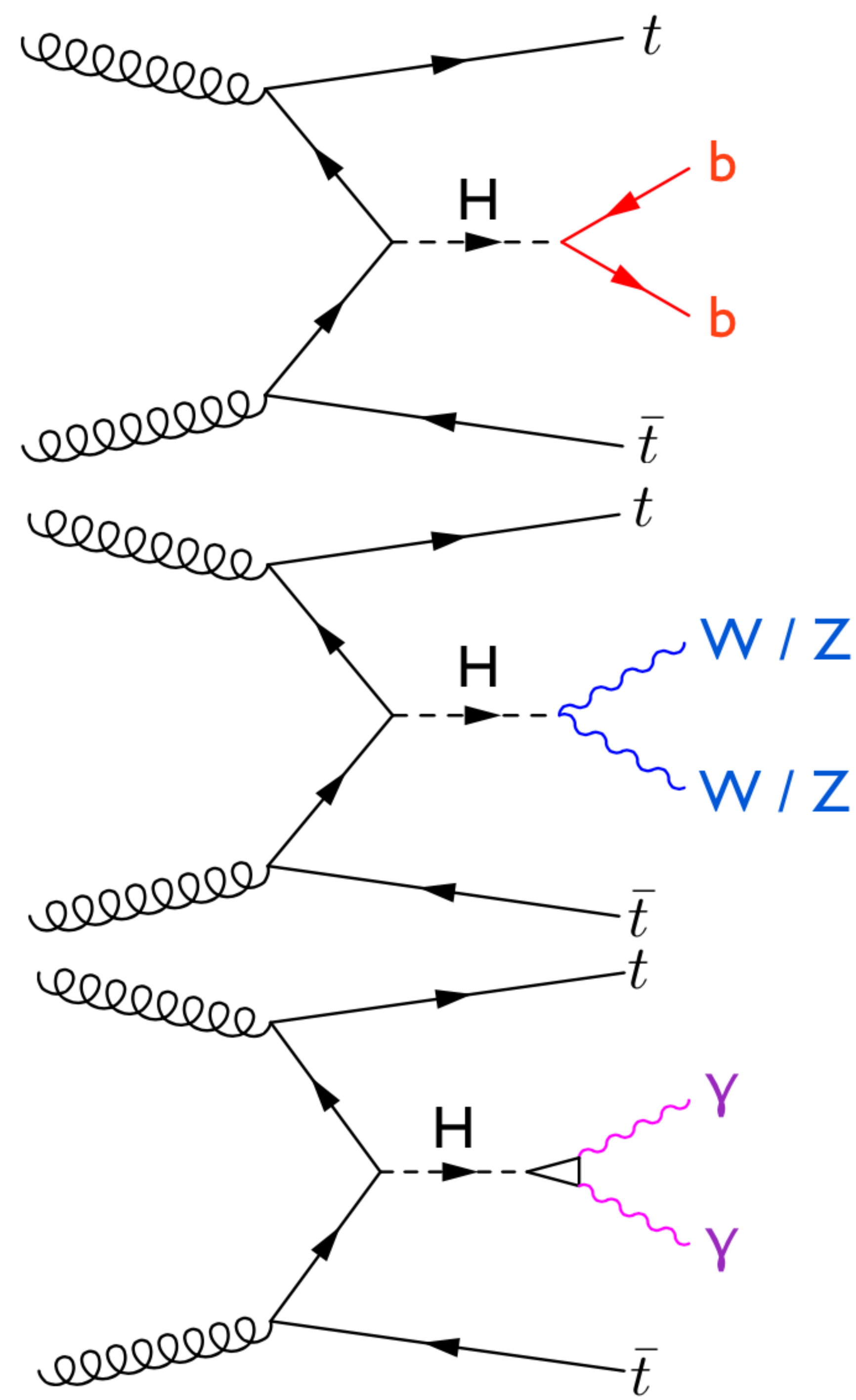
ATLAS
EXPERIMENT
Search for the associated production of a
Top Quark Pair
& Higgs Boson
at the LHC
Tamara Vázquez Schröder
McGill University
Particle Physics Highlight Seminar
University of Liverpool - 02/05/18
Run: 300571

Outline

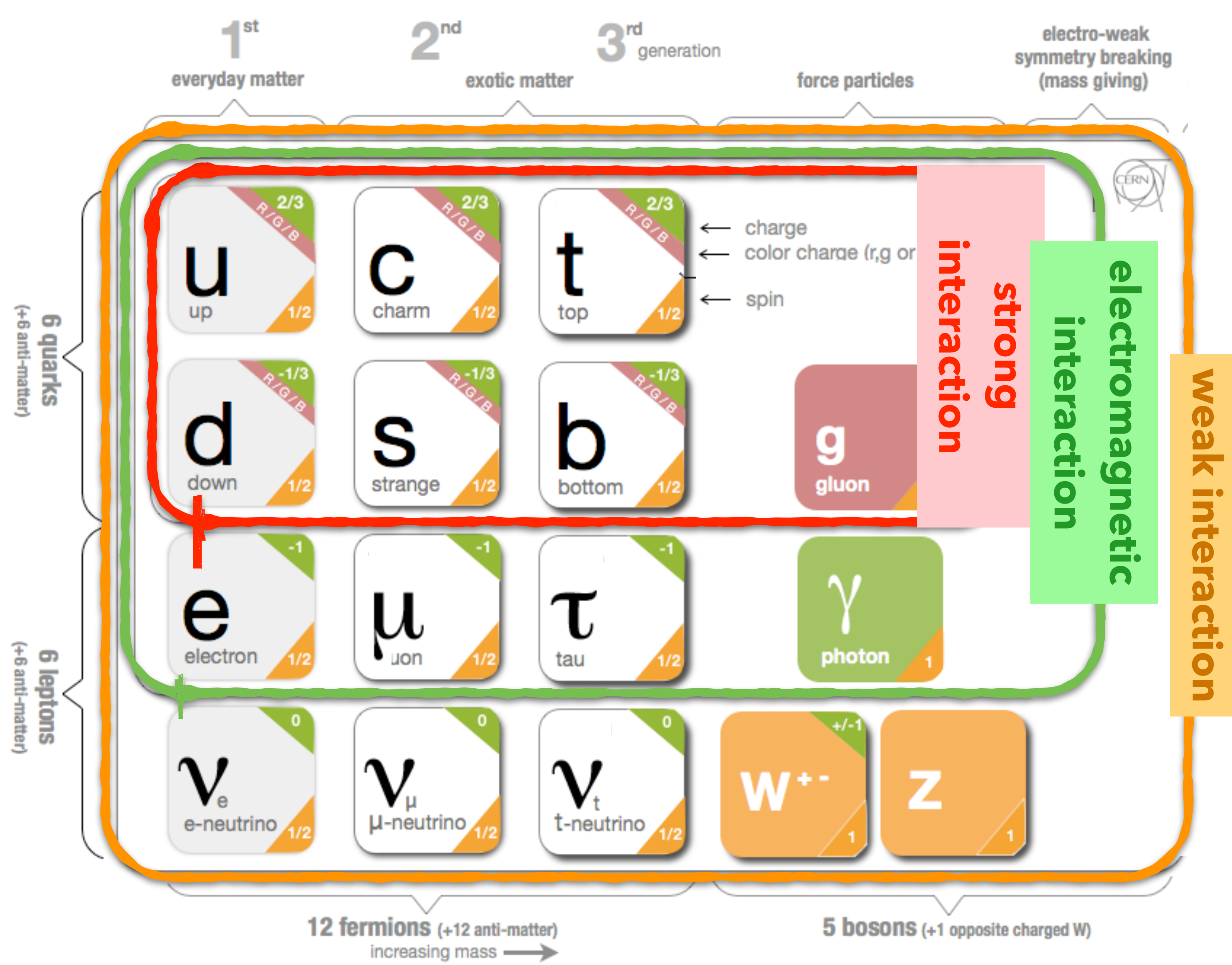
- The Standard Model
- The Giant: ATLAS
- The Heaviest: Top quark
- The Elusive: Higgs boson
- Top & H: $t\bar{t}H$ production
  - $H \rightarrow b\bar{b}$
  - $H \rightarrow WW, ZZ, \tau\tau$
  - $H \rightarrow \gamma\gamma$
- Conclusions



The Standard Model of Particle Physics

The SM provides unified picture of the electroweak (EW) and strong interactions

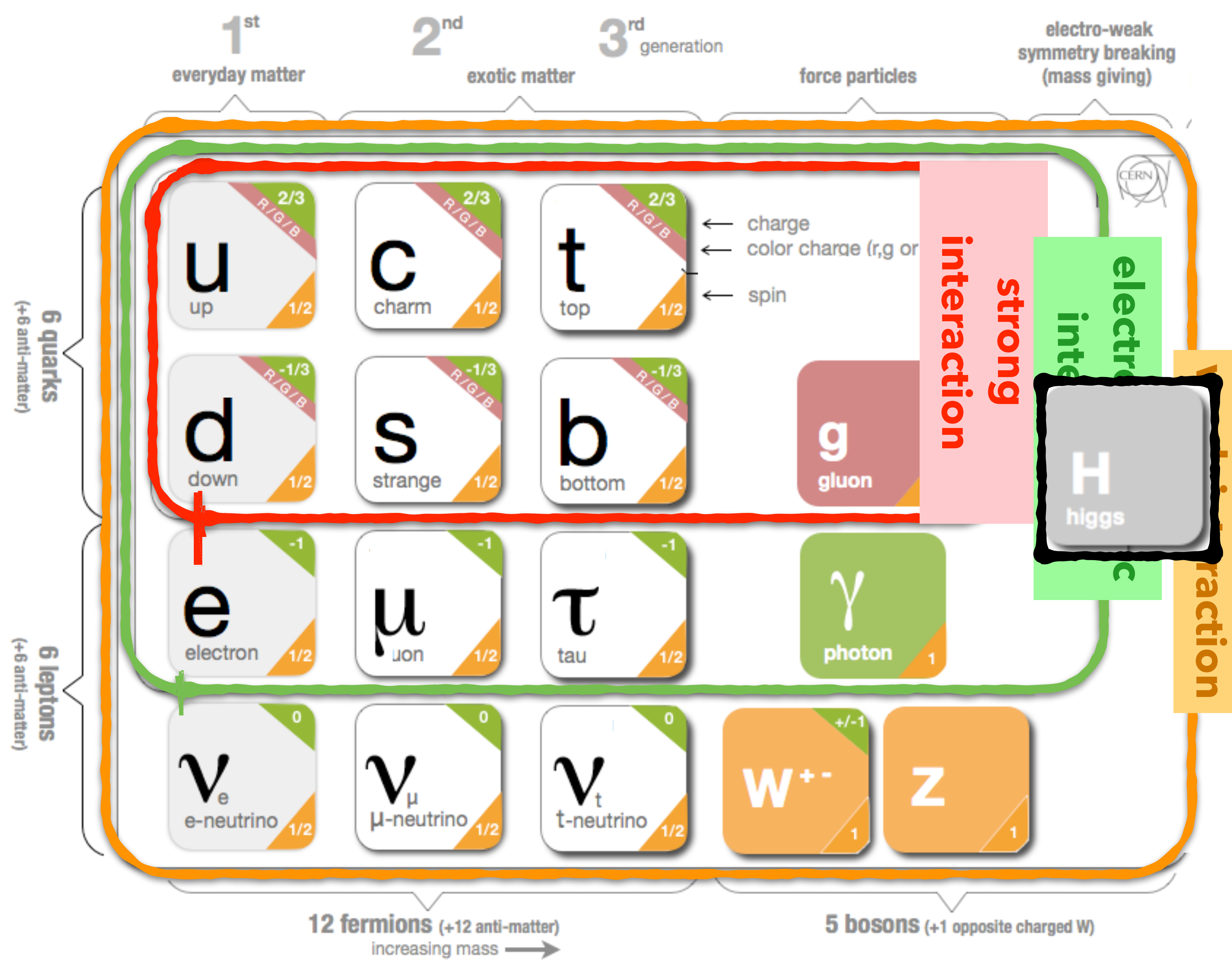
building blocks of matter: **fermions** (leptons and quarks)
force carriers: **bosons** (gluon, photon, W+-, Z)



The Standard Model of Particle Physics

The SM provides unified picture of the electroweak (EW) and strong interactions

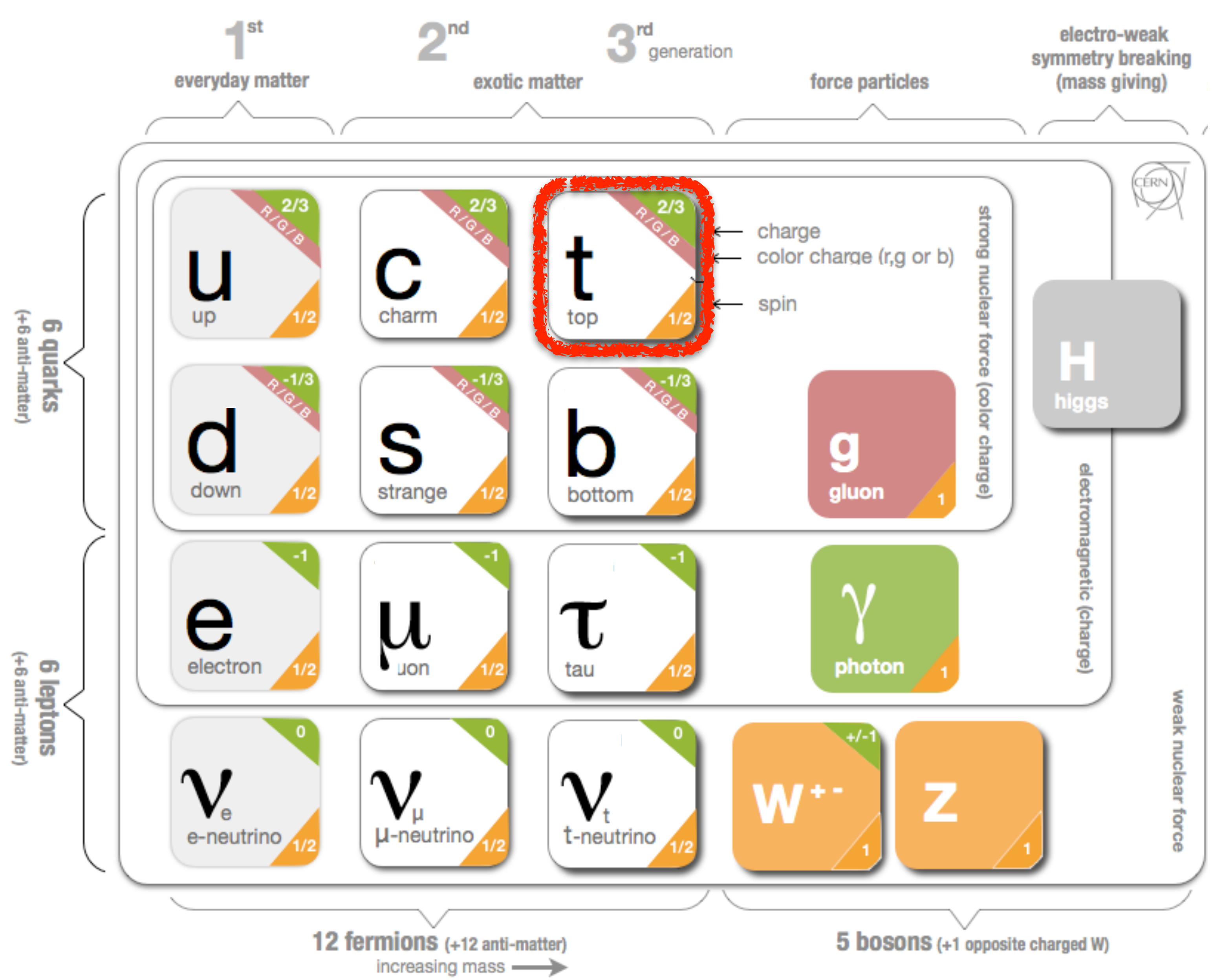
building blocks of matter: **fermions** (leptons and quarks)
force carriers: **bosons** (gluon, photon, W+-, Z)
Higgs field: added to the SM to generate the mass of EW bosons and fermions



The Standard Model of Particle Physics

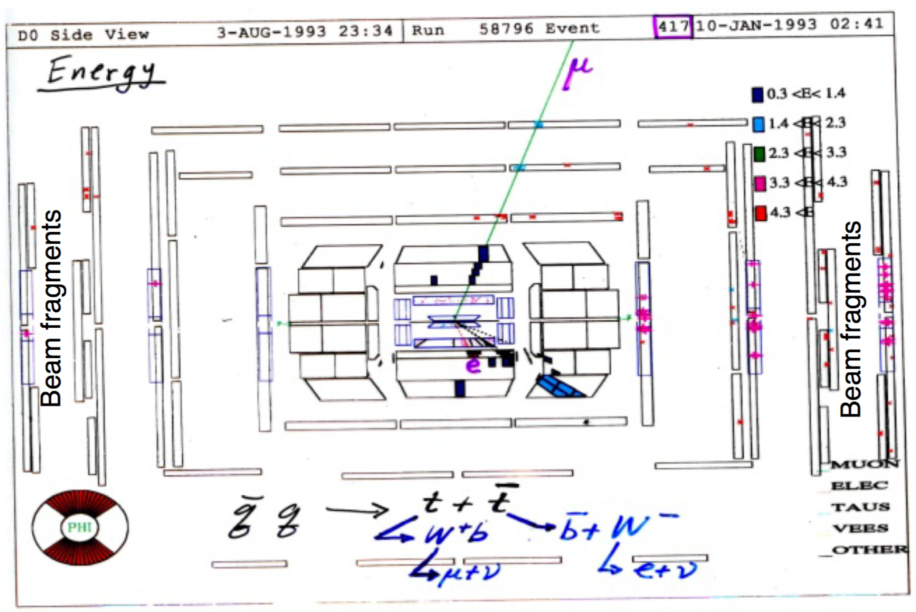
The Top Quark:

heaviest elementary particle in the SM: Yukawa coupling ≈ 1
decays before it can hadronise: study properties of a **'bare' quark**



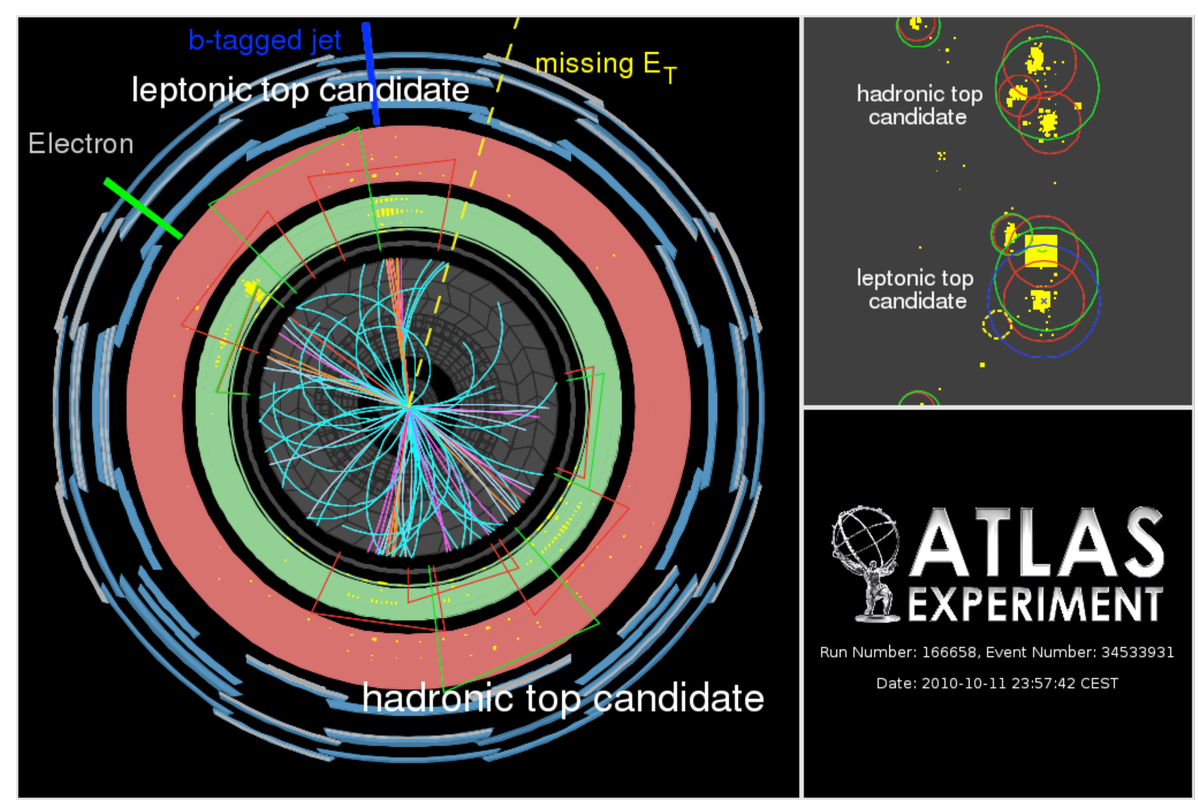
The Top Quark: > 20 Years Celebration!

Discovered in 1995 by the CDF and D0 experiments (Tevatron), completing the list of 6 fermions



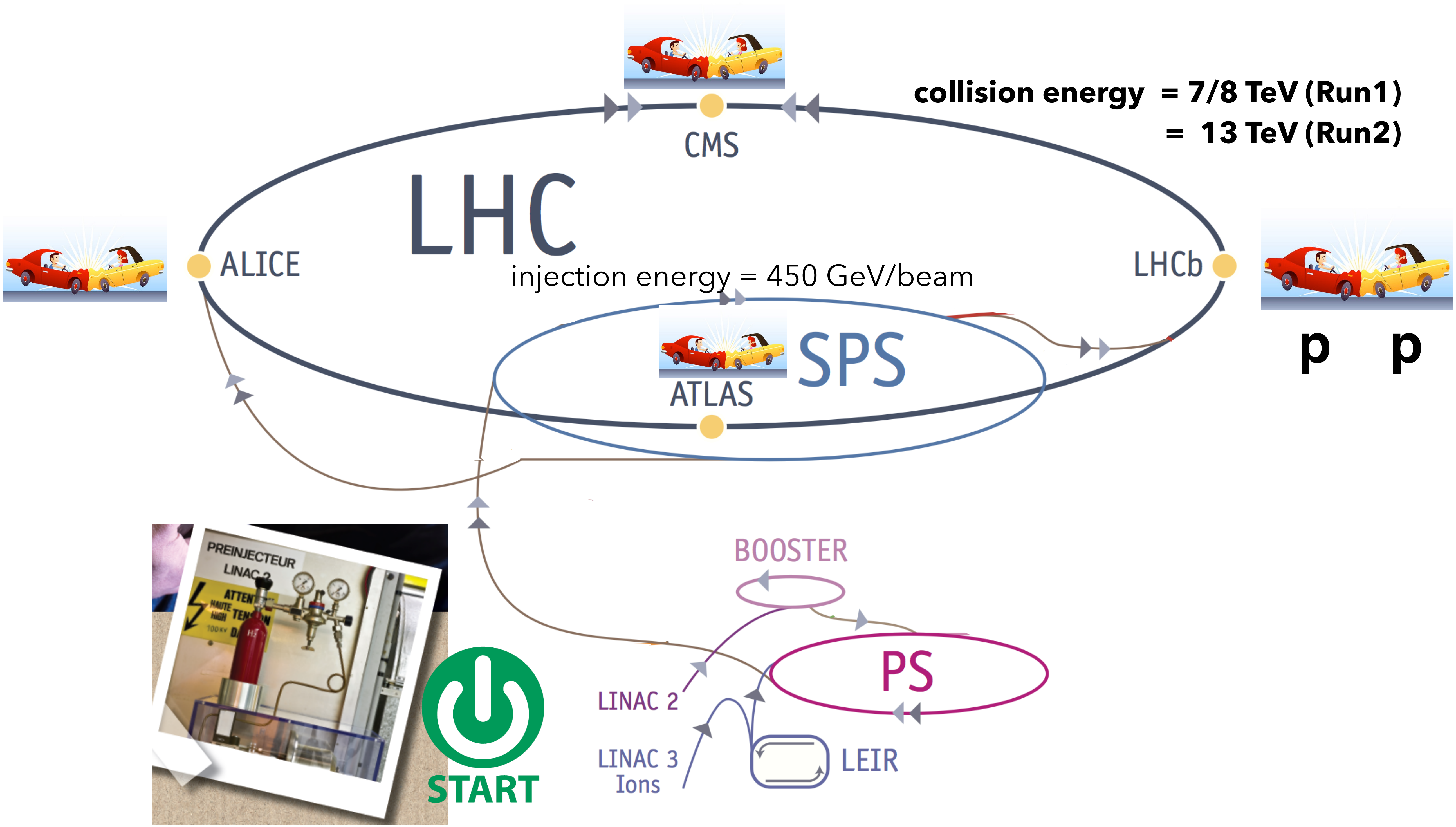
The famous Event #417: $t\bar{t}$ Candidate (1993) - D0 experiment

March 2, 1995: Joint CDF/DØ seminar announcing the top quark discovery



Last year we were celebrating 20 years since the birth of the heaviest fundamental particle… and still a hot topic in particle physics!

The Large Hadron Collider (LHC)

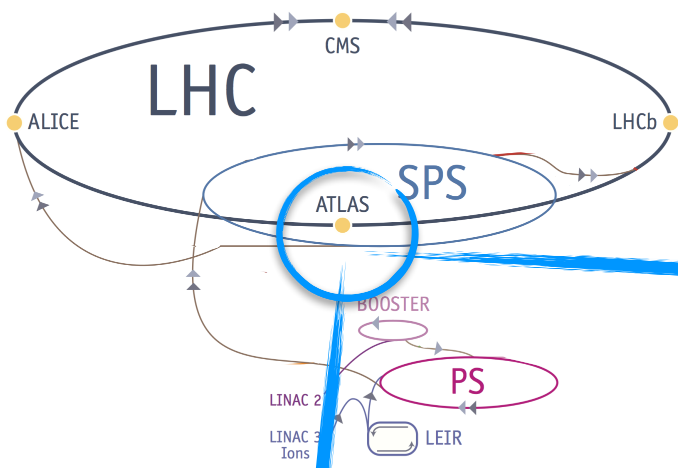


collision energy = 7/8 TeV (Run1)
= 13 TeV (Run2)

The CERN accelerator complex & the collider

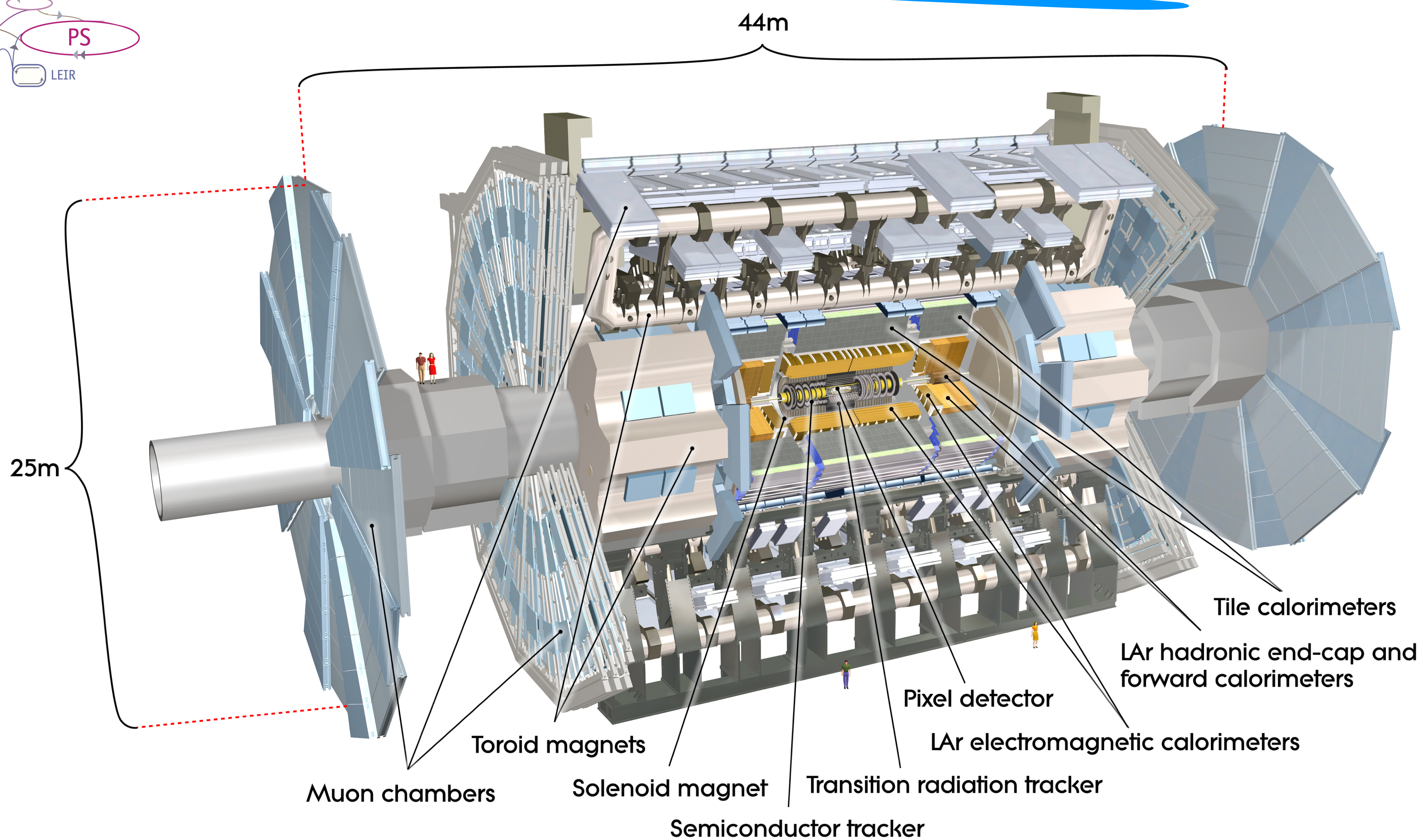
Tamara Vázquez Schröder (McGill University)

The ATLAS experiment

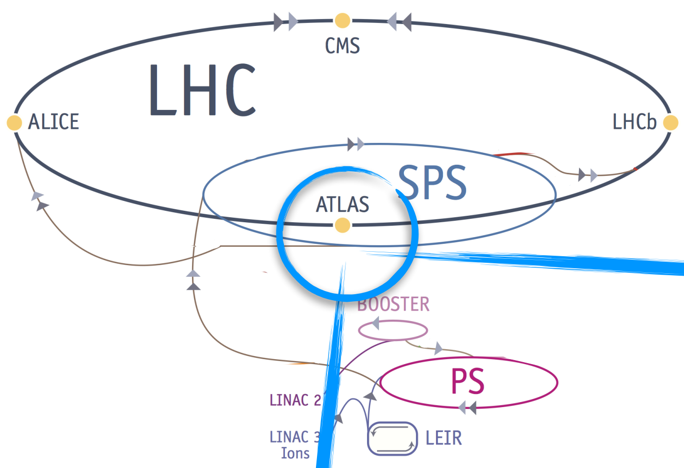


Multipurpose detector:

tracking detector + calorimeter + muon spectrometer

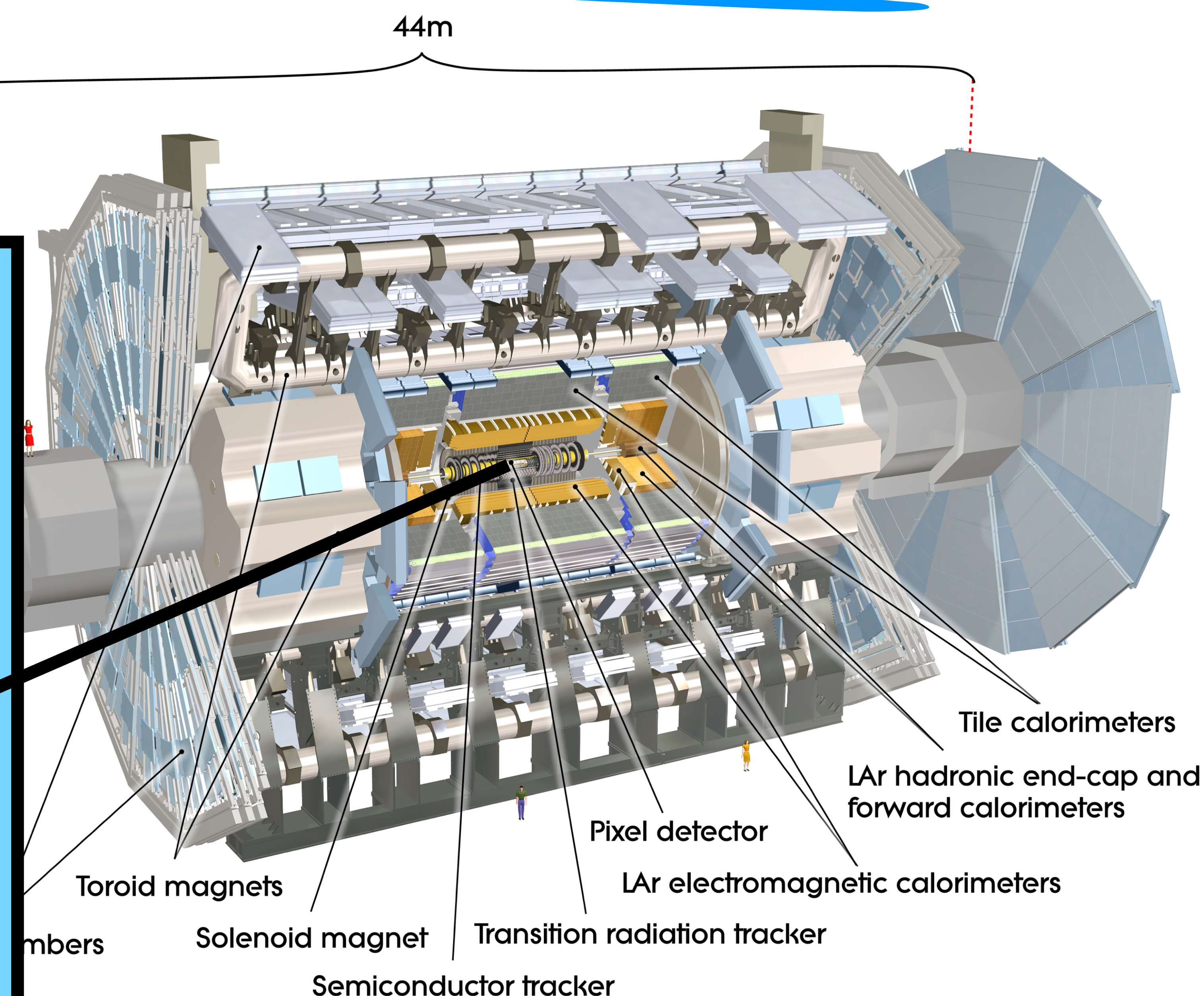


The ATLAS experiment



Multipurpose detector:

tracking detector + calorimeter + muon spectrometer

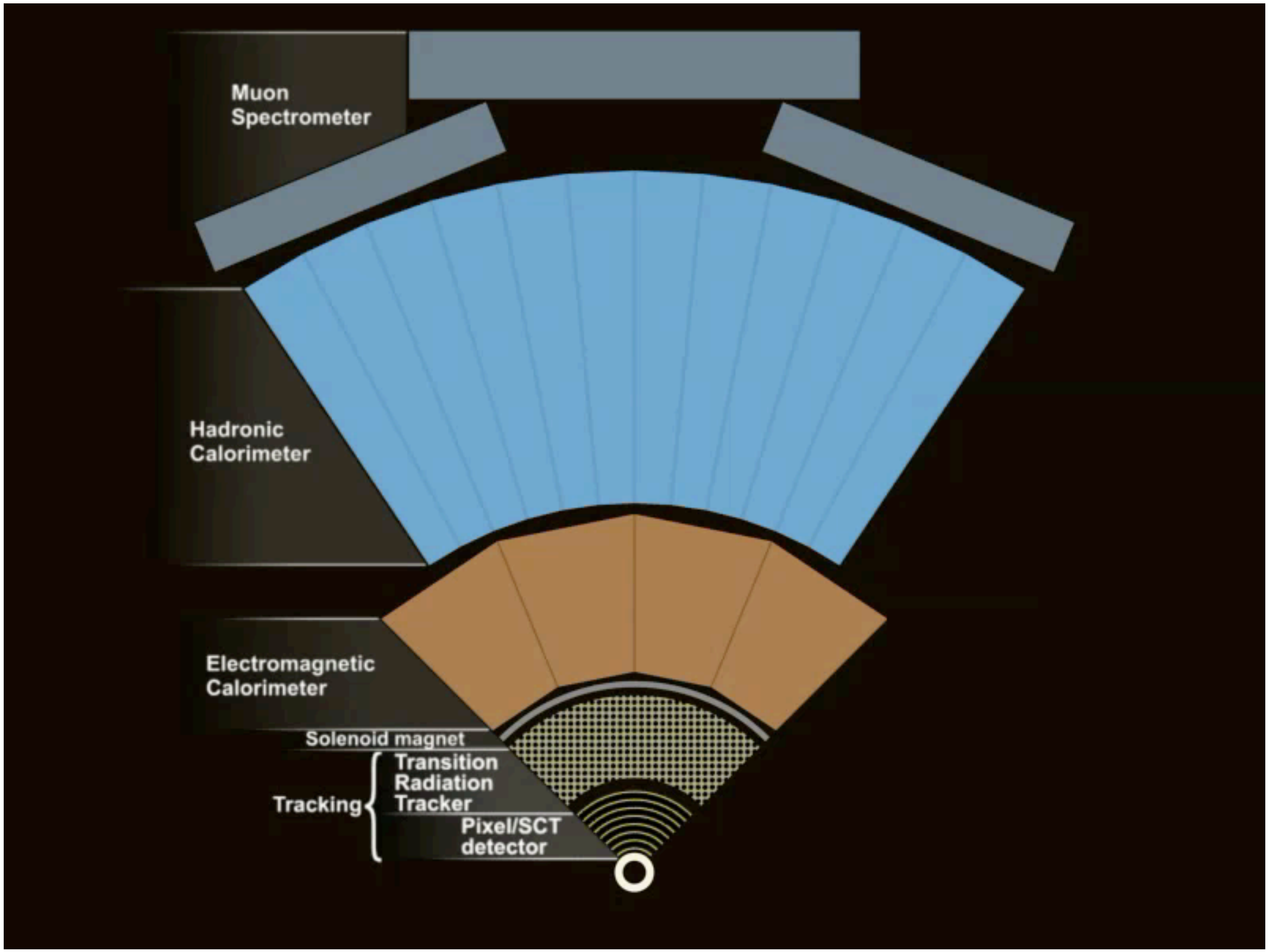


In Run2: new **Insertable B-Layer (IBL)** improved b-tagging performance!

