Study of the associated production of the Higgs boson with a top quark pair in a boosted regime in the ATLAS experiment at LHC

Silvia Biondi
from the PhD thesis work

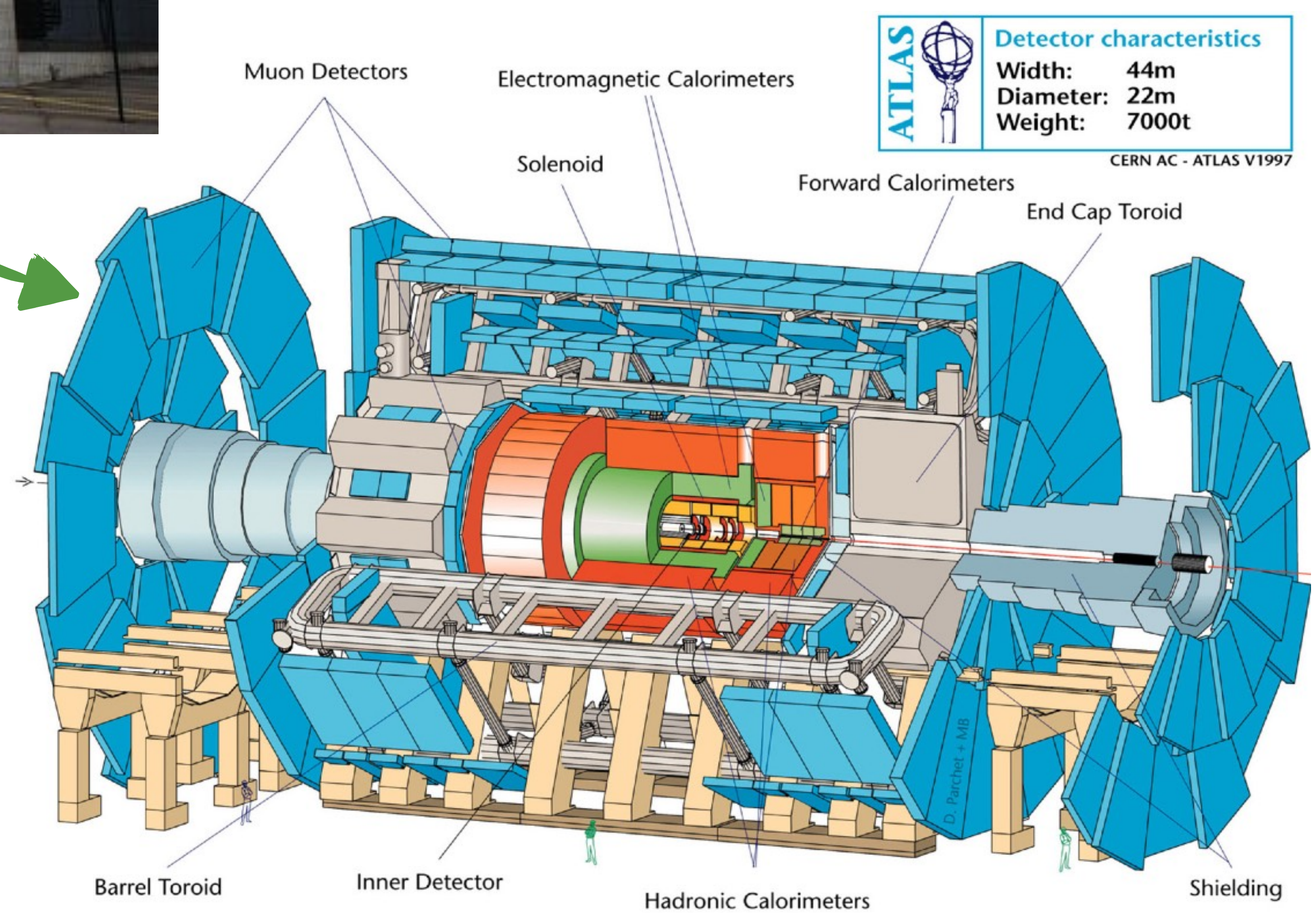
Aperitivo scientifico
Università di Bologna & INFN - 15.06.18

The Large Hadron Collider

- The **biggest particle accelerator in the world**:
  - **~100 m** under the ground of Geneva;
  - **27 km** ring of superconducting magnets;
  - in Run-2 (2015-2018), pp collisions at a center-of-mass energy of **13 TeV**;
  - 4 big experiments:
    - ALICE, **ATLAS**, CMS, LHCb;
  - other **smaller experiments**;
  - **active since 2009**.

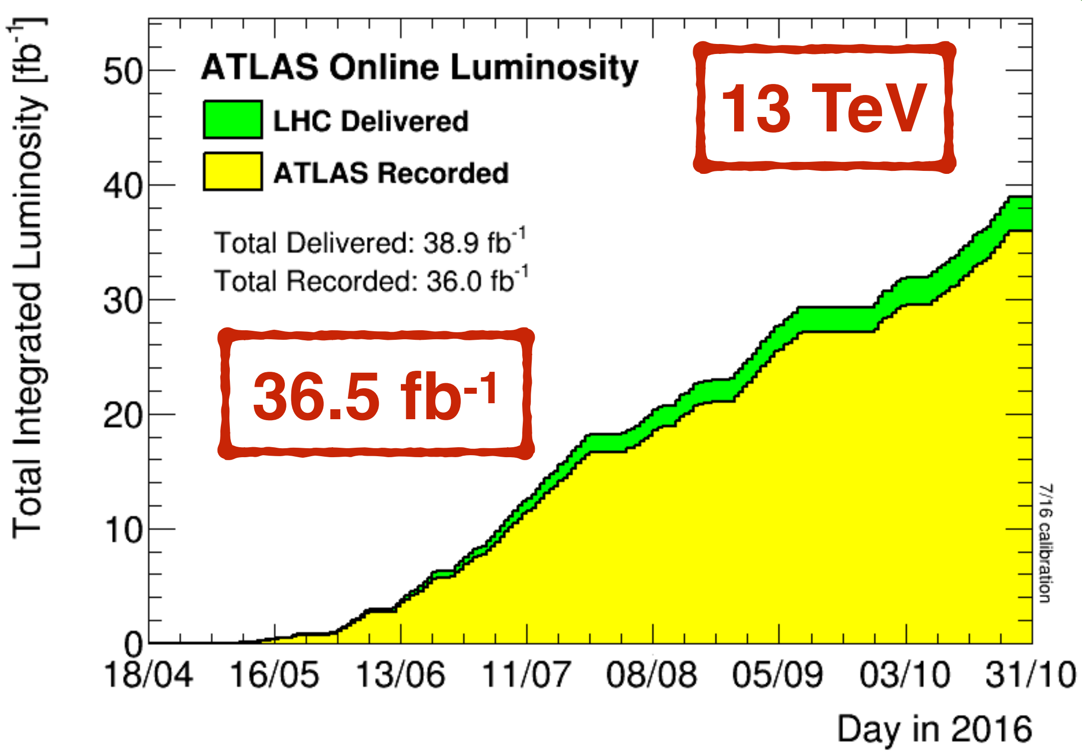
The ATLAS Experiment

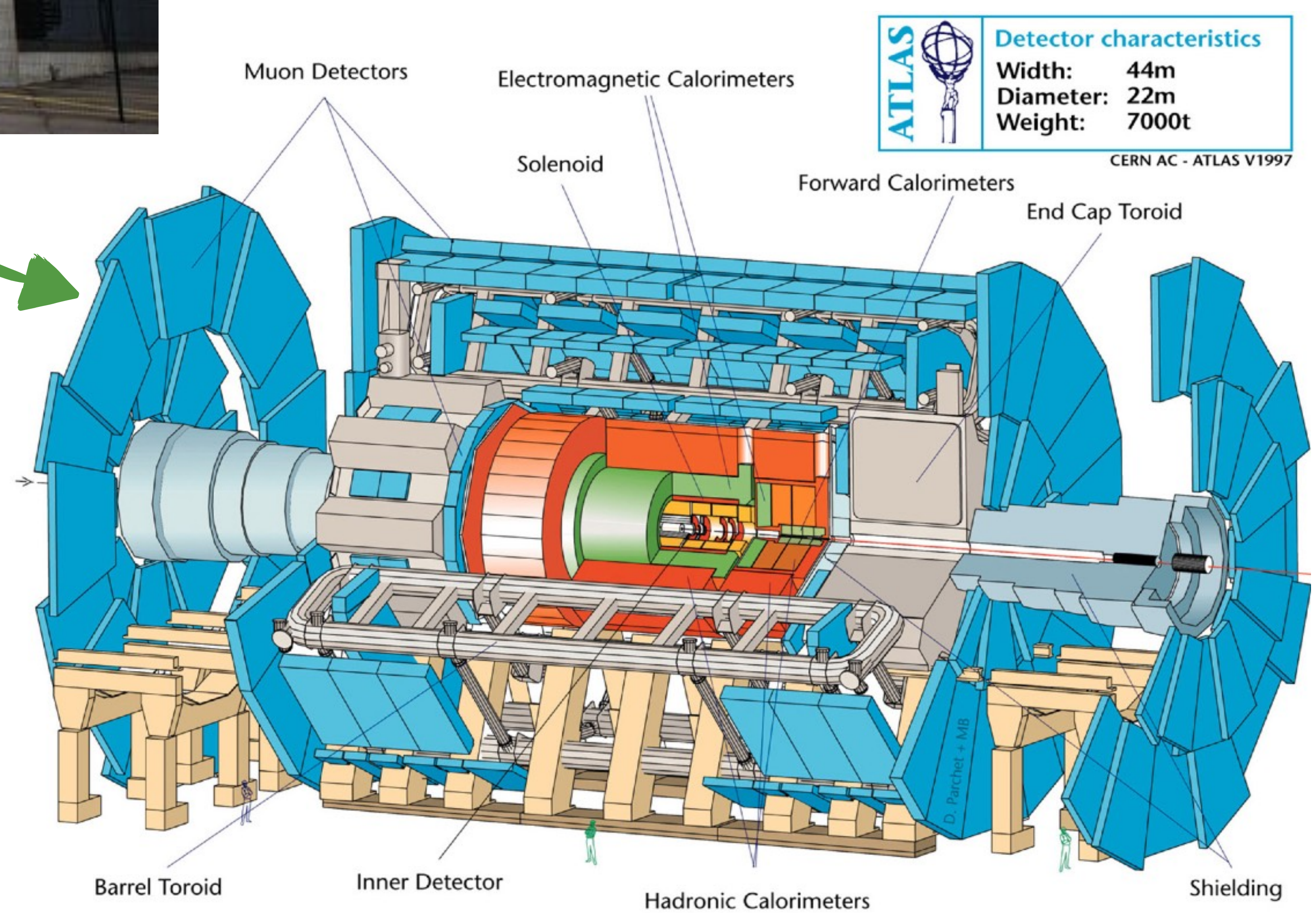
- **Multipurpose experiment** to discover signatures of new physics and to perform precise measurements of Standard Model.
- **Different technologies depending on the kind of physics to be detected**.



The ATLAS Experiment

- **Multipurpose experiment** to discover signatures of new physics and to perform precise measurements of Standard Model.
- **Different technologies depending on the kind of physics to be detected**.





Top Quark and Higgs Boson

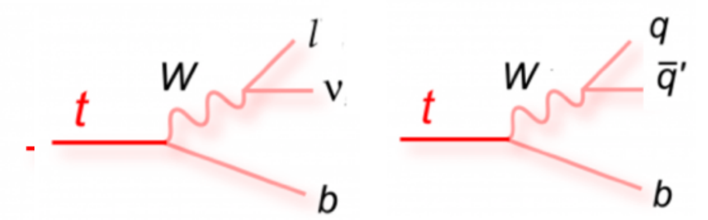
- The last quark discovered, only in **1995**.
- It is the **most massive** fundamental particle known:

reference

m_t = 173.34 ± 0.27 (stat) ± 0.71 (syst) GeV

- High m_t implies a **large Yukawa coupling with the Higgs boson** (~1), wrt other couplings ($<10^{-2}$).
- Due to its short lifetime, the top quark **decays before hadronizing** (detected as a **"jet"**: a cone of particles that goes through the detector).
- Unique opportunity to study properties of a bare quark.

Top decays ~100% of times in **W boson and b quark**



decay time ~ 10^{-24} s
hadronization ~ 10^{-23} s

Top Quark and Higgs Boson

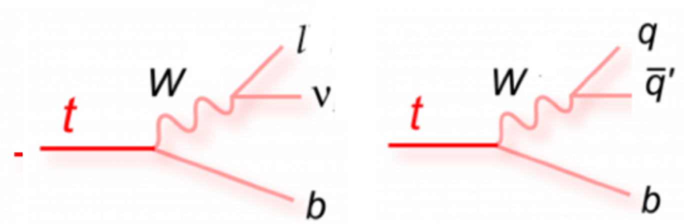
- The last quark discovered, only in **1995**.
- It is the **most massive** fundamental particle known:

reference

m_t = 173.34 ± 0.27 (stat) ± 0.71 (syst) GeV

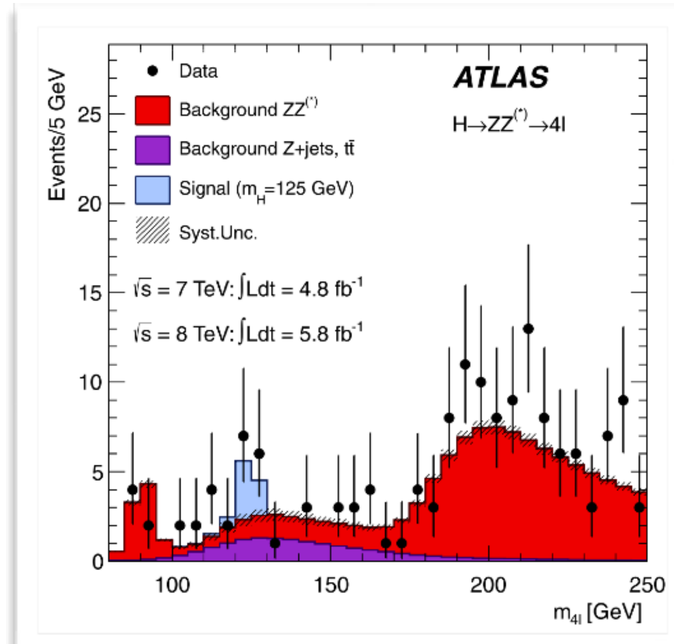
- High m_t implies a **large Yukawa coupling with the Higgs boson** (~1), wrt other couplings ($<10^{-2}$).
- Due to its short lifetime, the top quark **decays before hadronizing** (detected as a **"jet"**: a cone of particles that goes through the detector).
- Unique opportunity to study properties of a bare quark.

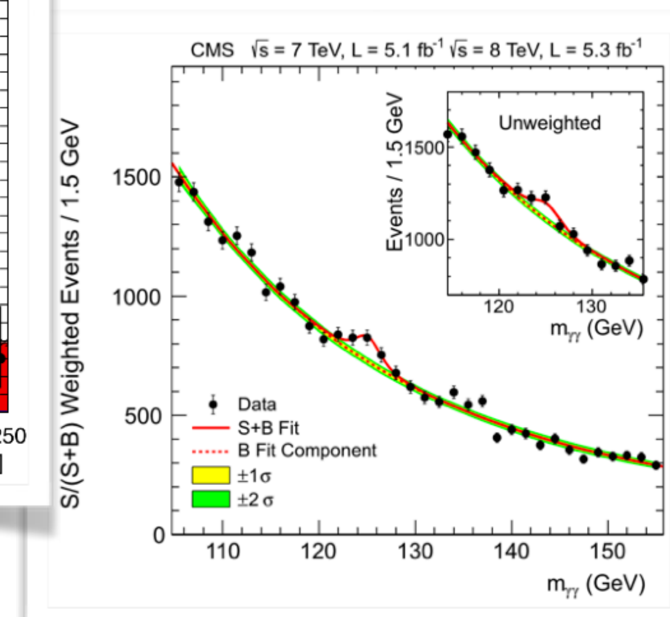
Top decays ~100% of times in **W boson and b quark**



decay time ~ 10^{-24} s
hadronization ~ 10^{-23} s

- Speculated in **1964** by **Higgs**, **Englert** and **Brout**; discovered in 2012 at CERN; **Nobel in 2013**.
- ATLAS and CMS collaborations observed **a neutral scalar particle of mass ~125 GeV**: Higgs boson.





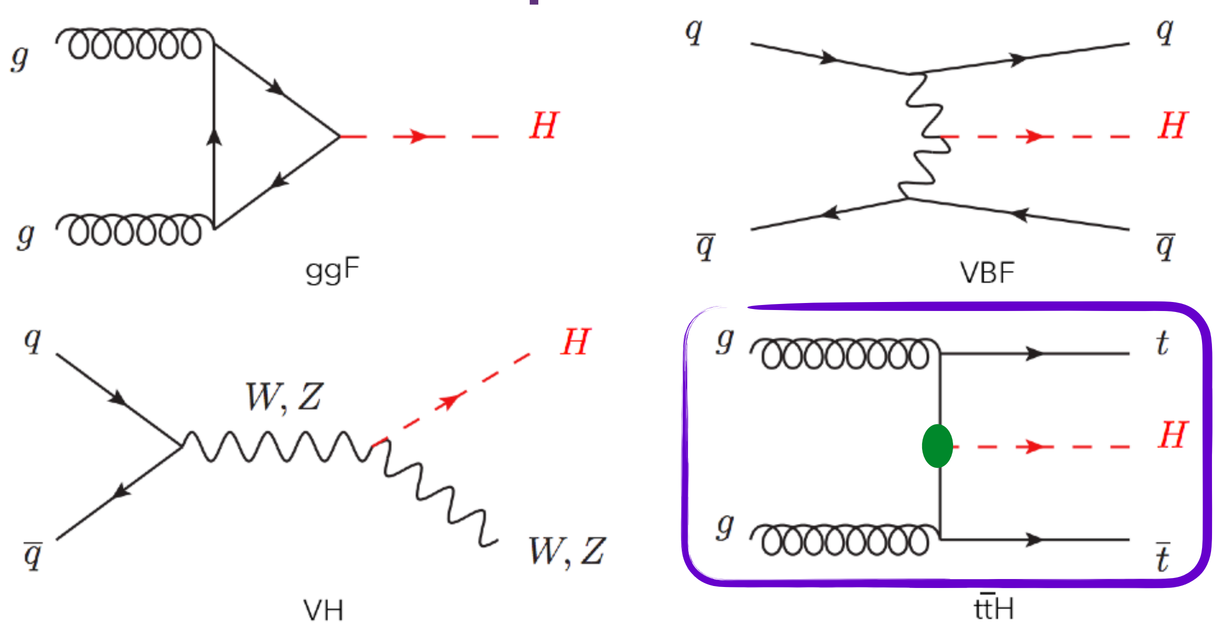
- Still **ongoing studies about its properties** (mass, spin, etc.).

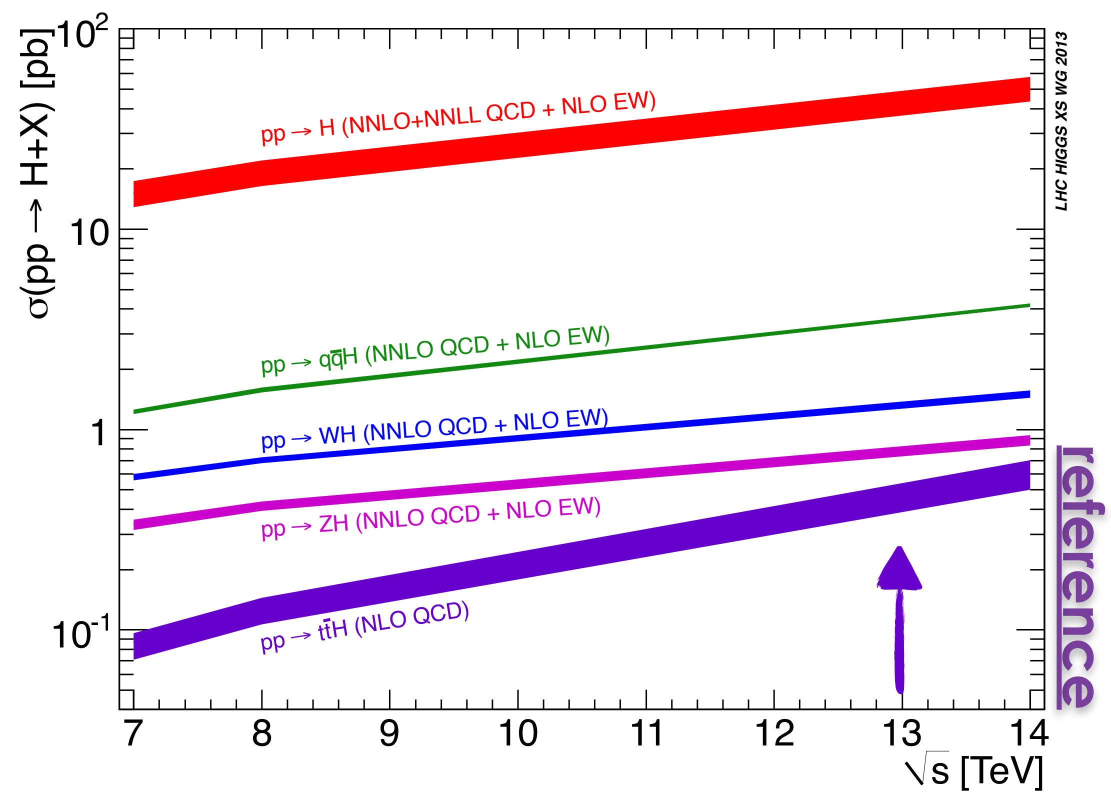
reference

m_H = 125.09 ± 0.21 (stat) ± 0.11 (syst) GeV

The associated Higgs production

The production

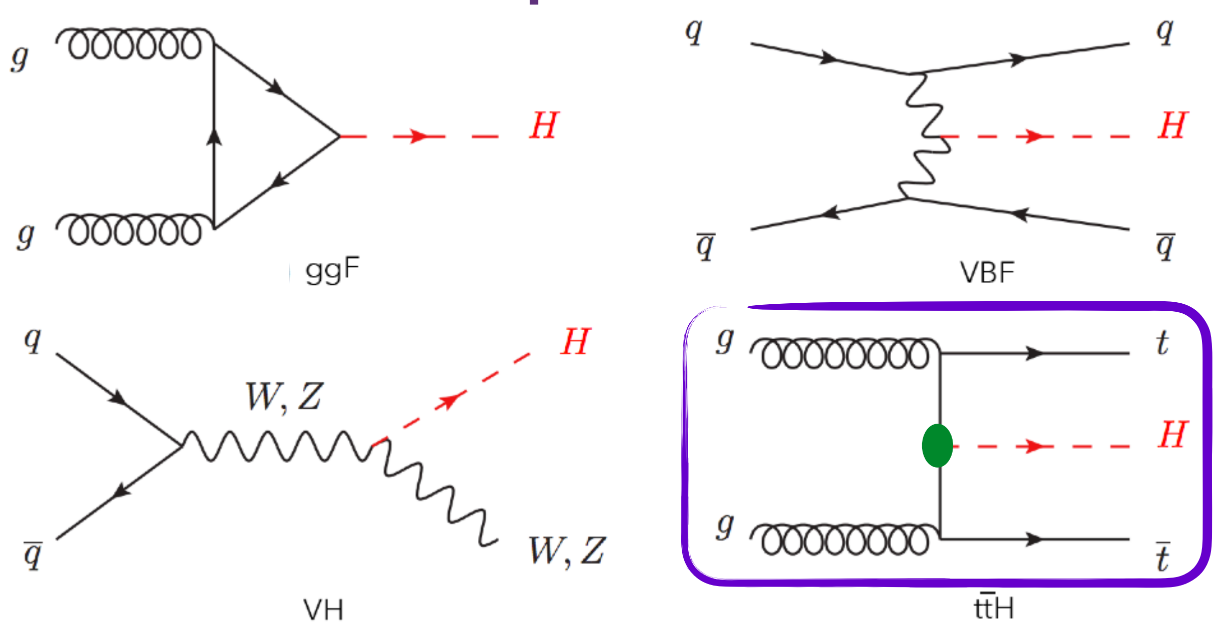


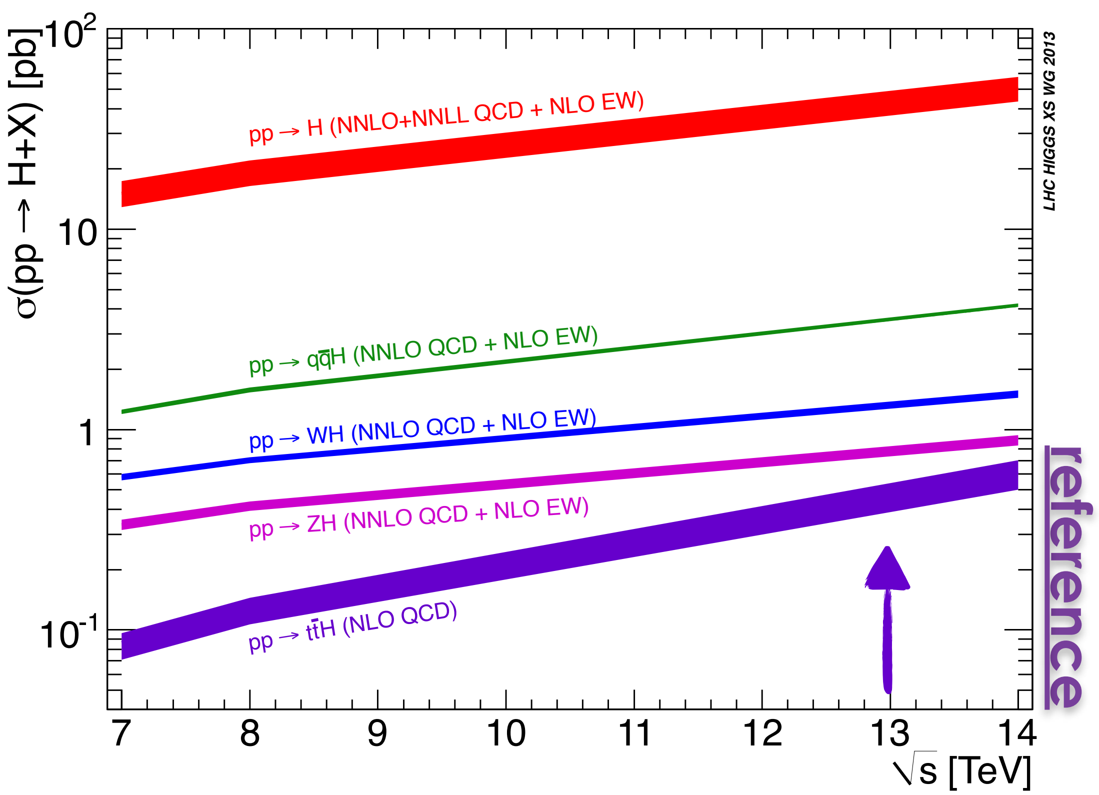


The **highest cross section increase as a function of energy** wrt other production modes

The associated Higgs production

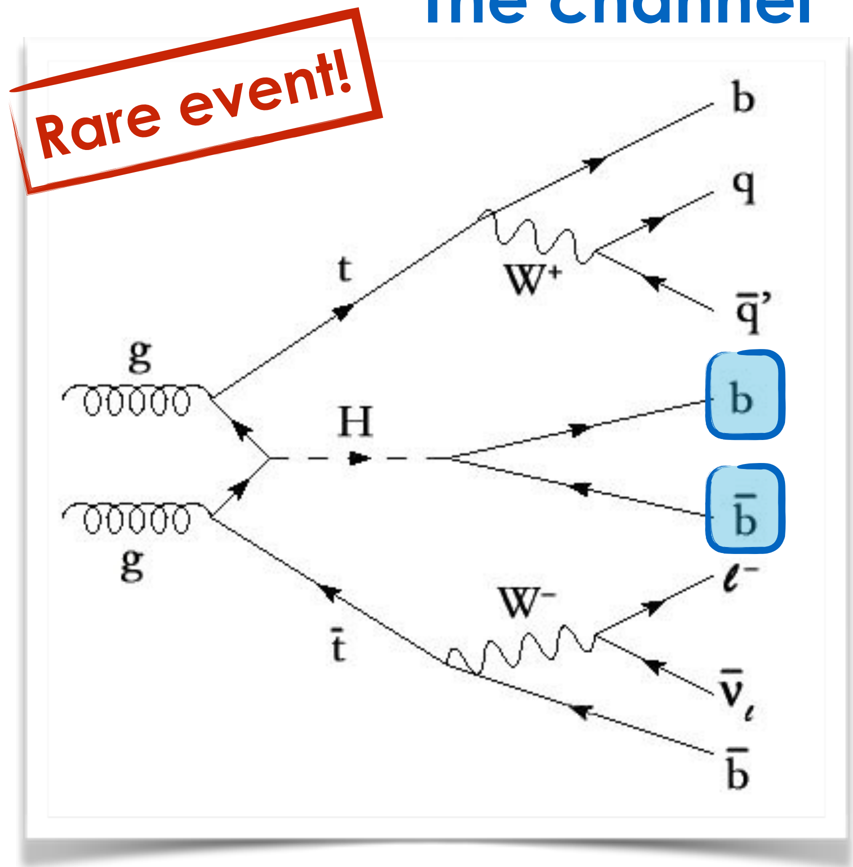
The production





The **highest cross section increase as a function of energy** wrt other production modes

The channel



semi-leptonic decay of tt system: BR (tt->lνqq) ~35%

hadronic decay of Higgs boson: BR (H->bb) ~60%

σ(**pp**) ~ 10^{11} pb (100 mb)

σ(pp->**H+X**) ~ 45 pb

σ(pp->**ttH**) ~ 0.5 pb

@ 13 TeV

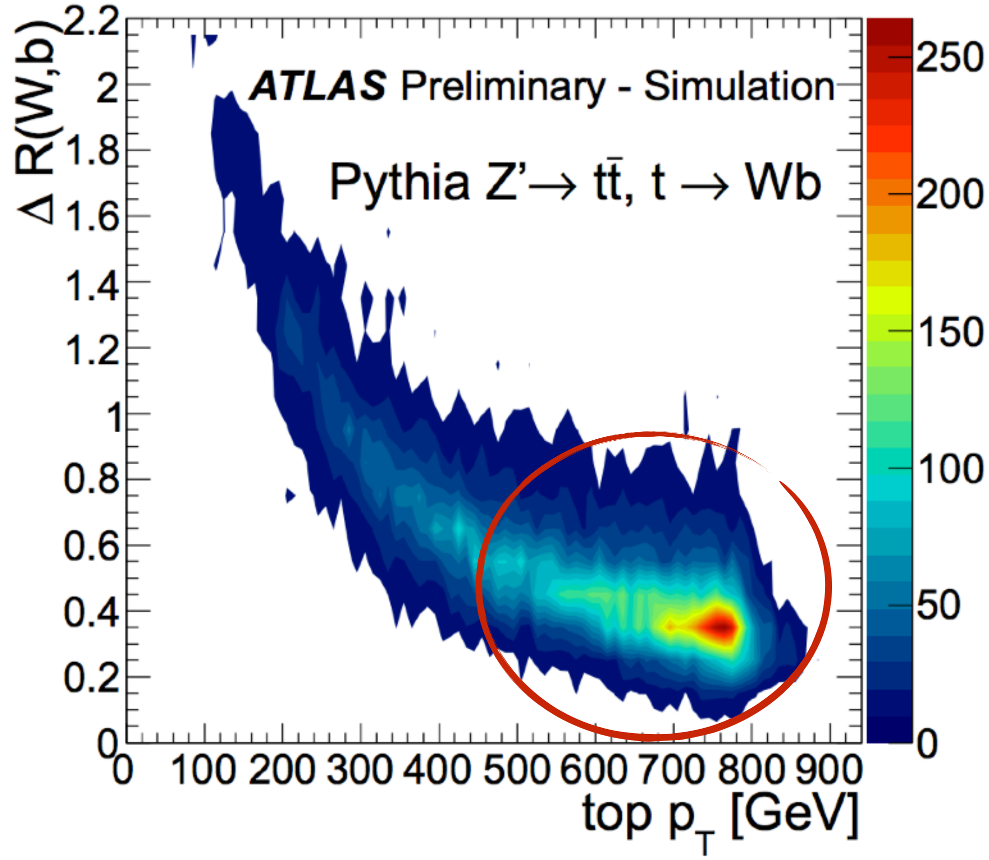
1 ttH each 10^{12} evts

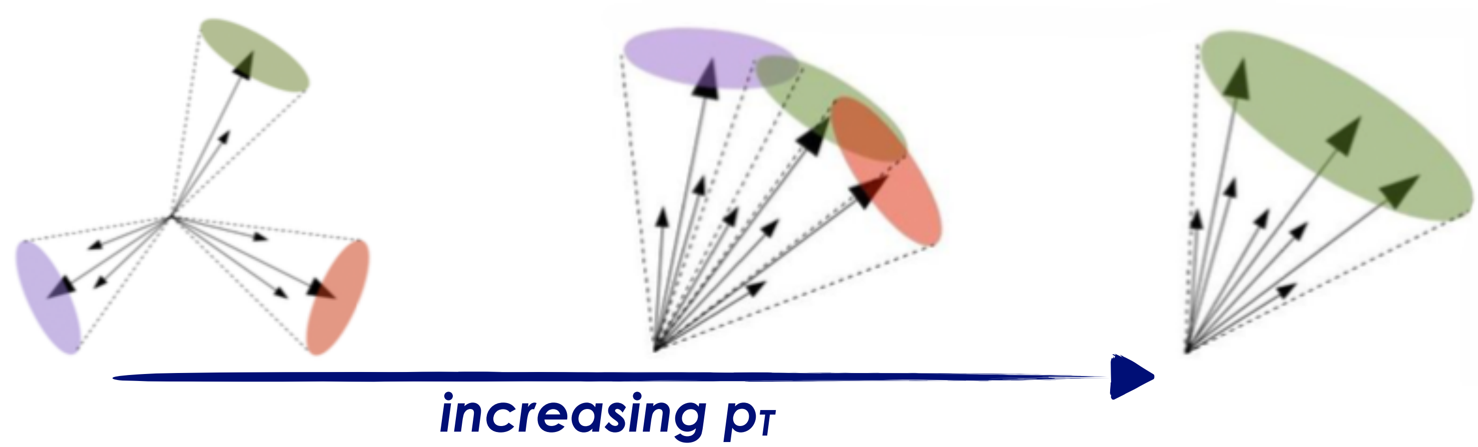
NB!

Aim: measure the signal strength μ = σ/σ$_{SM}$

The importance of the topology

- Increasing more and more the energy at LHC, a completely **new physics regime can be explored**:
  - heavy particles are often produced with large transverse momentum (**boosted particles**);
  - their decay products are **collimated to the decaying particle direction** in the detector rest frame;
  - the ability to resolve the individual hadronic decay products using **standard narrow-cone jet algorithms begins to degrade**.



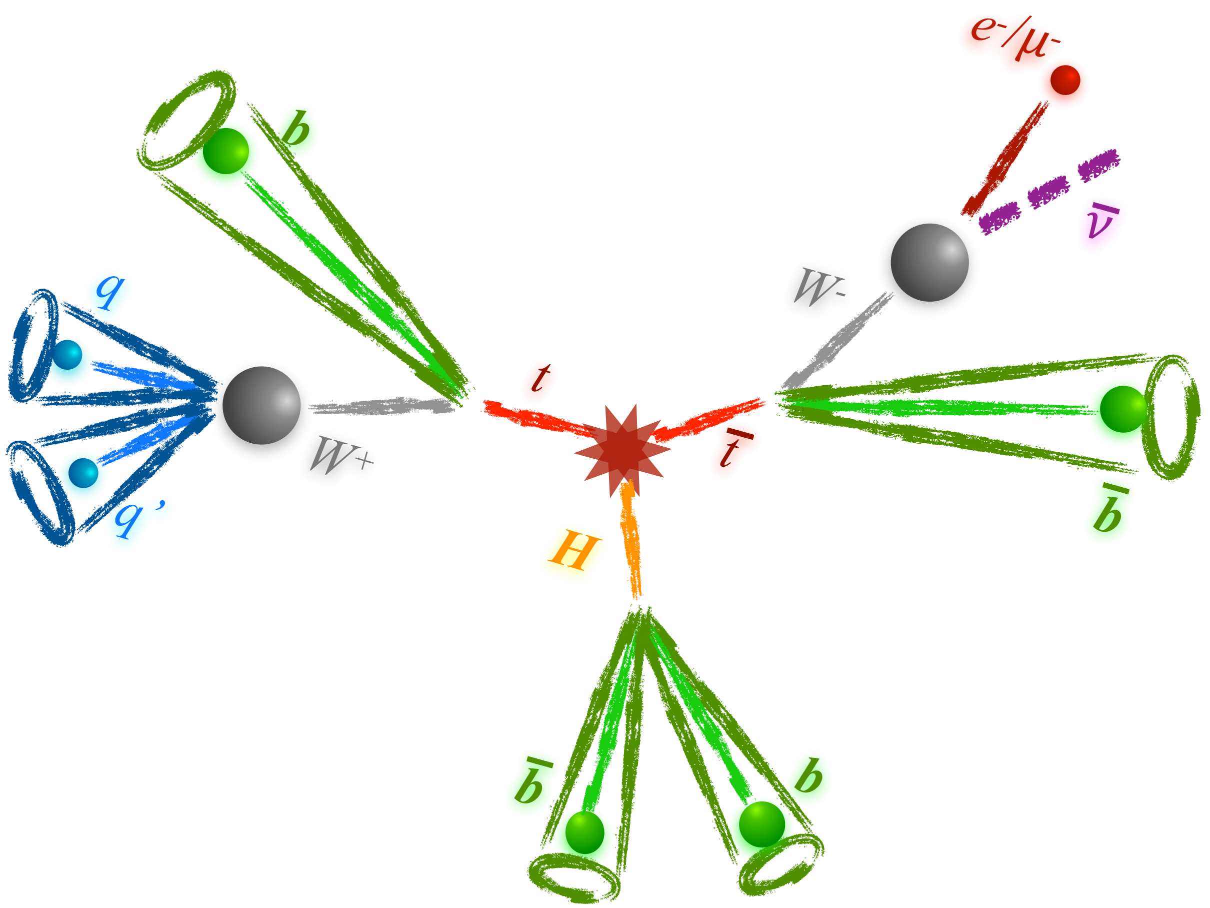


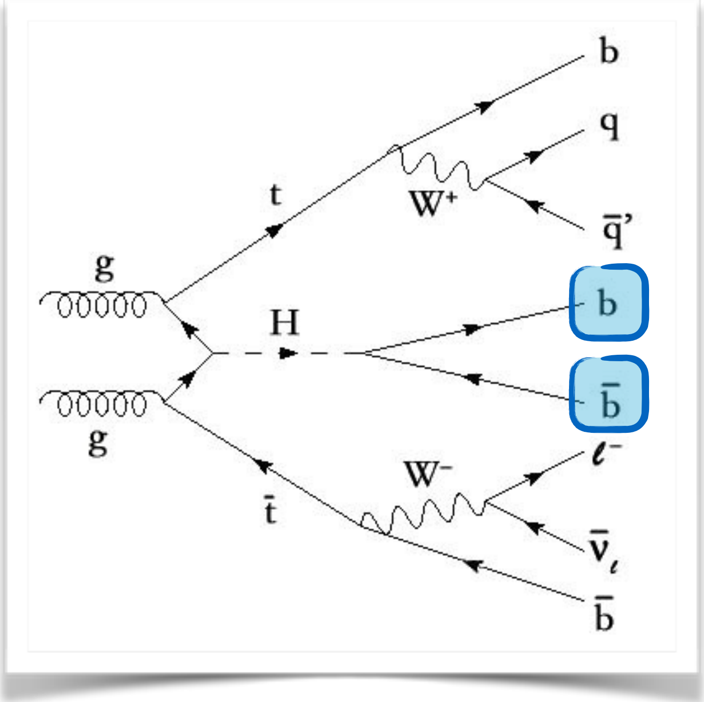
$$\Delta R \equiv \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2} \simeq 2m/p_T$$

Solution

The **decay products of a hadronically decaying object merge into a single, energetic and large radius jet** (***large-R jet***) with a characteristic substructure different from those initiated by a single parton.

Resolved analysis

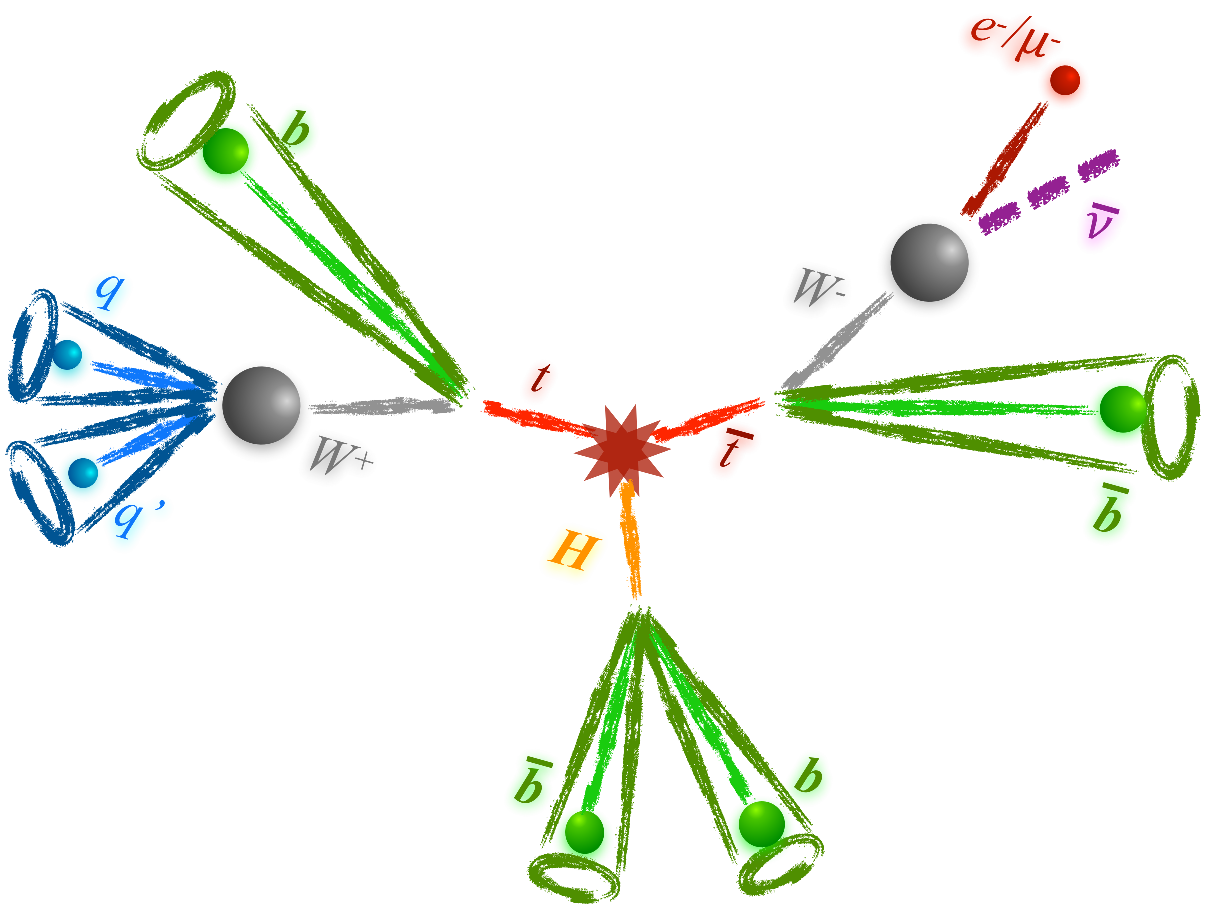




- **6** well separated small-R jets ("**resolved**"):
  - standard jet reconstruction algorithms (anti-k_t **ΔR < 0.4**);
  - significant combinatorial background.

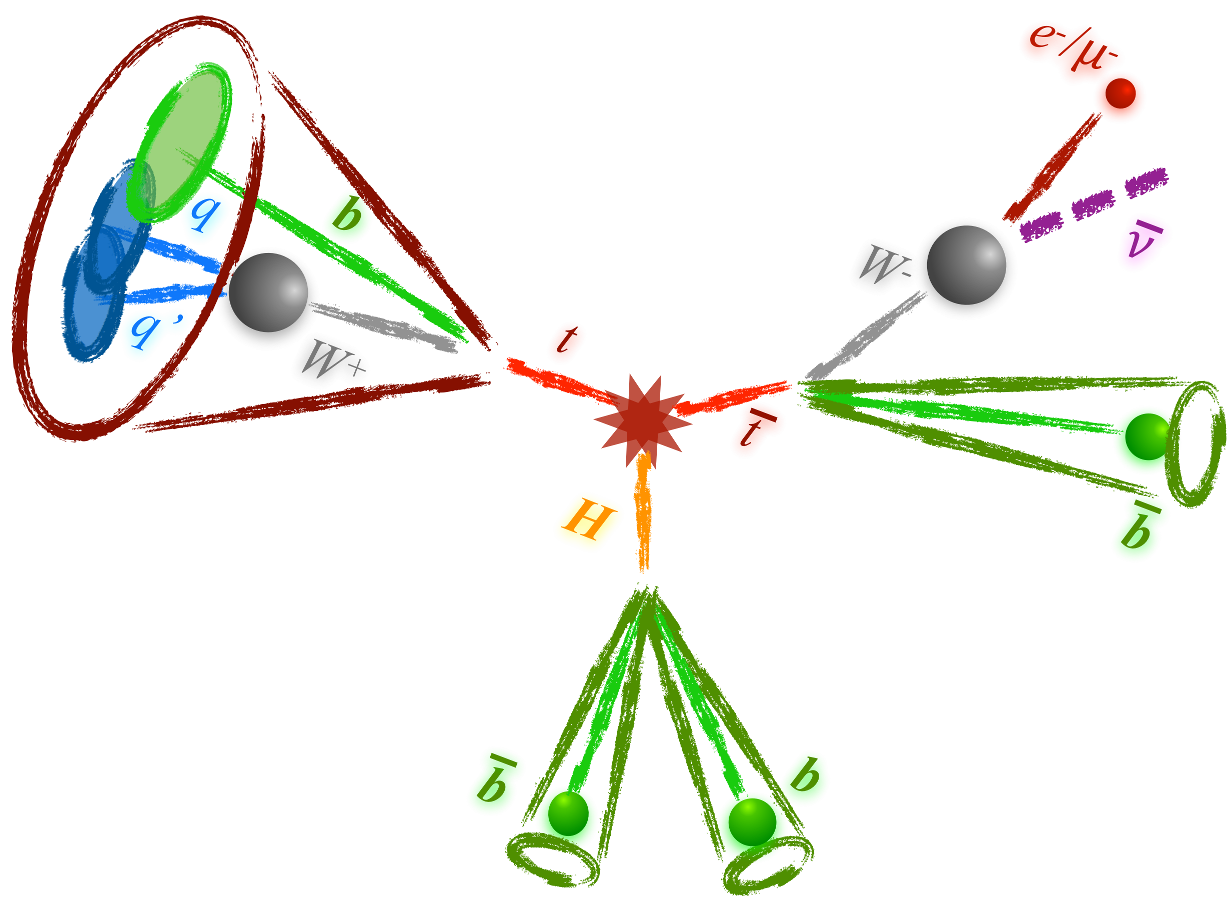
6j,≥4b

Combined analysis



- **6** well separated small-R jets ("**resolved**"):
  - standard jet reconstruction algorithms (anti-k_t **ΔR < 0.4**);
  - significant combinatorial background.

6j,≥4b



- At least 1 large-R jet ("**boosted**"):
  - standard jet reconstruction algorithms (anti-k_t **ΔR < 1.0**);
  - high background contamination into large-R jet;
  - additional algorithms for jet grooming and identification.

Large-R jets: reconstruction and grooming

both small-R jets and large-R jets

Reconstruction algorithms

Jet are reconstructed with an iterative algorithm which combines calo deposits inside a given radius R = 1.0

Large-R jets: reconstruction and grooming

both small-R jets and large-R jets

only large-R jets

Reconstruction algorithms

Jet are reconstructed with an iterative algorithm which combines calo deposits inside a given radius R = 1.0

Jets are then cleaned, with "**grooming**" algorithms, from contamination due to the high particles concentration

Large-R jets: reconstruction and grooming

both small-R jets and large-R jets

Reconstruction algorithms

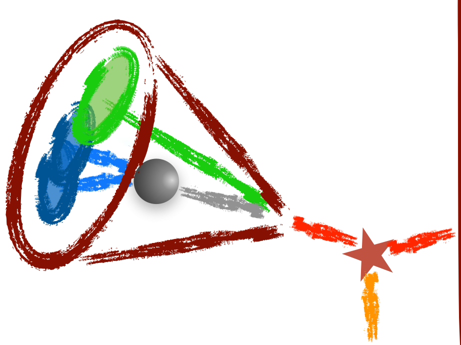
Jet are reconstructed with an iterative algorithm which combines calo deposits inside a given radius R = 1.0

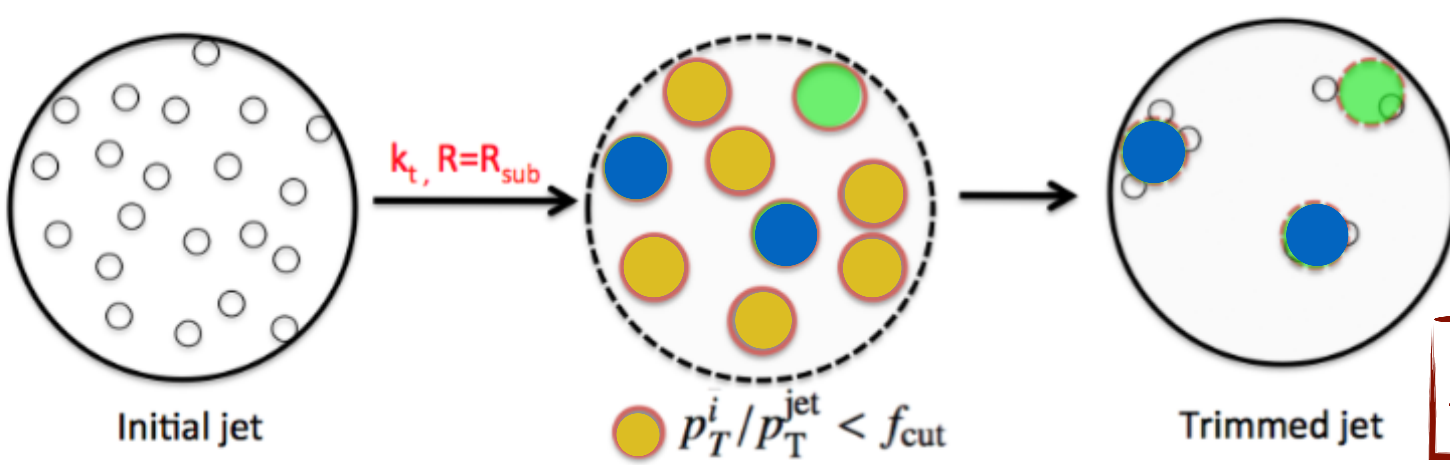
Jets are then cleaned, with "**grooming**" algorithms, from contamination due to the high particles concentration

only large-R jets

Trimming algorithm

Jet constituents are reconstructed again into jets with smaller radius $\mathbf{R_{sub}}$ (subjet). Subjets with lower p_T than a fraction $\mathbf{f_{cut}}$ of initial jet p_T are dropped off.
The final jet is reconstructed using only the remaining subjets.

