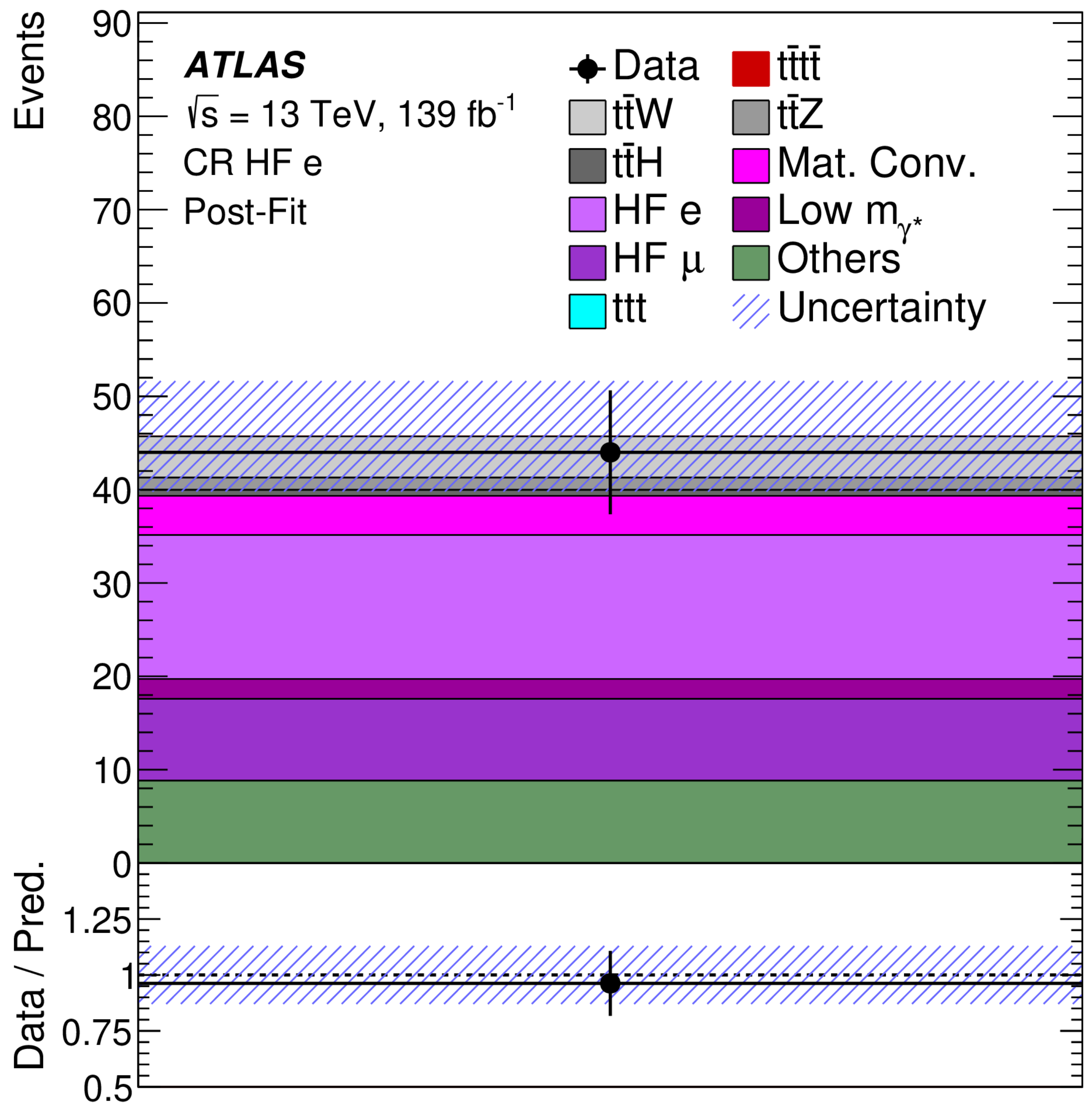
Distribution of e^+e^- mass at Primary Vertex fitted



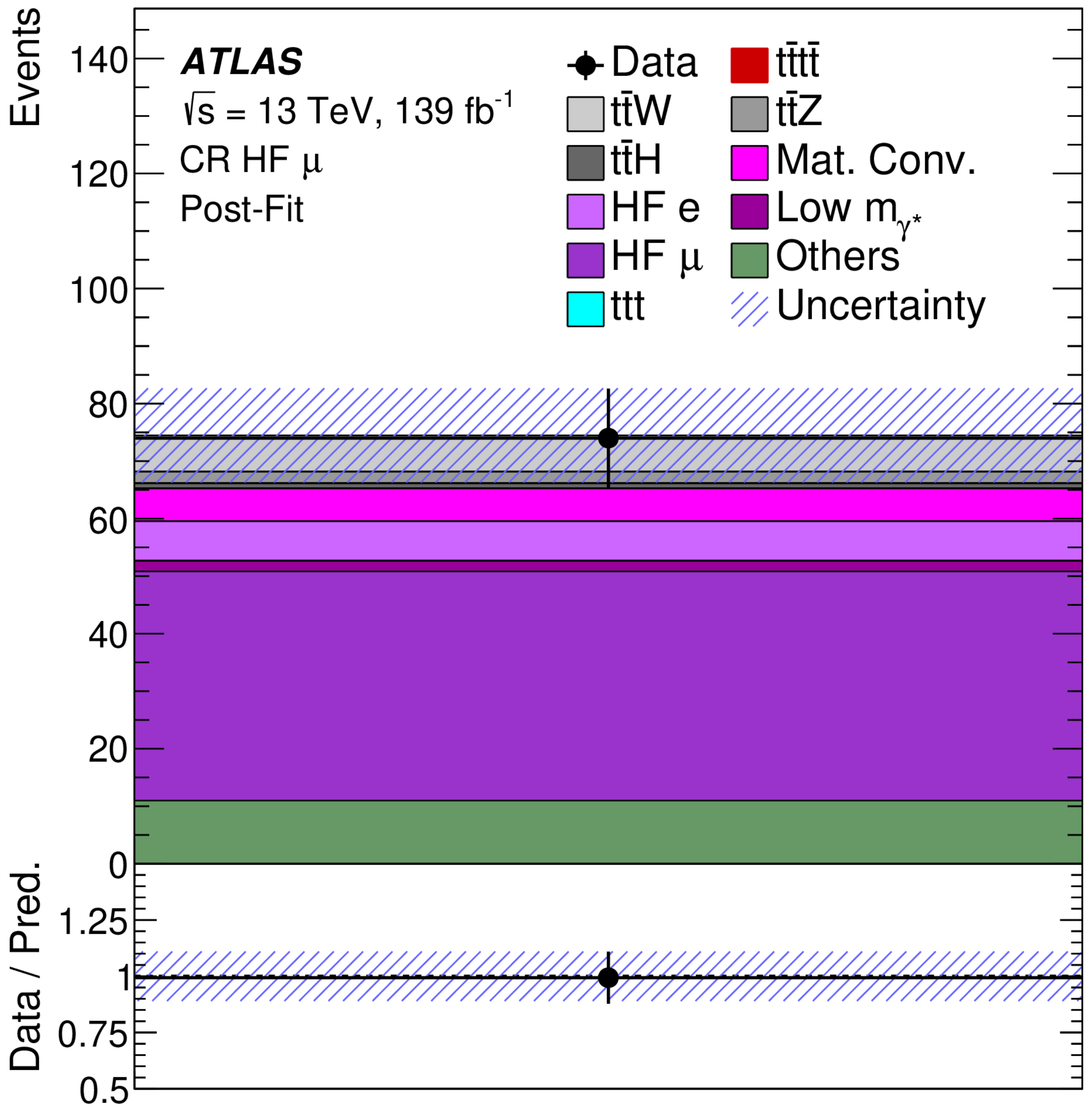
Leptons from heavy-flavours (HF) decays

One region per lepton flavour

Electrons from HF



Muons from HF



[SSML] Irreducible background estimation

EPJC 80 (2020) 1085

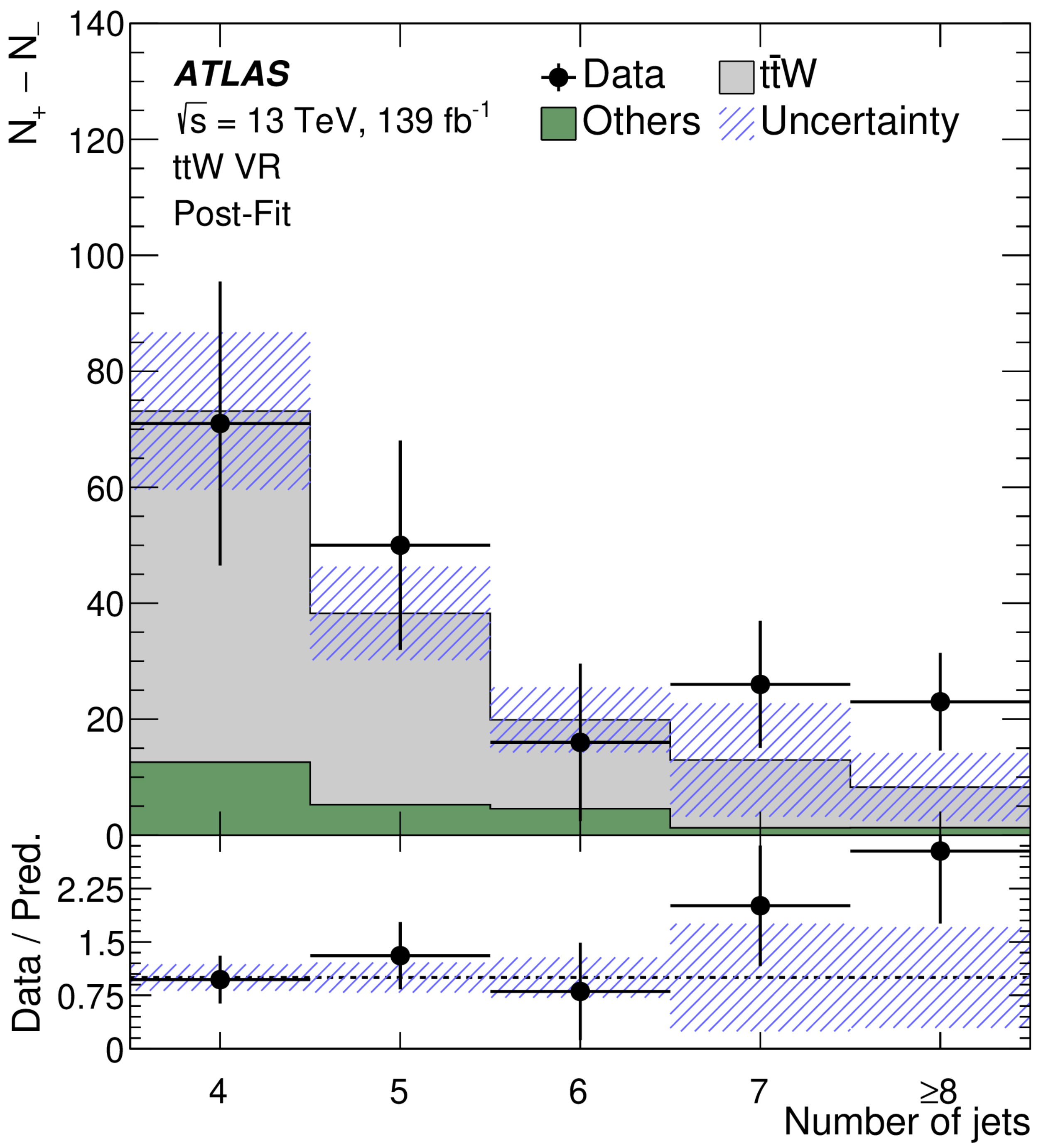
Dominant contribution among the irreducible background from $t\bar{t}W$ production: **template method**.

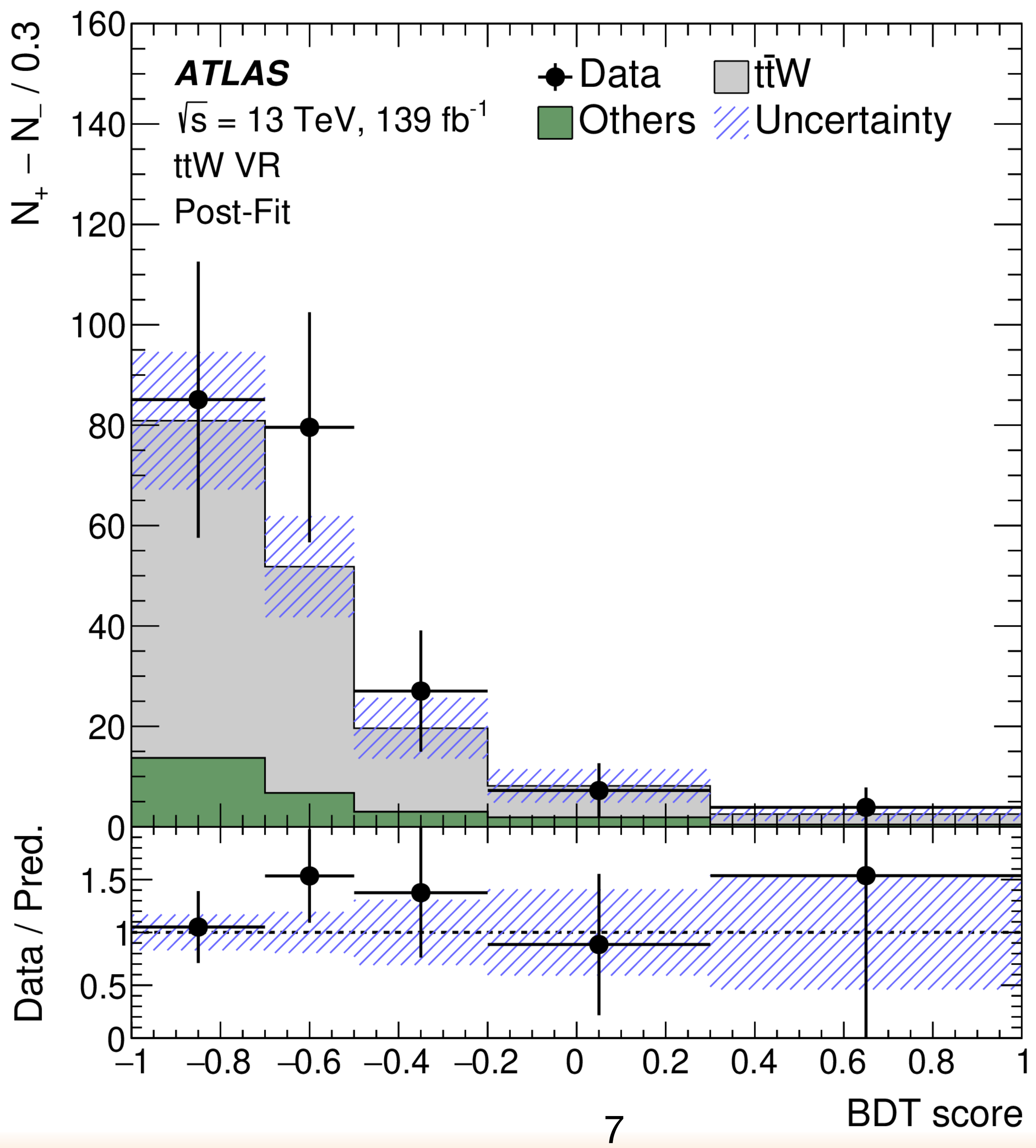
Dedicated control region (CR ttW) to measure $t\bar{t}W$ normalisation and constrain the modelling.