- needed to shrink the diameter of the beam pipe and insert it into the gap between the Pixel Detector and the pipe

- faster read-out chips and new silicon sensor technologies developed to cope with higher radiation and high particle occupancy

Particle identification

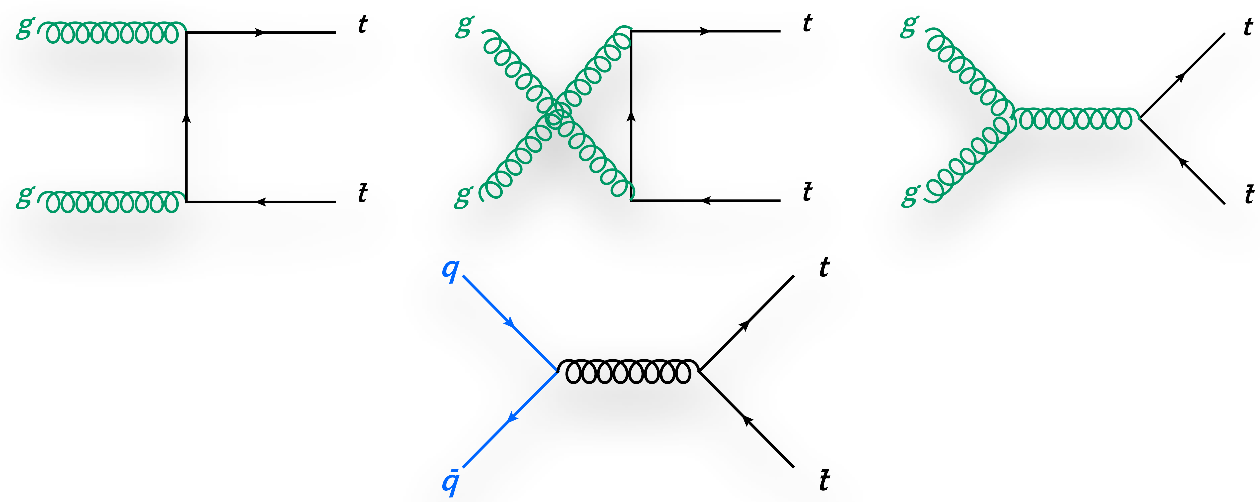


 http://www.atlas.ch/multimedia/#how-atlas-detects-particles

- Each layer, different interaction with particles, different targets
- Energy, momentum, measurements

The Top Quark: Production

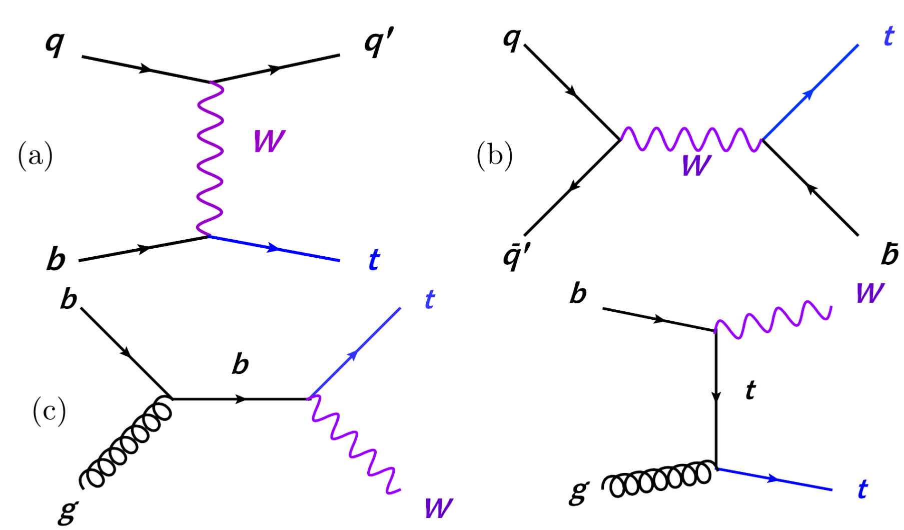
* Main production mode: **top quark pair** via the strong interaction



gg initiated
(80% LHC 7 TeV)

qq̄ initiated

* Smaller branching ratio: produced as **single top quark** via EW interaction



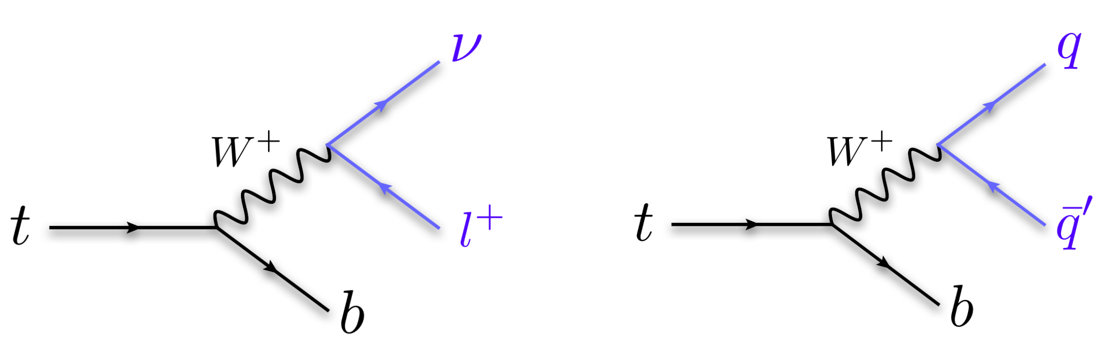
(a) t-channel

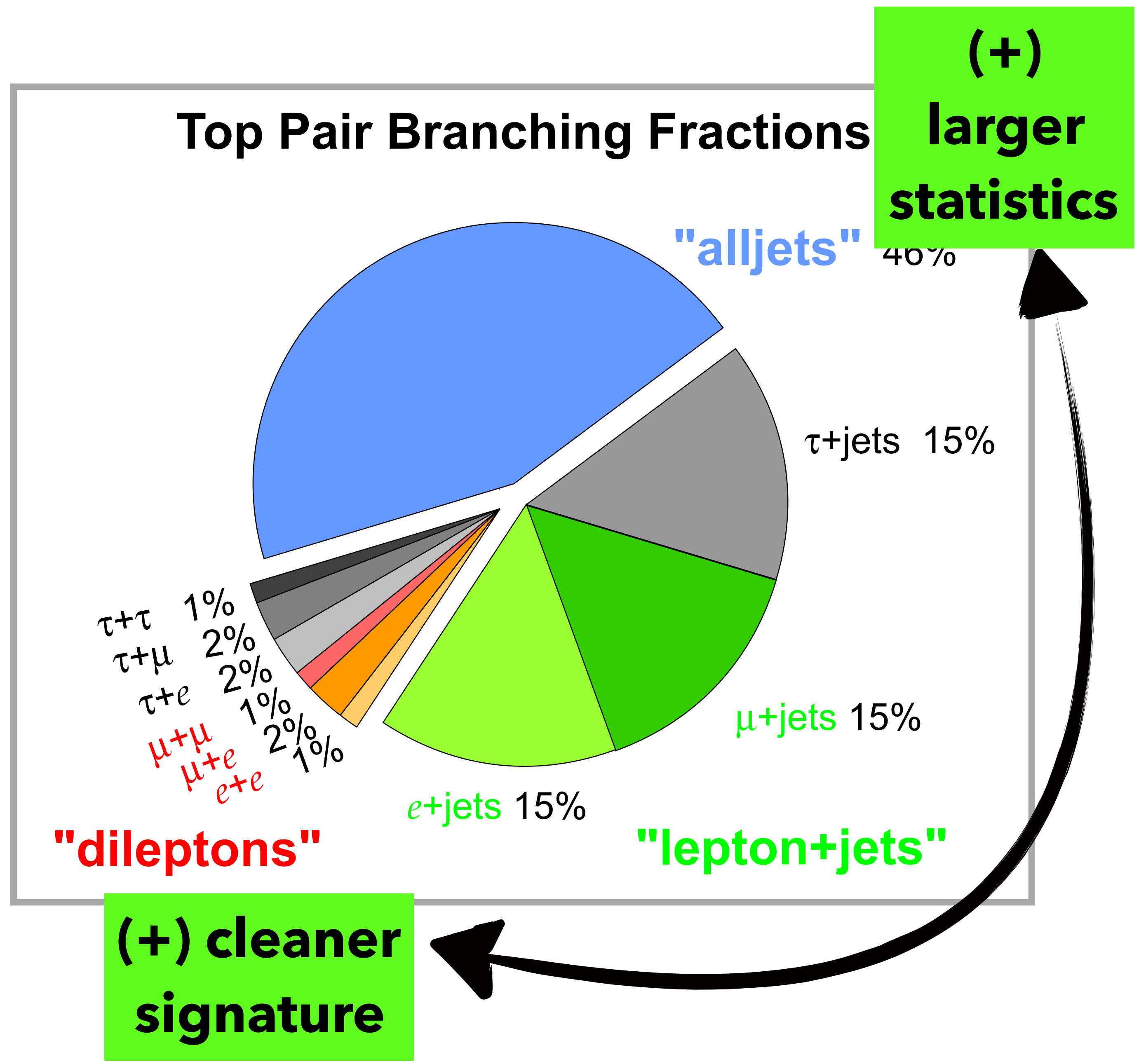
(b) s-channel

(c) Wt-channel

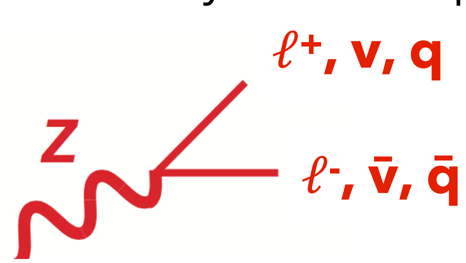
The Top Quark: Decay

- Decays into lighter particles via the EW interaction
  - Almost exclusively as t → Wb
  - W-boson decay modes: (**leptonic** or **hadronic**)



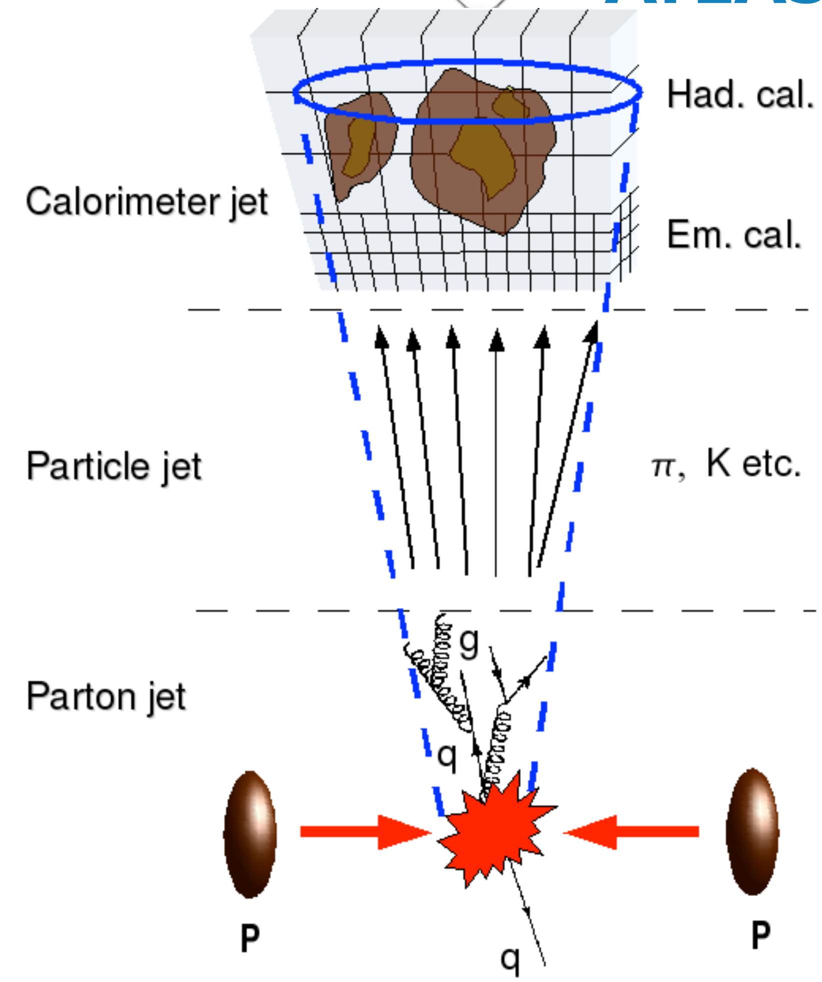


Z-boson decays for comparison

()

Experimental signatures

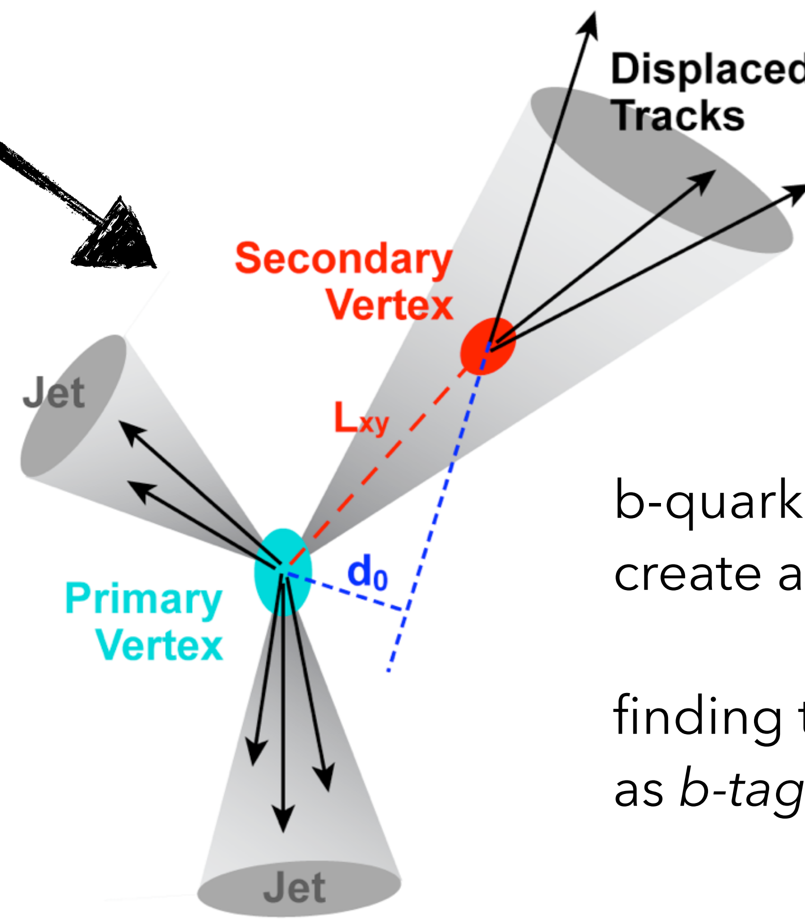
jets are a consequence
of the strong force



What to look for?

- (≥ 2, 4, 6) **jets**!
- ≥ 2 jets originating from b-quarks (**bjets**)
- **charged lepton**(s)
- **neutrino**(s) (missing transverse energy)

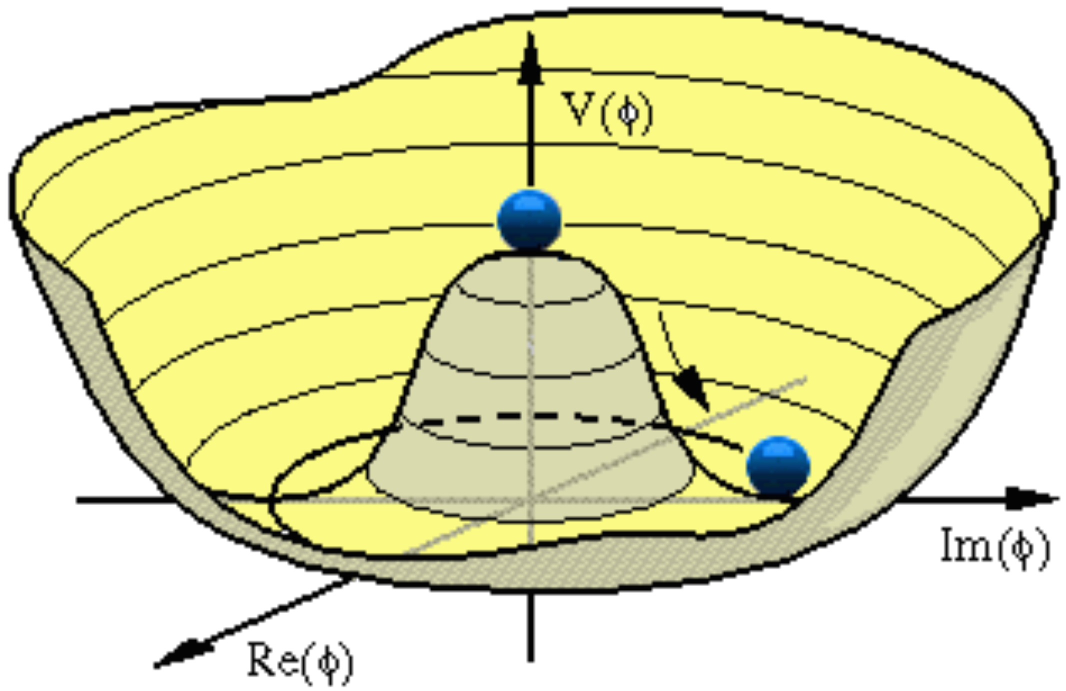
originate from b/c-quarks (heavy-flavour jets) or light quarks (light jets)



b-quarks live long enough (~ps) to create a secondary vertex at the decay

finding these jets from b-quarks is known as *b-tagging*

The Higgs Mechanism

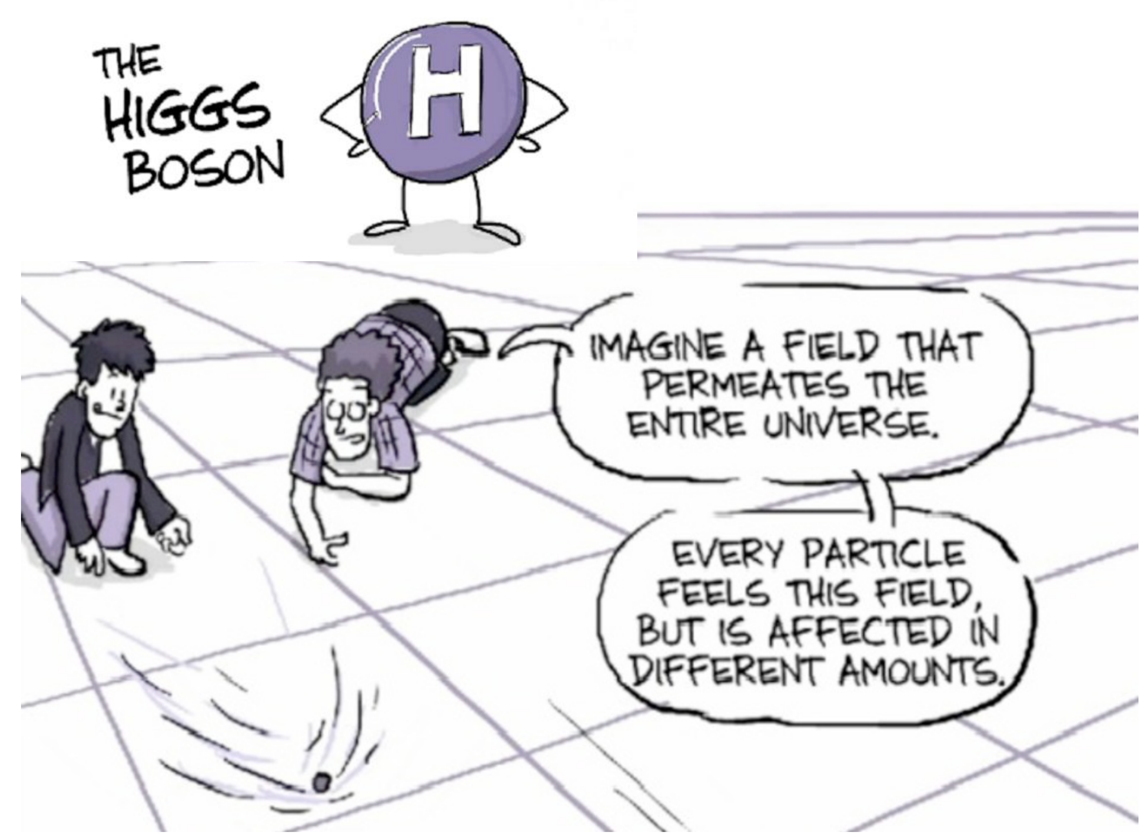


* In electroweak theory, W and Z Boson should not have mass - however we know they do!

* The Higgs Mechanism, explains the mass of the W and Z Boson, as well as the fermions, and the additional observation of a spin-0 boson

* The Higgs field couples to particles giving them mass

* The stronger the coupling, the heavier the particle (Yukawa coupling)



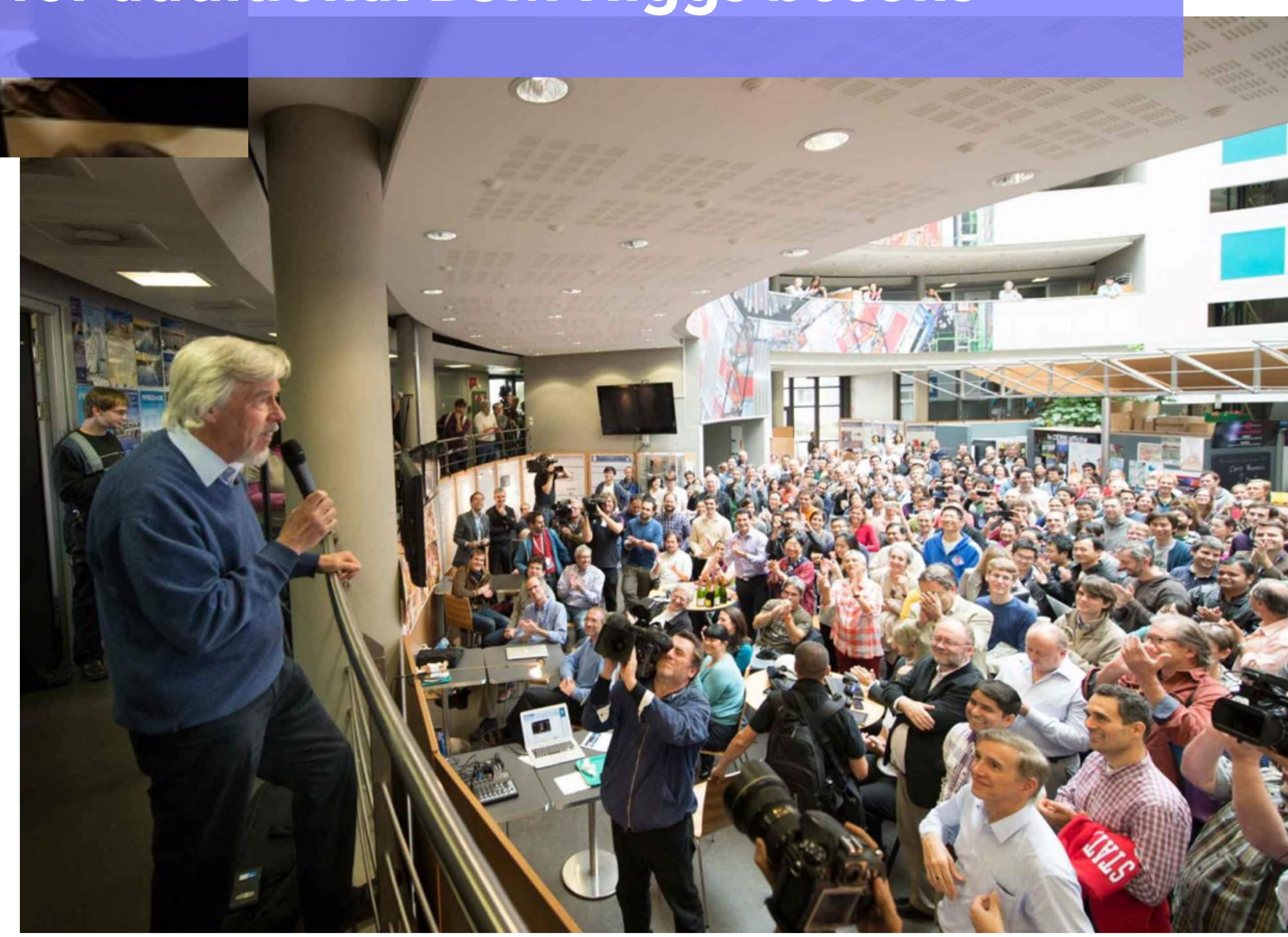
http://www.phdcomics.com/comics.php?f=1489

4th July 2012: Higgsdependence day!

The Nobel Prize in Physics 2013 was awarded jointly to François Englert and Peter W. Higgs *"for the theoretical discovery of a mechanism that contributes to our understanding of the origin of mass of subatomic particles, and which recently was confirmed through the discovery of the predicted fundamental particle, by the ATLAS and CMS experiments at CERN's Large Hadron Collider"*

4th July 2012: Higgsdependence day!

Since its discovery in 2012, focus on **precision measurements of production and decay** of the Higgs boson, and the **search for additional BSM Higgs bosons**

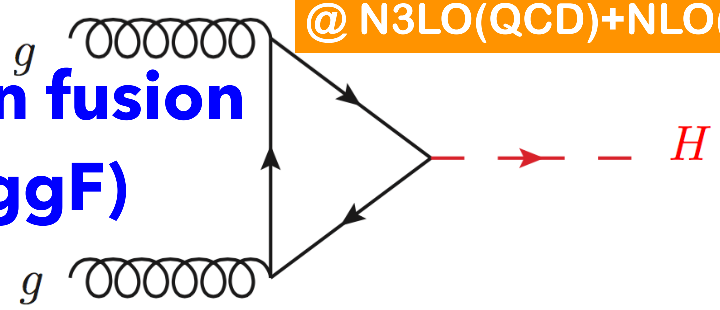


The Nobel Prize in Physics 2013 was awarded jointly to François Englert and Peter W. Higgs *"for the theoretical discovery of a mechanism that contributes to our understanding of the origin of mass of subatomic particles, and which recently was confirmed through the discovery of the predicted fundamental particle, by the ATLAS and CMS experiments at CERN's Large Hadron Collider"*

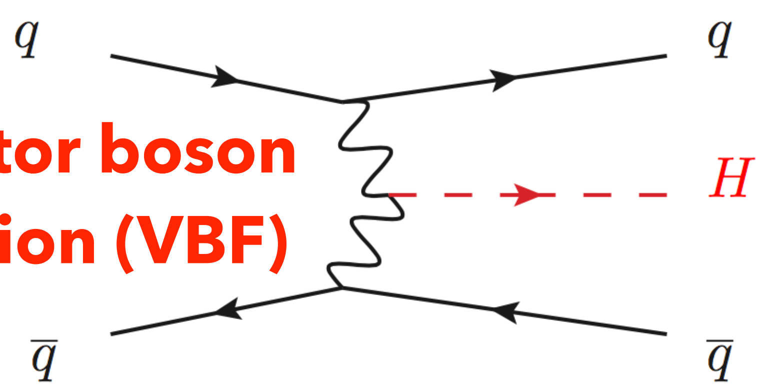
Higgs production modes

cross section calculation @ N3LO(QCD)+NLO(EW)

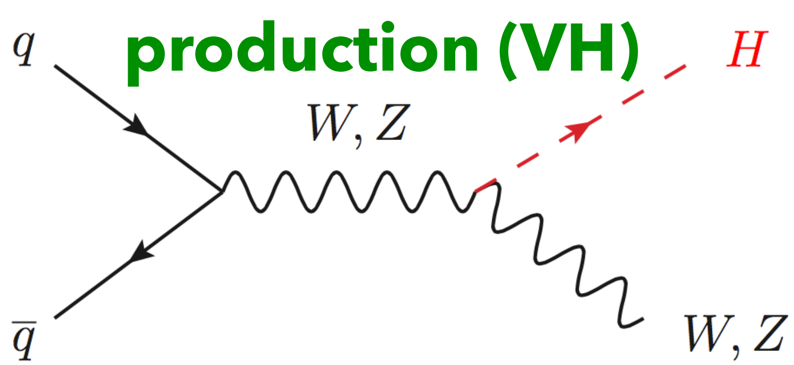
gluon fusion (ggF)



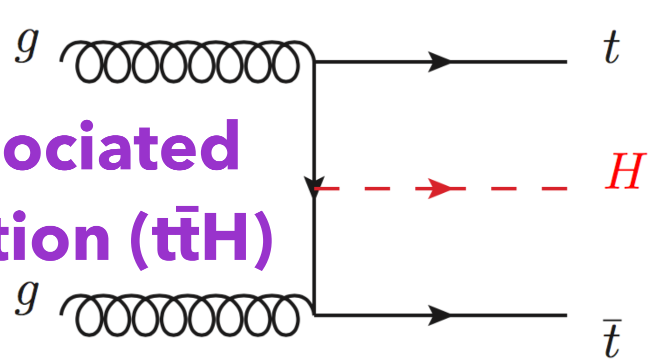
vector boson fusion (VBF)



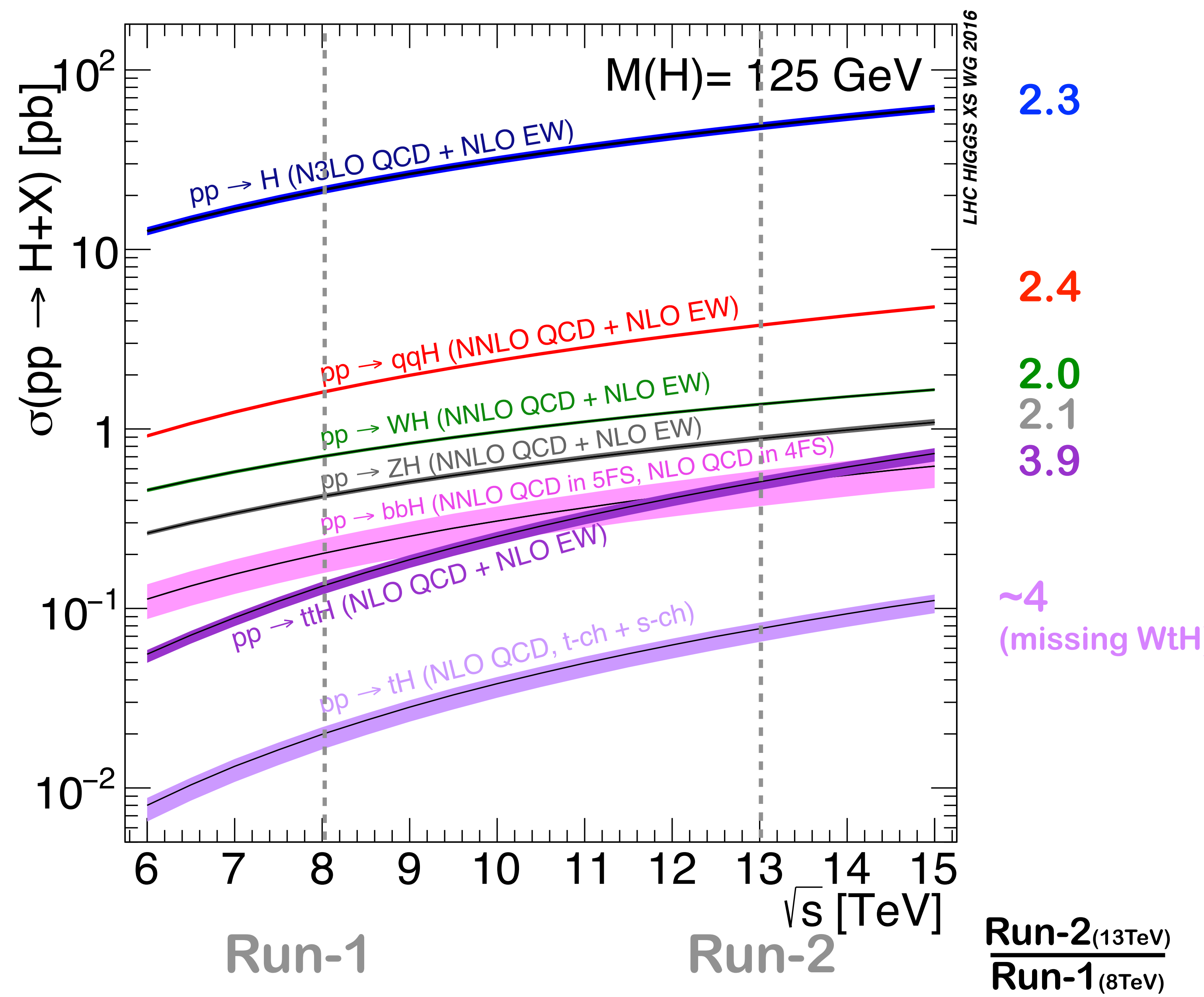
W, Z associated production (VH)



top associated production ($t\bar{t}H$)



- Gluon fusion has the largest production rate, order of magnitude higher than VBF or VH
- Large cross section increase from 8 to 13 TeV, especially for $t\bar{t}$H and tH

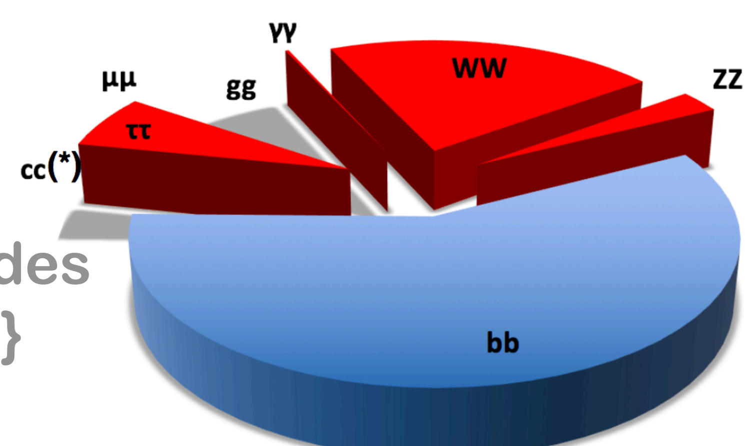


Tamara Vázquez Schröder (McGill University)

Higgs Decay Modes

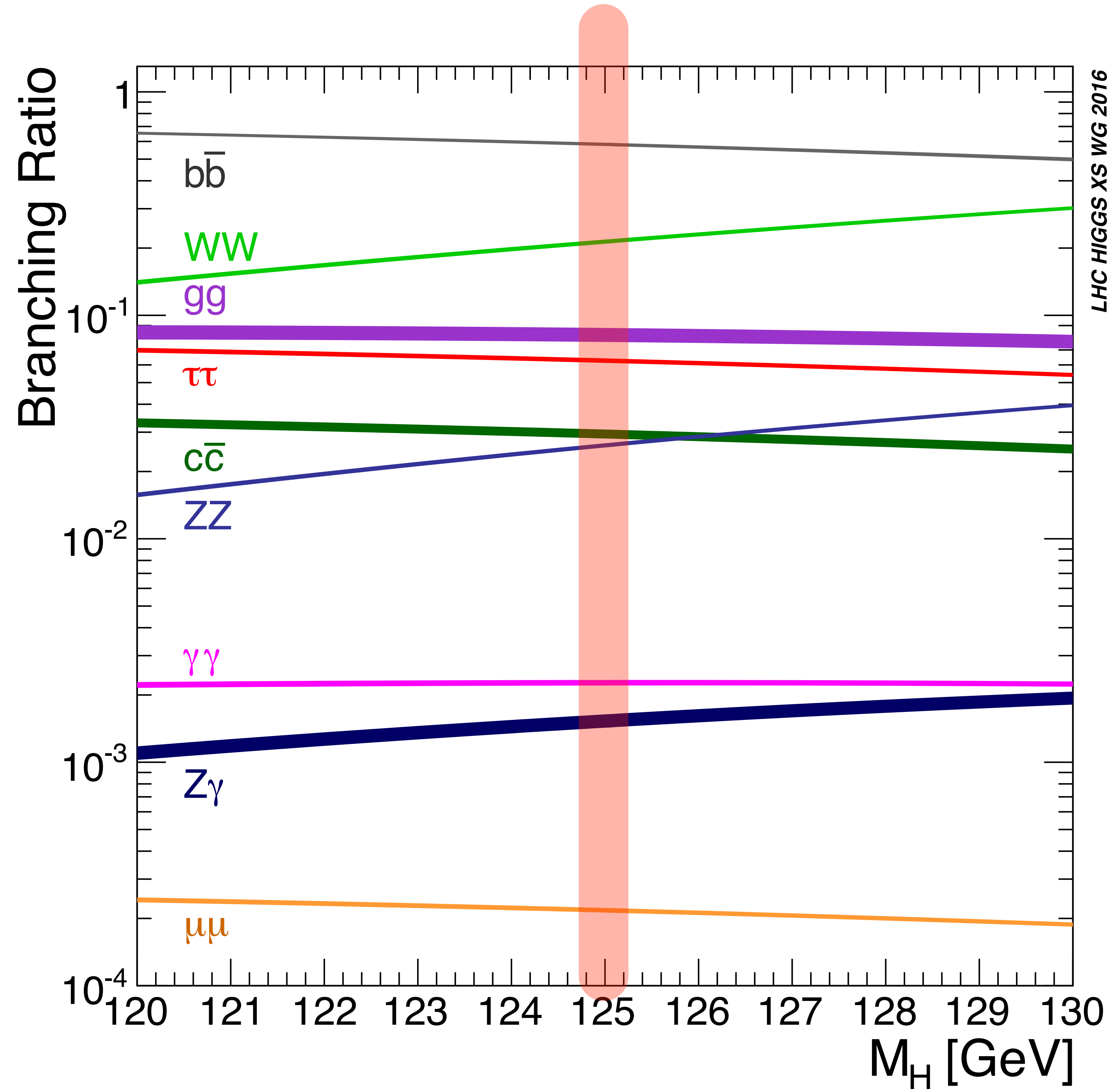
- At m_H = 125 GeV, Higgs decays to **$b\bar{b}$ is the most dominant** (~ 57 %)
- The second largest decay mechanism is H → WW (~22 %)
- Though $b\bar{b}$ decays are the most dominant, they are very **difficult to reconstruct**
- Also have broad mass resolution (contrary to γγ and ZZ(➜4ℓ))

observed decay modes (31%)