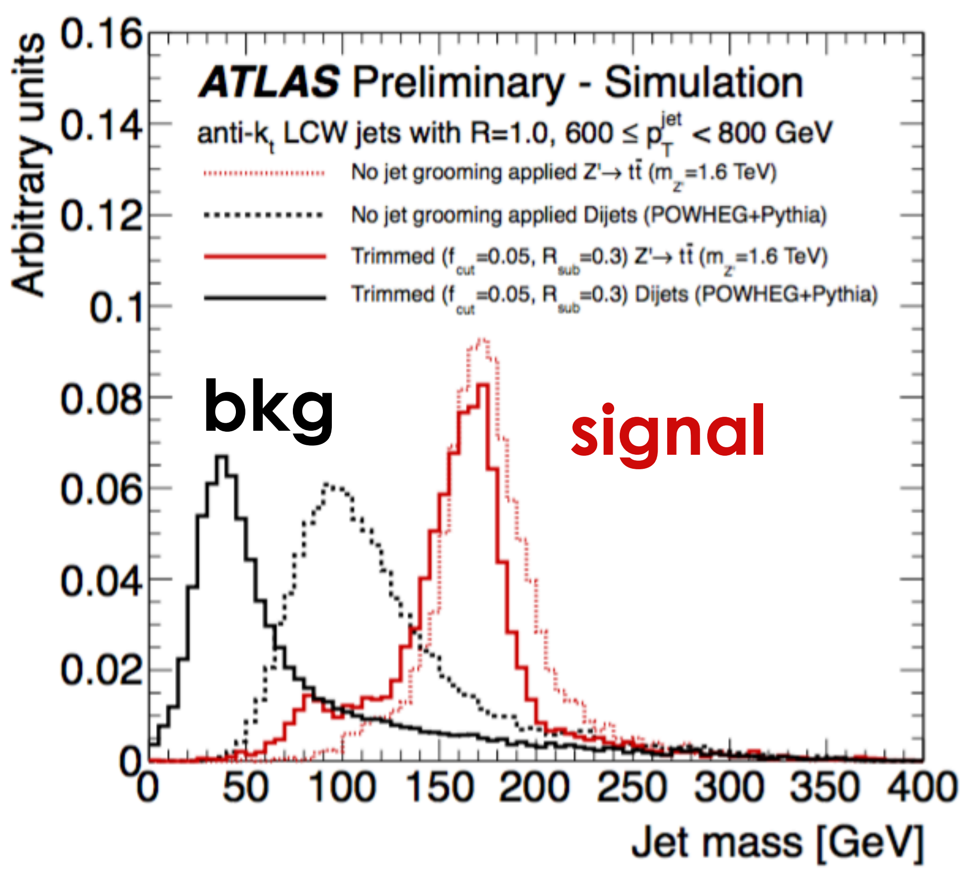


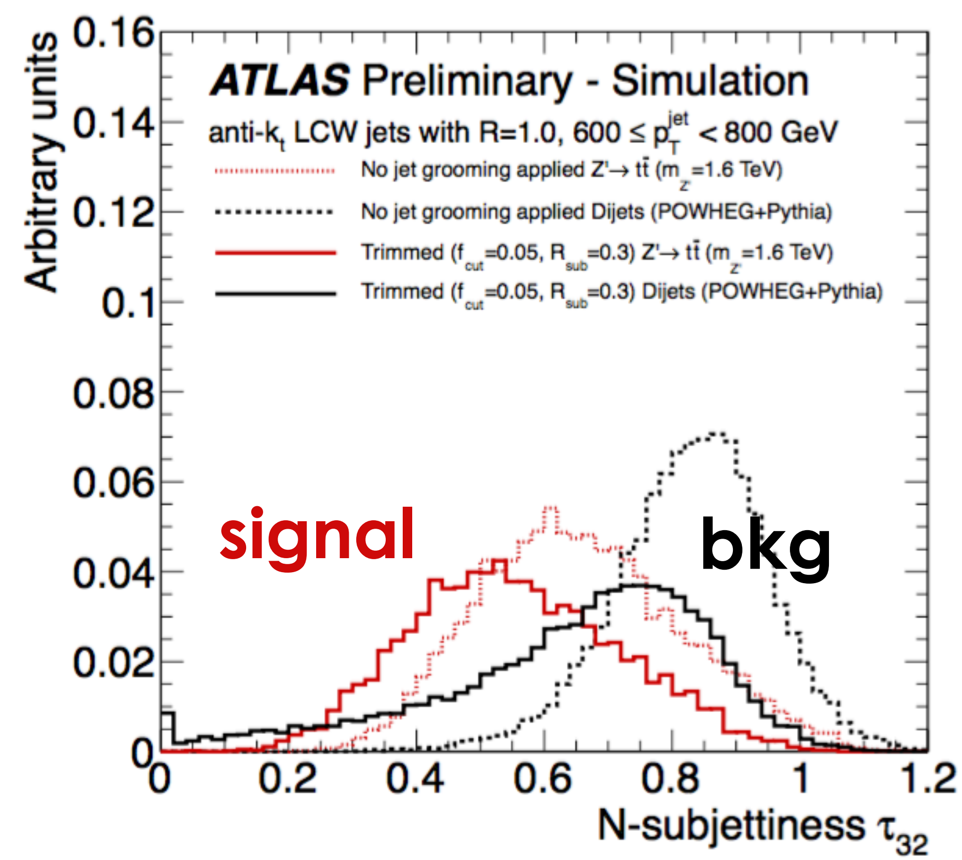


R_{sub} = 0.2
f_{cut} = 0.05

Top Tagging technique

- Used to **exploit all the substructure characteristics of the large-R jets** in boosted regimes.
- Substructure variables in the algorithm: **large-R jet mass** and **τ_{32} ratio**.





Performances

- Comparison between **efficiencies in tt (main background) and ttH (signal)**.
- Looking at:

Signal efficiency

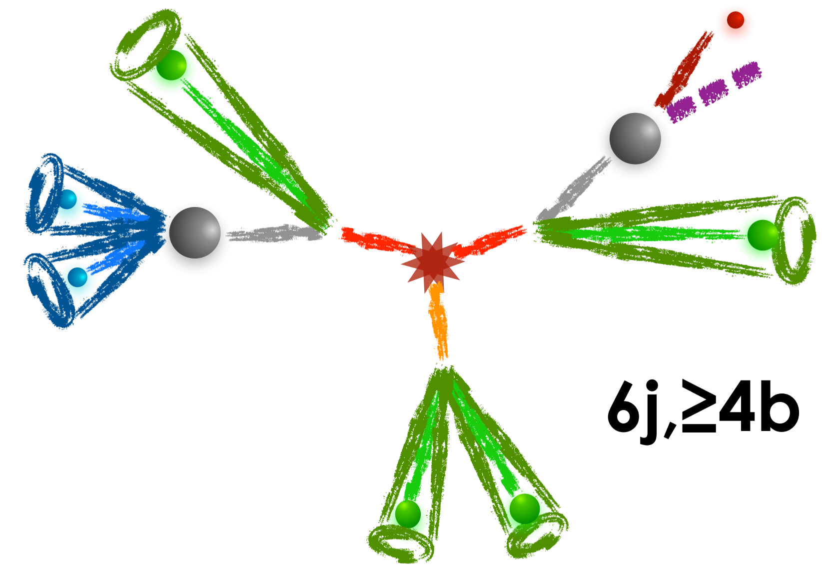
$$\epsilon = \left(\frac{N_{tagged}}{N_{total}}\right)_{signal}$$

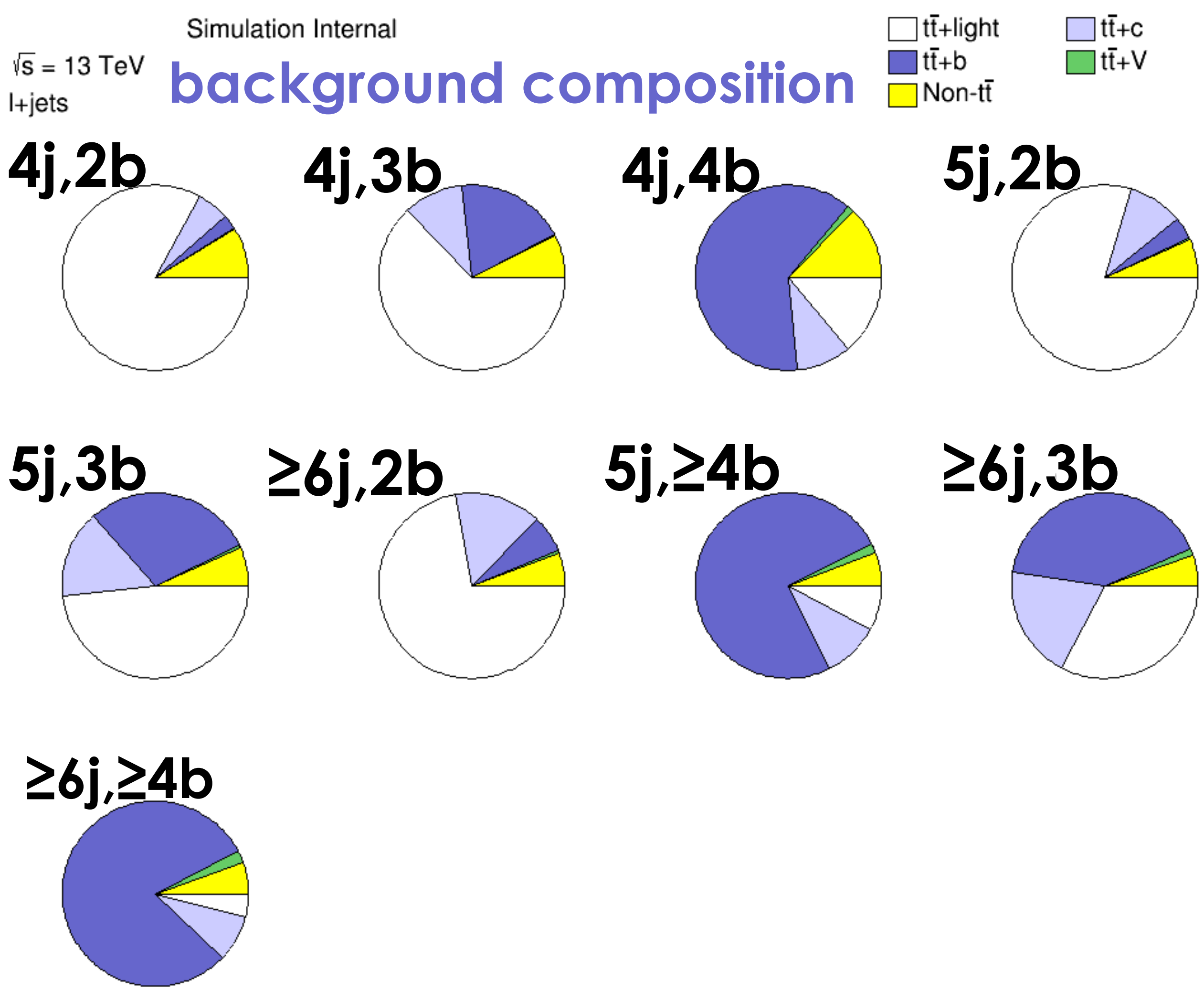
Background rejection

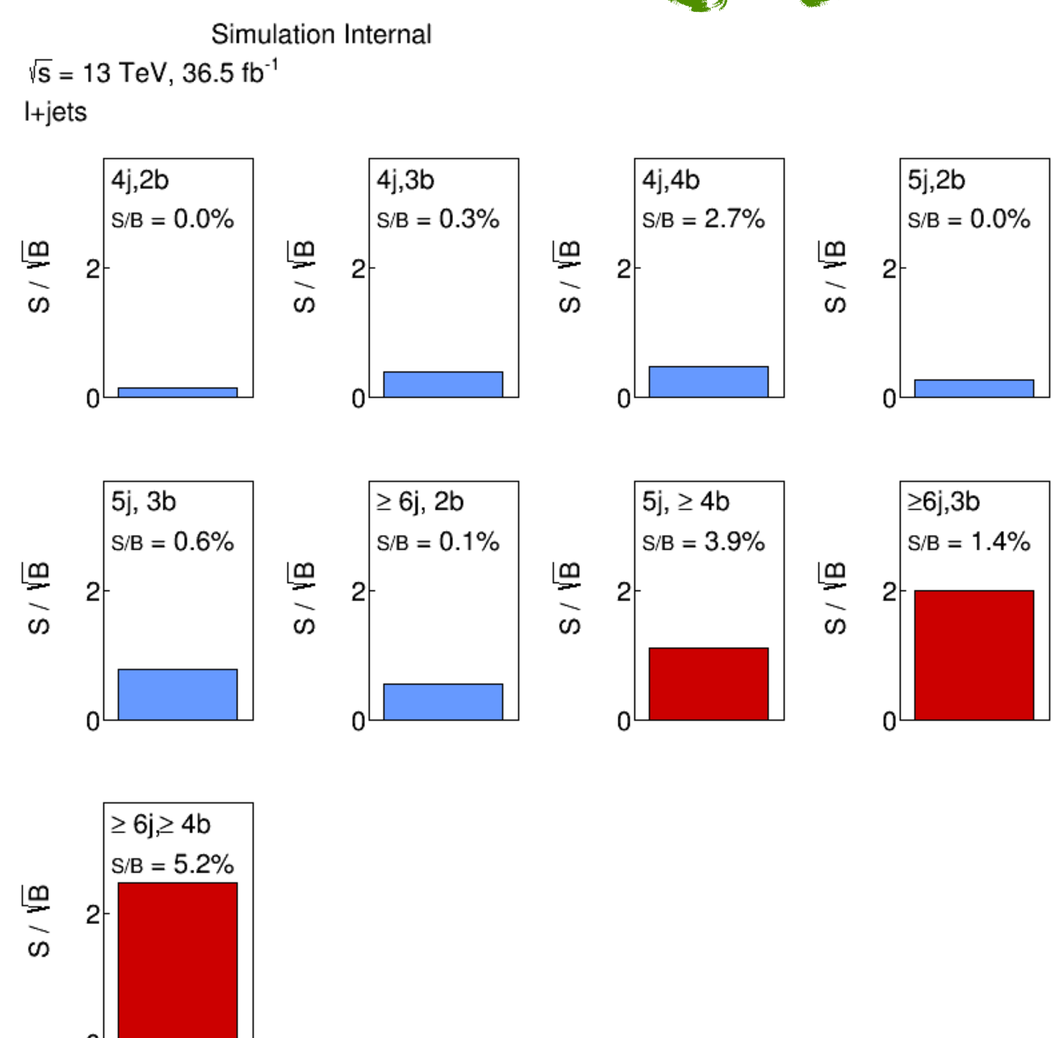
$$r = \left(\frac{N_{total}}{N_{tagged}}\right)_{background}$$

Control and signal regions

- **Common pre-selection**: exactly 1 lepton, ≥ 3 jets, ≥ 2 b-jets.
- To optimize the analysis sensitivity, **division of the phase space into several regions** (defined by the jet and b-jet multiplicity):
  - where **ttH is enhanced** wrt to background (high S/B and S/√B);
  - **dominated by background**, allowing a tighter constraint of backgrounds and systematic uncertainties in the analysis final step (fit).





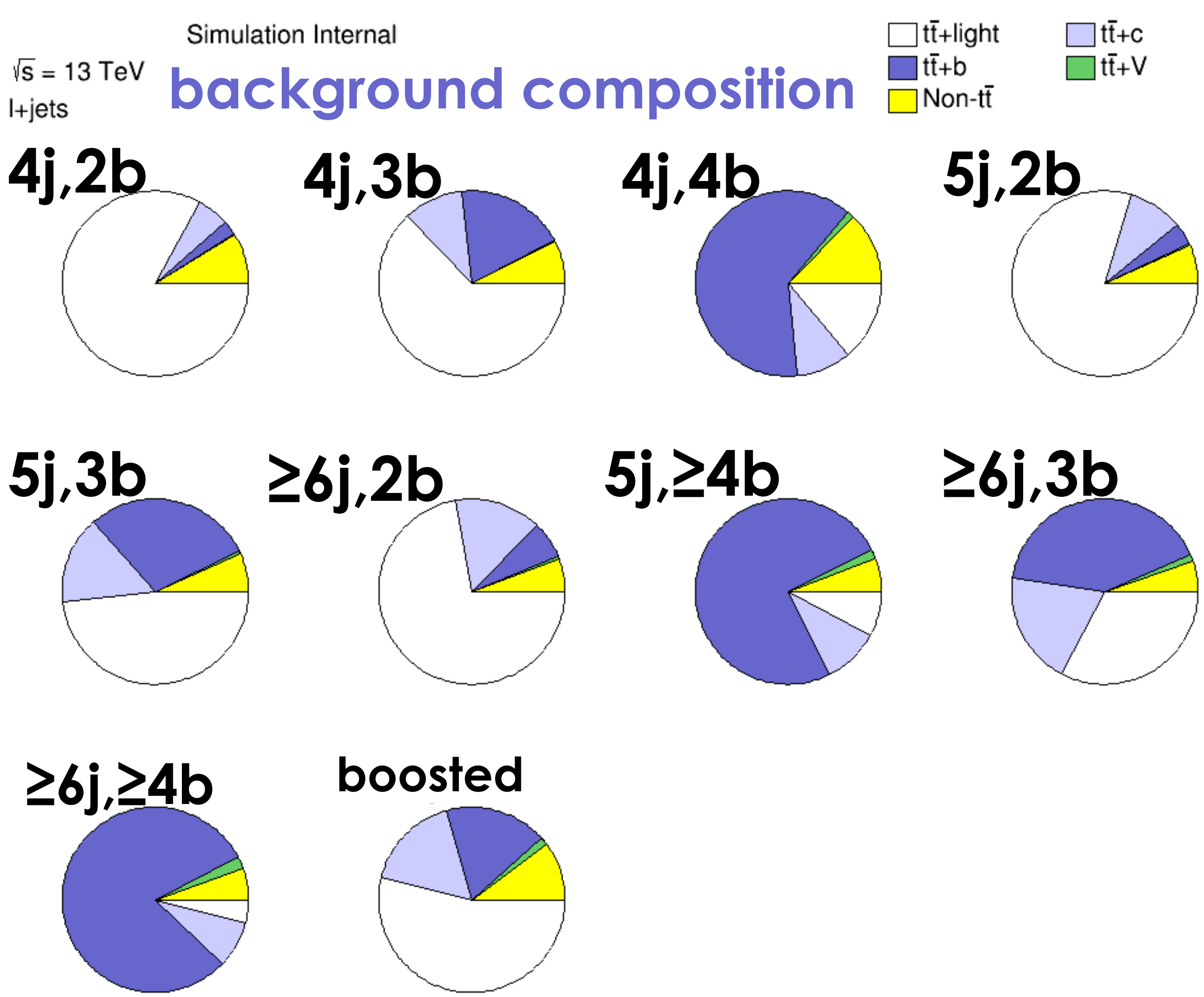


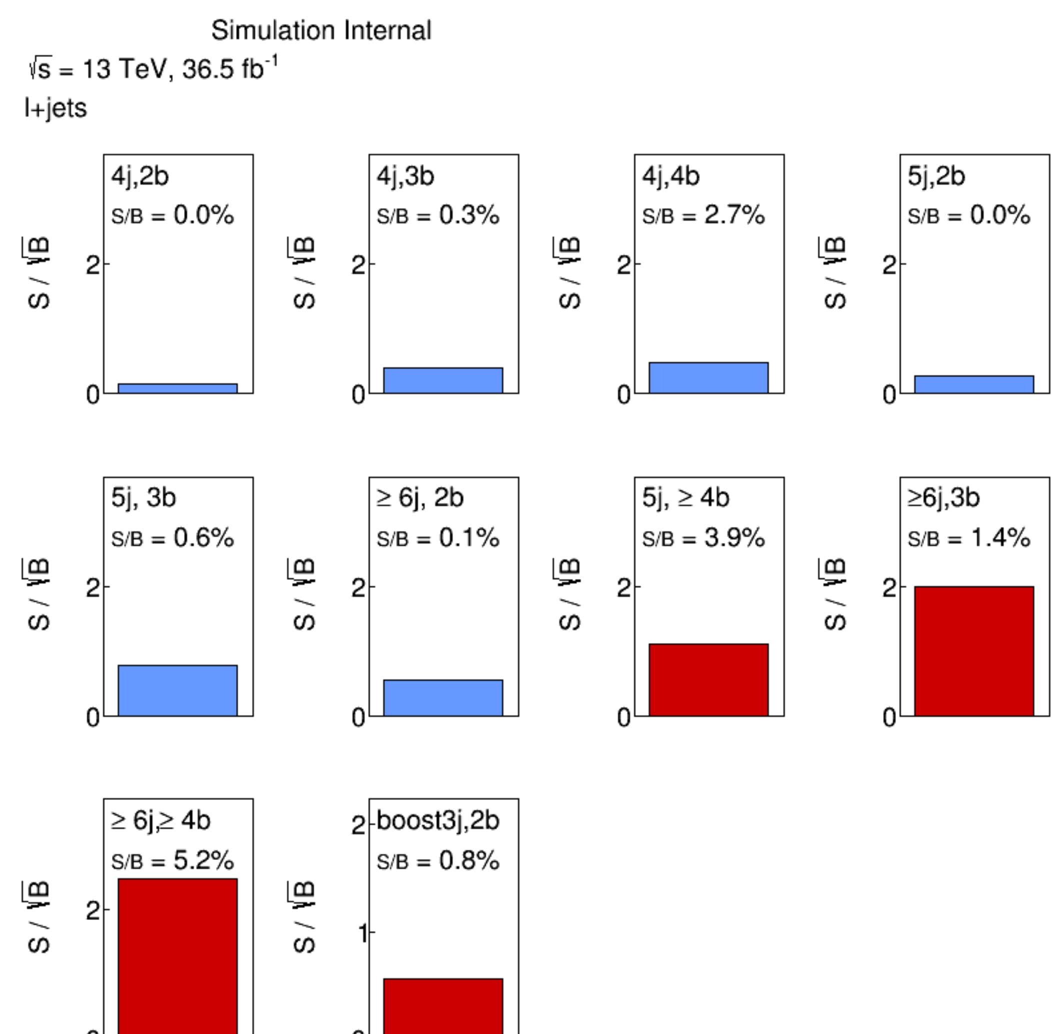
Control and signal regions

- **Common pre-selection**: exactly 1 lepton, ≥ 3 jets, ≥ 2 b-jets.
- To optimize the analysis sensitivity, **division of the phase space into several regions** (defined by the jet and b-jet multiplicity):
  - where **ttH is enhanced** wrt to background (high S/B and S/√B);
  - **dominated by background**, allowing a tighter constraint of backgrounds and systematic uncertainties in the analysis final step (fit).

Already done by the resolved analysis

now adding the boosted region!





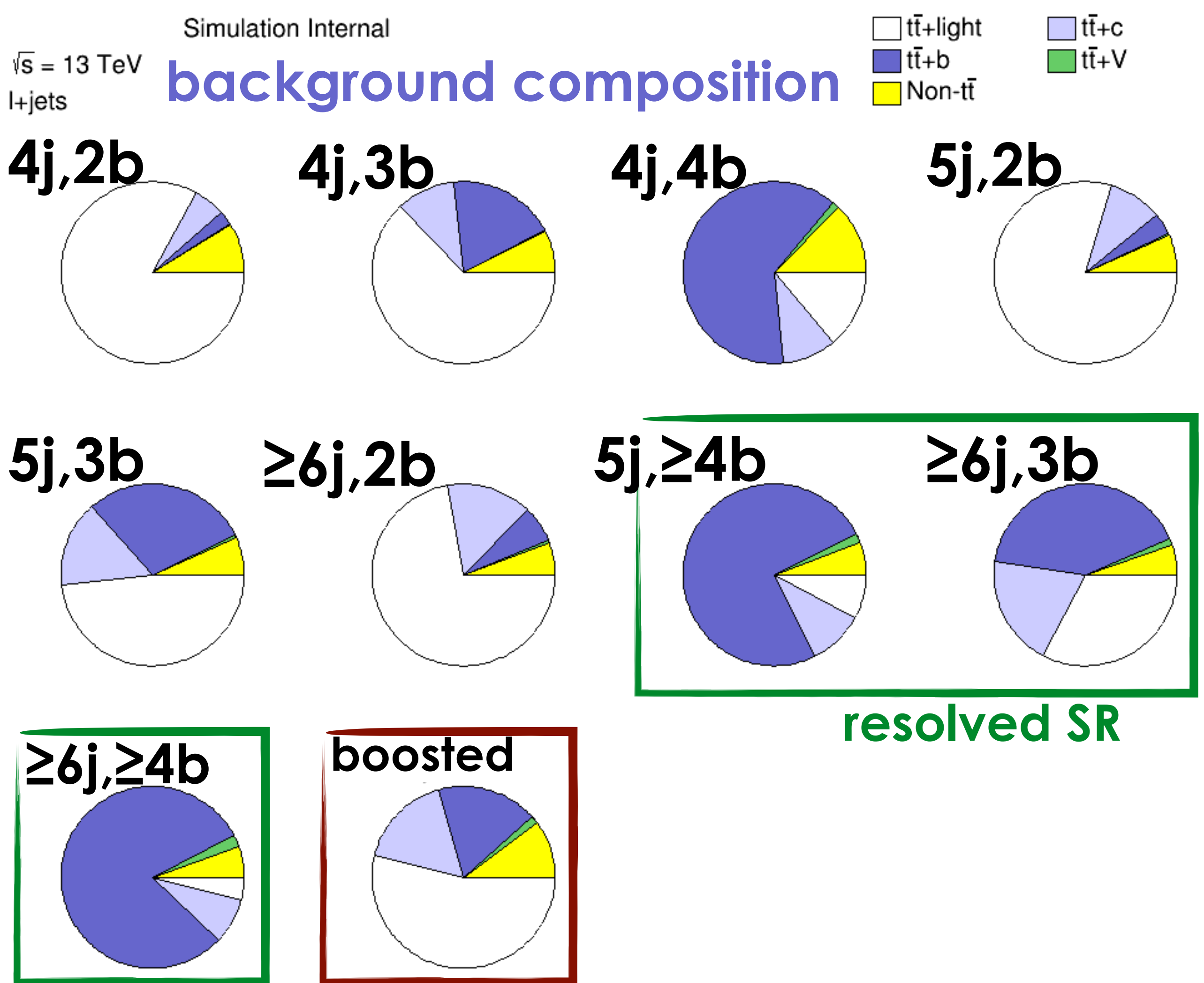
disjointed regions!

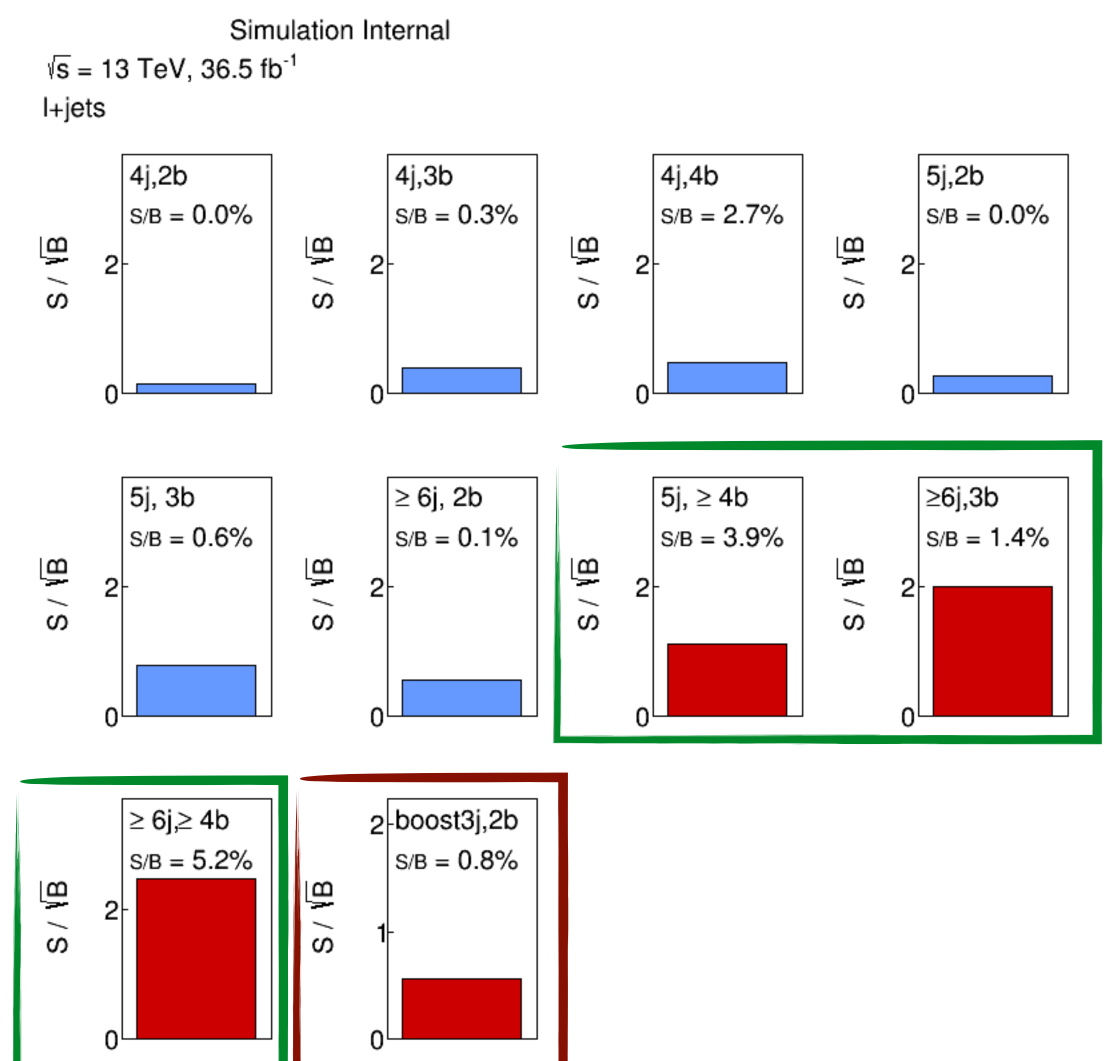
Control and signal regions

- **Common pre-selection**: exactly 1 lepton, ≥ 3 jets, ≥ 2 b-jets.
- To optimize the analysis sensitivity, **division of the phase space into several regions** (defined by the jet and b-jet multiplicity):
  - where **ttH is enhanced** wrt to background (high S/B and S/√B);
  - **dominated by background**, allowing a tighter constraint of backgrounds and systematic uncertainties in the analysis final step (fit).

Already done by the resolved analysis

now adding the boosted region!



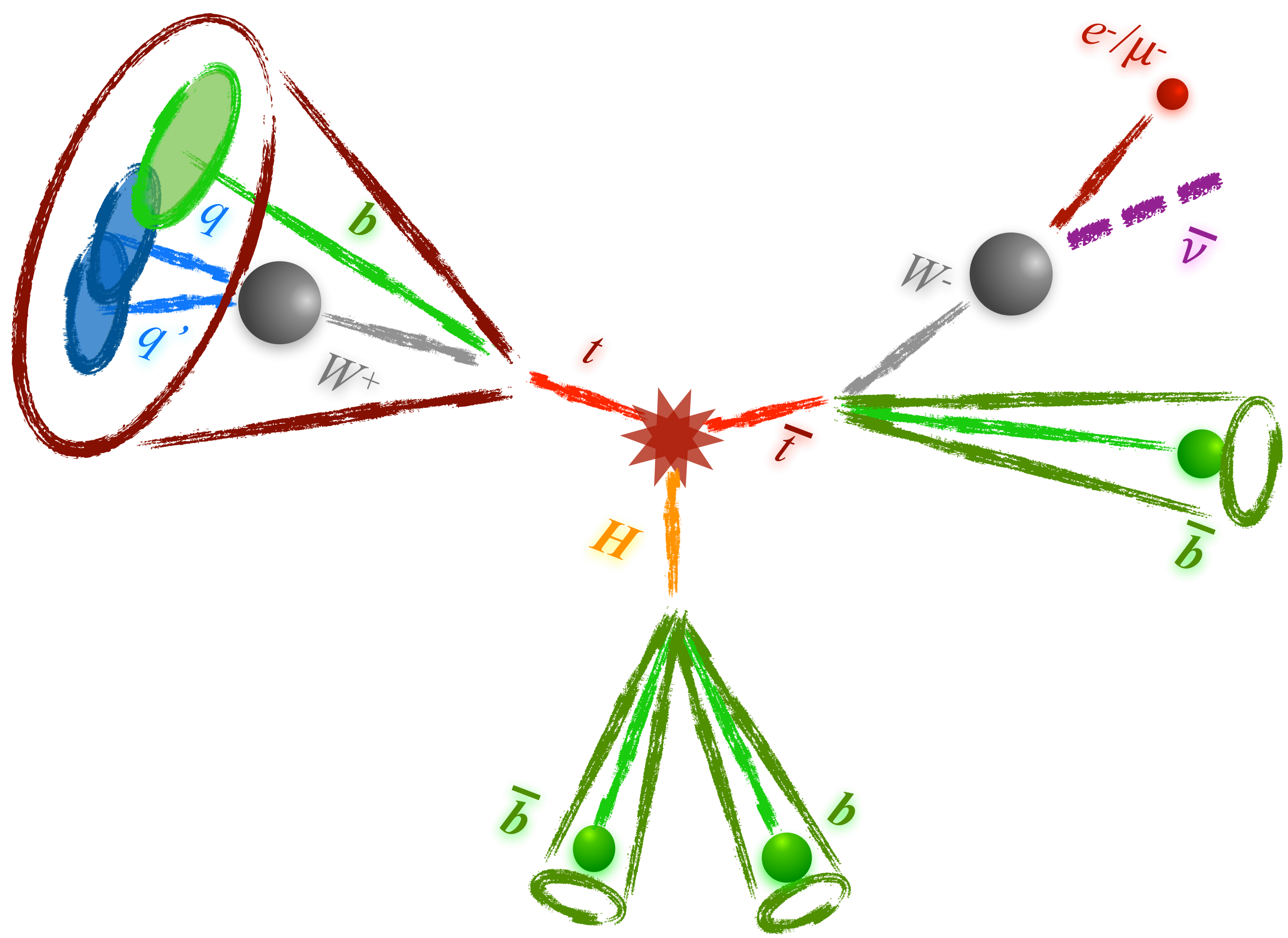


disjointed regions!

Boosted event selection

- Signal selection definition is based on **number of additional jets and b-jets**:
  - exactly **1 lepton** (e or μ);
  - ≥ **1 large-R jet** ($p_T \geq 250$ GeV);
  - ≥ **1 top-tagged large-R jet (80%)**;
  - overlap removal between top and electron: **ΔR(J,el) < 1.0** ;
  - ≥ **3 additional small-R jets** ($p_T \geq 25$ GeV);
  - ≥ **2 additional b-tagged jets (70%)**.

outside any top-tagged jet:
ΔR(j,T) > 1.0

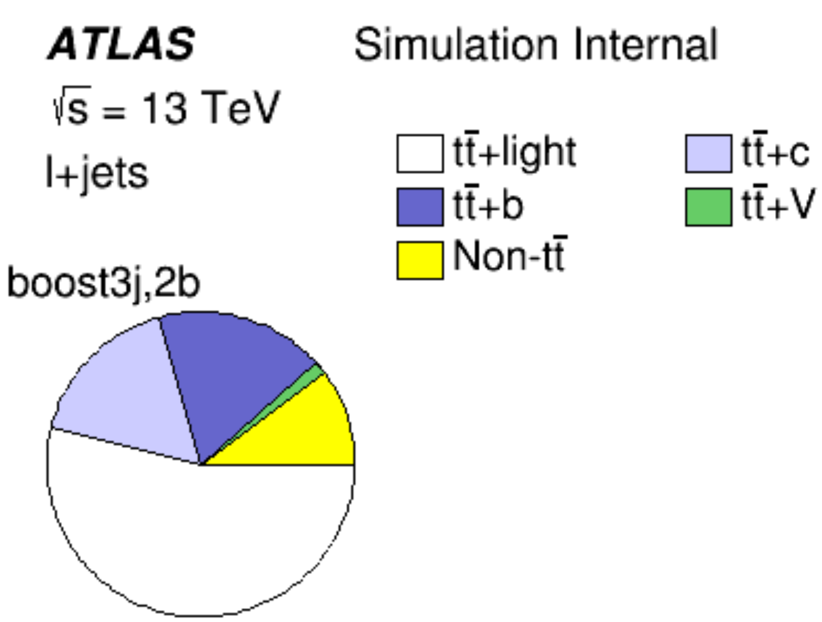


Boosted event selection

- Signal selection definition is based on **number of additional jets and b-jets**:
  - exactly **1 lepton** (e or μ);
  - ≥ **1 large-R jet** ($p_T \geq 250$ GeV);
  - ≥ **1 top-tagged large-R jet (80%)**;
  - overlap removal between top and electron: **$\Delta R(J,el) < 1.0$** ;
  - ≥ **3 additional small-R jets** ($p_T \geq 25$ GeV);
  - ≥ **2 additional b-tagged jets (70%)**.

outside any top-tagged jet:
$\Delta R(j,T) > 1.0$

Background composition



36.5 fb^{-1}

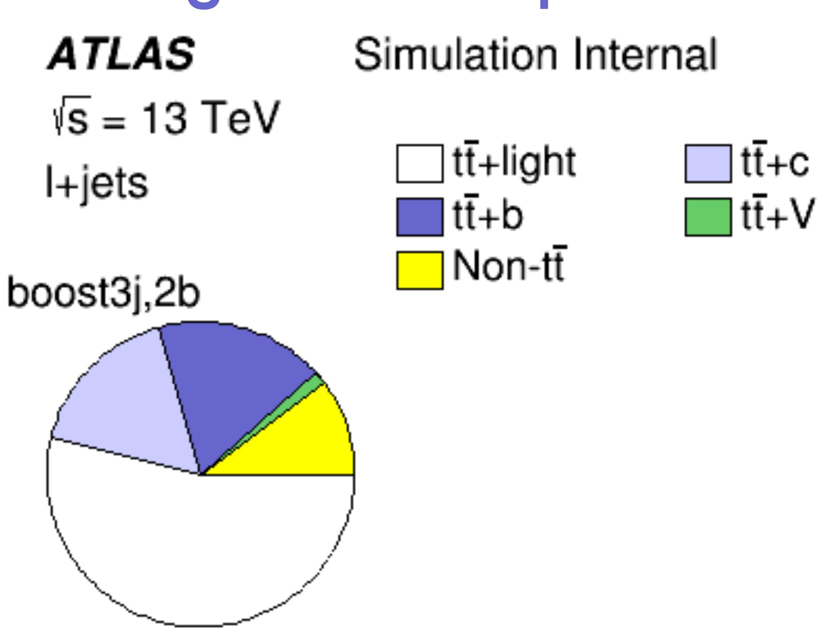
PROCESS	YIELDS
ttH	139
tt+ ≥ 1b	2632
tt+ ≥ 1c	2506
tt+light	7988
tt+V	201
single top	1059
V+jets	383
diboson	51
tX	29
data	16763

Boosted event selection

- Signal selection definition is based on **number of additional jets and b-jets**:
  - exactly **1 lepton** (e or μ);
  - ≥ **1 large-R jet** ($p_T \geq 250$ GeV);
  - ≥ **1 top-tagged large-R jet (80%)**;
  - overlap removal between top and electron: **$\Delta R(J,el) < 1.0$** ;
  - ≥ **3 additional small-R jets** ($p_T \geq 25$ GeV);
  - ≥ **2 additional b-tagged jets (70%)**.

outside any top-tagged jet: **$\Delta R(j,T) > 1.0$**

Background composition



36.5 fb^{-1}

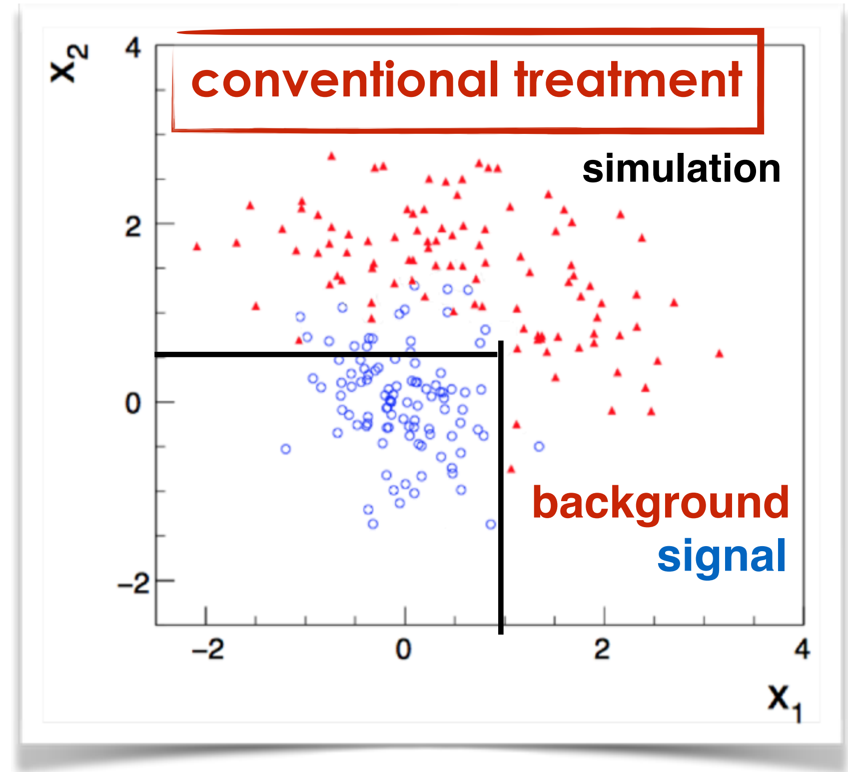
PROCESS	YIELDS
ttH	**139**
tt+ ≥ 1b	**2632**
tt+ ≥ 1c	**2506**
tt+light	**7988**
tt+V	201
single top	1059
V+jets	383
diboson	51
tX	29
data	**16763**

MultiVariate Analysis

Problem

ttH = 139/16763 !!

- Analysis aim: to identify **events** that are both **rare** and **overwhelmed** by a wide variety of processes that mimic the signal.
- Conventional approach by using cuts on individual kinematic variables **far from be optimal**!



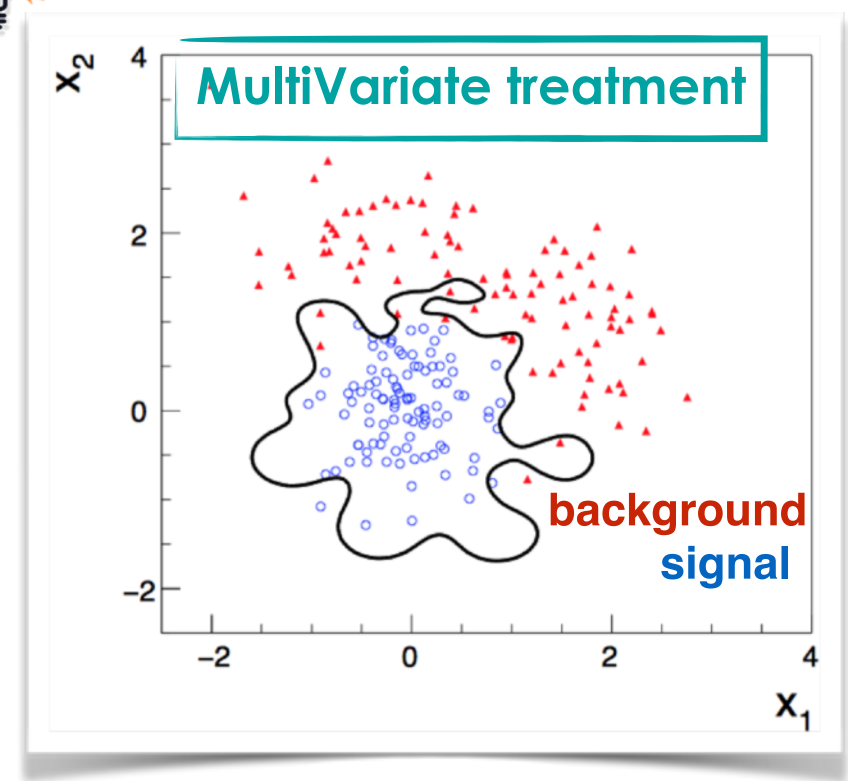
MultiVariate Analysis

Problem

- Analysis aim: to identify **events** that are both **rare** and **overwhelmed** by a wide variety of processes that mimic the signal.
- Conventional approach by using cuts on individual kinematic variables **far from be optimal**!

Solution: MultiVariate Analysis (MVA)

1. choice of **set of variables**, characterising an event;
2. application of **non-linear cuts** on signal and background samples;
3. define a function (**classifier**) that, using the **discriminating variables**, is able to identify each event of the real data belonging to the signal or to the background category.



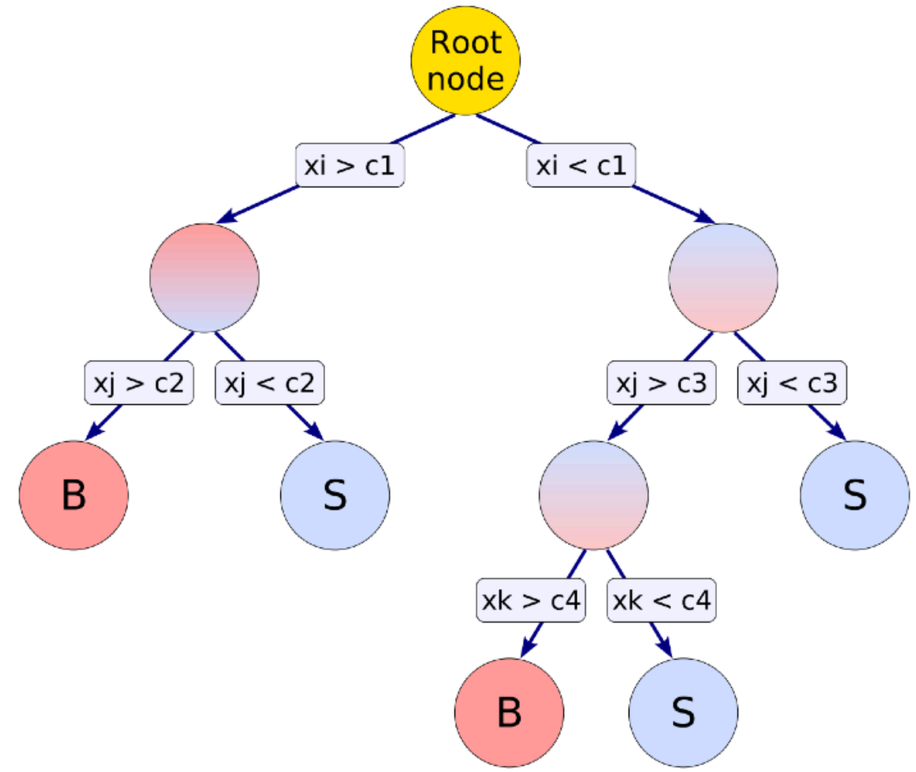
The algorithm "**learns**" signal and background characteristics (**training**) and assigns a weight to each event (~ probability that event is signal or background).

Many different algorithms available (Neural Networks, Boosted Decision Tree, Likelihood, …).

MVA in the ttH channel

Boosted Decision Tree (BDT)

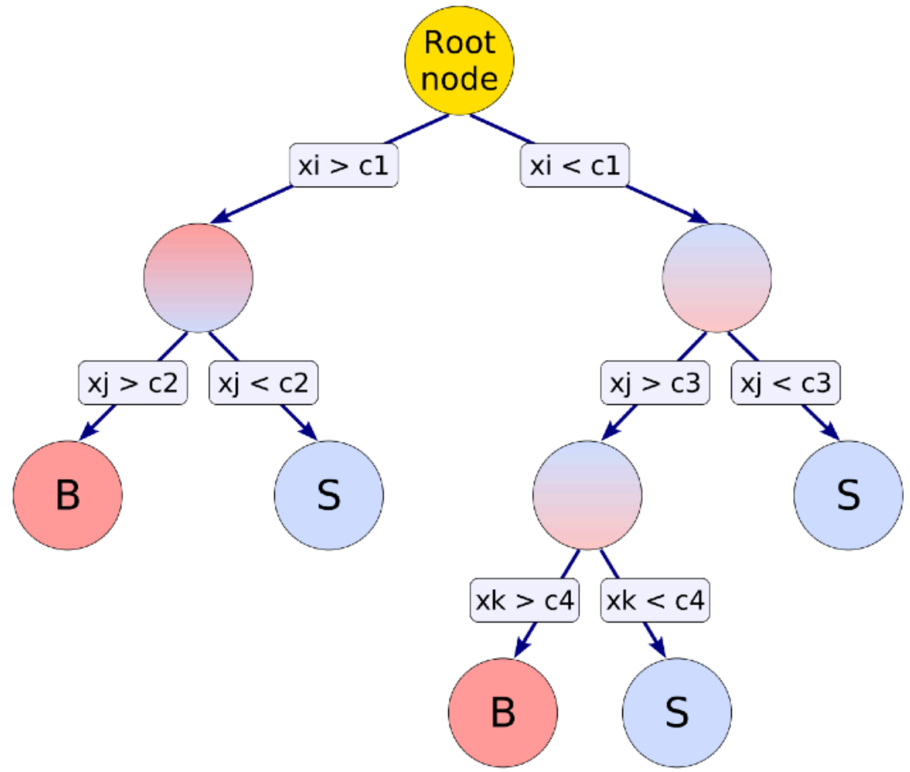
- **Decision Tree** = **sequence of binary splits** applied to the data, using discriminating variables.
- In order to improve the performance of the algorithm, a "forest" of **binary tree structured classifier** is considered.
- The **final leaves are labelled as signal or background**, depending on the majority of events in the respective node.



MVA in the ttH channel

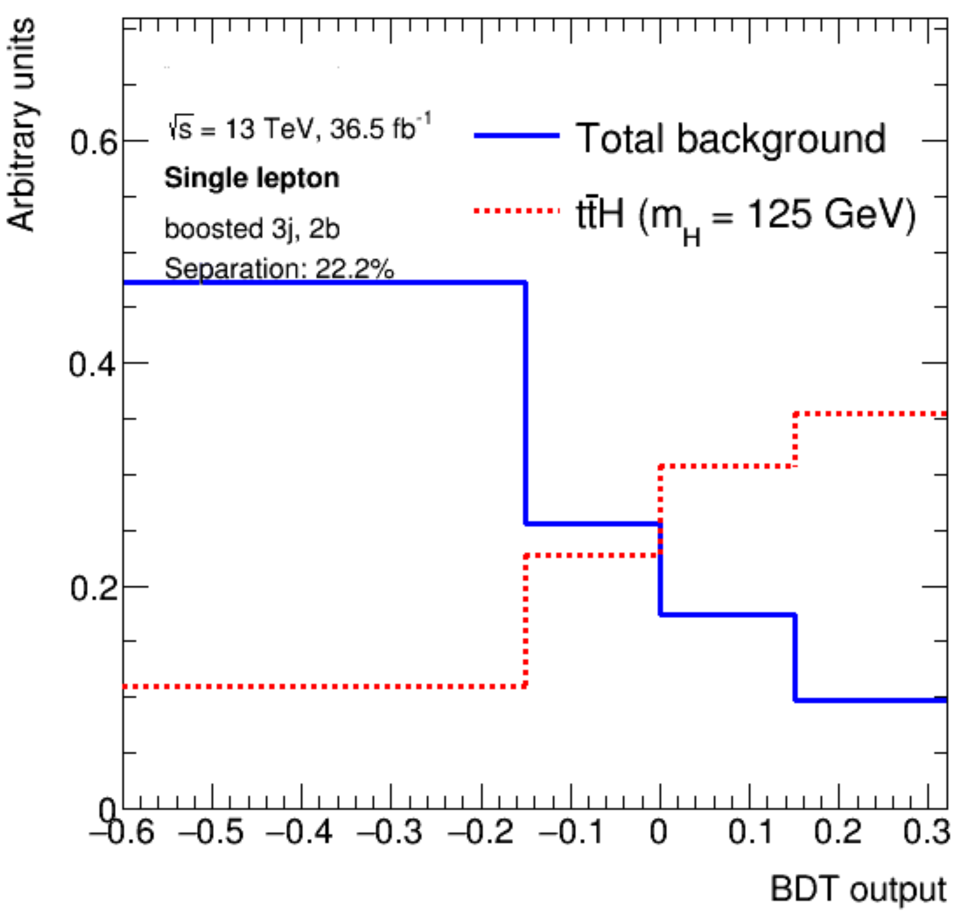
Boosted Decision Tree (BDT)

- **Decision Tree** = **sequence of binary splits** applied to the data, using discriminating variables.
- In order to improve the performance of the algorithm, a "forest" of **binary tree structured classifier** is considered.
- The **final leaves are labelled as signal or background**, depending on the majority of events in the respective node.



Procedure in the ttH channel

1. Choice of a set of **10 kinematic variables**:
    1. optimised from a large set of initial variables, looking at the **separation power**, **importance ranking** and **correlation** amongst them;
    2. **substructure**, **Higgs reconstruction** and **global event topology** related observables;
2. **training** on signal and background events;
3. results: **BDT discrimination output** (gives the best separation between signal and background on the real data as well).