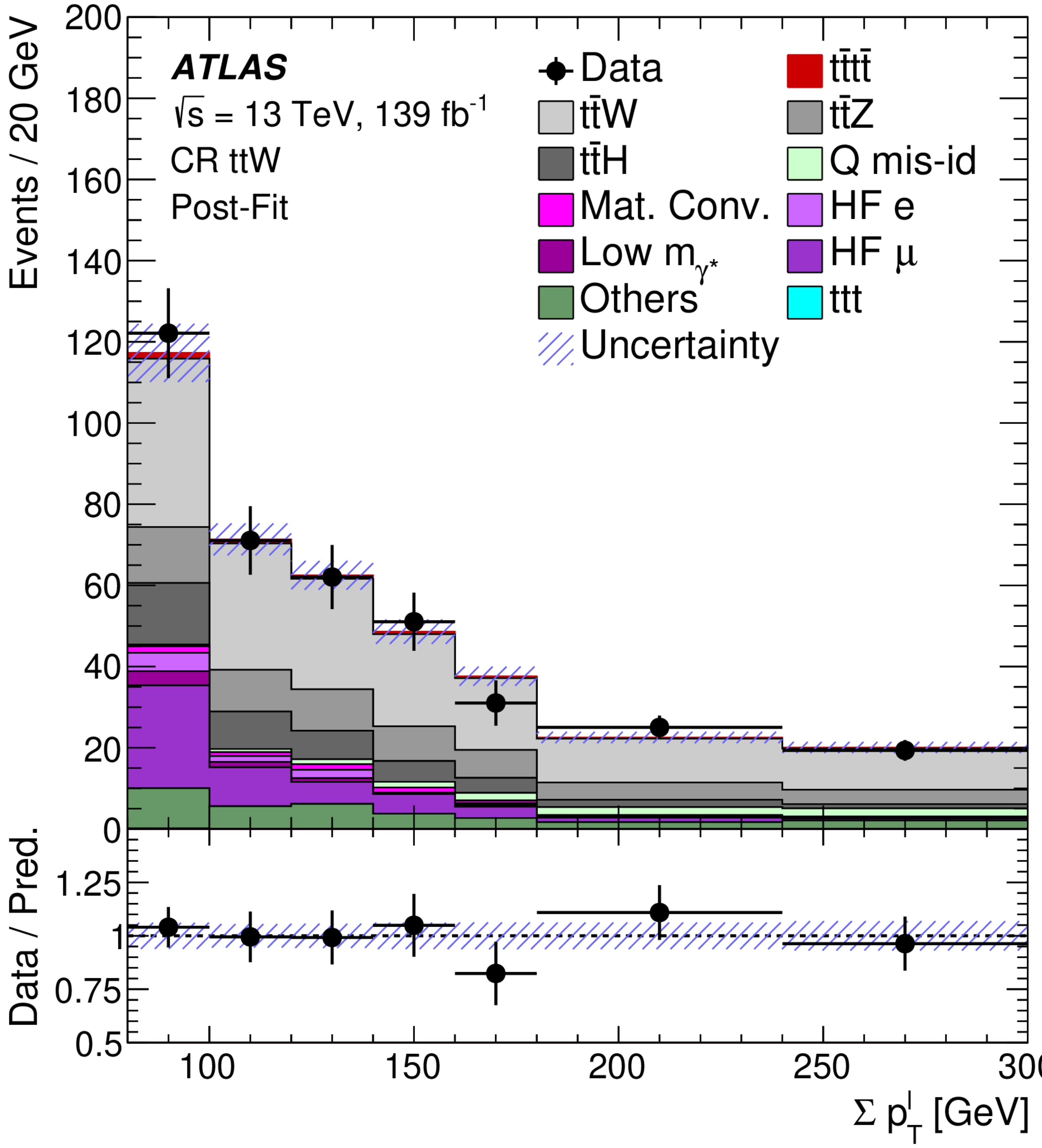
Goodness of $t\bar{t}W$ modelling checked in a looser region, exploiting $t\bar{t}W$ charge asymmetry to make a pure validation region.

- Mismodeling observed at ≥ 7 jets —> additional systematics
  - ttW with 7 (≥8) jets: 130 (300) %
  - ttW with 3 (≥4) b-jets: 50 %

Good post-fit modelling!







[SSML] Results

EPJC 80 (2020) 1085

A profile-likelihood fit is performed in control and signal regions simultaneously.

Normalization factors of the background sources are fitted together with the signal strength $\mu_{t\bar{t}t\bar{t}}$

- Only ttW NF significantly higher than expected, but compatible with previous ATLAS results, eg ttH(ML) [ATLAS-CONF-2019-045]
- Other NFs not significantly far from unity.

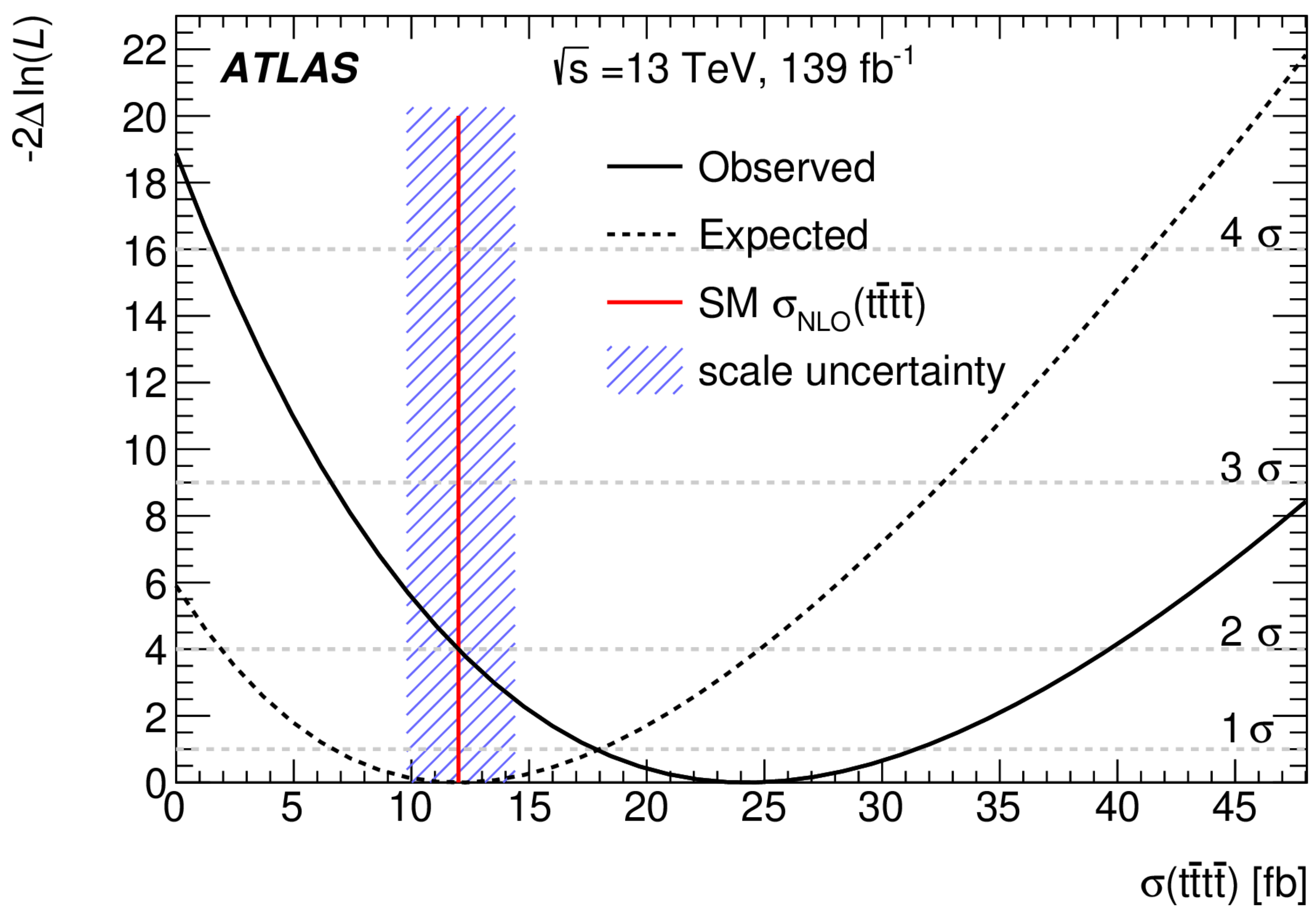
$$\sigma_{t\bar{t}t\bar{t}}^{SSML} = 24\,^{+5}_{-5}\,(\text{stat.})\,^{+5}_{-4}\,(\text{syst.}) = 24\,^{+7}_{-6}\,\text{fb}$$

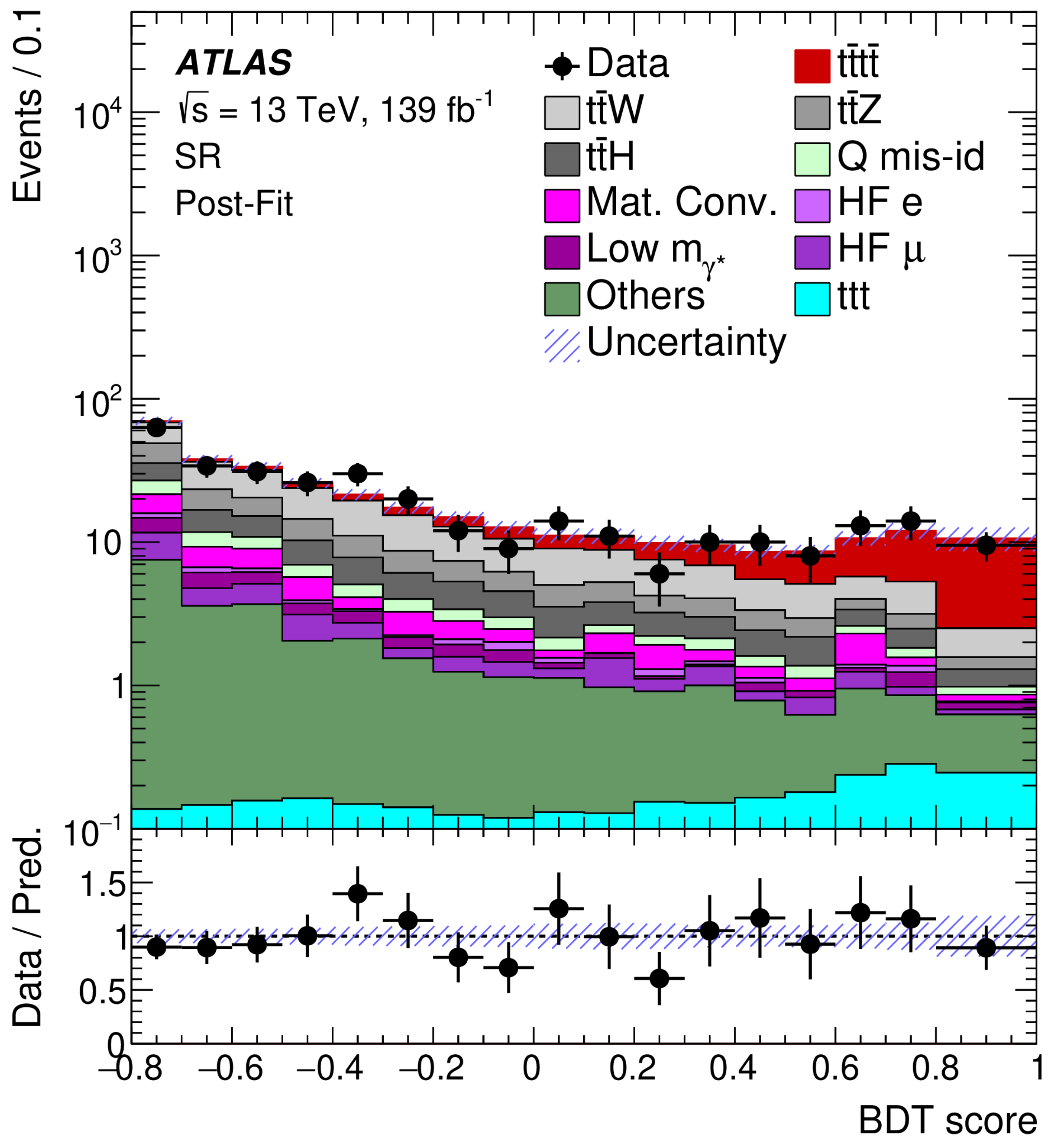
Parameter	$\text{NF}_{t\bar{t}W}$	$\text{NF}_{\text{Mat. Conv.}}$	$\text{NF}_{\text{Low } m_{\gamma^*}}$	$\text{NF}_{\text{HF } e}$	$\text{NF}_{\text{HF } \mu}$
Value	1.6 ± 0.3	1.6 ± 0.5	0.9 ± 0.4	0.8 ± 0.4	1.0 ± 0.4