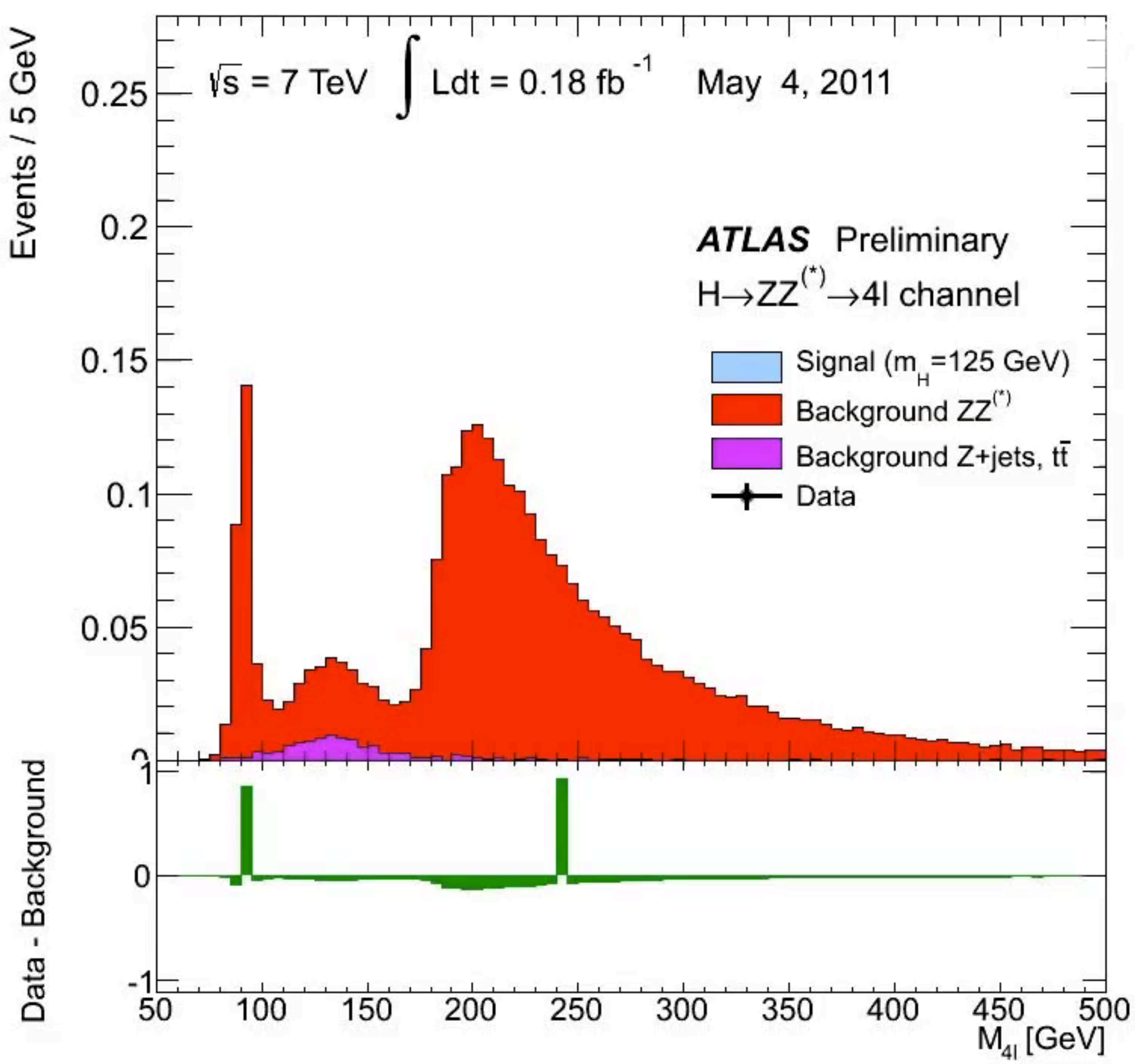
‘difficult’ modes {cc, gg, μμ} (11%) ➜ long term

evidence (58%)

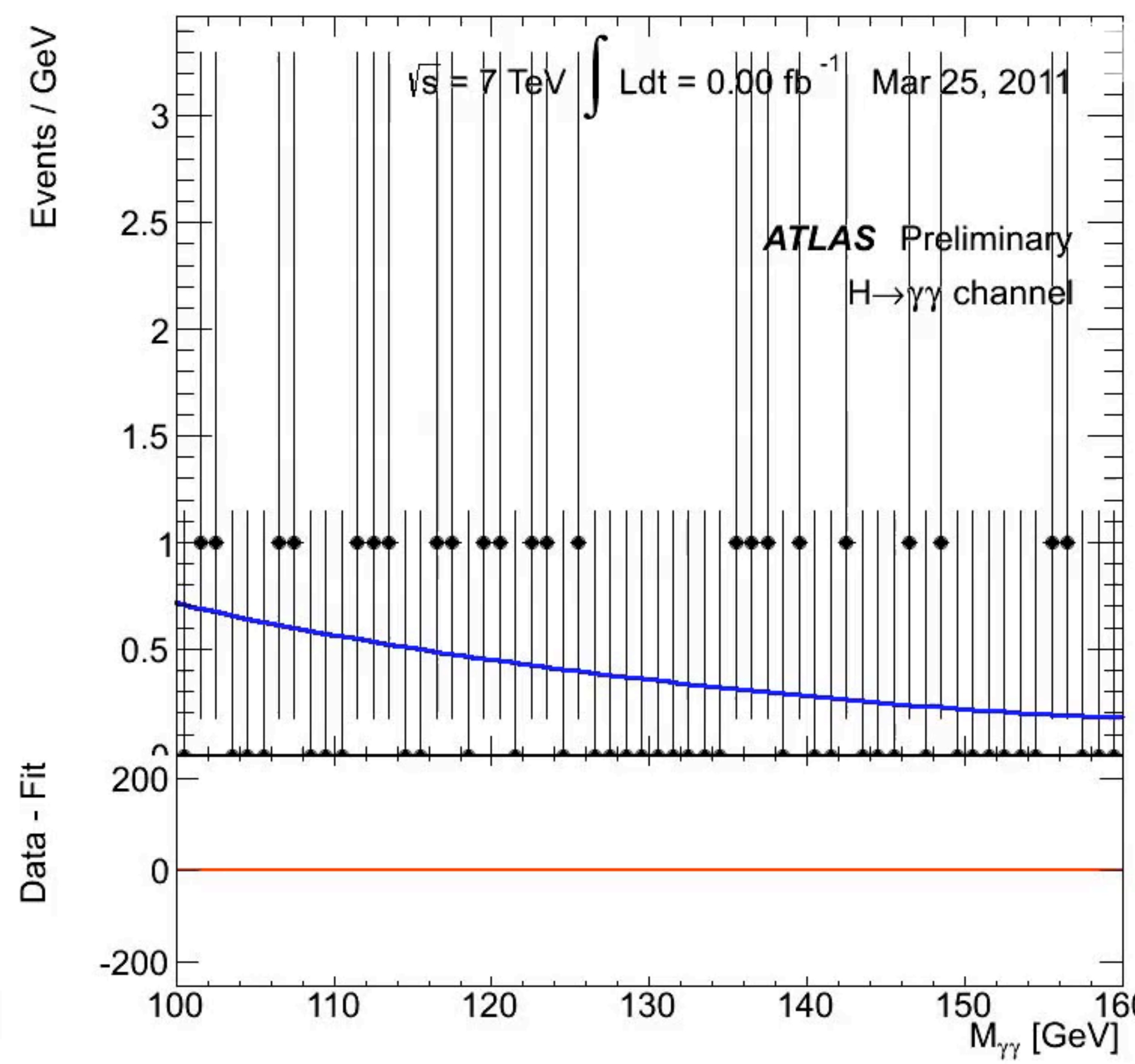


Where did we find it?

Evolution of the signal from both 7 and 8 TeV data



https://twiki.cern.ch/twiki/pub/AtlasPublic/HiggsPublicResults//4l-FloatingScale-NoMuProf2.gif



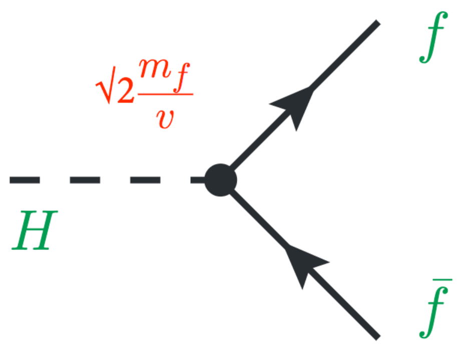
https://twiki.cern.ch/twiki/pub/AtlasPublic/HiggsPublicResults/Hgg-FloatingScale-Short2.gif

Higgs decay to ZZ → 4ℓ *(Golden Channel)*

Higgs decay to γγ

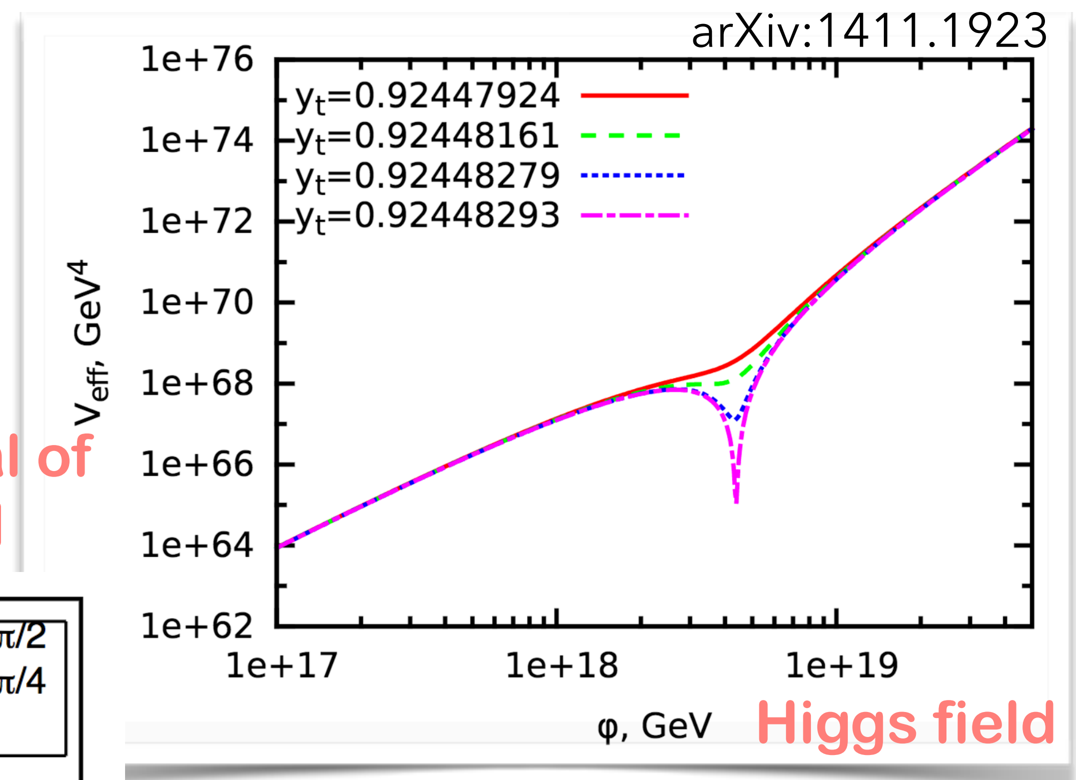
Top Yukawa coupling… why should we care?

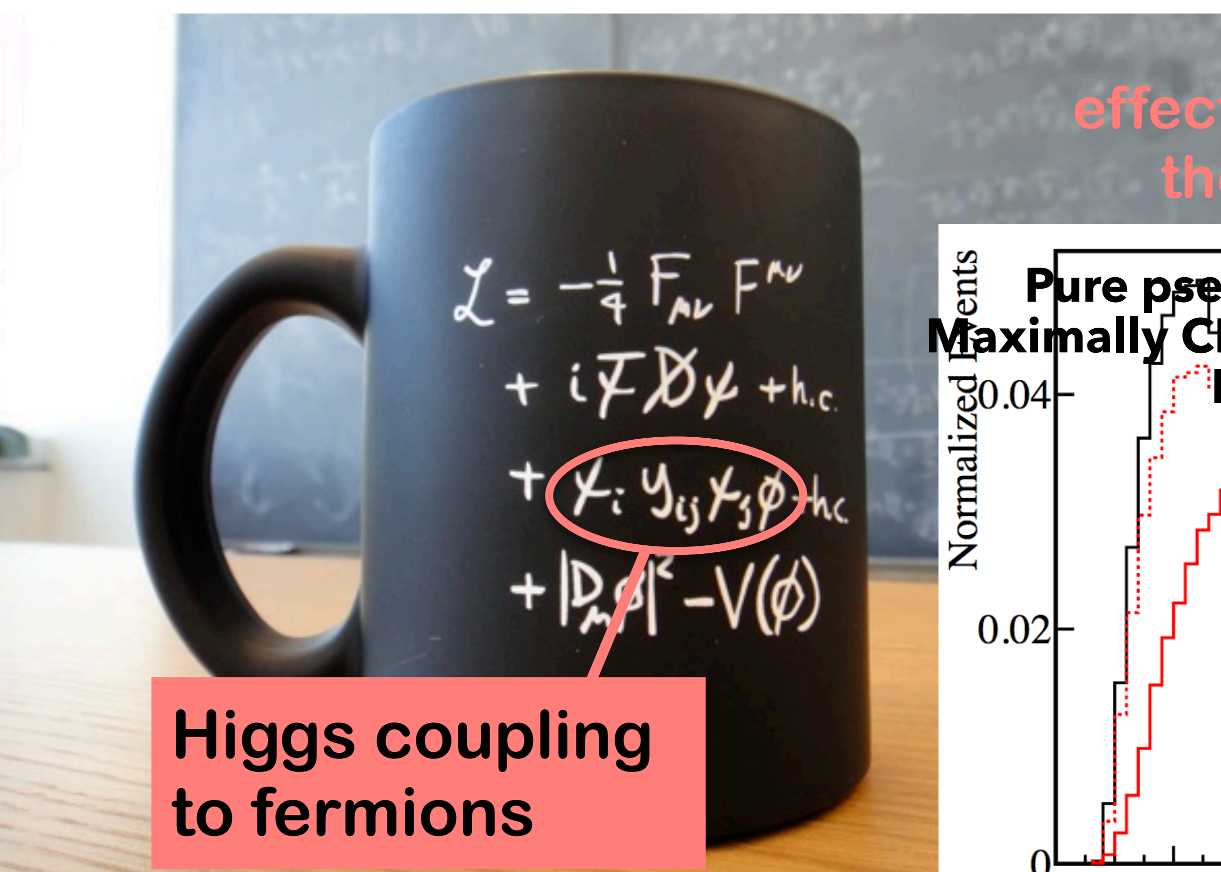
Top quark is the heaviest fermion in the SM ➜ Largest Yukawa coupling

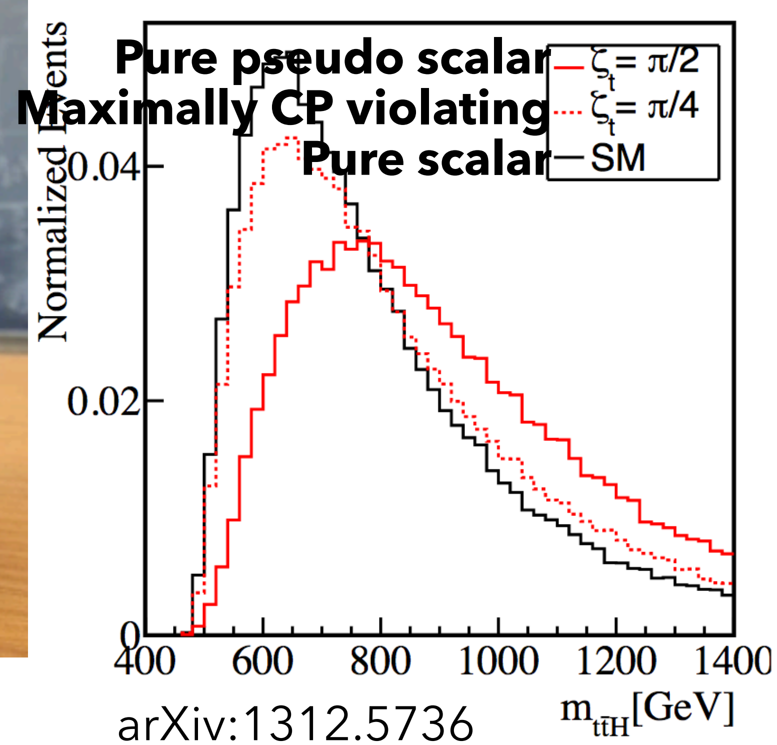


- The only fermion with such a natural coupling
- Does this point to a special role in electroweak symmetry breaking or beyond the SM physics?
- Top quark Yukawa coupling tells us about the stability of Universe and the required energy scale for new physics

Is the Universe stable or only metastable?



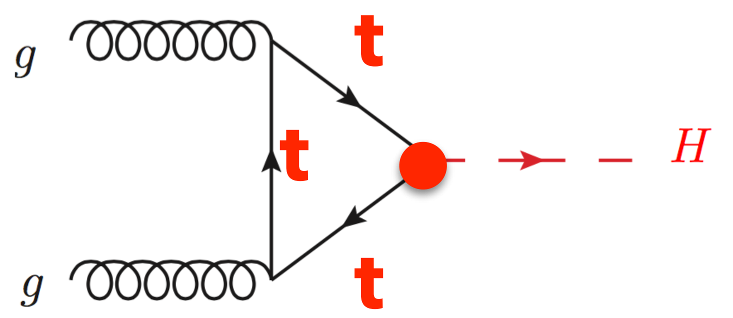


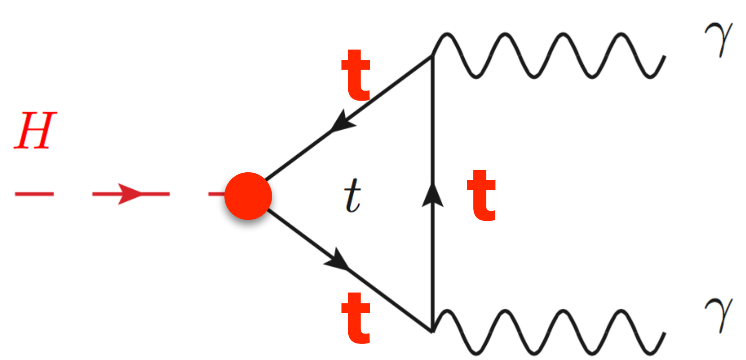


What is the CP nature of the Higgs boson?

- A CP admixture is still allowed
- Maybe $t\bar{t}H$ production can help us disentangle the BSM component

Top & Higgs

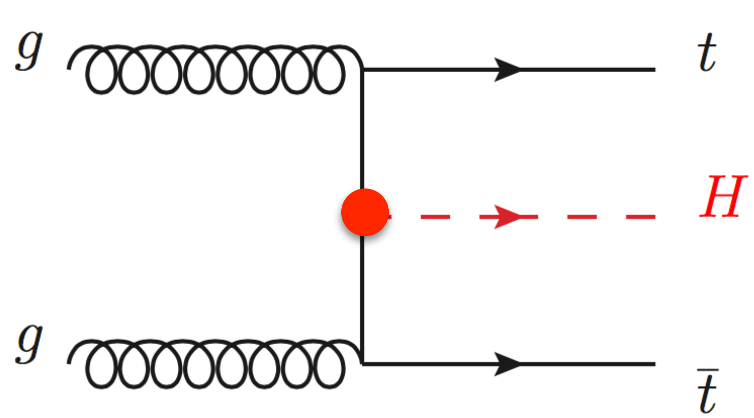


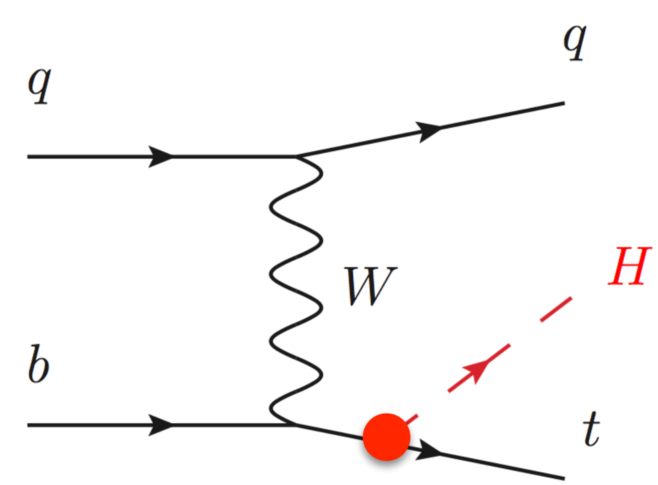


indirect top Yukawa coupling constraints from gluon fusion production and γγ decay…

… assuming no additional heavy particles which could couple to the Higgs boson!

direct top Yukawa coupling measurement only possible at the LHC via $t\bar{t}H$ and tH

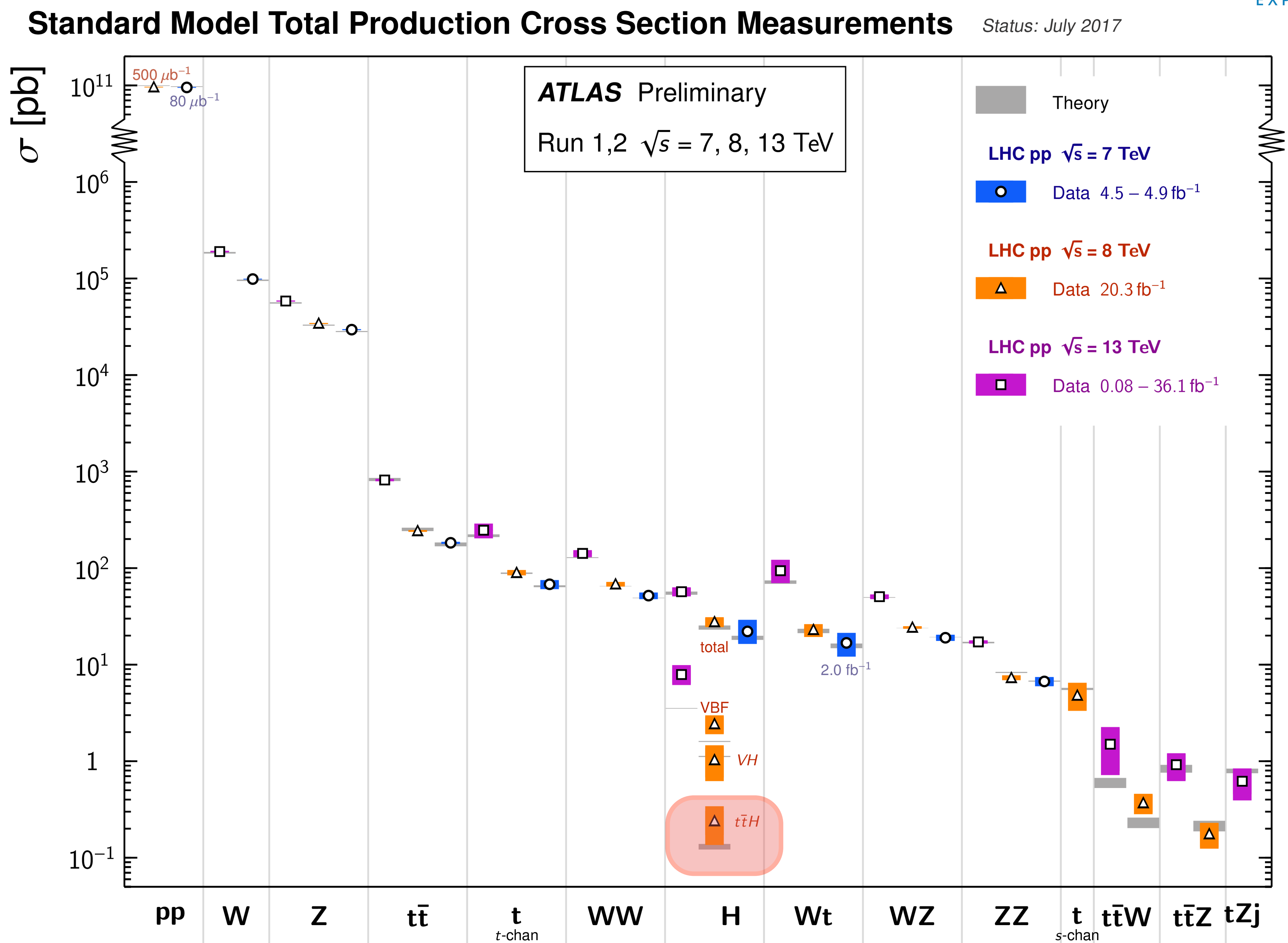




Similar signature is visible in SUSY searches, VLQ, black holes or heavy charged Higgs

If such new physics scenarios exist, will see significant deviations from SM prediction

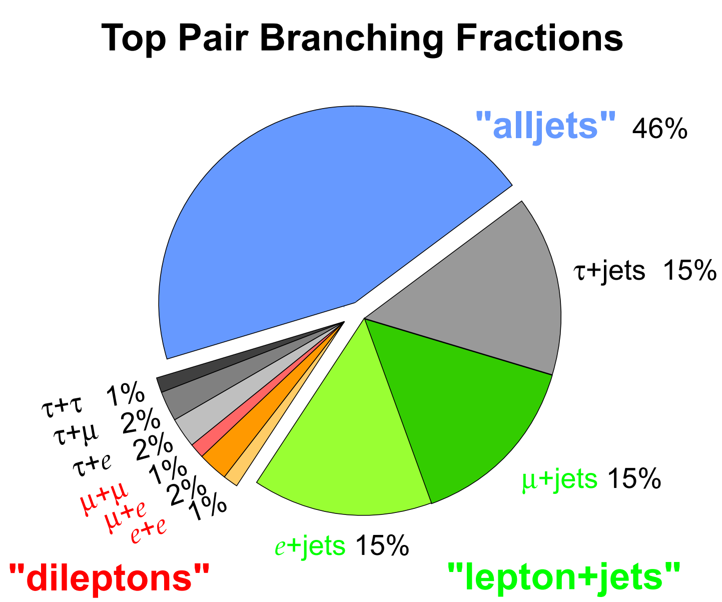
Measure one of the **tiniest**!



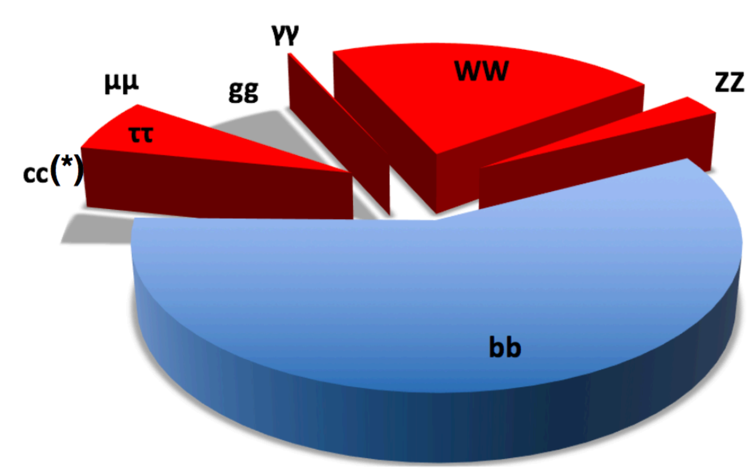
Where to look for $t\bar{t}H$ production?

- $t\bar{t}H$ production (~500 fb @ 13TeV) is:
  - **two orders** of magnitude smaller than ggF Higgs production
  - **three orders** of magnitude smaller than $t\bar{t}$ production

- Look for $t\bar{t}H$ in final states with distinctive signatures and features
  - Combination of top quark x Higgs boson decay modes



x



= ...

$t\bar{t}H$ analysis channels

ttH (H➔bb)	ttH (H➔WW, ττ, ZZ) 'multilepton'	ttH (H➔γγ, ZZ(➔4ℓ))

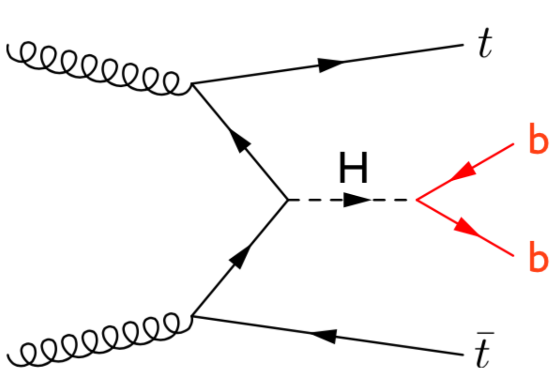
Low signal/background (need MVA) ⟷ Clear peak (clean bump hunt)

Large branching ratio (yields) ⟷ Small branching ratio

Difficult background modelling ⟷ Simple background

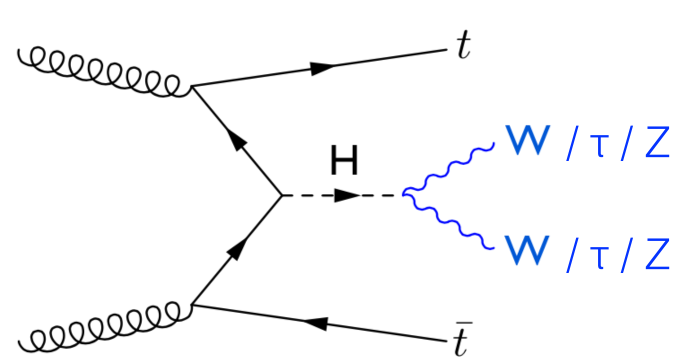
ttH (H➔bb)

- large irreducible $t\bar{t}$+jets (HF) background
- final states with **multiple b-jets**



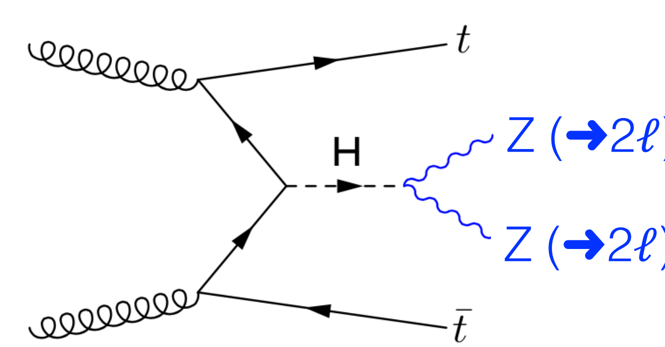
ttH (H➔WW, ττ, ZZ) 'multilepton'

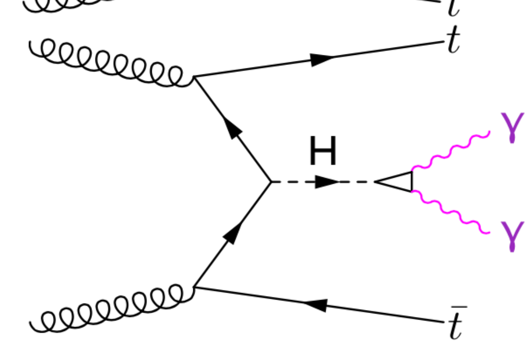
- leptonic decays of W / Z bosons and tau decays can give distinct **multilepton** signatures
- main background from $t\bar{t}$Z/W and non-prompt leptons



ttH (H➔γγ, ZZ(➔4ℓ))

- **resonant** channels





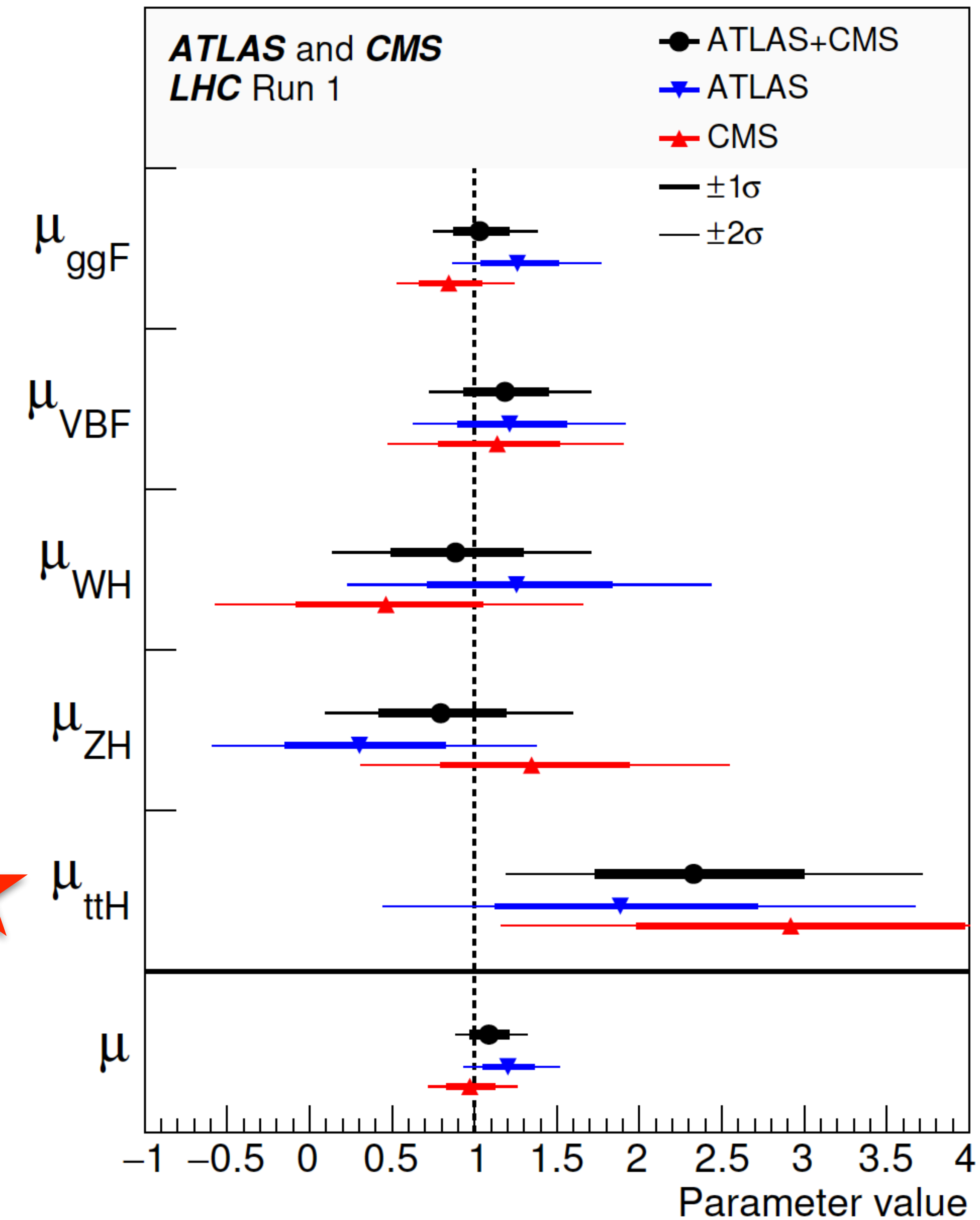
(ZZ➔4ℓ at m_H selection vetoed)

motivation ⟵ challenge

$t\bar{t}H$ **interest**: from Run-1 to Run-2

JHEP08(2016)045

- Run-1 ATLAS+CMS Higgs combination:
  - $t\bar{t}H$ significance of 4.4 σ (2.0 σ expected)
- Excess in both ATLAS and CMS $\mu_{t\bar{t}H} = \sigma/\sigma_{SM}$
  - Originating from $t\bar{t}H$ multilepton analyses



- Big leap **(x4)** for $t\bar{t}H$ SM cross section from 8 to 13 TeV (*) and **high statistics** of top quark samples collected by the LHC make this SM search **extremely interesting** to be studied in Run-2!

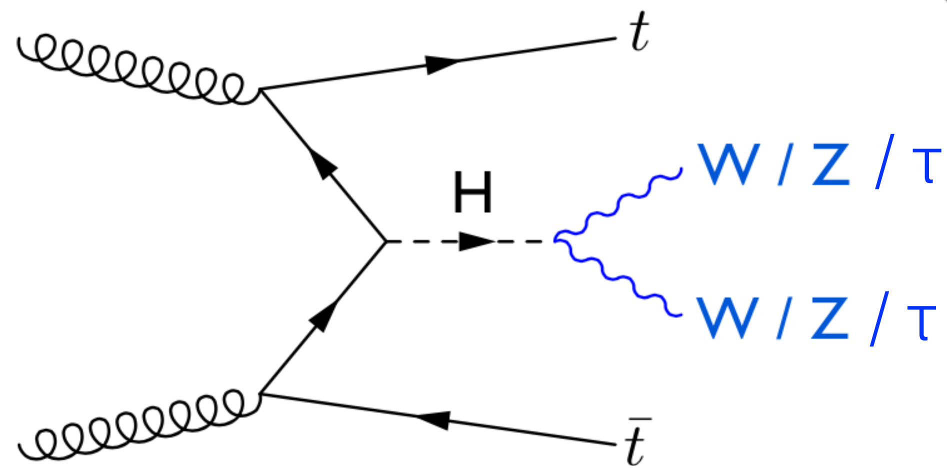
(*) Other background contributions cross section do not increase as much, but different kinematics at higher energies!