Signal measurement technique

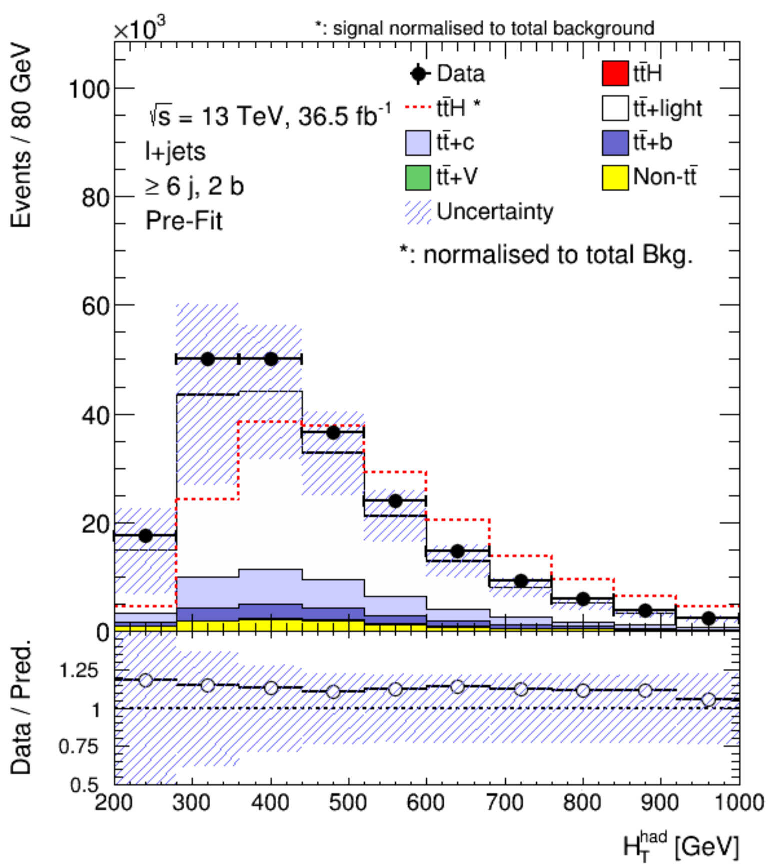
Ingredients

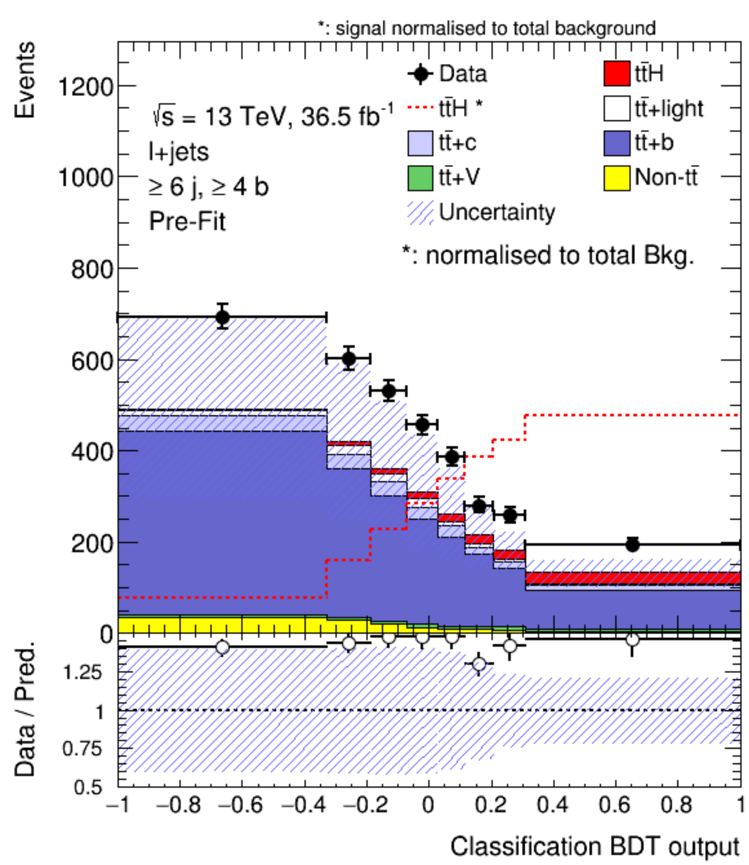
- **BDT** distributions in SRs;
- **H_T** ($= \Sigma_i\, p_{Ti}$) distributions in CRs;
- **P** depends on estimated number of events in each bin (function of μ);
- set of parameters to model the systematics uncertainties (**Nuisance Parameters**);
- **hypothesis**: S+B (μ = 1) or onlyB (μ = 0).

Recipe

In order to **test for signal presence** in the channel:

1. build a **likelihood** as a product of P terms over all the bins of the distributions;
2. perform a **fit** in the signal and control regions;
3. find a best-fit value of the **signal strength $\mu = \sigma/\sigma_{SM}$**;
4. put a **upper limit on μ @ 95% CL**.





Systematics included:

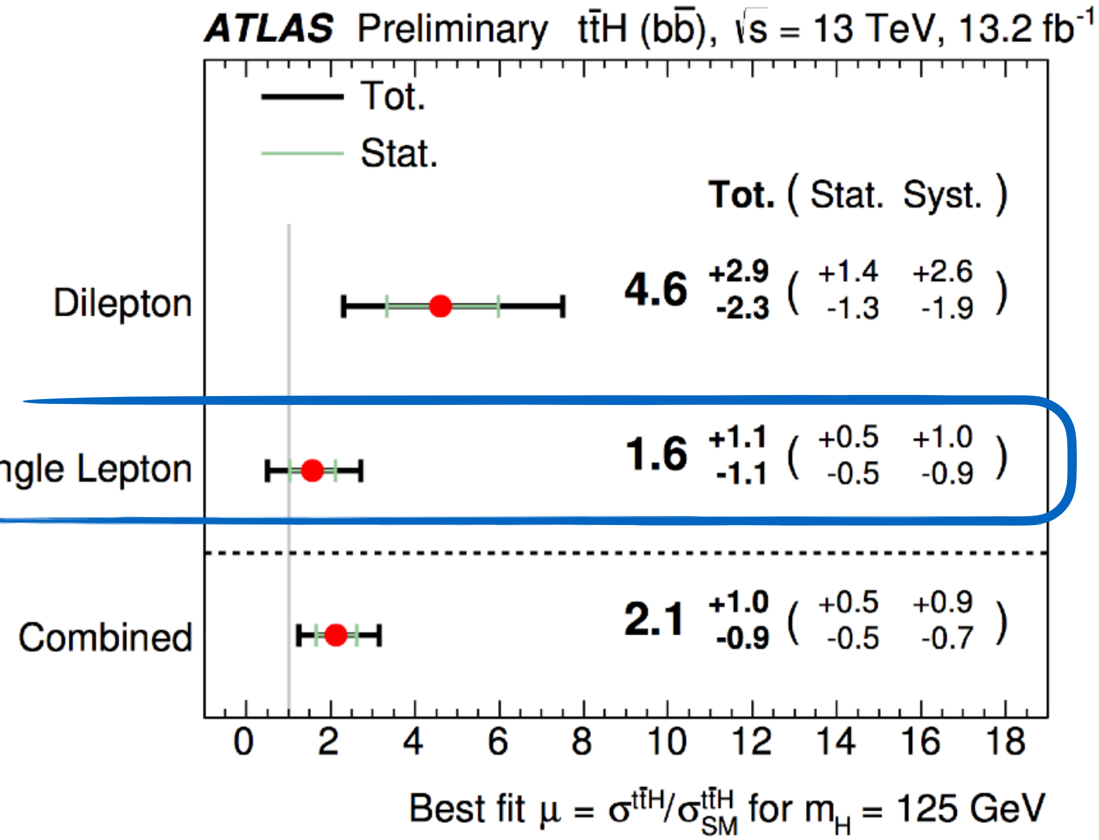
- luminosity (2.1% for 2015+2016);
- JES and JER;
- Jet Flavour Tagging;
- Light leptons;
- Large-R jets;
- Signal modelling;
- Background modelling.

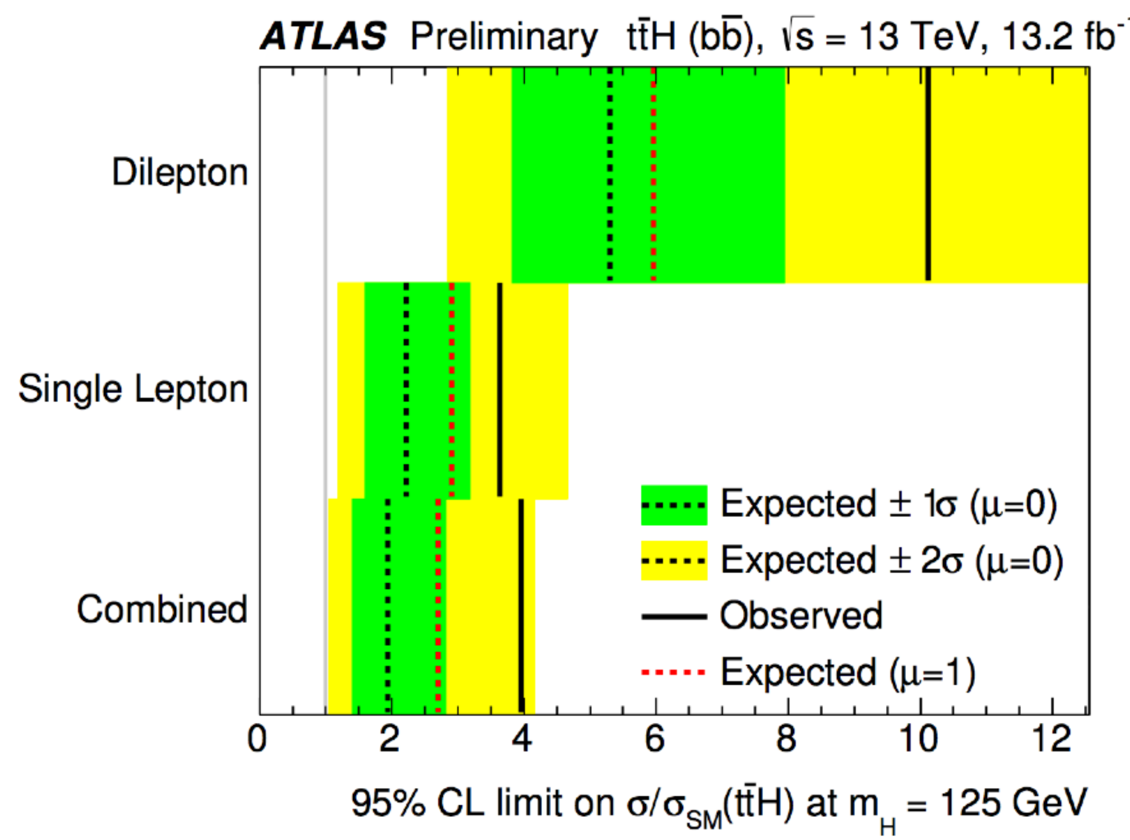
Results of resolved analysis

before my thesis results

- **First Run-2 results**: only resolved channels, only **13.2 fb⁻¹**.

ATLAS-COM-CONF-2016-044





- **The observed signal strength is μ = 1.6 ± 1.1, in the single lepton channel.**
- **A μ higher than 3.4 is excluded at 95% CL.**

- **Limits of the measurement:**
  - statistical and systematic uncertainties (**~70% in the single lepton channel**);
  - reconstruction uncertainty (**only 12%** of time Higgs and top **correctly reconstructed**!).

- **Proposal:**
  - adding the boosted region to improve the sensitivity of the analysis.

Results: resolved and combined analyses (I)

- **Resolved and combined analyses**, with **36.5 fb^{-1}**:
  - best-fit values of the signal strength and the free-floating normalisation factors;
  - upper limit @ 95% CL on the signal strength.

Upper limits @ 95% CL

	Observed (data)	Expected ($\mu = 0$) Median	$\pm 1\sigma$	$\pm 2\sigma$
Resolved*	2.3	1.0	[0.7,1.4]	[0.5,1.9]
Combined	2.0	0.9	[0.7,1.3]	[0.5,1.9]

Best-fit values $\mu = 1$

	Resolved	Combined
μ_{ttH}	1.5 ± 0.5	1.2 ± 0.5
k(tt+ ≥ 1b)	1.1 ± 0.2	0.9 ± 0.1
k(tt+ ≥ 1c)	1.6 ± 0.3	0.6 ± 0.2

First estimate of the limit on μ improvement!

- **Very strong motivation** to continue on this way;
- the **increasing of statistics benefits mainly the boosted region** wrt the resolved one!

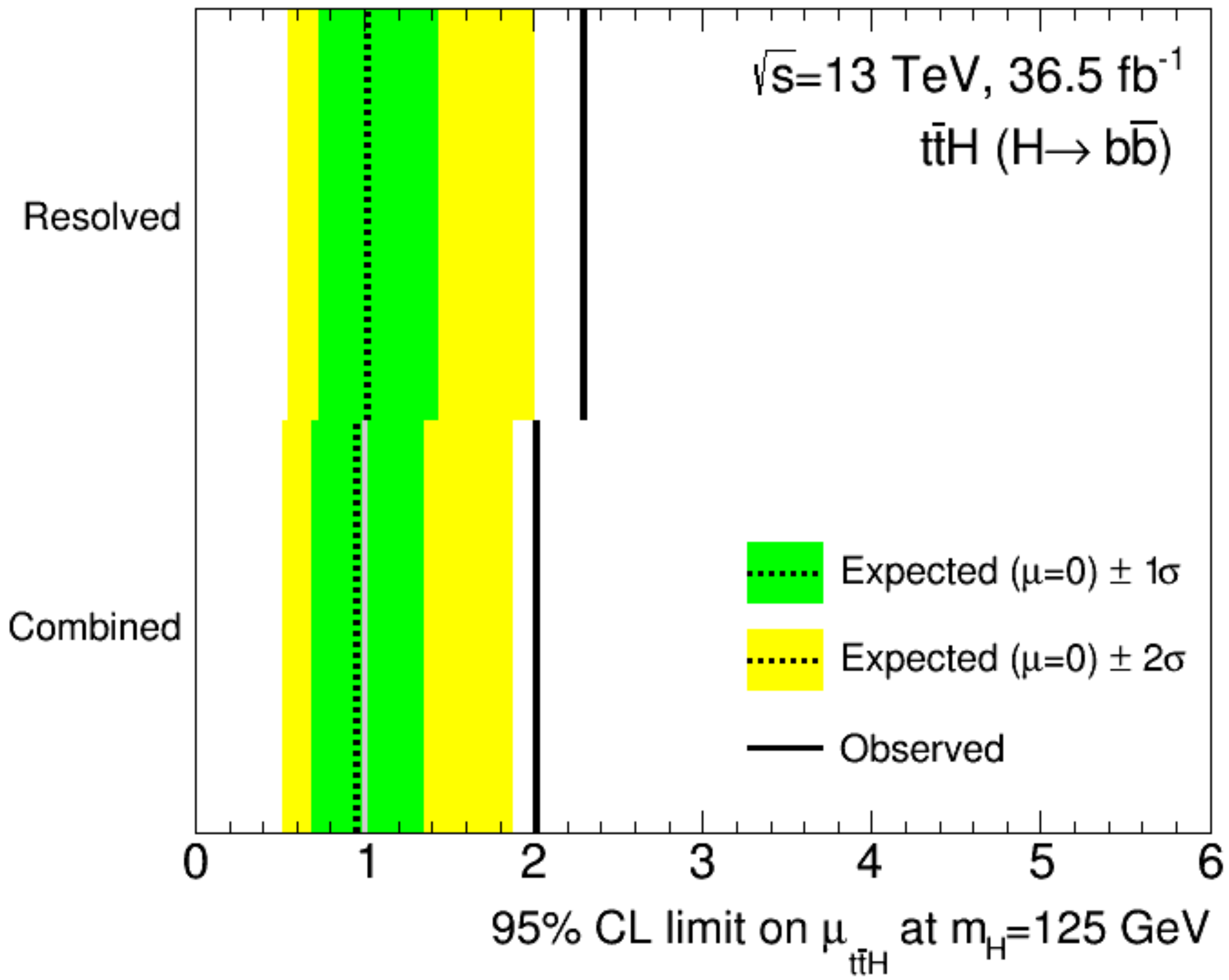
Systematic uncertainties NB!

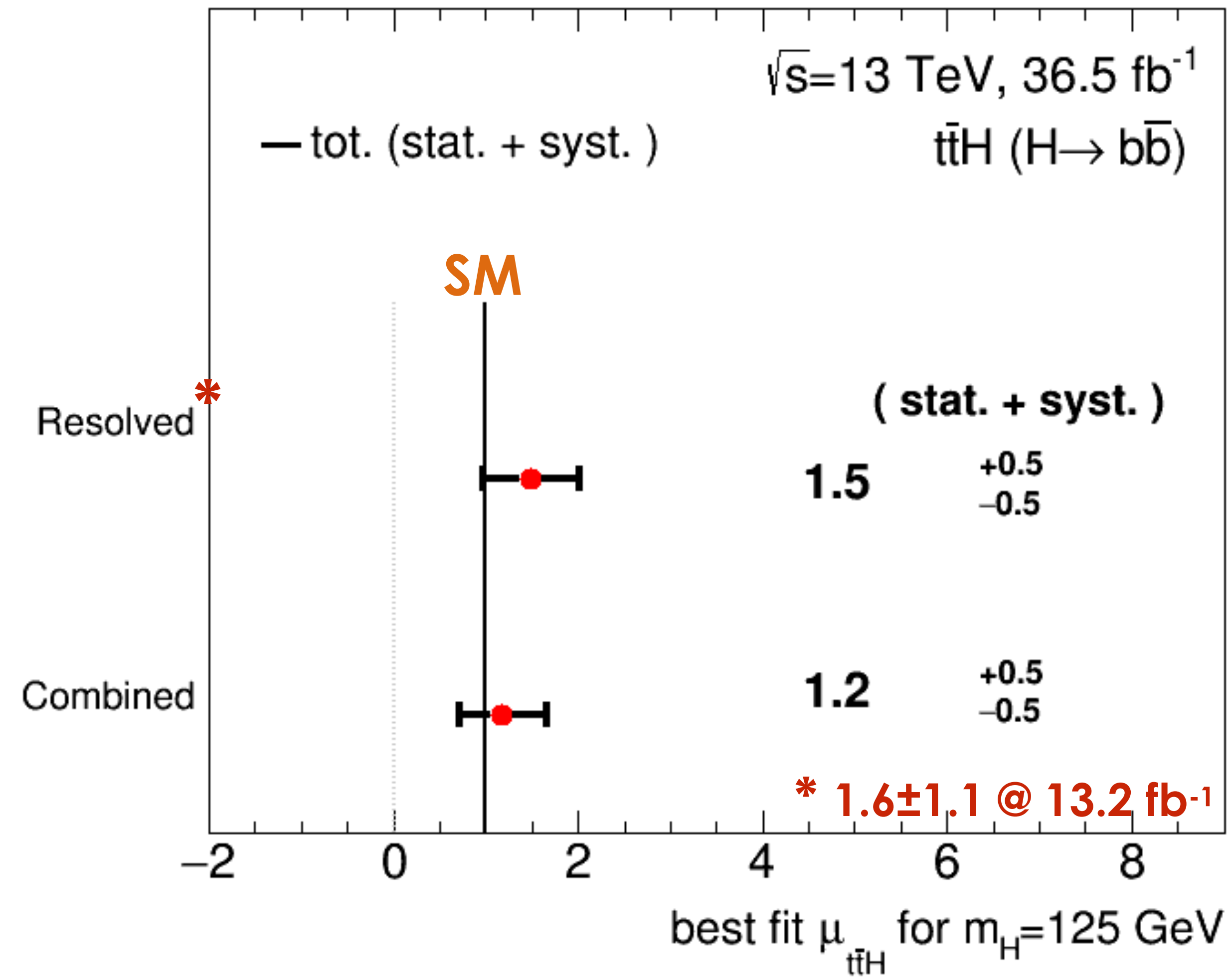
- Addition of systematics related to the **large-R jets reconstruction** and **top-tagging algorithm**.
- Nevertheless, **no increase of the overall uncertainty!**

*resolved l+jets channel only

Results: resolved and combined analyses (II)

- **Aims of the study:**
  - comparison of the resolved and the combined analyses results with the same configuration;
  - improvement of the sensitivity of the resolved analysis with the addition of a boosted region.





- **The observed signal strength is μ = 1.5 for resolved analysis and μ = 1.2 for the combined one.**
- **A μ higher than 2.3 and 2.0 is excluded at 95% CL, for the resolved and the combined analysis respectively.**

Conclusions (I)

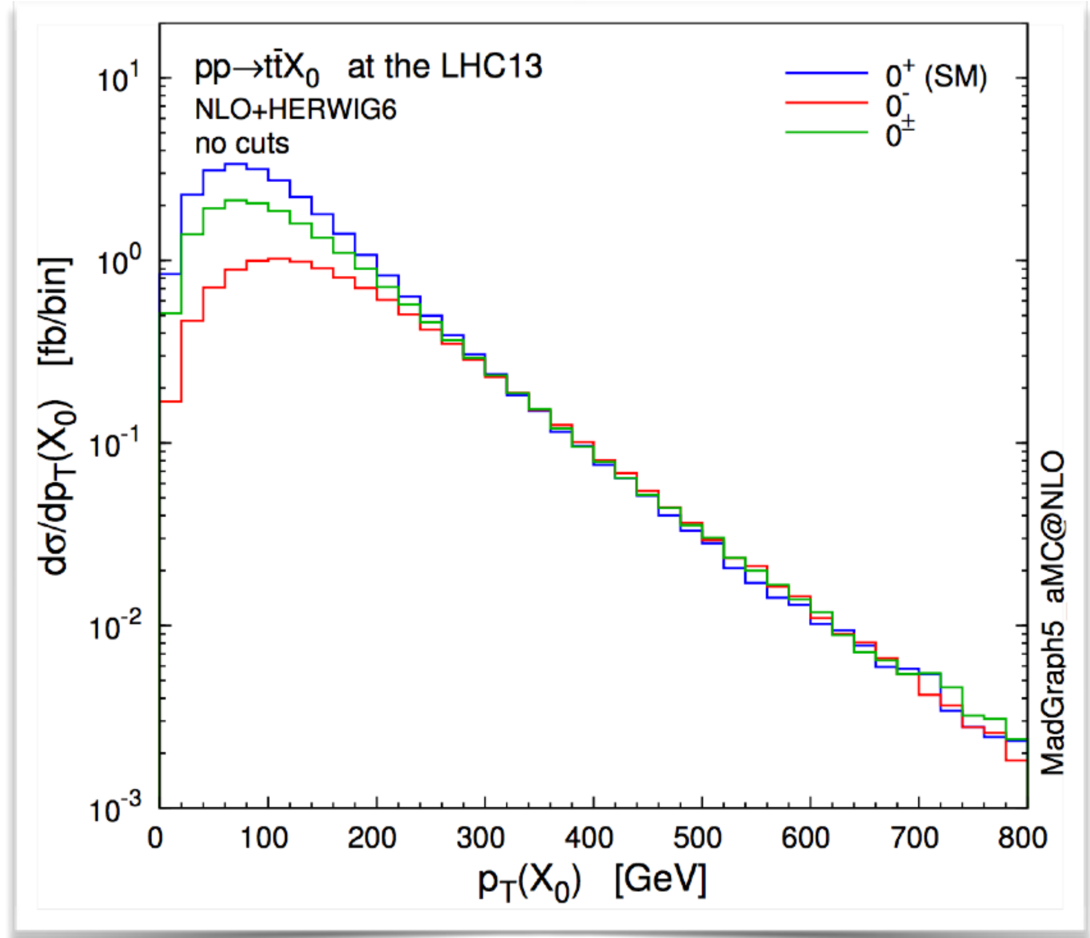
- **"The importance of being boosted $t\bar{t}H$"** - (~) Oscar Wilde
  - **$t\bar{t}H$ cross-section** shows the highest increase wrt the one of the other channels and allows a direct study to the Yukawa coupling;
  - **the "boosted" techniques** help to look at never observed kinematic regimes.
- **The results I have obtained:**
  - both the resolved and the combined results are **compatible with the prediction of the SM $t\bar{t}H$**, but **not enough sensitivity to exclude the null signal hypothesis**;
  - the addition of the **boosted channel constrains in a stronger way the upper limit on μ** to values that are closer to the SM predictions;
  - **paper publication in 2017** (in combination with the resolved analysis).

Conclusions (I)

- **"The importance of being boosted ttH"** - (~) Oscar Wilde
  - **ttH cross-section** shows the highest increase wrt the one of the other channels and allows a direct study to the Yukawa coupling;
  - **the "boosted" techniques** help to look at never observed kinematic regimes.
- **The results I have obtained:**
  - both the resolved and the combined results are **compatible with the prediction of the SM ttH**, but **not enough sensitivity to exclude the null signal hypothesis**;
  - the addition of the **boosted channel constrains in a stronger way the upper limit on μ** to values that are closer to the SM predictions;
  - **paper publication in 2017** (in combination with the resolved analysis).

Conclusions (II)

- **Future perspectives:**
  - with the actual statistics, **first CP studies of the Yukawa coupling** (equivalent to test the Higgs CP):
    - up to now, only through the **H->VV->4l channel**;
    - **measuring cross section in two disjointed regions** (resolved and boosted) could provide important information about the CP properties;
  - **more precise measurement of the signal strength** and **differential cross section** with more statistics (150 fb^{-1});
  - studies of the **structure and properties of the Yukawa Higgs-top coupling**, with even more statistics (300 fb^{-1}).



thanks to F. Maltoni for the collaboration

Post-thesis updates

Paper publication in 2017

arXiv:1712.08895v2

EUROPEAN ORGANISATION FOR NUCLEAR RESEARCH (CERN)

Phys. Rev. D. 97 (2018) 072016
DOI: 10.1103/PhysRevD.97.072016

CERN-EP-2017-291
15th May 2018

Search for the Standard Model Higgs boson produced in association with top quarks and decaying into a $b\bar{b}$ pair in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector

The ATLAS Collaboration

…it happened the 24th of December 2017!

Paper publication in 2017

Phys. Rev. D 97 (2018) 072003

PHYSICAL REVIEW D **97,** 072003 (2018)

Evidence for the associated production of the Higgs boson and a top quark pair with the ATLAS detector

M. Aaboud *et al.*[*]
(ATLAS Collaboration)

(Received 27 December 2017; published 9 April 2018)

A search for the associated production of the Higgs boson with a top quark pair ($t\bar{t}H$) is reported. The search is performed in multilepton final states using a data set corresponding to an integrated luminosity of 36.1 fb^{-1} of proton-proton collision data recorded by the ATLAS experiment at a center-of-mass energy $\sqrt{s} = 13$ TeV at the Large Hadron Collider. Higgs boson decays to WW^*, $\tau\tau$, and ZZ^* are targeted. Seven final states, categorized by the number and flavor of charged-lepton candidates, are examined for the presence of the Standard Model Higgs boson with a mass of 125 GeV and a pair of top quarks. An excess of events over the expected background from Standard Model processes is found with an observed significance of 4.1 standard deviations, compared to an expectation of 2.8 standard deviations. The best fit for the $t\bar{t}H$ production cross section is $\sigma(t\bar{t}H) = 790^{+230}_{-210}$ fb, in agreement with the Standard Model prediction of 507^{+35}_{-50} fb. The combination of this result with other $t\bar{t}H$ searches from the ATLAS experiment using the Higgs boson decay modes to $b\bar{b}$, $\gamma\gamma$ and $ZZ^* \rightarrow 4\ell$, has an observed significance of 4.2 standard deviations, compared to an expectation of 3.8 standard deviations. This provides evidence for the $t\bar{t}H$ production mode.

DOI: 10.1103/PhysRevD.97.072003

…**first evidence** (significance > 3σ) of the ttH production ever with the combination of all the ATLAS ttH analyses!

Paper publication in 2018

arXiv:1806.00425

EUROPEAN ORGANISATION FOR NUCLEAR RESEARCH (CERN)

ATLAS EXPERIMENT

CERN

Submitted to: PLB

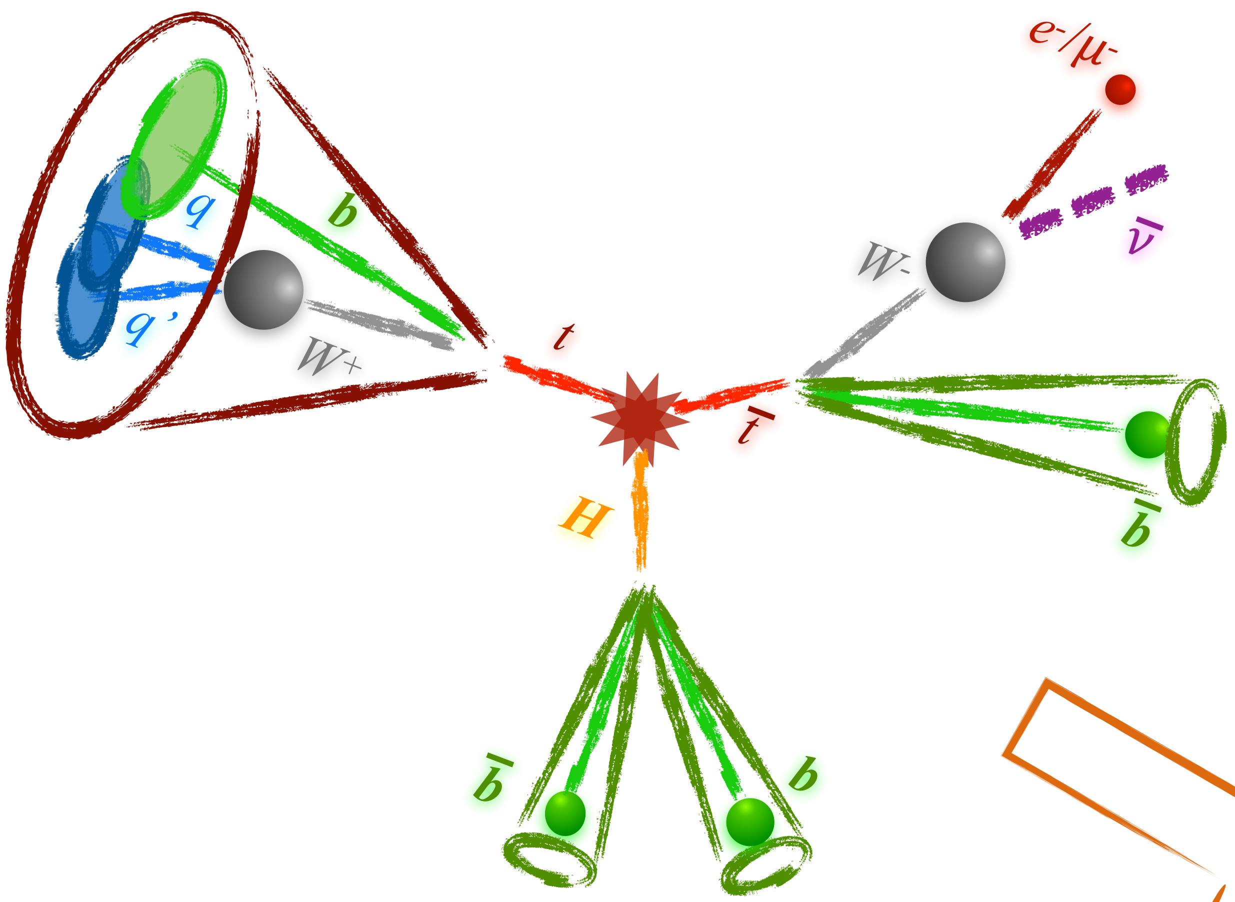
CERN-EP-2018-138
4th June 2018

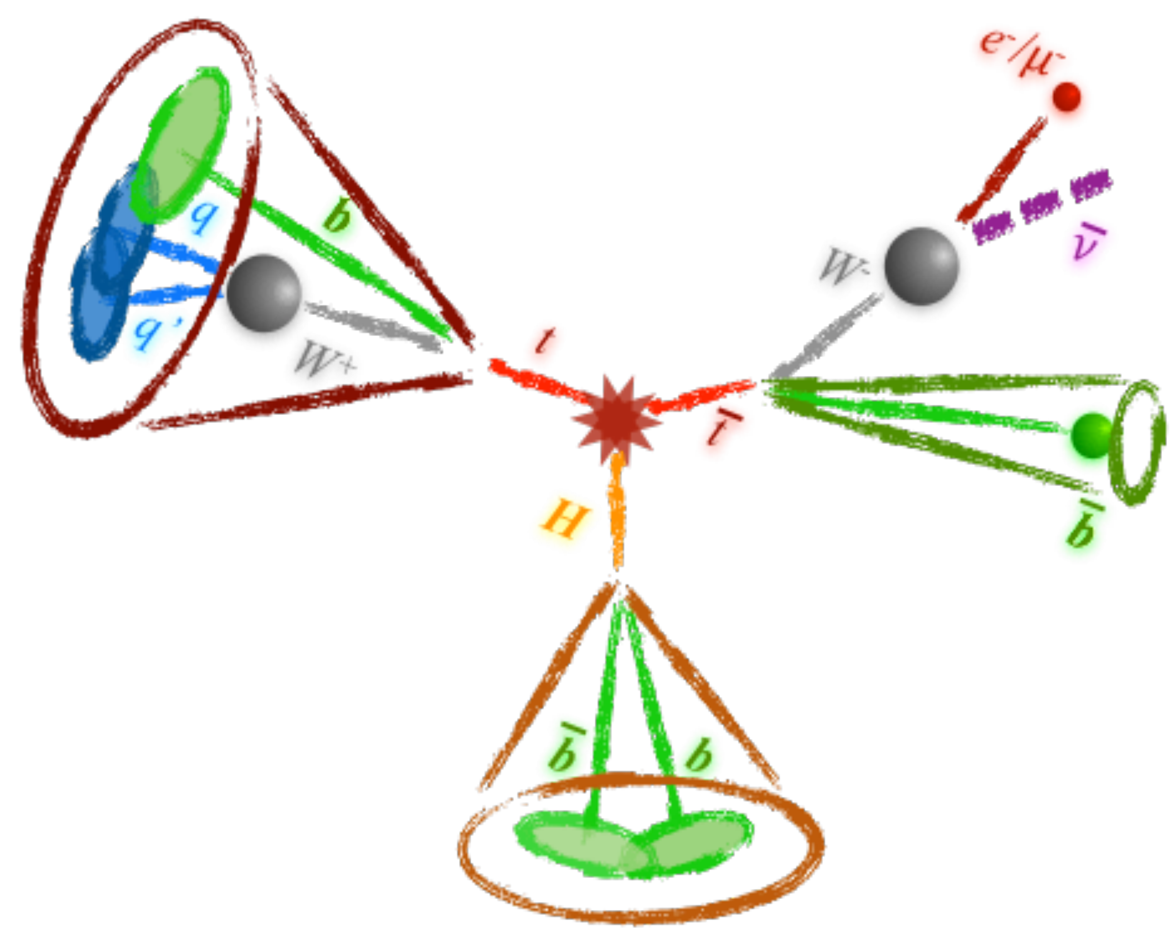
Observation of Higgs boson production in association with a top quark pair at the LHC with the ATLAS detector

The ATLAS Collaboration

...**first observation** (significance > 5σ) of the ttH production ever with the combination of all the ATLAS ttH analyses!

New signal region definition



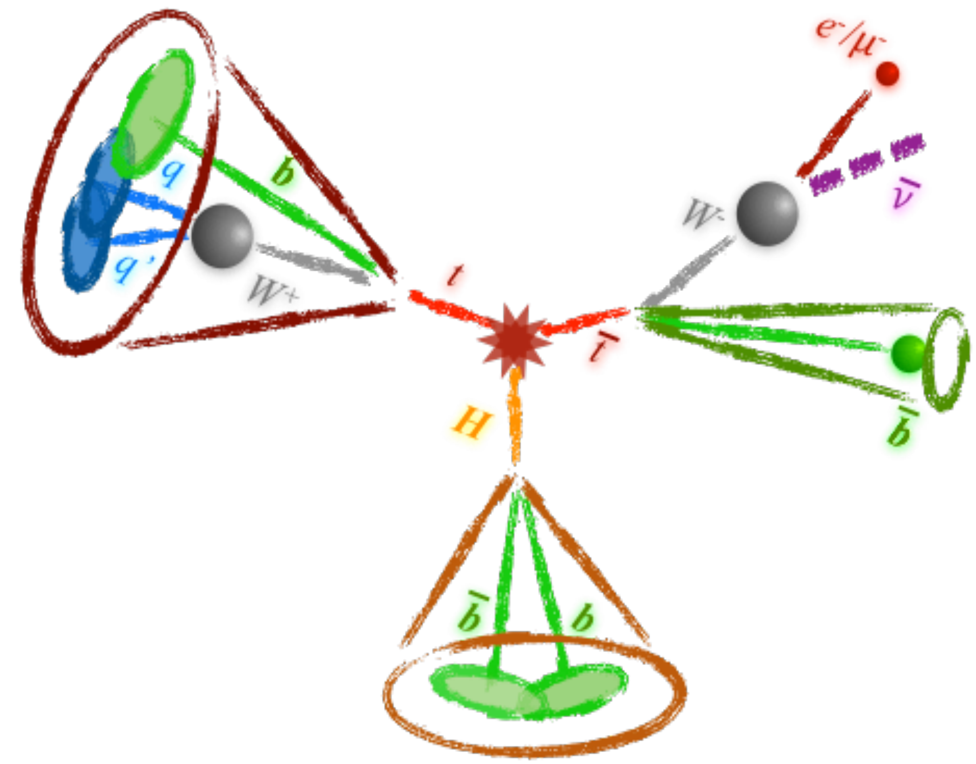


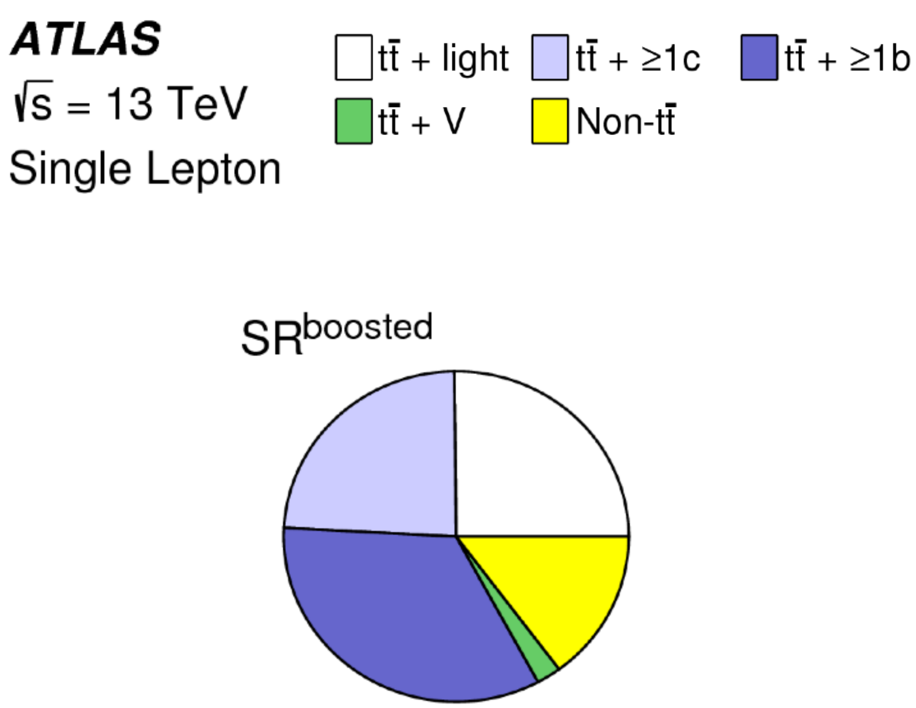
Improvement of the clustering algorithm

Re-clustering techniques (II)

New signal region definition

- exactly **one lepton**;
- **one Higgs candidate:** one reclustered jet (p_T > 200 GeV) with two associated b-tagged jets (85%);
- **one Top candidate:** one reclustered jet (p_T > 250 GeV) with one associated b-tagged jet (85%) and one non-b-tagged jet;
- **one b-tagged jet** outside the two reclustered jets.

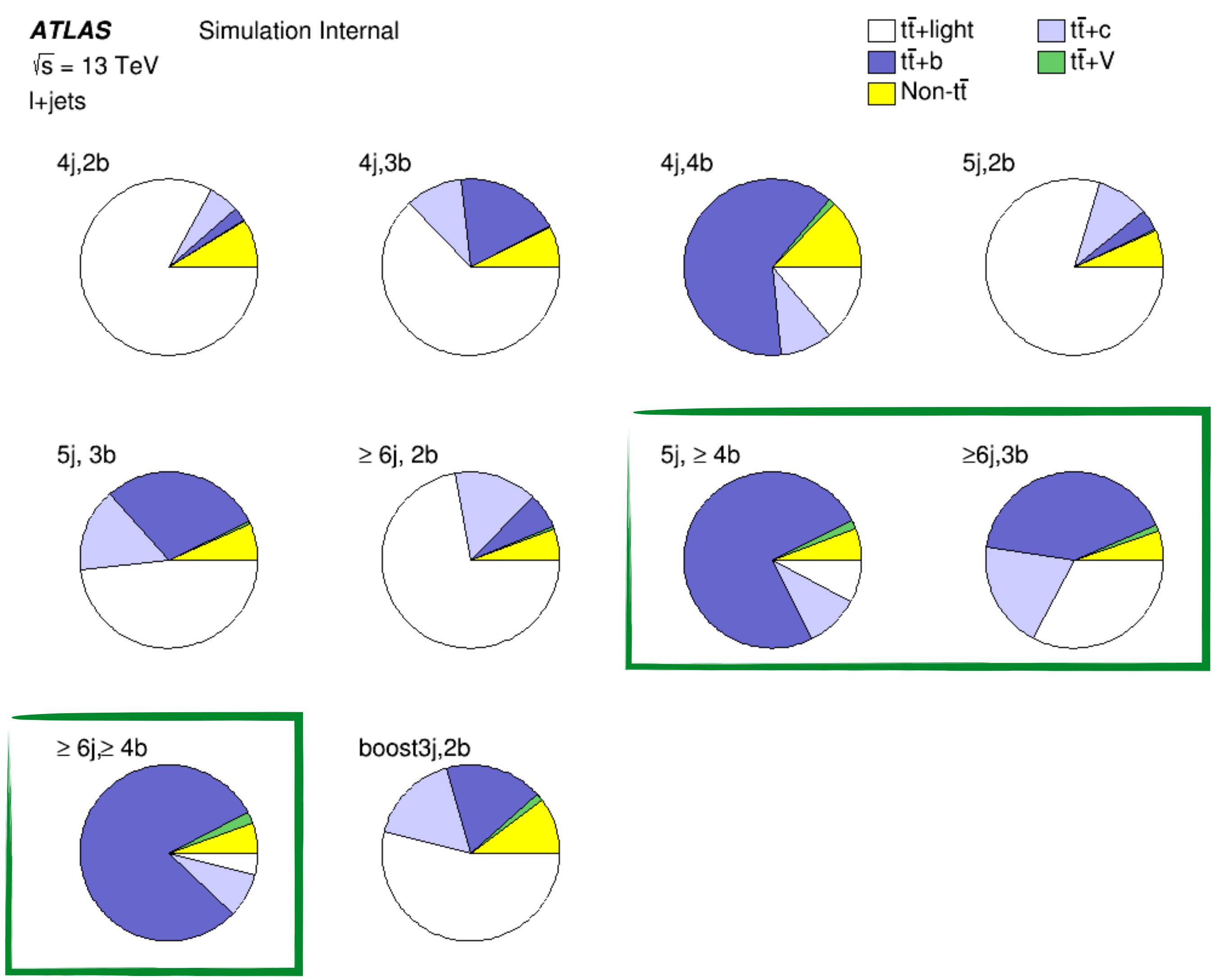




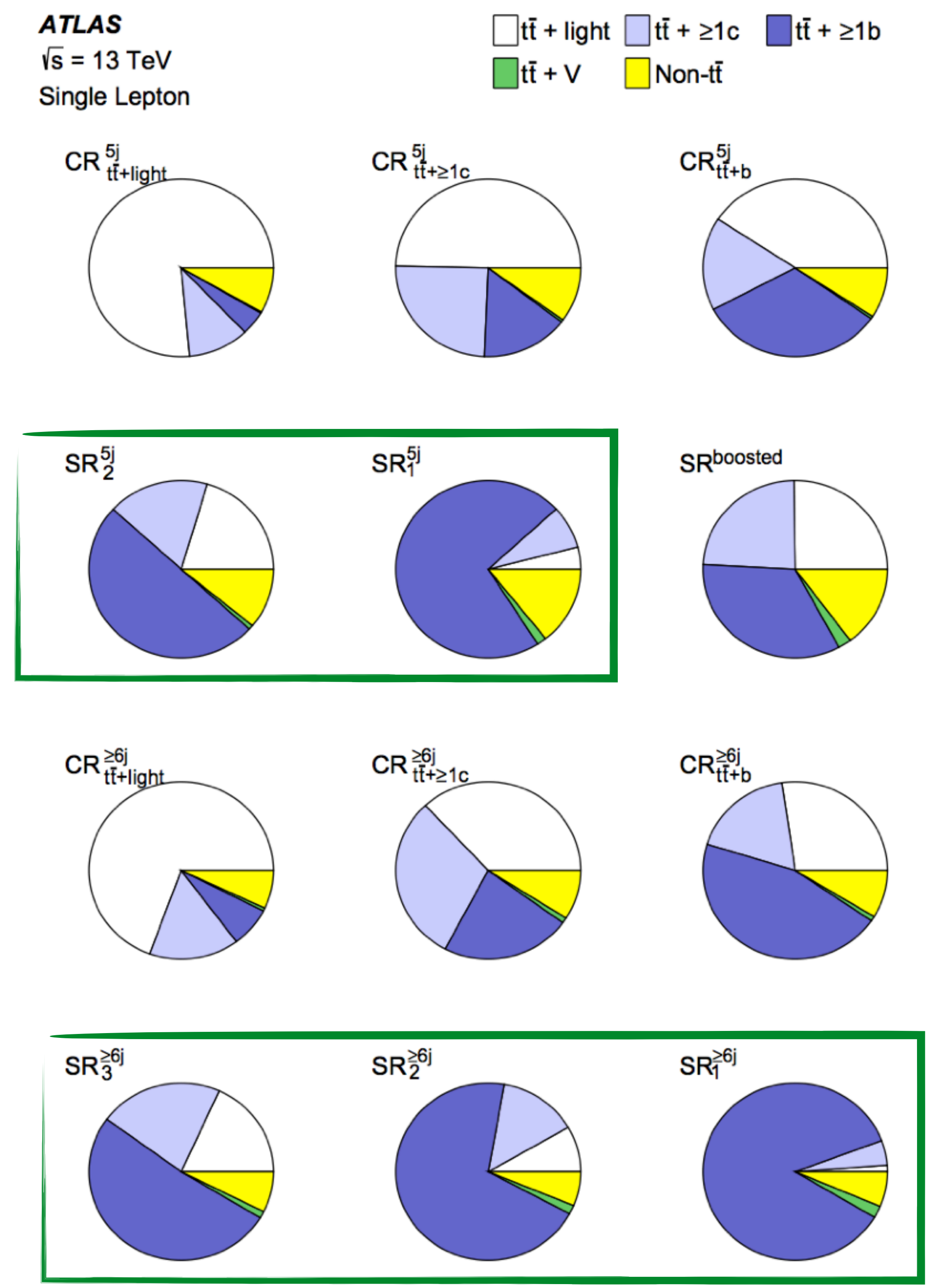
Process	Pre-fit Yield
tt+ light	180 ± 120
tt+ ≥1c	168 ± 70
tt+ ≥1b	236 ± 89
tt+V	16 ± 3
non-tt	104 ± 30
ttH	16 ± 2
data	740

Optimisation of resolved classification

BEFORE



AFTER



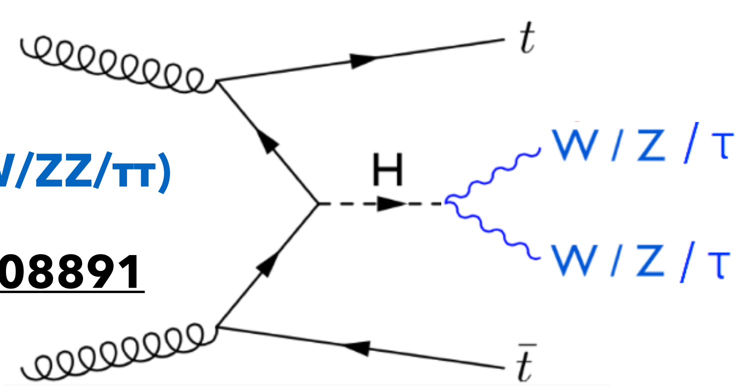
- classification on jets and b-tagged jets multiplicity;
- **new:** further split into SRs and UltraPure SRs, CRs splitting depends on background composition

First observation of ttH in ATLAS!

13 TeV, 36.1 fb^{-1}

ttH(WW/ZZ/ττ)

1712.08891



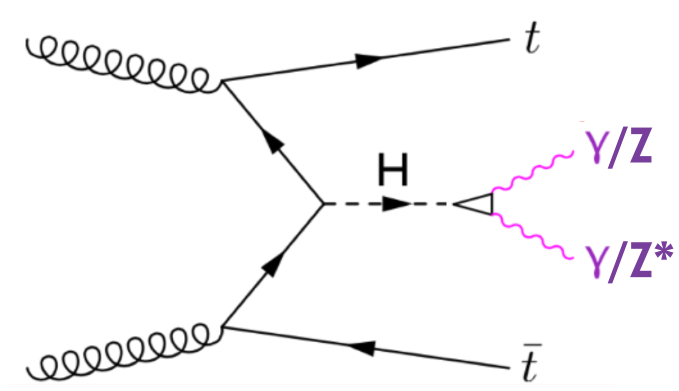
13 TeV, 79.8 fb^{-1}

ttH in H->γγ

1806.00425

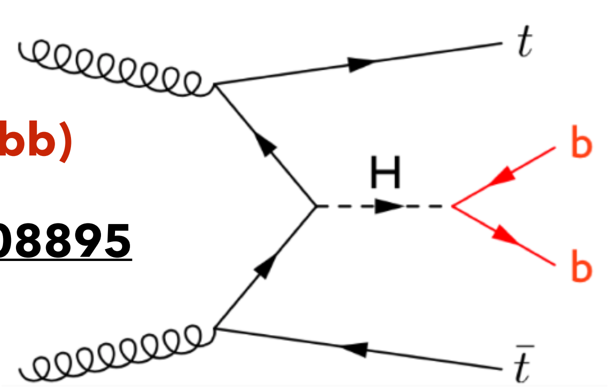
ttH(4l)

1806.00425



ttH(bb)

1712.08895



13 TeV, 36.1 fb^{-1}

13 TeV, 36.1 - 79.8 fb^{-1}

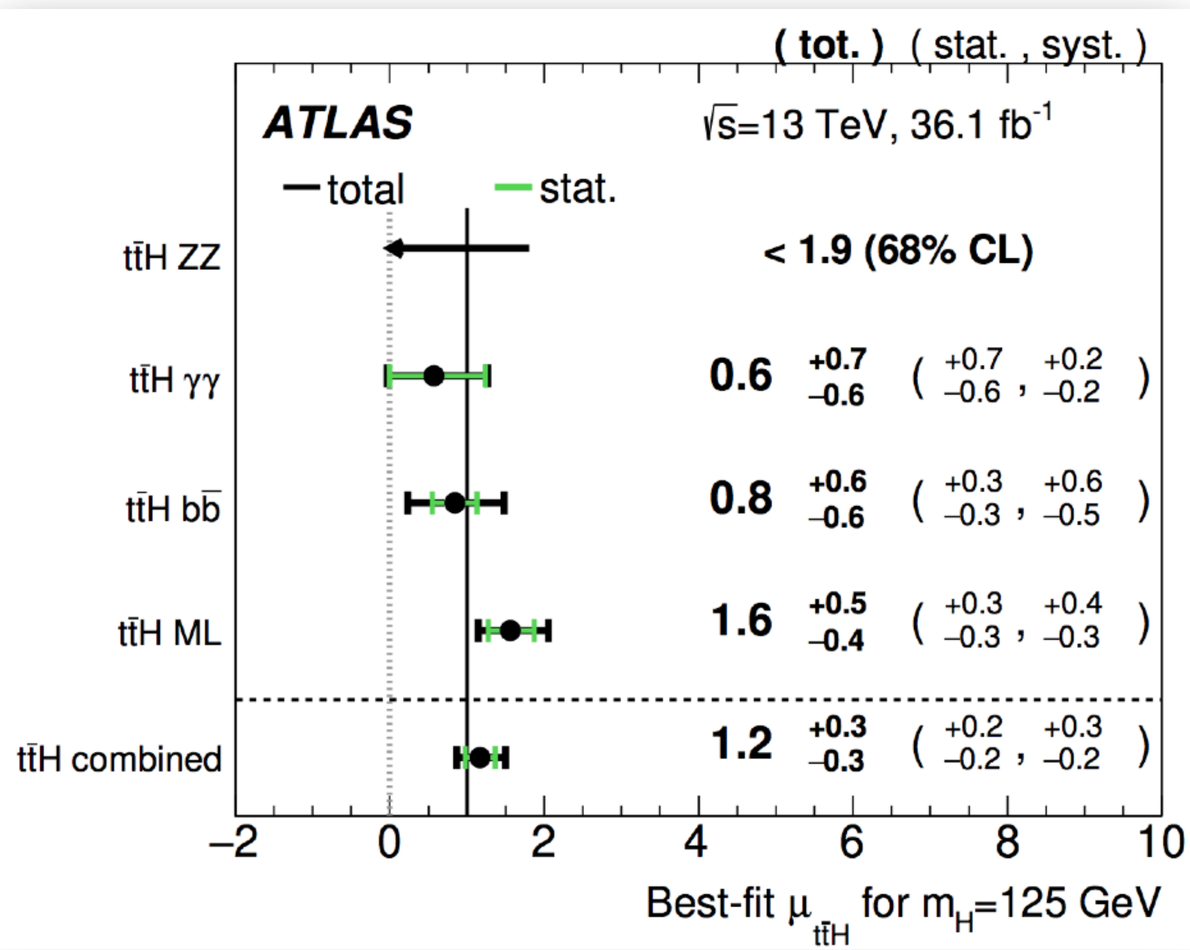
ttH combination

1806.00425

First evidence of ttH in ATLAS!