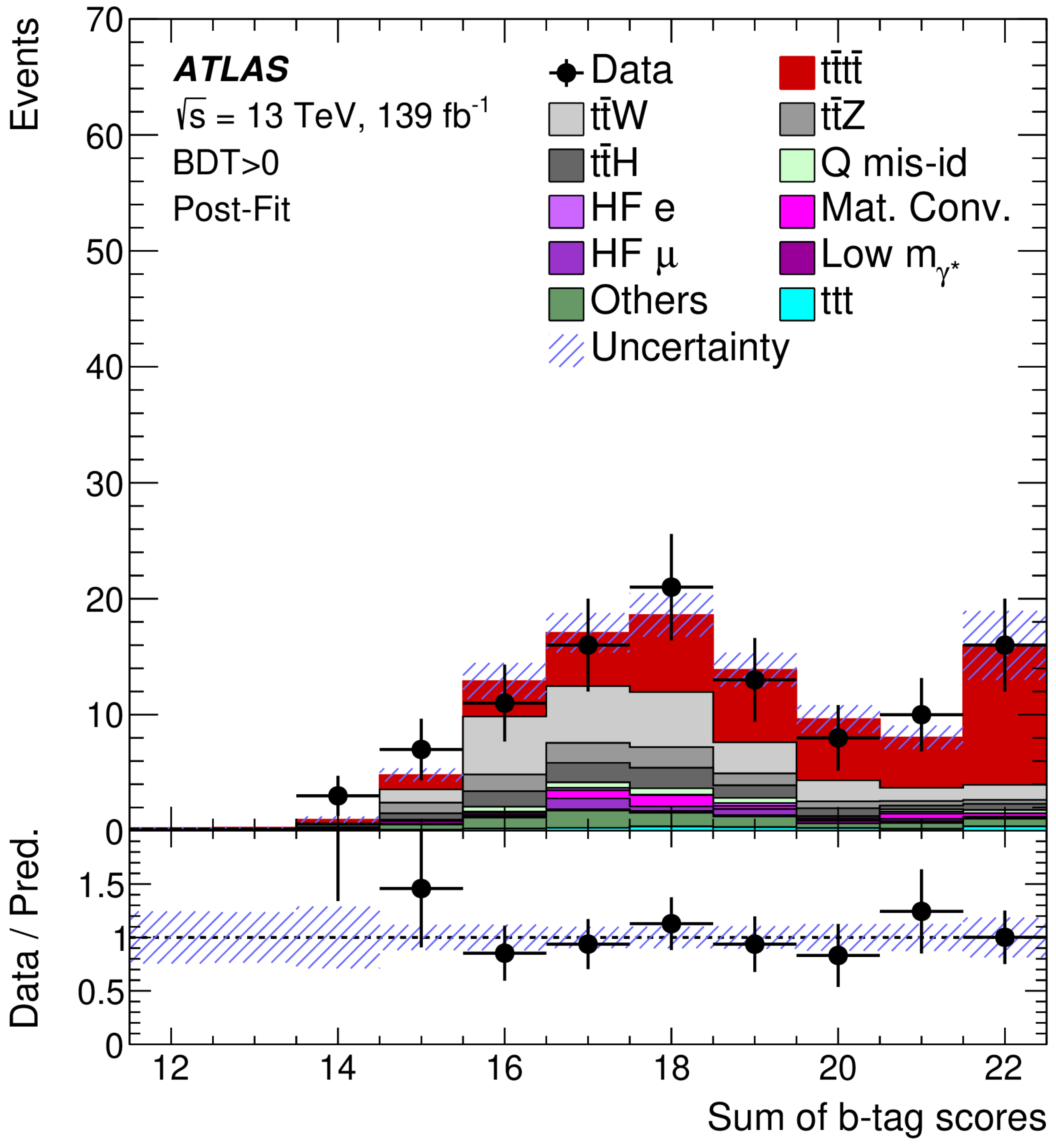
Observed (exp.) significance over background:

4.3 (2.4) σ

First evidence of four-top production!







[1L2LOS] Analysis strategy

2106.11683

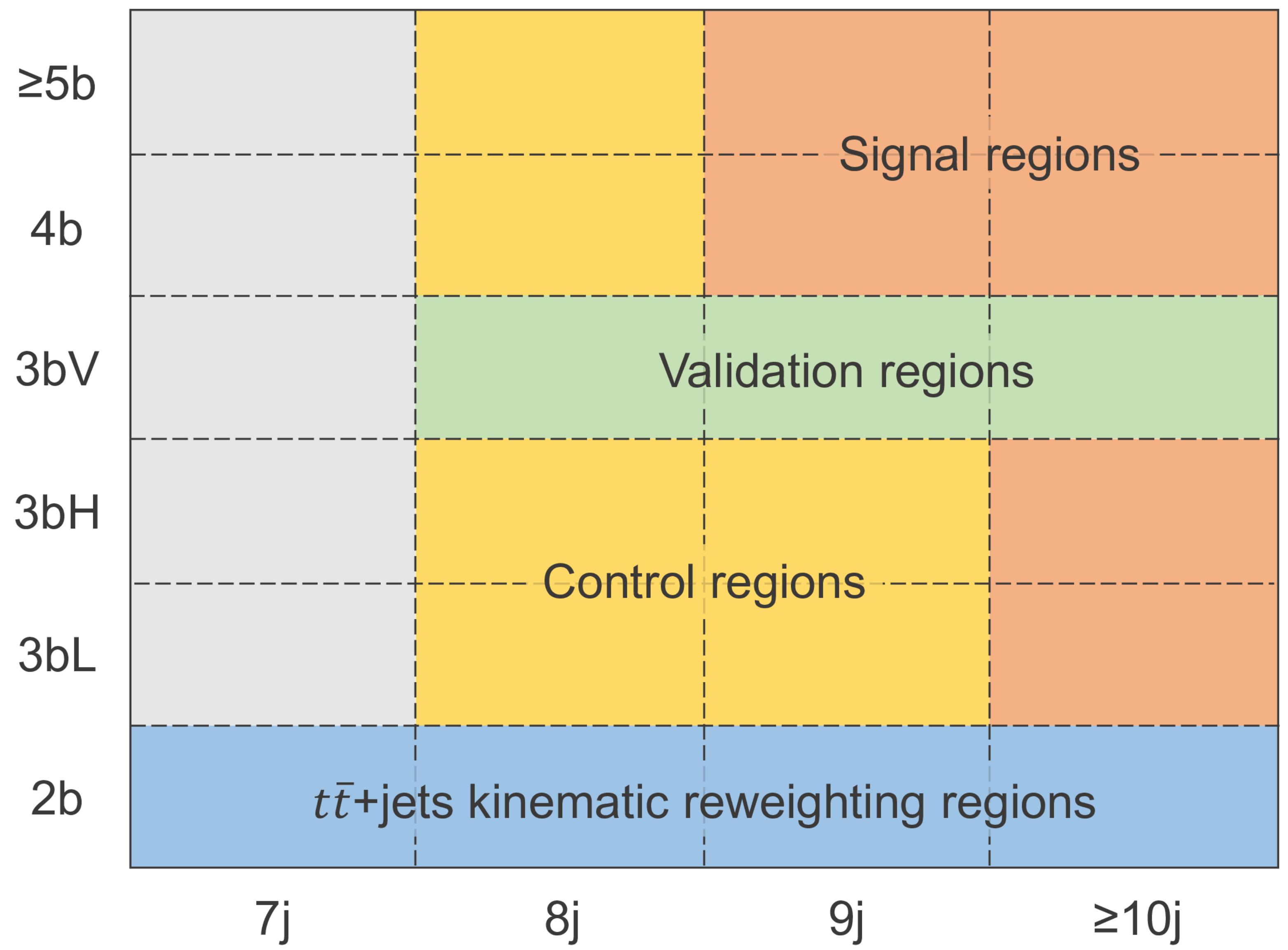
Analysis exploiting the full ATLAS Run 2 dataset (139 fb^{-1})

Selections for 1L (2LOS):

- 4-top signature: 10 (8) jets (of which 4 b-jets)
- ≥ 8 (6) jets of which (of which ≥3 b-tagged)
- 2-b-jets regions used for background estimation

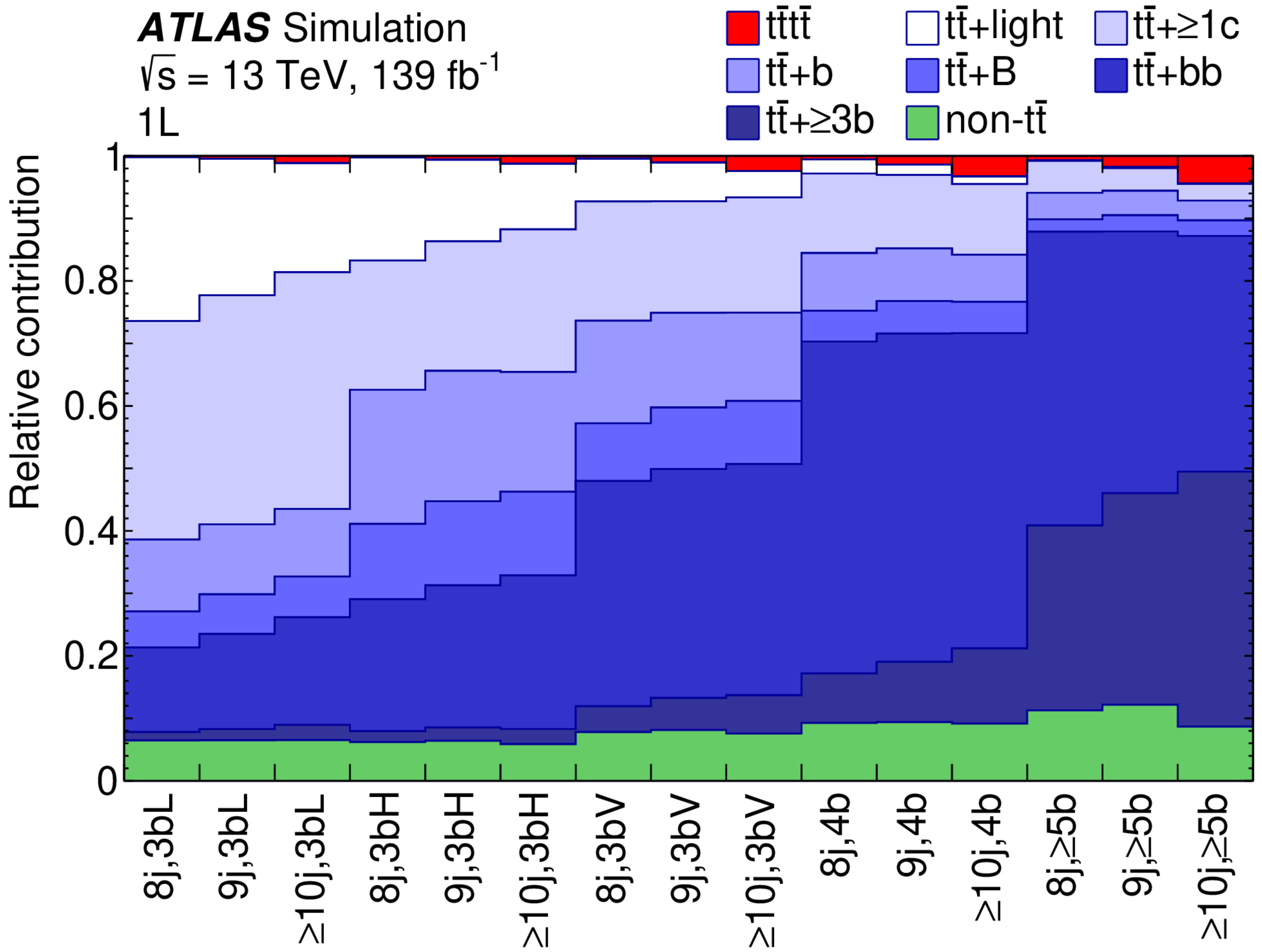
Regions are defined with increasing jet and b-jet multiplicity.

Here 1L, 2LOS in backup



Dominant background: top pair production associated to large (heavy-flavour) jet radiation — $t\bar{t}b\bar{b}$

Categorisation of $t\bar{t}$+jets in terms of flavour of the radiation, based on truth information.



3bL(oose)**, 3bH**(igh) **and 3bV**(alidation) are orthogonal and defined by using different b-tagging working points.
From 3bL to 3bV, $t\bar{t} + \geq 1b$ increases!