Most recent $t\bar{t}H$ results

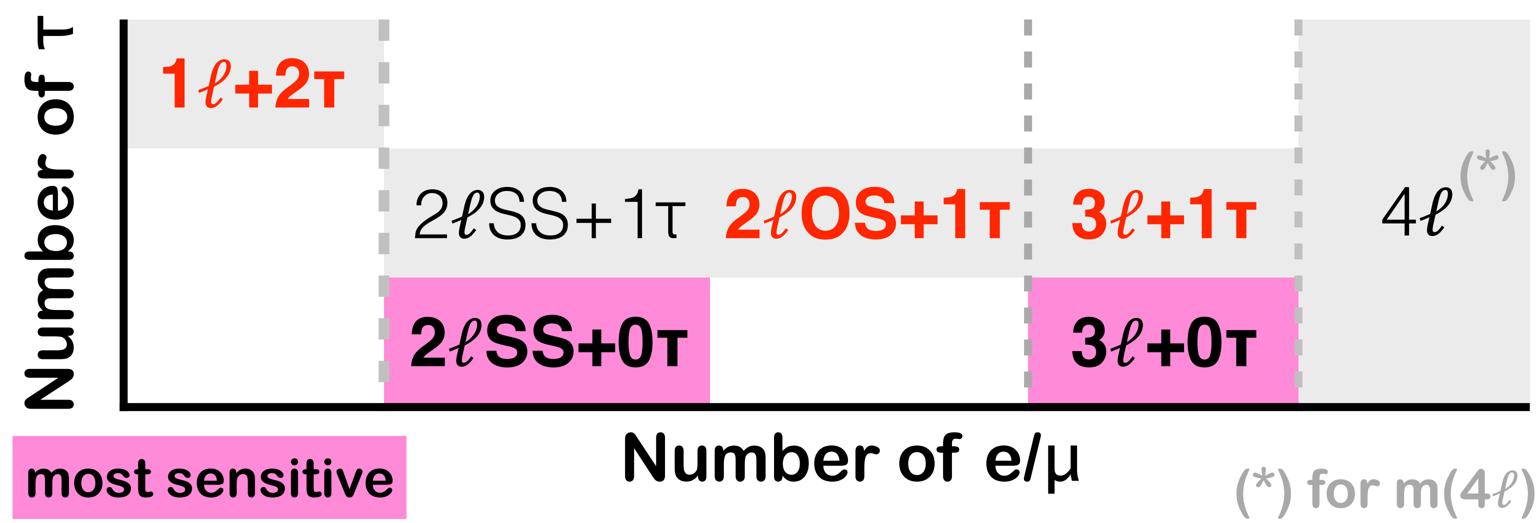
2015+2016 data [~36 fb^{-1}]

	ATLAS EXPERIMENT	CMS
ttH multilepton ($H \to WW/\tau\tau/ZZ$)	★ arXiv:1712.08891 submitted to PRD (including combination)	CMS-HIG-17-018 $\mu_{ttH} = 1.23^{+0.45}_{-0.43}$
ttH(bb)	★ arXiv:1712.08895 submitted to PRD (leptonic)	CMS-HIG-17-026 (leptonic) $\mu_{ttH} = 0.72 \pm 0.45$ CMS-HIG-17-022 (all-hadronic) $\mu_{ttH} = 0.9 \pm 1.5$
ttH($ZZ \to 4\ell$)	arXiv:1712.02304 submitted to JHEP $\mu_{ttH} < 7.1$	arXiv:1706.09936 $\mu_{ttH} < 1.18$
ttH($\gamma\gamma$)	ATLAS-CONF-2017-045 1.0σ (exp: 1.8σ) $\mu_{ttH} = 0.5 \pm 0.6$	CMS-PAS-HIG-16-040 3.3σ (exp: 1.5σ) $\mu_{ttH} = 2.2^{+0.9}_{-0.8}$
ATLAS+CMS Run1 combination	JHEP 1608 (2016) 045 4.4σ (exp: 2.0σ) $\mu_{ttH} = 2.3^{+0.7}_{-0.6}$	

$t\bar{t}H$ (multileptons): analysis strategy



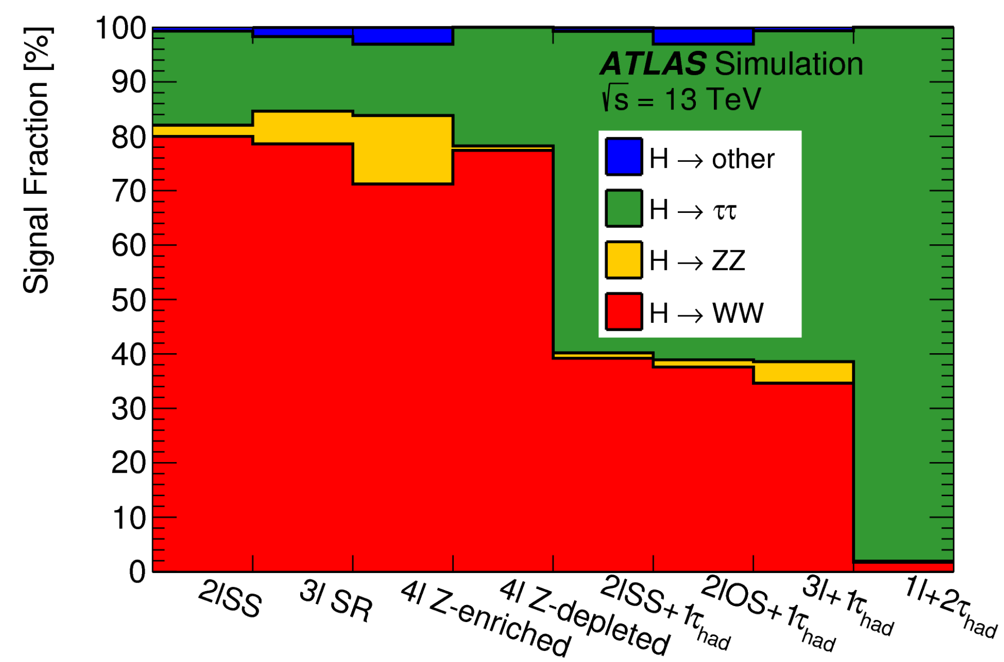
- **Target**: $t\bar{t}H$ with
  - $H \to WW/ZZ/\tau\tau \to \geq 1\ell$
  - $t\bar{t} \to (\ell$+jets, dilepton)



(*) for m(4ℓ) != Higgs mass window

- **High multiplicity** final state
- **Rare in SM:** same-sign 2ℓ, 3ℓ, 4ℓ
  - Exploit presence of hadronically decaying τ

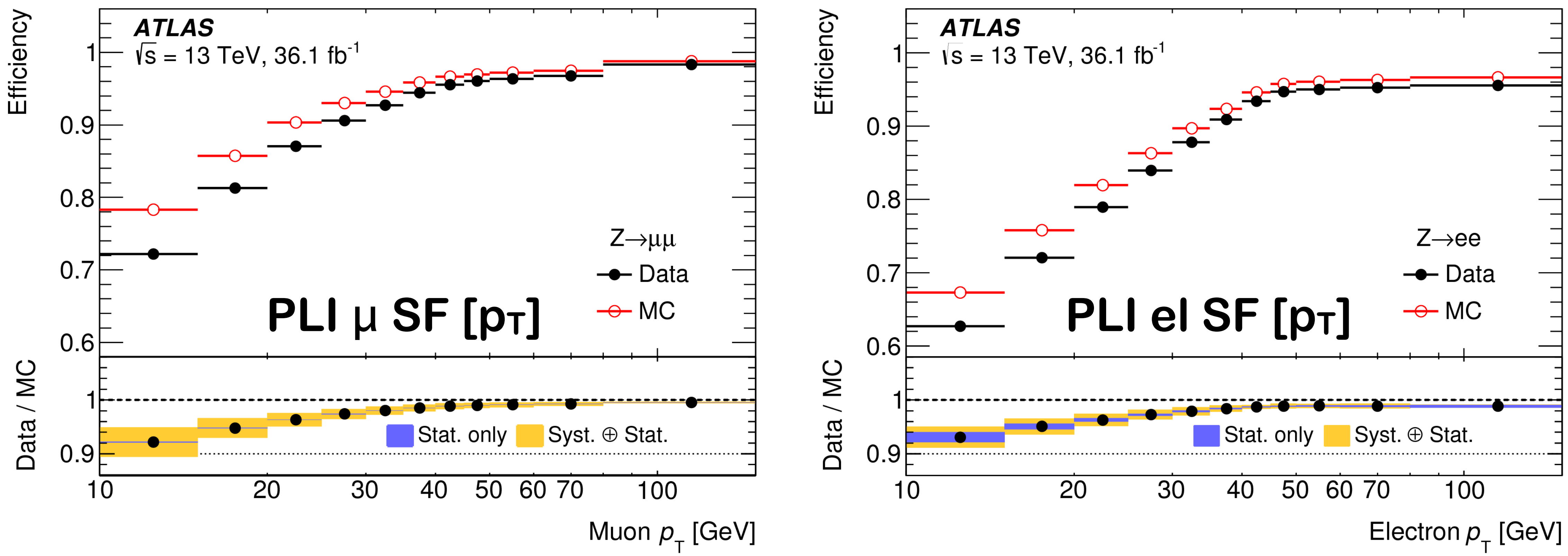
- Split in categories based on **number of e/μ** and **number of τ**
  - **Loose** lepton definition (no isolation, loose ID)
  - Dilepton and single lepton triggers



Tamara Vázquez Schröder (McGill University)
arXiv:1712.08891 submitted to PRD

$t\bar{t}H$ (multileptons): tight lepton definition

- **Common main/important background**: non-prompt leptons from semileptonic b-decay
- **New MVA lepton isolation (PromptLeptonIso=PLI)** to reject non-prompt ℓ based on:
  - lepton and overlapping **track jets** properties
  - lepton track/calorimeter **isolation** variables



- **SF values ~ 0.90 - 0.98**
- SF systematic uncertainties small (max ~3% @ low pT), w/ negligible impact in the analysis
- Factor $\mathcal{O}(20)$ rejection for leptons originating from b-hadrons

- **New MVA cut to reduce QMIsID** for 2ℓSS and $3\ell+0\tau$
  - Factor $\mathcal{O}(17)$ background rejection for a 95% signal efficiency

arXiv:1712.08891 submitted to PRD

$t\bar{t}H$ (multileptons): object definition summary

- Several "Loose" and "Tight" lepton definitions to optimise the event selection in each multilepton channel

	e					μ		
	L	L†	L*	T	T*	L	L†	L*/T/T*
Isolation	No	Yes				No	Yes	
Non-prompt lepton BDT	No		Yes			No		Yes
Identification	Loose			Tight		Loose		
Charge misassignment veto BDT	No				Yes	No		
Transverse impact parameter significance, $\lvert d_0\rvert/\sigma_{d_0}$	< 5					< 3		
Longitudinal impact parameter, $\lvert z_0 \sin\theta\rvert$	< 0.5 mm							

L = Loose
L† = + Loose isolated
L* = + PLI isolated

T = Tight (PLI isolated)
T* = + QMisID MVA veto (el only)

Thad
Medium BDT ID to reject jets (1M, 1T in 1ℓ+2τ)
$p_T > 25$ GeV
BDT to reject el faking τ
τ-μ overlap removal
b-jet veto
τ_{had} vertex is PV

Jets $p_T > 25$ GeV
BJets MV2c10 70% WP

- Minimum jet requirements: $N_{jets} \geq 2$; $N_{b\text{-}jets} \geq 1$

	2ℓSS	3ℓ	4ℓ	$1\ell+2\tau_{had}$	2ℓSS$+1\tau_{had}$	2ℓOS$+1\tau_{had}$	$3\ell+1\tau_{had}$
Light lepton	2T*	1L*, 2T*	2L, 2T	1T	2T*	2L†	1L†, 2T
τ_{had}	0M	0M	–	1T, 1M	1M	1M	1M
N_{jets}, N_{b-jets}	$\geq 4, = 1, 2$	$\geq 2, \geq 1$	$\geq 2, \geq 1$	$\geq 3, \geq 1$	$\geq 4, \geq 1$	$\geq 3, \geq 1$	$\geq 2, \geq 1$

$t\bar{t}H$ (multileptons): multivariate analysis strategy

- **Signal extraction**: fit or cut on **BDTs (boosted decision tree)** to discriminate signal against the main background processes [except in 3ℓ+1τ]
- **Pre-MVA region**: loose selection per channel to train MVA
  - **Input variables**: system reconstruction, pseudo-continuous b-tagging, kinematics *[full list in back-up]*
- **Final selection** per channel:
  - Either pre-MVA selection (e.g. 2ℓSS+0τ), tighter selection (e.g. 2ℓSS+1τ), or split pre-MVA region in categories (e.g. 3ℓ+0τ)
  - **2ℓSS0τ**: combination of two BDTs ($t\bar{t}H$ vs. $t\bar{t}$; $t\bar{t}H$ vs. $t\bar{t}V$)
  - **3ℓ0τ**: 5-dimensional multinominal BDTs mapped to 5 categories ($t\bar{t}H$, $t\bar{t}W$, $t\bar{t}Z$, $t\bar{t}$, VV)
  - **4ℓ (Z-enriched)**: $t\bar{t}H$ vs. $t\bar{t}Z$
  - **2ℓSS+1τ, 2ℓOS+1τ, 1ℓ+2τ**: $t\bar{t}H$ vs. $t\bar{t}$ (with fake τ)

	2ℓSS	3ℓ	4ℓ	$1\ell+2\tau_{\text{had}}$	$2\ell\text{SS}+1\tau_{\text{had}}$	$2\ell\text{OS}+1\tau_{\text{had}}$	$3\ell+1\tau_{\text{had}}$
BDT trained against	Fakes and $t\bar{t}V$	$t\bar{t}$, $t\bar{t}W$, $t\bar{t}Z$, VV	$t\bar{t}Z$ / -	$t\bar{t}$	all	$t\bar{t}$	-
Discriminant	2×1D BDT	5D BDT	Event count	BDT	BDT	BDT	Event count
Number of bins	6	5	1 / 1	2	2	10	1
Control regions	-	4	-	-	-	-	-

Tamara Vázquez Schröder (McGill University)
arXiv:1712.08891 submitted to PRD

$t\bar{t}H$ (multileptons): background composition

✱ **Non-prompt** lepton in $t\bar{t}$
- semileptonic b-decay
- γ conversions

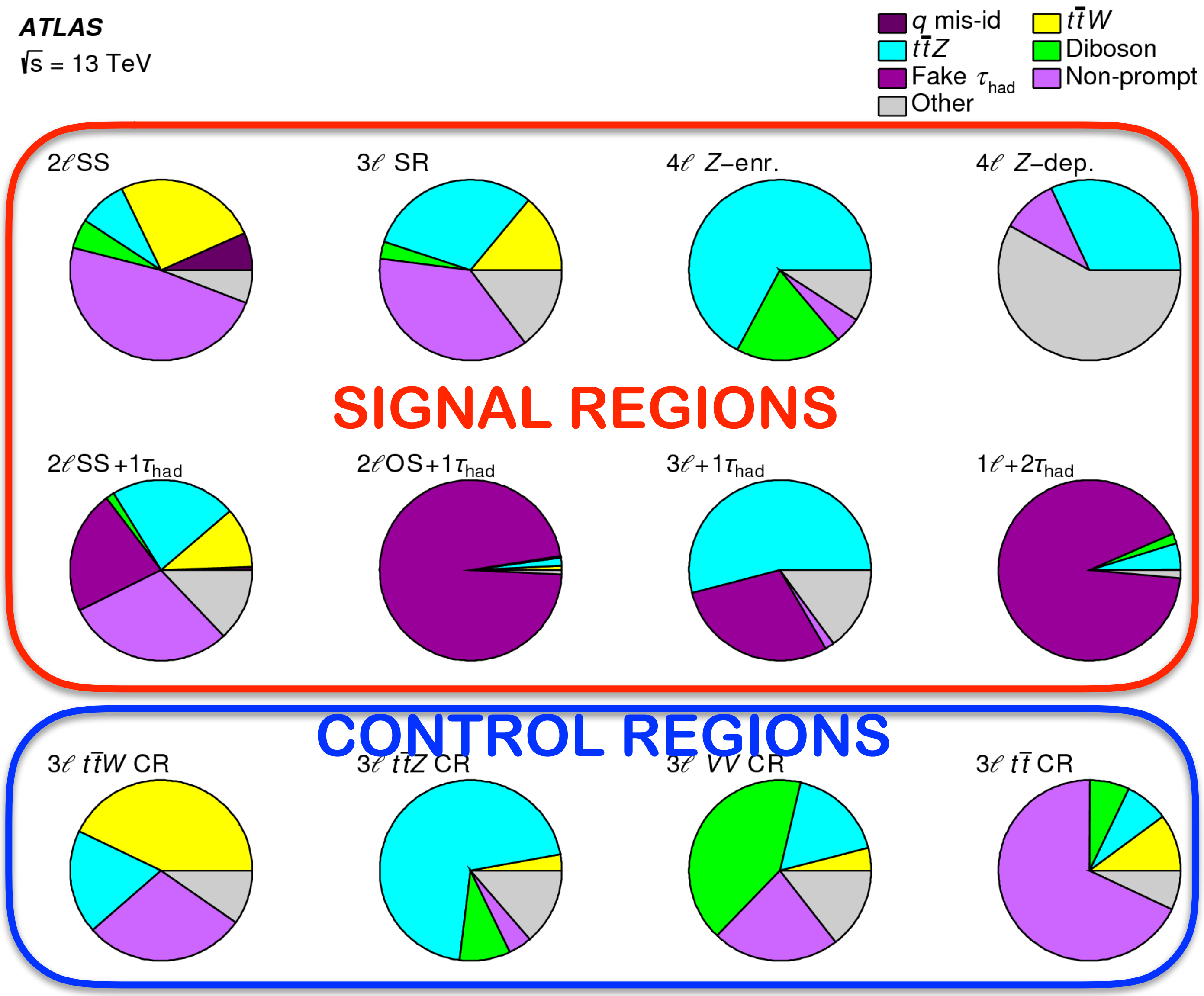
✱ **Fake τ** from light/b-jets

DATA-DRIVEN (DD):
MATRIX METHOD (MM), FAKE FACTOR (FF)

FF ~ matrix method except prompt background is taken from MC

✱ **Misidentified charge** lepton
- e.g. trident electrons (Bremsstrahlung)
- using **3D likelihood method** [pT, η, Tight/Loose]

DATA-DRIVEN (DD):
LIKELIHOOD FIT



✱ **Irreducible backgrounds** with prompt-leptons (**$t\bar{t}Z$, $t\bar{t}W$, VV**)

MC
(cross check: fit to data)

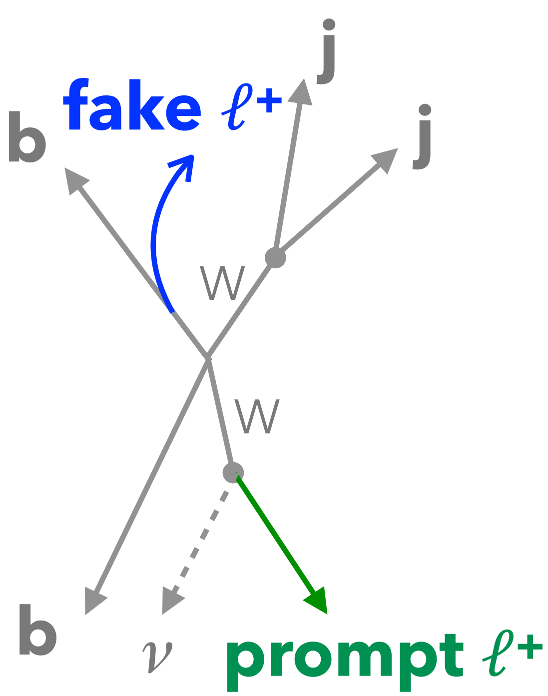
"Other": 4tops, $t\bar{t}WW$, tH, tZ

arXiv:1712.08891 submitted to PRD

$t\bar{t}H$ (multileptons): non-prompt light ℓ

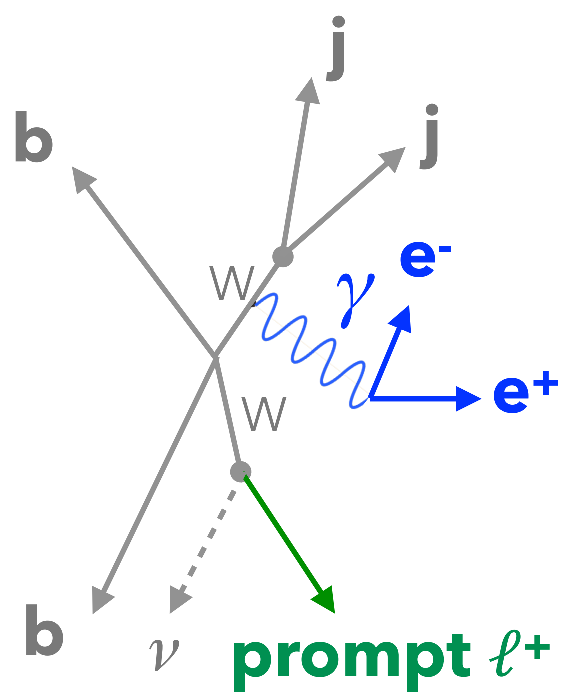
Method [parametr.]	$2\ell SS+0\tau$			$3\ell+0\tau$	4ℓ	$2\ell SS+1\tau$	Other τ channels
Non-prompt lepton	**DD (MM)** el: [p_T, NBjets] μ: [p_T, dR(μ,j)]				**pseudo-DD (Fake SF)**	**DD (FF)** el/μ: [p_T]	**MC** (very small)
DD/MC	ee: 2.0±0.5	eμ: 1.7±0.4	μμ: 1.5 ±0.5	SR: 1.8 ± 0.8			

Semileptonic b-decay



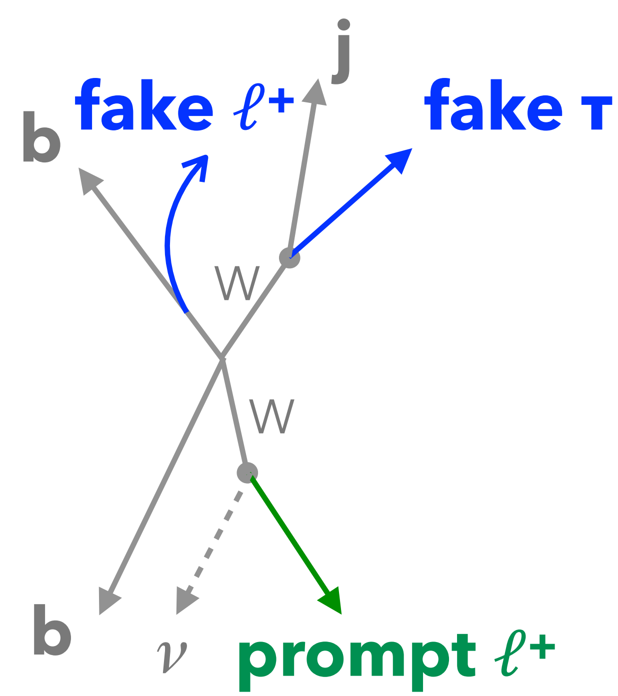
strongly reduced with PLI

Photon conversions



50% of the "fakes" in 3ℓ!

Non-prompt lepton & fake τ



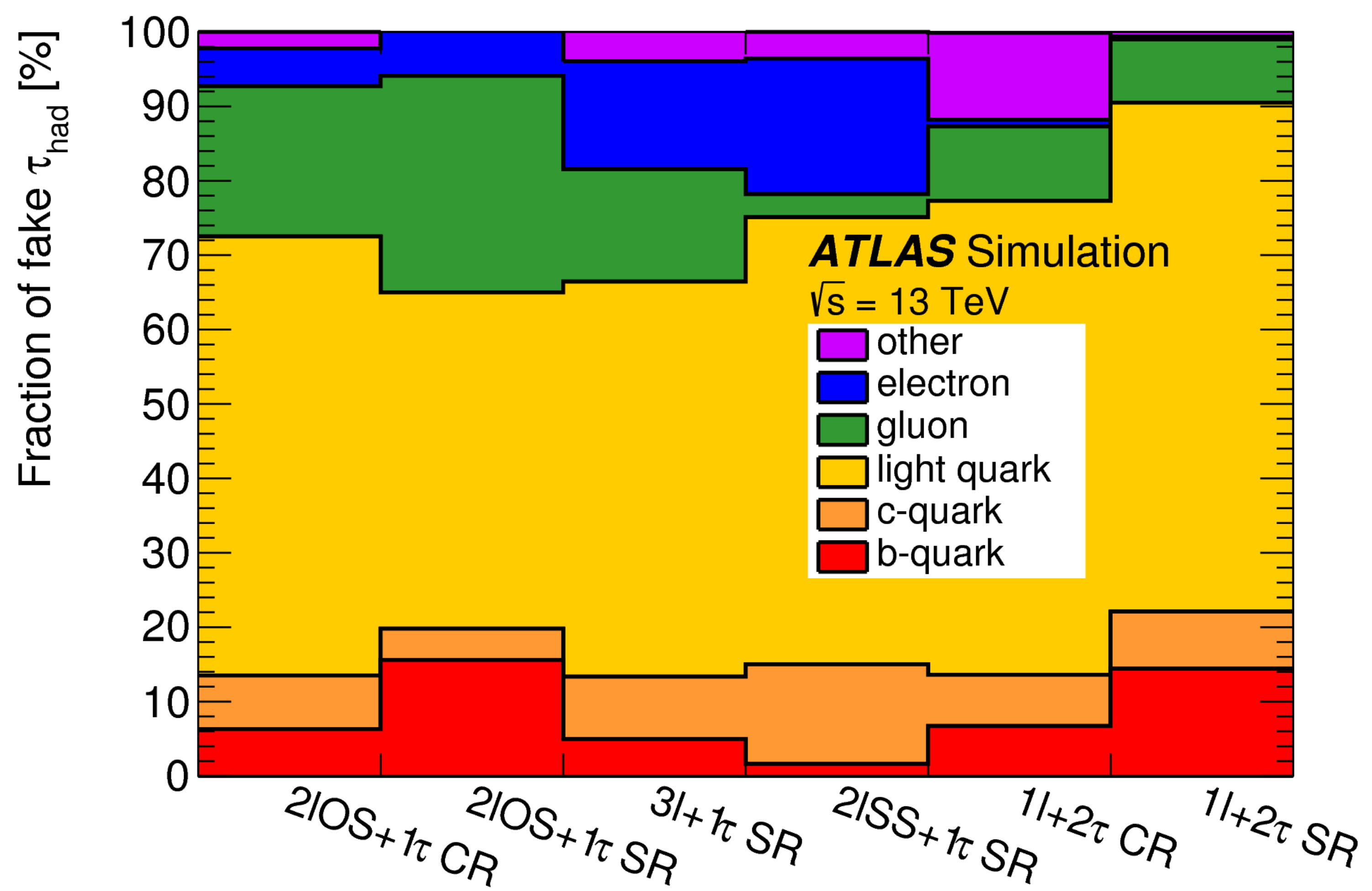
70% from $t\bar{t}$ in $2\ell SS+1\tau$

arXiv:1712.08891 submitted to PRD

$t\bar{t}H$ (multileptons): fake tau

Estimate method [parametrisation]	1ℓ+2τ	2ℓOS+1τ	2ℓSS+1τ	3ℓ+1τ
Fake tau	DD (SS data)	DD (FF) [p_T]	pseudo-DD (MC correction with 2ℓOS+1τ DD SF)	

- **1ℓ+2τ:** mostly $t\bar{t}$ with 1 or 2 fake τ
- **2ℓOS+1τ:** fake factor
  - Mainly $t\bar{t}$ with jet faking τ
  - Fake rates parametrised in τ pT
- **2ℓSS+1τ and 3ℓ+1τ:** MC correction with SF derived from **{DD(2ℓOS+1τ) / MC}**
  - Harmonised 1-fake-τ estimate for all channels, profit from large statistics from 2ℓOS+1τ
  - Final SF = 1.36 ± 0.16

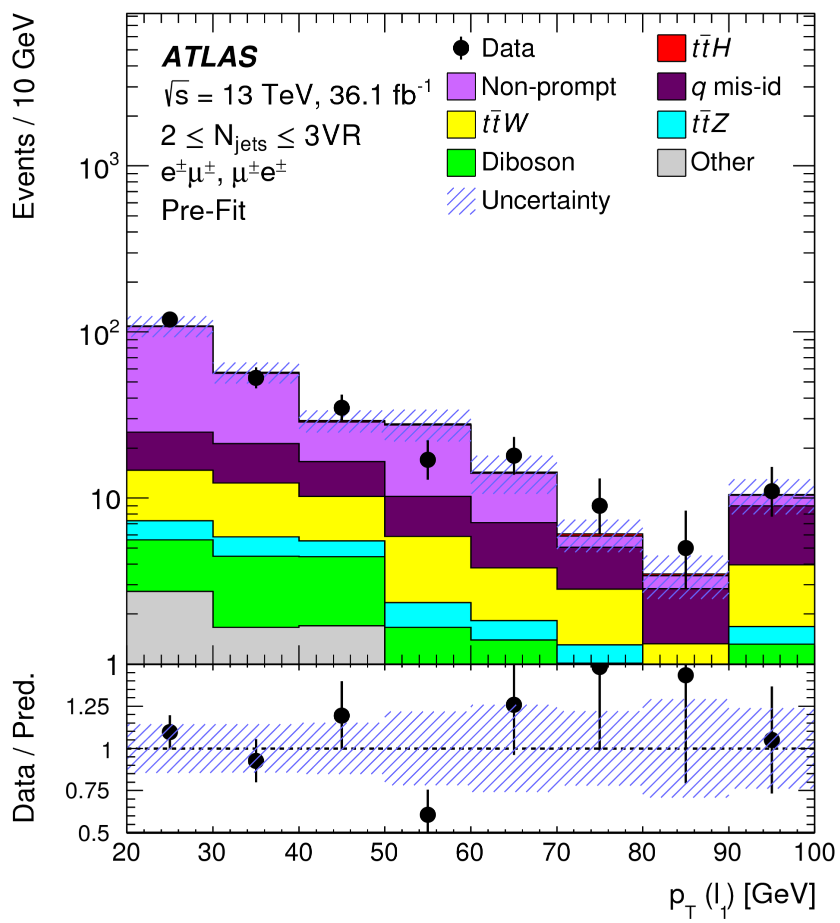


Tamara Vázquez Schröder (McGill University)
arXiv:1712.08891 submitted to PRD

$t\bar{t}H$ (multileptons): fakes/non-prompt ℓ validation

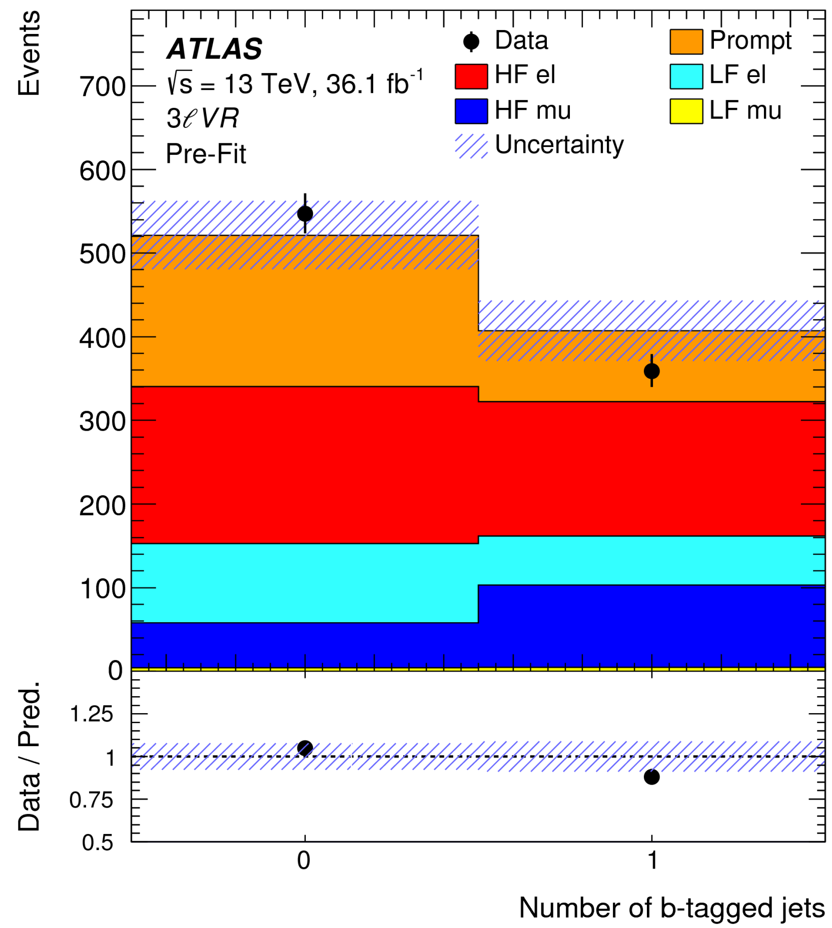
* Overall **reasonable data/prediction agreement** with estimates fakes in VRs

"Low N_{jets}" $2 \leq N_{jets} \leq 3$, $N_{jets} \geq 1$



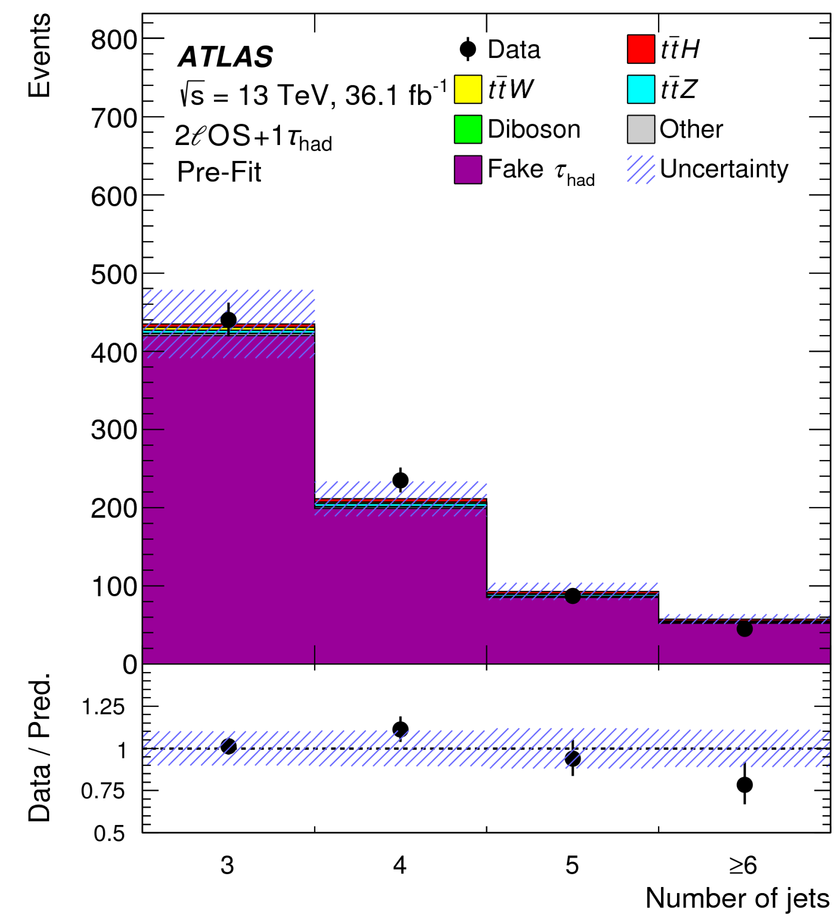
2ℓSS+0τ fake ℓ

"3 loose light ℓ" $N_{jets} \geq 3$



4ℓ fake ℓ

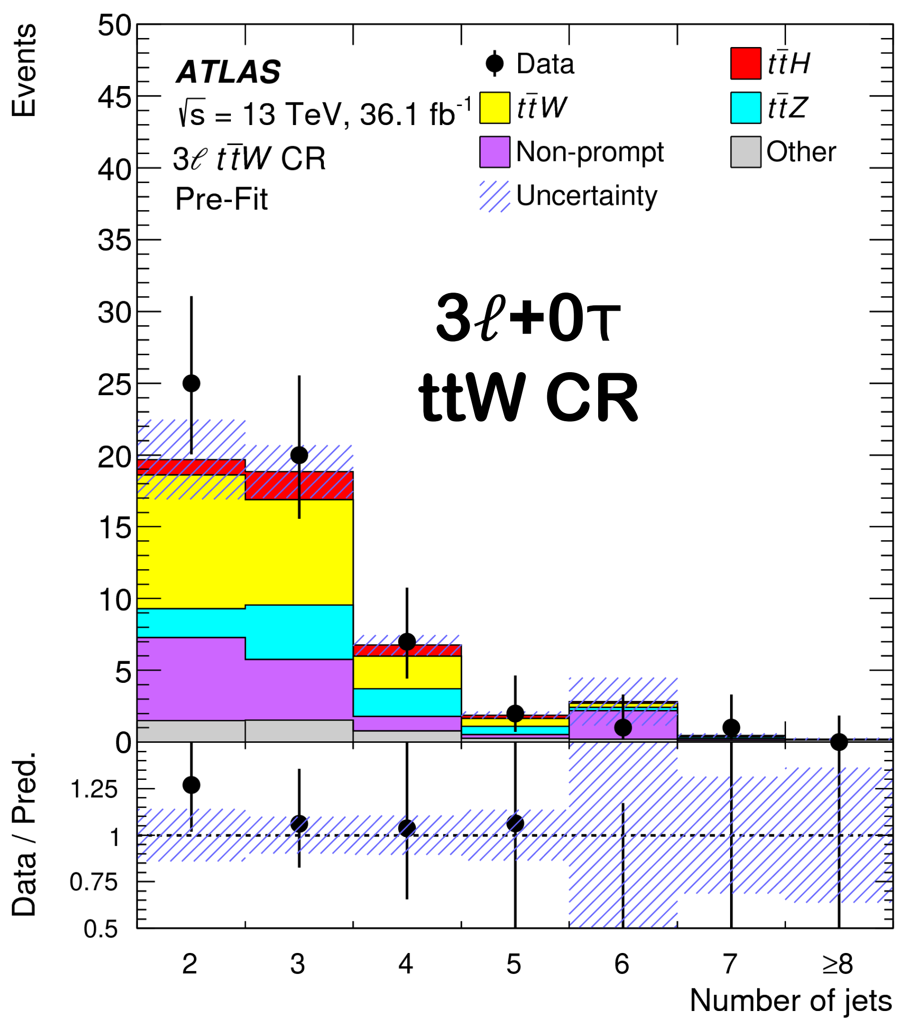
"Pre-selection" $N_{jets} \geq 3$

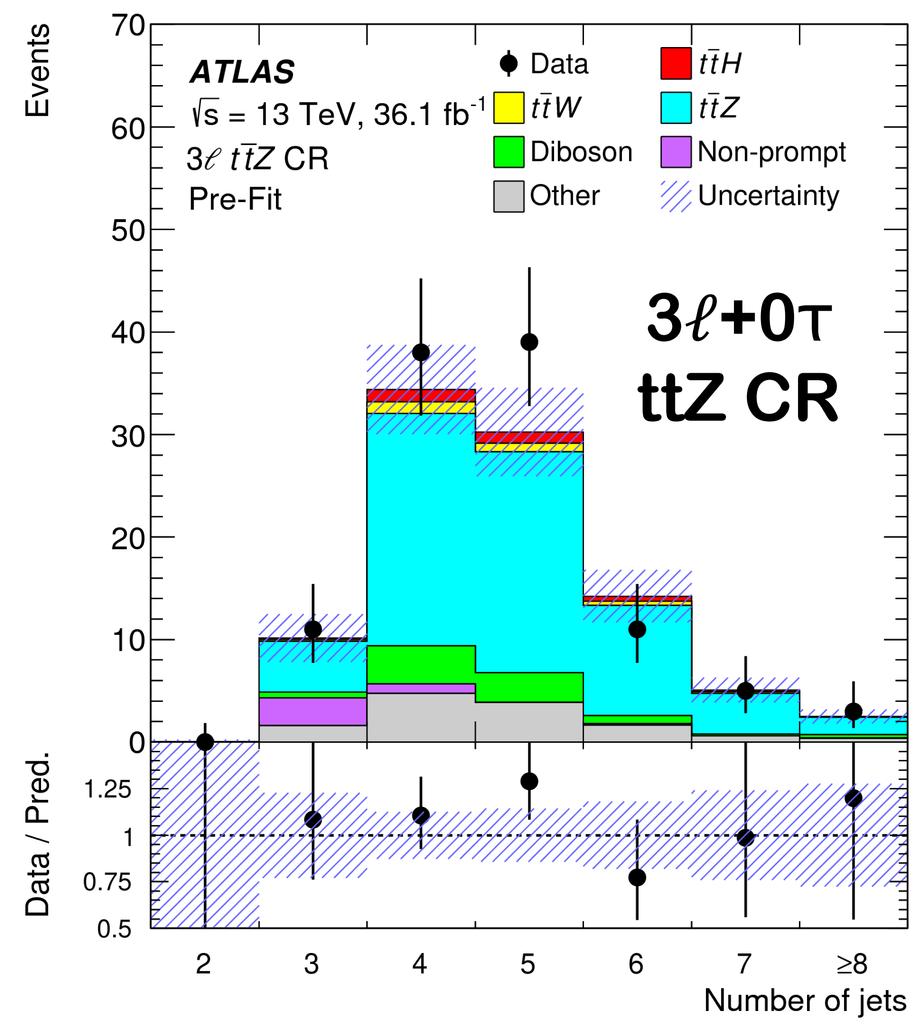


2ℓOS+1τ fake τ

$t\bar{t}H$ (multileptons): prompt ℓ background validation

- Largest irreducible backgrounds: **$t\bar{t}W$, $t\bar{t}Z$, diboson**
- Estimated using **NLO MC samples**, with theory/modelling uncertainties:
  - Cross-section uncertainties
  - Scale variations
  - Generator comparisons
- Validated in several regions, eg: 3ℓ $t\bar{t}W/Z$ CRs built using the multinomial BDT
- Overall **good data/prediction agreement** in $t\bar{t}V$-enriched CRs using MC simulation
  - Also good agreement in cut-based VRs





$t\bar{t}H$ (multileptons): profile likelihood fit

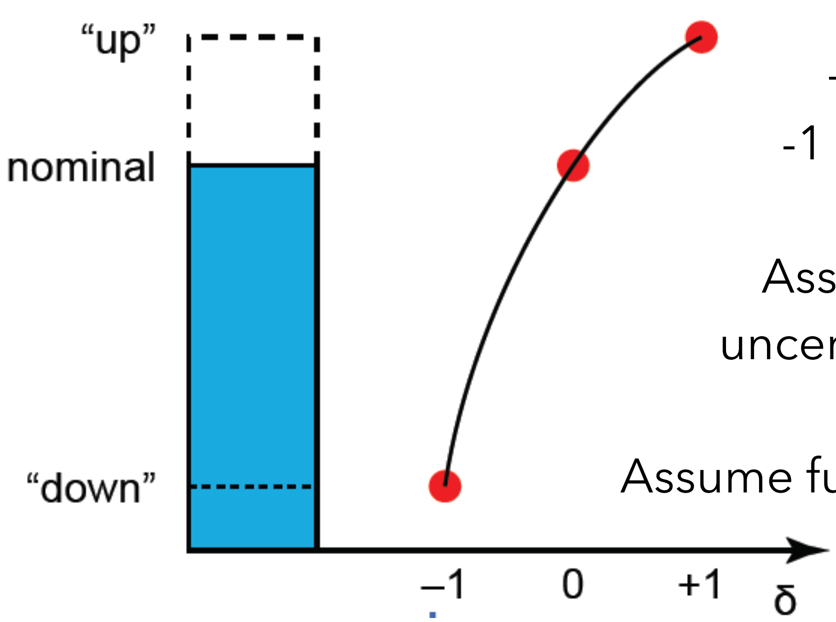
- Binned profile likelihood fit

$$L(\mu,\theta) = L_{Pois}(\mu,\theta)\cdot\prod_p \frac{1}{\sqrt{2\pi}}\exp\left(-\frac{\theta_p^2}{2}\right)$$

- **Parameter of interest**: signal strength

$$\mu_{t\bar{t}H} = \frac{\sigma_{t\bar{t}H}}{\sigma_{t\bar{t}H}^{SM}}$$

- **Systematic uncertainties** included in the fit as <u>nuisance parameters θ</u>
    - Need sufficiently flexible model of signal and background!