- Both systematic and statistical uncertainties limit the measurements;
- Significance of ttH production at **4.2σ** (exp 3.8σ)
- Best-fit: $\mu_{ttH}=1.17\pm0.19(stat)^{+0.27}_{-0.23}(syst)$

Channel	Best-fit μ Observed	Best-fit μ Expected	Significance Observed	Significance Expected
Multilepton	$1.6^{+0.5}_{-0.4}$	$1.0^{+0.4}_{-0.4}$	4.1σ	2.8σ
$H \to b\bar{b}$	$0.8^{+0.6}_{-0.6}$	$1.0^{+0.6}_{-0.6}$	1.4σ	1.6σ
$H \to \gamma\gamma$	$0.6^{+0.7}_{-0.6}$	$1.0^{+0.8}_{-0.6}$	0.9σ	1.7σ
$H \to 4\ell$	< 1.9	$1.0^{+3.2}_{-1.0}$	—	0.6σ
Combined	$1.2^{+0.3}_{-0.3}$	$1.0^{+0.3}_{-0.3}$	4.2σ	3.8σ

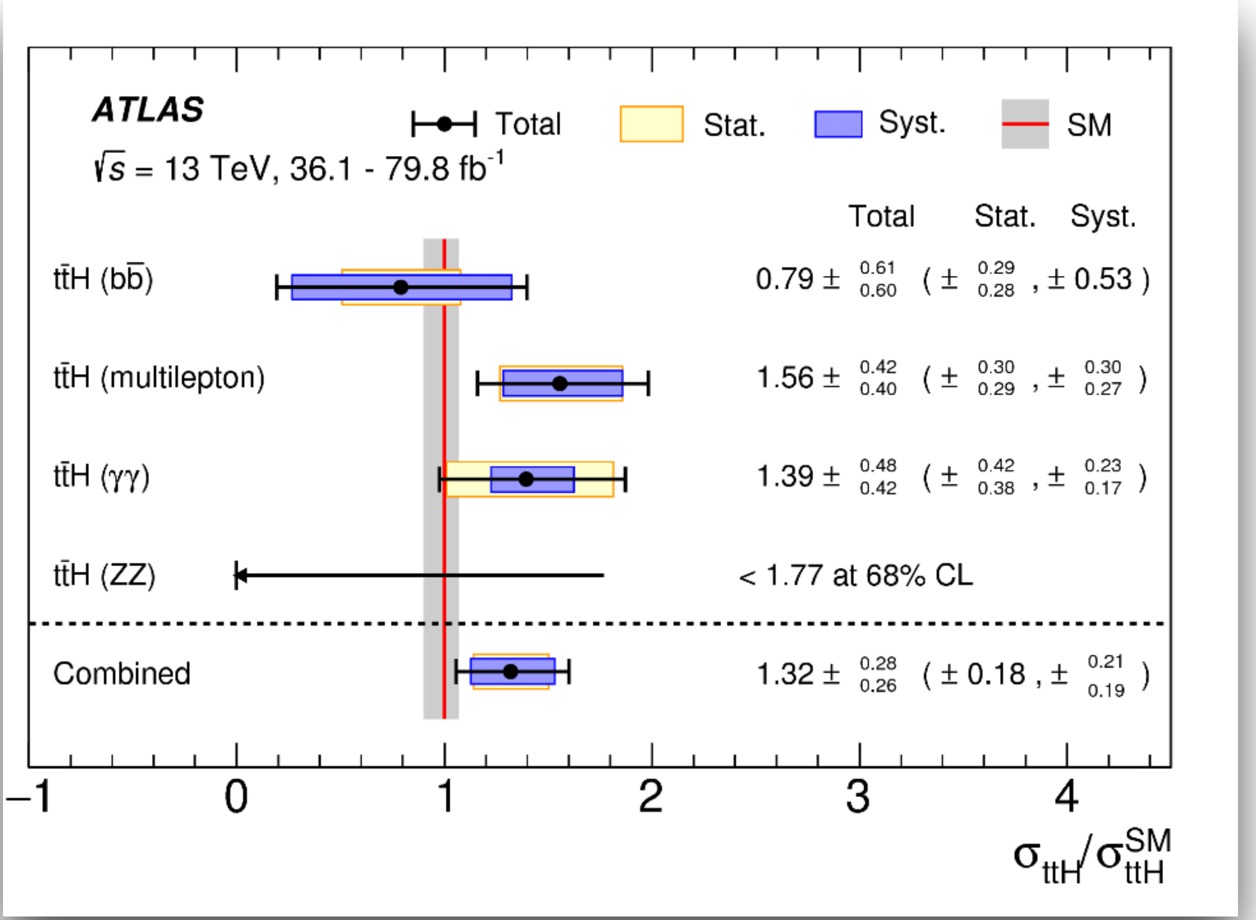


Uncertainty Source	$\Delta\mu$	
$t\bar{t}$ modeling in $H \to b\bar{b}$ analysis	+0.15	−0.14
$t\bar{t}H$ modeling (cross section)	+0.13	−0.06
Non-prompt light-lepton and fake τ_{had} estimates	+0.09	−0.09
Simulation statistics	+0.08	−0.08
Jet energy scale and resolution	+0.08	−0.07
$t\bar{t}V$ modeling	+0.07	−0.07
$t\bar{t}H$ modeling (acceptance)	+0.07	−0.04
Other non-Higgs boson backgrounds	+0.06	−0.05
Other experimental uncertainties	+0.05	−0.05
Luminosity	+0.05	−0.04
Jet flavor tagging	+0.03	−0.02
Modeling of other Higgs boson production modes	+0.01	−0.01
Total systematic uncertainty	+0.27	−0.23
Statistical uncertainty	+0.19	−0.19
Total uncertainty	+0.34	−0.30

First observation of ttH in ATLAS!

- Correlation scheme studied in detail
- Most sensitive channels limited by systematic uncertainties, mostly theoretical uncertainties. Other channels still statistically limited
- Significance of ttH production at **5.8σ** (exp 4.9σ)

Analysis	Integrated luminosity [fb^{-1}]	$t\bar{t}H$ cross section [fb]	Obs. sign.	Exp. sign.
$H \to \gamma\gamma$	79.8	$710\ ^{+210}_{-190}$ (stat.) $^{+120}_{-90}$ (syst.)	4.1 σ	3.7 σ
$H \to$ multilepton	36.1	790 ± 150 (stat.) $^{+150}_{-140}$ (syst.)	4.1 σ	2.8 σ
$H \to b\bar{b}$	36.1	$400\ ^{+150}_{-140}$ (stat.) $\pm$ 270 (syst.)	1.4 σ	1.6 σ
$H \to ZZ^* \to 4\ell$	79.8	<900 (68% CL)	0 σ	1.2 σ
Combined (13 TeV)	36.1–79.8	670 ± 90 (stat.) $^{+110}_{-100}$ (syst.)	5.8 σ	4.9 σ
Combined (7, 8, 13 TeV)	4.5, 20.3, 36.1–79.8	–	6.3 σ	5.1 σ

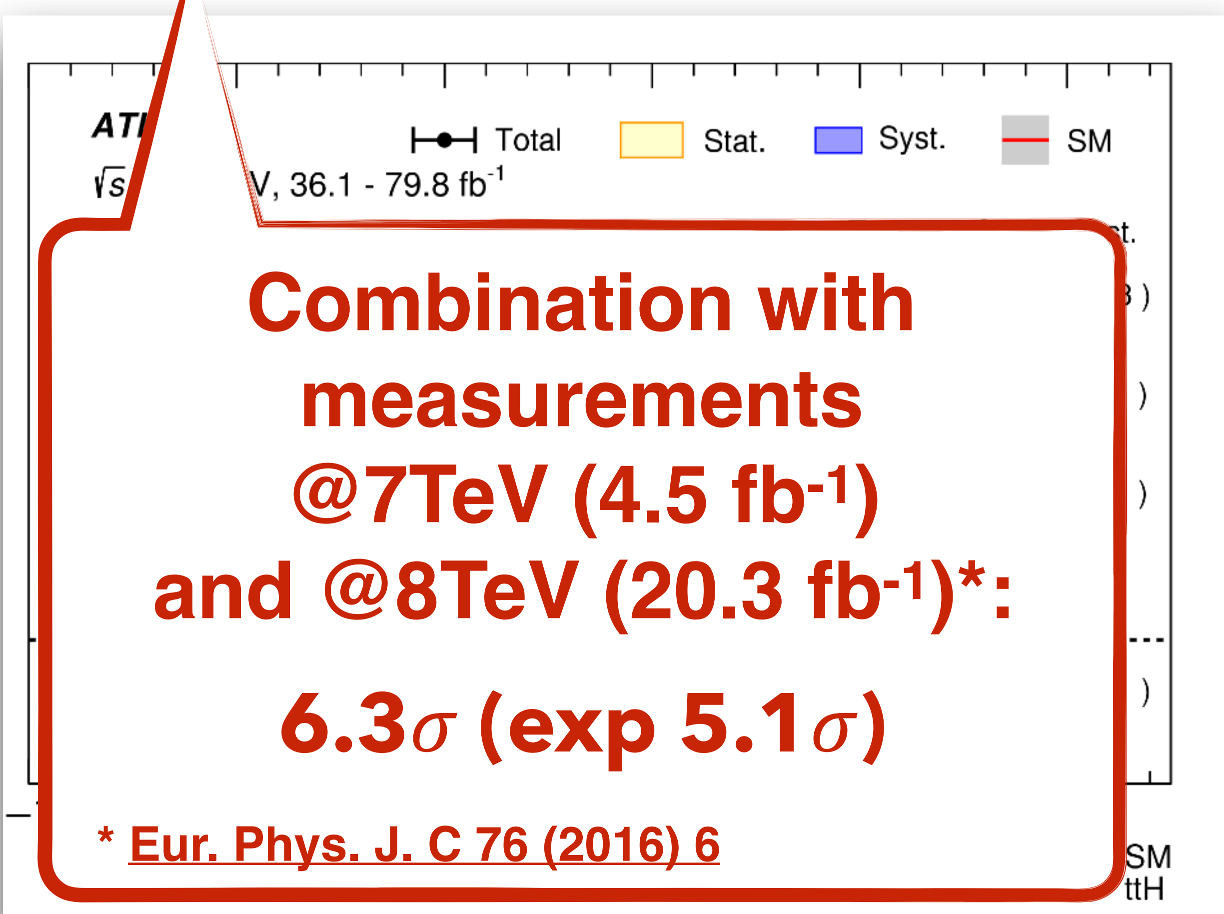


Uncertainty source	$\Delta\sigma_{t\bar{t}H}/\sigma_{t\bar{t}H}$ [%]
Theory uncertainties (modelling)	11.9
$t\bar{t}$ + heavy flavour	9.9
$t\bar{t}H$	6.0
Non-$t\bar{t}H$ Higgs boson production modes	1.5
Other background processes	2.2
Experimental uncertainties	9.3
Fake leptons	5.2
Jets, E_T^{miss}	4.9
Electrons, photons	3.2
Luminosity	3.0
τ-lepton	2.5
Flavour tagging	1.8
MC statistical uncertainties	4.4

First observation of ttH in ATLAS!

- Correlation scheme studied in detail
- Most sensitive channels limited by systematic uncertainties, mostly theoretical uncertainties. Other channels still statistically limited
- Significance of ttH production at **5.8σ** (exp 4.9σ)

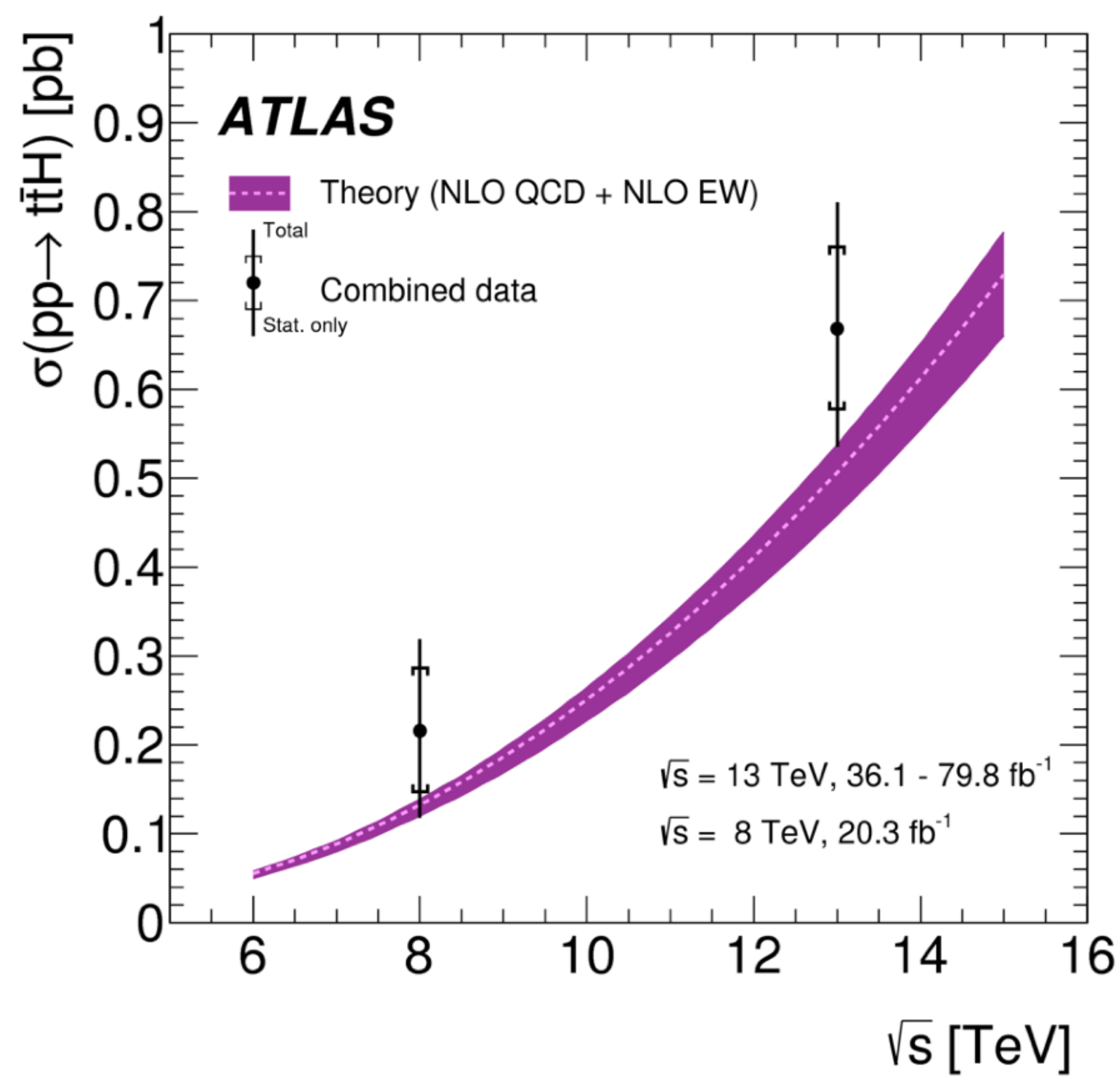
Analysis	Integrated luminosity [fb^{-1}]	$t\bar{t}H$ cross section [fb]	Obs. sign.	Exp. sign.
$H \to \gamma\gamma$	79.8	$710\ ^{+210}_{-190}$ (stat.) $^{+120}_{-90}$ (syst.)	4.1 σ	3.7 σ
$H \to$ multilepton	36.1	790 ± 150 (stat.) $^{+150}_{-140}$ (syst.)	4.1 σ	2.8 σ
$H \to b\bar{b}$	36.1	$400\ ^{+150}_{-140}$ (stat.) $\pm$ 270 (syst.)	1.4 σ	1.6 σ
$H \to ZZ^* \to 4\ell$	79.8	<900 (68% CL)	0 σ	1.2 σ
Combined (13 TeV)	36.1–79.8	670 ± 90 (stat.) $^{+110}_{-100}$ (syst.)	5.8 σ	4.9 σ
Combined (7, 8, 13 TeV)	4.5, 20.3, 36.1–79.8	–	6.3 σ	5.1 σ



Combination with measurements @7TeV (4.5 fb^{-1}) and @8TeV (20.3 fb^{-1})*:

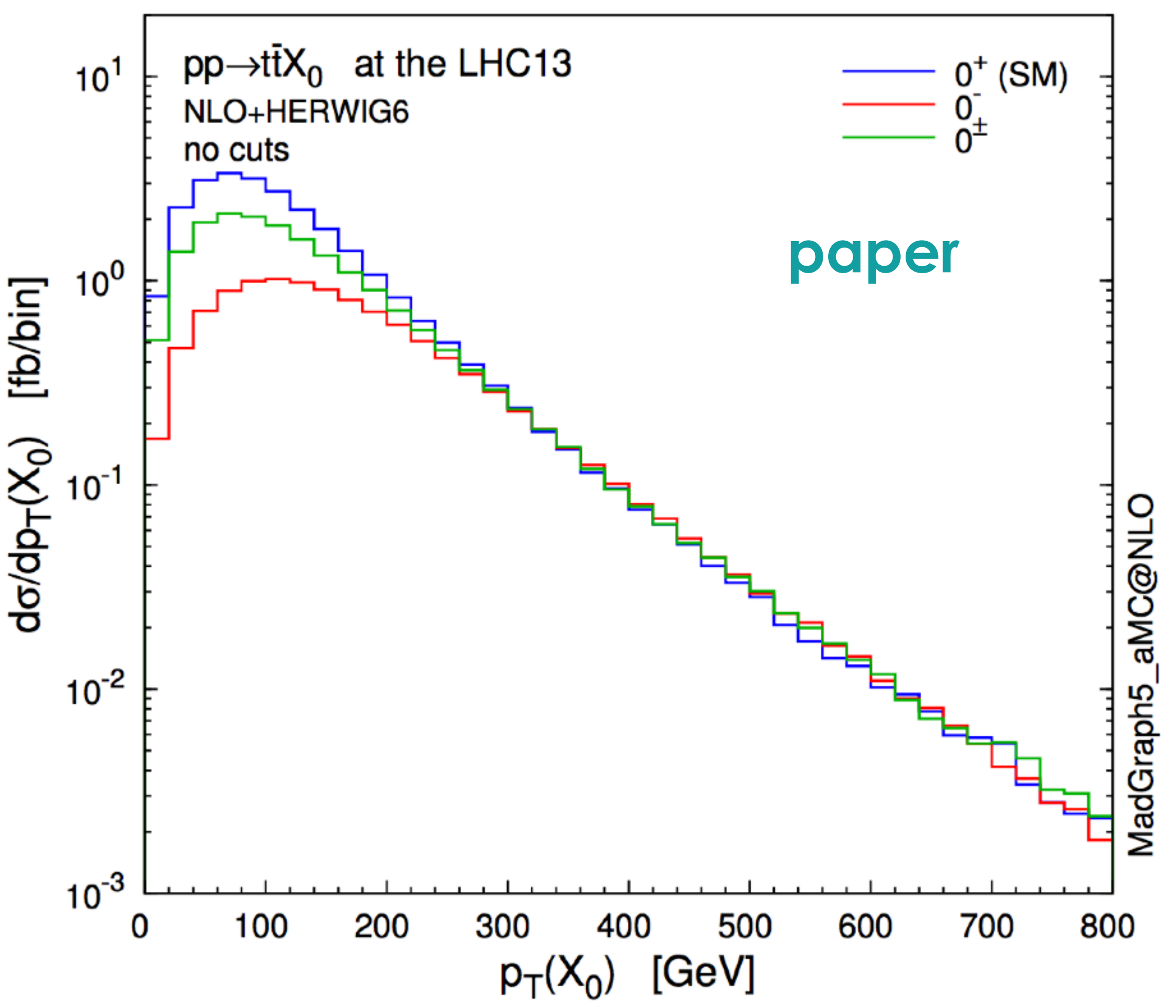
6.3σ (exp 5.1σ)

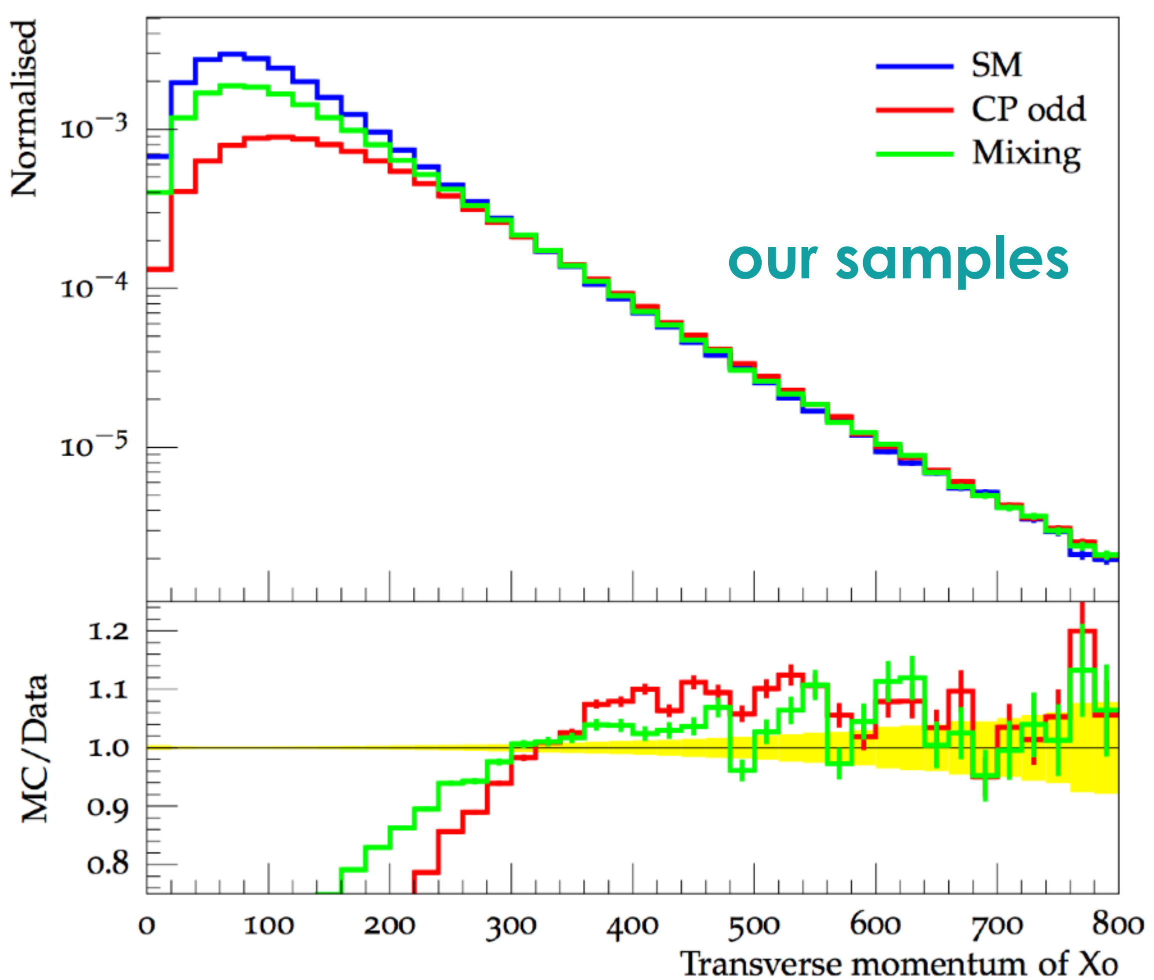
* Eur. Phys. J. C 76 (2016) 6



Now we go ahead!

- Testing the properties of ttH process: starting with CP structure
- Both resolved and boosted regimes are involved in the effort
- Production of samples with different models





to be continued

…this is not the end!

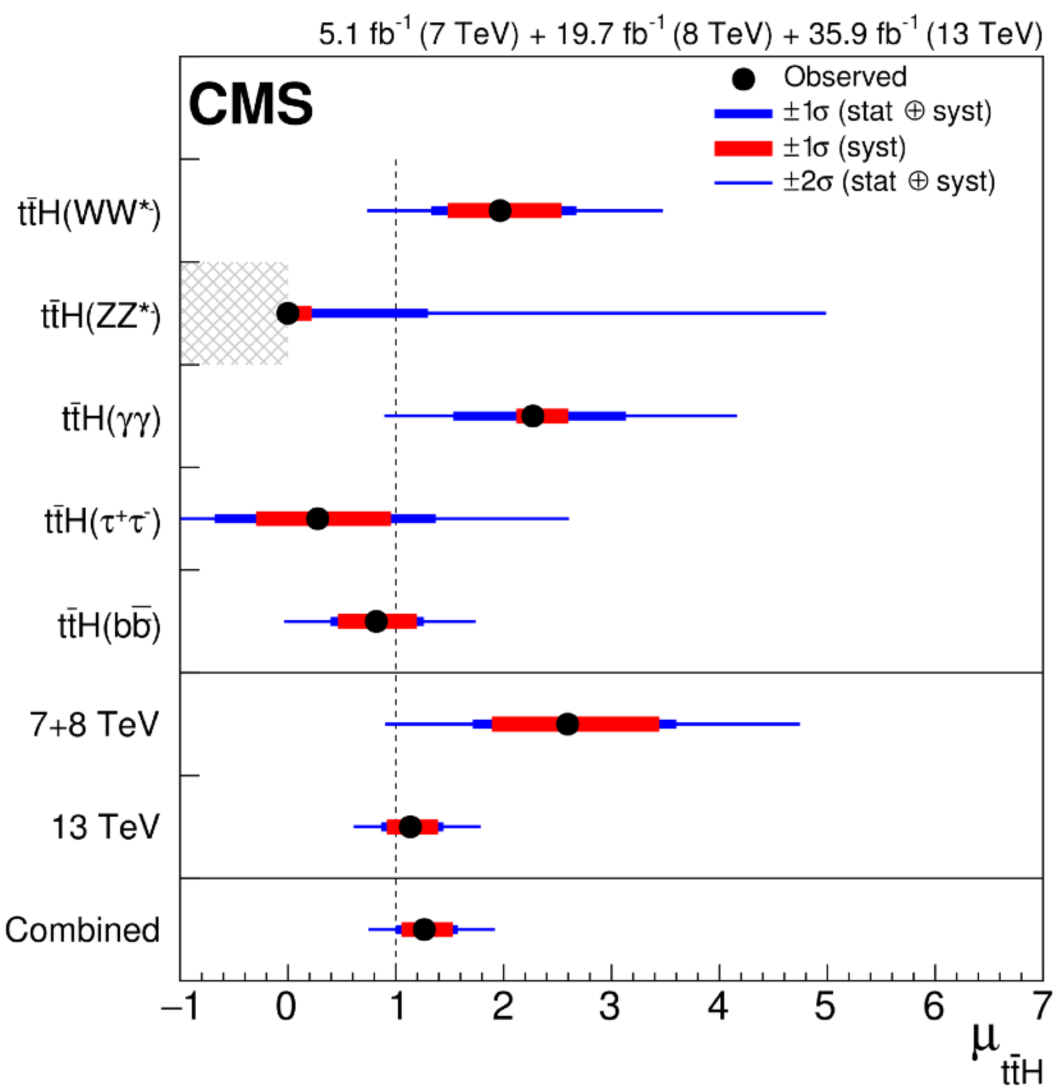
GRAZIE PER L'ATTENZIONE!

Supporting material

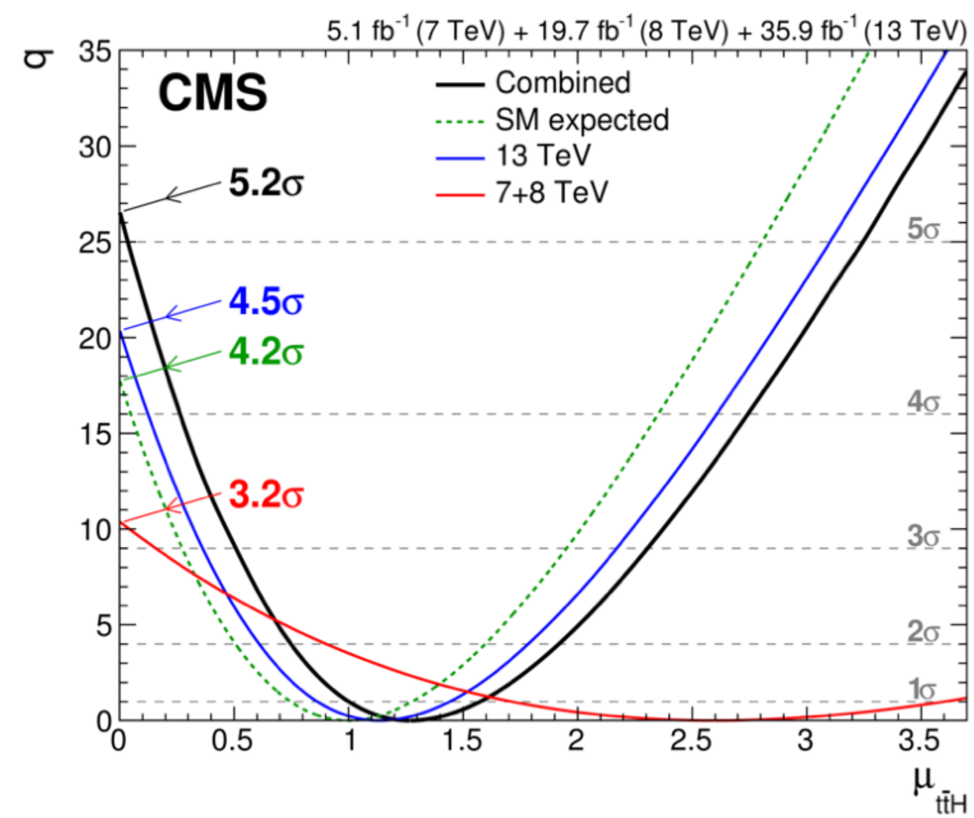
CMS observation

arxiv:1804.02610

- ttH and tt+bb theoretical uncertainties are dominant, as for the ATLAS analysis;
- significance of ttH production at **5.2**σ (exp 4.2σ)
- Best-fit: $\mu_{ttH}=1.26\pm0.19(\text{stat})^{+0.31}_{-0.26}(\text{syst})$



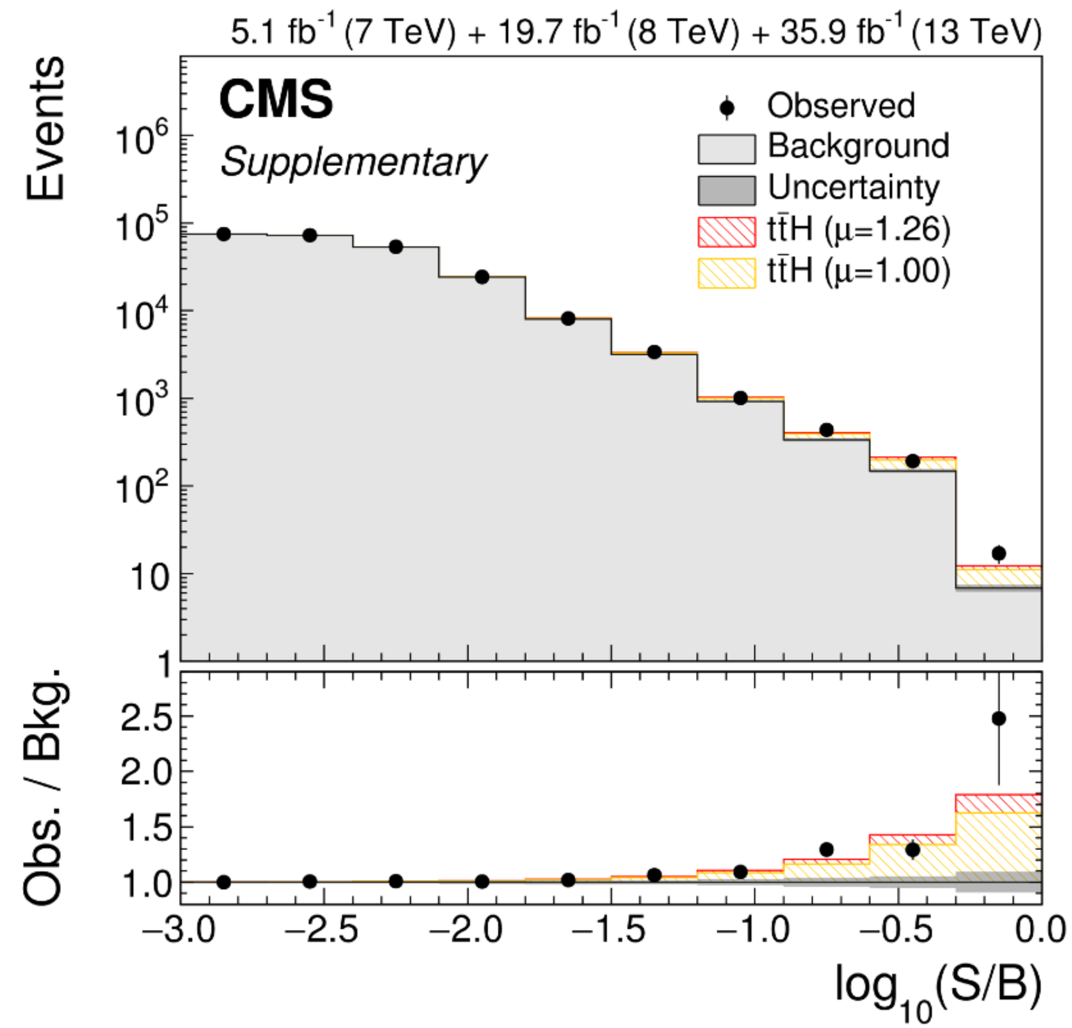
Uncertainty source	$\Delta\mu$	
Signal theory	+0.15	−0.07
Inclusive ttH normalisation (cross section and BR)	+0.15	−0.07
ttH acceptance (scale, pdf, PS and UE)	+0.004	−0.004
Other Higgs boson production modes	+0.002	−0.003
Background theory	+0.14	−0.13
tt + bb/cc prediction	+0.13	−0.11
tt + V(V) prediction	+0.06	−0.06
Other background uncertainties	+0.03	−0.03
Experimental	+0.17	−0.15
Lepton (inc. τ_h) trigger, ID and iso. efficiency	+0.08	−0.06
Misidentified lepton prediction	+0.06	−0.06
b-Tagging efficiency	+0.05	−0.04
Jet and τ_h energy scale and resolution	+0.04	−0.04
Luminosity	+0.04	−0.03
Photon ID, scale and resolution	+0.01	−0.01
Other experimental uncertainties	+0.01	−0.01
Finite number of simulated events	+0.08	−0.07
Statistical	+0.16	−0.16
Total	+0.31	−0.26



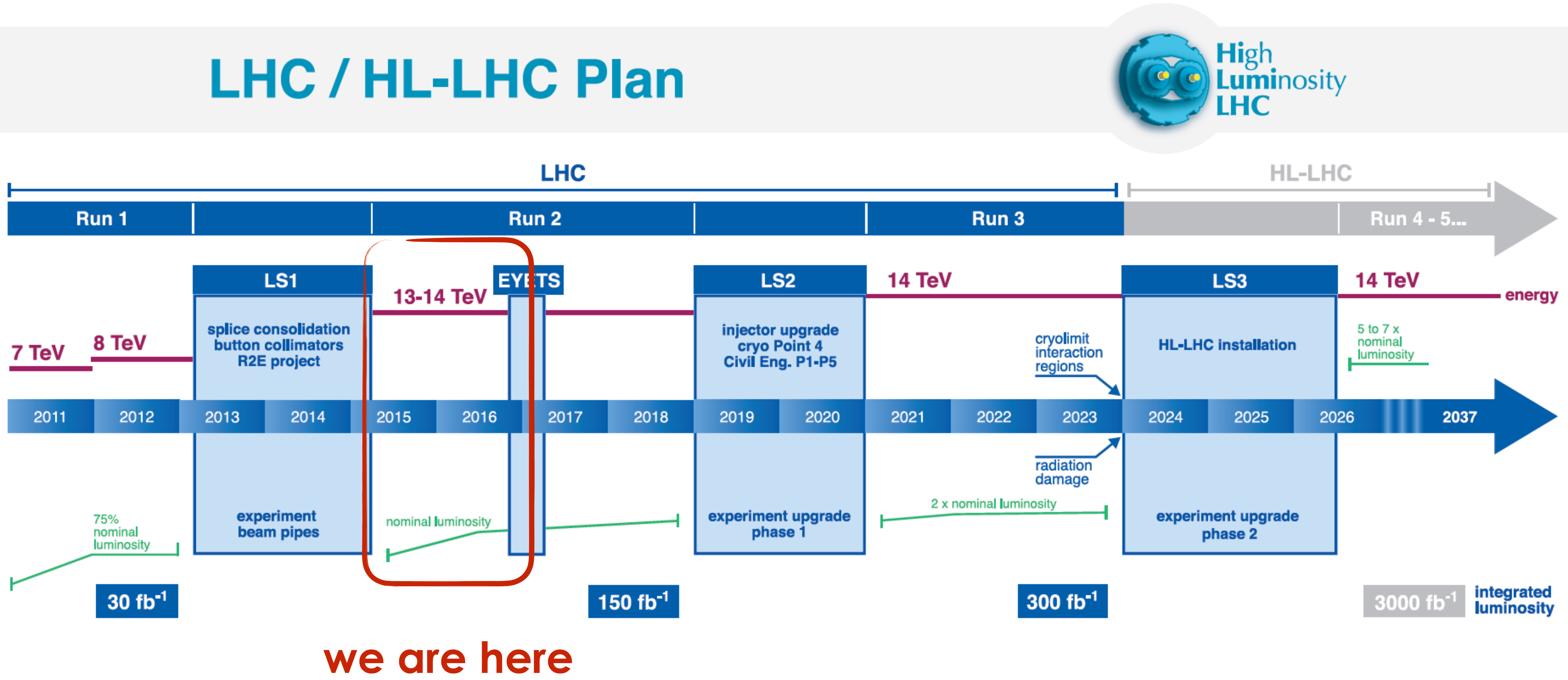
CMS observation

arxiv:1804.02610

- ttH and tt+bb theoretical uncertainties are dominant, as for the ATLAS analysis;
- significance of ttH production at **5.2**σ (exp 4.2σ)
- Best-fit: μ_{ttH}=1.26±0.19(stat)$^{+0.31}_{-0.26}$(syst)



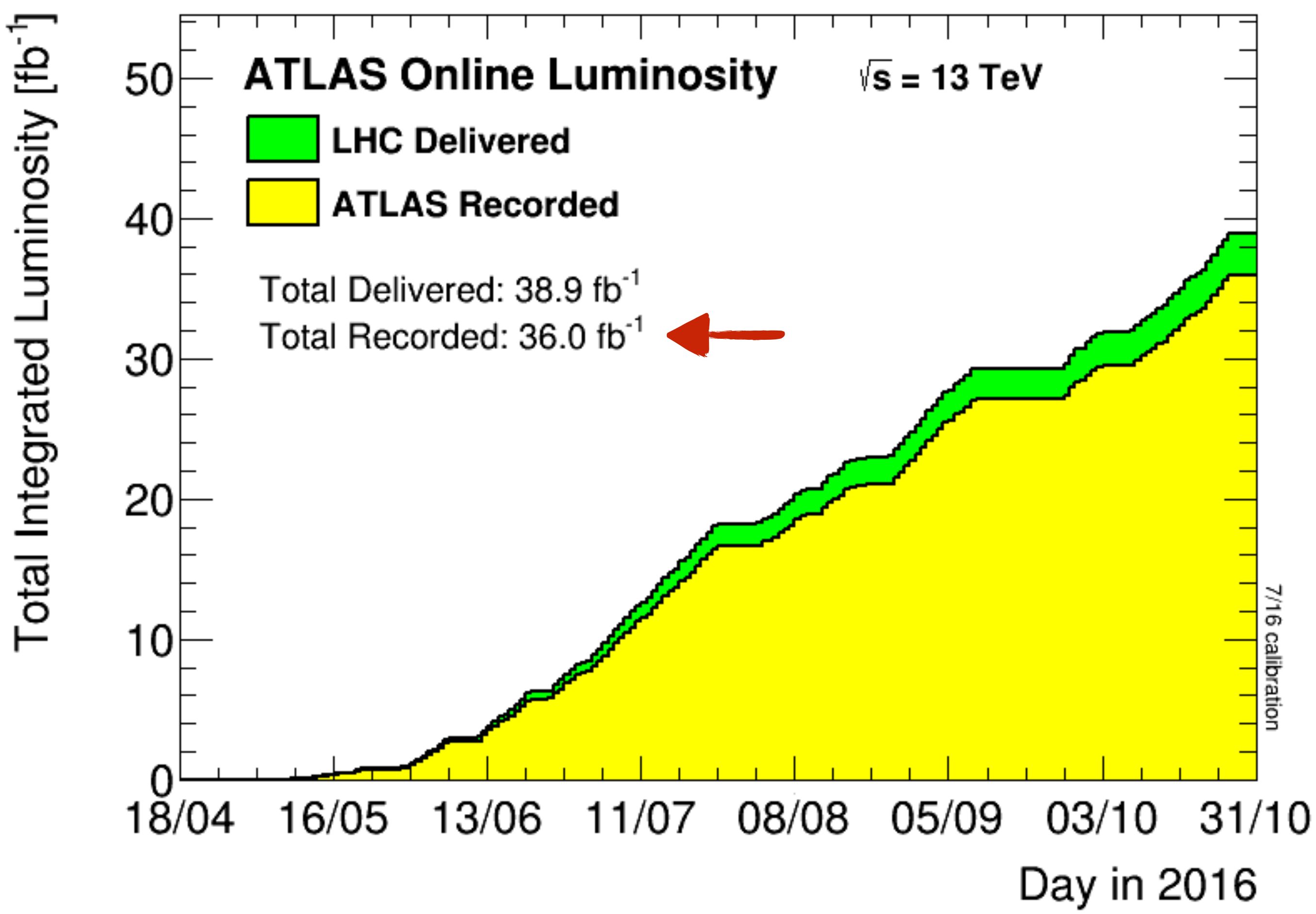
Time schedule of LHC

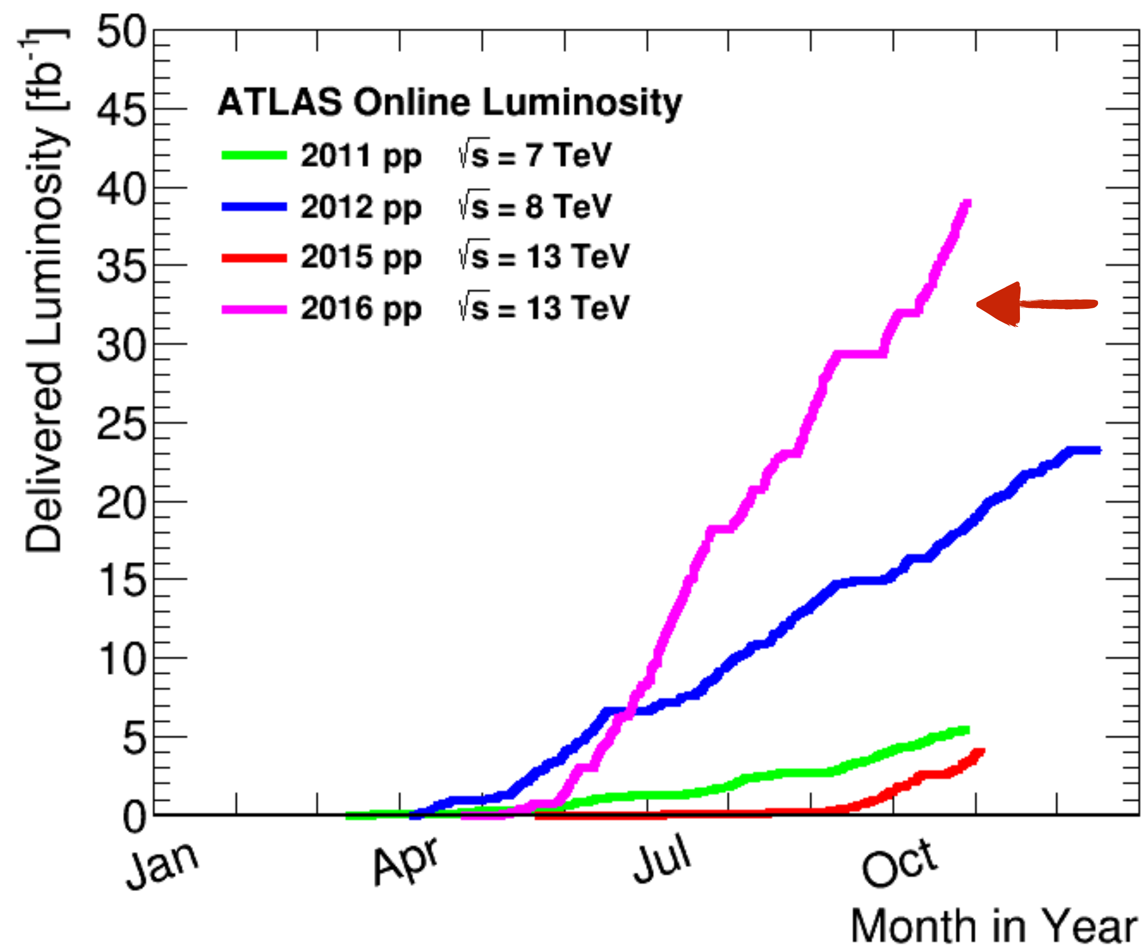


2015 and 2016 data!

- unexplored energies and unprecedented rates!

36.5 fb^{-1}



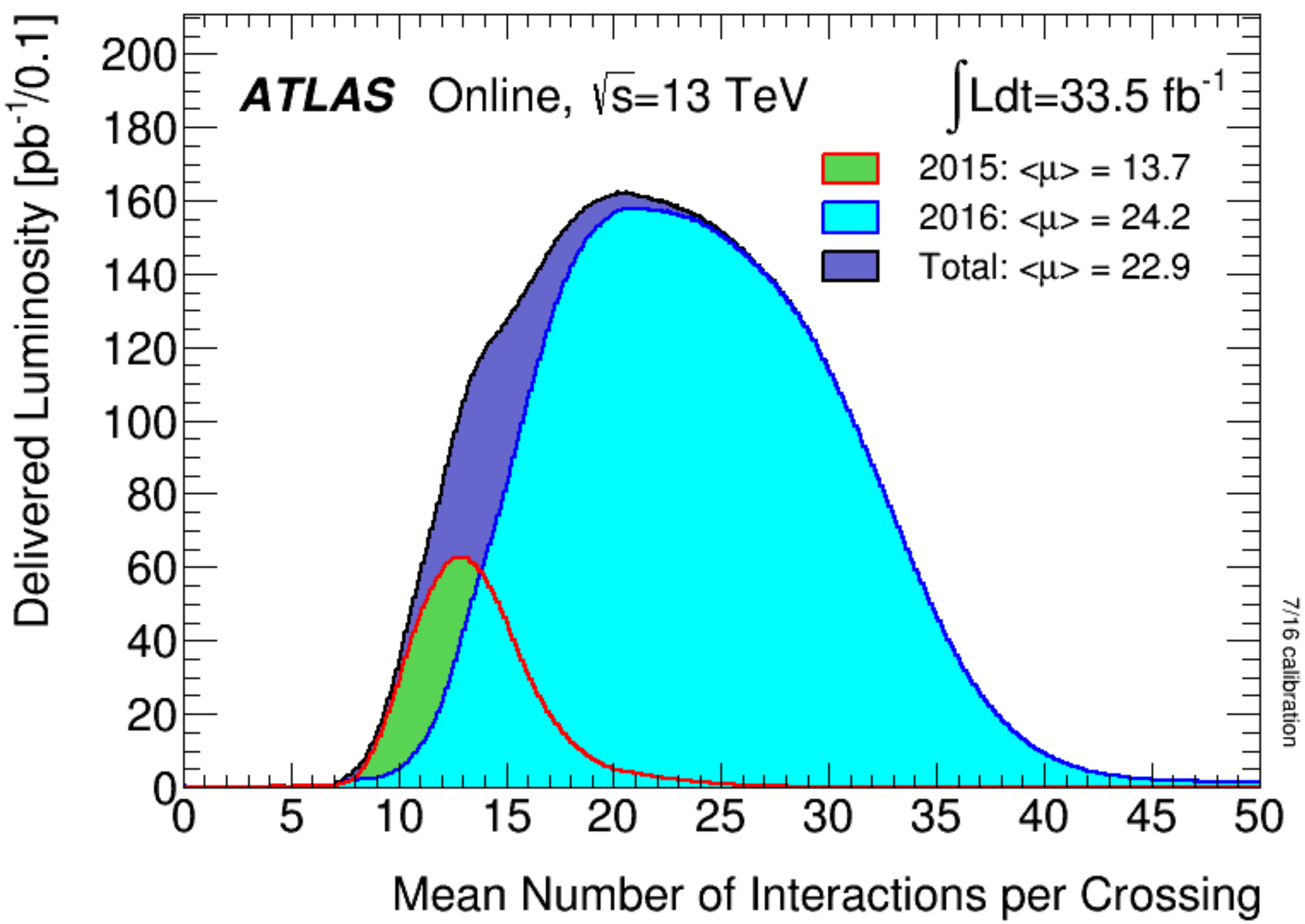


- **delivered luminosity** from the start of stable beams until ATLAS goes to standby mode for the beam dump;
- **recorded luminosity** reflects the data acquisition inefficiency.
- **during 2016**, ATLAS reached the highest luminosity ever in only 6 months!

reference

2015 and 2016 data!

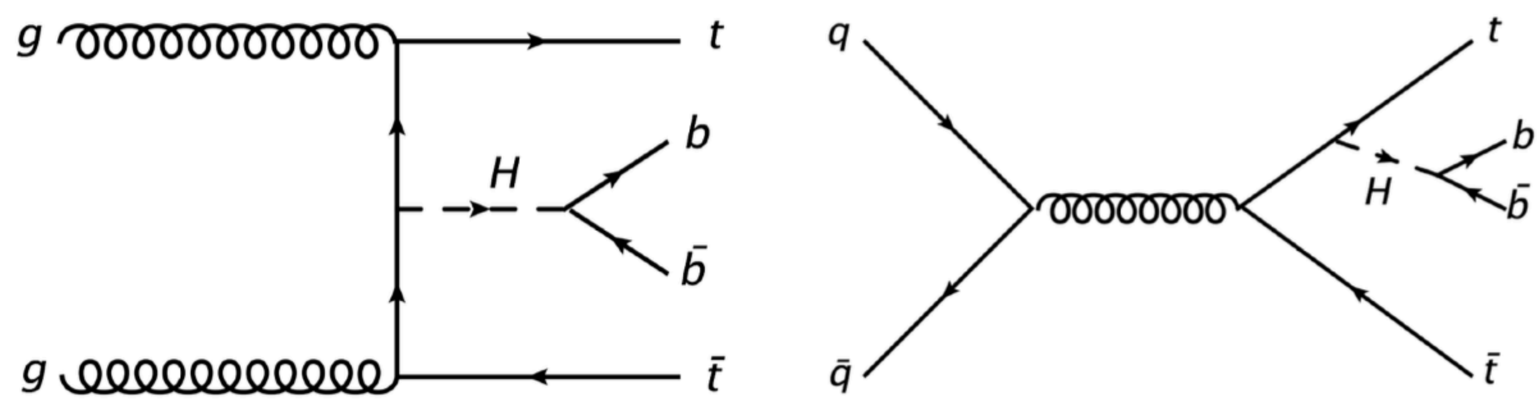
- unexplored energies and unprecedented rates!