[1L2LOS] Background estimation

2106.11683

Significant mismodelling of $t\bar{t}$ background in the analysis regions (both normalisation and shape)

1. Flavour rescaling

$t\bar{t}$ + light, $t\bar{t} + \geq 1c$ and $t\bar{t} + \geq 1b$ are fitted to data in looser regions with ≥8 (≥6) jets for 1L (2LOS)

Four regions used: **2b, 3bL, 3bH and ≥4b**

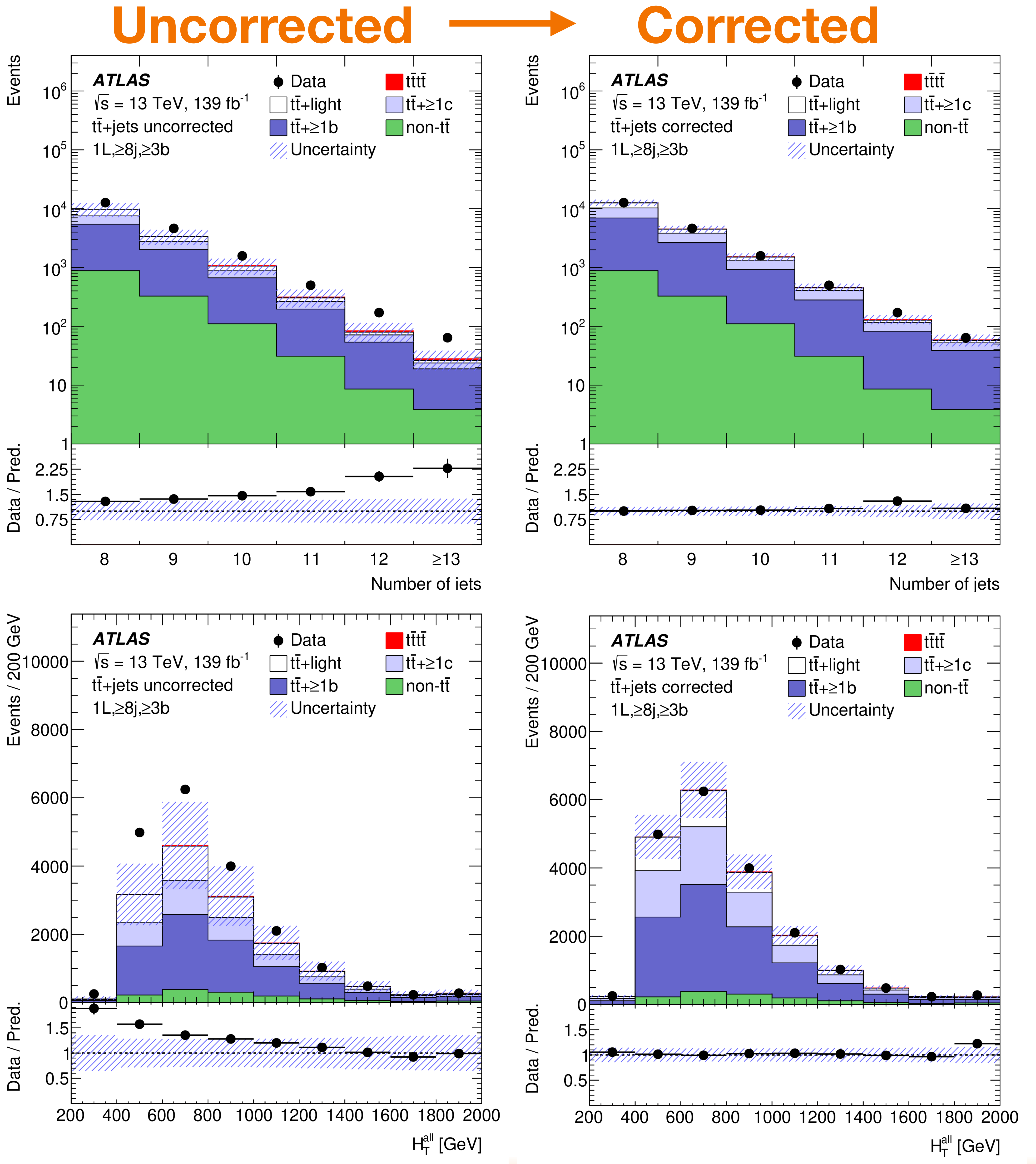
Large correction factors: $t\bar{t} + \geq 1c$ $(1b)$ get approx. 60% (30%) increase

2. Sequential kinematic reweighing

Total $t\bar{t}$ is reweighed to data in regions with 2 b-jets as a function of:

- Jet and reclustered jet mulitplicities
- $H_T^{all,red}$: sum of the p_T of all jets in the events "normalized" by the number of jets
- ΔR^{jj}_{avg} : average angular distance between jets

Improved agreement and reduced uncertainties.



[1L2LOS] Fit setup and systematic model

2106.11683

A total of **21** regions (12 in 1L + 9 in 2LOS) are considered in the fit.

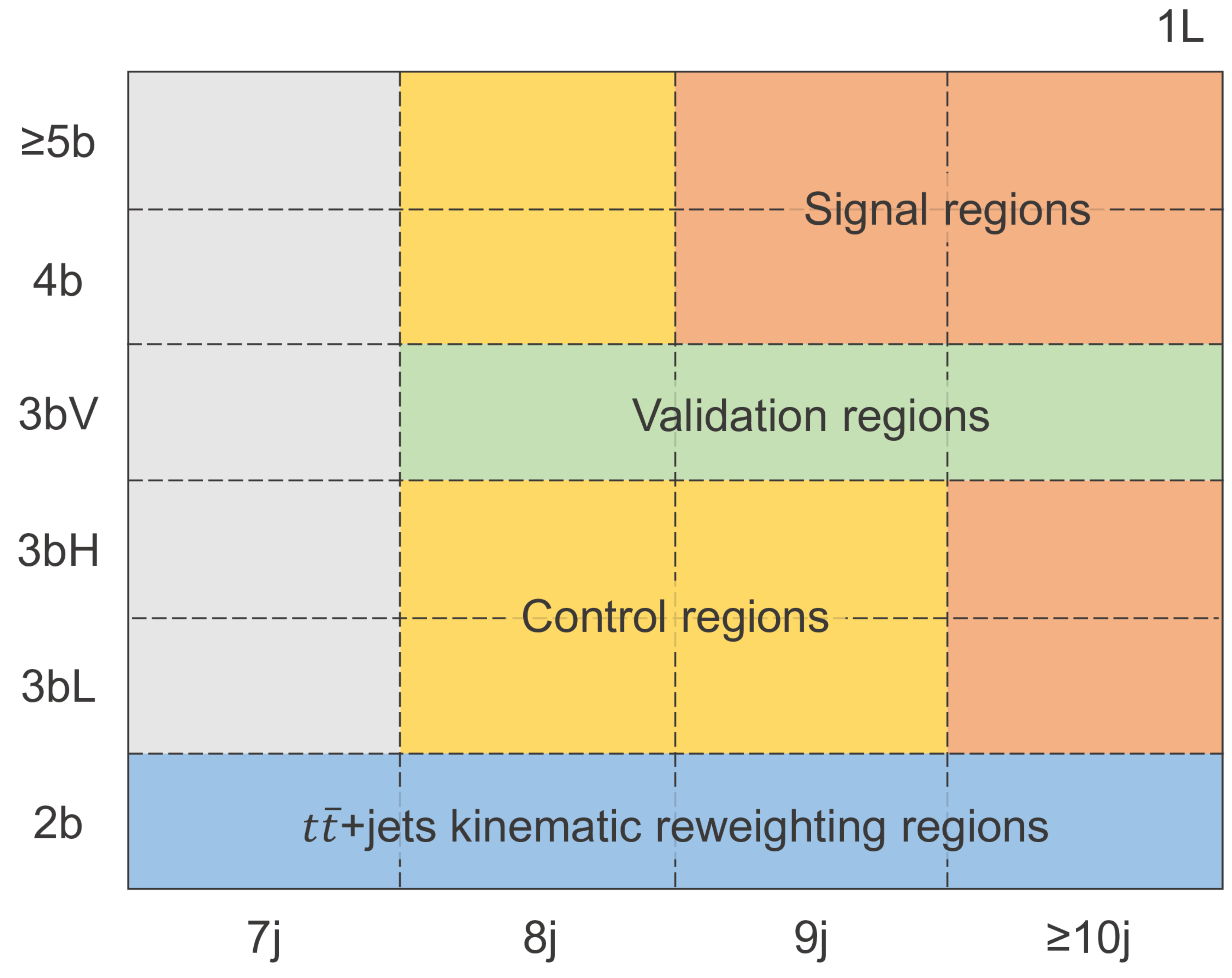
- Including **6** validation regions (3bV)

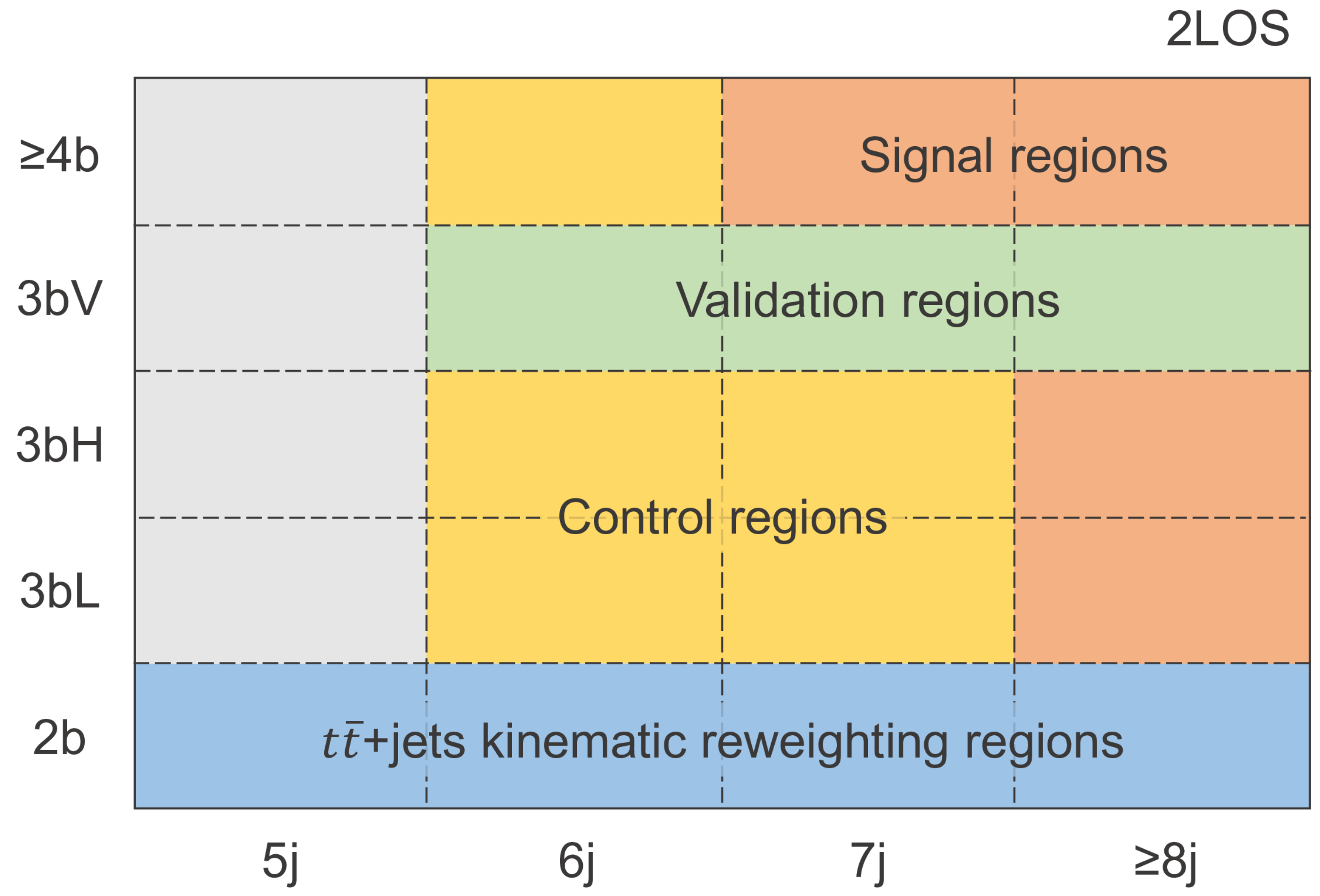
3bV regions are used to validate the $t\bar{t}$+jets modelling closest to the signal region.

H_T^{all} used for the profile likelihood fit in all control regions.
Boosted Decision Trees score is used in signal-regions.