Given:
+1σ, nominal, and
-1 σ of each systematic

Assume interpolation of
uncertainty between [-1,1] σ

Assume full correlation between bins

Constrain uncertainty in <u>control region</u>, propagate this knowledge to <u>signal region</u>

- Find best values for μ and θ from minimising the -log L
- Calculate experimental sensitivity in terms of the significance
    - Quantify level of disagreement between data and background-only hypothesis as Gaussian standard deviations (σ)

$t\bar{t}H$ (multileptons): systematic uncertainties

Systematic uncertainty	Type	Components
Luminosity	N	1
Pileup reweighting	SN	1
Physics Objects		
Electron	SN	6
Muon	SN	15
τ_{had}	SN	10
Jet energy scale and resolution	SN	28
Jet vertex fraction	SN	1
Jet flavor tagging	SN	126
$E_{\text{T}}^{\text{miss}}$	SN	3
Total (Experimental)	–	191
Data-driven non-prompt/fake leptons and charge misassignment		
Control region statistics	SN	38
Light-lepton efficiencies	SN	22
Non-prompt light-lepton estimates: non-closure	N	5
γ-conversion fraction	N	5
Fake τ_{had} estimates	N/SN	12
Electron charge misassignment	SN	1
Total (Data-driven reducible background)	–	83
$t\bar{t}H$ **modeling**		
Cross section	N	2
Renormalization and factorization scales	S	3
Parton shower and hadronization model	SN	1
Higgs boson branching fraction	N	4
Shower tune	SN	1
$t\bar{t}W$ **modeling**		
Cross section	N	2
Renormalization and factorization scales	S	3
Matrix-element MC event generator	SN	1
Shower tune	SN	1
$t\bar{t}Z$ **modeling**		
Cross section	N	2
Renormalization and factorization scales	S	3
Matrix-element MC event generator	SN	1
Shower tune	SN	1
Other background modeling		
Cross section	N	15
Shower tune	SN	1
Total (Signal and background modeling)	–	41
Total (Overall)	–	315

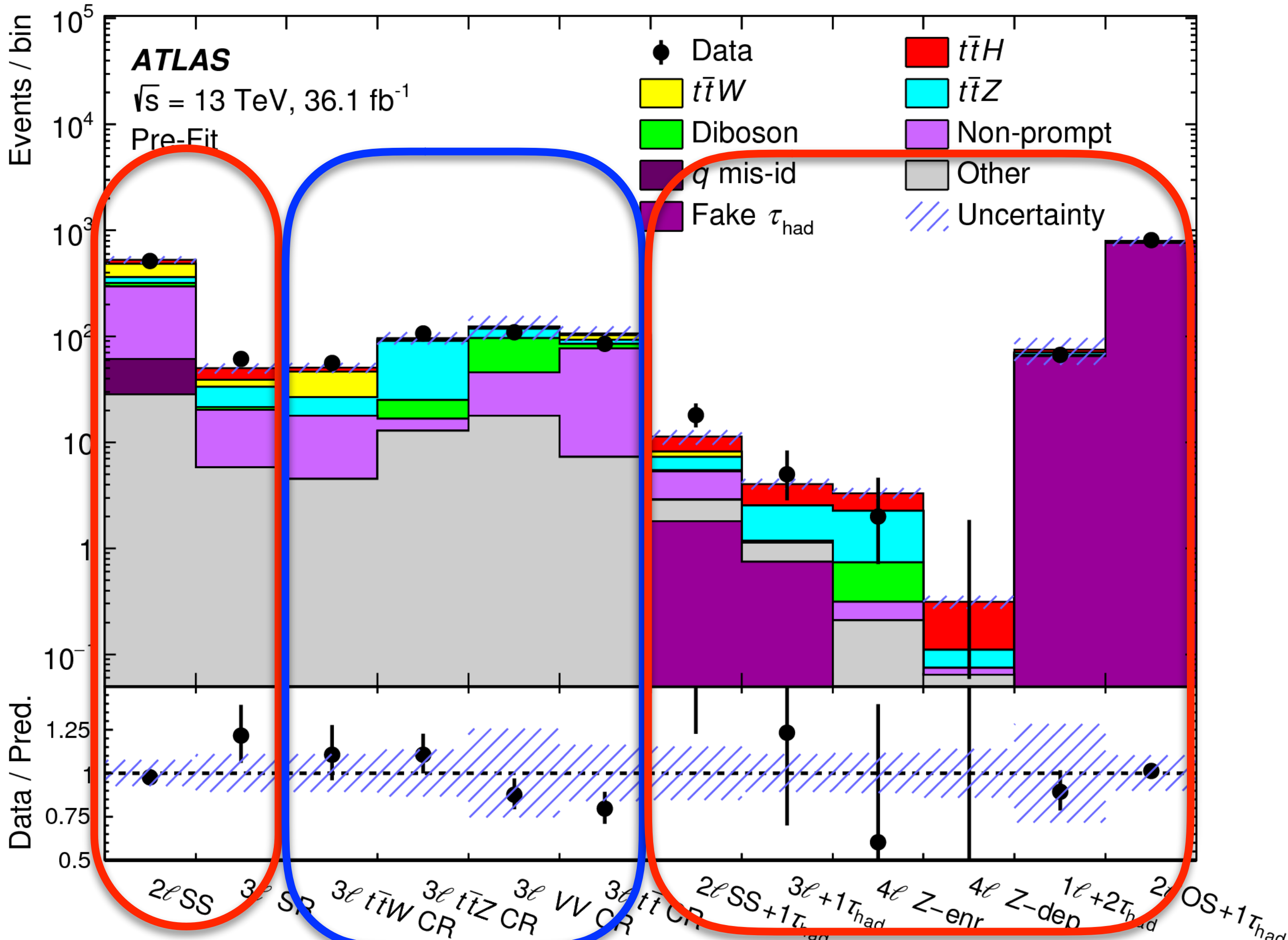
- **Instrumental/detector** uncertainties correlated across channels
  - Exception: JES Flavour Composition
- **DD fake estimate** uncertainties
- **Prompt background modelling** uncertainties correlated across channels

One parameter of interest: $\mu(t\bar{t}H)$

315 nuisance parameters

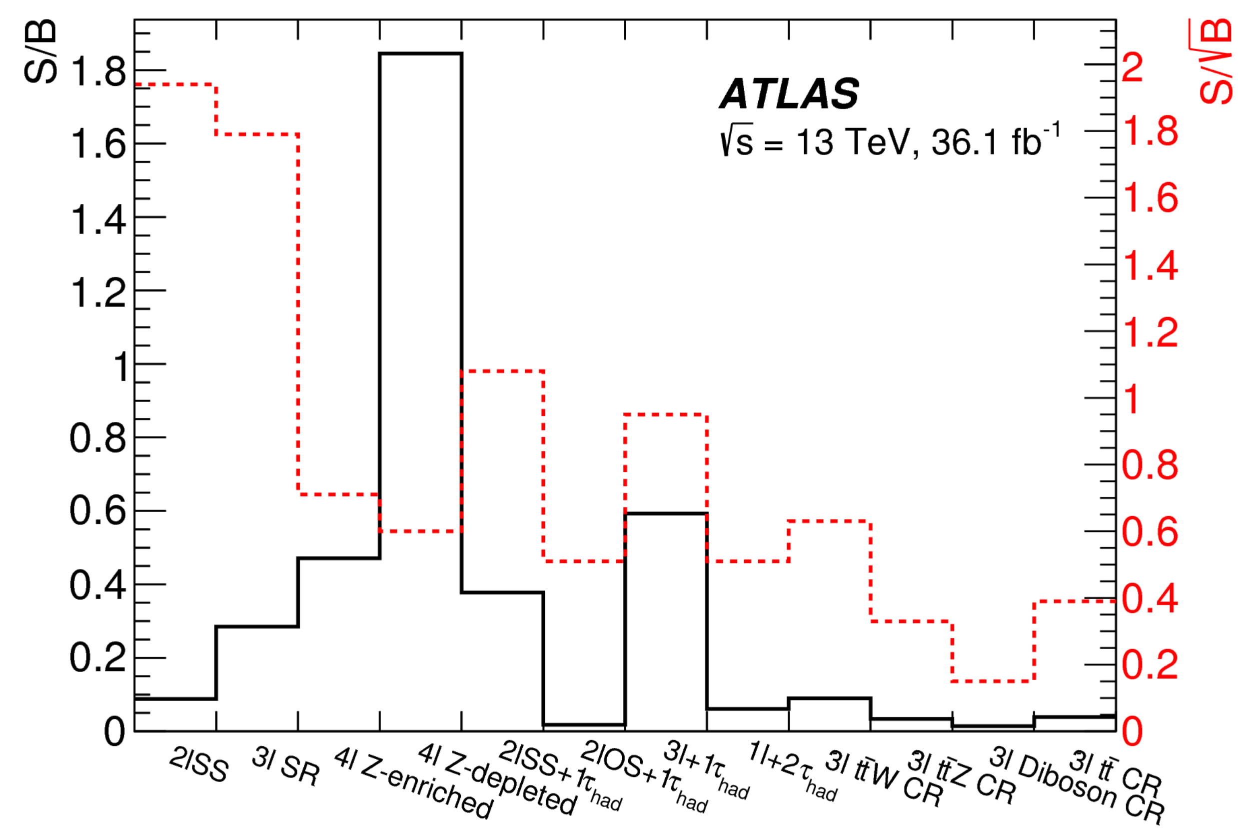
$t\bar{t}H$ (multileptons): pre-fit summary

- **Most statistically sensitive** to $t\bar{t}H$: $2\ell SS+0\tau$ and $3\ell+0\tau$
- **Purest** but lowest statistics: 4ℓ
- **Largest pre-fit excess** per fit category: $2\ell SS+1\tau$



CONTROL REGIONS

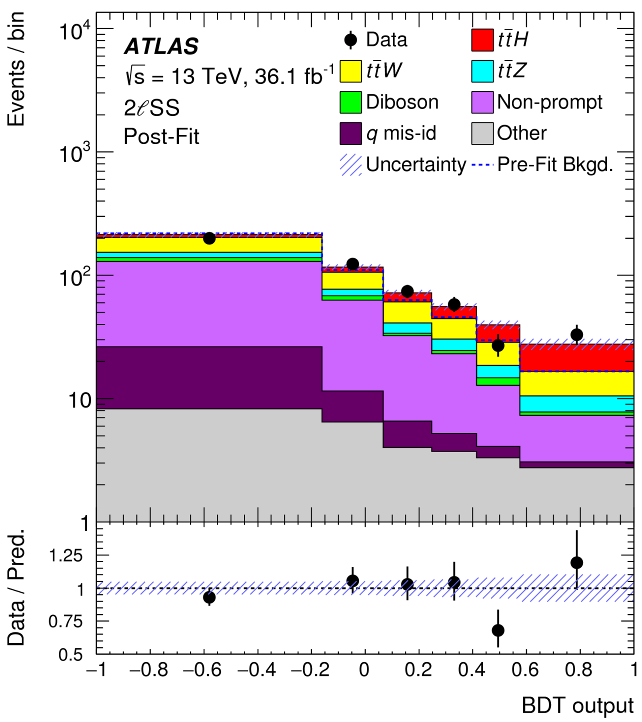
SIGNAL REGIONS



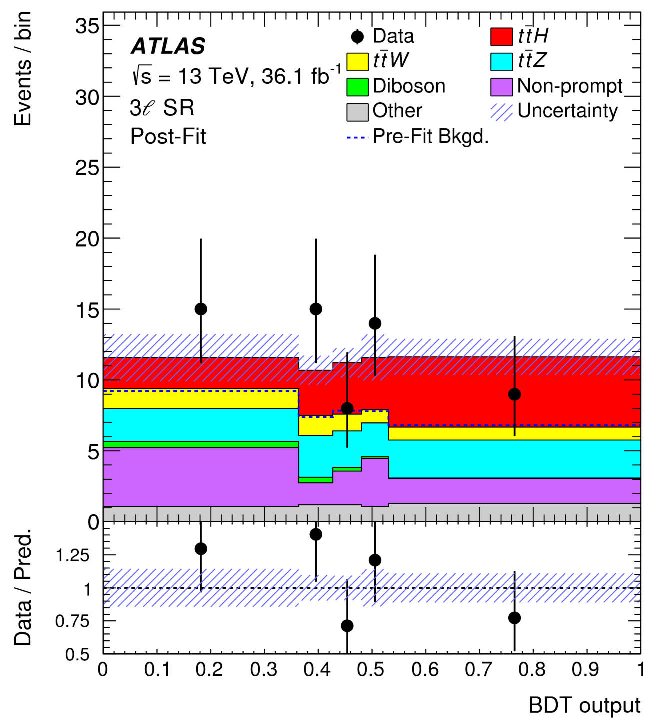
"Other": 4tops, $t\bar{t}WW$, tH, tZ

$t\bar{t}H$ (multileptons): post-fit SRs

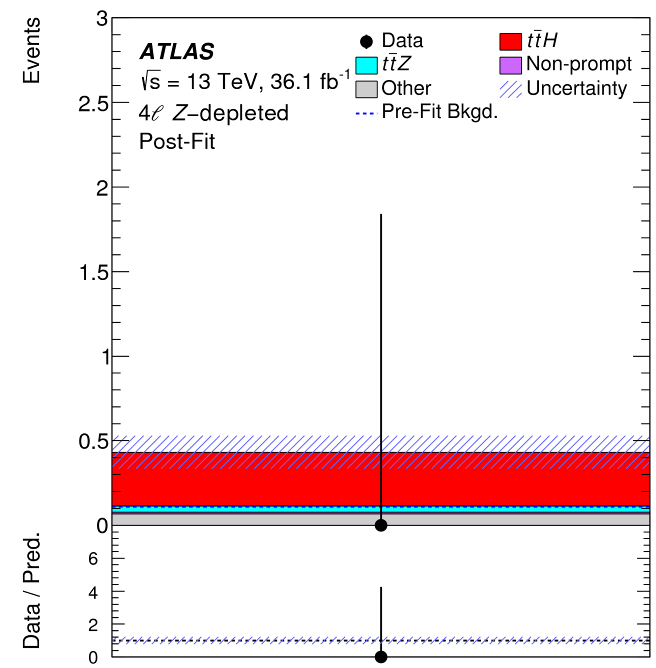
2ℓSS+0τ



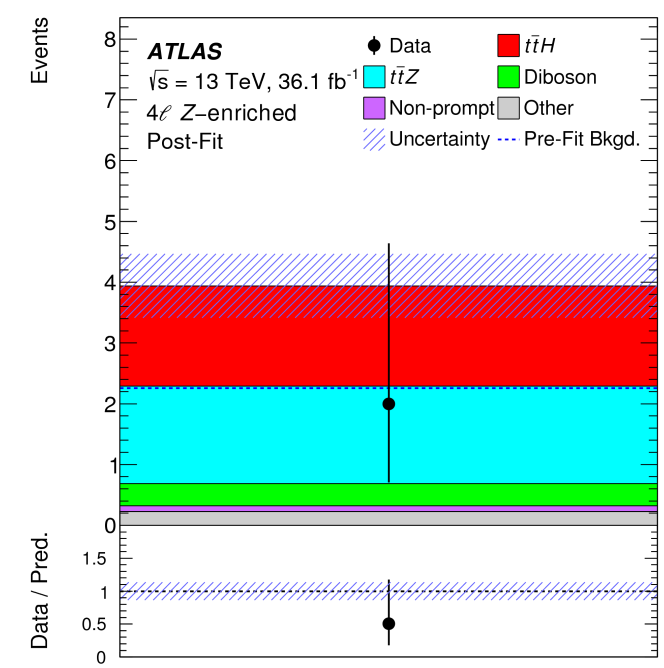
3ℓ+0τ



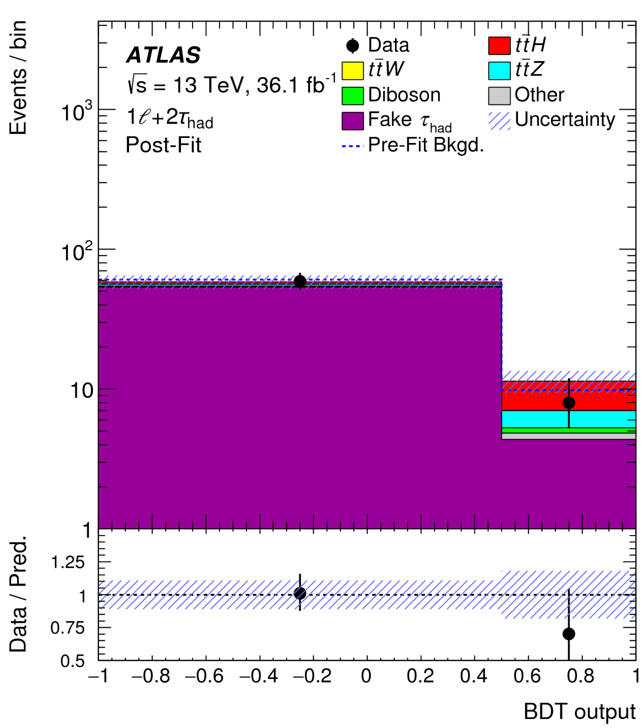
4ℓ (Z-depleted)



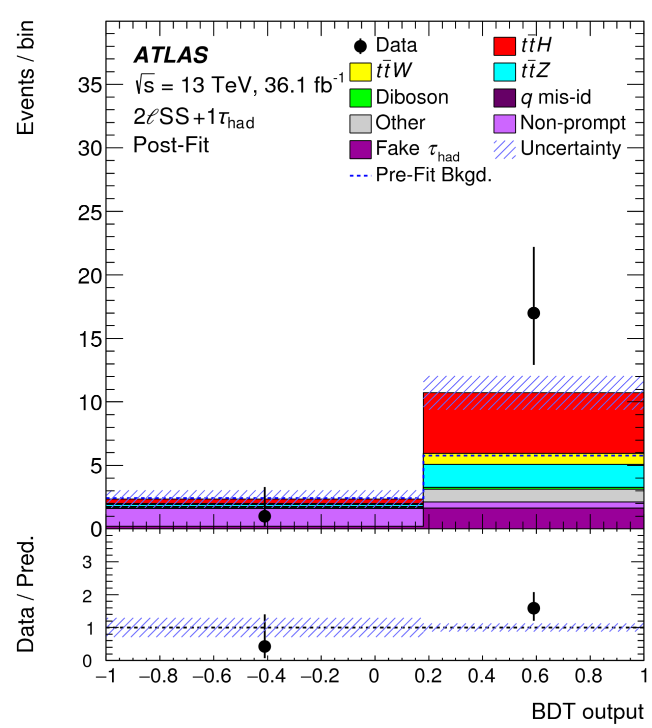
4ℓ (Z-enriched)



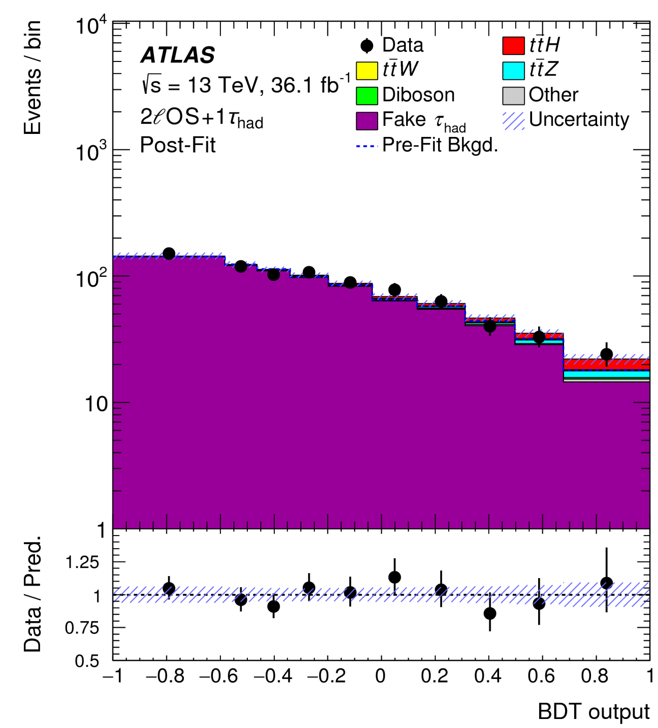
1ℓ+2τ



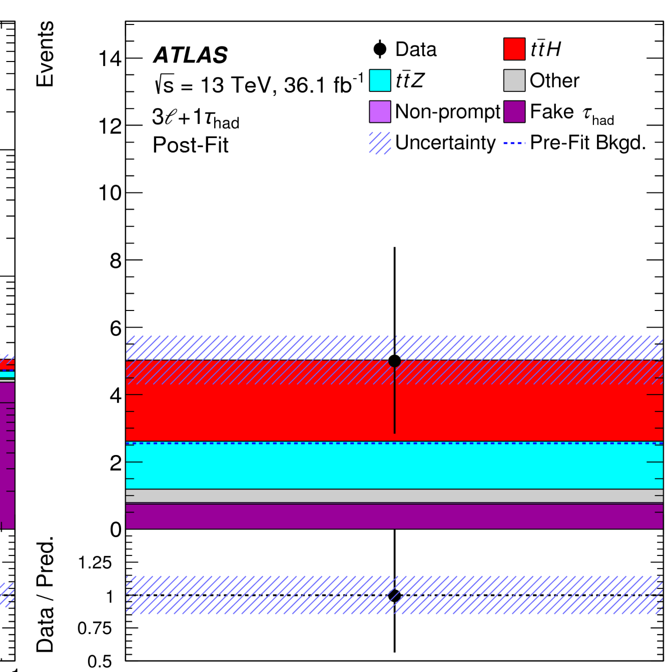
2ℓSS+1τ



2ℓOS+1τ

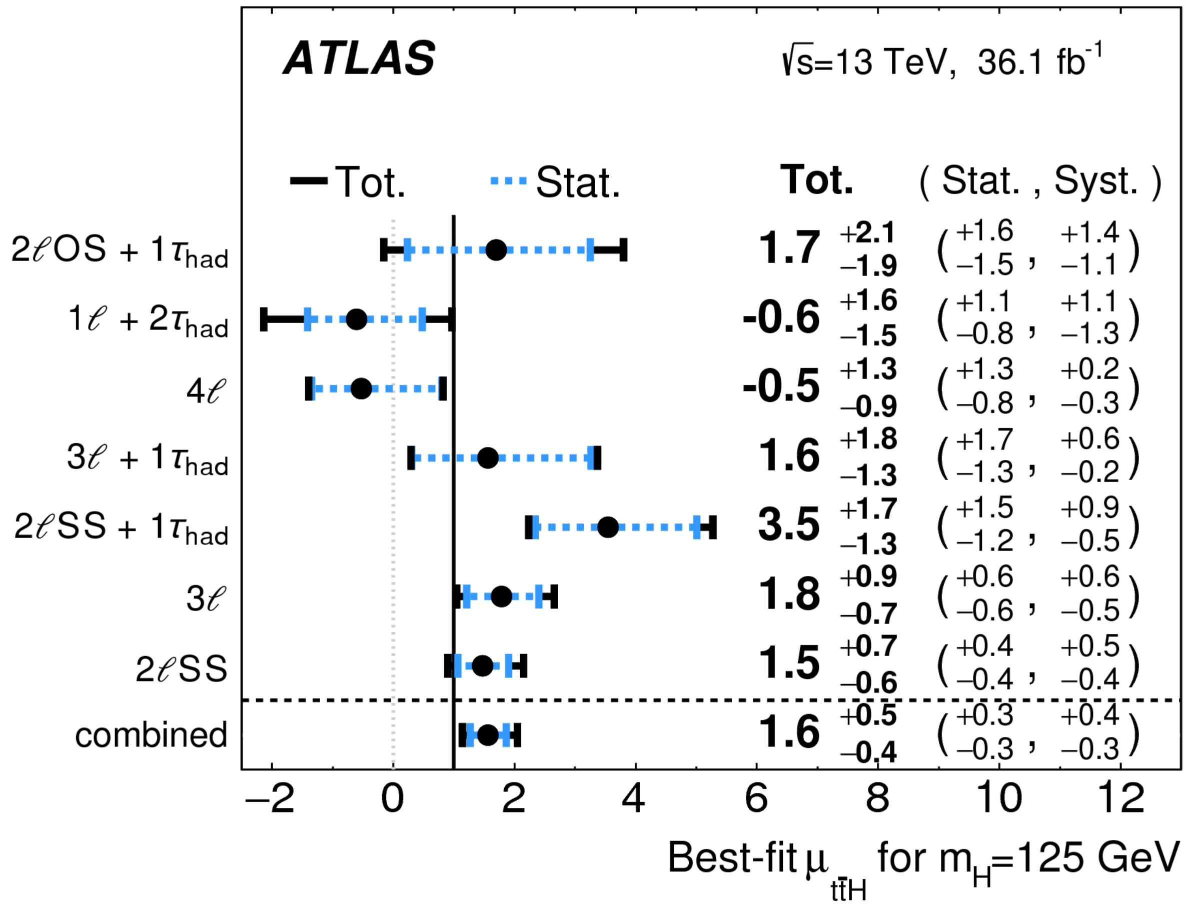


3ℓ+1τ



arXiv:1712.08891 submitted to PRD

$t\bar{t}H$ (multileptons): fit results (II)



Channel	Significance	
	Observed	Expected
2ℓOS+$1\tau_{had}$	0.9σ	0.5σ
1ℓ+$2\tau_{had}$	-	0.6σ
4ℓ (*)	-	0.8σ
3ℓ+$1\tau_{had}$	1.3σ	0.9σ
2ℓSS+$1\tau_{had}$	3.4σ	1.1σ
3ℓ	2.4σ	1.5σ
2ℓSS	2.7σ	1.9σ
Combined	4.1σ	2.8σ

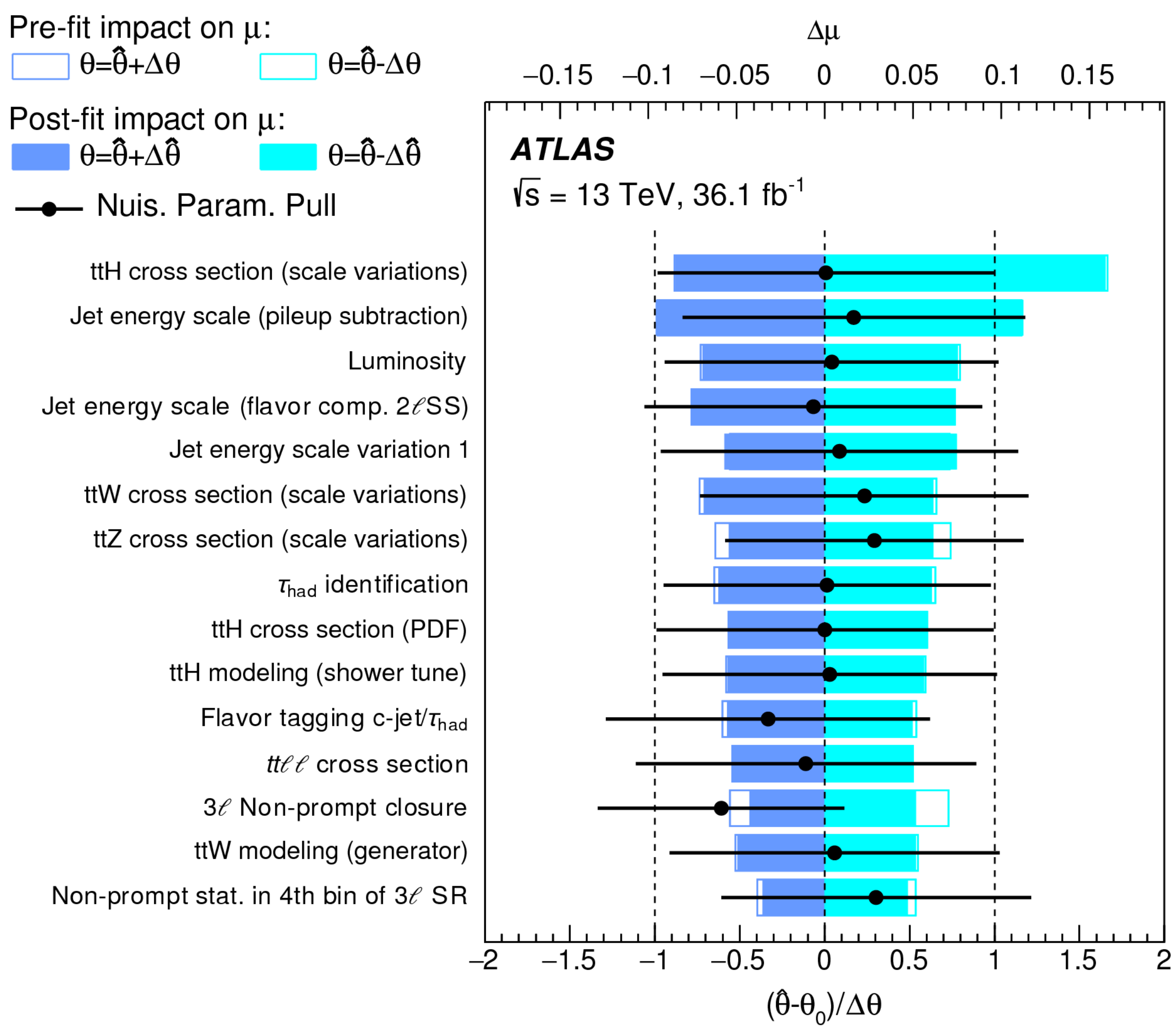
- Cross-section extrapolated to the inclusive phase space:
  - **σ(ttH)=790 ±150 (stat.) $^{+170}_{-150}$ (syst.) fb**
- Significance with respect to background-only hypothesis = **4.1 σ (2.8 σ) obs (exp)**
- Compatible with SM (within 1.4σ)
- Compatibility (7 chan.) = 34%
- **Alternative fit**: $t\bar{t}Z$ and $t\bar{t}W$ normalisation free-floating
  - 15% loss in sensitivity: $\mu(t\bar{t}H) = 1.57^{+0.57}_{-0.50}$
  - $\mu(t\bar{t}Z/W)$ in agreement with SM: $\mu_{t\bar{t}W} = 0.92 \pm 0.32$; $\mu_{t\bar{t}Z} = 1.17^{+0.25}_{-0.22}$

(*) for m(4ℓ) != Higgs mass window

$t\bar{t}H$ (multileptons): fit results (I)

- **Largest (grouped) impact on μ($t\bar{t}H$)**:
  - signal modelling, JES and JER, and the non-prompt light ℓ estimates
- **No major constraints or pulls** of nuisance parameters

Uncertainty Source	$\Delta\mu$	
$t\bar{t}H$ modeling (cross section)	+0.20	−0.09
Jet energy scale and resolution	+0.18	−0.15
Non-prompt light-lepton estimates	+0.15	−0.13
Jet flavor tagging and τ_{had} identification	+0.11	−0.09
$t\bar{t}W$ modeling	+0.10	−0.09
$t\bar{t}Z$ modeling	+0.08	−0.07
Other background modeling	+0.08	−0.07
Luminosity	+0.08	−0.06
$t\bar{t}H$ modeling (acceptance)	+0.08	−0.04
Fake τ_{had} estimates	+0.07	−0.07
Other experimental uncertainties	+0.05	−0.04
Simulation sample size	+0.04	−0.04
Charge misassignment	+0.01	−0.01
Total systematic uncertainty	+0.39	−0.30



arXiv:1712.08891 submitted to PRD

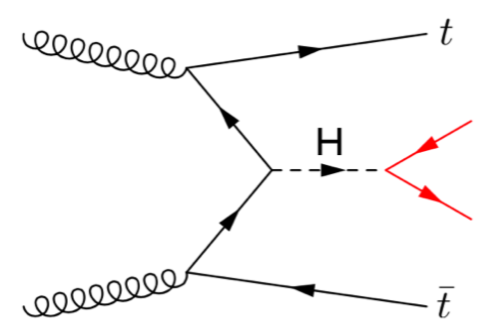
$t\bar{t}H(b\bar{b})$: analysis strategy

- **Biggest challenge**: good modelling of the $t\bar{t}$+HF (≥1b, ≥1c) background
  - Nominal sample: 5-flavour scheme
  - Relative contribution of $t\bar{t}$+≥1b sub-components reweighted to $t\bar{t}+b\bar{b}$ predictions by Sherpa+OpenLoops (4-flavour scheme)

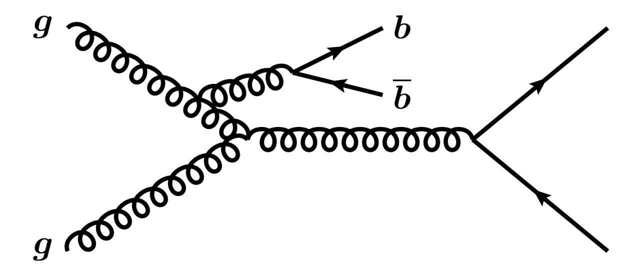
- **Channel categorisation** based on
  - Number of ℓ (1 or 2 opposite-sign)
  - Number of jets
  - Requirements on the b-tagging discriminant (4 calibrated working points)
  - Resolved or boosted, for single lepton channel

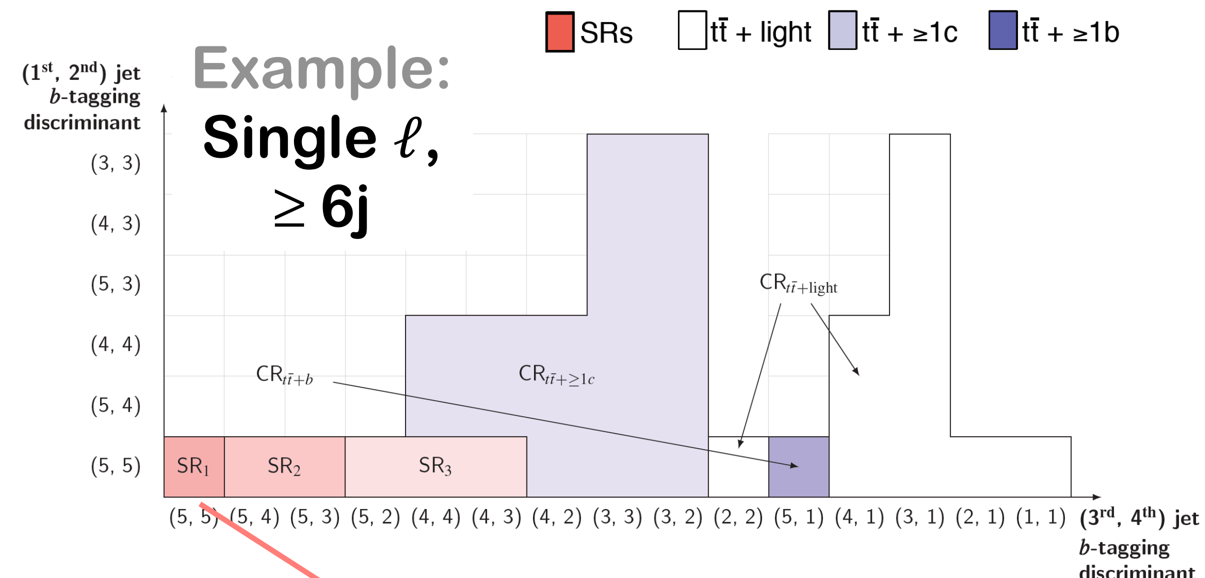
- **MVA analysis** needed to discriminate signal from the overwhelming background
  - The '**classification BDT**' includes as input variables: kinematic variables, reconstruction BDTs (resolved), likelihood and matrix element method discriminants (where available), discrete btagging discriminant