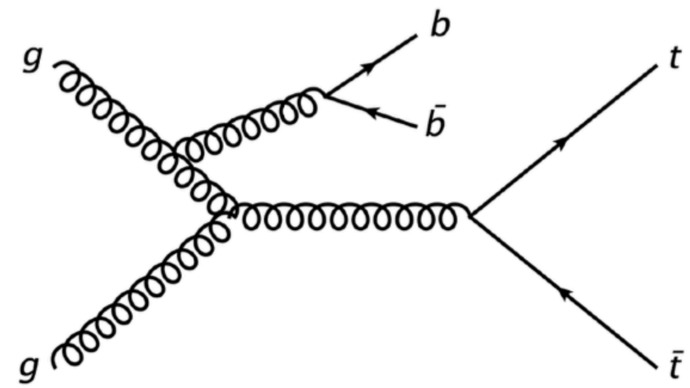
- luminosity-weighted distribution of the **mean number of interactions per crossing** for the 2015+2016 pp collision data recorded

- **$\mu = L_{bunch} \times \sigma_{inel} / f_r$**
  - L_{bunch} = per bunch instantaneous luminosity;
  - σ_{inel} = inelastic cross section (80mb);
  - f_r = LHC revolution frequency.

The associated Higgs production

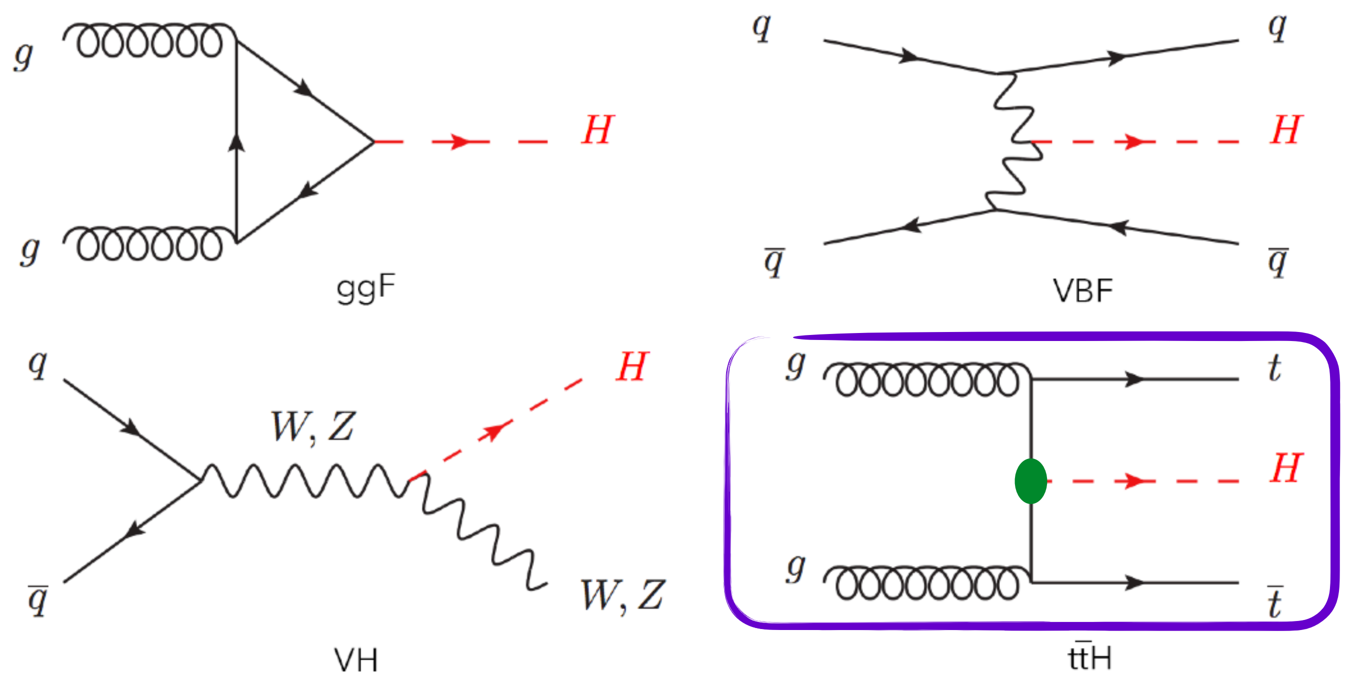


Representative tree-level Feynman diagrams for the production of the Higgs boson in association with a top-quark pair (ttH) and the subsequent decay of the Higgs to bb

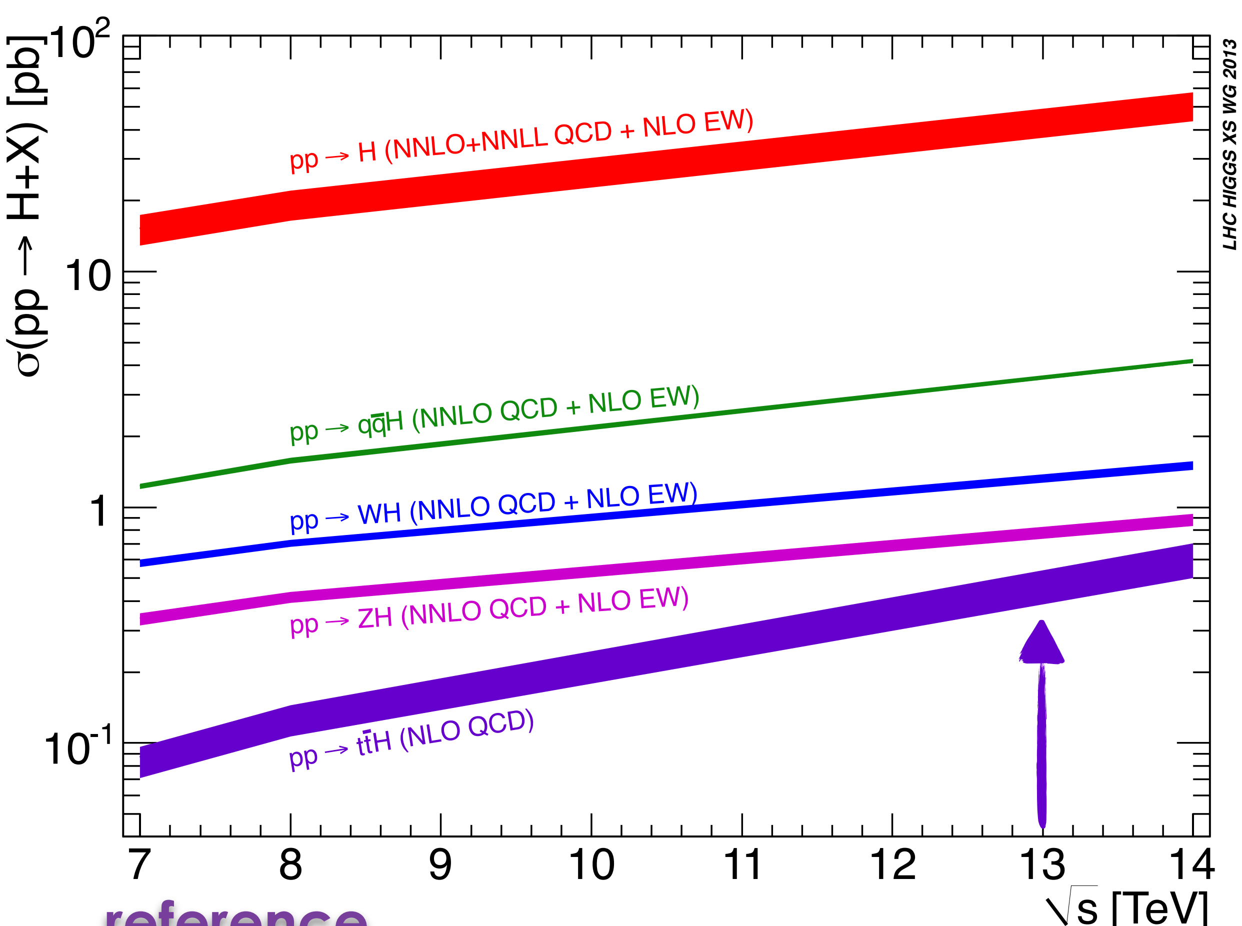


Representative tree-level Feynman diagram for the main background tt+bb

The associated Higgs production



expected cross section (pb)			
channel	7 TeV	8 TeV	13 TeV
ggH	15	19	43
VBF	1.2	1.6	3.7
WH	0.5	0.7	1.4
ZH	0.3	0.4	0.9
ttH	0.09	0.13	0.50

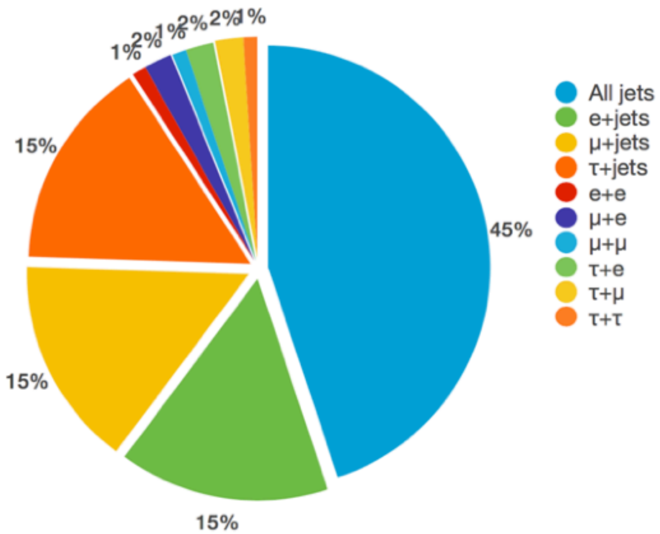


reference

Some motivations

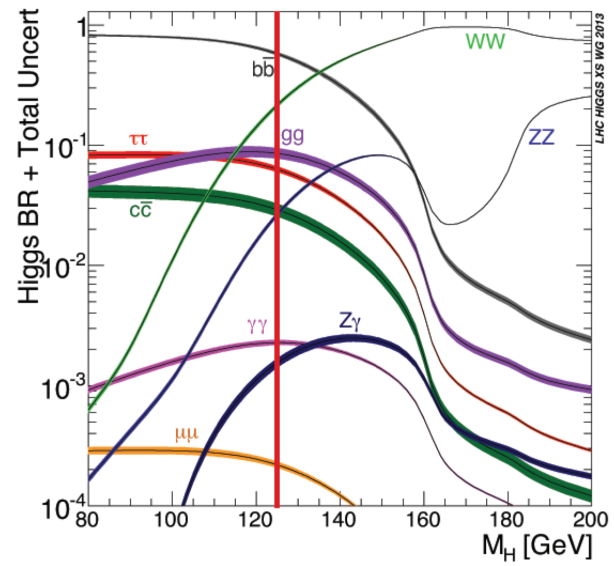
- Higgs produced in association with a top quark pair (ttH) allows **direct access to Yukawa coupling of Higgs boson to top quark**;
- ttH shows the **highest cross section increase as a function of energy** wrt other production modes;
- any deviation in the cross-section measurements could be an **hint of new physics**!

The ttH channel



BR (tt->lνqq) ~35%

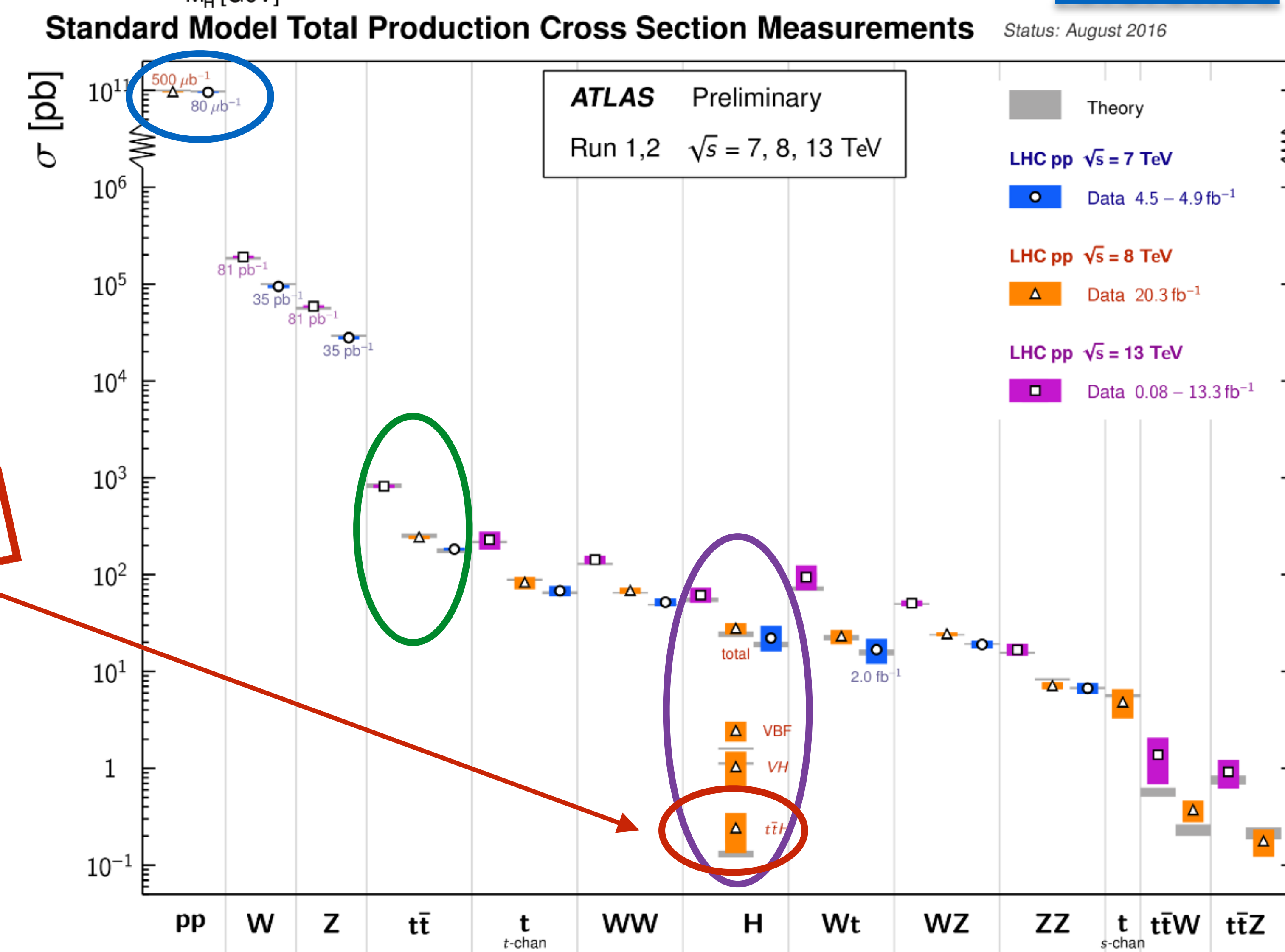
+



BR (H->bb) ~60%

reference

reference



σ(pp->**tt**) ~ 8x10³ pb

80 events each billion

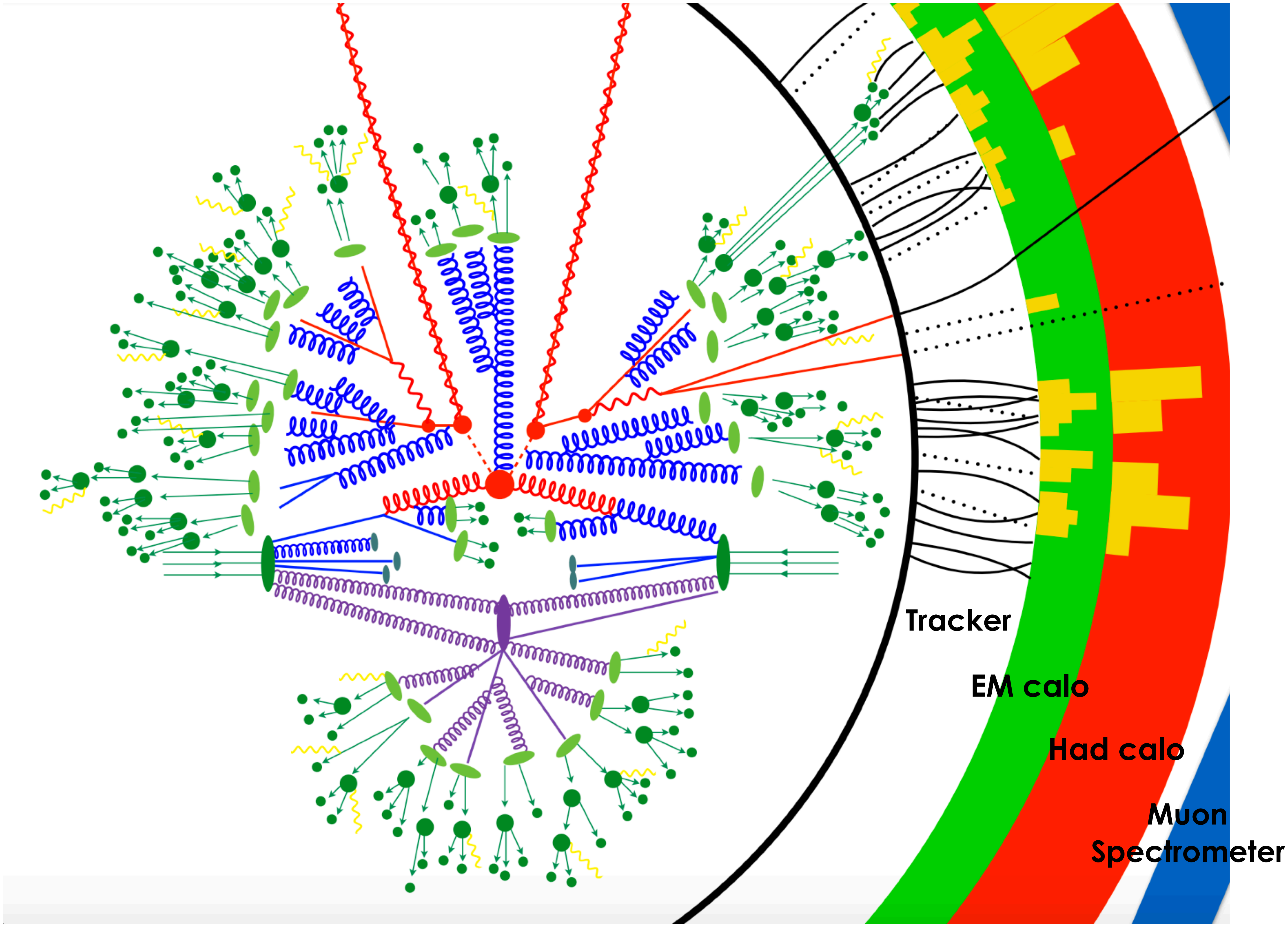
NB! the main background contribution to our channel

σ(pp->**HX**) ~ 45 pb

1 event each 2 billions

Rare event!

Reconstruction



Anti-k_t and k_t algorithms

- The iterative recombination procedure works by first cleaning a list of all objects (either hadrons, topo-clusters or tracks) in an event.
- The ordering of the list is irrelevant and proto-jets are built from these objects.
- Two distance measures in y-ϕ-space are associated to each member of the list, between the proto-jet and its closest neighbor:

$$\rho_{ij} = \min\left(p_{Ti}^{2p}, p_{Tj}^{2p}\right)\frac{(\Delta R_{ij})^2}{R^2}$$

measure of the opening angle between the two constituents

$$\Delta R_{ij} = \sqrt{(y_i - y_j)^2 + (\phi_i - \phi_j)^2}$$

and between the proto-jet and the beam:

$$\rho_{iB} = p_{Ti}^{2p}$$

- ➡ If $\varrho_{iB} < \varrho_{ij}$: the proto-jet is closer to the beam than it is to any other proto-jet in the event, so it is defined as a jet and removed from the list.
- ➡ If $\varrho_{iB} > \varrho_{ij}$: the two proto-jets i and j are combined into one, thereby forming a new proto-jet.

This procedure continues through all proto-jets in the event.

- ✦ If p = +1 ⟶ k_t algorithm: proto-jets with the smallest p_T tend to be clustered first, so that the highest p_T proto-jets are clustered last.
- ✦ If p = -1 ⟶ anti-k_t algorithm: proto-jets with the largest p_T are clustered first. A consequence of this is that isolated anti-k_t jets tend to be very close to circular in η-ϕ space, because the axis of the jet is relatively fixed after the first few steps of recombination. This stability makes anti-k_t jets more robust than k_t jets in high multiplicity environments.

Tagging techniques

- Used to exploit all the substructure characteristics of the large-R jets in boosted regimes;
- **Top-tagging:** simple algorithm which provides cuts on two large-R jet substructure variables:

Calibrated mass of the jet

$$(m_{\text{jet}}^{\text{uncalib}})^2 = \left(\sum_i E_i\right)^2 - \left(\sum_i \vec{p}_i\right)^2$$

calibration needed because of low p_T and large angle contributions

Ratio τ_{32} between N-subjettiness variables

$$\tau_N = \frac{1}{d_0} \sum_k p_{Tk} \times \min(\delta R_{1k}, \delta R_{2k}, \ldots, \delta R_{Nk})$$

number of reconstructed subjets

$$d_0 = \sum_k p_{Tk} \times R$$

distance between subjet i and constituent k

$$\tau_{32} = \tau_3 / \tau_2$$

- **Higgs-tagging:** very similar to top-tagging but for the second substructure variable:

Energy correlation D_2

$$D_2^{(\alpha,\beta)} = \frac{e_3^{(\alpha)}}{(e_2^{(\beta)})^{3\alpha/\beta}}$$

e_2, e_3 = energy correlation functions: take into account all p_T, combinations and boost-invariant angle ΔR between constituents

α, β =1

- **Taggers performances** are given by two values, calculated in the same way:

Signal efficiency

$$\epsilon = \left(\frac{N_{tagged}}{N_{total}}\right)_{signal}$$

Background rejection

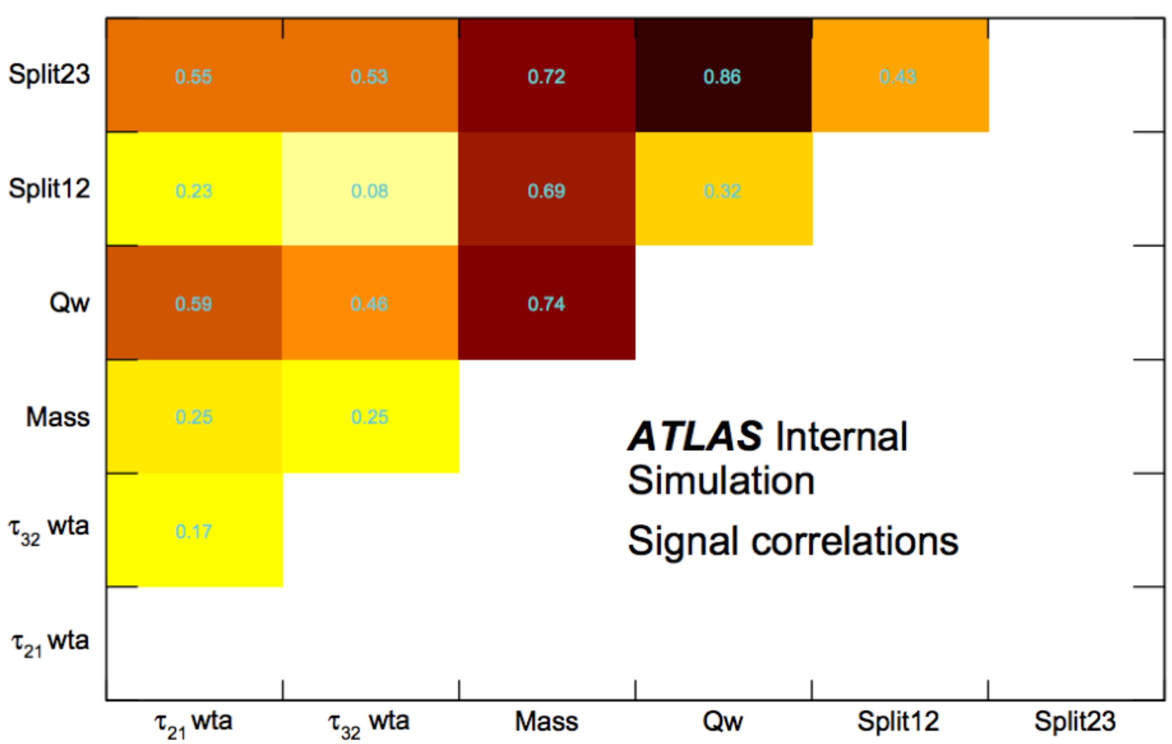
$$r = \left(\frac{N_{total}}{N_{tagged}}\right)_{background}$$

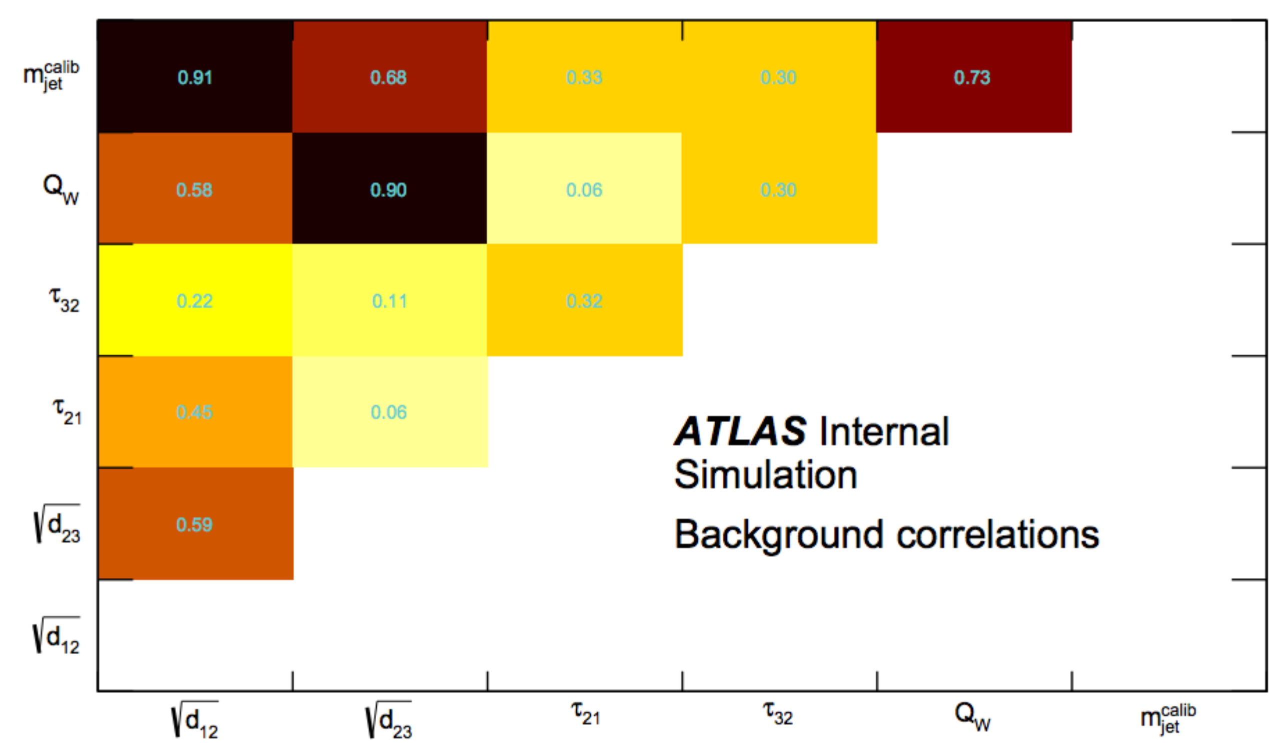
Top Tagging

The two variables $\mathbf{m_{jet}^{calib}}$ and $\boldsymbol{\tau_{32}}$ were chosen from a set of substructure variables, including other N-subjettiness ratio (τ_{21}), splitting scale variables ($\sqrt{d_{12}}$,$\sqrt{d_{23}}$) and the minimum dijet mass from the three subjets (Qw).

- The two chosen variables show a **good background rejection** at 50% and 80% signal efficiency.
- The τ_{ij} variables are **uncorrelated** with respect to the mass and energy scale variables.

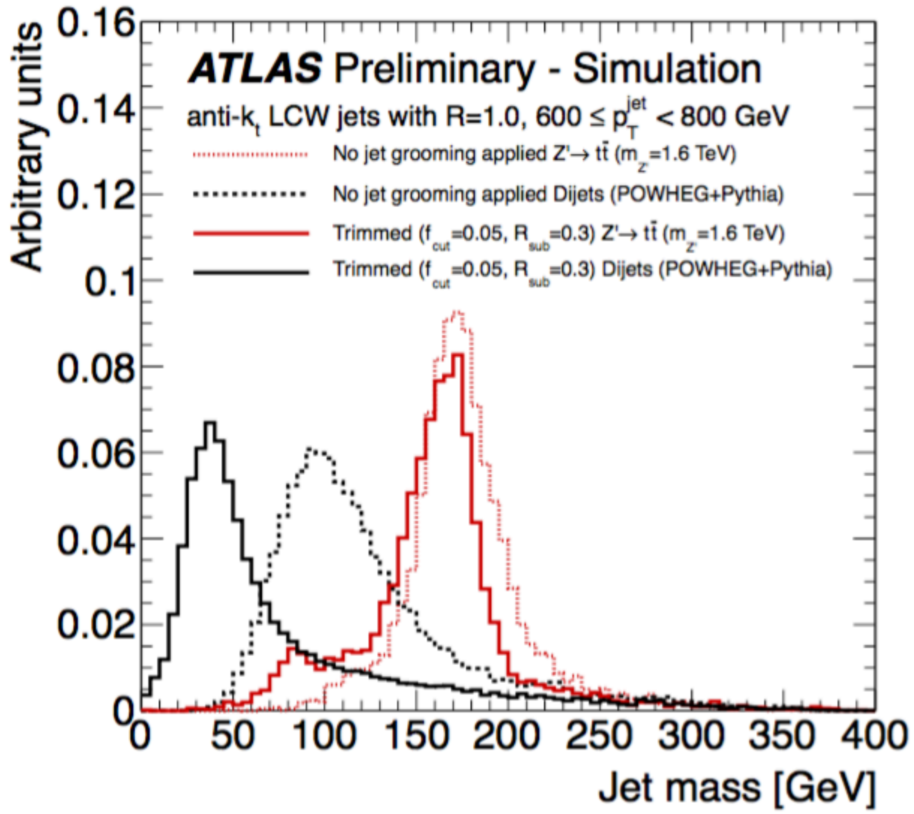
This combination of strong performance and lack of correlated behavior motivates the choice of tagging variables.

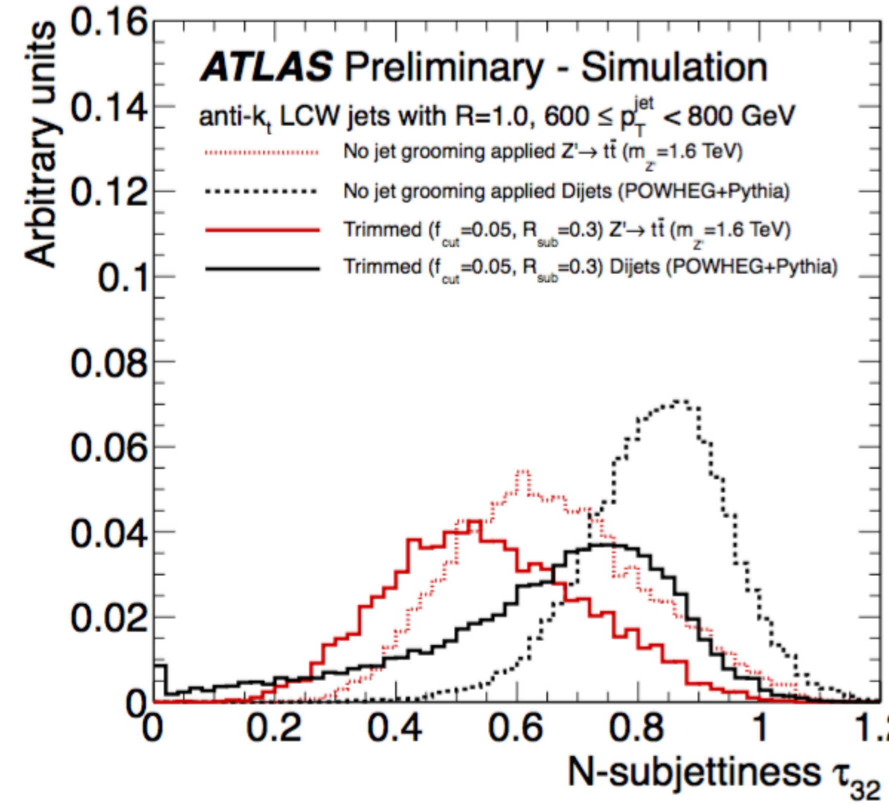




Top Tagging technique

- Used to **exploit all the substructure characteristics of the large-R jets** in boosted regimes;
- substructure variables in the algorithm: **large-R jet mass** and **τ_{32} ratio**.





Performances

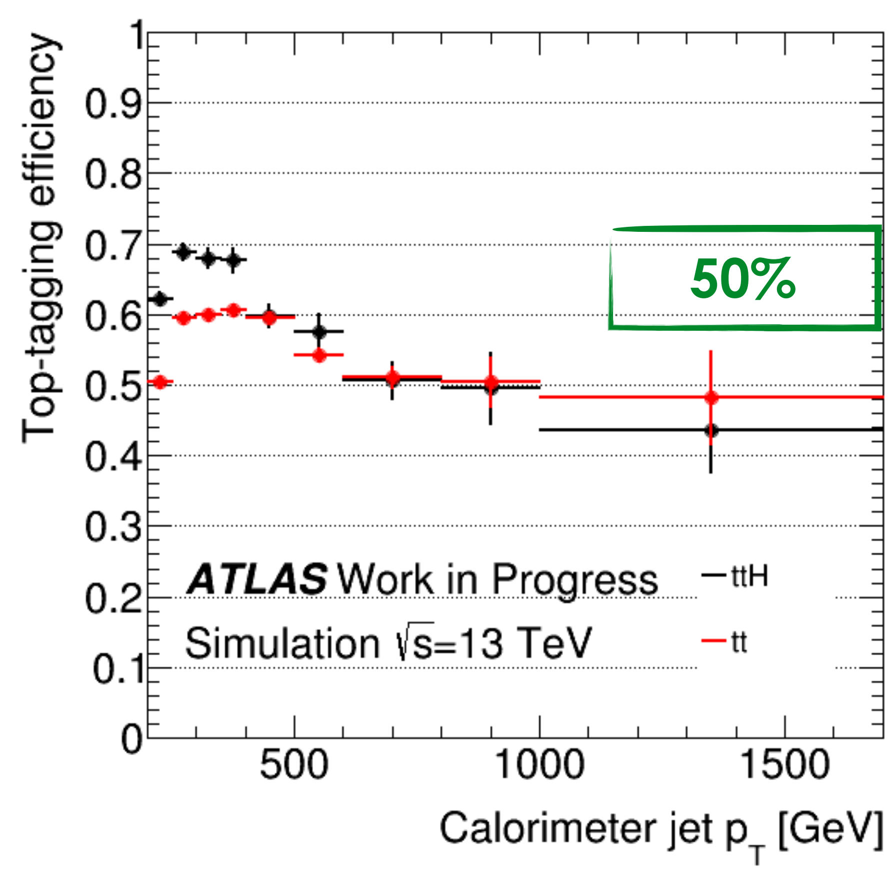
- Comparison between **efficiencies in tt and ttH**;
- looking at:

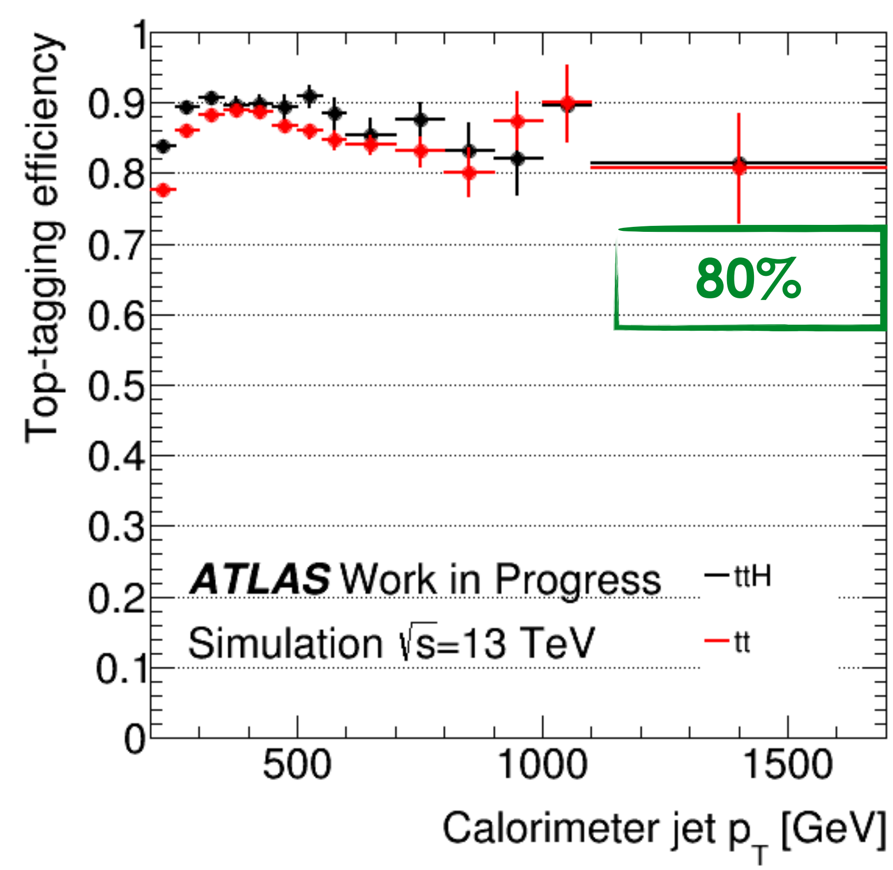
Signal efficiency

$$\epsilon = \left(\frac{N_{tagged}}{N_{total}}\right)_{signal}$$

Background rejection

$$r = \left(\frac{N_{total}}{N_{tagged}}\right)_{background}$$

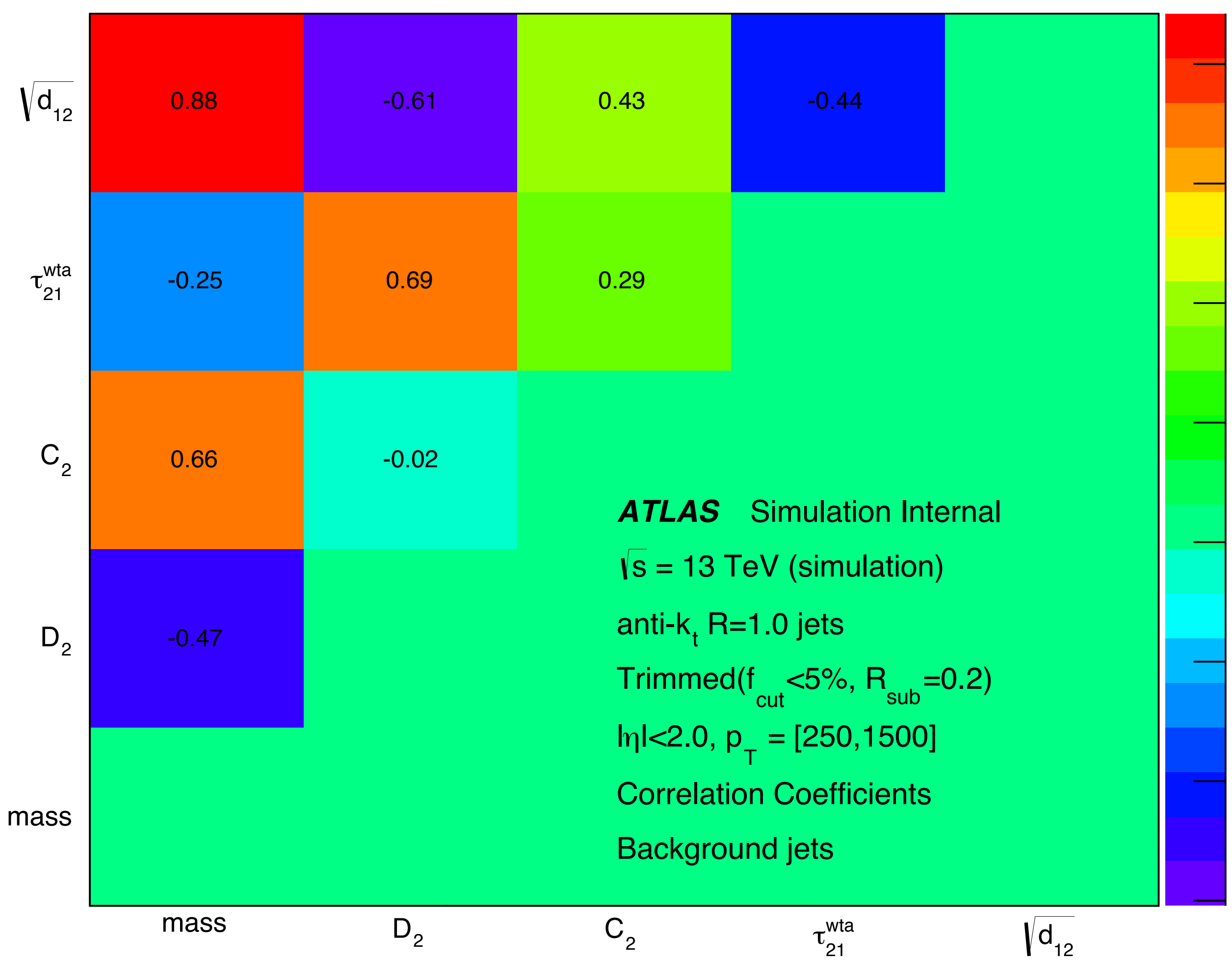




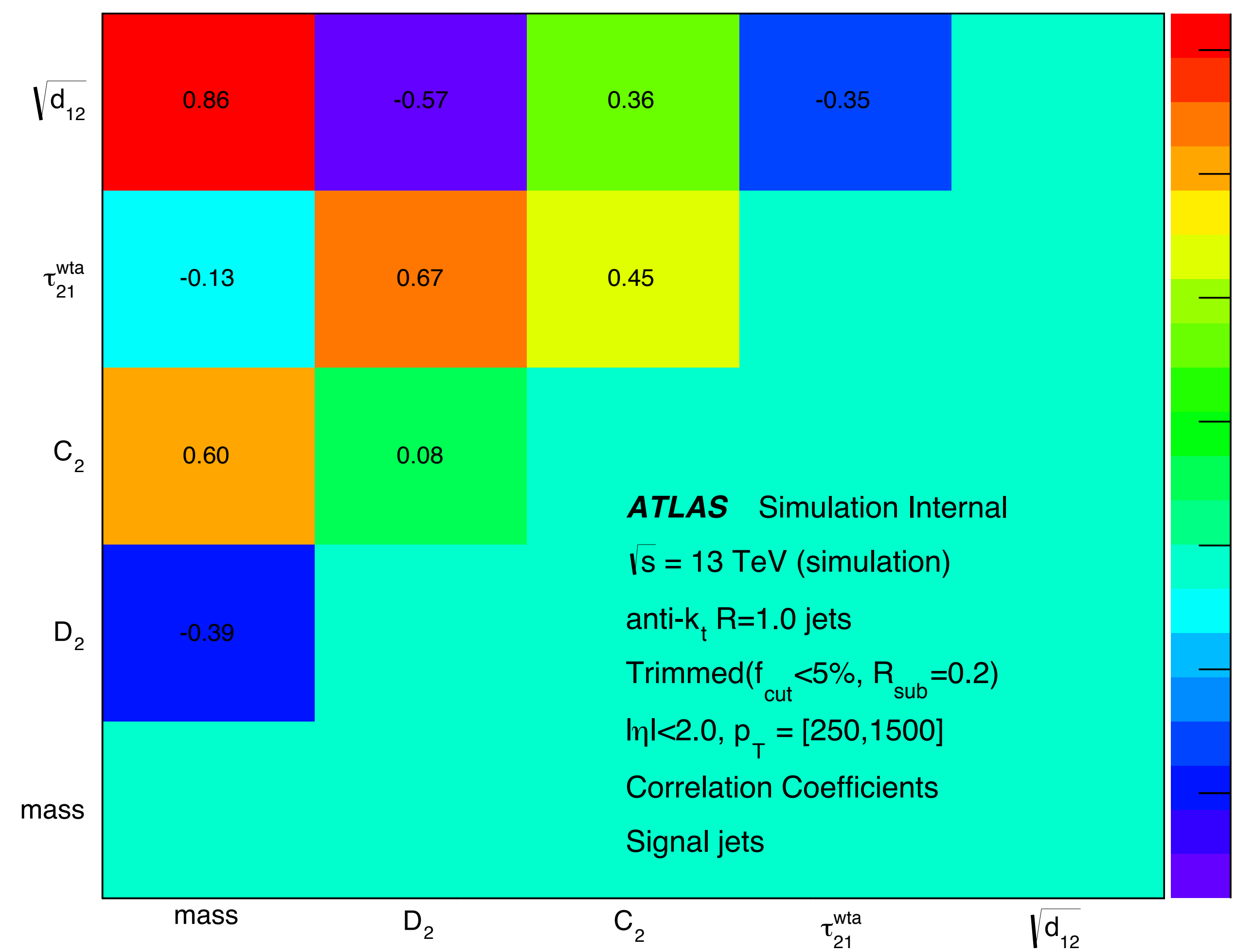
Higgs Tagging

- Correlation matrices:
  - 5 variables studied;
  - considered all the combinations.

Background

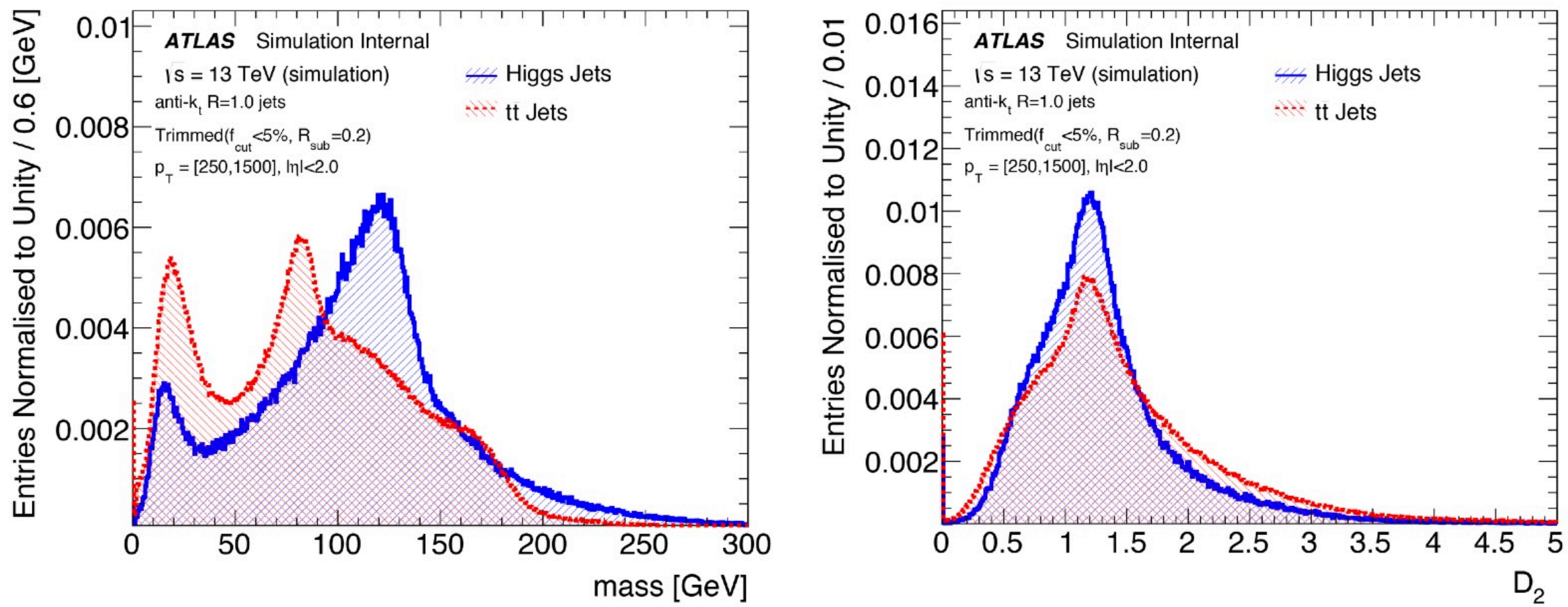


Signal



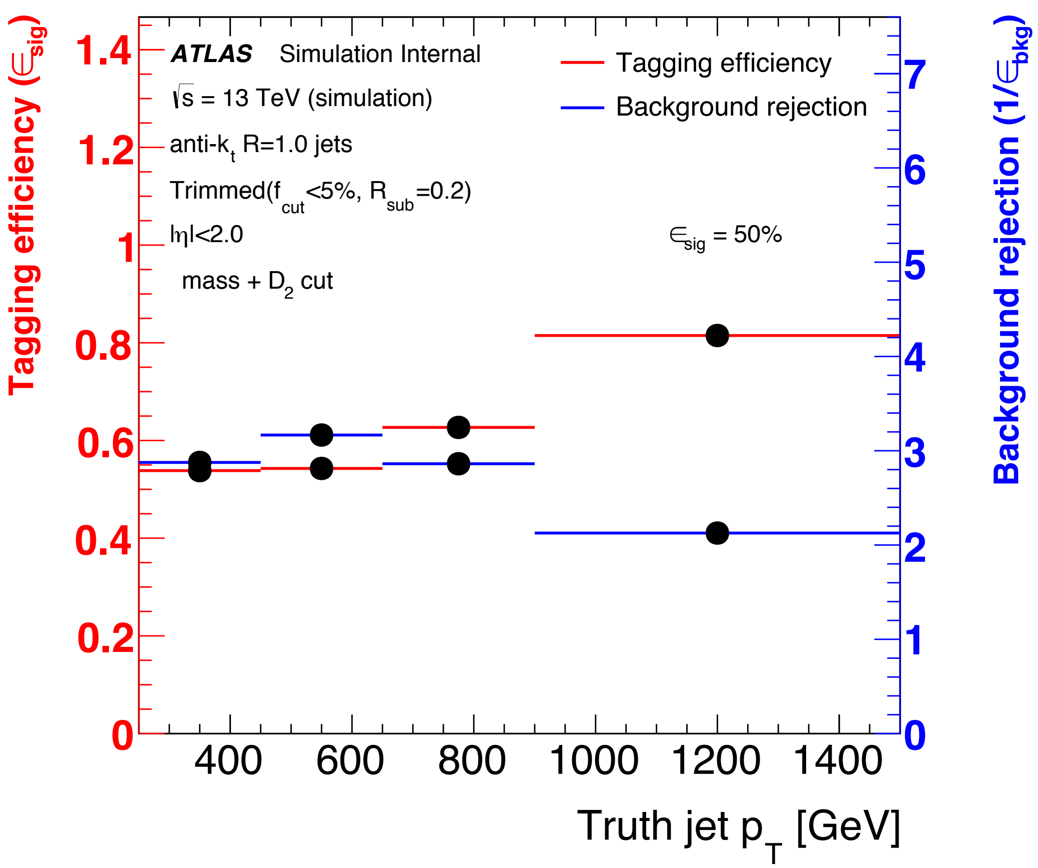
Higgs Tagging

Substructure variables
in the algorithm



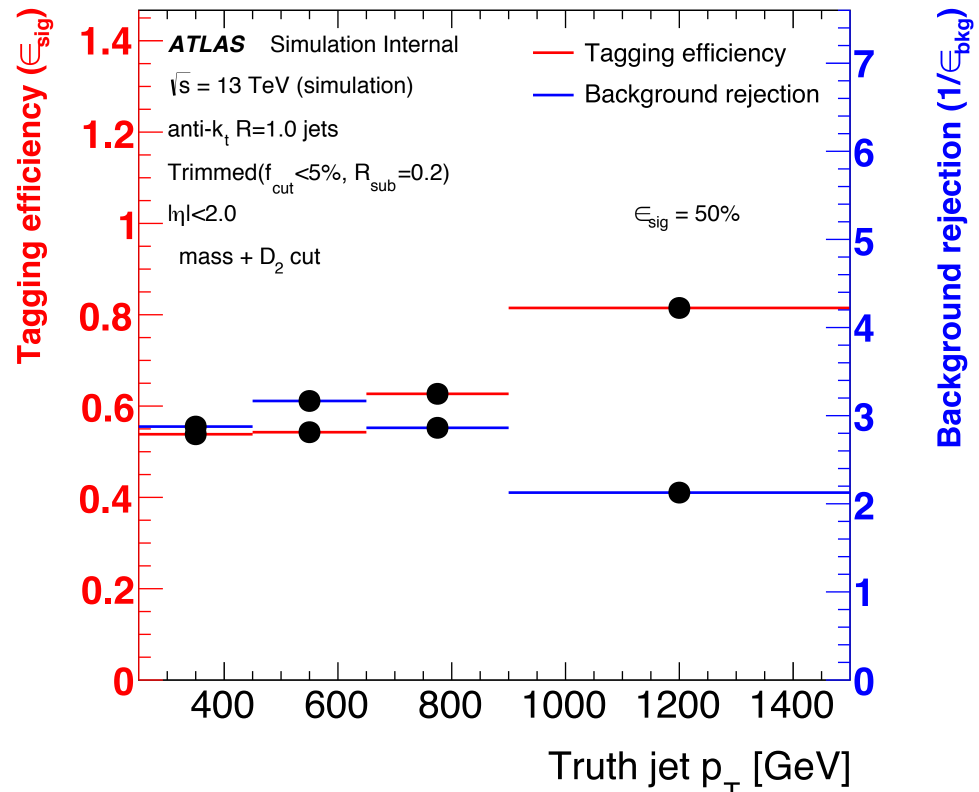
Performances

- Signal efficiency on ttH and background rejection on tt;
- **only one working point**: 50% signal efficiency.