Very complex **systematic model** to account for mismodelling in this extreme phase-space (more in the backup)

- All instrumental systematic uncertainties are included.
- Inflated modelling uncertainties for non-$t\bar{t}$ background
- Detailed systematic model for $t\bar{t}$ modelling (45 nuisance parameters)
  - Additional uncertainty associated to heavy-flavour radiation.
  - 4FS vs 5FS additional uncertainty
  - All uncertainties split in terms of radiation flavour.
  - 2-point systematics split in shape and migration





[1L2LOS] Results

2106.11683

Excellent post-fit agreement with data in all regions

$$\sigma^{1L2LOS}_{t\bar{t}t\bar{t}} = 24\,^{+8}_{-8}\,(\text{stat.})\,^{+15}_{-13}\,(\text{syst.}) = 24\,^{+17}_{-15}\,\text{fb}$$

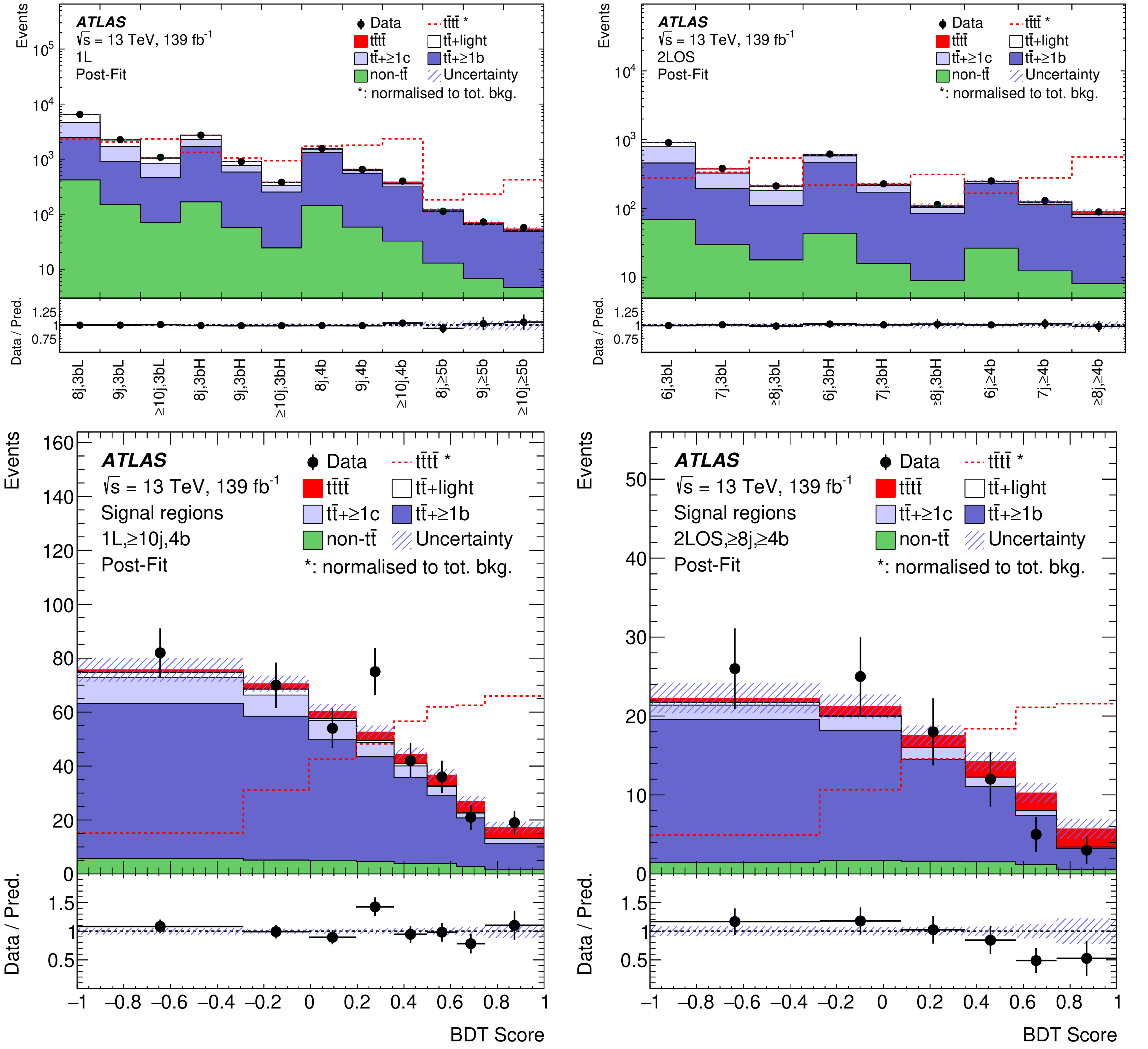
Corresponding to **2.2** times the SM cross-section (compatible within 1 std. dev.)

Observed (exp.) significance of the signal over background: **1.9 (1.0)**

Well compatible with SSML results.

Most important systematics:

- Signal modelling (parton-shower & cross-section)
- $t\bar{t}b\bar{b}$ 4FS/5FS & $t\bar{t} + \geq 1c$ normalisation uncertainties
- Light jet mis-tag rate



Combination

The two results from SSML and 1L2LOS are combined.

Most dominating systematics are totally different: limited impact of systematics correlations.

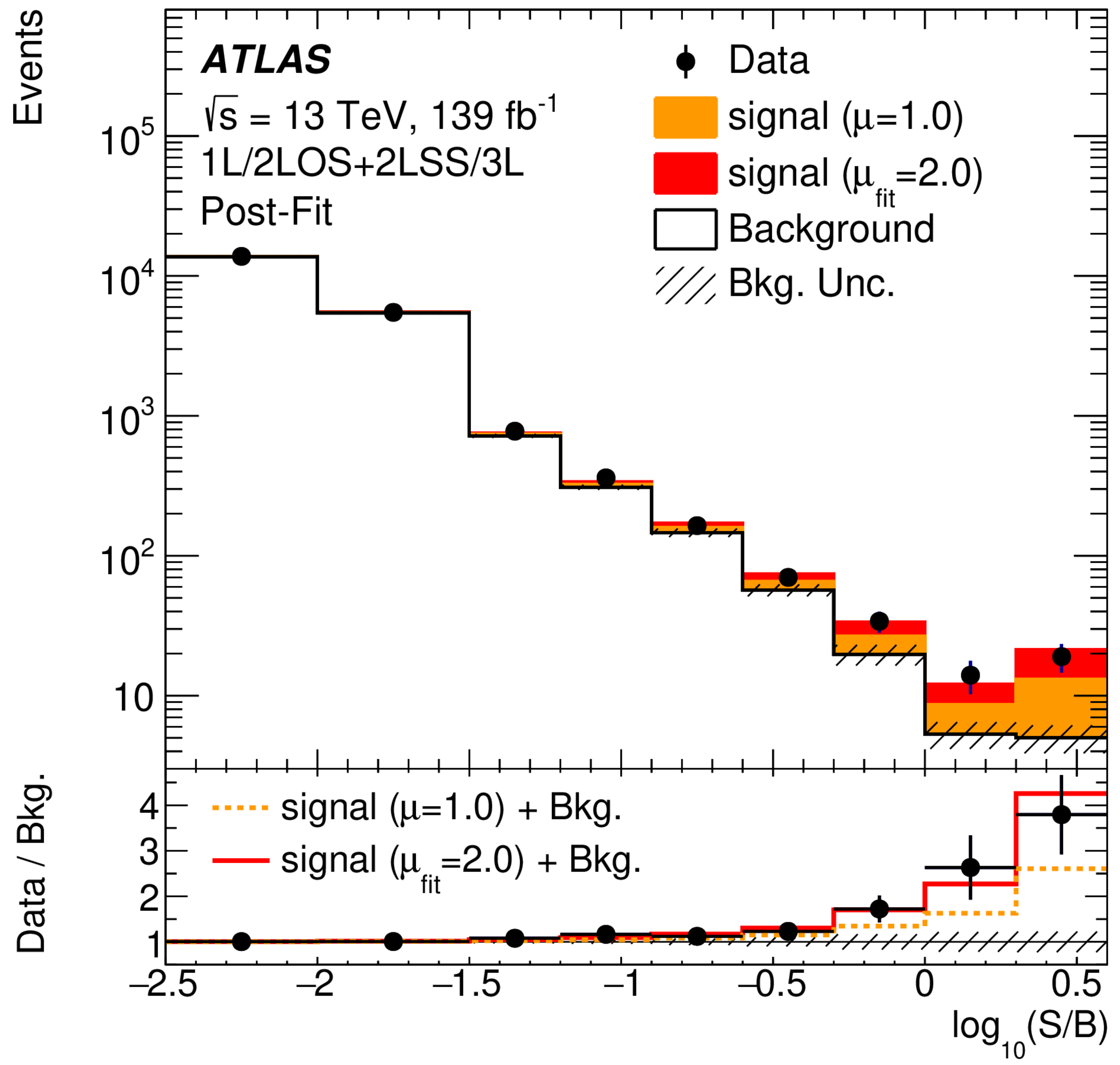
Instrumental systematics are fully correlated, as well as non-$t\bar{t}$ and non-$t\bar{t}W$ modelling uncertainties

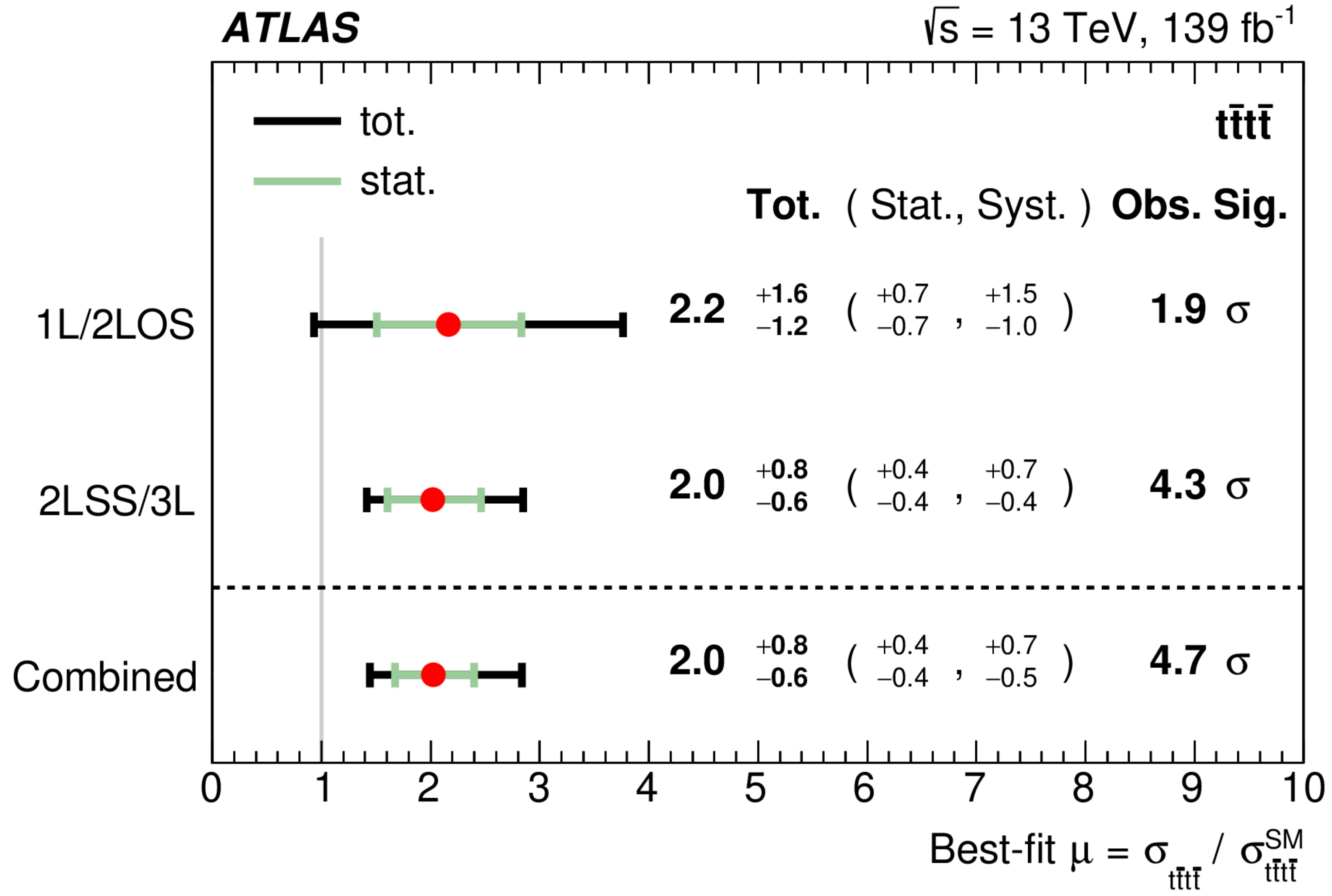
SSML result dominates in the combination.

$$\sigma_{t\bar{t}t\bar{t}}^{1L2LOS/SSML} = 24\,^{+4}_{-4}\,(\text{stat.})\,^{+5}_{-4}\,(\text{syst.}) = 24\,^{+7}_{-6}\,\text{fb}$$

Observed (expected) significance 4.7 (2.6) std. deviation.

Improved result with respect the single channels.





Conclusions

Thanks a lot for the attention!!

The latest ATLAS results on the measurements of SM four top-quark production are presented.

Two analyses performed with the full Run2 dataset in different channels

Same-sign dilepton and trilepton channel (SSML)

First observed **evidence** of the four-top quark production, with a significance of **4.3** std. deviations.

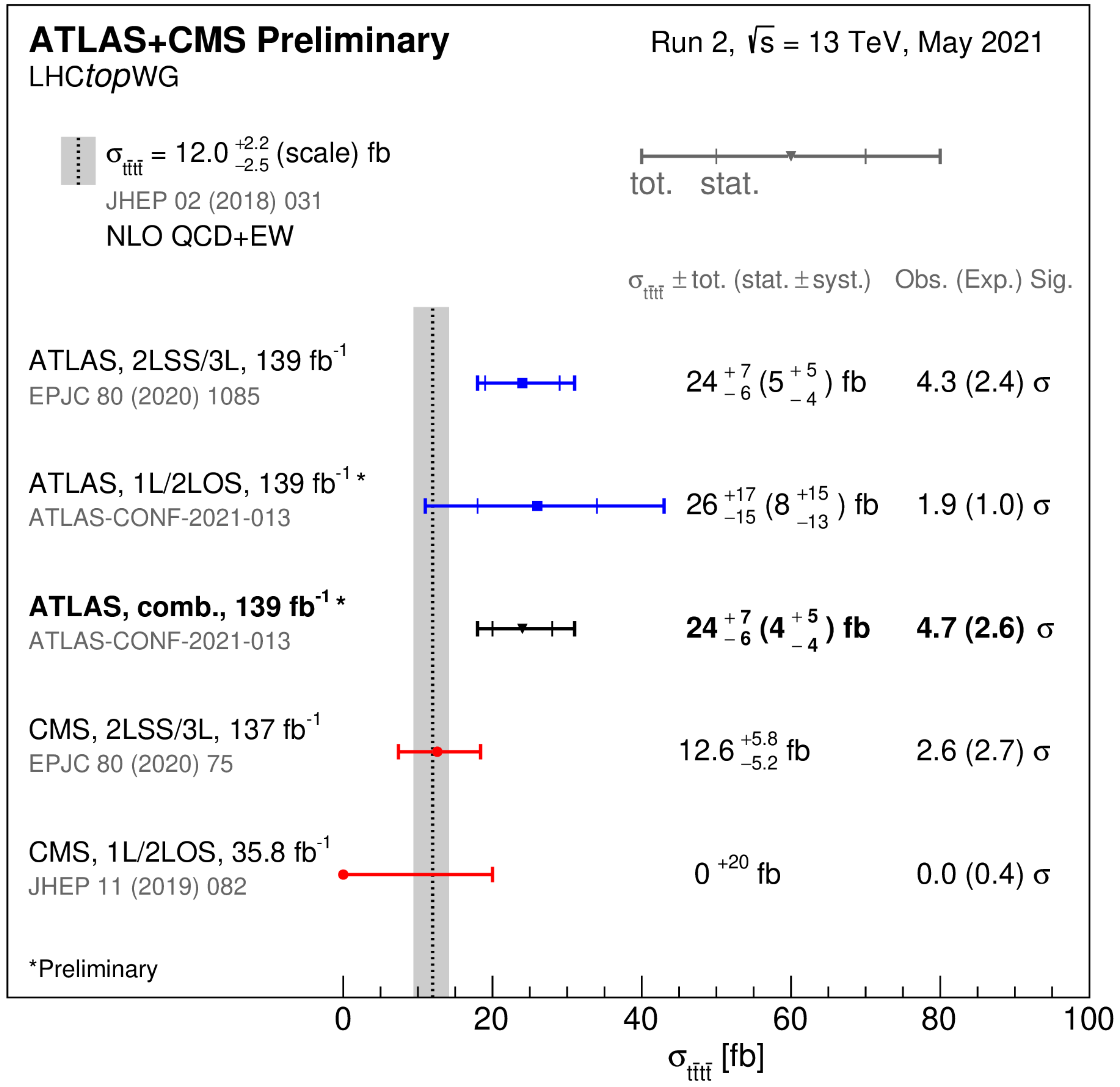
Single lepton and opposite-sign dilepton channel (1L2LOS)

Observed significance of **1.9** std. deviations, in line with the SSML results.