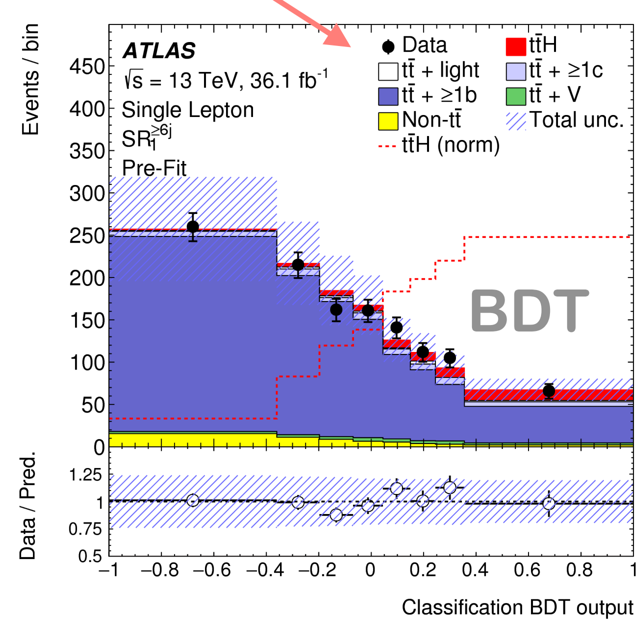
signal (ttH)

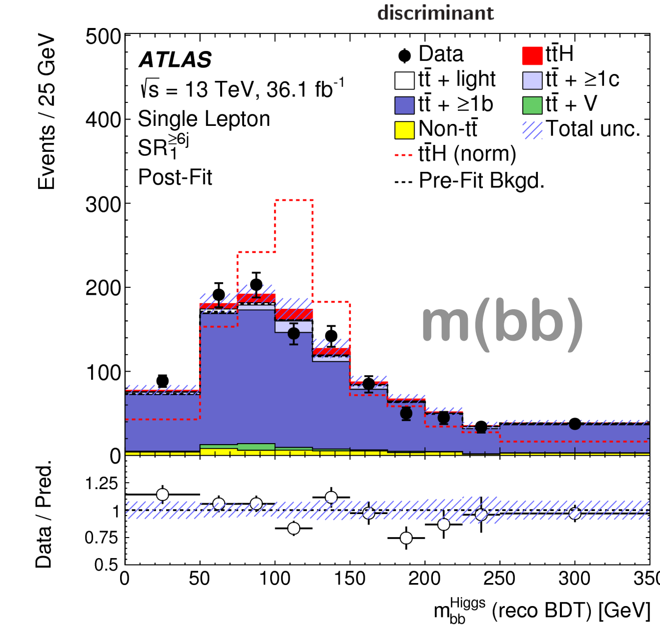


ttbb



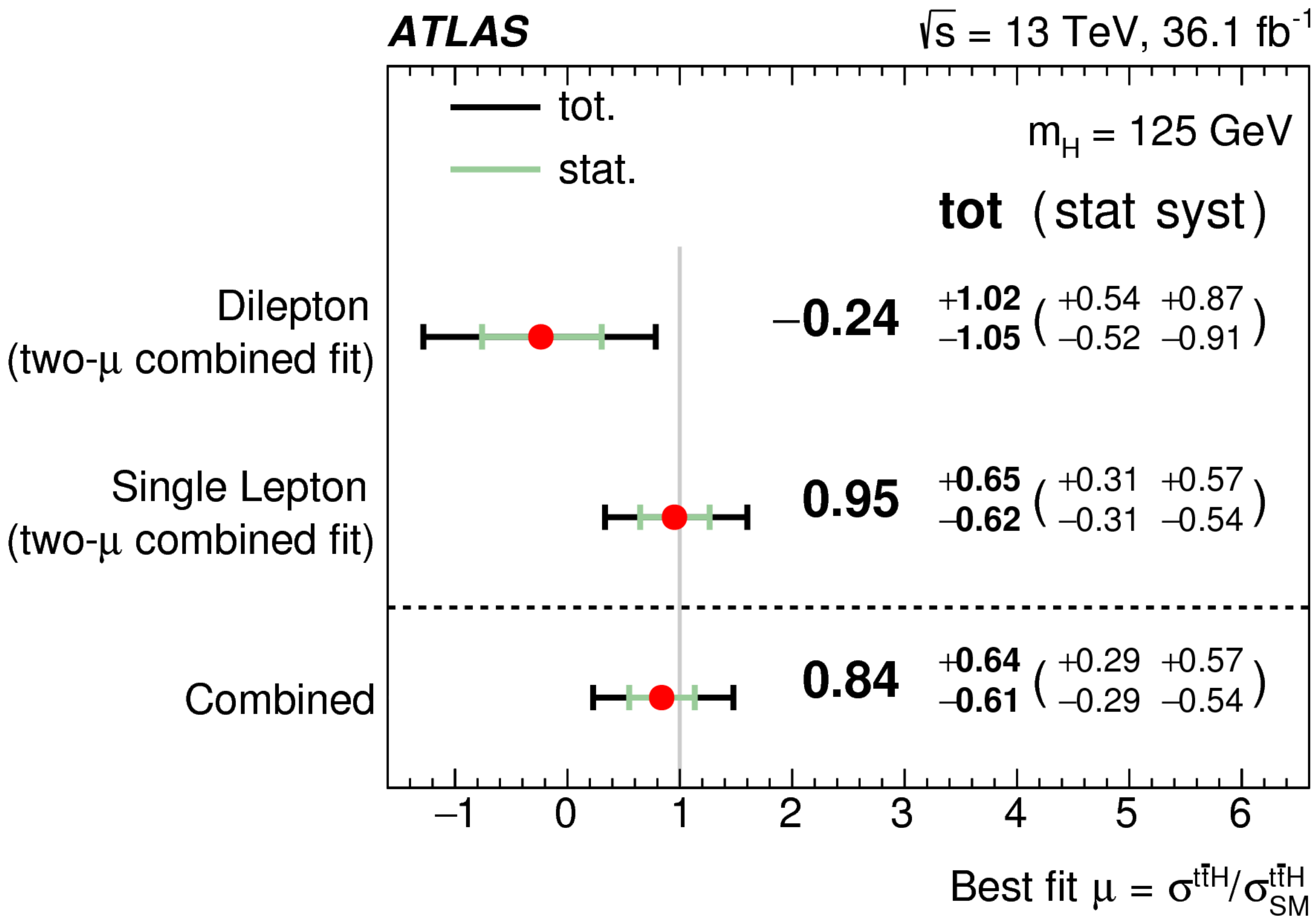






arXiv:1712.08895 submitted to PRD

$t\bar{t}H(b\bar{b})$: results



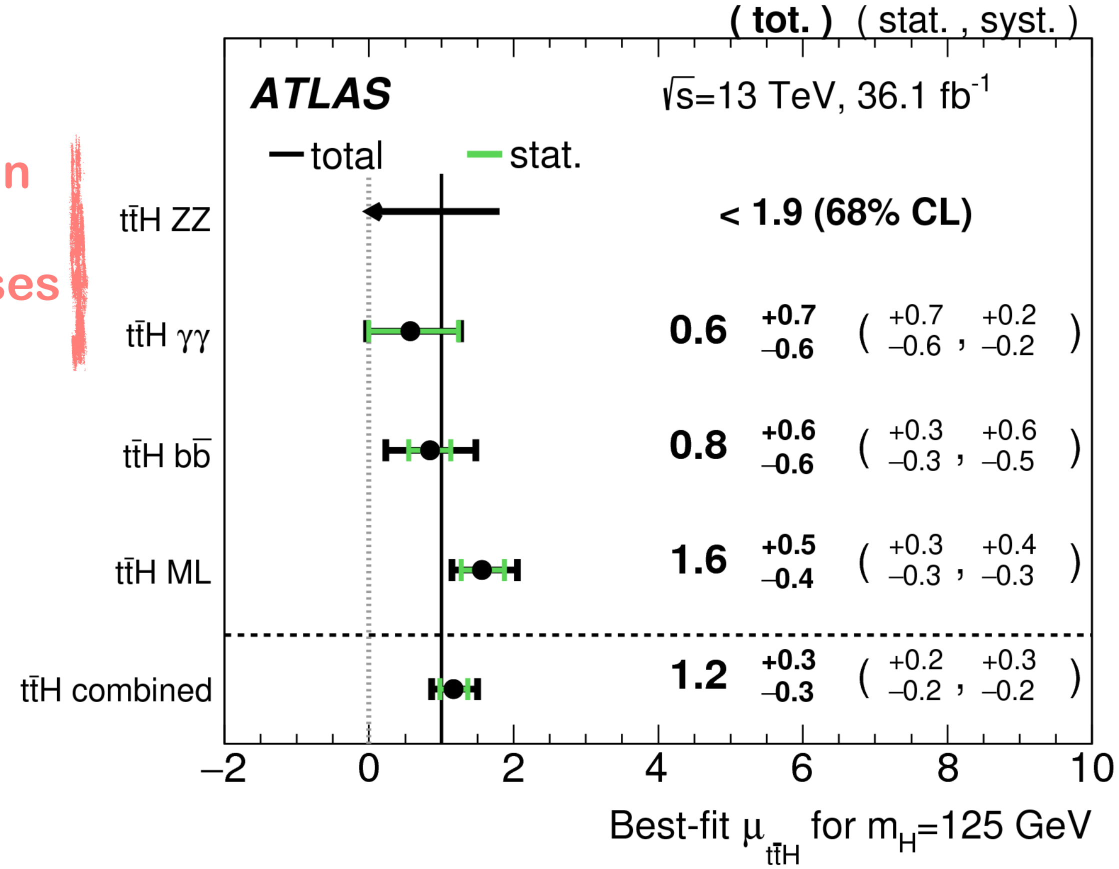
Uncertainty source	$\Delta\mu$	
$t\bar{t}+\geq 1b$ modeling	+0.46	−0.46
Background-model stat. unc.	+0.29	−0.31
b-tagging efficiency and mis-tag rates	+0.16	−0.16
Jet energy scale and resolution	+0.14	−0.14
$t\bar{t}H$ modeling	+0.22	−0.05
$t\bar{t}+\geq 1c$ modeling	+0.09	−0.11
JVT, pileup modeling	+0.03	−0.05
Other background modeling	+0.08	−0.08
$t\bar{t}$ + light modeling	+0.06	−0.03
Luminosity	+0.03	−0.02
Light lepton (e, μ) id., isolation, trigger	+0.03	−0.04
Total systematic uncertainty	+0.57	−0.54
$t\bar{t}+\geq 1b$ normalization	+0.09	−0.10
$t\bar{t}+\geq 1c$ normalization	+0.02	−0.03
Intrinsic statistical uncertainty	+0.21	−0.20
Total statistical uncertainty	+0.29	−0.29
Total uncertainty	+0.64	−0.61

- Normalisation factors for $t\bar{t}+\geq 1b$ and $t\bar{t}+\geq 1c$ left free-floating in the fit:
  - NF($t\bar{t}+\geq 1b$) = 1.24 ± 0.10
  - NF($t\bar{t}+\geq 1c$) = 1.63 ± 0.23
- Most relevant uncertainties related to $t\bar{t}+\geq 1b$ background modelling
- Analysis is **dominated by systematic** uncertainties
- Significance w.r.t background-only hypothesis: **1.4σ (1.6σ) obs (exp)**

arXiv:1712.08895 submitted to PRD

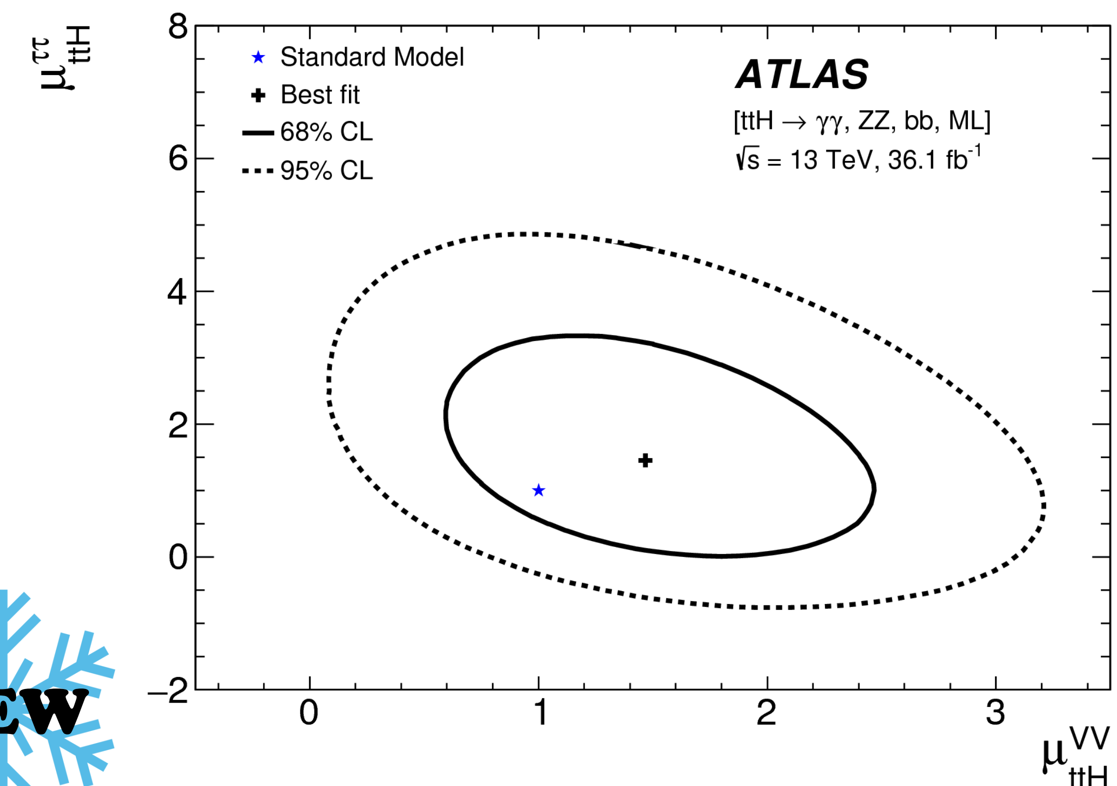
$t\bar{t}H$ combination

dedicated categories in H→γγ and H→ZZ analyses (resonant channels)



Channel	Significance	
	Observed	Expected
Multilepton	4.1σ	2.8σ
$H \to b\bar{b}$	1.4σ	1.6σ
$H \to \gamma\gamma$	0.9σ	1.7σ
$H \to 4\ell$	—	0.6σ
Combined	4.2σ	3.8σ

Bonus: measure μ_{ttH} for different decay modes



- Combination of multilepton, $b\bar{b}$, γγ, and ZZ→4ℓ $t\bar{t}$H analyses
- Results in agreement with the SM predictions
  - **σ(ttH) =590 $^{+160}_{-150}$ fb**
  - **σ_{SM}(ttH) =507 $^{+35}_{-50}$ fb**
- Significance w.r.t background-only hypothesis: **4.2σ (3.8σ) obs (exp)**
  - **Evidence for $t\bar{t}$H production!**

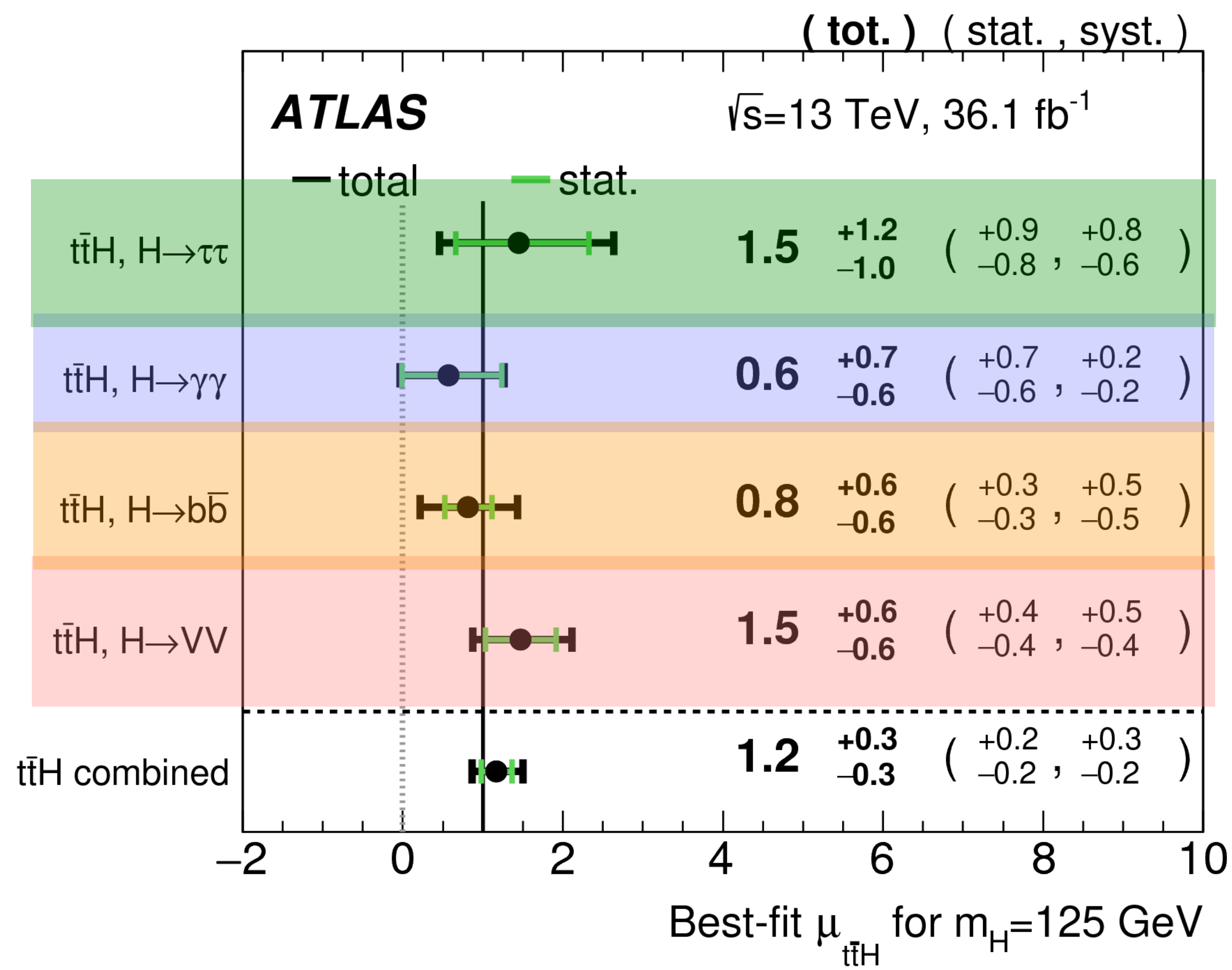
NEW

arXiv:1712.08891 submitted to PRD

Comparison with CMS (I)

- CMS claimed observation of $t\bar{t}H$ combining Run1+Run2 [2015+2016 dataset]
  - Significance = 5.2 (**4.2**) σ observed (expected)
- How different are the sensitivities?

Parameter	Best fit	Stat
$\mu_{t\bar{t}H}^{WW^*}$	$1.97^{+0.71}_{-0.64}$ $\left(^{+0.57}_{-0.54}\right)$	$^{+0.42}_{-0.41}$ $\left(^{+0.39}_{-0.38}\right)$
$\mu_{t\bar{t}H}^{ZZ^*}$	$0.00^{+1.30}_{-0.00}$ $\left(^{+2.89}_{-0.99}\right)$	$^{+1.28}_{-0.00}$ $\left(^{+2.82}_{-0.99}\right)$
$\mu_{t\bar{t}H}^{\gamma\gamma}$	$2.27^{+0.86}_{-0.74}$ $\left(^{+0.73}_{-0.64}\right)$	$^{+0.80}_{-0.72}$ $\left(^{+0.71}_{-0.64}\right)$
$\mu_{t\bar{t}H}^{\tau^+\tau^-}$	$0.28^{+1.09}_{-0.96}$ $\left(^{+1.00}_{-0.89}\right)$	$^{+0.86}_{-0.77}$ $\left(^{+0.83}_{-0.76}\right)$
$\mu_{t\bar{t}H}^{b\bar{b}}$	$0.82^{+0.44}_{-0.42}$ $\left(^{+0.44}_{-0.42}\right)$	$^{+0.23}_{-0.23}$ $\left(^{+0.23}_{-0.22}\right)$
$\mu_{t\bar{t}H}^{7+8\,\text{TeV}}$	$2.59^{+1.01}_{-0.88}$ $\left(^{+0.87}_{-0.79}\right)$	$^{+0.54}_{-0.53}$ $\left(^{+0.51}_{-0.49}\right)$
$\mu_{t\bar{t}H}^{13\,\text{TeV}}$	$1.14^{+0.31}_{-0.27}$ $\left(^{+0.29}_{-0.26}\right)$	$^{+0.17}_{-0.16}$ $\left(^{+0.16}_{-0.16}\right)$
$\mu_{t\bar{t}H}$	$1.26^{+0.31}_{-0.26}$ $\left(^{+0.28}_{-0.25}\right)$	$^{+0.16}_{-0.16}$ $\left(^{+0.15}_{-0.15}\right)$



- All channels comparable
  - Except $t\bar{t}H$ ($H\to b\bar{b}$), CMS ~33% smaller uncertainty!

Tamara Vázquez Schröder (McGill University)

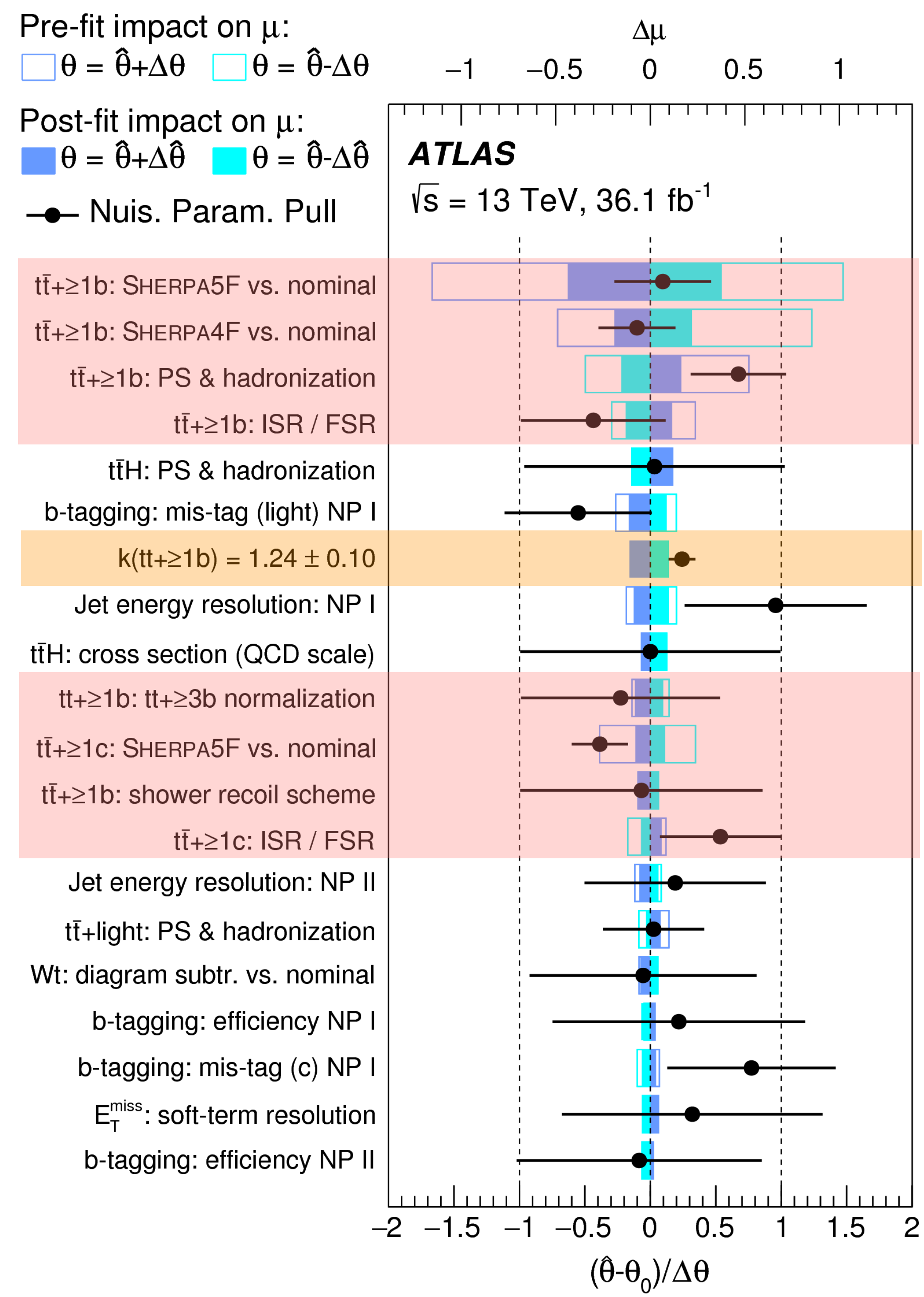
Comparison with CMS (II)

- Stat-only CMS uncertainty slightly better: use full b-tagging discriminant shape, deep neural networks in single ℓ channel

- **Largest** difference coming from **systematic** uncertainties:

Modelling uncertainties on $t\bar{t}+\geq 1b$ from comparison of different generators, parton showers, … **largest impact** on uncertainty on signal strength and **only present in ATLAS**!

CMS **only** includes renormalisation/ factorisation scale and PDF variations **shape** uncertainties for $t\bar{t}$



Open questions to CMS…

- Which systematic increase (decrease) the $t\bar{t}+\geq 1c/b$ ($t\bar{t}$+light) yields?
- What are the pulls that decrease $t\bar{t}$+light?
- Are $t\bar{t}+\geq 1c/b$ yields used to correct the shapes?

Conclusions

* New results presented for $t\bar{t}H$ production search in ATLAS with 36.1 fb^{-1}

* Challenging and complex final states
  - Sensitivity driven by usage of multivariate analysis techniques and precise background modelling

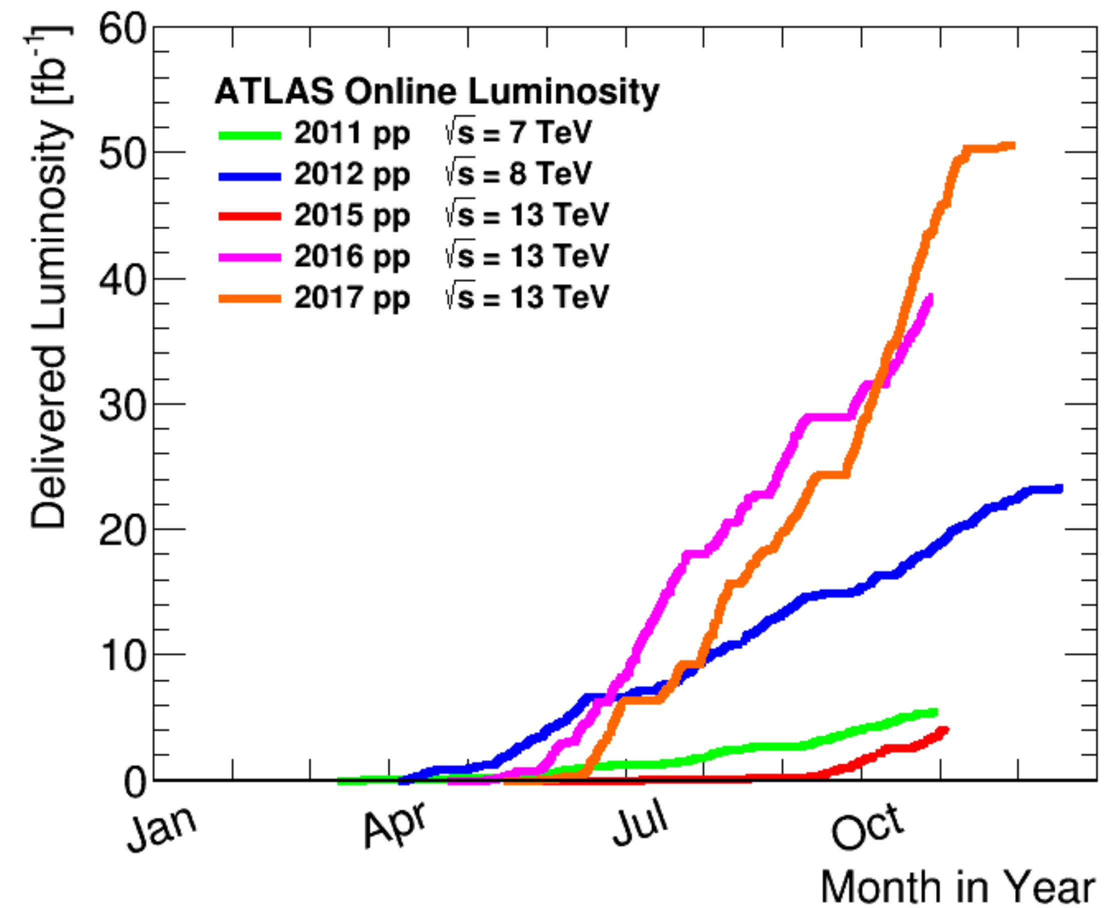
* Four channels are better than one!
  - **Evidence for $t\bar{t}H$ production** when combining all available channels
    - **4.2σ (3.8σ) obs (exp) significance**
  - Results consistent with the SM predictions
    - **σ(ttH) $=590^{+160}_{-150}$ fb** σ_{SM}(ttH) $=507^{+35}_{-50}$ fb

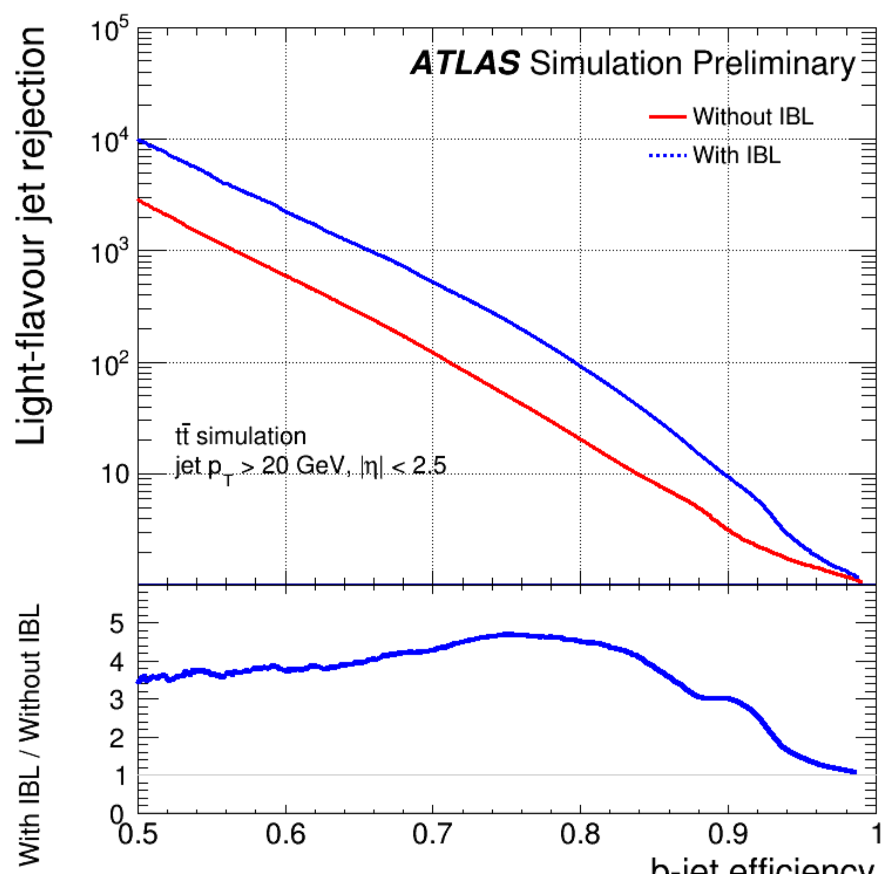
* 2017 pp collisions data on tape: more than what has been analysed so far!
  - **Stay tuned for updated results!**

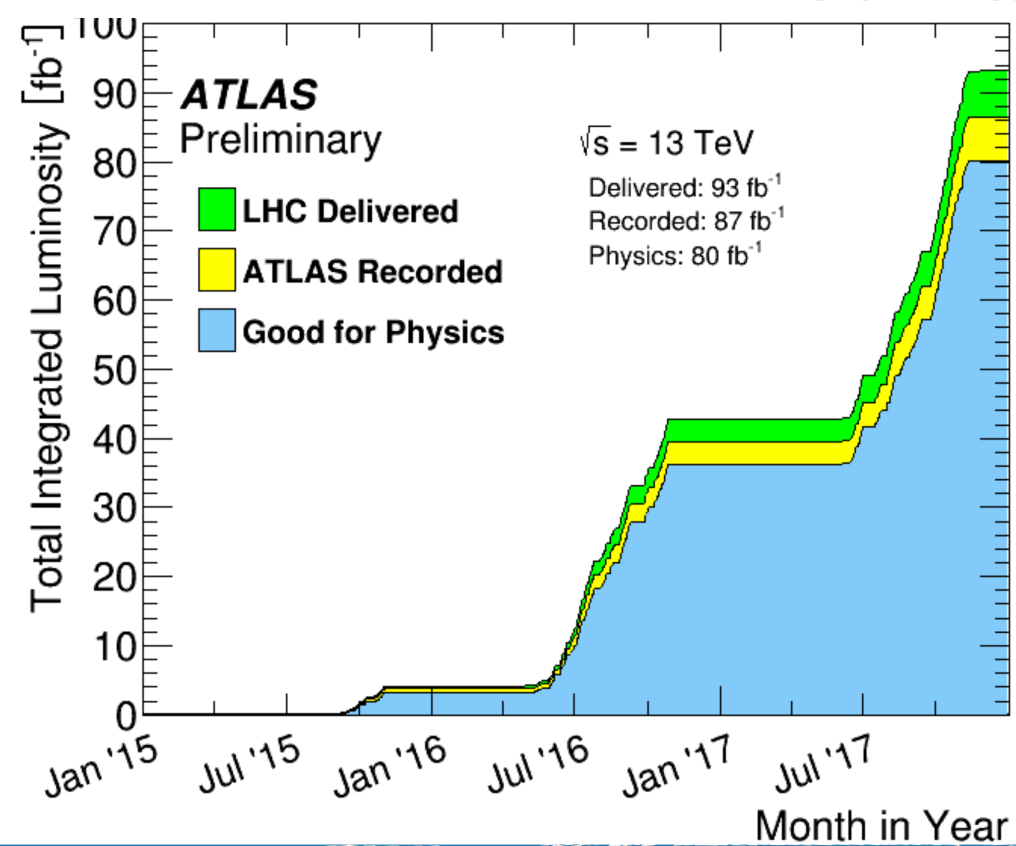
Back-up slides

Detector performance (I)

- **Excellent performance of LHC and ATLAS in Run 2 so far**:
  - Record instantaneous luminosity for pp interactions in 2017: $2.06 \times 10^{34}\ cm^{-2}s^{-1}$
    - **double** the LHC design!
  - 80 fb^{-1} good for physics from 87 fb^{-1} recorded by ATLAS
- **Improved b-tagging** performance with the inclusion of **IBL (Insertable B-Layer)** for Run 2