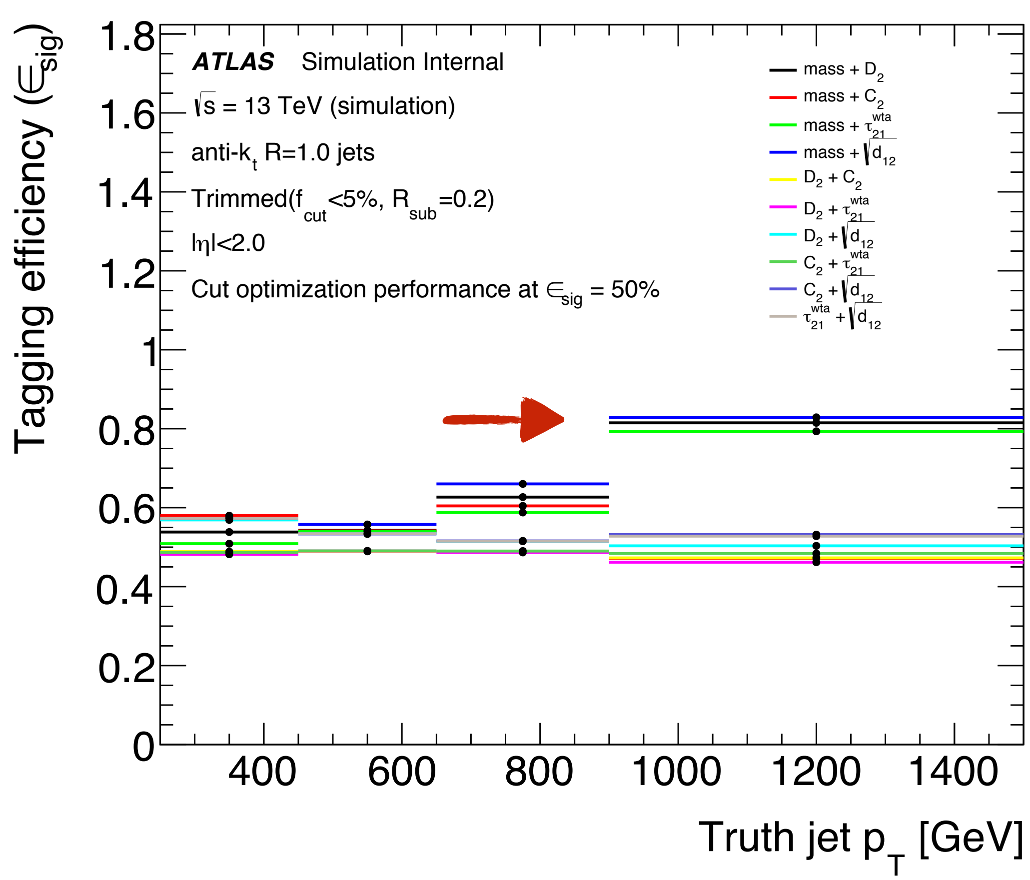


Higgs Tagging performances

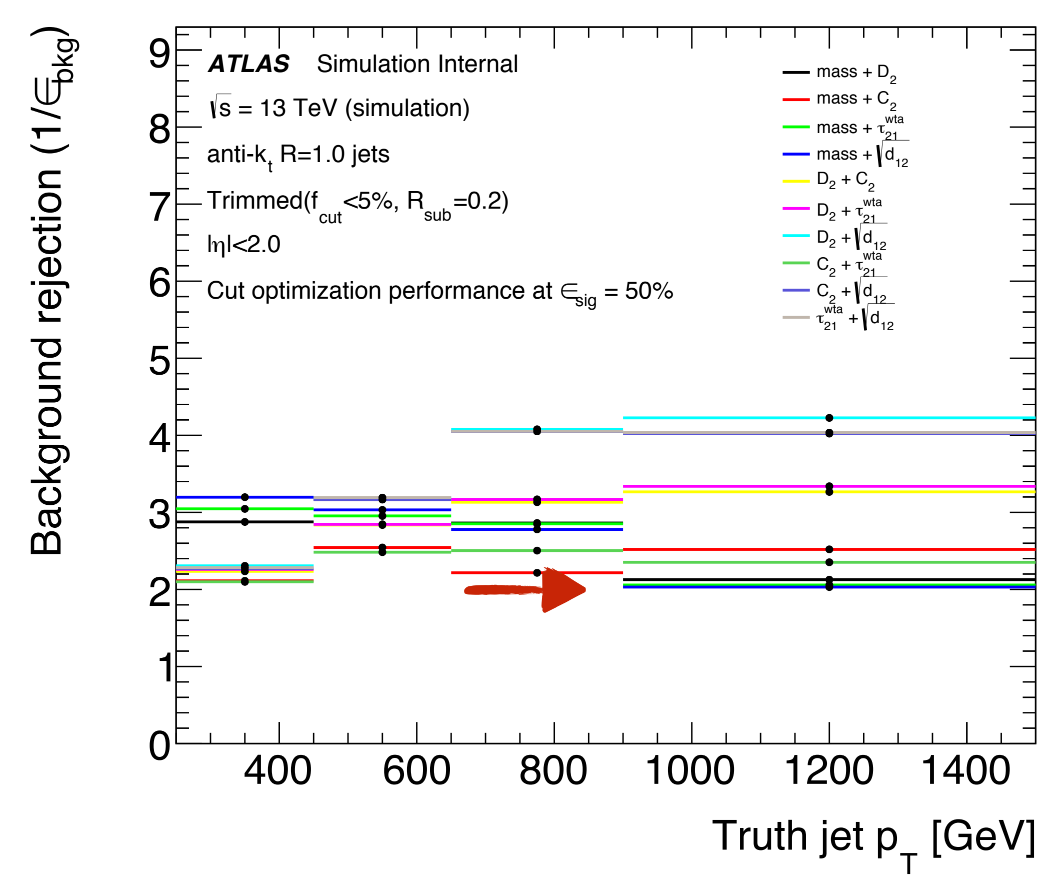
Aim: **discriminate ttH from tt.**



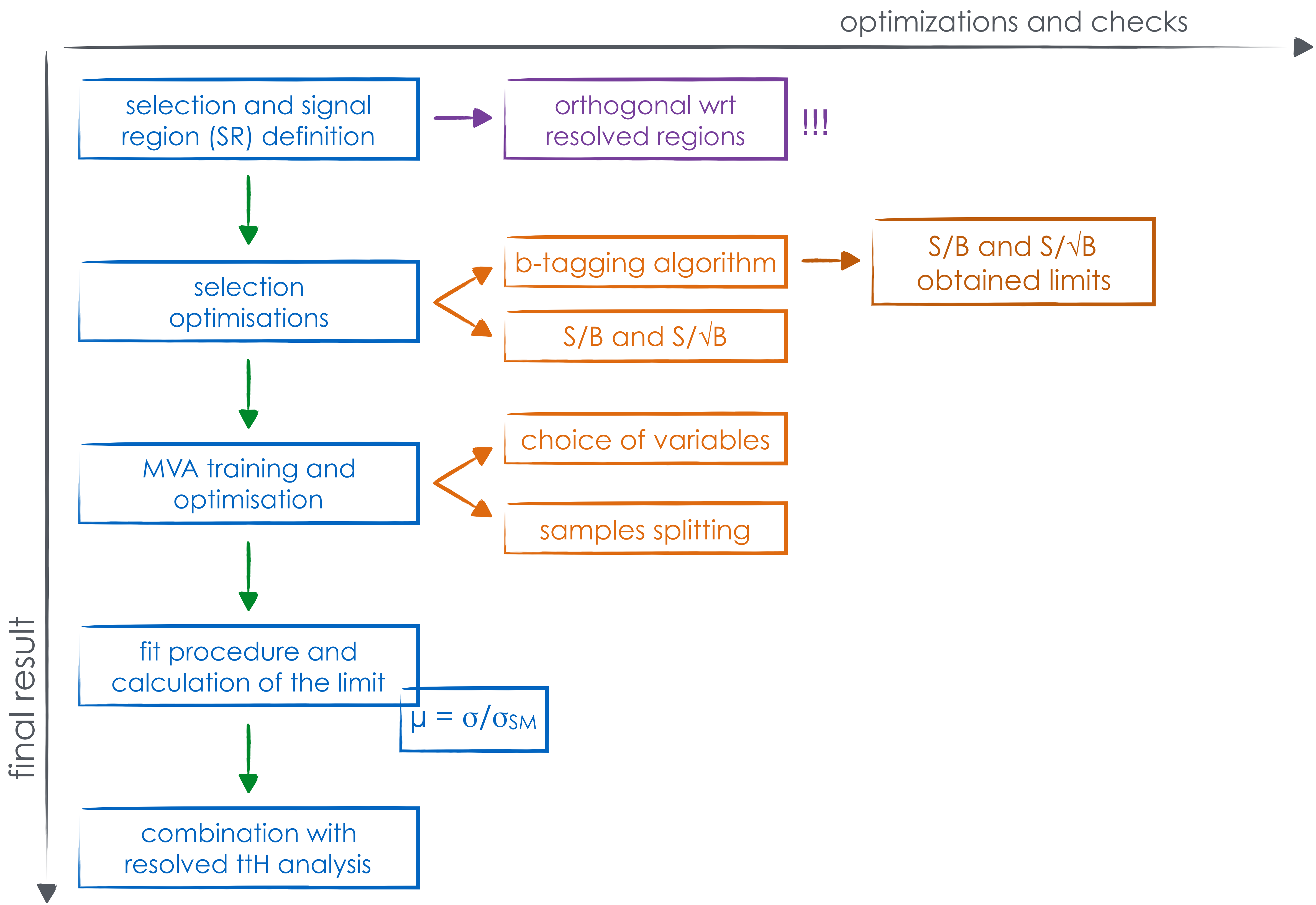
- Different variables combinations studied:
  - correlations, cuts, signal efficiencies and background rejections;
  - **m-D_2** seems the best choice (slightly better than **m-τ_{21}**).



Only one working point (50%).

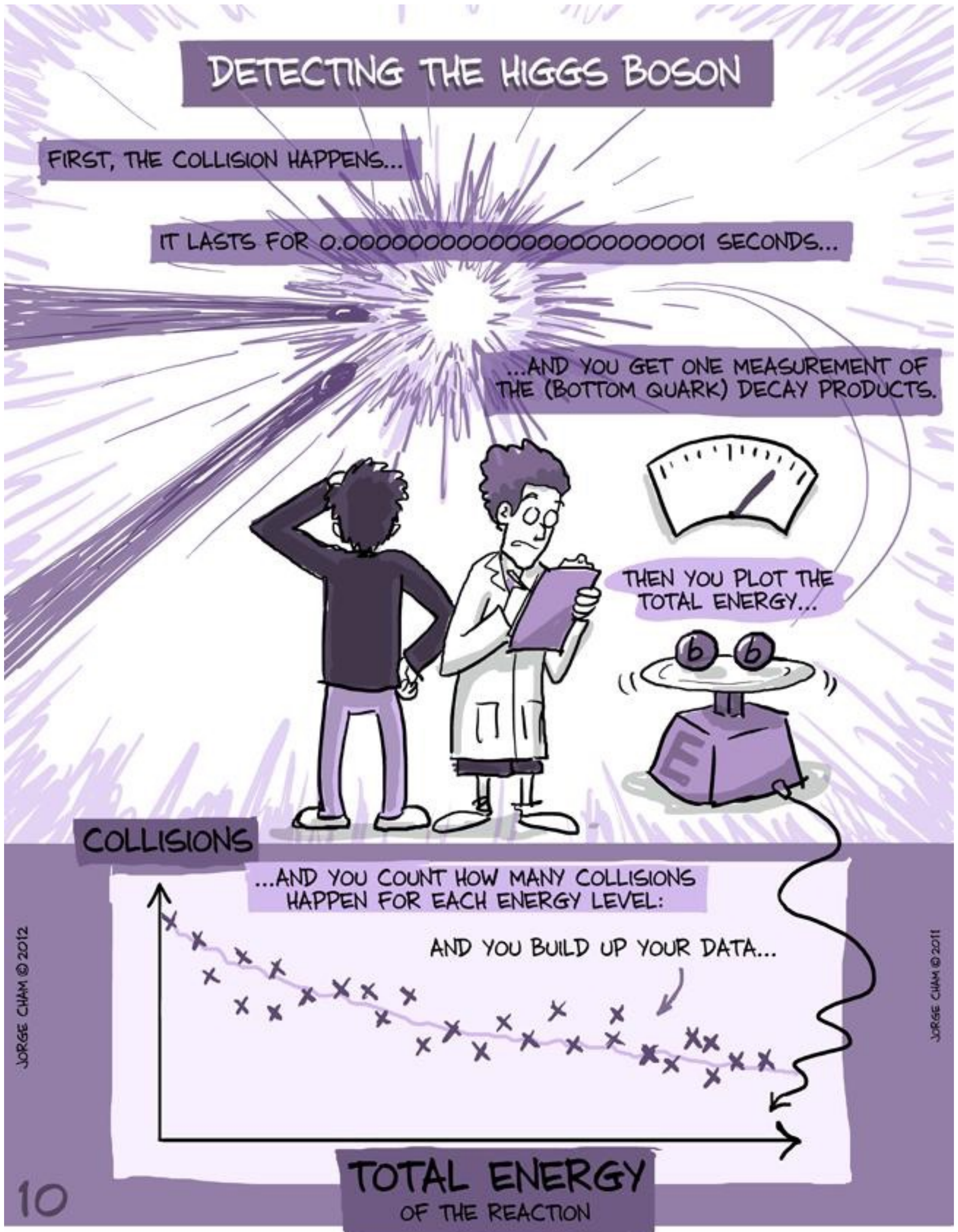


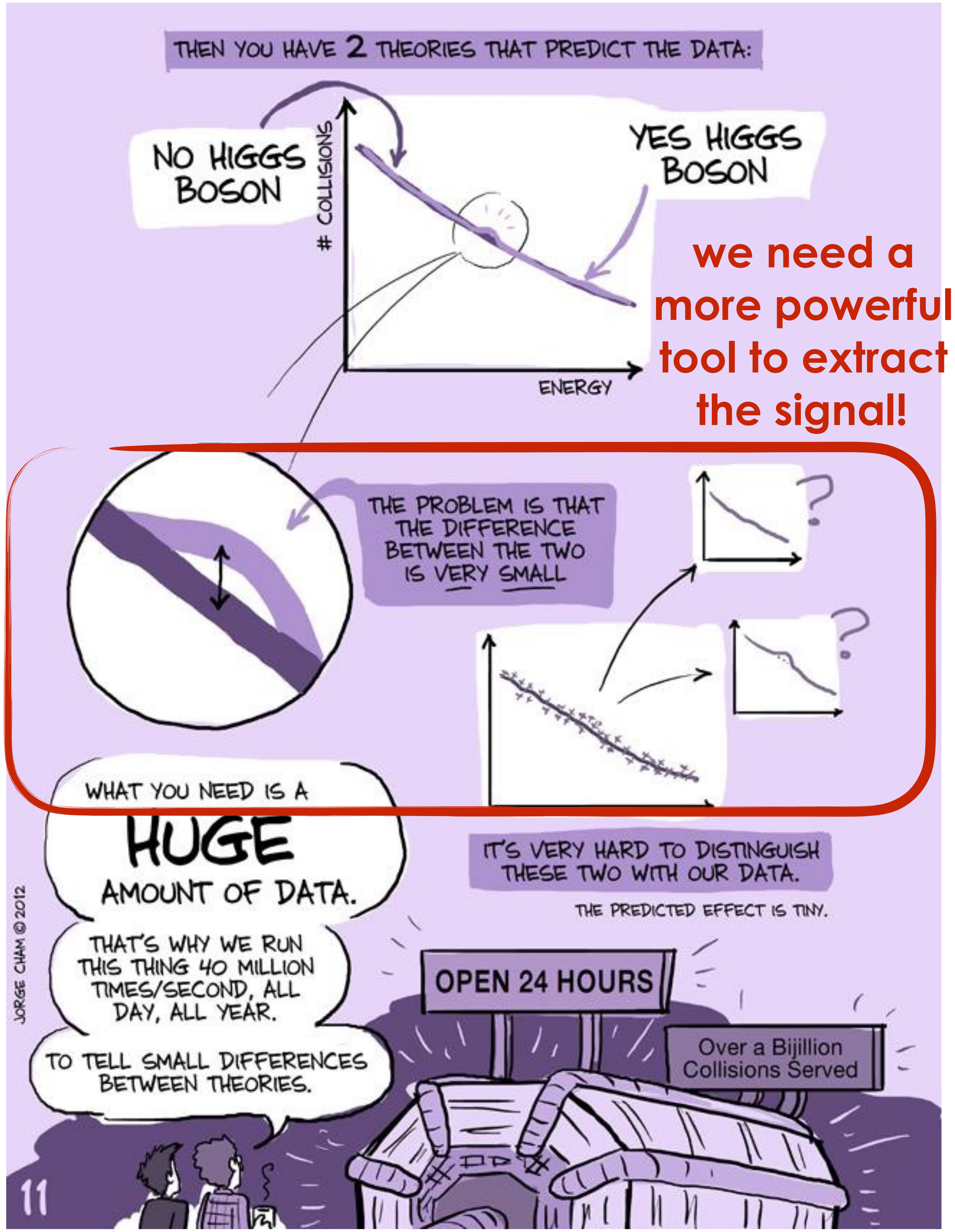
Analysis strategy



The problem is…

http://www.phdcomics.com/

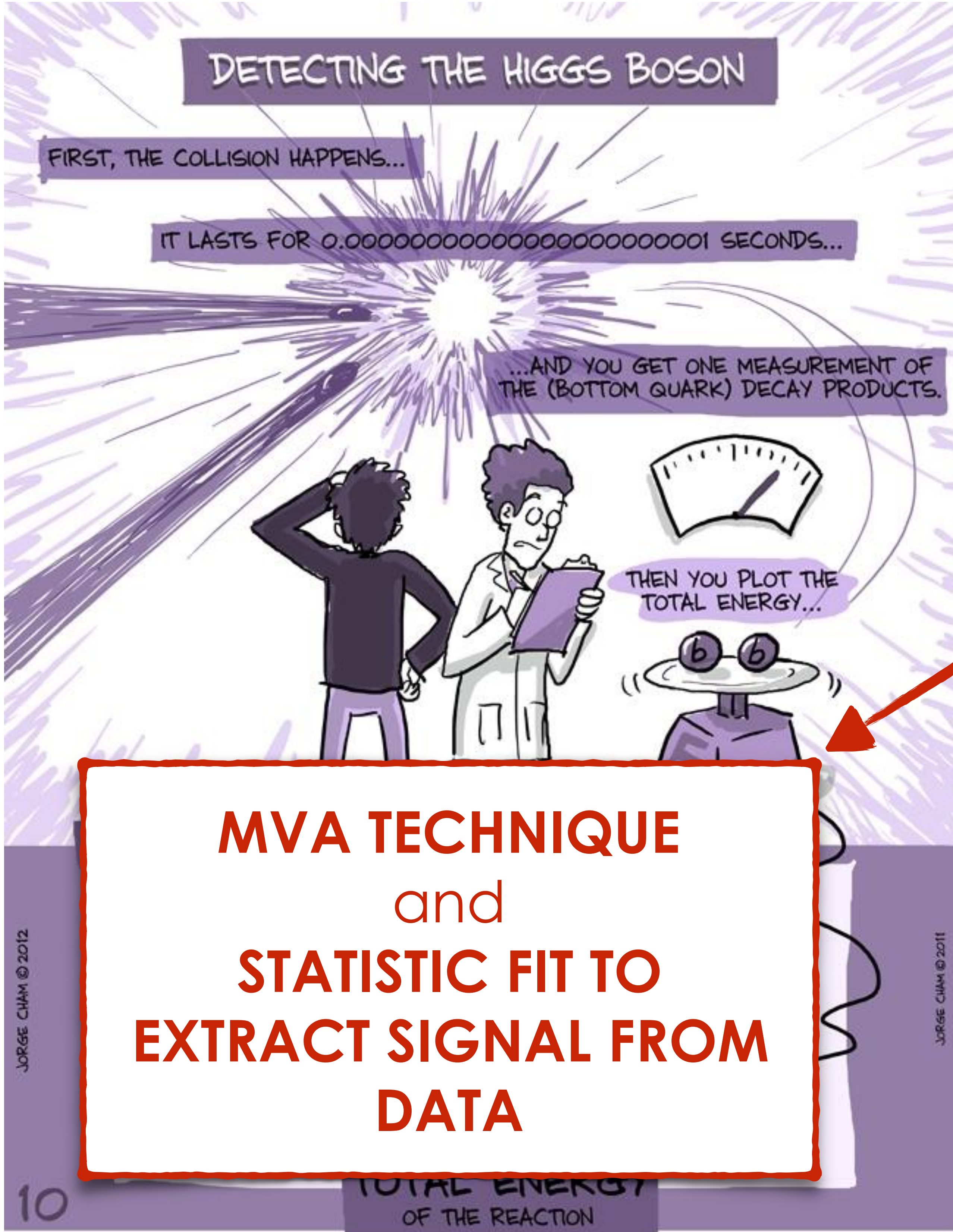




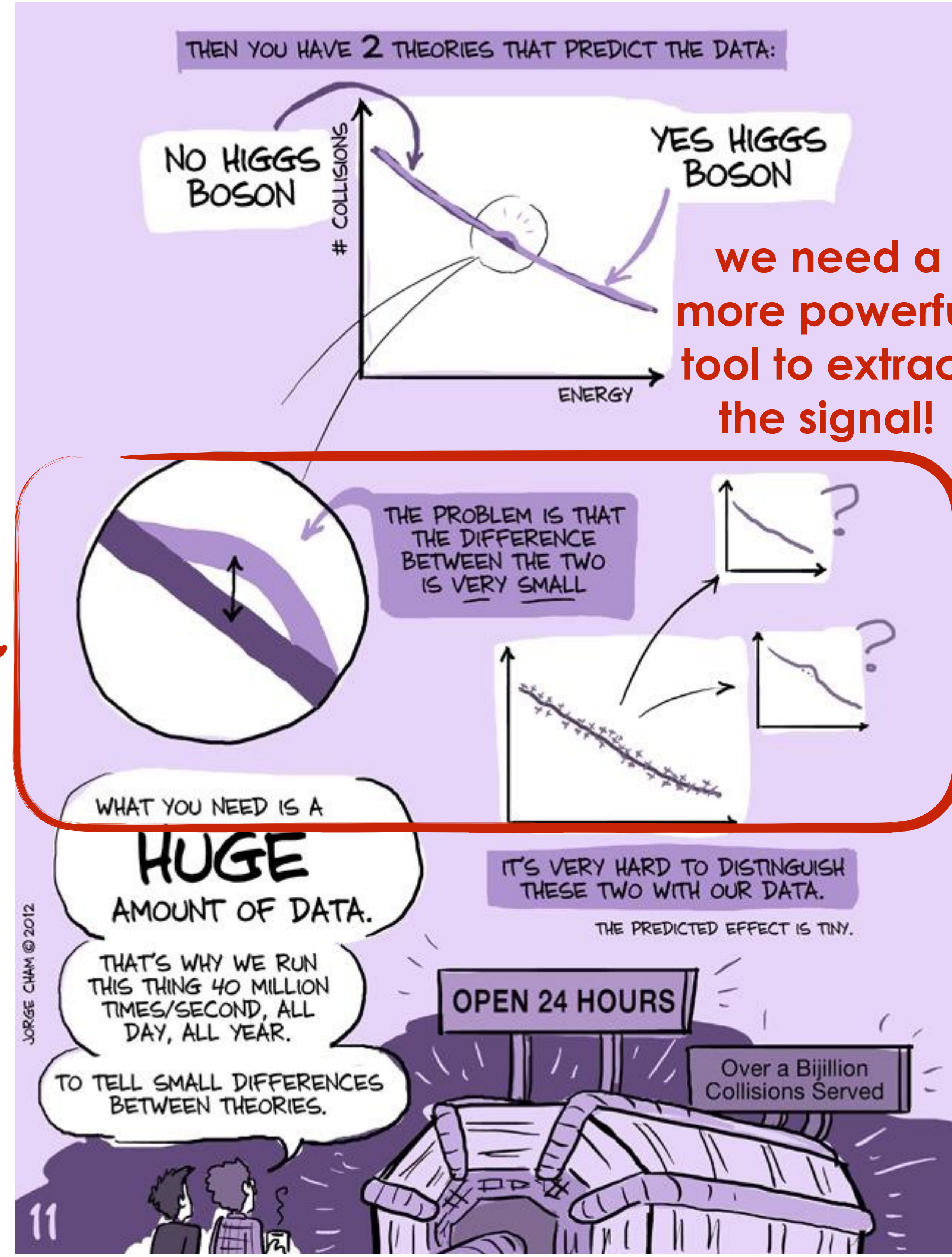
we need a more powerful tool to extract the signal!

...but we have

http://www.phdcomics.com/



MVA TECHNIQUE
and
STATISTIC FIT TO EXTRACT SIGNAL FROM DATA



we need a more powerful tool to extract the signal!

MC simulation samples

PDF, Parton Shower and Normalisation choice

Sample	Generator	PDF	Parton Shower	Normalisation
$t\bar{t}H$	aMC@NLO	NNPDF3.0NLO	Pythia 8.2	(N)NLO
$t\bar{t}$+jets	Powheg	CTEQ6L1	Pythia 6.428	NNLO+NNLL
W/Z + jets	Sherpa	CT10	Sherpa 2.1.1	NNLO
Single top (s-, Wt-channels)	Powheg	CT10	Pythia 6.428	aNNLO
Single top (t-channel)	Powheg	CT10f4	Pythia 6.428	aNNLO
$t\bar{t} + V$	Madgraph	CTEQ6L1	Pythia 6.425	NLO
Diboson	Sherpa	CT10	Sherpa 2.1.1	NLO

Matching method choice

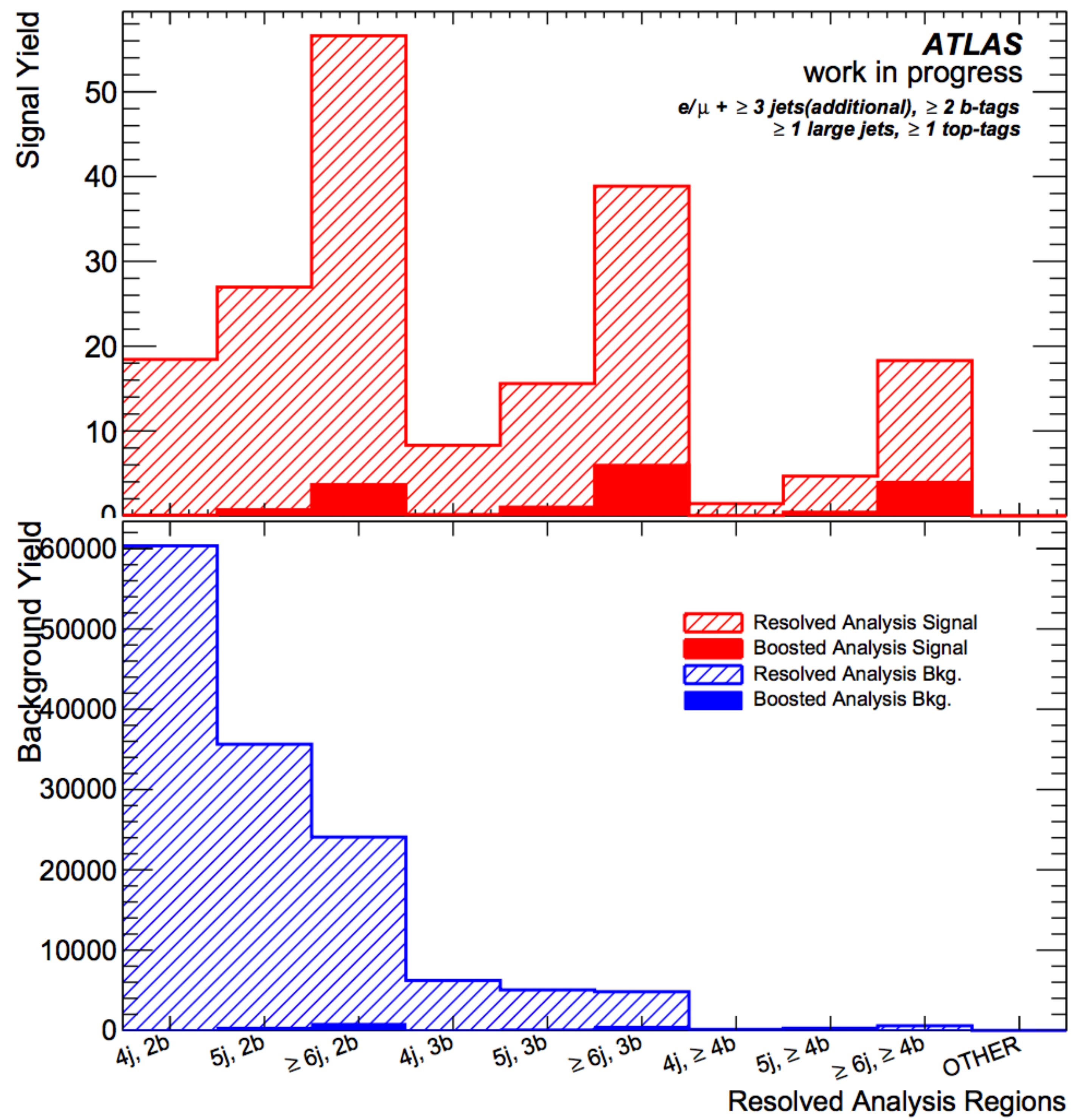
Sample	Generator	Matching method
$t\bar{t}H$	aMC@NLO	NLO+PS
$t\bar{t}$+jets	Powheg	NLO+PS
W/Z + jets	Sherpa	ME+PS (LO)
Single top (s-, Wt-channels)	Powheg	NLO+PS
Single top (t-channel)	Powheg	NLO+PS
$t\bar{t} + V$	Madgraph	ME+PS (LO)
Diboson	Sherpa	ME+PS (LO)

"Resolved safe" analysis

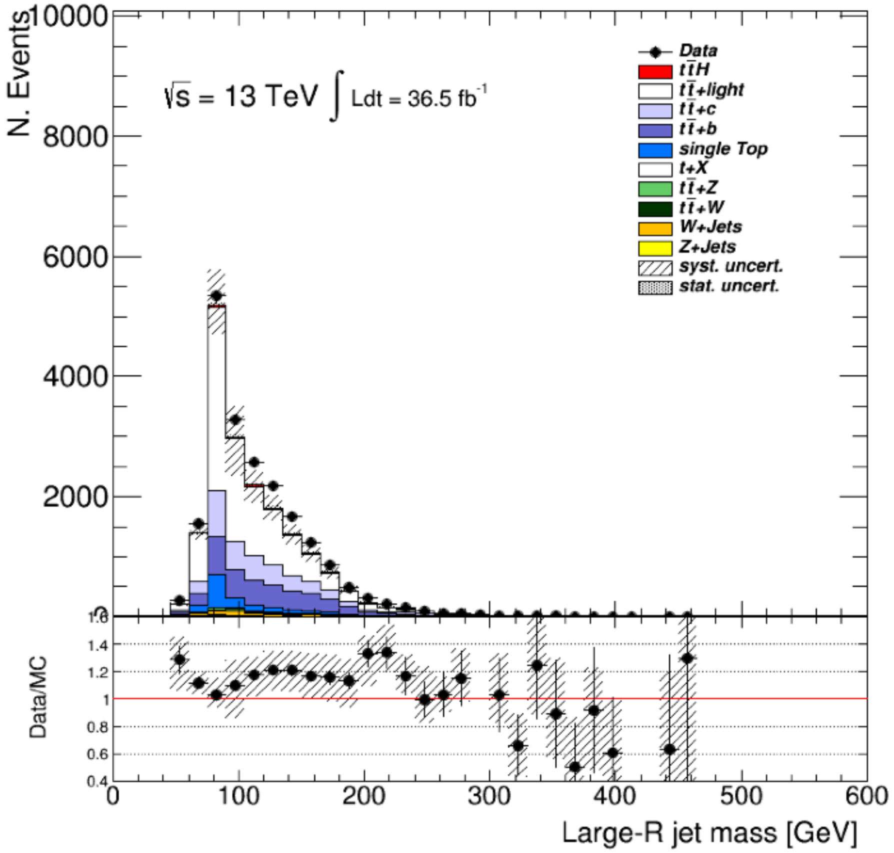
- **Goal:**
  - combination resolved+boosted;
- **Why:**
  - increasing the analysis sensitivity;
  - increasing the phase space;
- **How:**
  - using orthogonal and independent regions, otherwise risk of correlation amongst measurements.

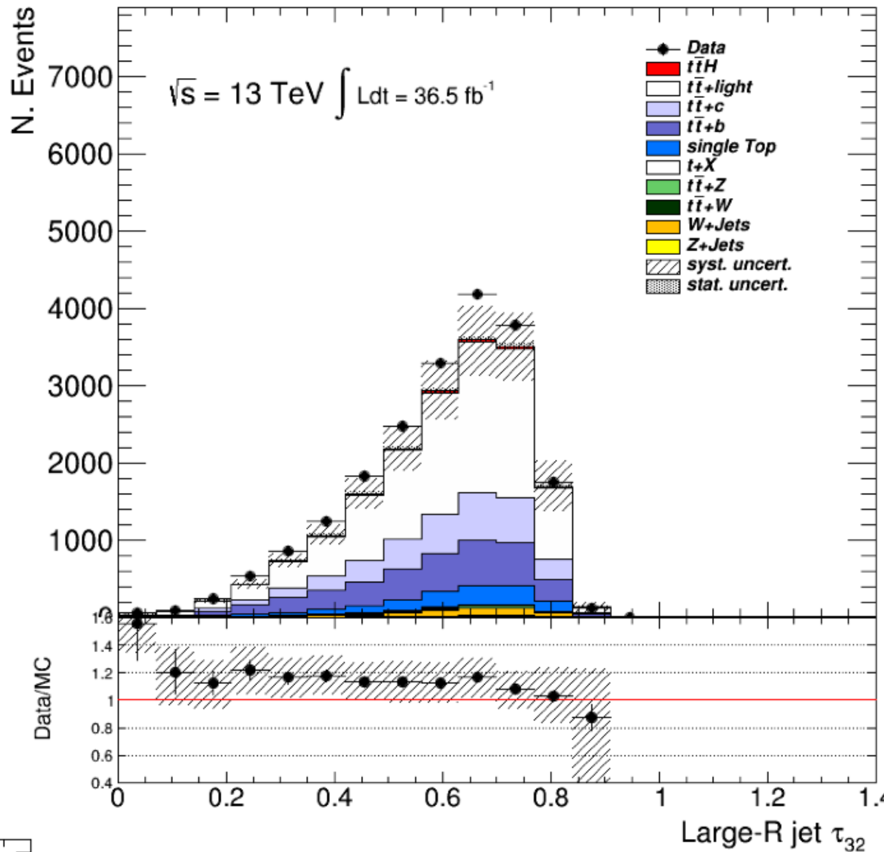
Overlap of boosted events into resolved regions

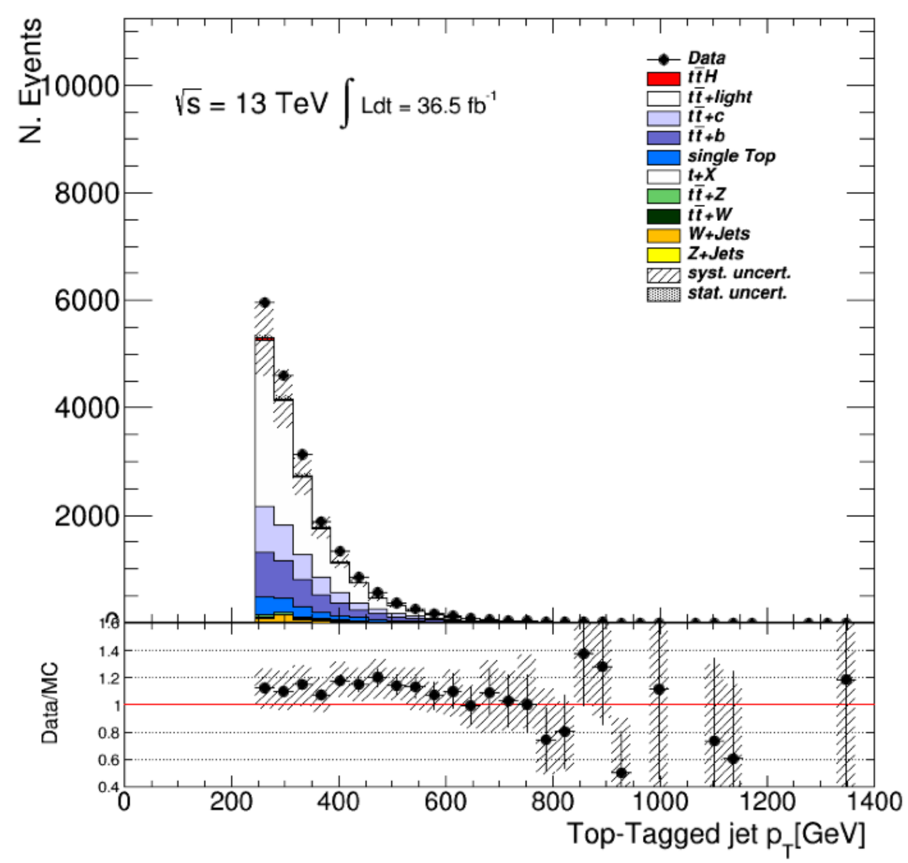
In resolved signal (SR) e control (CR) regions: veto on events passing the boosted selection



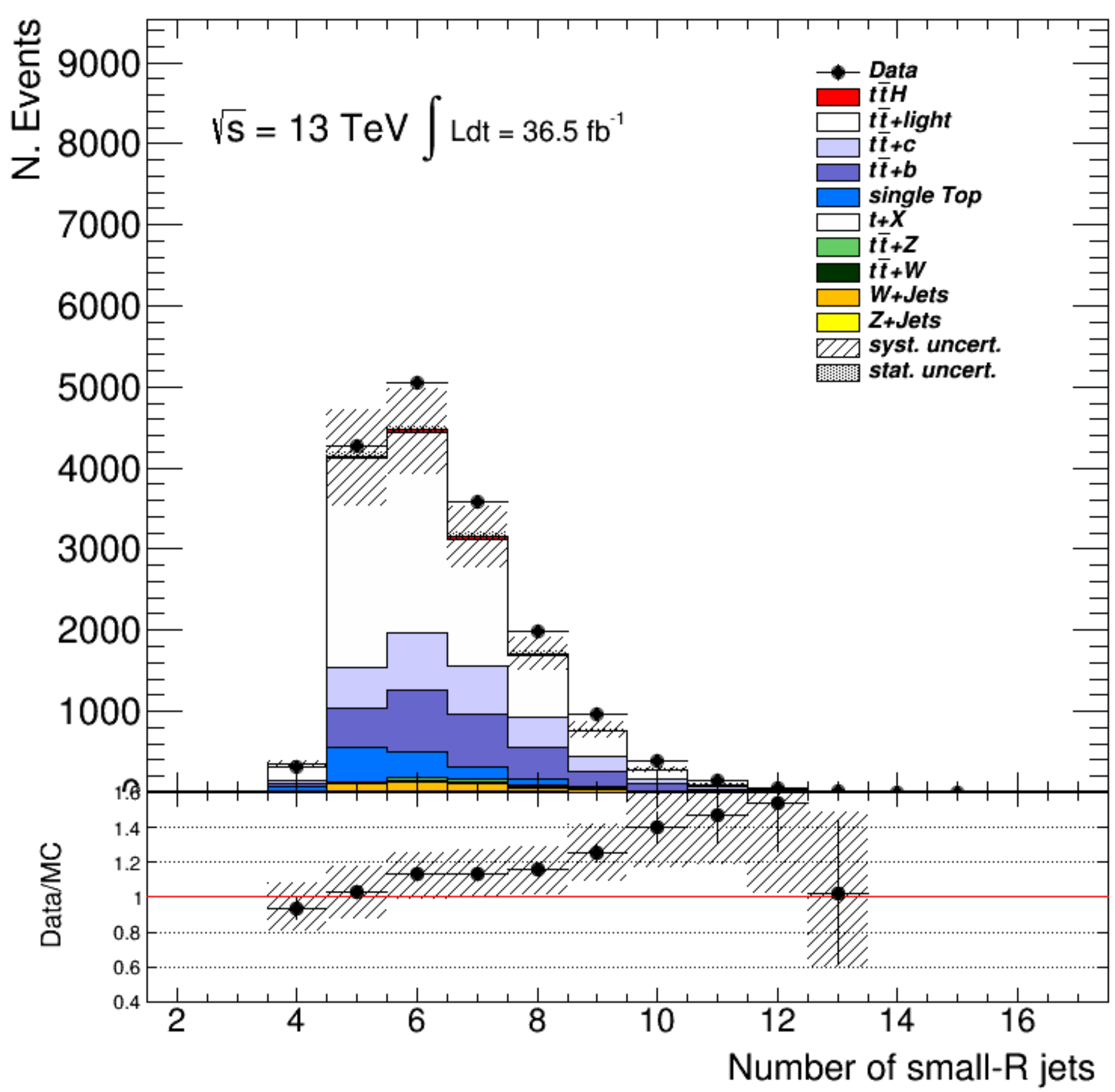
Boosted selection: distributions

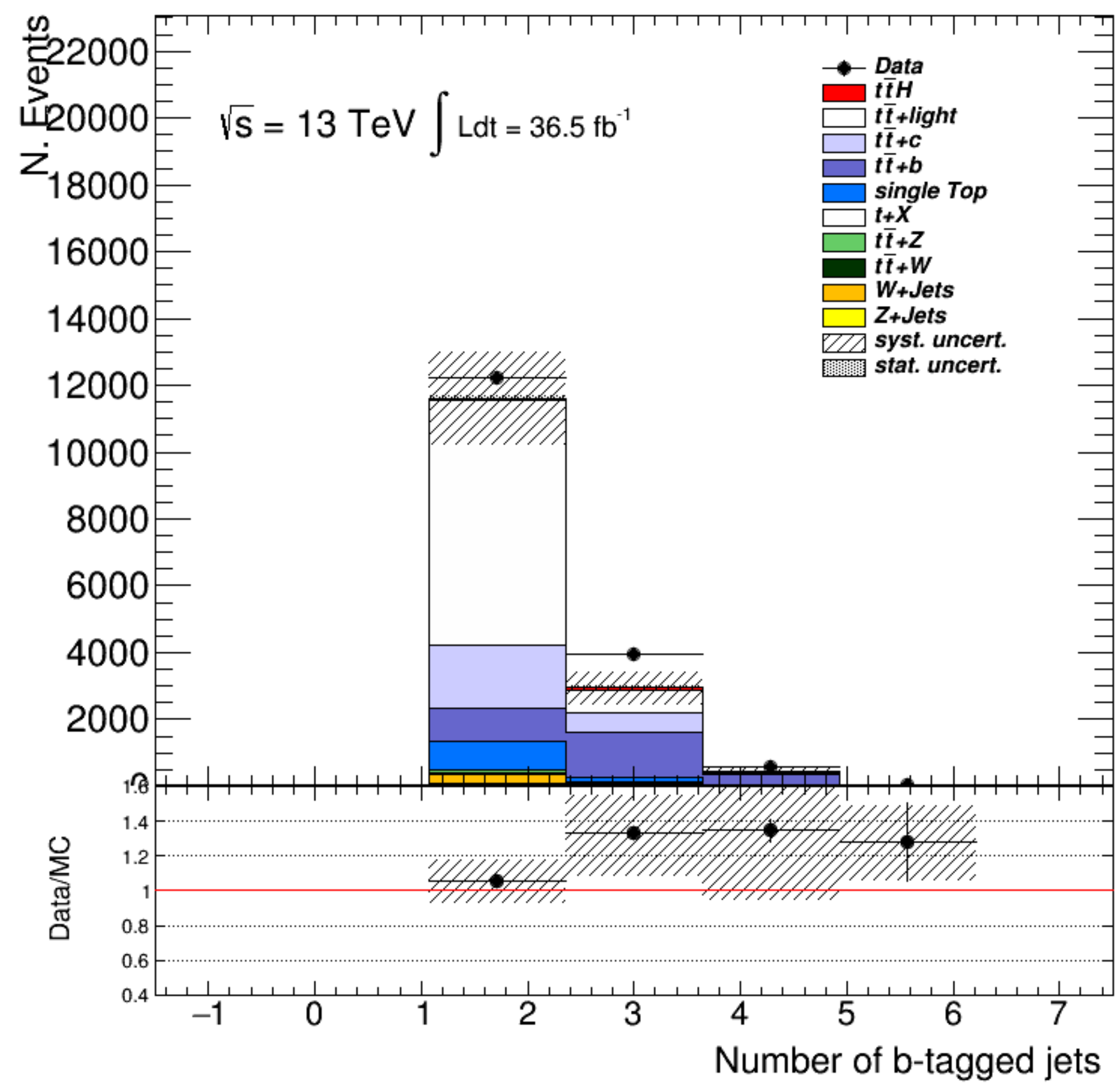




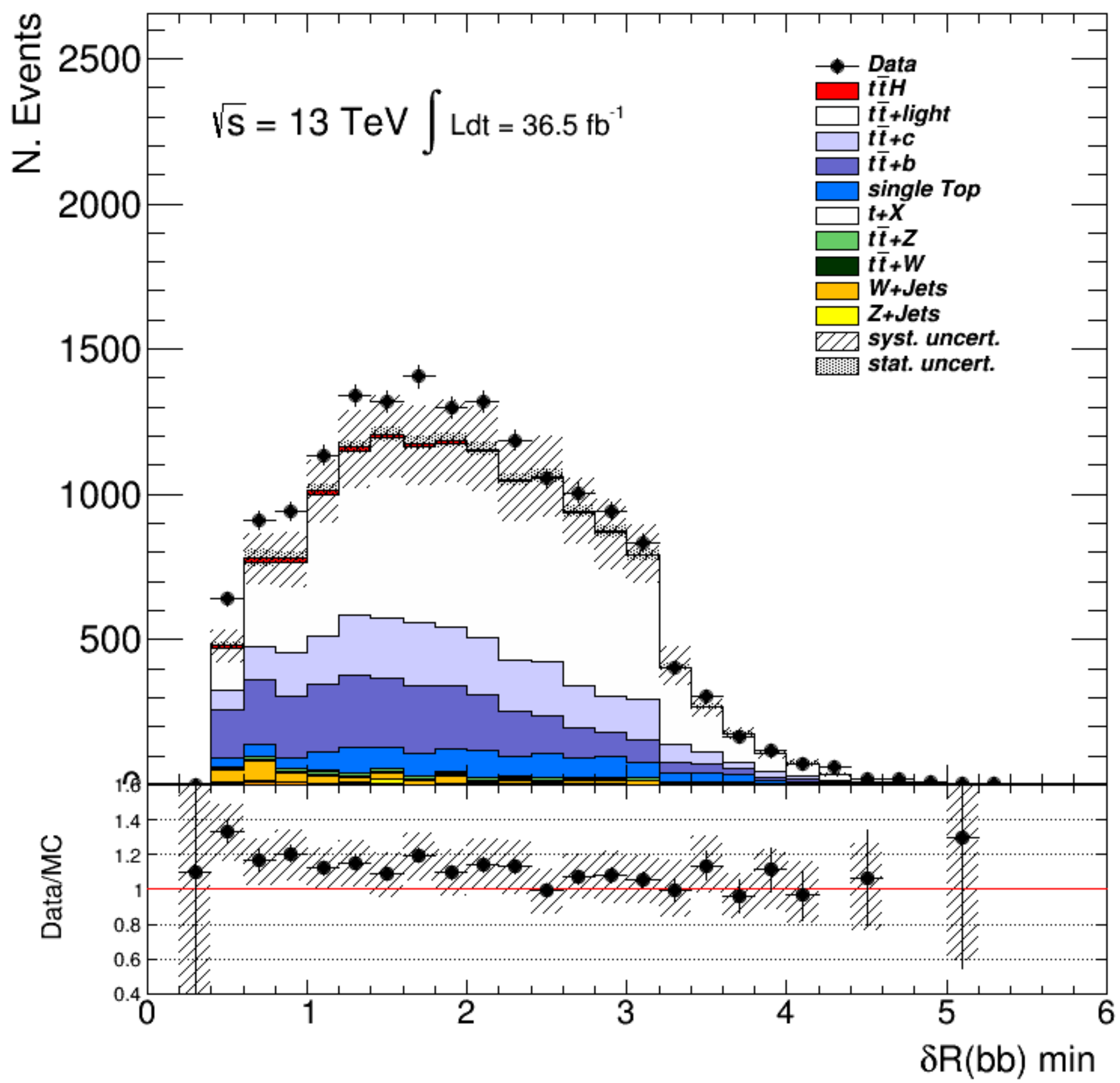


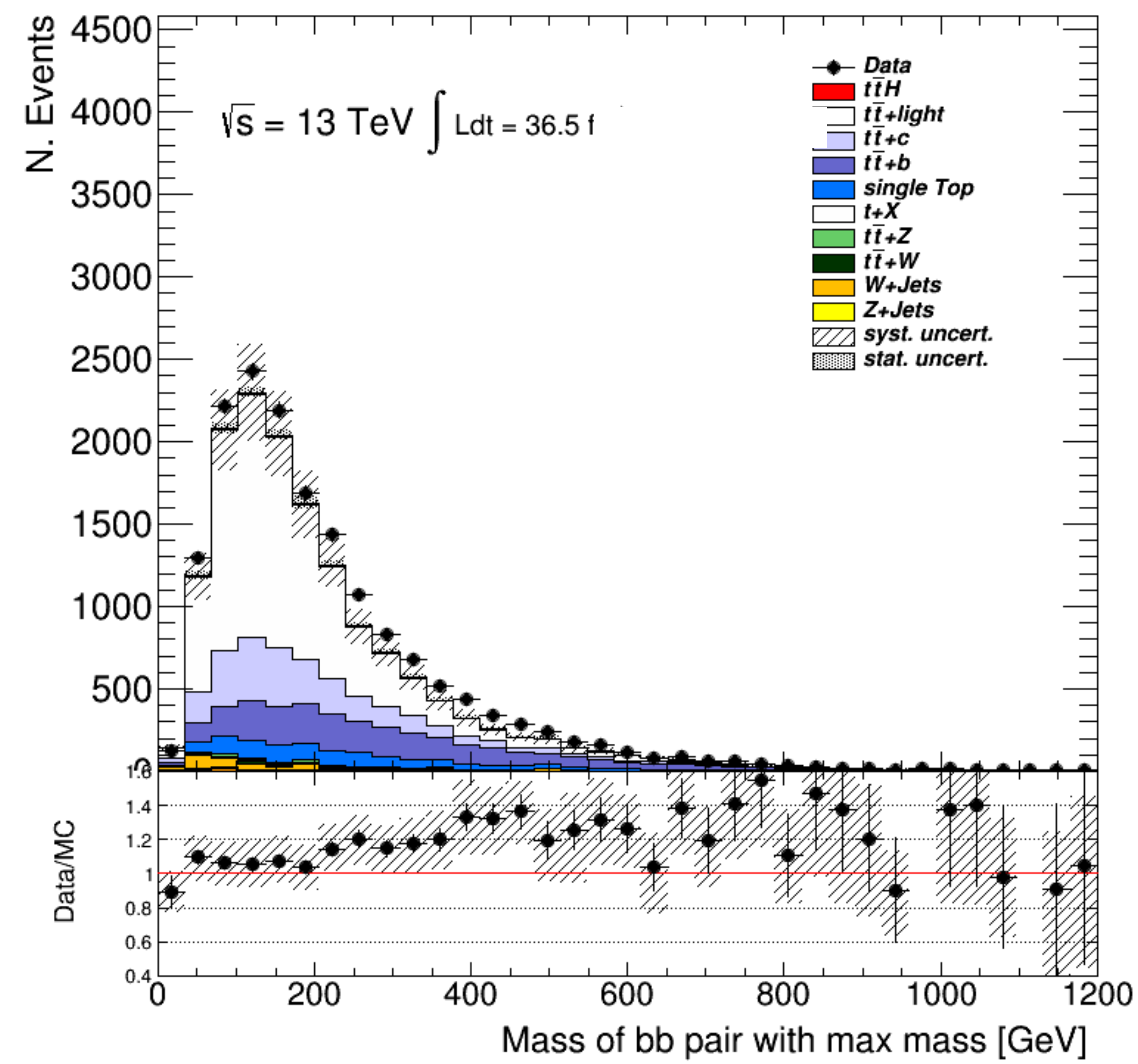
Boosted selection: distributions





Boosted selection: distributions





Yields: Control regions

Post-fit

	4j, 2b	4j, 3b	4j, $\geq$ 4b	5j, 2b	5j, 3b	$\geq$ 6j, 2b
$t\bar{t}H$ ($H \rightarrow b\bar{b}$)	120 ± 50	70 ± 30	11 ± 4	170 ± 70	130 ± 50	290 ± 120
$t\bar{t}H$ ($H \rightarrow WW$)	30 ± 11	2 ± 1	0.02 ± 0.01	50 ± 20	4 ± 1	170 ± 60
$t\bar{t}H$ ($H \rightarrow other$)	35 ± 14	3 ± 1	0.08 ± 0.03	60 ± 20	6 ± 2	130 ± 50
$t\bar{t}$+light	439000 ± 6000	20700 ± 700	160 ± 30	262000 ± 5000	16200 ± 600	161000 ± 5000
$t\bar{t}$+$\geq$1c	18000 ± 6000	1600 ± 500	0 ± 6	19300 ± 5400	2300 ± 600	26000 ± 5000
$t\bar{t}$+$\geq$1b	10700 ± 1800	4000 ± 600	230 ± 20	10700 ± 1700	5700 ± 700	11900 ± 1800
$t\bar{t}$+W	250 ± 30	15 ± 2	0.4 ± 0.1	340 ± 40	27 ± 4	530 ± 70
$t\bar{t}$+Z	290 ± 30	40 ± 5	4 ± 1	370 ± 40	70 ± 9	690 ± 80
Single Top	28500 ± 1700	1500 ± 100	47 ± 7	14000 ± 1000	1110 ± 80	7900 ± 700
Diboson	1300 ± 600	110 ± 50	12 ± 9	800 ± 400	60 ± 30	600 ± 300
W+jets	18800 ± 1400	1110 ± 140	6 ± 2	8800 ± 800	880 ± 170	4600 ± 500
Z+jets	4700 ± 2100	200 ± 90	2 ± 1	2000 ± 1000	380 ± 190	1400 ± 600
t+X	96 ± 4	22 ± 1	2.2 ± 0.3	68 ± 3	13 ± 1	84 ± 5
Total	522000 ± 9000	29000 ± 1000	480 ± 40	319000 ± 8000	27000 ± 1000	215000 ± 7000
Data	521749	29398	530	318964	26905	214822