Combination of 1L2LOS + SSML

Dominated by SSML channel result: $\sigma_{t\bar{t}t\bar{t}}^{1L2LOS/SSML} = 24\,^{+7}_{-6}$ fb

Cross-section is compatible within 2 std. dev. with SM predictions



The $t\bar{t}t\bar{t}$ is not a mystery anymore: time for refined analyses, more stats (Run3) and interpretation!

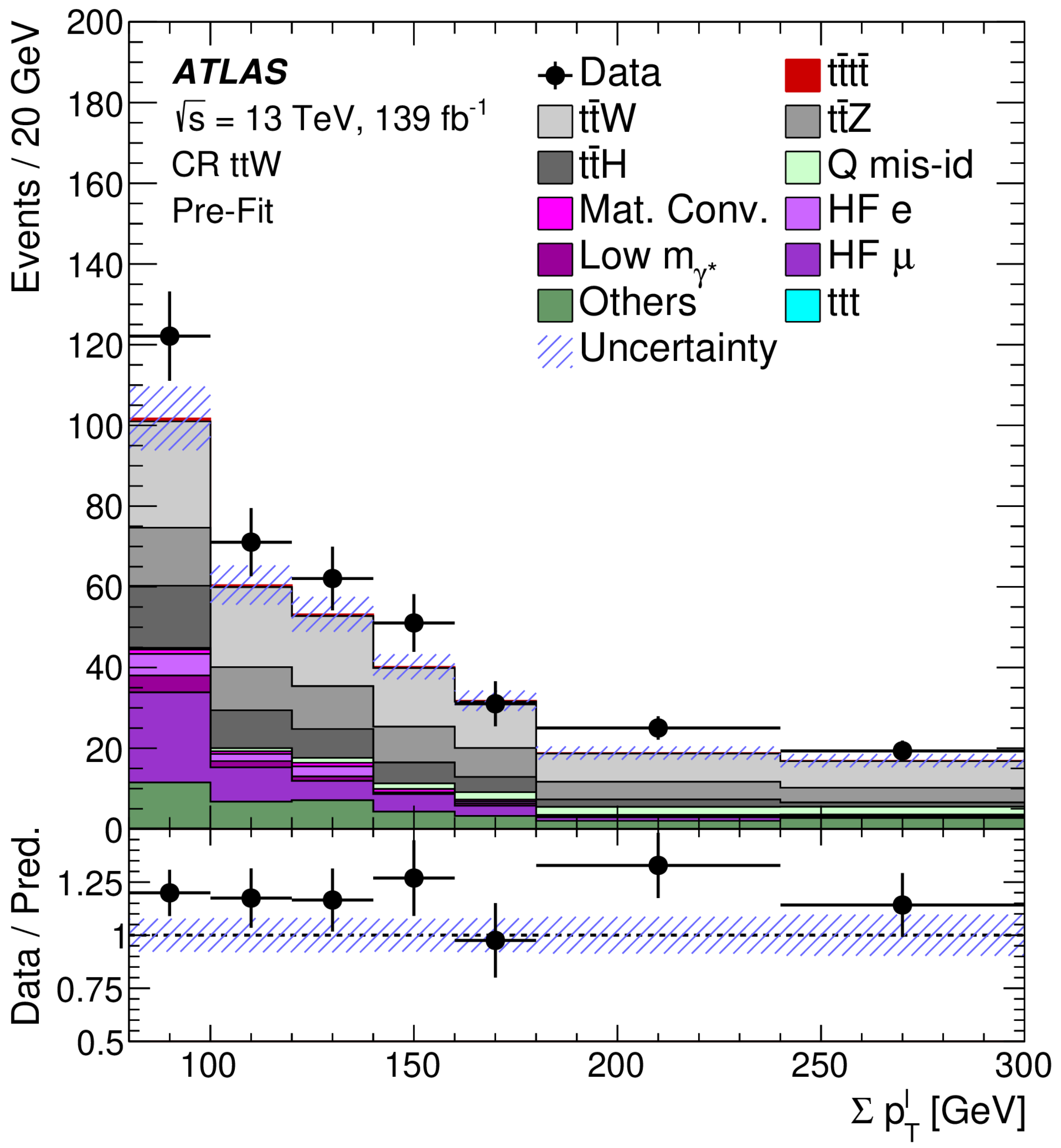
Backup

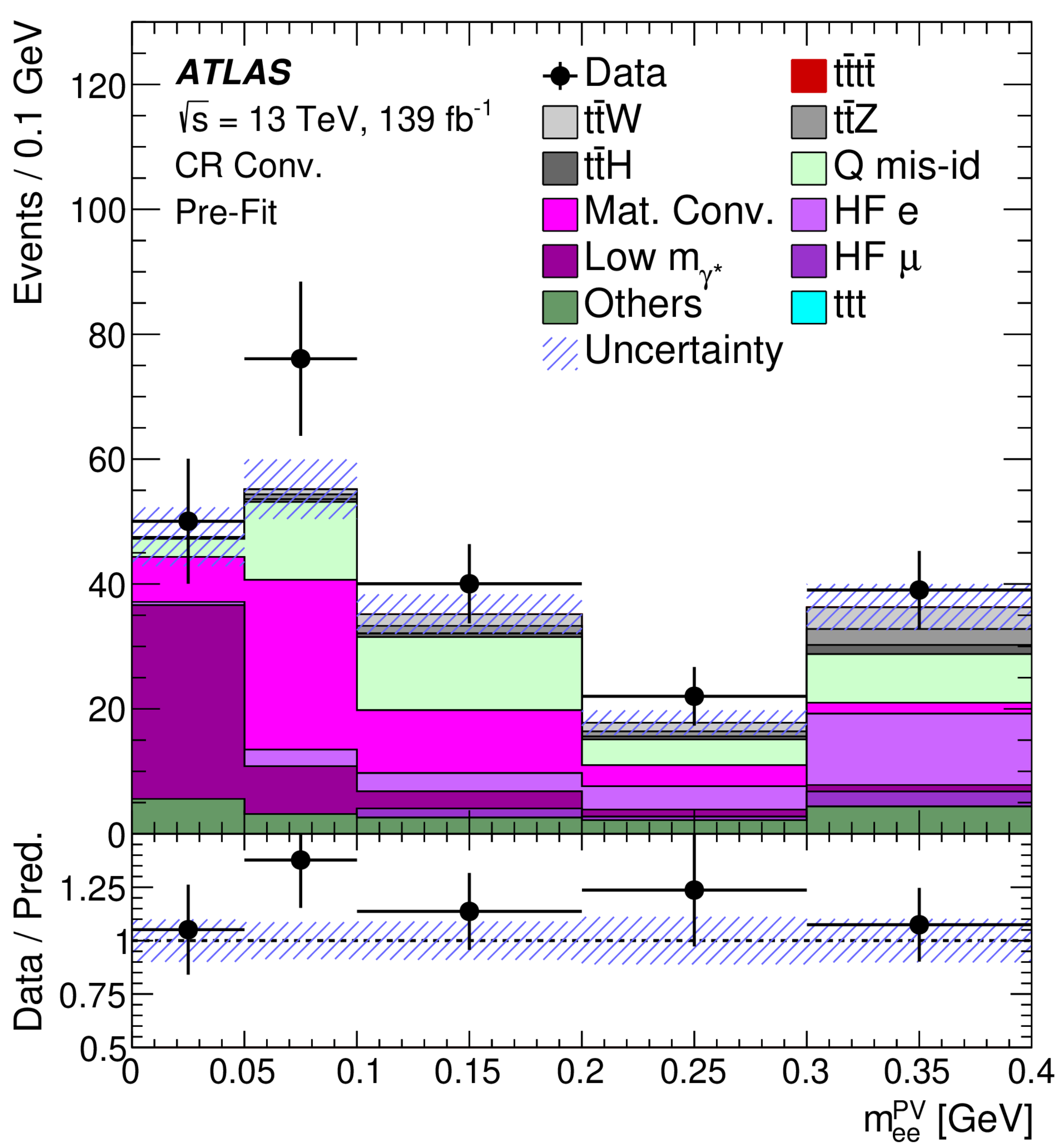
[SSML] Region definition

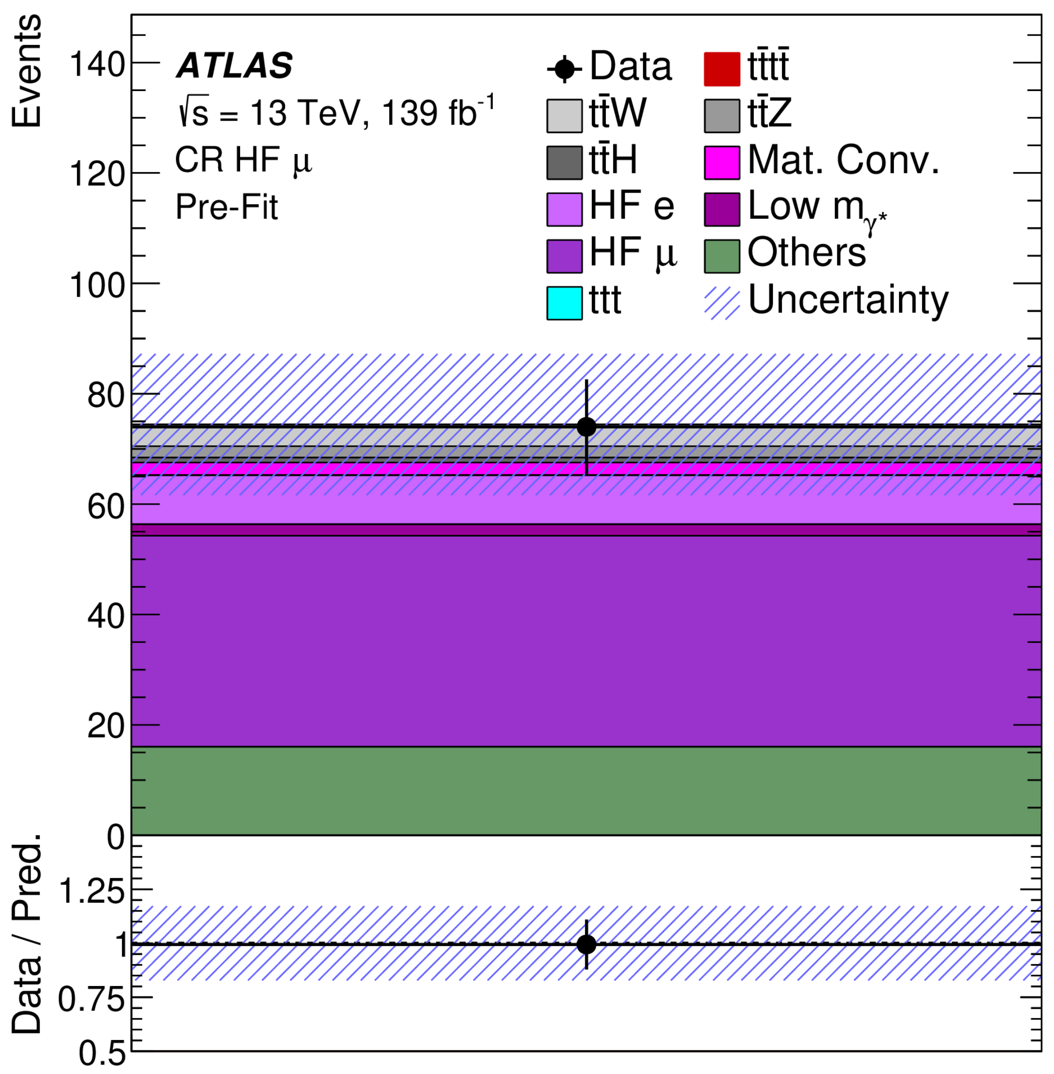
EPJC 80 (2020) 1085

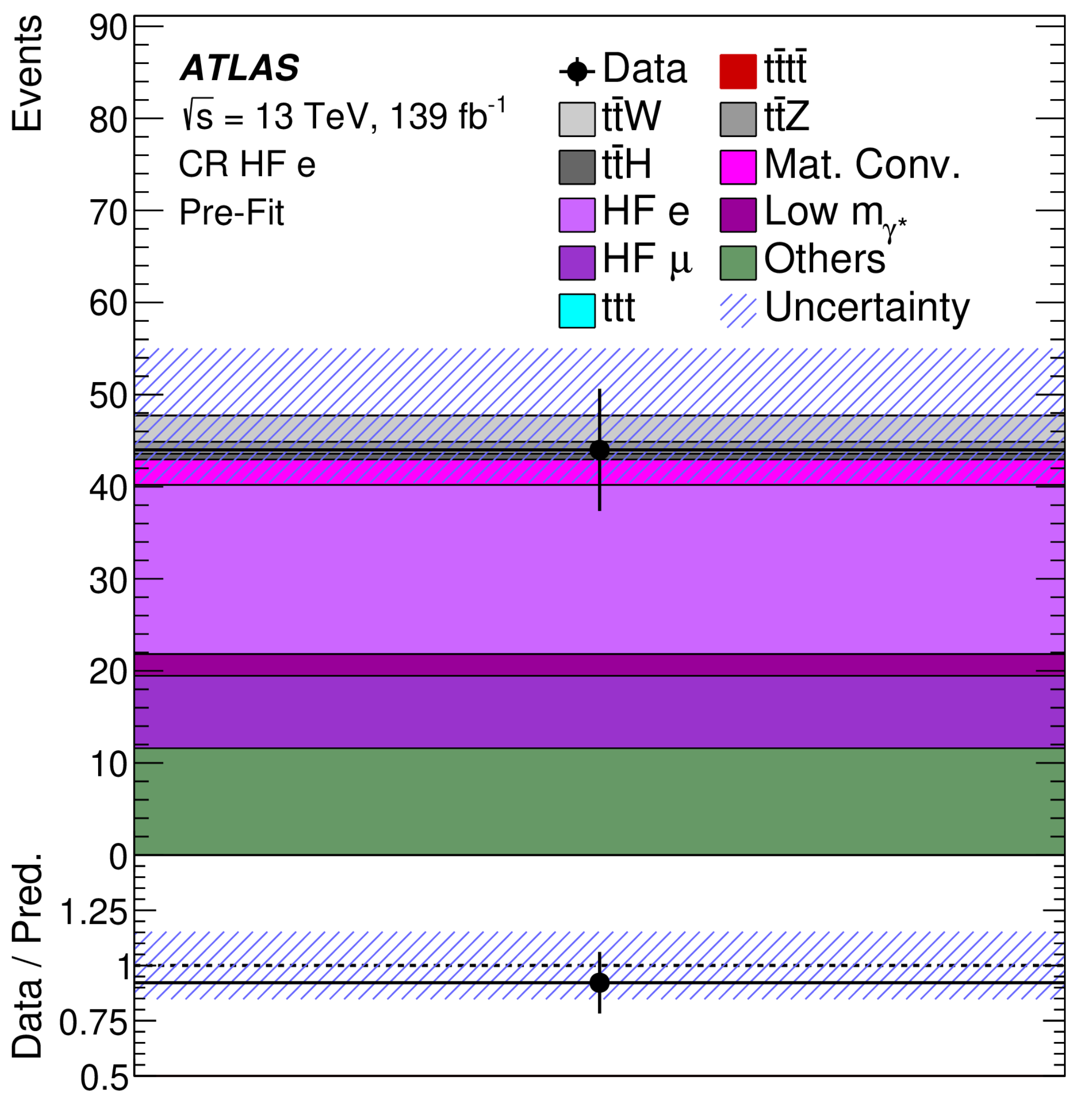
Region	Channel	N_j	N_b	Other requirements	Fitted variable
SR	2LSS/3L	≥ 6	≥ 2	$H_T > 500$	BDT
CR Conv.	$e^\pm e^\pm \| e^\pm \mu^\pm$	$4 \leq N_j < 6$	≥ 1	$m_{ee}^{CV} \in [0, 0.1\ \text{GeV}]$ $200 < H_T < 500$ GeV	m_{ee}^{PV}