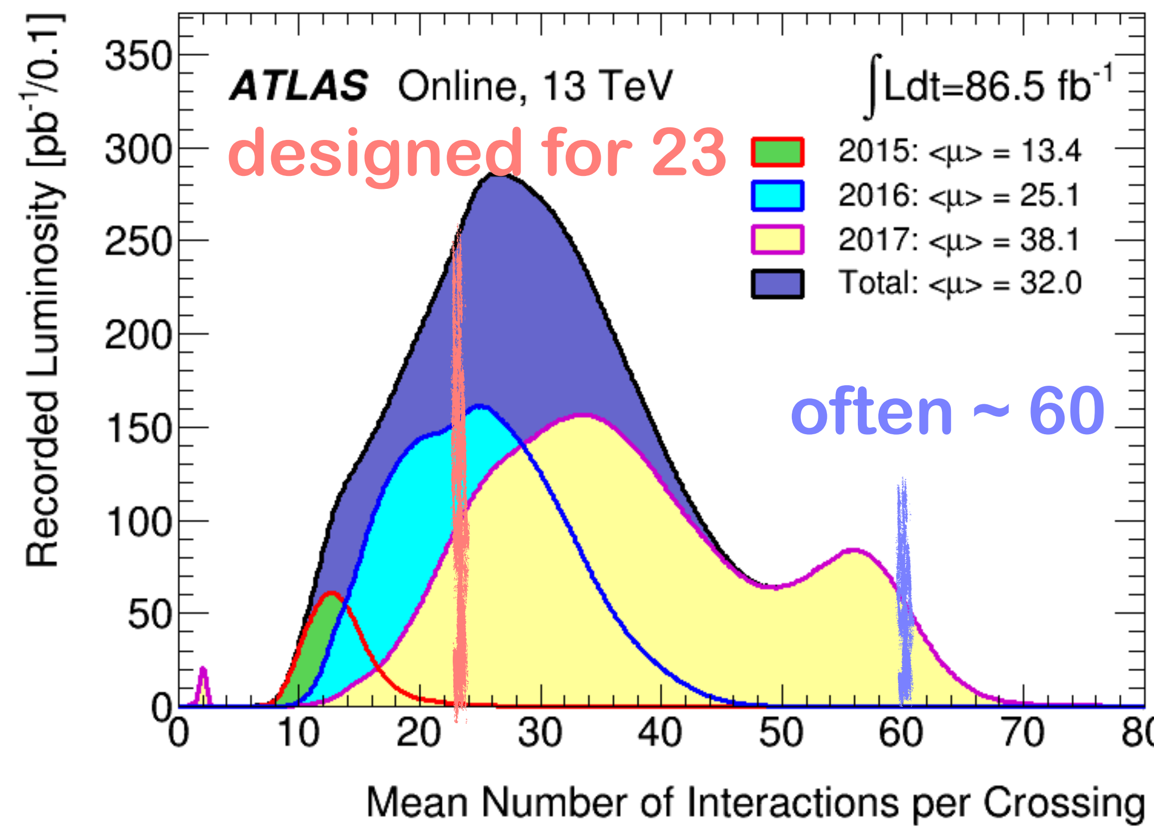




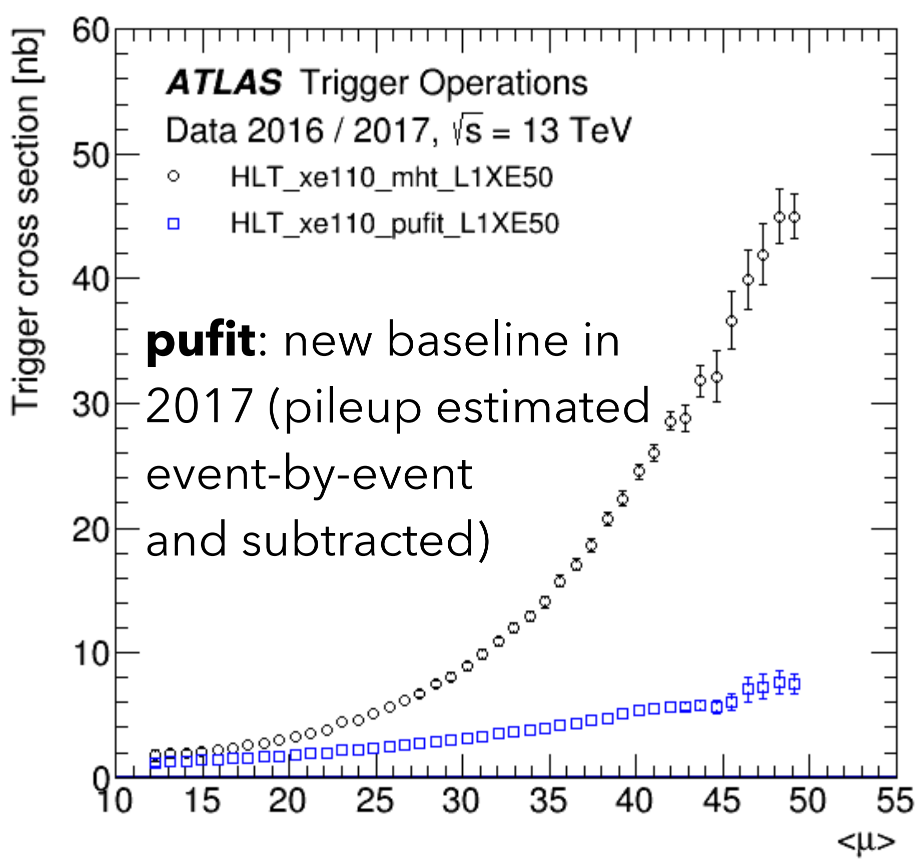
http://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/PLOTS/FTAG-2017-005/

Detector performance (II)

- **Biggest challenge**: robustness against pile-up
  - Controlling trigger rates at high interaction per bunch crossing
  - Online and offline reconstruction performance maintained even at the highest pile-up

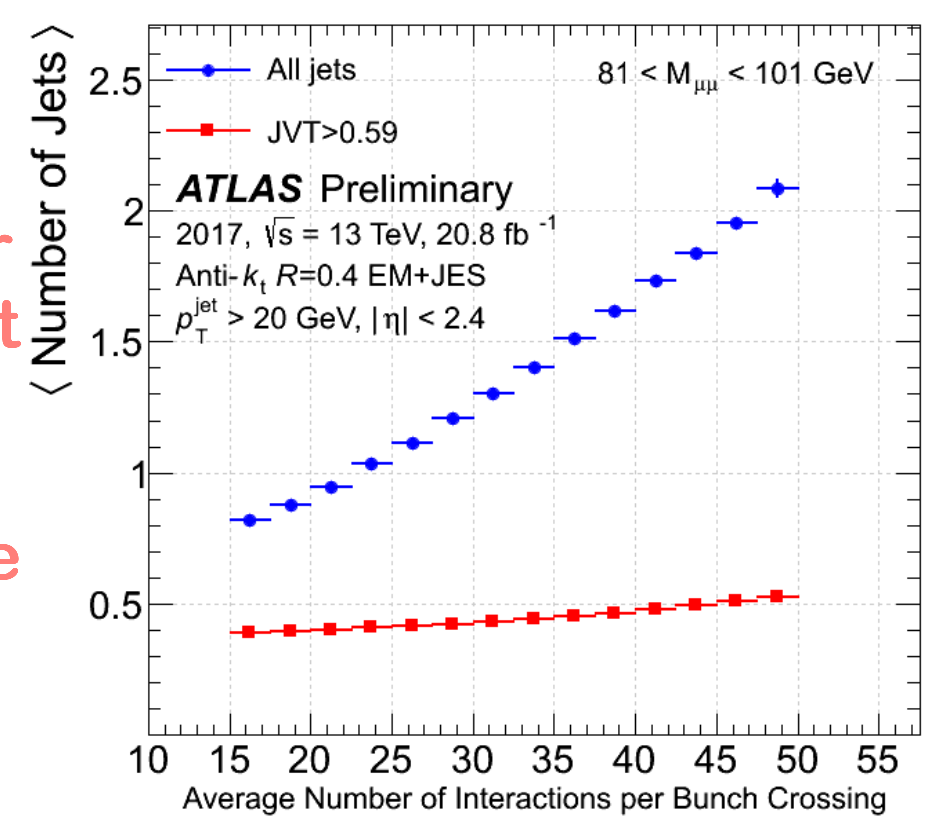


improved HLT algorithms to suppress dependence of E_T^{miss} trigger rates on pile-up



pufit: new baseline in 2017 (pileup estimated event-by-event and subtracted)

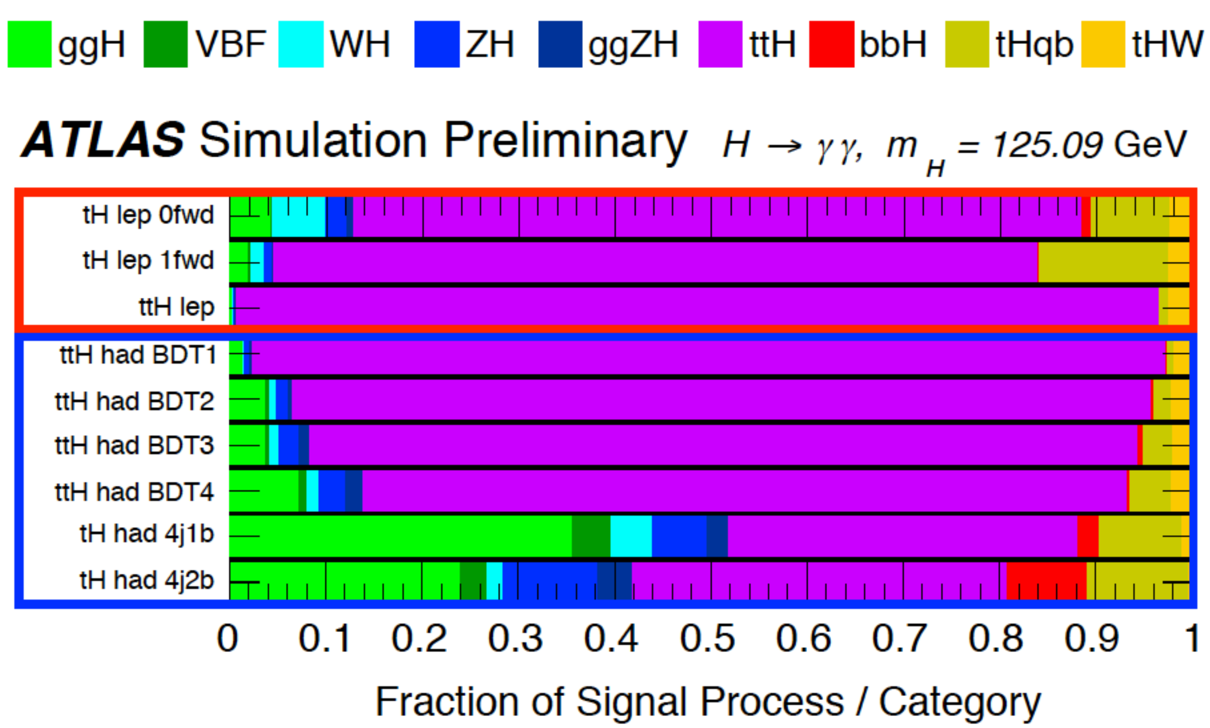
Jet Vertex Tagger (JVT) requirement applied to jets with $20<p_T<60$ GeV to reduce the fake jet rate

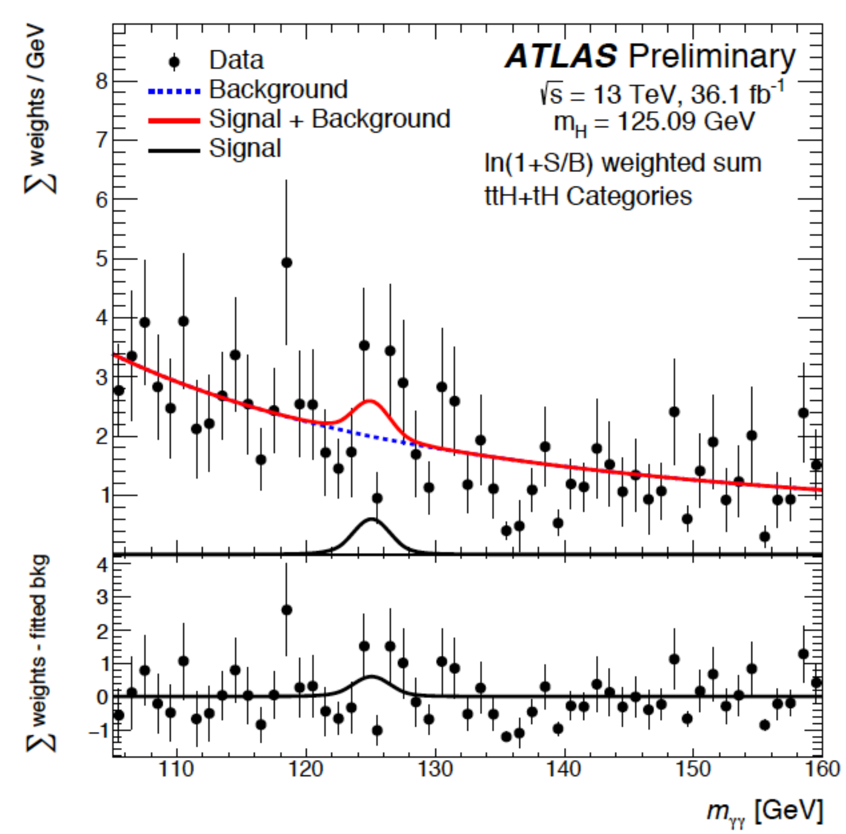


Tamara Vázquez Schröder

$t\bar{t}H(H\rightarrow\gamma\gamma)$

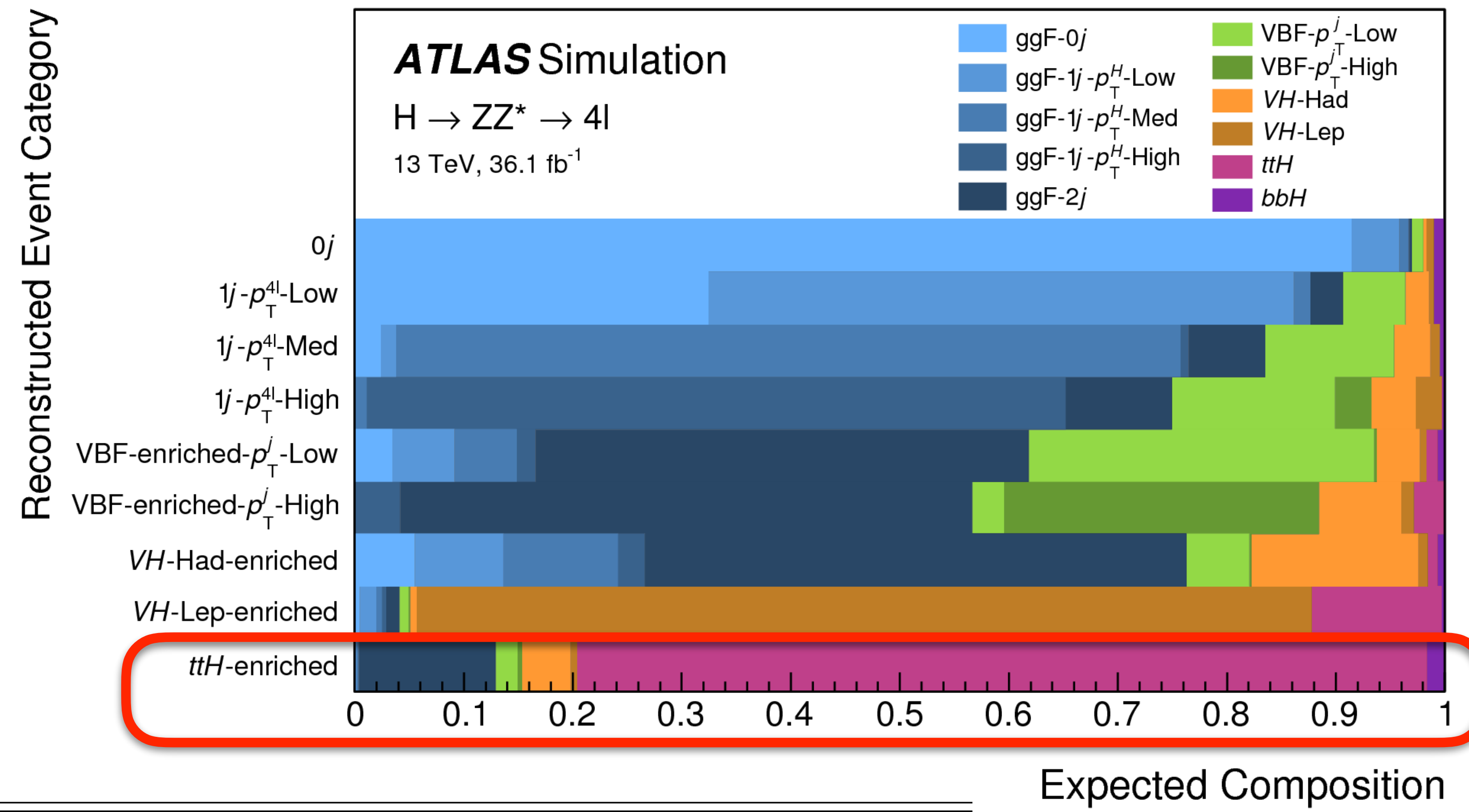
- Several categories optimised for ttH, tHqb and tHW, with 0-≥1ℓ from $t\bar{t}$ decays
- **Leptonic channel:**
  - One category for ttH
  - Two categories for tH (with/without forward jets)
- **Hadronic channel:**
  - ≥3 jets, ≥1 b-tagged jets: BDT to identify ttH against ggH and multijet background
  - 4 central jets, 1/≥2 b-tagged jets





$t\bar{t}H(H\to ZZ\to 4\ell)$ resonant

- Higgs boson candidates with $118 < m(4\ell) < 129$ GeV
- $t\bar{t}H$ enriched category:
  - ≥1 b-tagged jet
  - ≥4 jets or 1ℓ + ≥2 jets
- No events observed → Upper limits on $t\bar{t}H$



Reconstructed event category	Signal	ZZ^* background	Other backgrounds	Total expected	Observed
$0j$	26.8 ± 2.5	13.7 ± 1.0	2.23 ± 0.31	42.7 ± 2.7	49
$1j$-$p_{\mathrm{T}}^{4\ell}$-Low	8.8 ± 1.1	3.1 ± 0.4	0.53 ± 0.07	12.5 ± 1.2	12
$1j$-$p_{\mathrm{T}}^{4\ell}$-Med	5.4 ± 0.7	0.88 ± 0.12	0.38 ± 0.05	6.7 ± 0.7	9
$1j$-$p_{\mathrm{T}}^{4\ell}$-High	1.47 ± 0.24	0.139 ± 0.022	0.045 ± 0.007	1.65 ± 0.24	3
VBF-enriched-p_{T}^{j}-Low	6.3 ± 0.8	1.08 ± 0.32	0.40 ± 0.04	7.7 ± 0.9	16
VBF-enriched-p_{T}^{j}-High	0.58 ± 0.10	0.093 ± 0.032	0.054 ± 0.006	0.72 ± 0.10	3
VH-Had-enriched-$p_{\mathrm{T}}^{4\ell}$-Low	2.9 ± 0.5	0.63 ± 0.16	0.169 ± 0.021	3.7 ± 0.5	3
VH-Had-enriched-$p_{\mathrm{T}}^{4\ell}$-High	0.64 ± 0.09	0.029 ± 0.008	0.0182 ± 0.0022	0.69 ± 0.09	0
VH-Lep-enriched	0.318 ± 0.019	0.049 ± 0.008	0.0137 ± 0.0019	0.380 ± 0.020	0
ttH-enriched	0.39 ± 0.04	0.014 ± 0.006	0.07 ± 0.04	0.47 ± 0.05	0
Total	54 ± 4	19.7 ± 1.5	3.9 ± 0.5	77 ± 4	95

$t\bar{t}H$ (multileptons): non-prompt light ℓ (II)

- **2ℓSS/3ℓ+0τ:** Matrix Method

events in pre-MVA signal region with SS **loose** leptons
(in 3ℓ, lep_0 (OS to SS pair) is prompt in 98% of the times)

$$N^f_{TT} = w_{TT}N^{TT} + w_{T\not{T}}N^{T\not{T}} + w_{\not{T}T}N^{\not{T}T} + w_{\not{T}\not{T}}N^{\not{T}\not{T}}$$

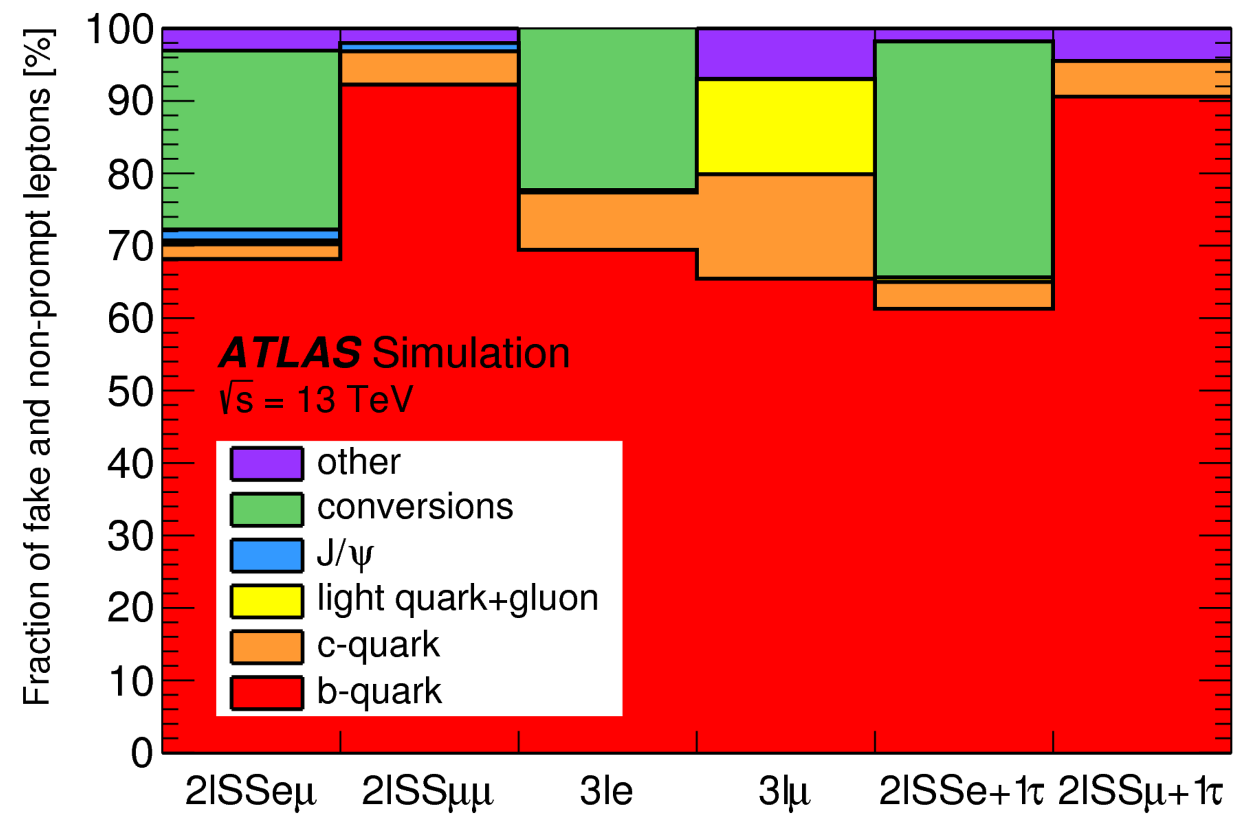
f(εr, εf)

via tag&probe method in $t\bar{t}$ events

Channel	Region	Selection criteria
2ℓSS		$2 \leq N_{\text{jets}} \leq 3$ and $N_{b\text{-jets}} \geq 1$
(3ℓ)		One tight, one loose light lepton with $p_{\text{T}} > 20$ (15) GeV
		Zero τ_{had} candidates
	ϵ_{real}	Opposite charge, opposite flavour
	ϵ_{fake}	Same charge and $e\mu$ or $\mu\mu$

- **electrons ε_f**: 2D ($N_{b\text{-tags}}$, p_T) parametrisation
  - more conversions in 2bj than 1bj
- **muons ε_f**: 2D (minΔR(μ,j), p_T) parametrisation

- Fakes composition in the **CR** is ~ representative of the fakes in **pre-MVA SR**



- **4ℓ**: correct MC with **3 fake factors** (λ^{el}_{heavy}, λ^{el}_{light}, λ^{μ}_{all}) derived from fit to data in 4 CRs
- **2ℓSS+1τ**: fake factor [pT], estimates the fakes originating from $t\bar{t}$

$t\bar{t}H$ (multileptons): fake tau (II)

2ℓOS+1τ: fake factor

FF method	τ	anti-τ
apply in B:	**A**	B
extract FF:	C	D

Regions		nJets	nBJets	Z cut	Used for
A, B	2LOS+tau selection	≥ 3	≥ 1	$\lvert M_{ee/\mu\mu} - M_Z \rvert > 10$ GeV	to be estimated
C, D	ZVeto 3j0b	≥ 3	0	$\lvert M_{ee/\mu\mu} - M_Z \rvert > 10$ GeV	nominal FF
C, D	OnZ 3j0b	≥ 3	0	$\lvert M_{ee/\mu\mu} - M_Z \rvert < 10$ GeV	systematics (Z+jets enriched)
C, D	exc2j1b	2	≥ 1	$\lvert M_{ee/\mu\mu} - M_Z \rvert > 10$ GeV	systematics (ttbar enriched)

- τ/anti-τ definition based on BDT score of jet-vs-τ
- Reasonable agreement of yield and shape of DD estimate with data

2ℓSS+1τ and 3ℓ+1τ: MC correction with SF derived from **{DD(2ℓOS+1τ) / MC}**

- Harmonised 1-fake-τ estimate for all channels, profit from large statistics from 2ℓOS+1τ, composition uncertainties to cover wide range of b-faking-τ content
- Final SF = 1.36 ± 0.16

Fakes/Non-prompt ℓ: uncertainties

Systematic uncertainty	Values	Type	Comp
Data-driven fake leptons and electron charge mis-assignment			
Control region statistics		SN	21
Real lepton efficiencies		SN	1
Fake lepton rate		SN	6 (μ), 2 (e), 3 (bkg sub.)
Non-promt lepton estimate: non-closure	Tab. 15	N	4
γ-conversion fraction	Tab. 14	N	4
Electron charge mis-assignment		SN	1
4l fake lepton rate		SN	1
TOTAL (Data driven reducible background		–	43

Correlation across channels:

- (blue) Correlated between 2ℓSS+0τ and 3ℓ+0τ
- (dark green) Correlated between 2ℓSS+1τ and 3ℓ+1τ
- (light green) Correlated between 2ℓSS+1τ, 3ℓ+1τ, and 2ℓOS+1τ

Systematic uncertainty	Type	Comp
Data-driven fake taus		
$1\ell+2\tau_{had}$ SS data: CR statistics	SN	2
$1\ell+2\tau_{had}$ SS data: non-closure	N	1
$1\ell+2\tau_{had}$ SS data: shape	S	1
2ℓOS$+1\tau_{had}$ FF: CR statistics	SN	10
2ℓOS$+1\tau_{had}$ FF: statistics	SN	5
2ℓOS$+1\tau_{had}$ FF: real tau sub	SN	1
2ℓSS$+1\tau_{had}$ and $3\ell+1\tau_{had}$ SF correction: statistics	N	1
2ℓSS$+1\tau_{had}$ fake lepton rate	SN	10
2ℓSS$+1\tau_{had}$ QmisID: CR statistics	SN	2
2ℓSS$+1\tau_{had}$ bkg sub: statistics	SN	2
2ℓSS$+1\tau_{had}$ γ-conversion fraction	N	1
tau had fakes: composition	SN	2
$3\ell+1\tau_{had}$ SF correction: MC statistics	N	1
TOTAL	–	40

MC samples

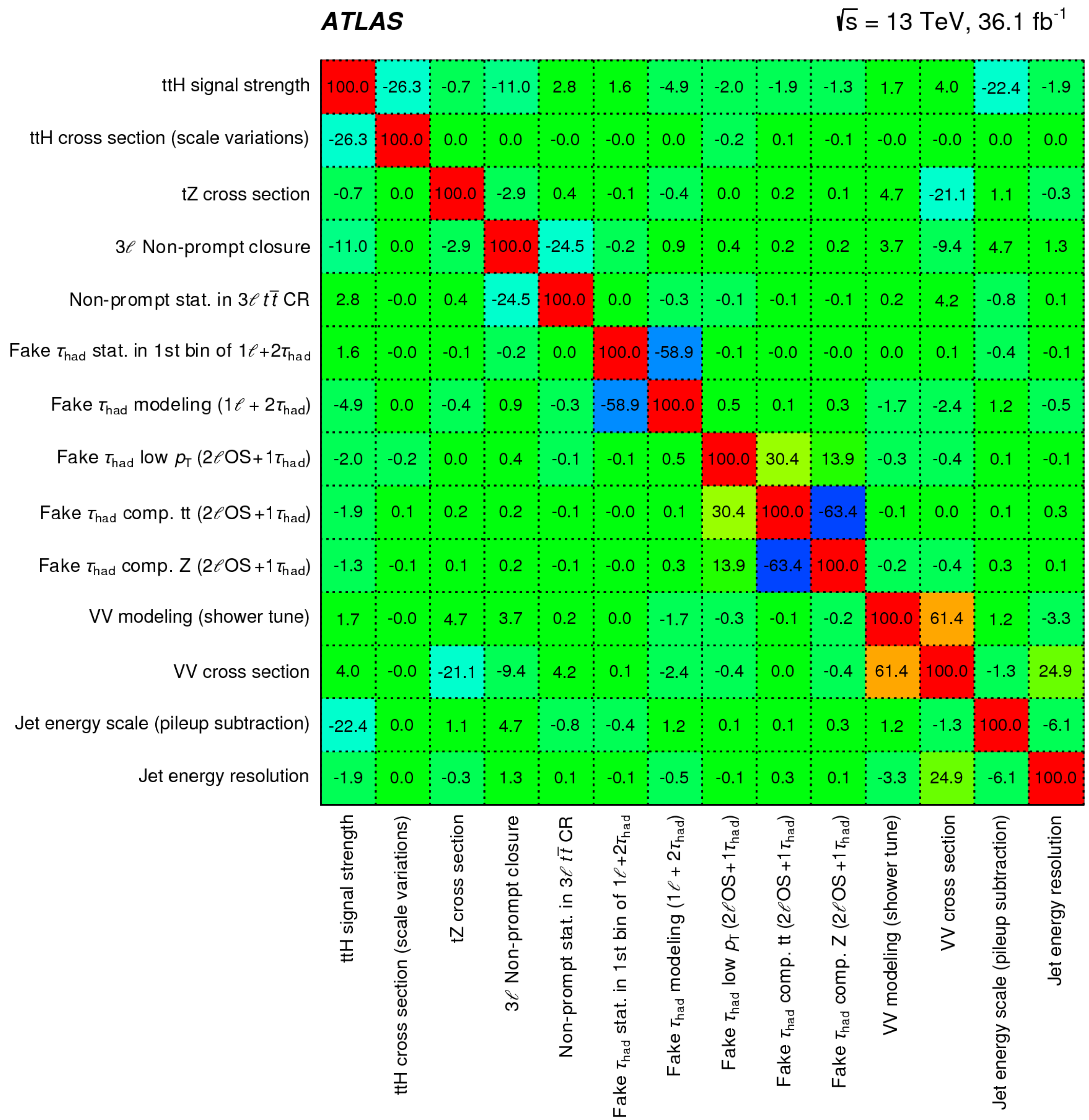
Process	Event generator	ME order	Parton Shower	PDF	Tune
$t\bar{t}H$	MG5_aMC	NLO	Pythia 8	NNPDF 3.0 NLO [71]	A14
	(MG5_aMC)	(NLO)	(Herwig++)	(CT10 [72])	(UE-EE-5)
$tHqb$	MG5_aMC	LO	Pythia 8	CT10	A14
tHW	MG5_aMC	NLO	Herwig++	CT10	UE-EE-5
$t\bar{t}W$	MG5_aMC	NLO	Pythia 8	NNPDF 3.0 NLO	A14
	(Sherpa 2.1.1)	(LO multileg)	(Sherpa)	(NNPDF 3.0 NLO)	(Sherpa default)
$t\bar{t}(Z/\gamma^* \to ll)$	MG5_aMC	NLO	Pythia 8	NNPDF 3.0 NLO	A14
	(Sherpa 2.1.1)	(LO multileg)	(Sherpa)	(NNPDF 3.0 NLO)	(Sherpa default)
tZ	MG5_aMC	LO	Pythia 6	CTEQ6L1	Perugia2012
tWZ	MG5_aMC	NLO	Pythia 8	NNPDF 2.3 LO	A14
$t\bar{t}t$, $t\bar{t}t\bar{t}$	MG5_aMC	LO	Pythia 8	NNPDF 2.3 LO	A14
$t\bar{t}W^+W^-$	MG5_aMC	LO	Pythia 8	NNPDF 2.3 LO	A14
$t\bar{t}$	Powheg-BOX v2 [73]	NLO	Pythia 8	NNPDF 3.0 NLO	A14
$t\bar{t}\gamma$	MG5_aMC	LO	Pythia 8	NNPDF 2.3 LO	A14
s-, t-channel, Wt single top	Powheg-BOX v1 [74,75,76]	NLO	Pythia 6	CT10	Perugia2012
$VV(\to llXX)$, $qqVV$, VVV	Sherpa 2.1.1	MEPS NLO	Sherpa	CT10	Sherpa default
$Z \to l^+l^-$	Sherpa 2.2.1	MEPS NLO	Sherpa	NNPDF 3.0 NLO	Sherpa default

Input variables for MVA

	Variable	$2\ell SS$	3ℓ	4ℓ	$1\ell+2\tau_{\mathrm{had}}$	$2\ell SS+1\tau_{\mathrm{had}}$	$2\ell OS+1\tau_{\mathrm{had}}$
Lepton properties	Leading lepton p_{T}		×				
	Second leading lepton p_{T}	×	×			×	
	Third lepton p_{T}		×				
	Dilepton invariant mass (all combinations)	×	×*				×
	Three-lepton invariant mass		×				
	Four-lepton invariant mass			×			
	Best Z-candidate dilepton invariant mass			×			
	Other Z-candidate dilepton invariant mass			×			
	Scalar sum of all leptons p_{T}			×			×
	Second leading lepton track isolation					×	
	Maximum $\lvert\eta\rvert$ (lepton 0, lepton 1)	×				×*	
	Lepton flavor	×*	×*				
	Lepton charge		×				
Jet properties	Number of jets	×*	×*		×	×	×
	Number of b-tagged jets	×*	×*		×	×	×
	Leading jet p_{T}						×
	Second leading jet p_{T}		×			×*	
	Leading b-tagged jet p_{T}		×				
	Scalar sum of all jets p_{T}		×		×	×	×
	Scalar sum of all b-tagged jets p_{T}						×
	Has leading jet highest b-tagging weight?		×				
	b-tagging weight of leading jet		×				
	b-tagging weight of second leading jet		×			×	
	b-tagging weight of third leading jet					×	
	Pseudorapidity of fourth leading jet					×	
τ_{had}	Leading τ_{had} p_{T}				×		×
	Second leading τ_{had} p_{T}				×		
	Di-τ_{had} invariant mass				×		
	Invariant mass τ_{had}–furthest lepton					×	
Angular distances	ΔR(lepton 0, lepton 1)		×				
	ΔR(lepton 0, lepton 2)		×				
	ΔR(lepton 0, closest jet)	×	×				
	ΔR(lepton 0, leading jet)		×			×	
	ΔR(lepton 0, closest b-jet)		×				
	ΔR(lepton 1, closest jet)	×	×				
	ΔR(lepton 2, closest jet)		×				
	Smallest ΔR(lepton, jet)		×				×
	Smallest ΔR(lepton, b-tagged jet)						×
	Smallest ΔR(non-tagged jet, b-tagged jet)						×
	ΔR(lepton 0, τ_{had})						×
	ΔR(lepton 1, τ_{had})						×
	Minimum ΔR between all jets				×		
	ΔR between two leading jets					×	
$\vec{p}_{\mathrm{T}}^{\,\mathrm{miss}}$	Missing transverse momentum $E_{\mathrm{T}}^{\mathrm{miss}}$	×		×			
	Azimuthal separation $\Delta\phi$(leading jet, $\vec{p}_{\mathrm{T}}^{\,\mathrm{miss}}$)		×				
	Transverse mass leptons (H/Z decay) - $\vec{p}_{\mathrm{T}}^{\,\mathrm{miss}}$			×			
	Pseudo-Matrix-Element			×			

The variables used in cross-check analyses are indicated by a $*$

Correlation NPs



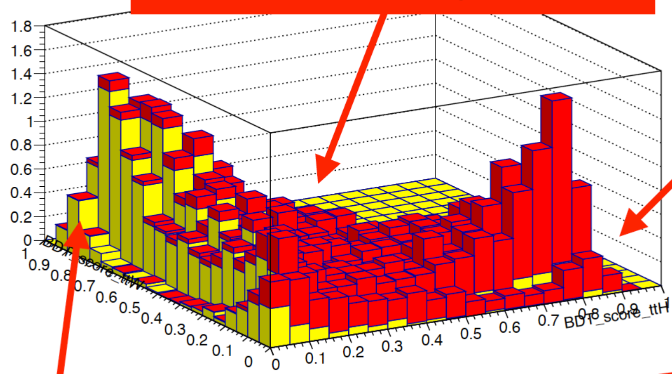
Correlation min threshold = 20%

Multinomial classification

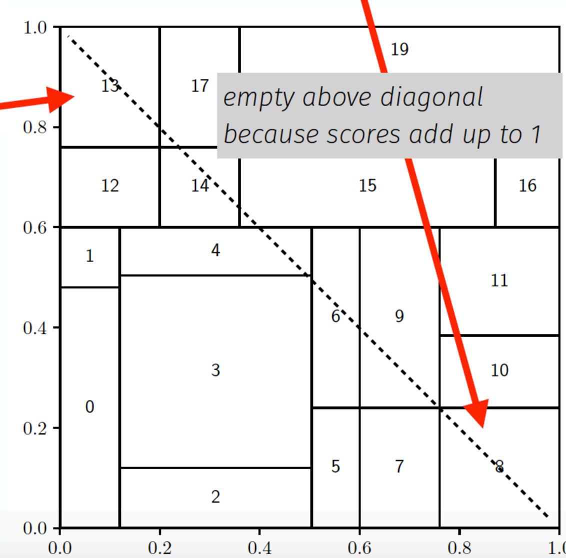
- Explore multinomial classifiers to simultaneously define signal and control regions
  - Processes are separated in the space of a multiD observable
  - Define CRs and VRs with a topology similar to the SR

"Rest" (not ttW, not ttH)

ttH like



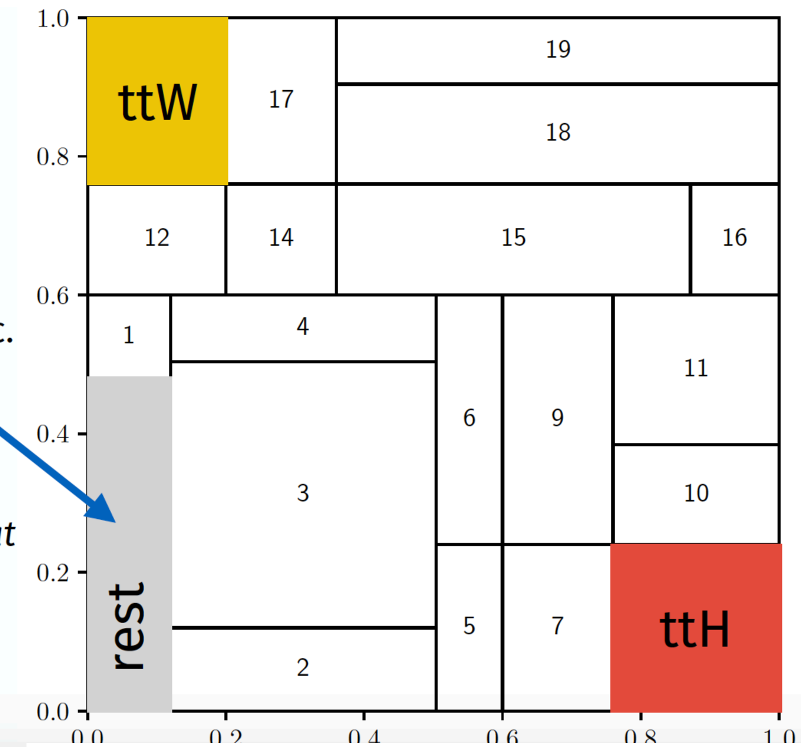
ttW like

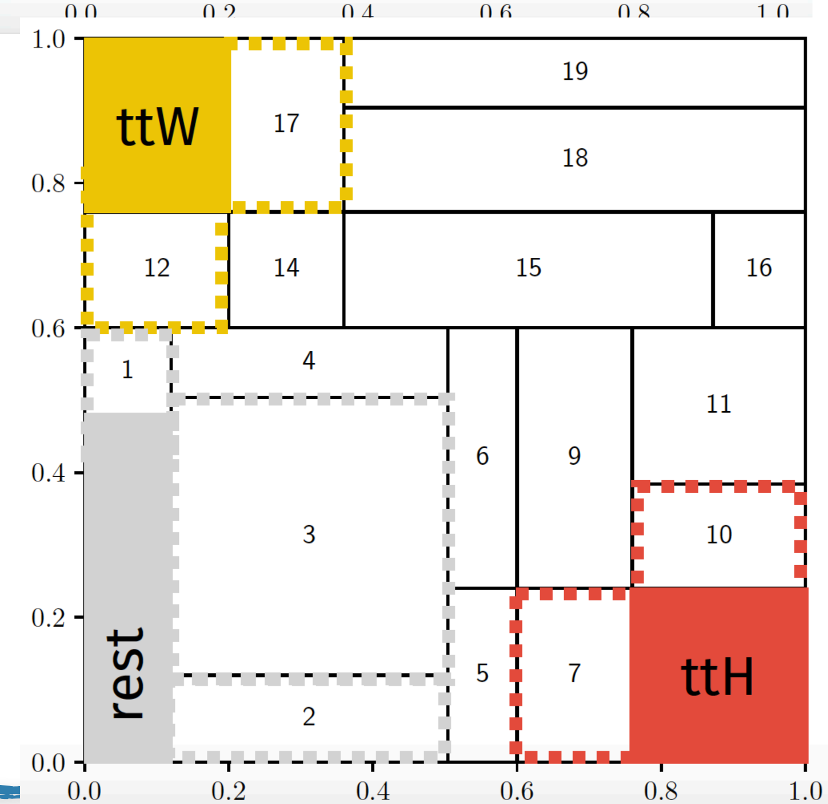


- Approximate the n-dimensional space of the BDT output scores by bins of variable size
- In practice, start with 1 large bin and split it according to the density

"rare" processes, etc.