Yields: Signal regions

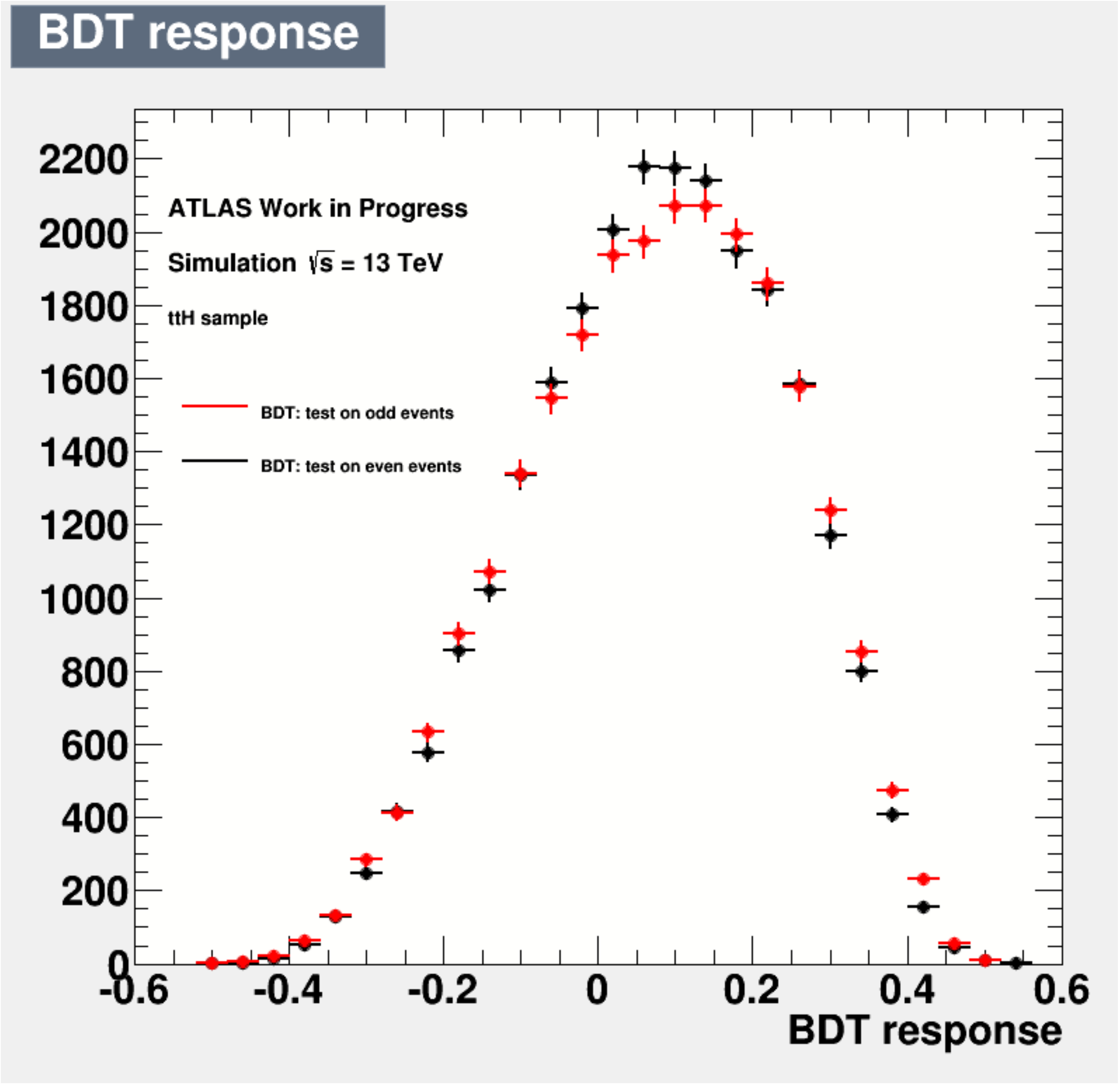
Post-fit

	5j, $\geq$ 4b	$\geq$ 6j, 3b	$\geq$ 6j, $\geq$ 4b	boosted 3j, 2b
$t\bar{t}H$ ($H \to b\bar{b}$)	42 $\pm$ 15	380 $\pm$ 140	170 $\pm$ 60	130 $\pm$ 50
$t\bar{t}H$ ($H \to WW$)	0.10 $\pm$ 0.04	28 $\pm$ 8	3 $\pm$ 1	18 $\pm$ 6
$t\bar{t}H$ ($H \to other$)	0.5 $\pm$ 0.2	26 $\pm$ 8	4 $\pm$ 1	17 $\pm$ 6
$t\bar{t}$+light	370 $\pm$ 70	14500 $\pm$ 800	400 $\pm$ 80	10100 $\pm$ 300
$t\bar{t}$+$\geq$1c	80 $\pm$ 30	4700 $\pm$ 900	400 $\pm$ 100	1600 $\pm$ 500
$t\bar{t}$+$\geq$1b	660 $\pm$ 50	9330 $\pm$ 970	2160 $\pm$ 140	2800 $\pm$ 300
$t\bar{t}$+W	0.8 $\pm$ 0.1	90 $\pm$ 13	7 $\pm$ 1	80 $\pm$ 10
$t\bar{t}$+Z	15 $\pm$ 2	220 $\pm$ 30	63 $\pm$ 9	151 $\pm$ 19
Single Top	64 $\pm$ 9	1190 $\pm$ 110	128 $\pm$ 14	1170 $\pm$ 90
Diboson	2 $\pm$ 1	80 $\pm$ 40	10 $\pm$ 5	70 $\pm$ 30
W+jets	10 $\pm$ 5	580 $\pm$ 80	40 $\pm$ 13	540 $\pm$ 60
Z+jets	2 $\pm$ 2	90 $\pm$ 40	9 $\pm$ 4	90 $\pm$ 40
t+X	3.1 $\pm$ 0.4	42 $\pm$ 2	15 $\pm$ 1	30 $\pm$ 2
Total	1250 $\pm$ 90	31200 $\pm$ 1500	3400 $\pm$ 200	16700 $\pm$ 700
Data	1235	31401	3398	16763

MultiVariate Analysis

- test to check if our training is stable
  - compatibility between BDT output from two halves of the total samples

- signal sample only



MVA in ttH

- **Separation**: $$< S^2 >= \frac{1}{2}\int \frac{(\hat{y}_S(y) - \hat{y}_B(y))^2}{\hat{y}_S(y) + \hat{y}_B(y)} dy$$
  - y_S and y_B are the signal and background probability density functions of y, respectively;
  - zero for identical signal and background shapes and 1 for shapes with no overlap.

- **Correlation**: $$\rho(X,Y) = \frac{cov(X,Y)}{\sigma_X \sigma_Y}$$
  - two random variables X and Y;
  - cov is the covariance and sigma(X) (sigma(Y)) is the variance of X (Y).

- **Importance ranking**:
  - by evaluating the number of times the variables are used to split decision tree nodes;
  - by weighting each split occurrence (by using the same variable) by the separation achieved and by the number of events in the node.

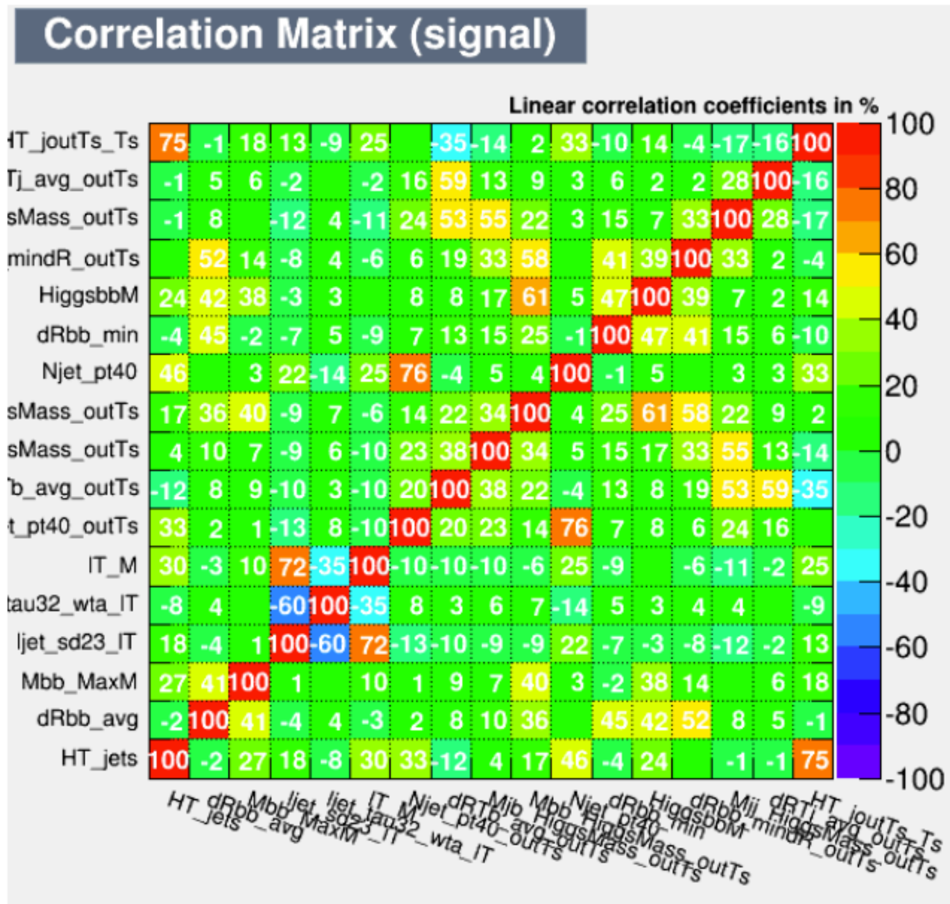
MVA training: definitions

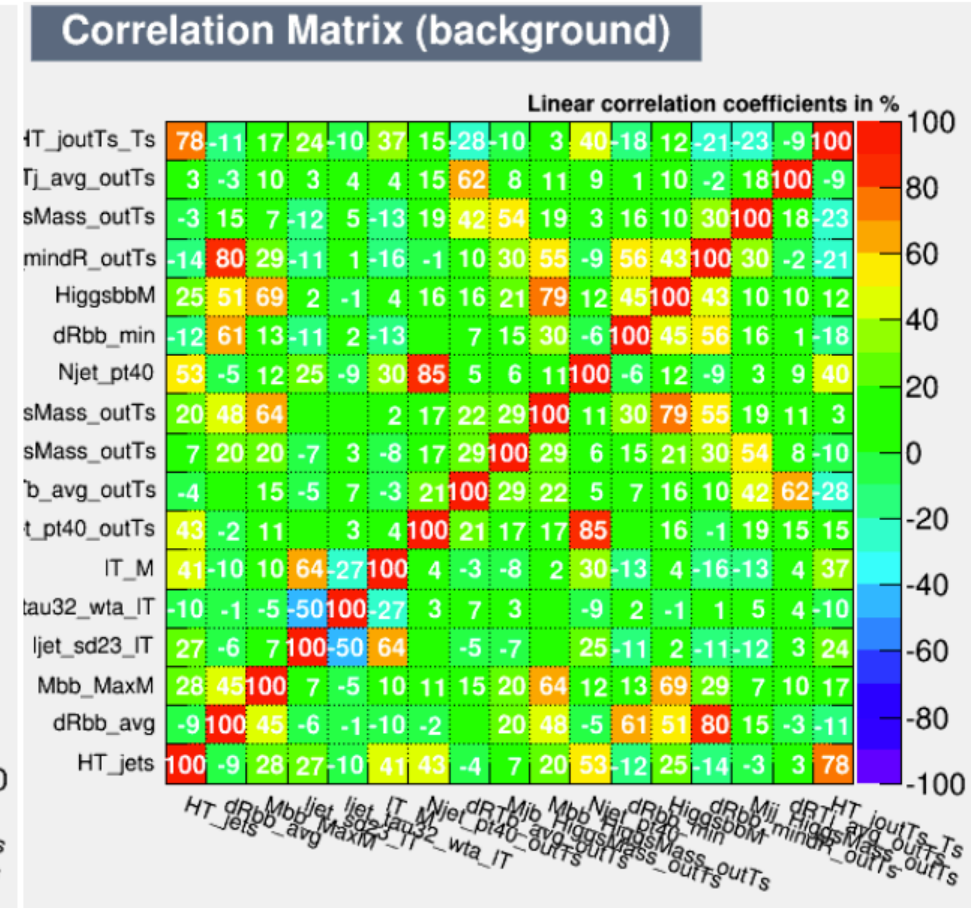
- $\boldsymbol{m_{top}^{lead}}$: mass of the leading top-tagged large-R jet in the event;
- $\boldsymbol{\tau_{32,top}^{lead}}$: N-subjettiness ratio τ_{32} (see Chapter 4) of the leading top-tagged large-R jet in the event;
- $\boldsymbol{\sqrt{d_{23,top}^{lead}}}$: splitting scale $\sqrt{d_{23}}$ (see Chapter 4) of the leading top-tagged large-R jet in the event;
- $\boldsymbol{\Delta R_{bb}^{min}}$: minimum ΔR between any b-jets in the event;
- $\boldsymbol{\Delta R_{bb}^{avg}}$: average ΔR between any b-jets in the event;
- $\boldsymbol{\Delta R_{(add)bb}^{min}}$: minimum ΔR between b-jets that do not overlap with any top-tagged large-R jet in the event ("add" means additional);
- $\boldsymbol{\Delta R_{top(add)b}^{avg}}$: average ΔR between the top-tagged large-R jets and b-jets that do not overlap in the event;
- $\boldsymbol{\Delta R_{top(add)j}^{avg}}$: average ΔR between the top-tagged large-R jets and jets that do not overlap in the event;
- $\boldsymbol{m_{bb}^{max}}$: maximum mass of a pair of b-jets in the event;
- $\boldsymbol{m_{(add)bb}^{H}}$: invariant mass of a pair of b-jets (that do not overlap with any top-tagged large-R jet) that is closest to the mass of the Higgs boson ($m_H \equiv 125$ GeV);
- $\boldsymbol{m_{(add)bj}^{H}}$: invariant mass of a pair of one b-jet and one jet (that do not overlap with any top-tagged large-R jet) that is closest to the mass of the Higgs boson ($m_H \equiv 125$ GeV);
- $\boldsymbol{m_{(add)jj}^{H}}$: invariant mass of a pair of jets (that do not overlap with any top-tagged large-R jet) that is closest to the mass of the Higgs boson ($m_H \equiv 125$ GeV);
- $\boldsymbol{m_{H,bb}^{reco}}$: invariant mass of a pair of closest (minimum ΔR) b-jets in the event;
- $\boldsymbol{N_{j}^{40}}$: number of jets with $p_{\mathrm{T}} > 40$ GeV in the events;
- $\boldsymbol{N_{(add)j}^{40}}$: number of jets with $p_{\mathrm{T}} > 40$ GeV that do not overlap with any top-tagged large-R jet in the events;
- $\boldsymbol{H_{T}^{jet}}$: scalar sum of the p_{T} of all the jets in the event;
- $\boldsymbol{H_{T}^{(add)jet}}$: scalar sum of the p_{T} of all top-tagged large-R jets and all jets that do not overlap with the top-tagged large-R jets in the event.

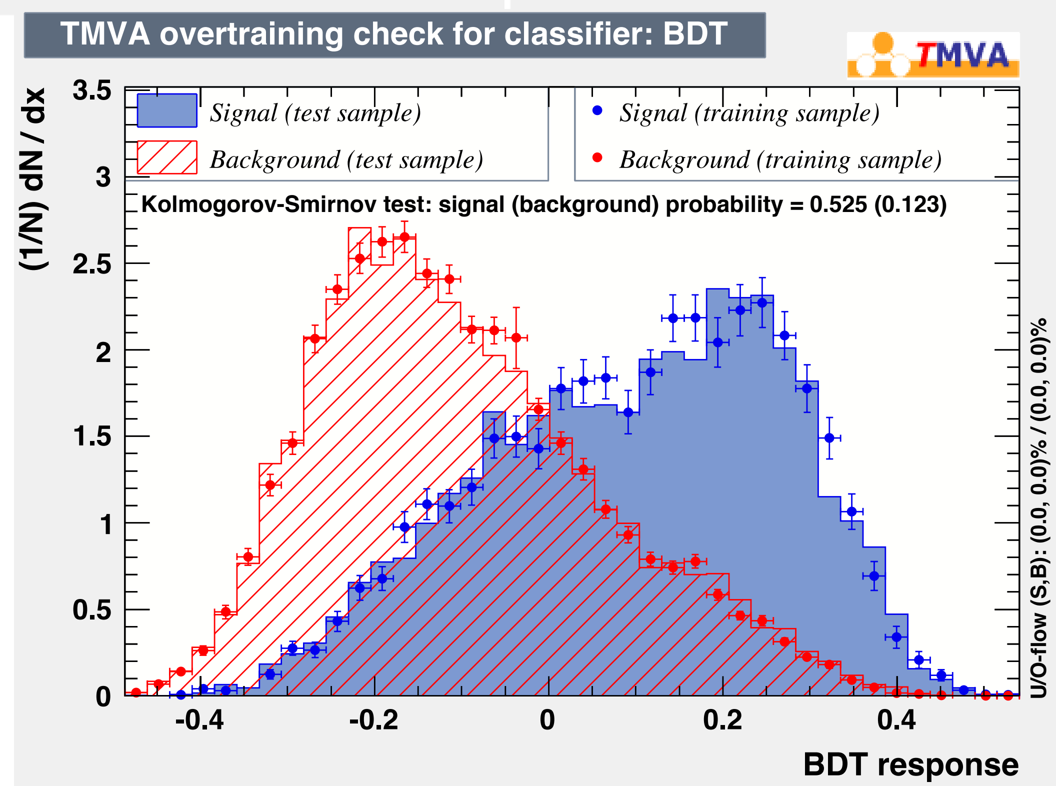
MVA training: results

Ranking	Variable	Separation power
1	ΔR_{bb}^{min}	0.0906
2	ΔR_{bb}^{avg}	0.0615
3	m_{top}^{lead}	0.0610
4	N_{j}^{40}	0.0483
5	m_{bb}^{max}	0.0476
6	$\Delta R_{(add)bb}^{min}$	0.0456
7	$m_{(add)bb}^{H}$	0.0362
8	H_{T}^{jet}	0.0311
9	$\tau_{32,top}^{lead}$	0.0261
10	$\Delta R_{top(add)b}^{avg}$	0.0180

Ranking	Variable	Importance
1	ΔR_{bb}^{min}	0.1335
2	ΔR_{bb}^{avg}	0.1257
3	m_{top}^{lead}	0.1105
4	$\Delta R_{top(add)b}^{avg}$	0.1103
5	$m_{(add)bb}^{H}$	0.1011
6	$\Delta R_{(add)bb}^{min}$	0.0947
7	N_{j}^{40}	0.0867
8	m_{bb}^{max}	0.0808
9	$\tau_{32,top}^{lead}$	0.0794
10	H_{T}^{jet}	0.0775

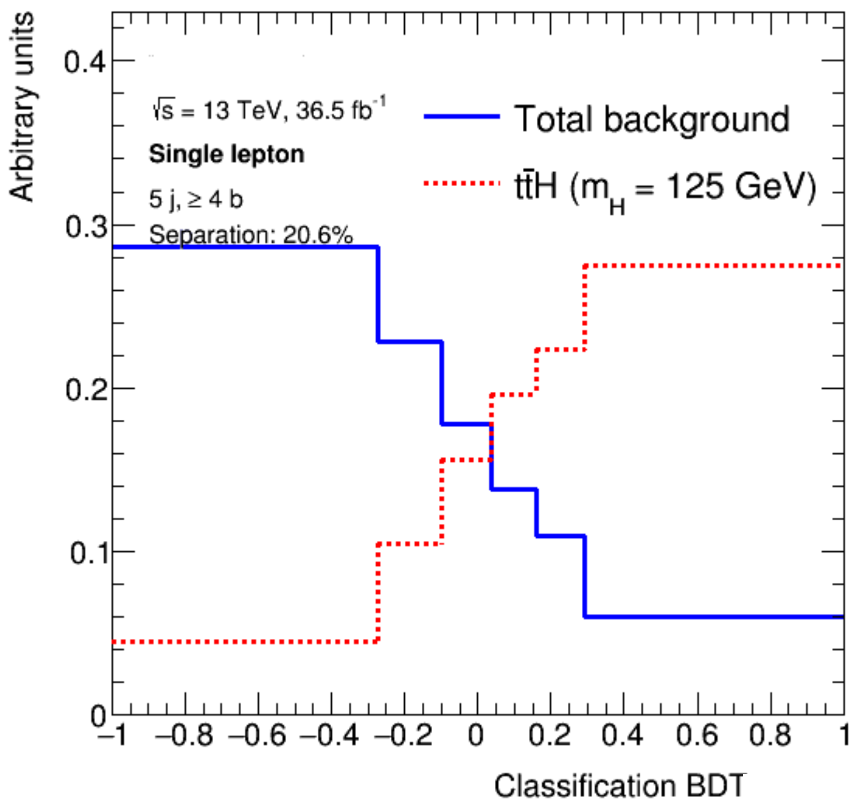


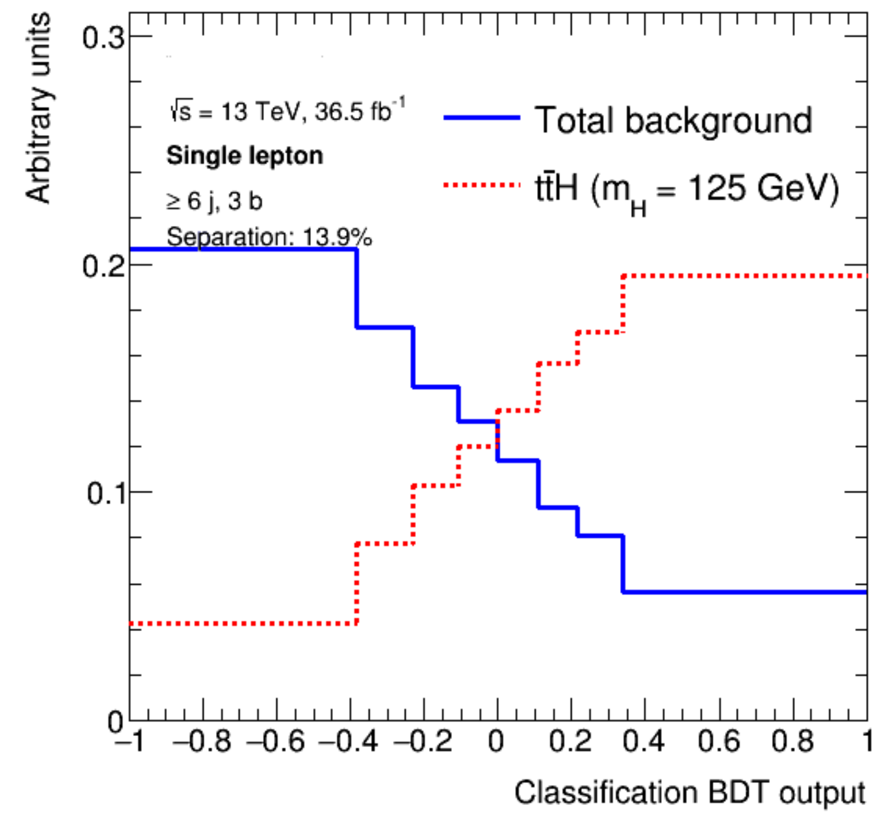


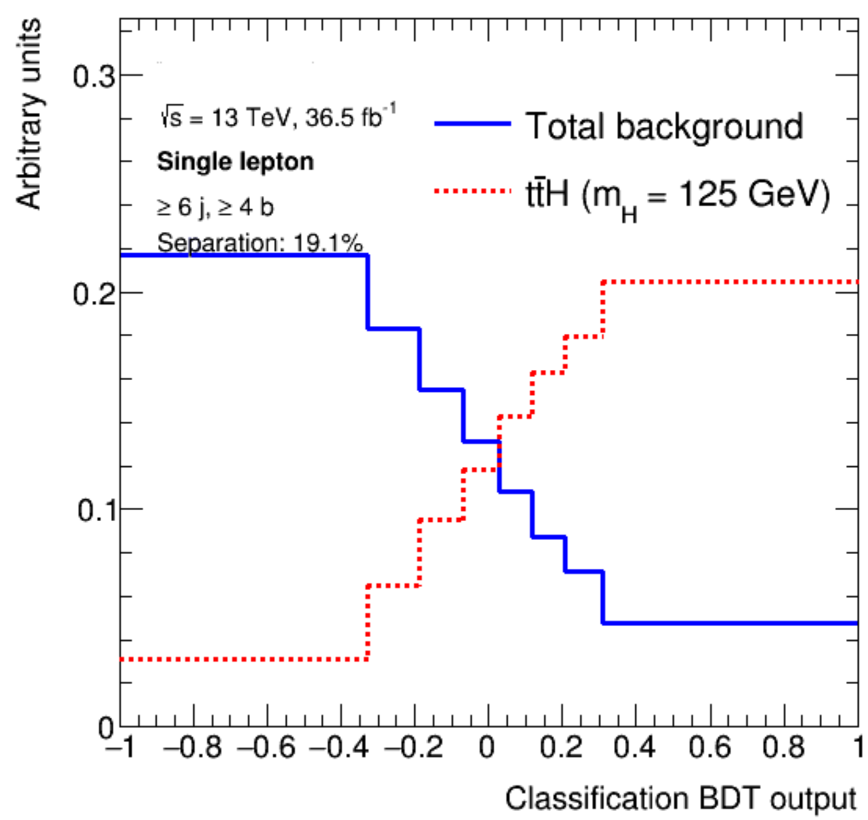


Resolved-boosted BDT outputs

Combined







Likelihood and test statistic

- **Likelihood function**:
  - defined as the product of Poisson probabilities for all bins:

$$s_i = s_{tot} \int_{\text{bin } i} f_s(x; \boldsymbol{\theta}_s) dx$$

$$b_i = b_{tot} \int_{\text{bin } i} f_b(x; \boldsymbol{\theta}_b) dx$$

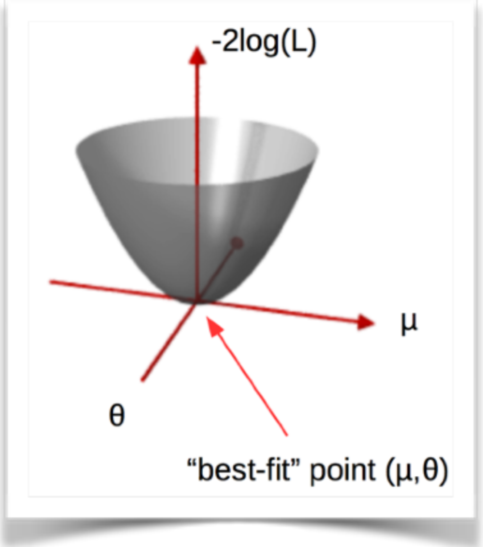
$$\mathcal{L}(\mu, \boldsymbol{\theta}) = \prod_{j=1}^{N} \frac{(\mu s_j + b_j)^{n_j}}{n_j!} e^{-(\mu s_j + b_j)} \prod_{k=1}^{M} \frac{u_k^{m_k}}{m_k!} e^{-u_k}$$

  - To test a hypothesized value of μ, the **profile likelihood ratio** is considered: $\lambda(\mu) = \dfrac{\mathcal{L}(\mu, \hat{\hat{\boldsymbol{\theta}}})}{\mathcal{L}(\hat{\mu}, \hat{\boldsymbol{\theta}})}$

- **Test statistic q_μ**:
  - for the purpose of establishing an upper limit on the strength parameter μ, it is defined as

$$q_\mu = \begin{cases} 0, & \mu < \hat{\mu} \\ -2\ln\lambda(\mu), & \mu \geq \hat{\mu} \end{cases}$$

  - Higher values of q_μ represent greater incompatibility between the data and the hypothesized value of μ.



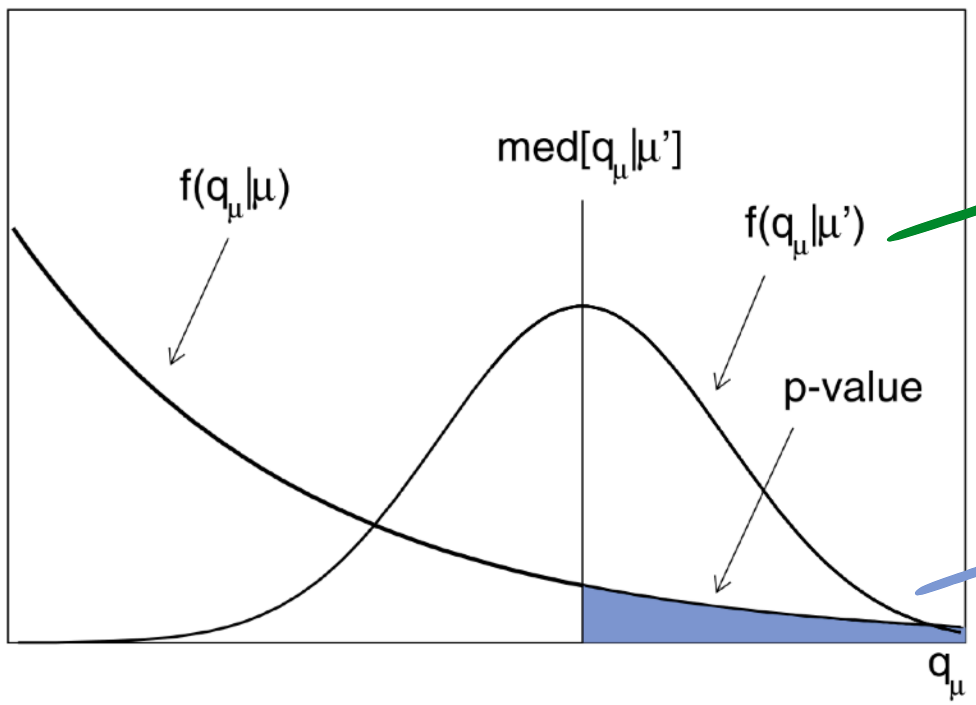
- **P-value**:
  - quantifies the level of agreement between the data and the hypothesized μ.

$$p_\mu = \int_{q_{\mu,obs}}^{\infty} f(q_\mu|\mu) dq_\mu$$

  - the value of μ, for which the median p-value is equal to 0.05, gives the median upper limit on μ at 95% confidence level.

Likelihood and test statistic

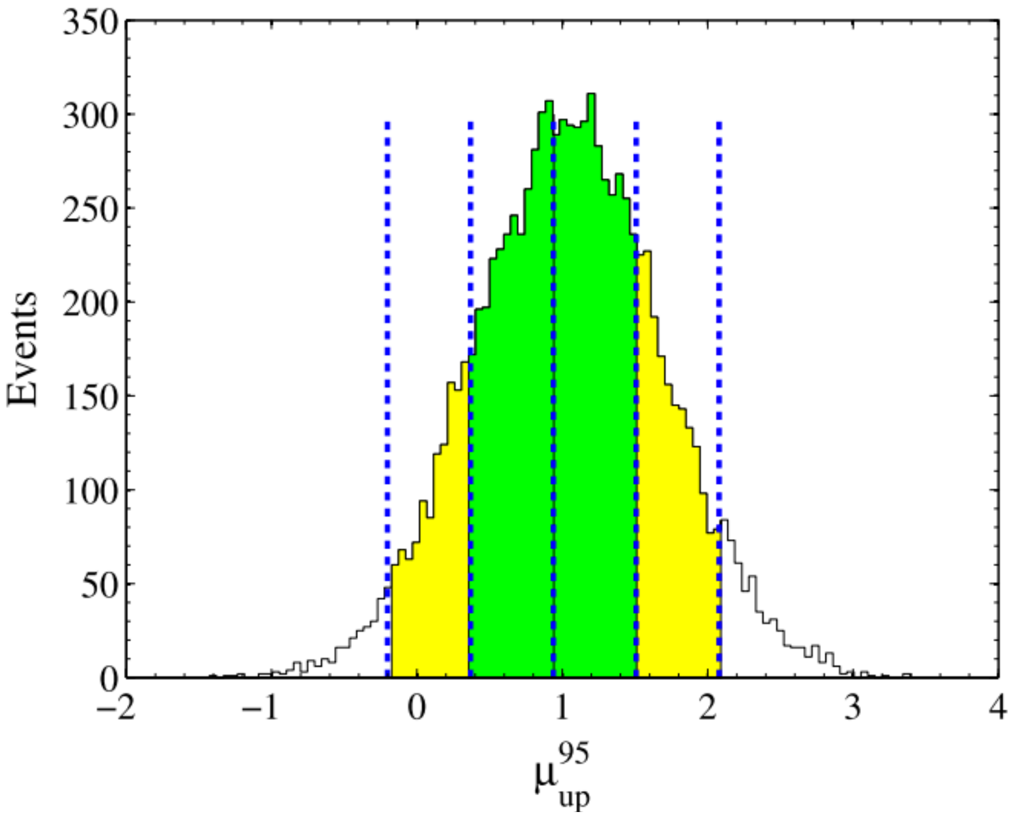
pdf for q_μ assuming both a strength parameter μ and a different value μ'



f(q_μ | μ′) is shifted to **higher value of q_μ**, corresponding on average to **lower p-values**

The **sensitivity of an experiment** can be characterised by giving the **p-value corresponding to the median q_μ assuming the alternative value μ'**

The upper limit on μ at a confidence level CL=1–α is the value of μ for which the p-value is p_μ =α.



By simulating the experiment many times with Monte Carlo, it is possible to obtain a histogram of the upper limits on μ at 95% CL.

The ±1σ and ±2σ error bands are obtained from the MC pseudo-experiments.

The vertical lines indicate the error bands as estimated directly without Monte Carlo simulation.

CL_s method

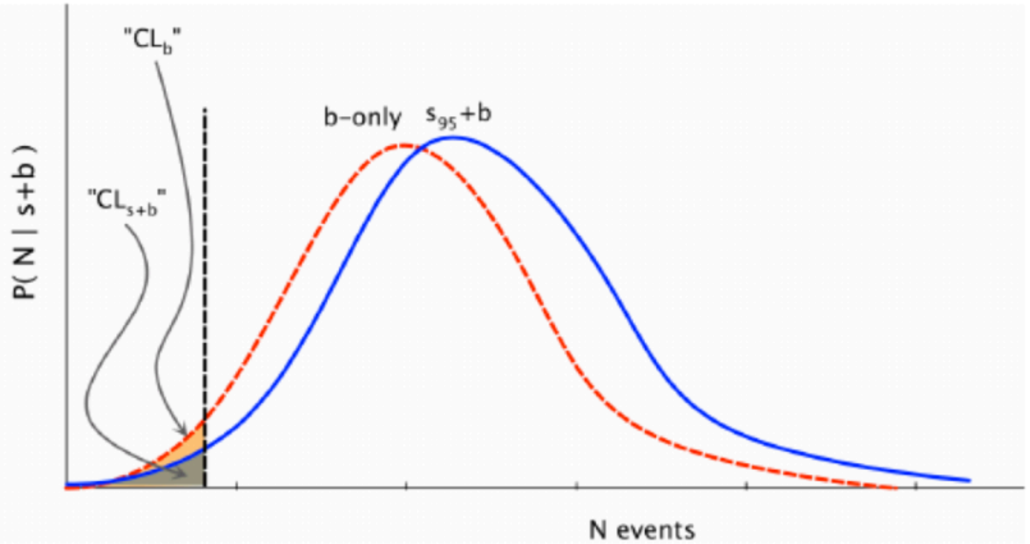
- **Modified Frequentist CL_s method**:
  - confidence level for **excluding the possibility of signal on top of background** (the s+b hypothesis):

$$\alpha_{\mathrm{s+b}} = P_{\mathrm{s+b}}(q_\mu \leq q_{\mu,obs})$$

  - **probability**, assuming the presence of both signal and background at their hypothesized levels, **that the test statistic would be less than or equal to that observed in the data**;
  - confidence level for the background alone:

$$\alpha_{\mathrm{b}} = P_{\mathrm{b}}(q_\mu \leq q_{\mu,obs})$$

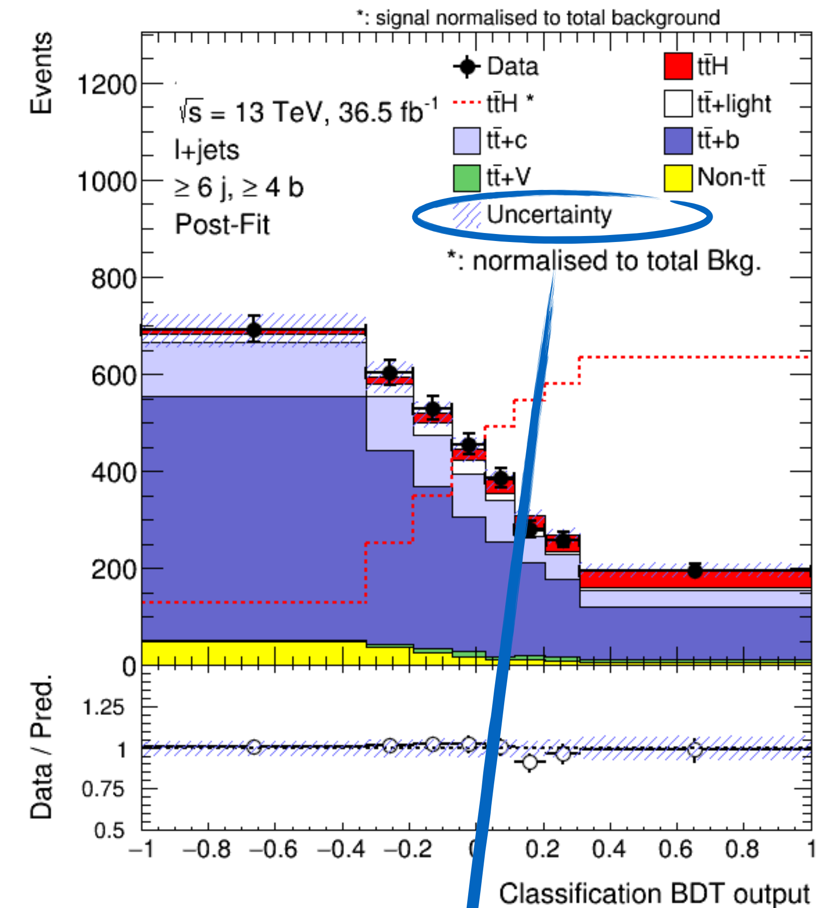
  - probability assumes the presence of the background only. This confidence level has been suggested to quantify the confidence of a potential discovery, as it expresses the **probability that background processes would give a number of events smaller than or equal to the number of observed candidates**.



$$CL_{\mathrm{s}} = \alpha_{\mathrm{s+b}} / \alpha_{\mathrm{b}}$$

Systematic uncertainties

- Signal and background predictions are **affected by systematic uncertainties**;
- inside the fit procedure, **described by parameters** for which the true values are unknown.



- **Detector systematics uncertainties**;
- **signal theory uncertainties** (QCD scale, PDF) **on inclusive cross sections**:
  - correlated between experiments, uncorrelated between processes;
- **signal theory uncertainties on acceptance and selection efficiency**:
  - uncorrelated between experiments (different method to estimate them);
- **PDF uncertainties on signal cross sections**:
  - correlated for a given process across experiments;
- **background theory uncertainties**:
  - treated differently by the experiments.

Fit to Asimov data set

- The **sensitivity of an experiment** is characterised in the **expected median significance to reject different values of μ**, from a chosen one (**μ = 0**).
- **Asimov data set:**
  - **replaces the ensemble of real data**;
  - helps to **understand any background mis-modelling and problems with the systematic uncertainties**;
  - provides **important information on the final uncertainty and on the sensitivity of the analysis.**

Ingredients

- **all the same control and signal regions;**
- **same MVA discriminants**;
- **upper limit with onlyB hypothesis (μ = 0);**
- **fit** of parameters **with S+B hypothesis (μ = 1):**
  - **NPs** and **μ**;
  - **normalisation factors for tt+≥1b and tt+≥1c free-floating**, because of the lack of precision in the tt background simulation.

Upper limit @ 95% CL — μ = 0

	Expected ($\mu = 0$)		
	Median	$\pm 1\sigma$	$\pm 2\sigma$
Resolved-only*	0.83	[0.60,1.18]	[0.45,1.64]
Combined	0.83	[0.60,1.17]	[0.44,1.63]

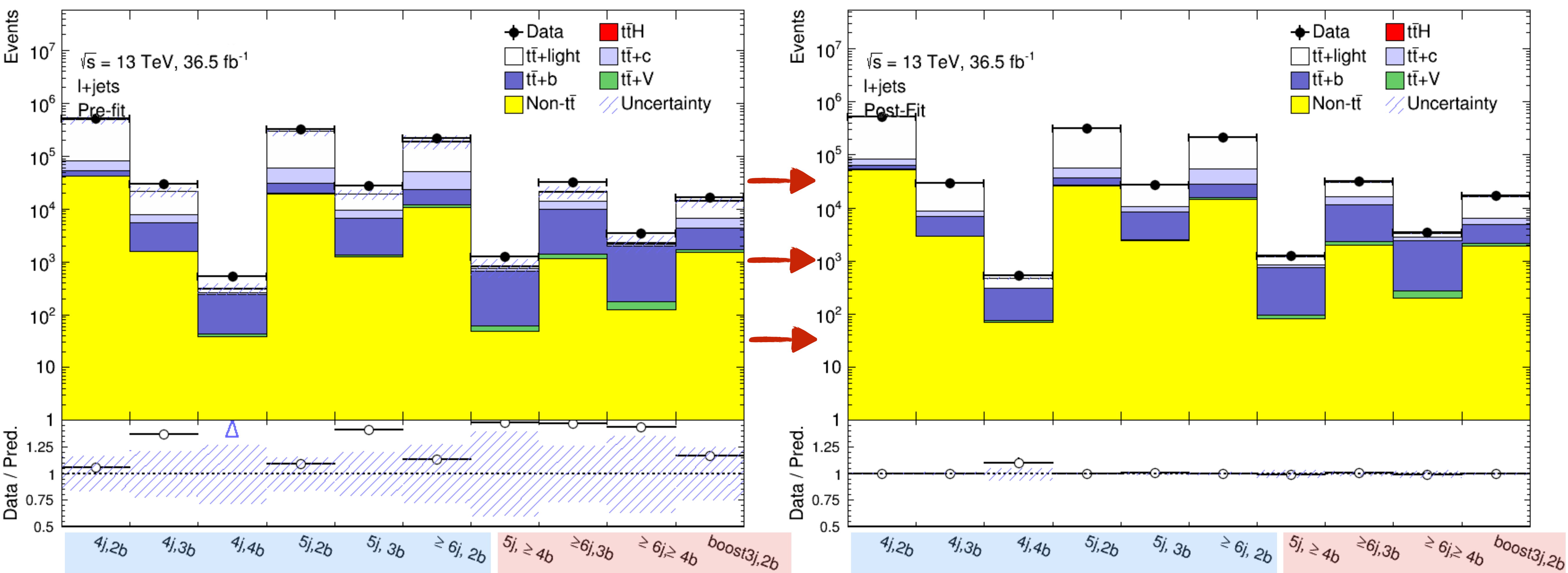
Best-fit values — μ = 1

	Resolved-only	Combined
μ_{ttH}	1.00 ± 0.44	1.00 ± 0.44
k(tt+b)	1.00 ± 0.14	1.00 ± 0.13
k(tt+c)	1.00 ± 0.02	1.00 +0.02

*resolved l+jets channel only

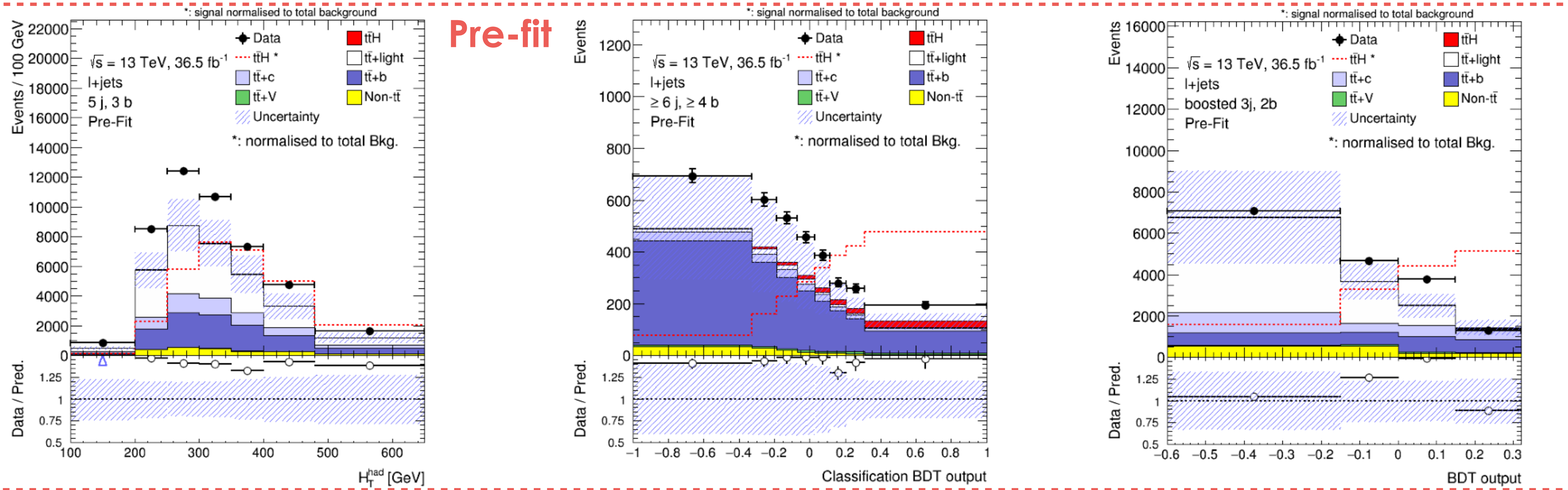
Fit to real data: toward the results

- **Fit effect on distributions and systematic uncertainties:**
  - significant improvement in data/MC agreement from the pre-fit plot to the post-fit one;
  - estimated number of events is in agreement with the number of data in all the regions;
  - the systematics bands are reduced significantly with the fit.