CR HF e	$eee \| ee\mu$	-	$= 1$	$100 < H_T < 250$ GeV	counting
CR HF μ	$e\mu\mu \| \mu\mu\mu$	-	$= 1$	$100 < H_T < 250$ GeV	counting
CR ttW	$e^\pm \mu^\pm \| \mu^\pm \mu^\pm$	≥ 4	≥ 2	$m_{ee}^{CV} \notin [0, 0.1\ \text{GeV}]$, $\lvert\eta(e)\rvert < 1.5$ for $N_b = 2$, $H_T < 500$ GeV or $N_j < 6$ for $N_b \geq 3$, $H_T < 500$ GeV	Σp_T^ℓ

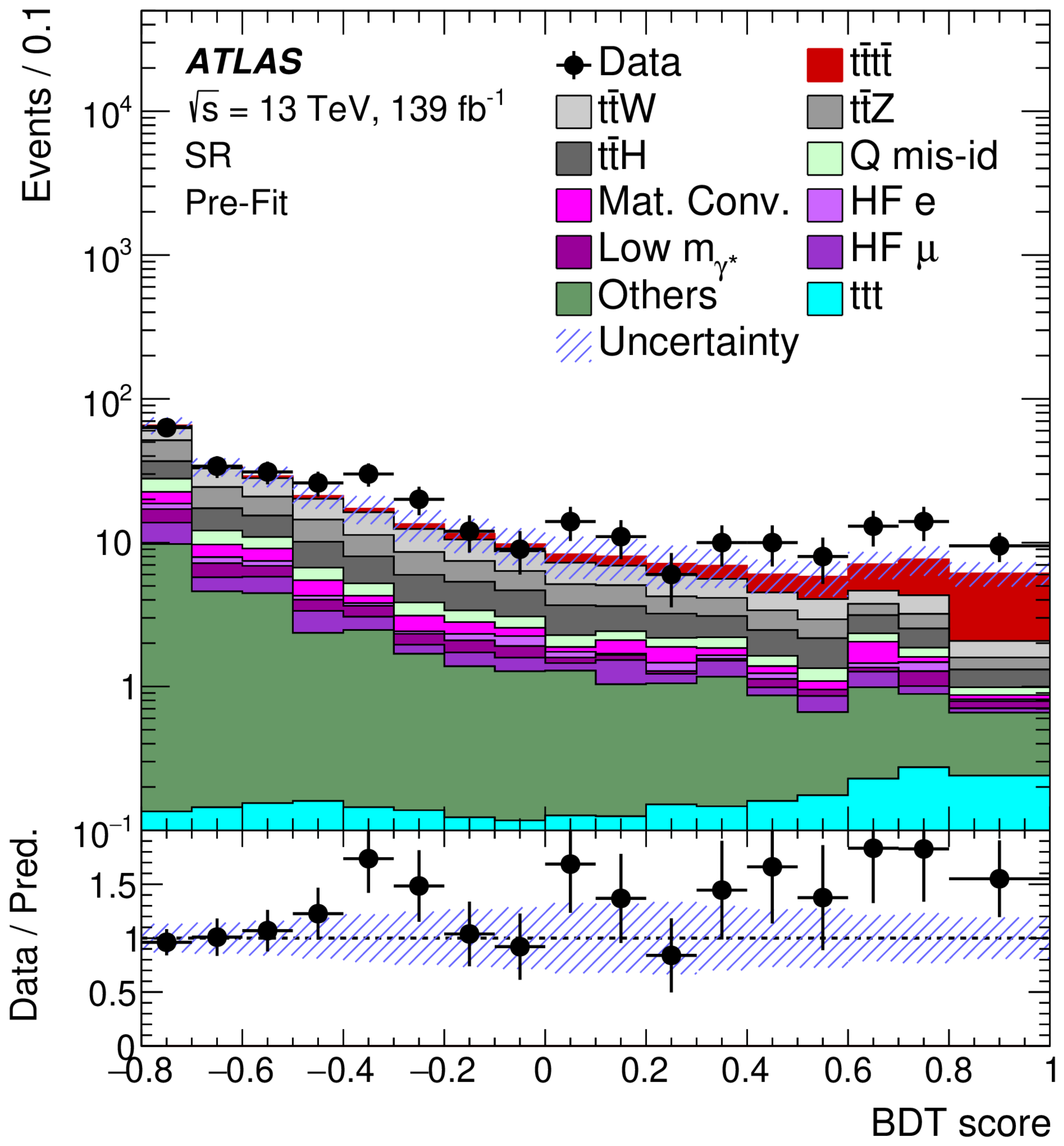
[SSML] Pre-fit plots











[SSML] ttW enhancement

2106.11683

	Pre-fit		Post-fit
Number of ttW events in SR (BDT > 0):	12.4 ± 8.8	$\longrightarrow$	23.2 ± 10.1

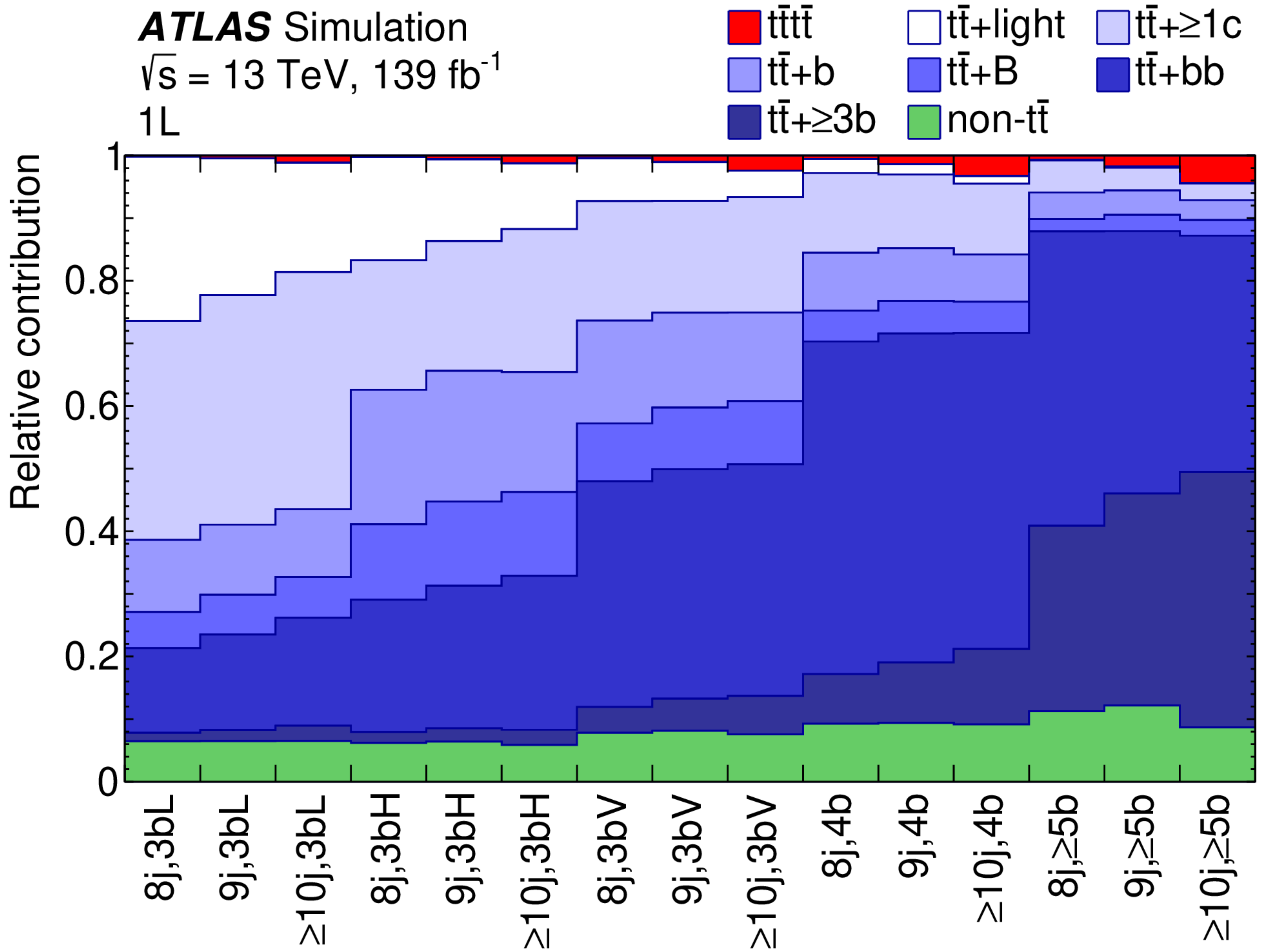
NP	Value	ttW increase
NF	1.6 ± 0.3	60%
ttW w/ 7 jets	$0.18^{+0.73}_{-0.61}$	22%
ttW w/ ≥8 jets	$0.22^{+0.56}_{-0.42}$	65%