Rest groups bins that do not contribute to ttH and ttW and speeds up the algorithm

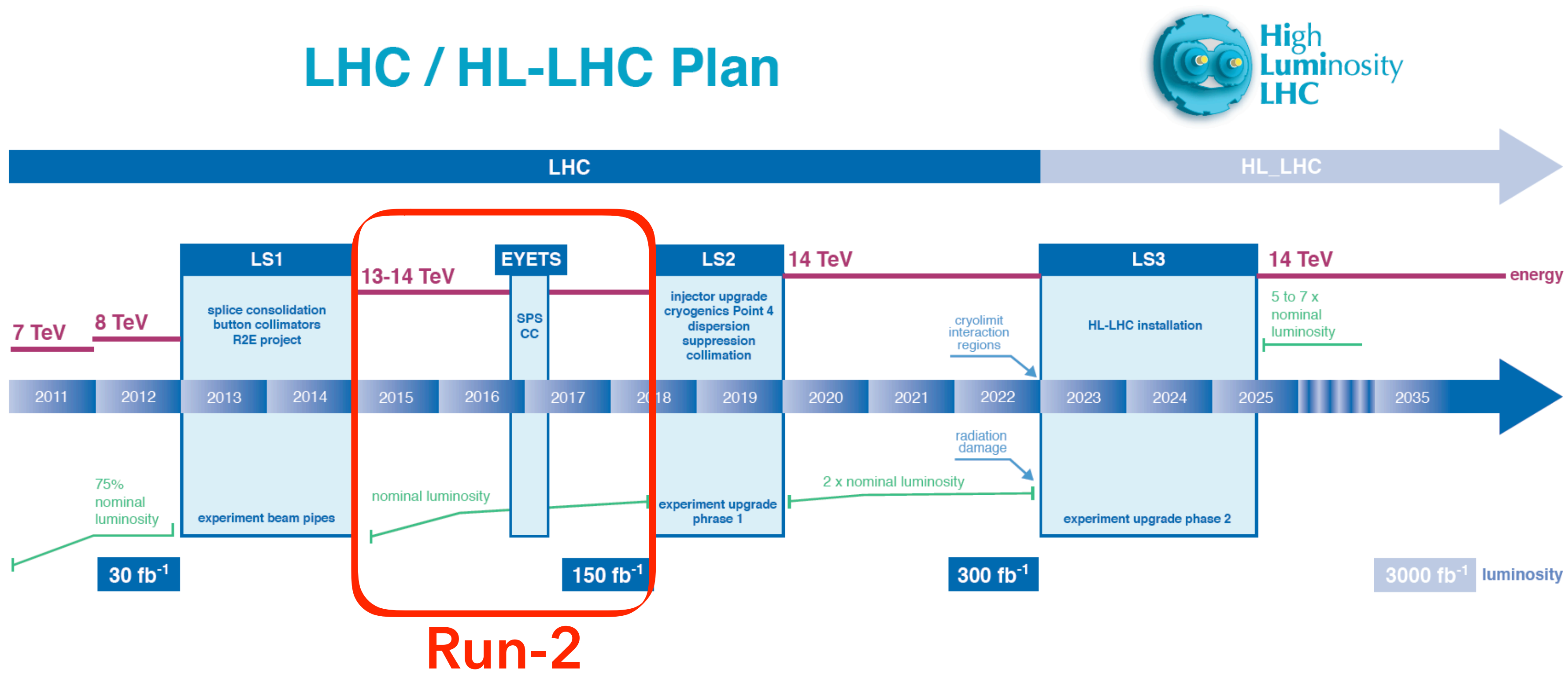




- Clustering: add a single neighbouring bin to the seed and compute analytically the significance again; add the cell giving the largest improvement

$t\bar{t}H$: new ideas for Run-2 and beyond

- **$t\bar{t}H(H\rightarrow c\bar{c})$:**
  - BR ~ 3 %, $t\bar{t}H$ is very difficult production search since also has 2 bjets from $t\bar{t}$ (it would need both b- and c-taggers)
- **$t\bar{t}H/t\bar{t}Z$** to measure top Yukawa coupling
  - ttH theory systematics 3rd leading in ttH(bb) and $t\bar{t}H$(ML)!
  - $t\bar{t}H$ and $t\bar{t}Z$: **identical production mechanisms** + **mZ~mH** → **correlated** QCD corrections, scale dependence, alpha_S dependence, and PDF systematics
  - For a given y_{top}, σ(ttH)/σ(ttZ) can be predicted theoretically with a **much better precision**
- **CP mixture states of the Higgs**



CP nature of ttH coupling

- Program to probe CP nature of the discovered Higgs already underway (e.g. H→ZZ)
- However, Higgs-fermion couplings provide **more sensitive probe of a CP-mixed state** than Higgs-bosons
- $t\bar{t}$H production **most direct probe** of the **Higgs-top coupling** and the **Higgs CP nature**
- Parametrise Higgs coupling to fermions:
  - **Scalar** (a_f=1, b_f=0)
  - **Pseudo-scalar** (a_f=0, b_f=1)
  - **Mixed** CP properties ($a_f \neq 0$, $b_f \neq 0$)

$$\mathcal{L}_{hf\bar{f}} = -\sum_f \frac{m_f}{v} h\bar{f}(a_f + i b_f \gamma_5) f$$

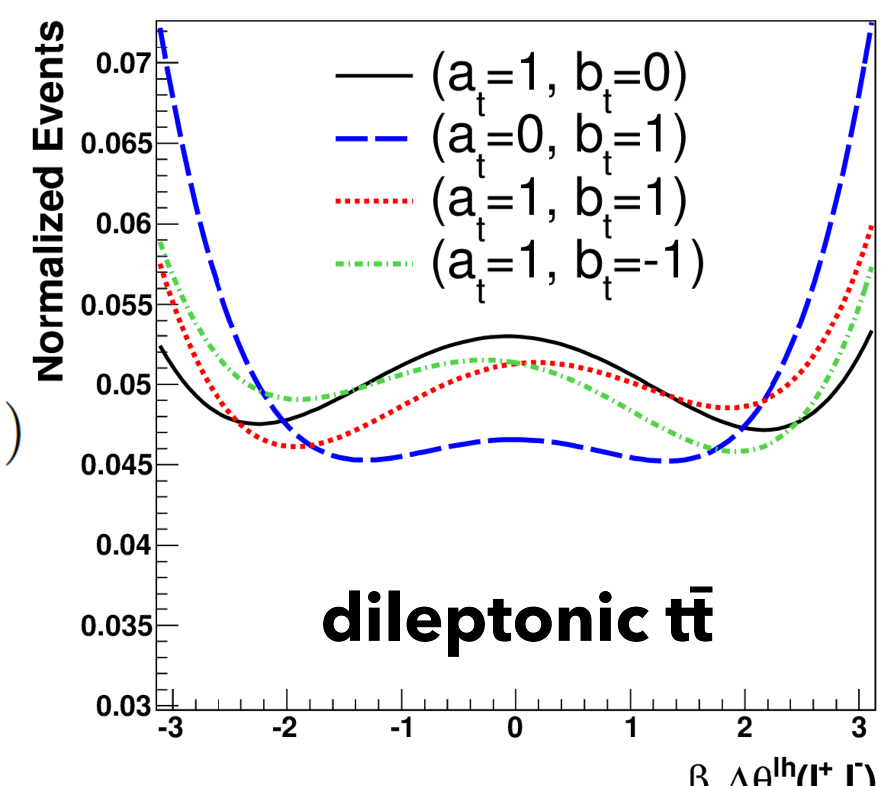
- A measurement of the $t\bar{t}$H production alone is not sufficient to determine the vertex (i.e. does not provide information about the **sign** of b_t)
- Discriminating observables based on the kinematics of the $t\bar{t}$H production to probe the nature of the Higgs-top coupling: $M_{t\bar{t}H}$, p_T^H, and $\Delta\Phi(t,\bar{t})$

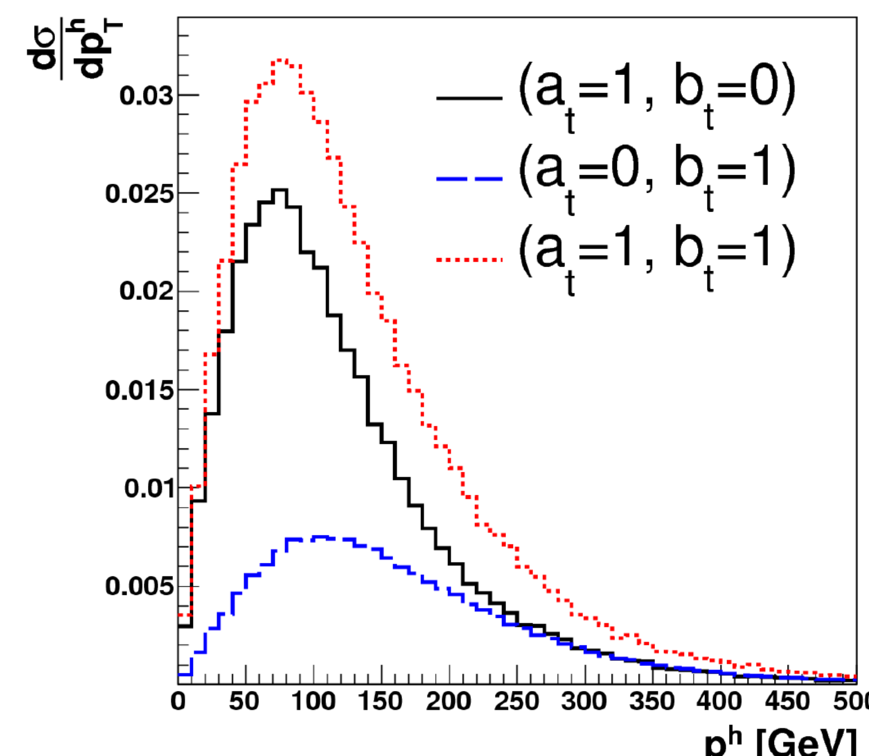
arXiv:1501.03157

- An observable sensitive to CP violation must be odd under CP transformations
  - e.g.

$$\beta \equiv \text{sgn}\left((\vec{p}_b - \vec{p}_{\bar{b}}) \cdot (\vec{p}_{\ell^-} \times \vec{p}_{\ell^+})\right)$$

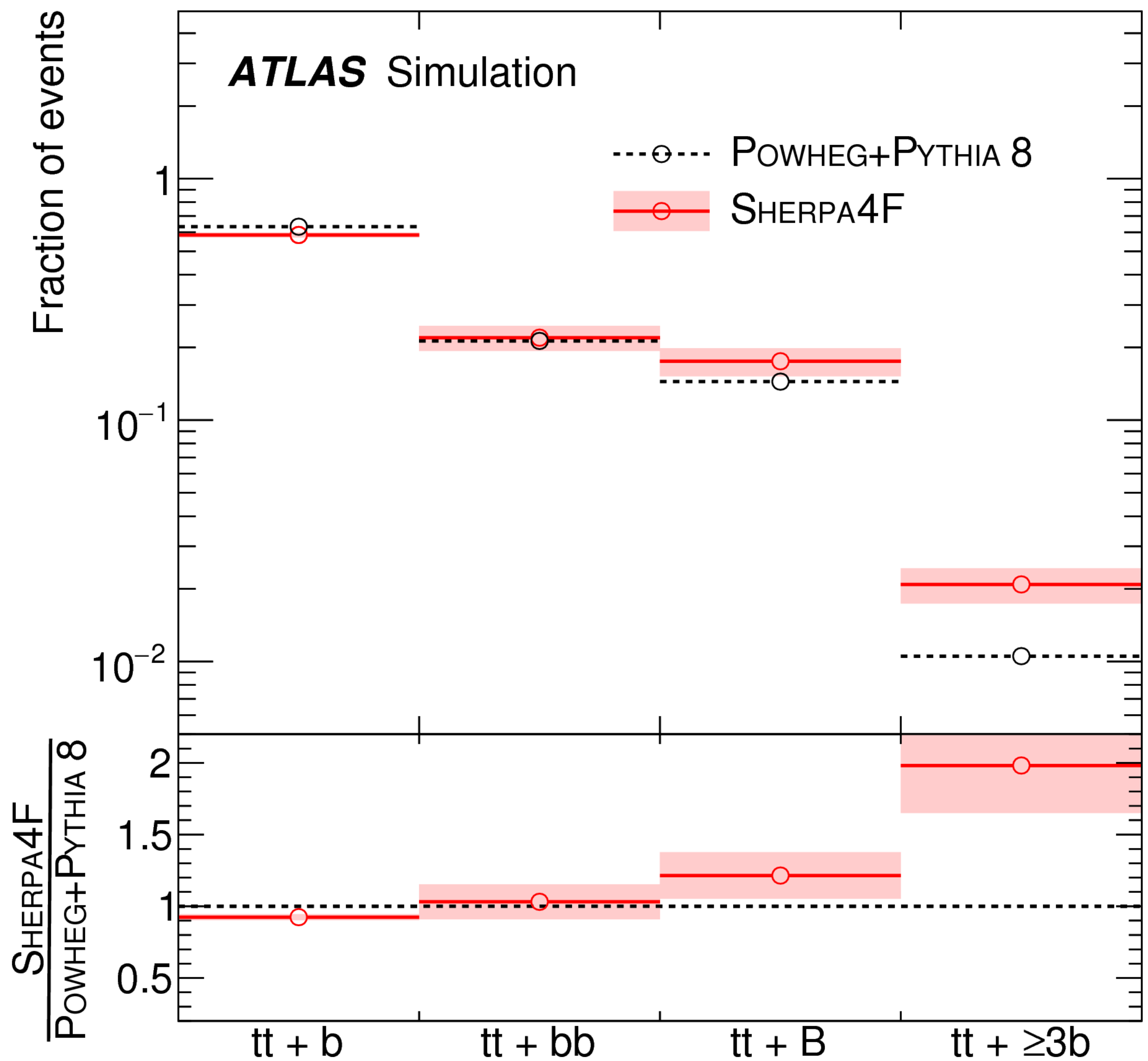
Longer term projects, but we can start thinking about it!

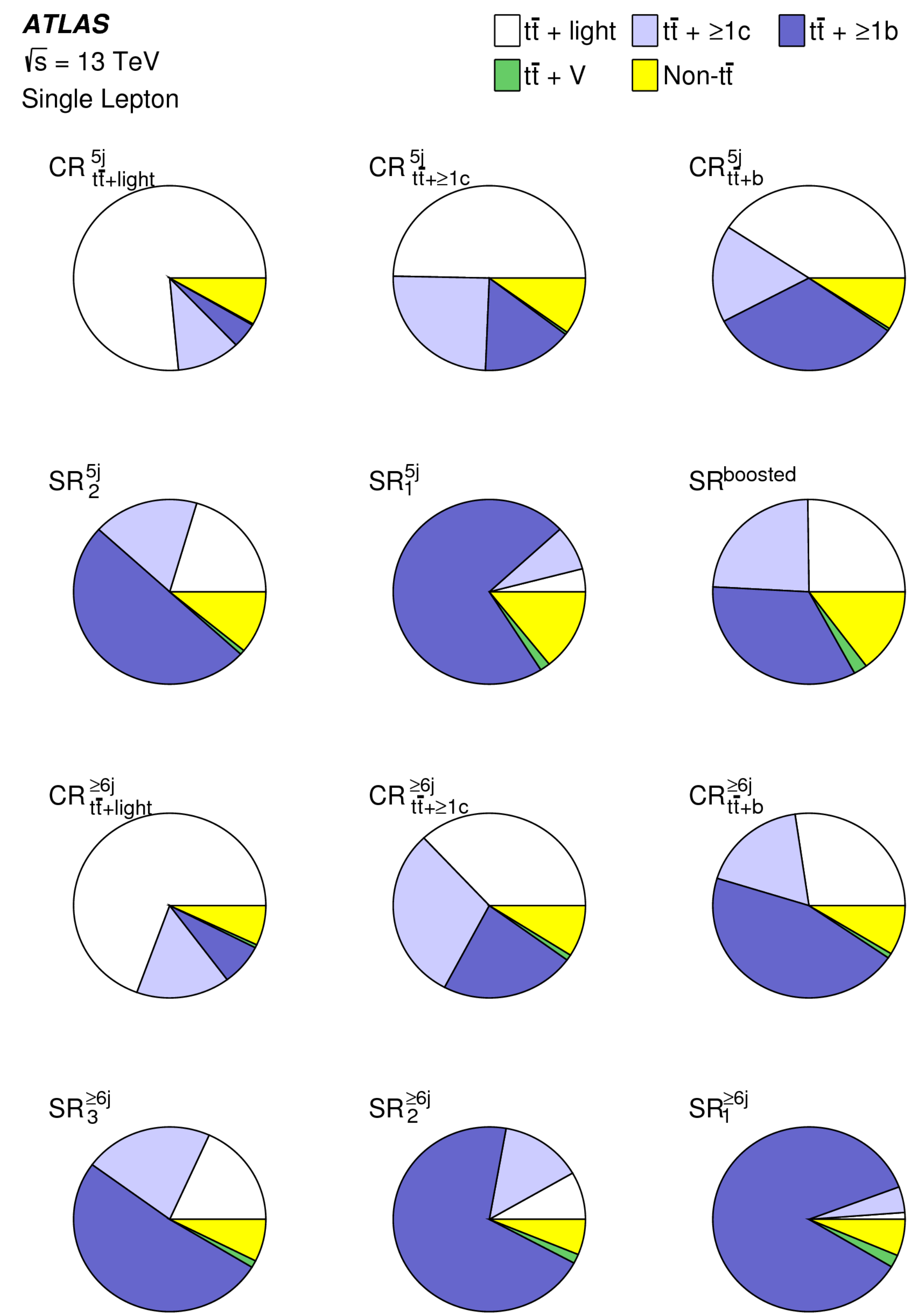




$\Delta\theta^{\ell h}(\ell^+, \ell^-)$ = the angle between the two lepton momenta projected onto the plane perpendicular to the h direction in the lab frame

$t\bar{t}H$ ($H\rightarrow b\bar{b}$): $t\bar{t}$ modelling



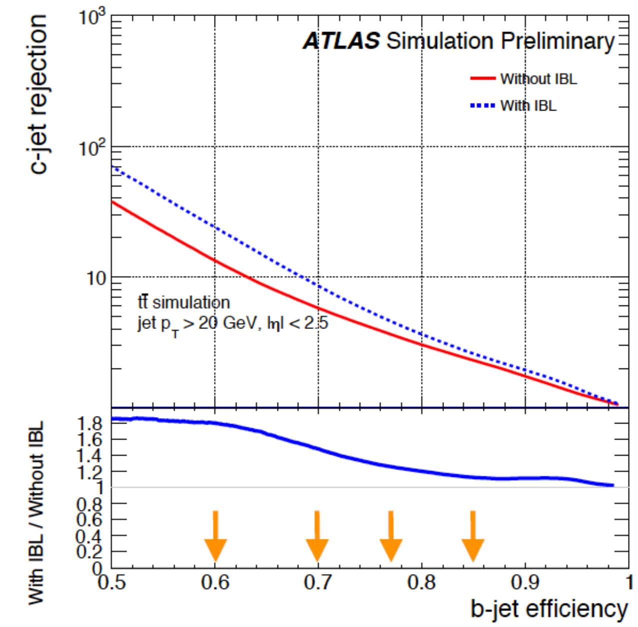


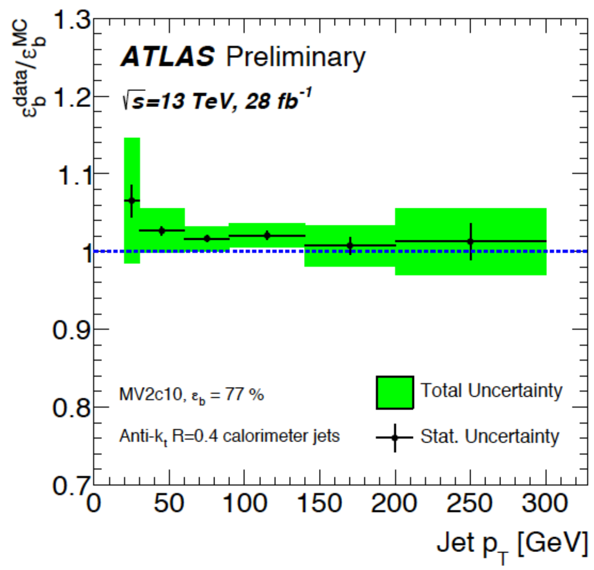
- MC sample split in number of HF jets at particle level
  - $t\bar{t}+\geq 1b$: jets matched to 1(b) or 2(B) b hadrons
  - Extra b-jets from MPI or FSR
  - $t\bar{t}+\geq 1c$: analogous to $t\bar{t}+\geq 1b$
  - $t\bar{t}$+light

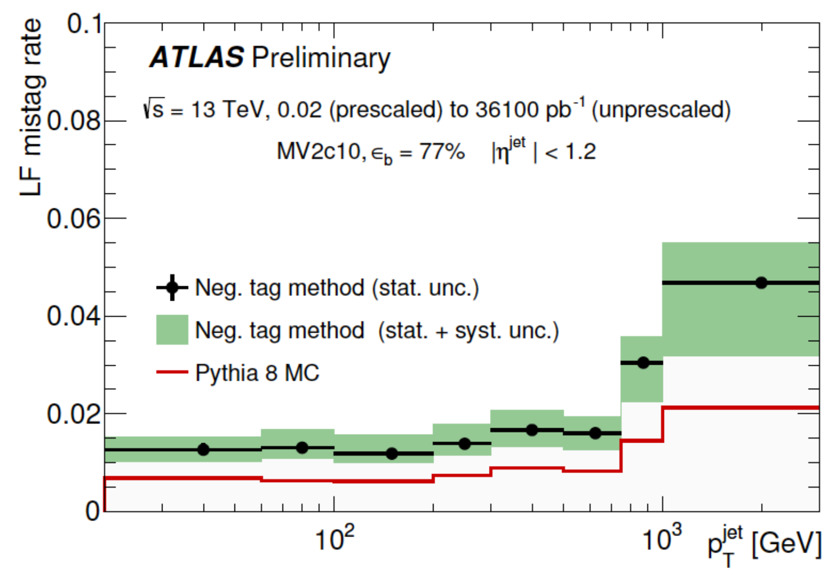
Tamara Vázquez Schröder

b-tagging discriminant

- Large improvement in *b*-tagging performance in Run-2 due to the additional Insertable B-Layer (radius: 3.3 cm)
- Calibration derived from data:
  - *b*-jet efficiency: dileptonic $t\bar{t}$ (2-10% uncert.)
  - *c*-jet mistag: semileptonic $t\bar{t}$ ($W \to cs$), $W+c$ (5-20% uncert.)
  - Light-flavour mistag: dijet events (10-50% uncert.)
- **Four different working points**, also calibration for discrete *b*-tagging discriminant combining all working points used in $t\bar{t}H(b\bar{b})$







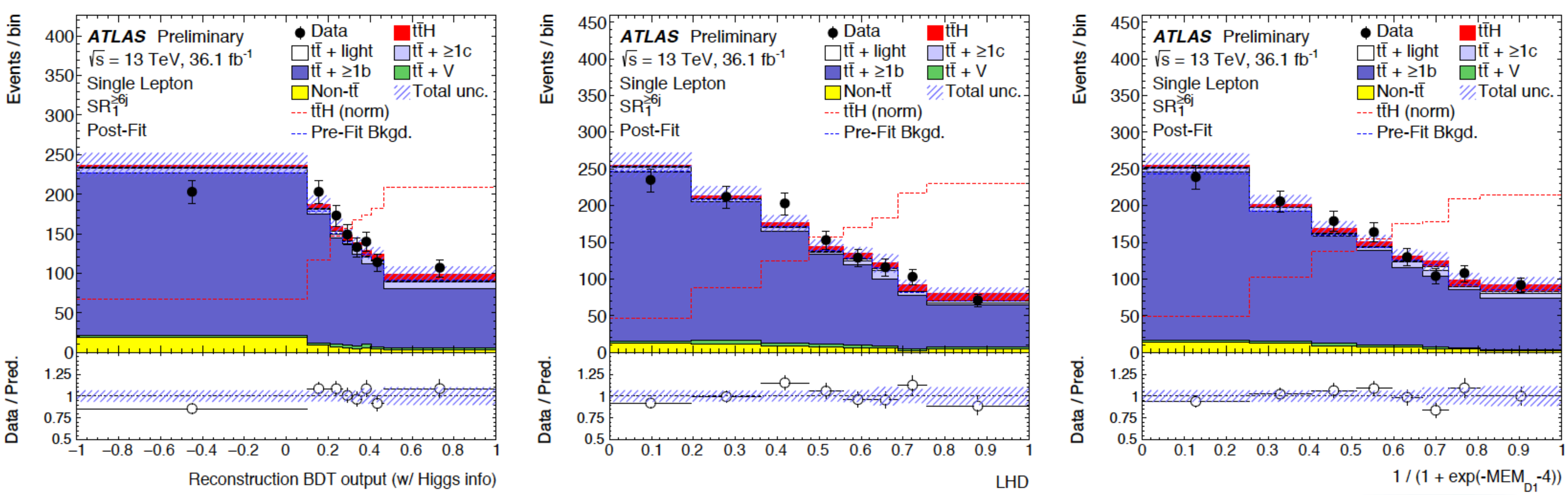
- *b*-tagging:
  - Considering 4 working points: *loose*, *medium*, *tight*, *very-tight*
  - Efficiency for *b*-jets: 85% → 60%
  - Rejection factor for *c*-jets [light jets]: 3→35 [30→1500]
  - *b*-tagging discriminant built as:

	none	*loose*	*medium*	*tight*	*very-tight*
Efficiency	-	85%	77%	70%	60%
Discriminant value	1	2	3	4	5

$t\bar{t}H$ ($H\to b\bar{b}$): MVA analysis

- Sensitivity enhanced using multivariate techniques to discriminate signal from backgrounds
- ‘Reconstruction’ BDT (all resolved SRs):
  - Combination of jets as originating from H/top decays to reconstruct the $t\bar{t}H(b\bar{b})$ system
- Likelihood discriminator (LHD) (1ℓ resolved SRs only):
  - Probability for signal/background ($t\bar{t}+\geq 2b$, $t\bar{t}+1b$) hypotheses using 1D distributions of discriminating variables (invariant mass, angular distributions, etc.)
- Matrix Element Method (MEM) ($\mathrm{SR}_1^{\geq 6j}$ only):
  - Signal/background probability using matrix element calculations at parton level

Reconstuction BDT | Likelihood discriminant | MEM discriminant

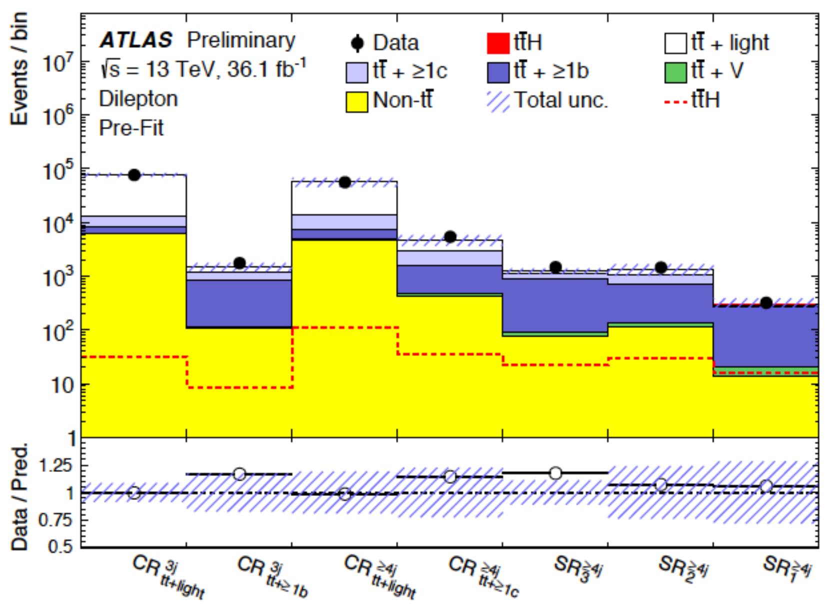


$t\bar{t}H$ ($H\to b\bar{b}$): fit model

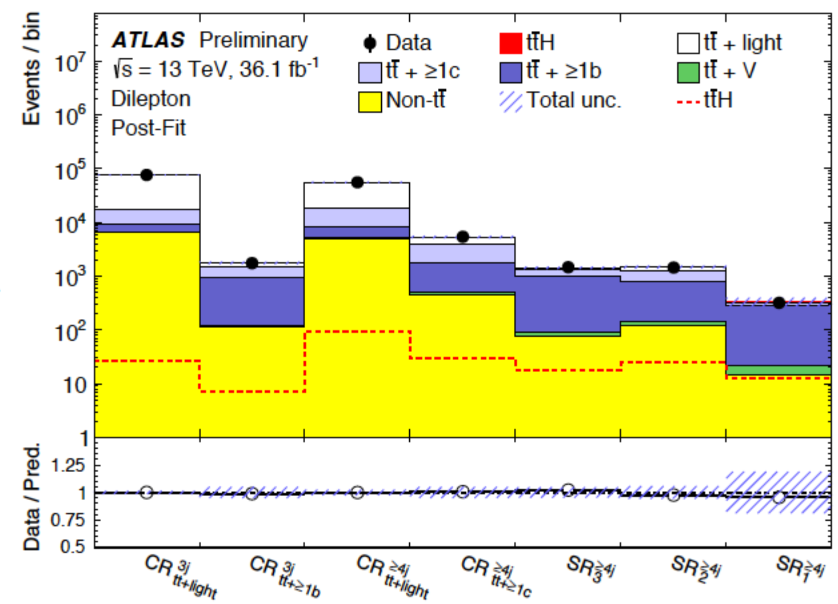
- Simultaneous profile likelihood fit to all SRs and CRs
  - SRs binned in 'classification BDT'
  - CRs: single bin, except $t\bar{t}+\geq 1c$ 1ℓ-CRs (binned in $H_T = \sum_{jet} p_T^{jet}$)

Pre-fit **Post-fit**

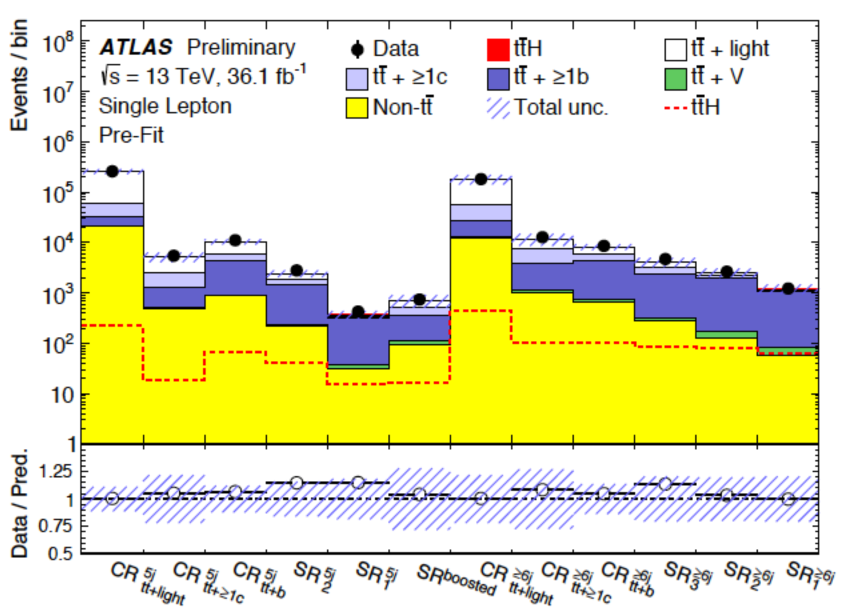
2ℓ



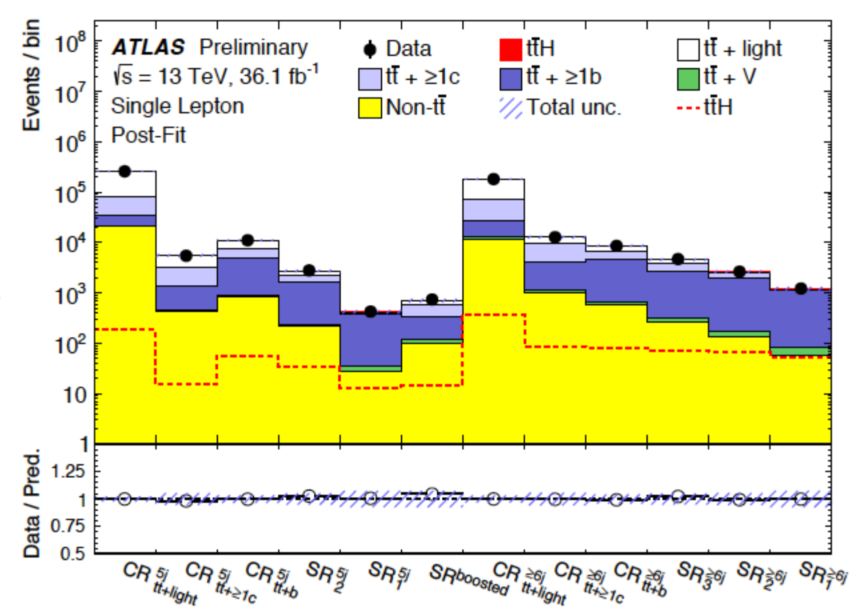
→



1ℓ



→



$t\bar{t}H$ ($H \to b\bar{b}$): tt modelling uncertainties

Systematic source	Description	$t\bar{t}$ categories
$t\bar{t}$ cross-section	Up or down by 6%	All, correlated
$k(t\bar{t}+\geq 1c)$	Free-floating $t\bar{t}+\geq 1c$ normalisation	$t\bar{t}+\geq 1c$
$k(t\bar{t}+\geq 1b)$	Free-floating $t\bar{t}+\geq 1b$ normalisation	$t\bar{t}+\geq 1b$
SHERPA5F vs. nominal	Related to the choice of the NLO generator	All, uncorrelated
PS & hadronisation	POWHEG-BOX+HERWIG 7 vs. POWHEG-BOX+PYTHIA 8	All, uncorrelated
ISR / FSR	Variations of μ_R, μ_F, h_{damp} and A14 Var3c parameters	All, uncorrelated
$t\bar{t}+\geq 1c$ ME vs. inclusive	MG5_aMC@NLO+HERWIG++: ME prediction (3F) vs. incl. (5F)	$t\bar{t}+\geq 1c$
$t\bar{t}+\geq 1b$ SHERPA4F vs. nominal	Comparison of $t\bar{t}+b\bar{b}$ NLO (4F) vs. POWHEG-BOX+PYTHIA 8 (5F)	$t\bar{t}+\geq 1b$
$t\bar{t}+\geq 1b$ renorm. scale	Up or down by a factor of two	$t\bar{t}+\geq 1b$
$t\bar{t}+\geq 1b$ resumm. scale	Vary μ_Q from $H_T/2$ to μ_{CMMPS}	$t\bar{t}+\geq 1b$
$t\bar{t}+\geq 1b$ global scales	Set μ_Q, μ_R, and μ_F to μ_{CMMPS}	$t\bar{t}+\geq 1b$
$t\bar{t}+\geq 1b$ shower recoil scheme	Alternative model scheme	$t\bar{t}+\geq 1b$
$t\bar{t}+\geq 1b$ PDF (MSTW)	MSTW vs. CT10	$t\bar{t}+\geq 1b$
$t\bar{t}+\geq 1b$ PDF (NNPDF)	NNPDF vs. CT10	$t\bar{t}+\geq 1b$
$t\bar{t}+\geq 1b$ MPI	Up or down by 50%	$t\bar{t}+\geq 1b$
$t\bar{t}+\geq 3b$ normalisation	Up or down by 50%	$t\bar{t}+\geq 1b$

- Many sources of modelling uncertainty considered:
  - Generator: Powheg+Pythia8 vs. Sherpa (5F)
  - Parton shower: Powheg+Pythia8 vs. Powheg+Herwig7
  - 5F vs. 4F in Sherpa+OpenLoops
  - Scale variations in Sherpa+OpenLoops
- All $t\bar{t}$+jets modelling uncertainties uncorrelated between $t\bar{t}+\geq 1b/\geq 1c/$light
- Scale variation uncertainties correlated across each $t\bar{t}+\geq 1b$ sub-component