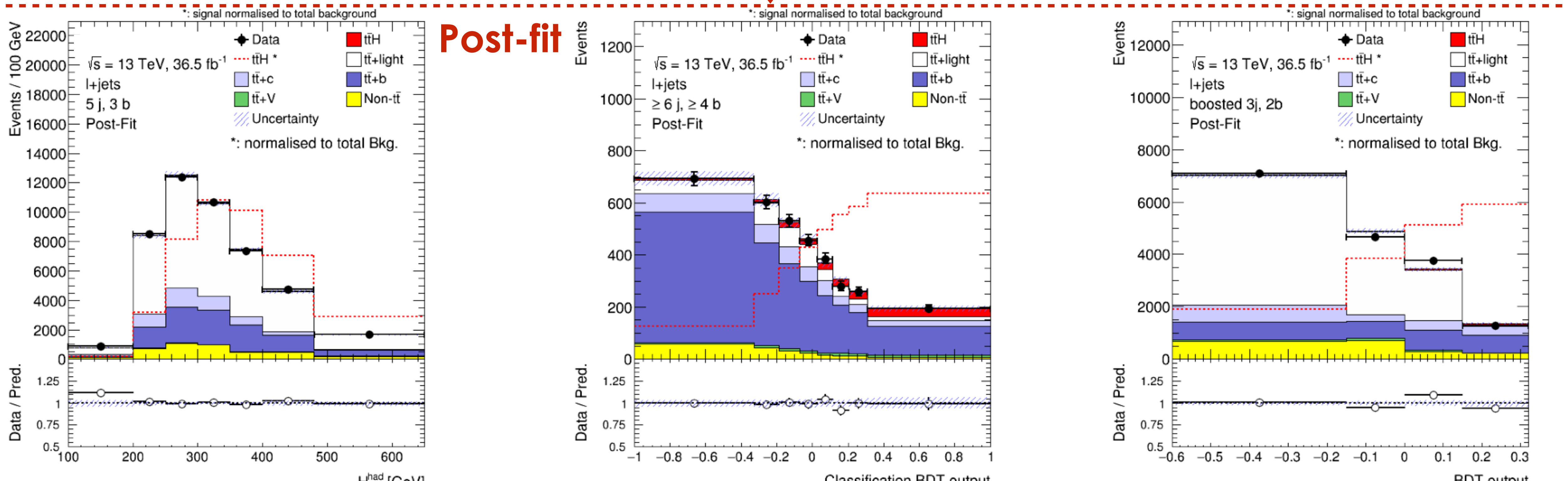


Fit to real data: toward the results

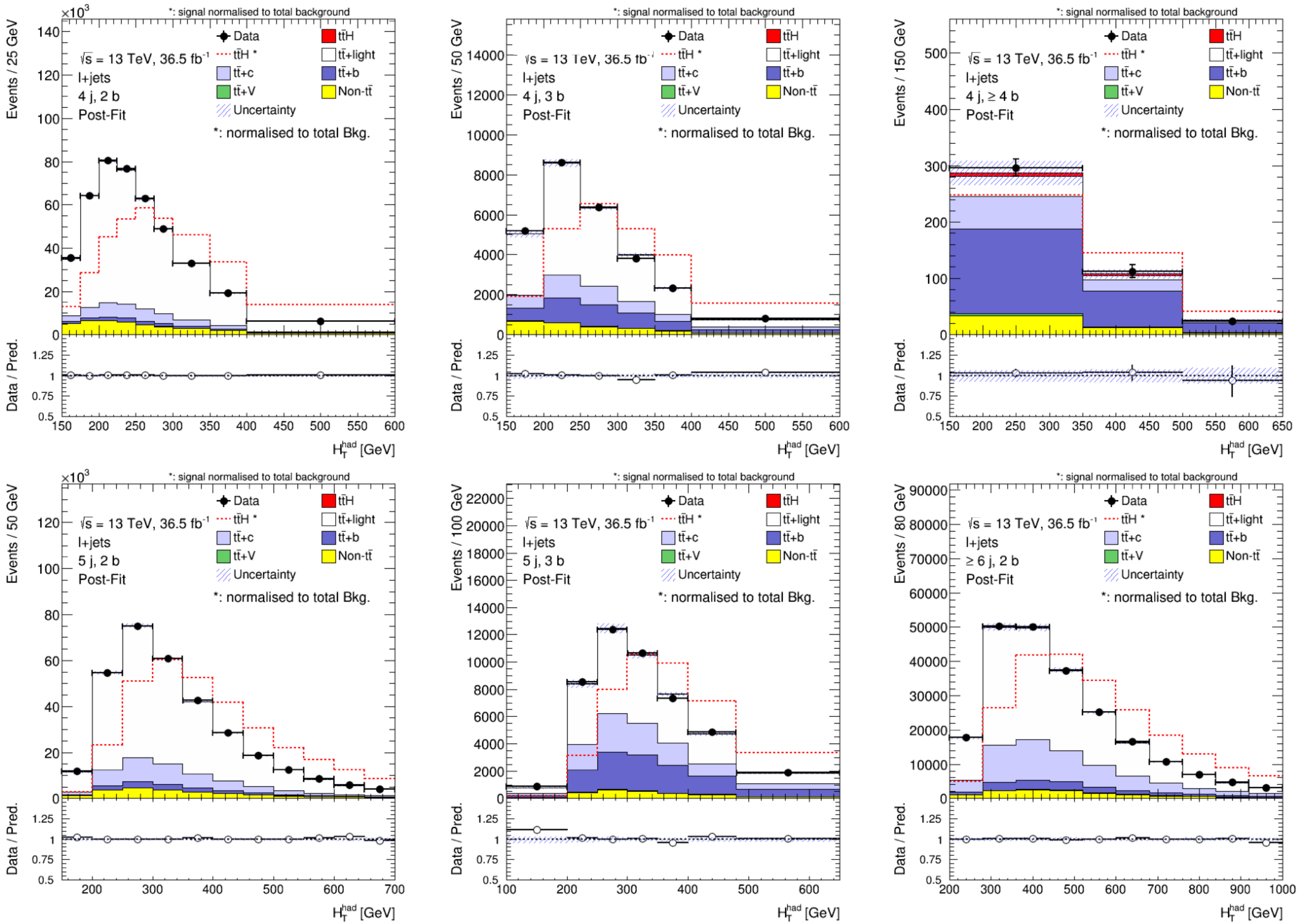
Pre-fit



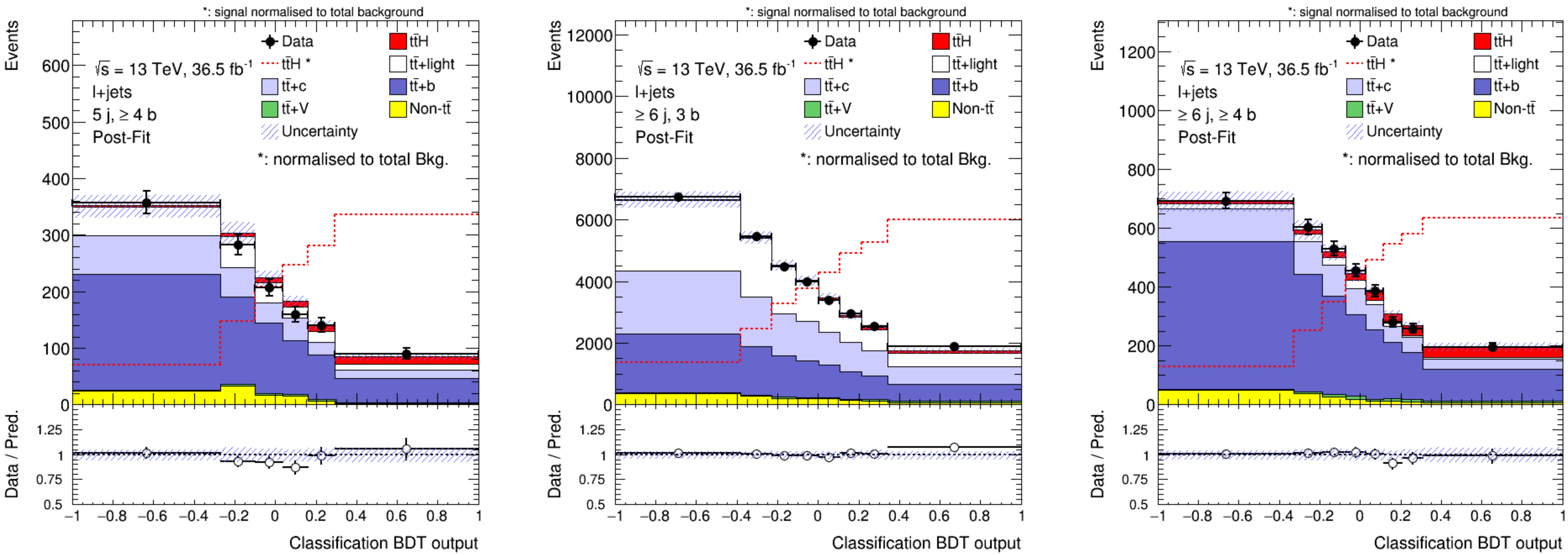
Post-fit



Resolved analysis: H_T



Resolved analysis: BDT

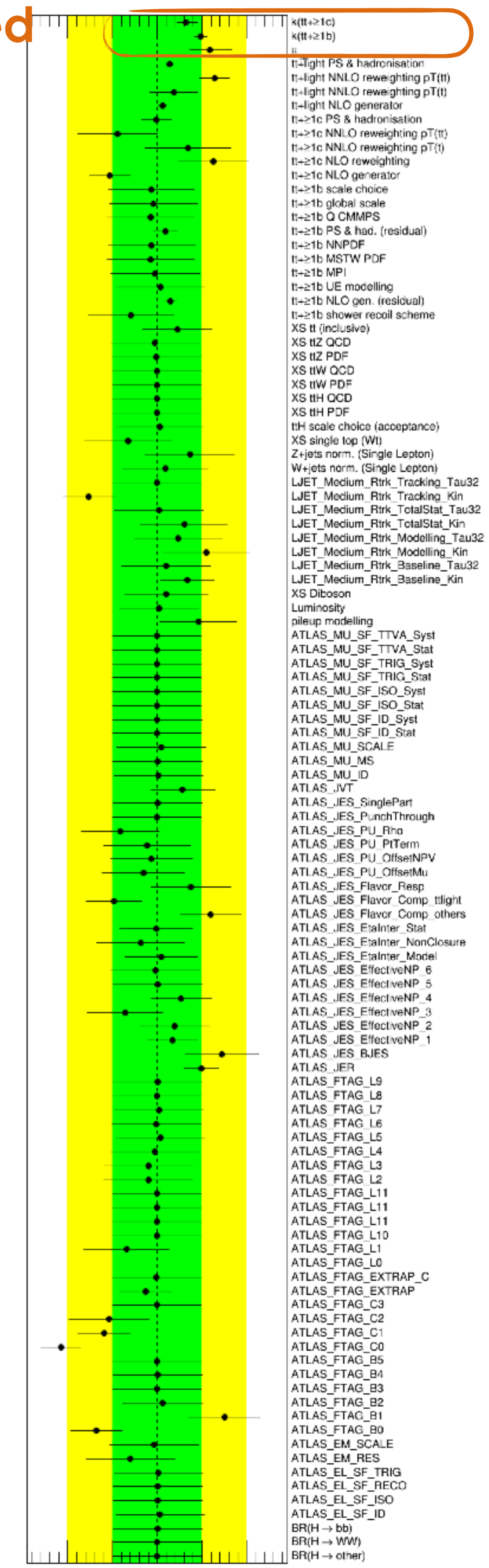


Pull plots: explanation and comparison

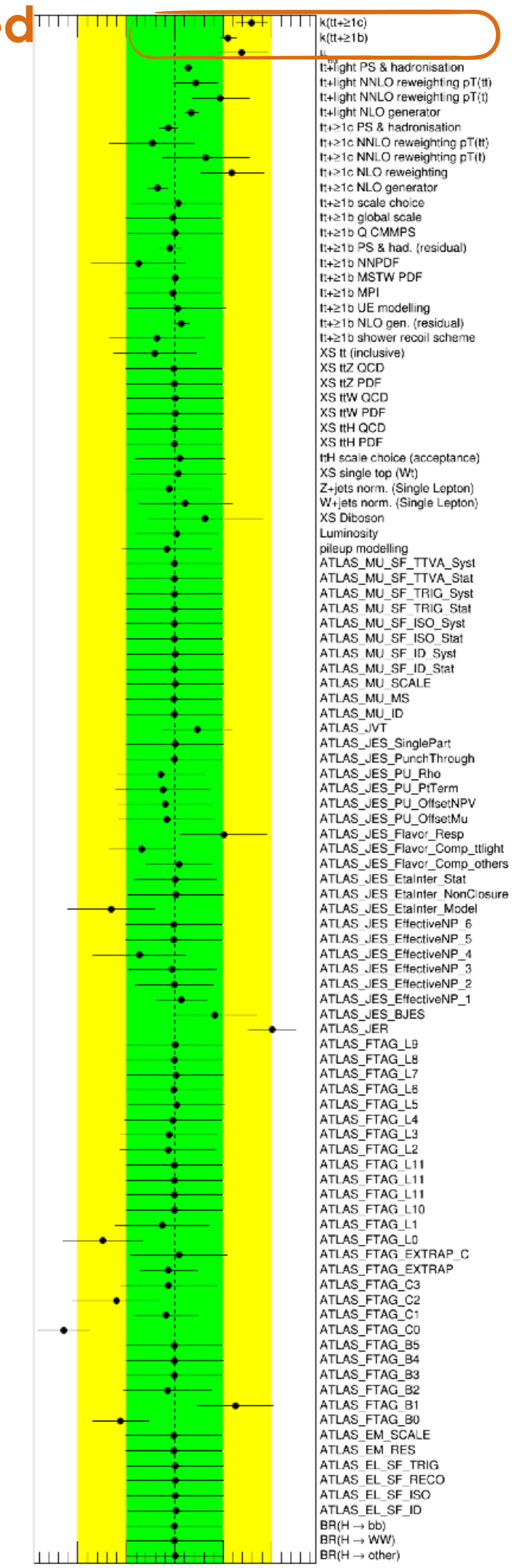
- The dots are the **nuisance parameters with their uncertainties**, normalised to the expected position (0 or 1 depending on the specific parameter);
- an eventual shift of the best-fit value (**"pulled" parameter**) should be due to a compensation operated by the fit for some data/MC disagreements;
- **"constrained" parameters** show a lower uncertainty (wrt 1, as expected in the Asimov test), evidencing a too large variation of that systematic wrt the statistical power of data.

- Comparison:
  - **constrained:** tt+jets PS&had and NLO generator, JER and top (tt) p_T reweighing related NPs;
  - **pulled:** several NPs from tt+jets modelling, flavour tagging and jets related systematics.

Combined

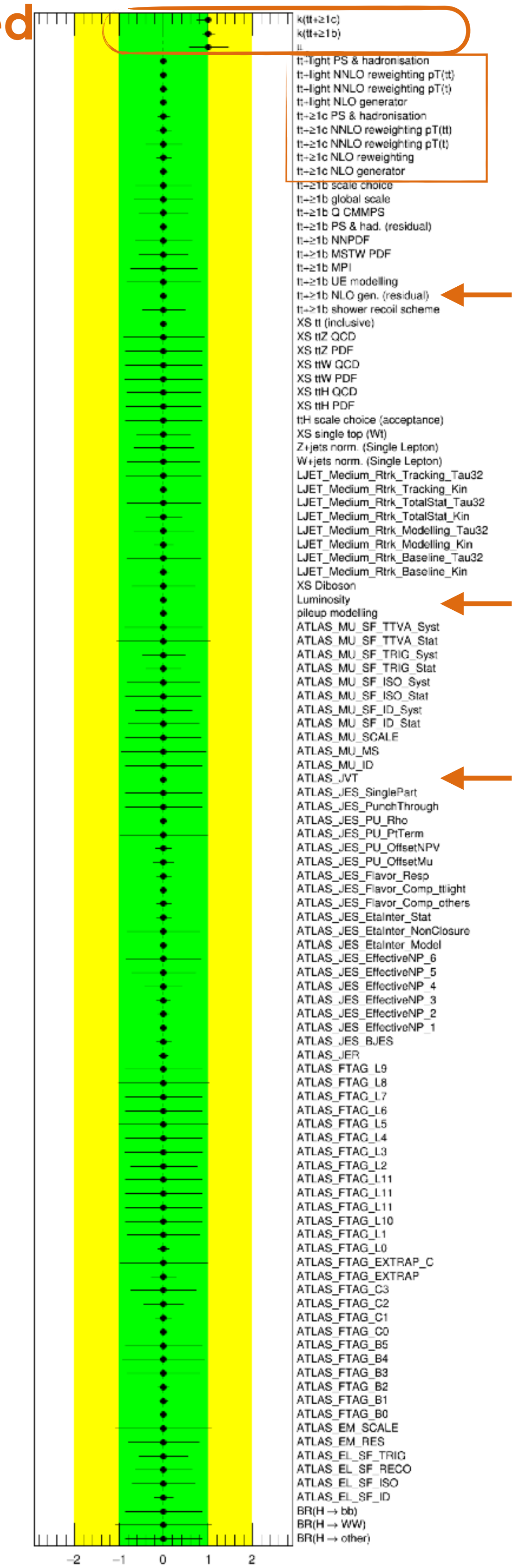


Resolved

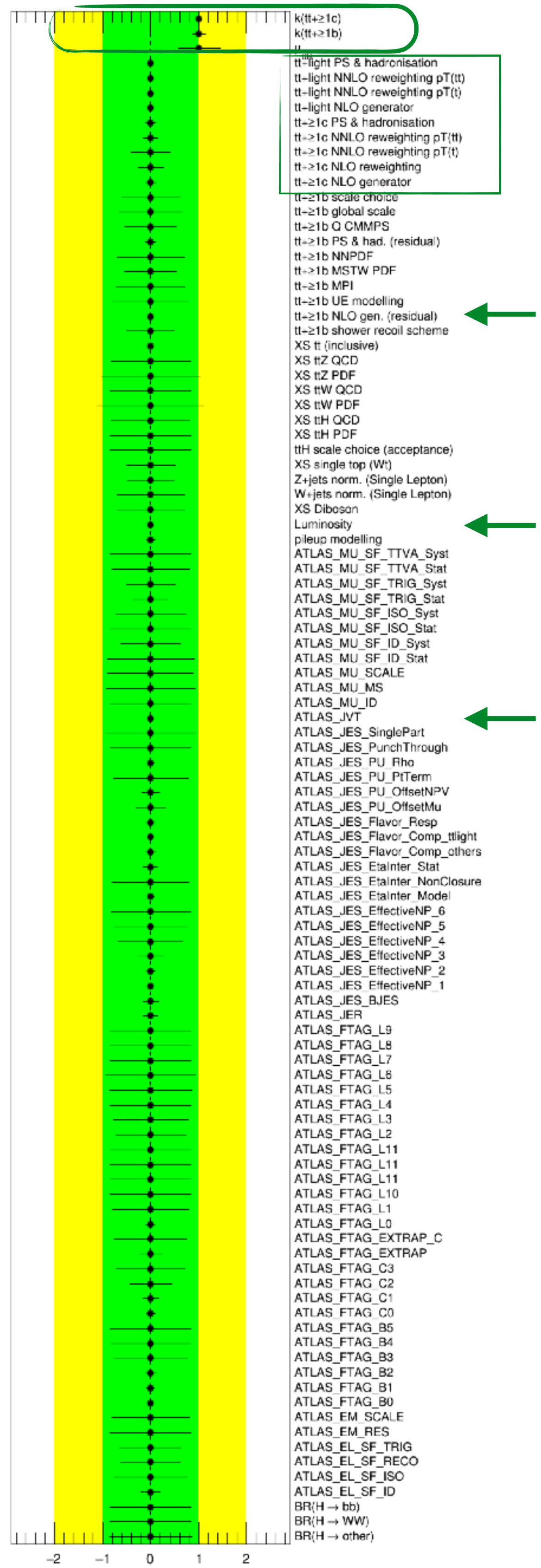


Pull plots: Asimov data set

Combined



Resolved



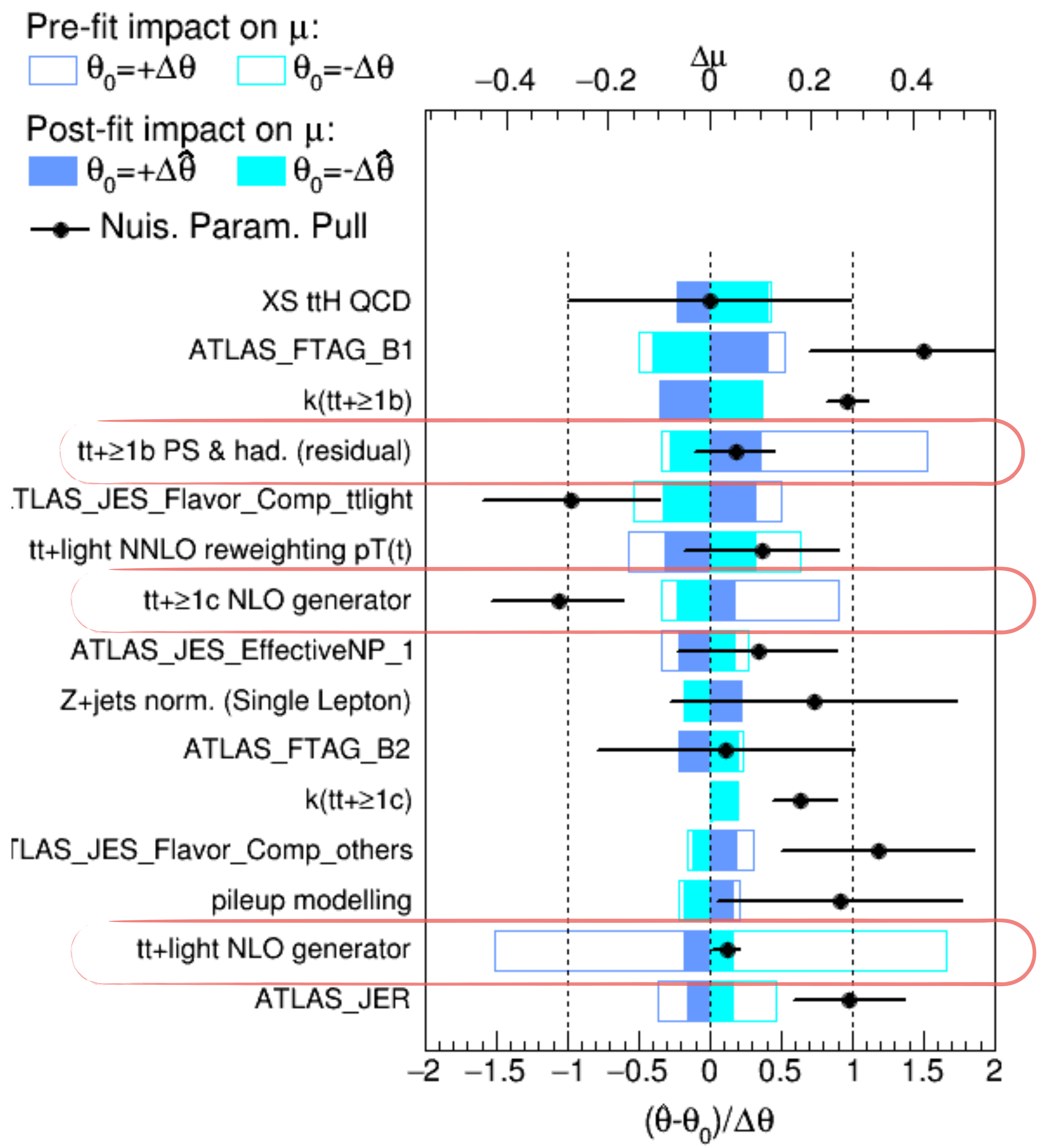
Some common features:

- luminosity, pu modelling, JVT, **tt+light PS&had**, **tt NLO generator** very constrained;
- **JES flavour Composition** split in tt+light & other samples: this has a significant effect on the fit;
- flavour tagging components not constrained very much;
- k factors and μ with a **compatible uncertainty between combined and resolved analyses**.

Fit to real data: Ranking plots

- **Systematic uncertainties and their effect on μ:**
  - tt+jets modelling has a very large effect on the μ best-fit;
  - flavour tagging, x-sec and pile-up modelling systs are quite important.

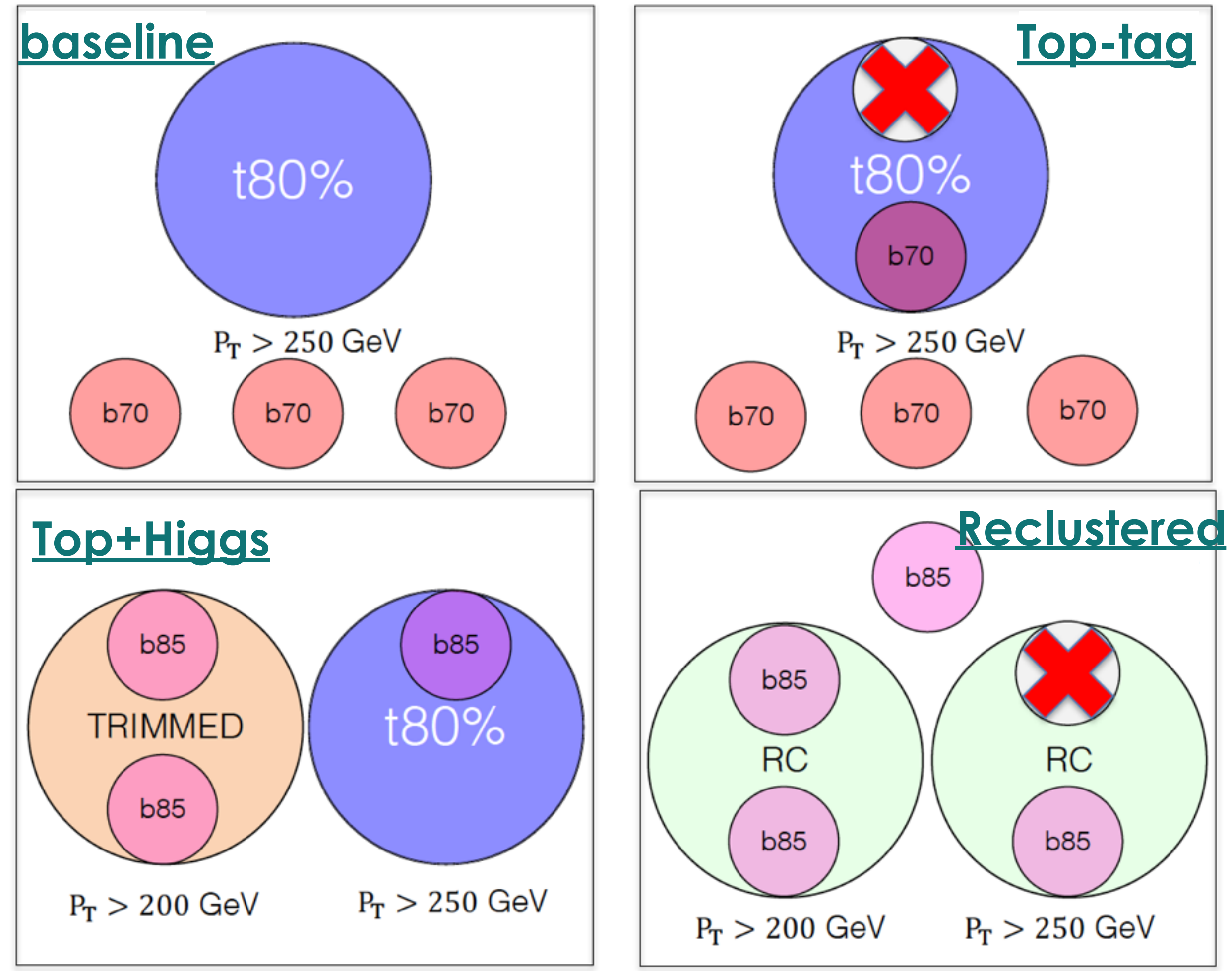
- **Importance ranking** of the first 15 NPs according to the impact on μ;
- **empty bands:** impact of the systematic uncertainty on μ (Δμ) before the fit;
- **full bands:** impact of the systematic uncertainty on μ (Δμ) after the fit;
- **black dots:** pull of the NPs, as explained for the pull plots.



Choosing the final SR

Looked at different factors to make a decision:

- **significance:** very small difference among the candidates, to be considered together with the statistics;
- **overlap with resolved (ICHEP):** T+H and Recl have the lowest overlap;
- **BDT training and separation:** Recl has the highest S/B separation;
- **Stat-only and full systematics (ICHEP) limits:** both boosted only and combined fit with the ICHEP analysis strategy for the resolved;
- **Systematics model:** the full set of uncertainties to be used has been considered for each option (Recl does not need the large-R jets syst).

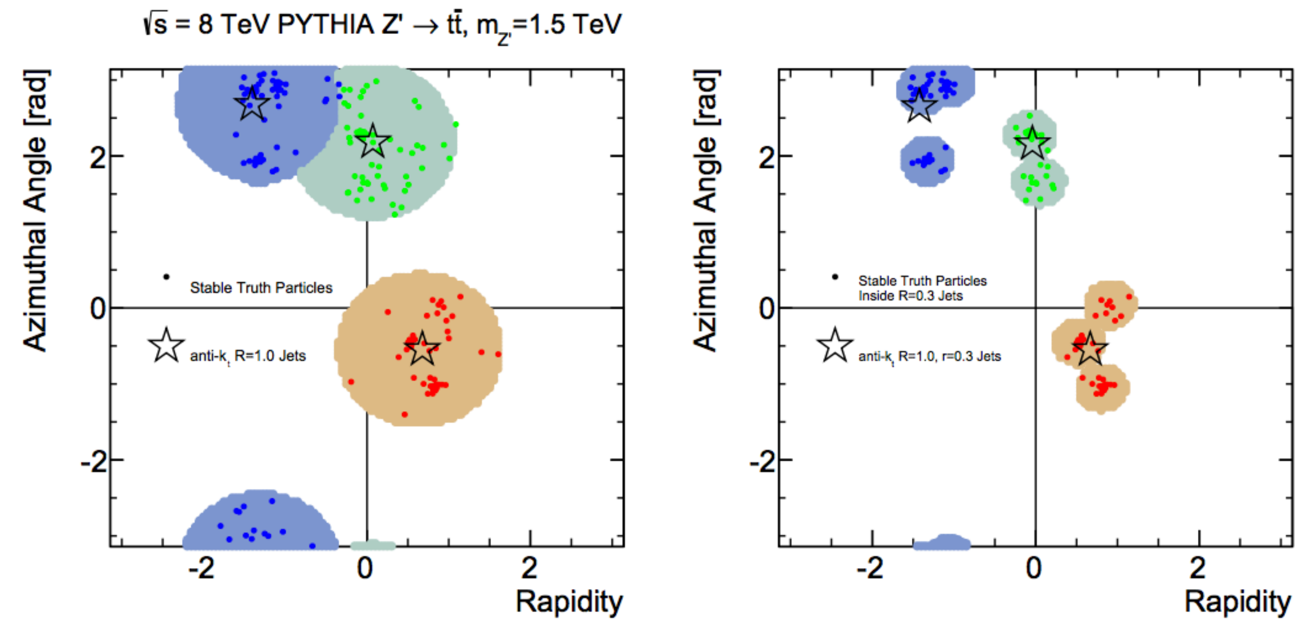


numbers from HTop Workshop (March2017)

	Baseline	Top-tag	Top+Higgs	Reclust
S/B	3%	7%	3%	4%
S/√B	1.0	1.1	0.8	0.8
BDT Separation	12%	22%	23%	24%
Boosted limit	$3.73^{+5.33}_{-2.69}$	$2.54^{+3.67}_{-1.86}$	$5.95^{+7.25}_{-4.28}$	$3.57^{+5.06}_{-2.57}$
Combined limit	$0.99^{+1.43}_{-0.71}$	$0.83^{+1.18}_{-0.60}$	$0.82^{+1.16}_{-0.59}$	$1.00^{+1.44}_{-0.72}$

→ **A compromise among all these factors has been reached in order to choose the final boosted SR**

Re-clustering techniques (I)



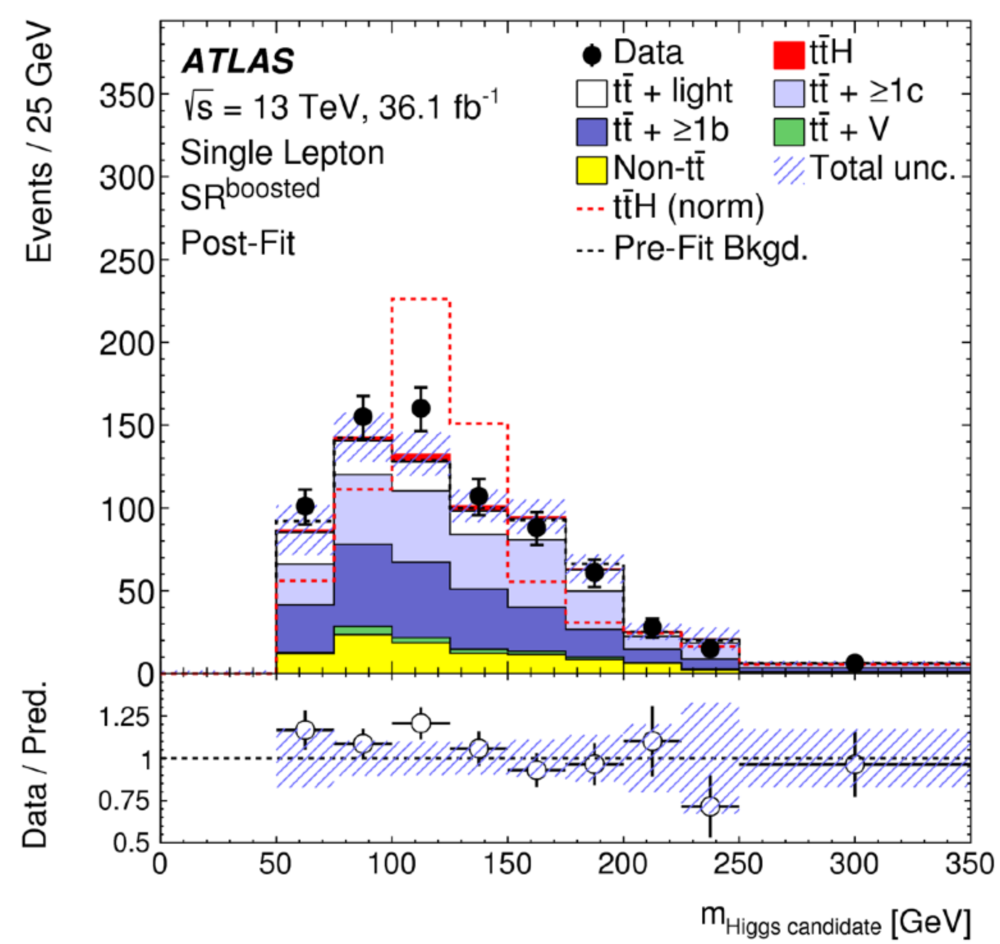
Problem

Traditionally only a few choices of radius parameter R are used for all analyses, because **every jet configuration must be calibrated** to account unmeasured energy deposits and other experimental effects.

Solution: re-clustering technique

- Allows a much broader class of algorithms and jet radius parameters to be selected by analyses;
- **no additional calibrations required** for different re-clustered jet radius;
- **anti-k_T small-R jets (R=0.4) used to re-cluster the large-R jets** (R = 1.0, 200 < p_T < 1500 GeV, $|\eta|$ < 2, m > 50 GeV) in this analysis.

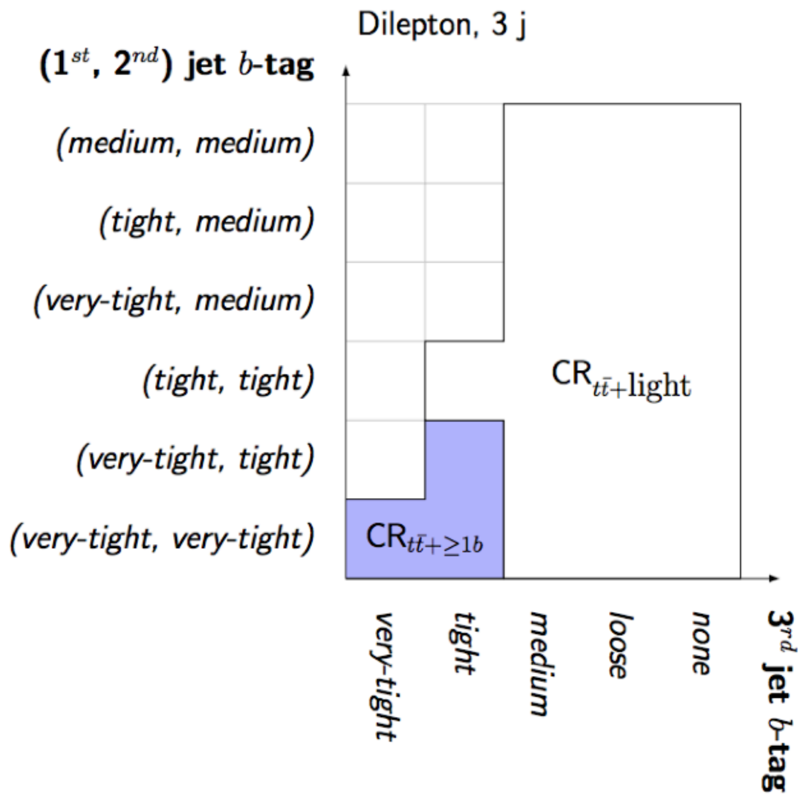
Higgs candidate mass distribution, reconstructed by re-clustering technique.

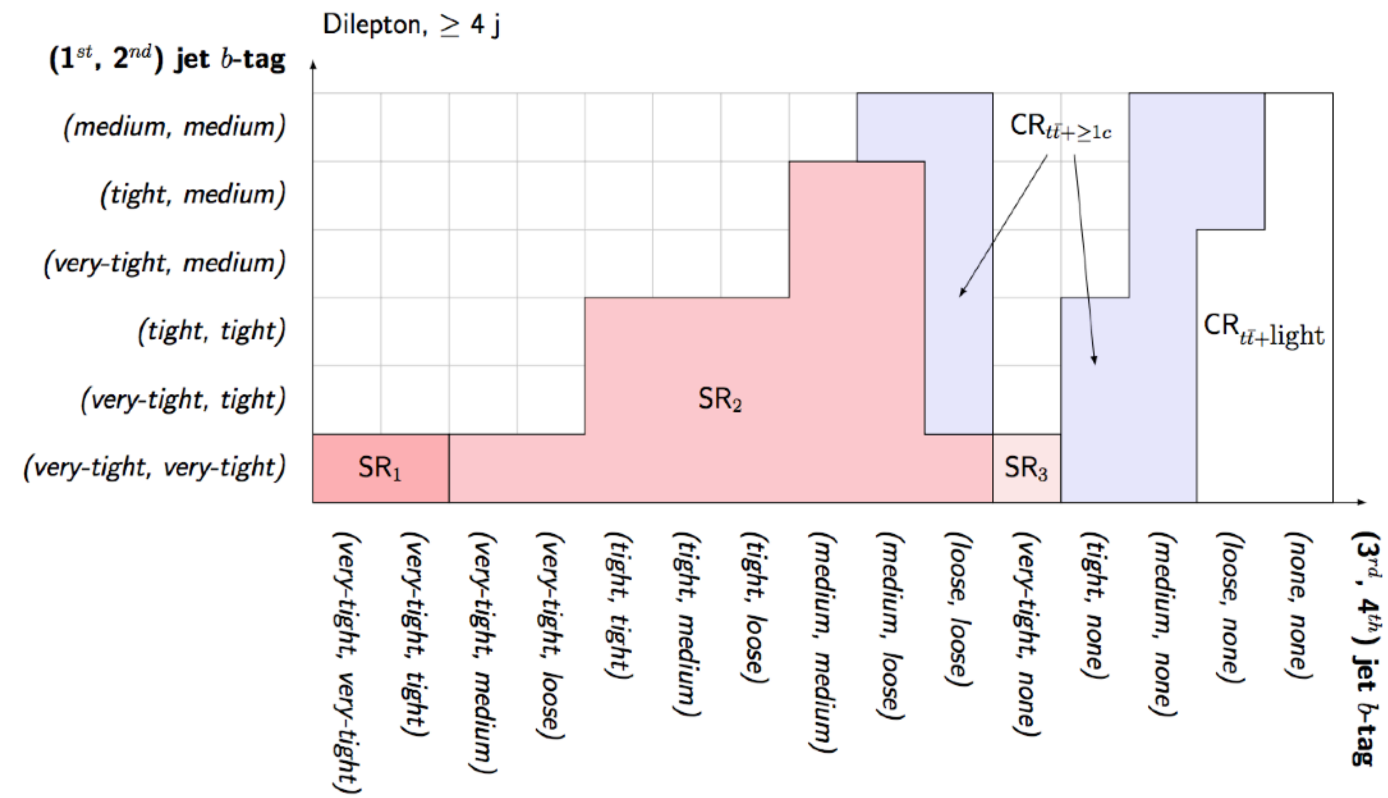


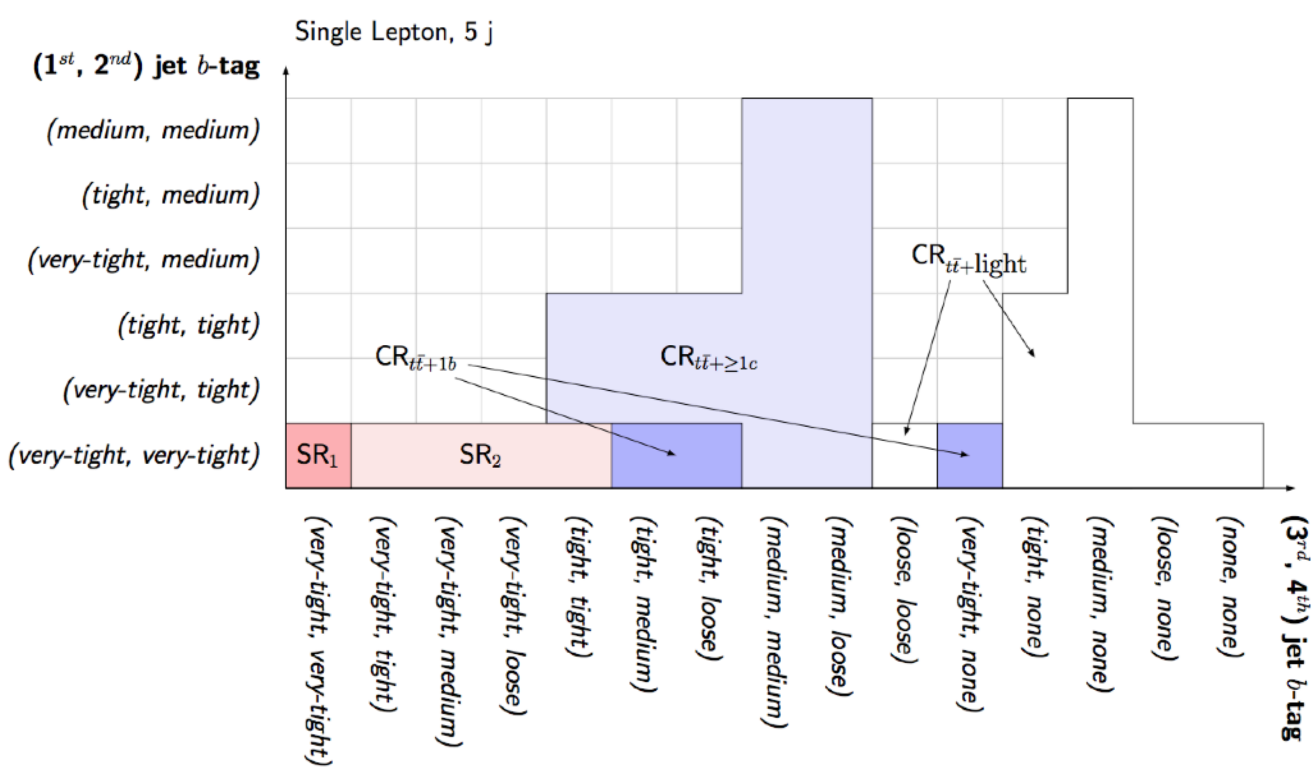
Resolved classification

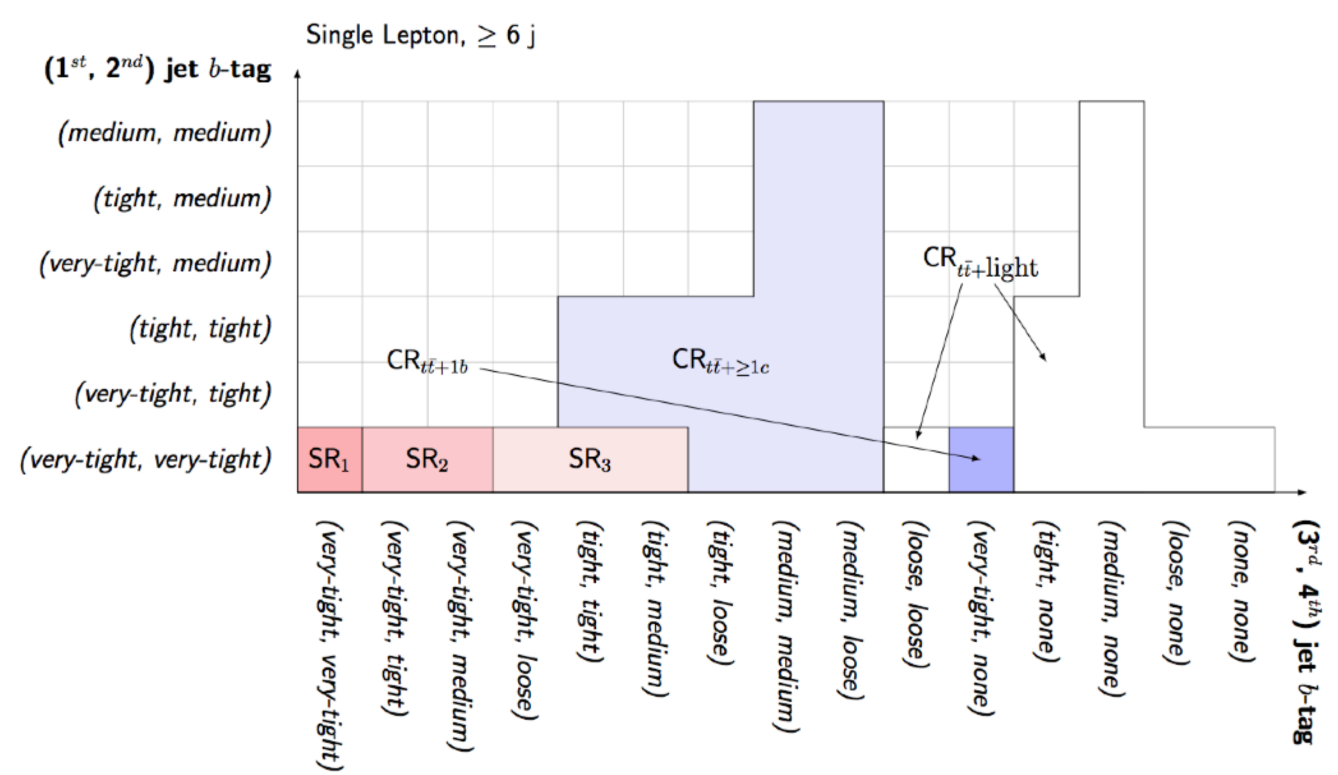
The jets are ordered according to their value of the b-tagging discriminant in descending order.

1 = no b-tagging criteria fulfilled
2 = loose = 85%
3 = medium = 77%
4 = tight = 70%
5 = very tight = 60%







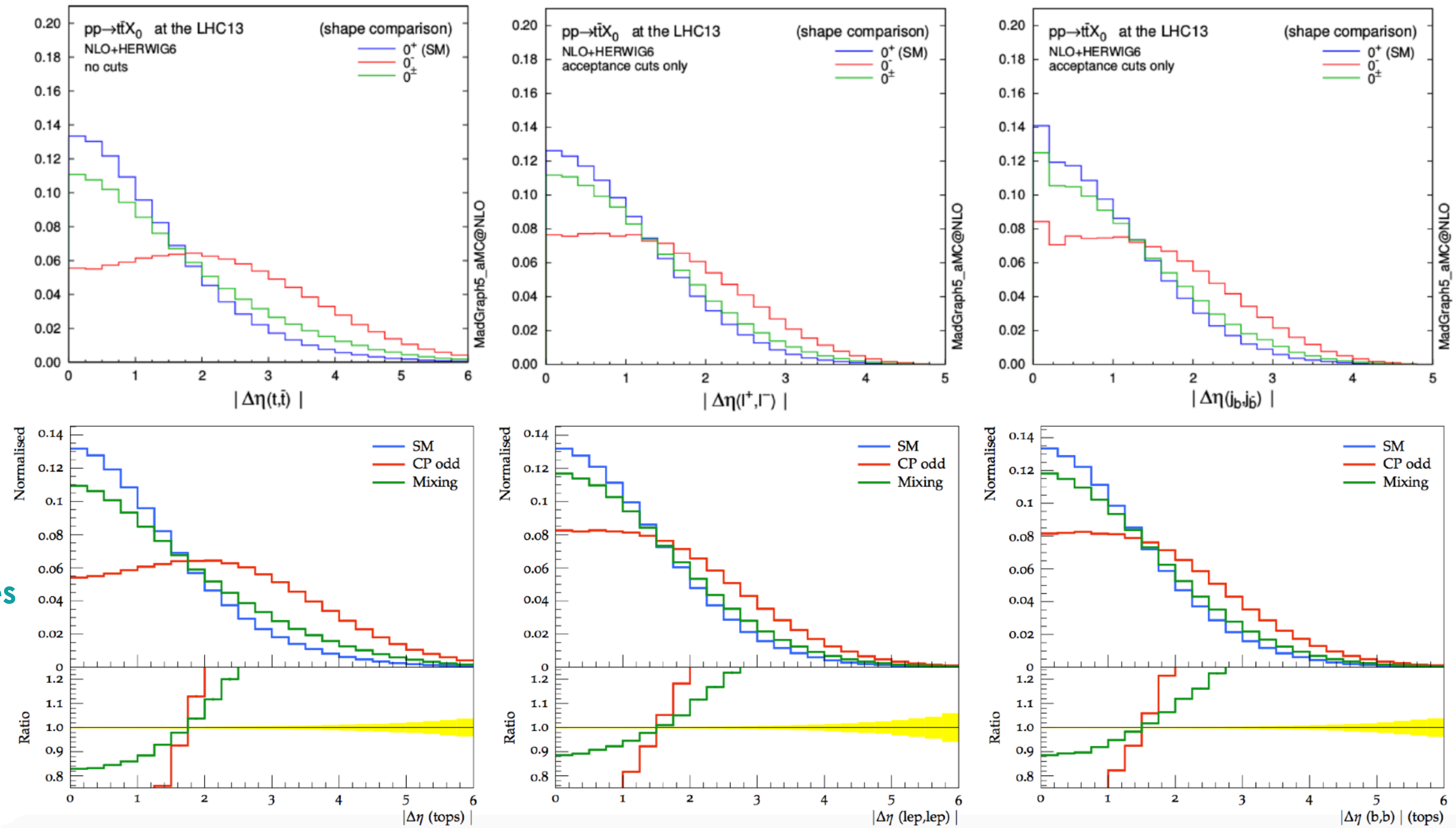


y-axis: values of the b-tagging discriminant for the first two jets, x-axis: values for the third jet or the third and fourth jets.

Waiting for the o… SPOILER! …n…

paper

our samples

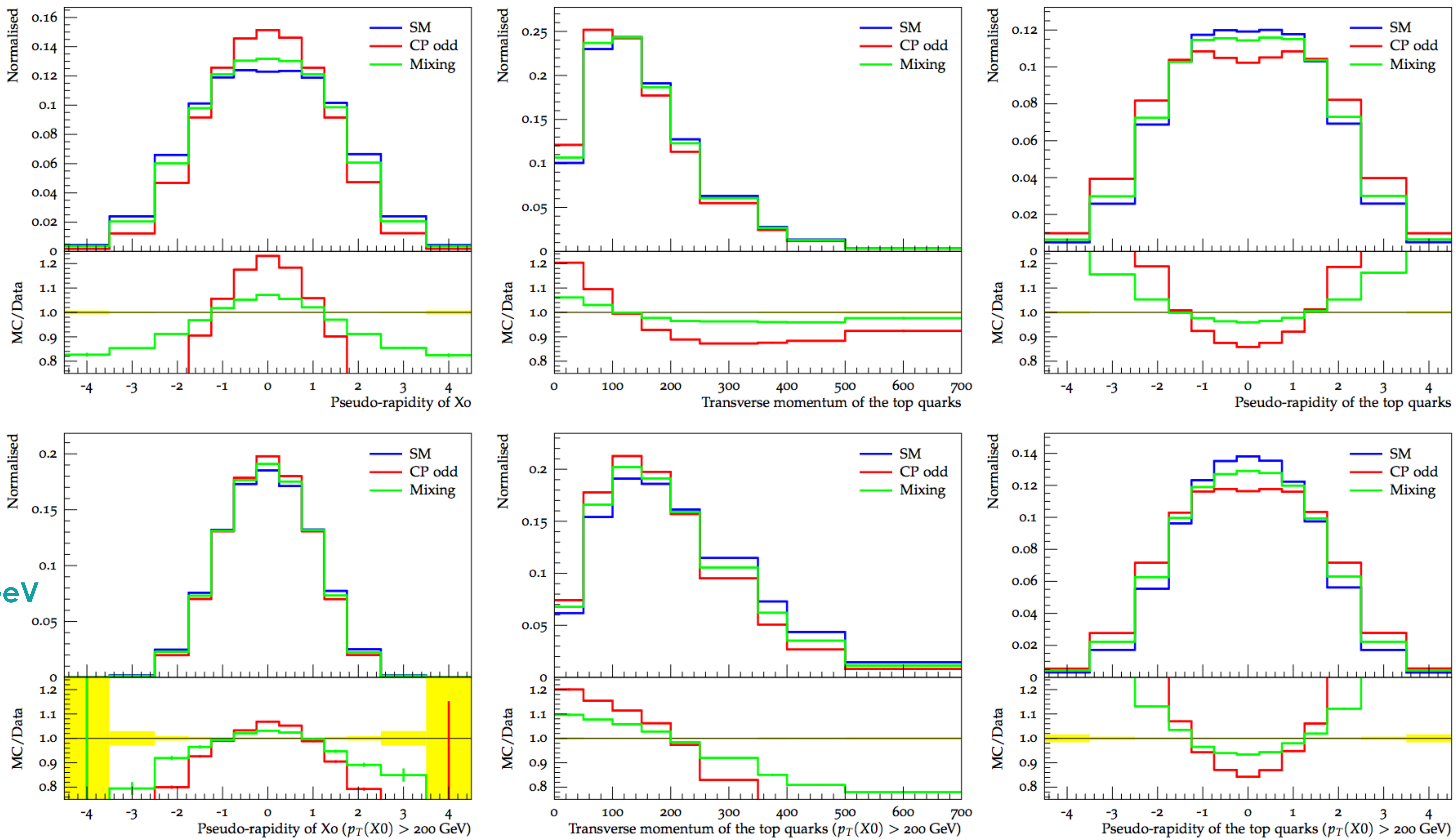


- testing different observables describing the topology of the ttH system
- distributions quite sensitive to different models!

Now we go ahead!

no cuts

$p_T(X_0)$>200 GeV



- testing different observables describing the topology of the ttH system
- distributions quite sensitive to different models!