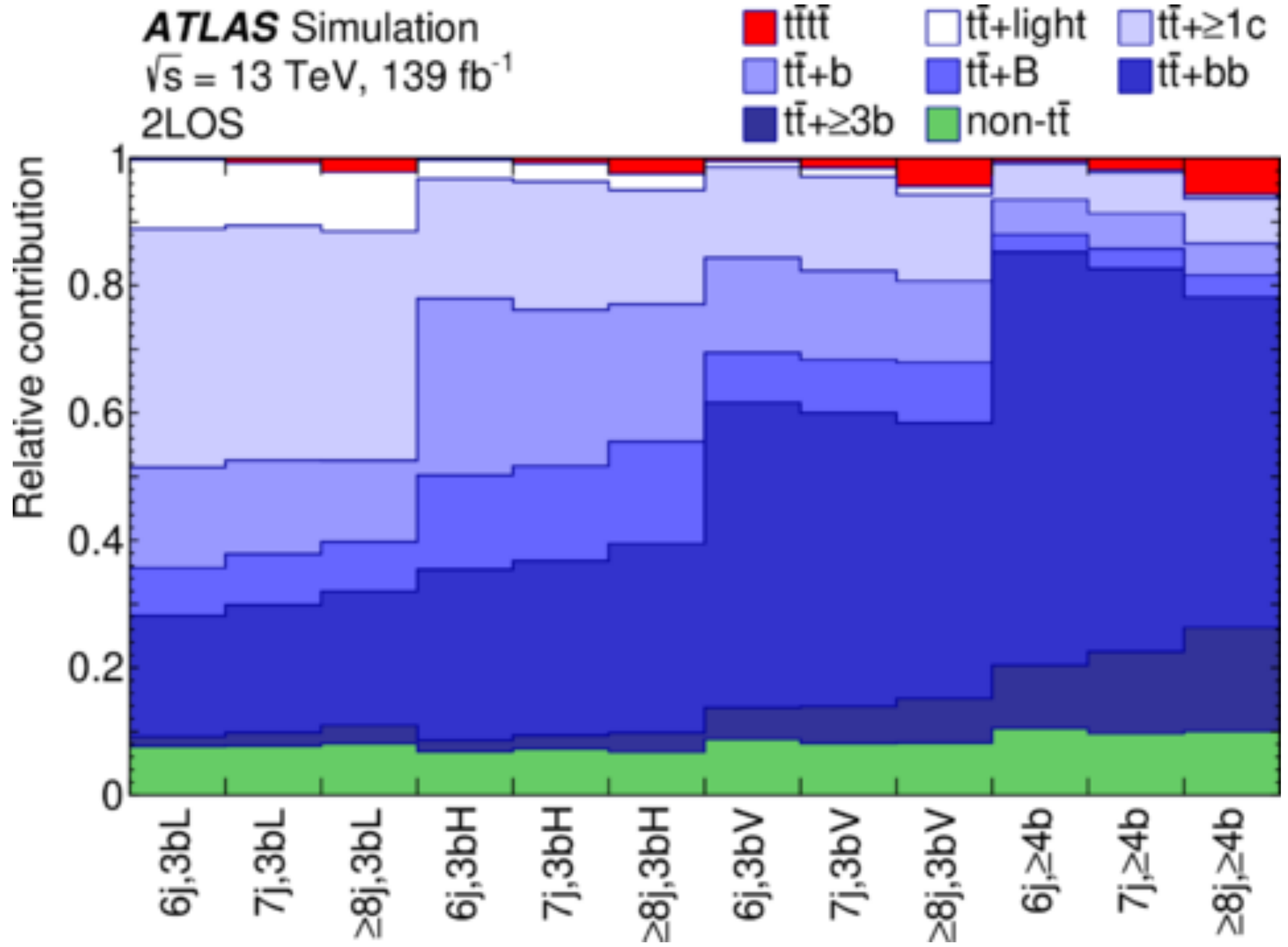
[1L2LOS] Radiation categorisation

2106.11683

Categorisation of $t\bar{t}$+jets in terms of flavour of the radiation:

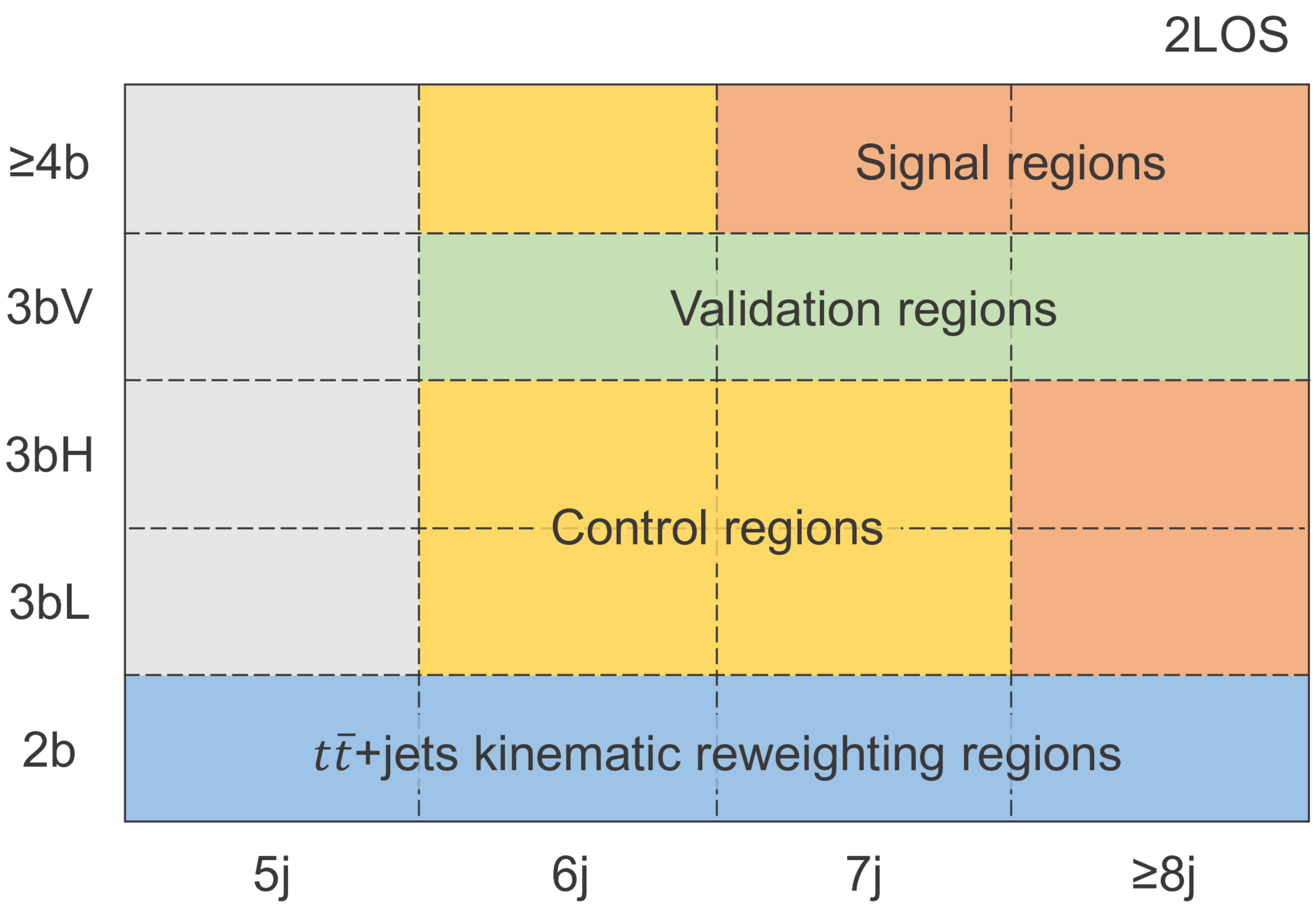
- $t\bar{t} + bb$: the jet radiation contains at ≥2 particle jets matching with 2 b hadrons
- $t\bar{t} + B$: the jet radiation contains 1 particle jets matching with 2 b hadrons
- $t\bar{t} + b$: the jet radiation contains 1 particle jets matching with 1 b hadron
- $t\bar{t} + \geq 3b$: all the other cases with ≥1 b hadron in the jet radiation
- $t\bar{t} + \geq 1c$: the jet radiation contains ≥1 particle jet matching with ≥ 1 c-hadron and no b-hadrons
- $t\bar{t} + \geq 1c$: no particle jets are matched with b- or c-hadrons

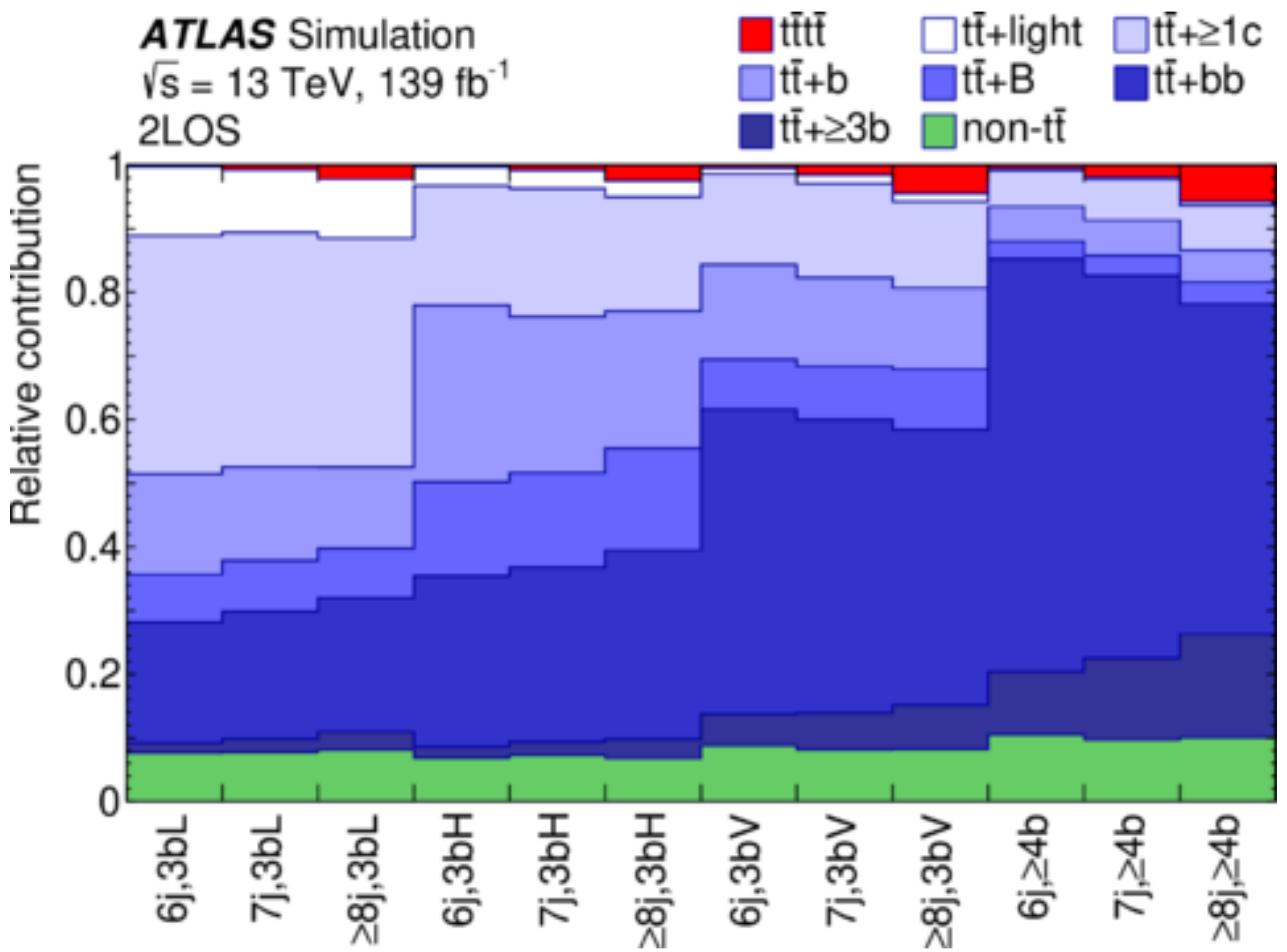




[1L2LOS] 2LOS regions and composition

2106.11683





[1L2LOS] Top pair systematics model

2106.11683

Uncertainty source	Description	Components (number)
$t\bar{t}+\geq 1b$ normalisation	$\pm 50\%$	$t\bar{t}+b$, $t\bar{t}+b\bar{b}$, $t\bar{t}+B$, $t\bar{t}+\geq 3b$ (4)
$t\bar{t}+\geq 1c$ normalisation	$\pm 50\%$	$t\bar{t}+\geq 1c$ (1)
Generator choice	POWHEG vs MADGRAPH5_AMC@NLO	($t\bar{t}$+light, $t\bar{t}+\geq 1c$, $t\bar{t}+b$, $t\bar{t}+b\bar{b}$, $t\bar{t}+B$, $t\bar{t}+\geq 3b$) $\otimes$ (shape, migration) (12)
PS choice	PYTHIA 8 vs HERWIG 7	($t\bar{t}$+light, $t\bar{t}+\geq 1c$, $t\bar{t}+b$, $t\bar{t}+b\bar{b}$, $t\bar{t}+B$, $t\bar{t}+\geq 3b$) $\otimes$ (shape, migration) (12)
Renormalisation scale	Varying μ_r in POWHEG	$t\bar{t}$+light, $t\bar{t}+\geq 1c$, $t\bar{t}+\geq 1b$ (3)
Factorisation scale	Varying μ_f in POWHEG	$t\bar{t}$+light, $t\bar{t}+\geq 1c$, $t\bar{t}+\geq 1b$ (3)
ISR	Varying α_S^{ISR}(PS) in PYTHIA 8	$t\bar{t}$+light, $t\bar{t}+\geq 1c$, $t\bar{t}+\geq 1b$ (3)
FSR	Varying μ_f (PS) in PYTHIA 8	$t\bar{t}$+light, $t\bar{t}+\geq 1c$, $t\bar{t}+\geq 1b$ (3)
5FS vs 4FS	POWHEGBOXRES (4FS) vs POWHEGBOX (5FS)	$t\bar{t}+b$, $t\bar{t}+b\bar{b}$, $t\bar{t}+B$, $t\bar{t}+\geq 3b$ (4)

[1L2LOS] 3b region flavour splitting

2106.11683

Classification made by making use of different working point of the b-tagging algorithm

- **3bL**: 3 b-tagged jets at 70% but with low scores of the b-tagging classifier
- **3bH**: 3 b-tagged jets at 70% but with intermediate scores of the b-tagging classifier
- **3bV**: 3 b-tagged jets at 70% but with high scores of the b-tagging classifier

Name	$N_b^{60\%}$	$N_b^{70\%}$	$N_b^{85\%}$
2b	-	$= 2$	-
3bL	≤ 2	$= 3$	-
3bH	$= 3$	$= 3$	$= 3$
3bV	$= 3$	$= 3$	≥ 4
≥4b (2LOS)	-	≥ 4	-
4b (1L)	-	$= 4$	-
≥5b (1L)	-	≥ 5	-