Hunting for new physics with dibosons at the LHC

Inês Ochoa

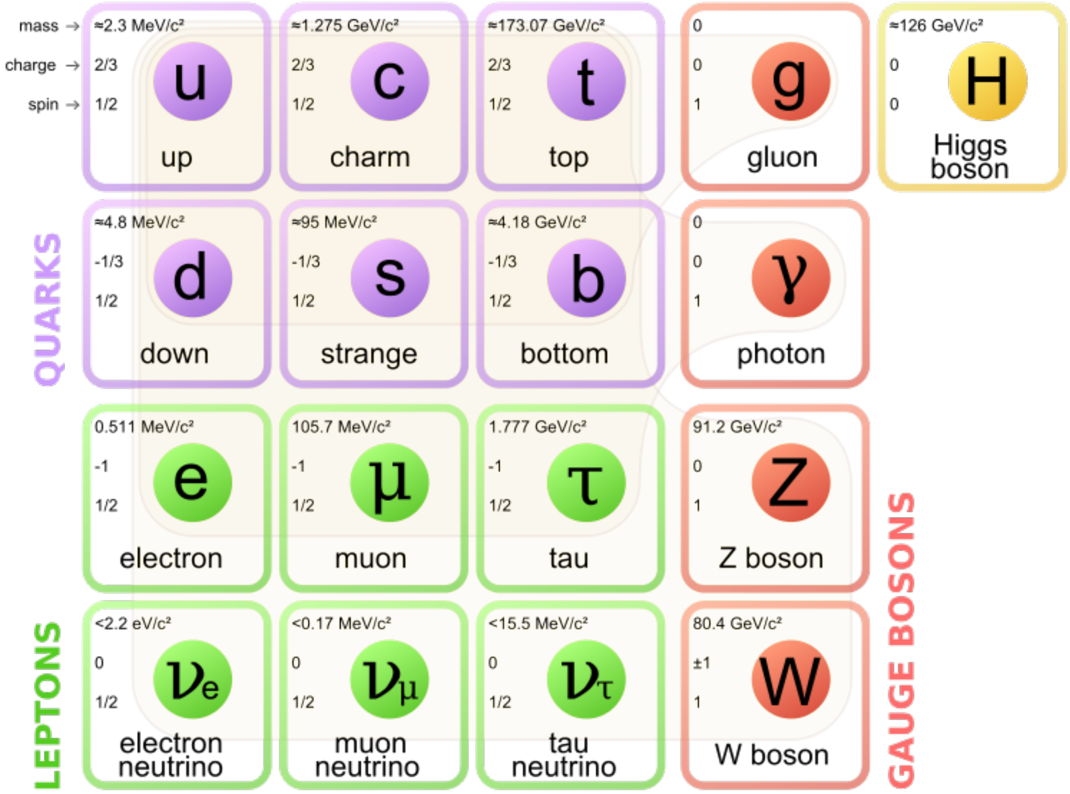
November 12th

Outline

- Motivation:
  - Why search for diboson resonances?
- Latest ATLAS diboson resonance searches
  - Boosted boson tagging
  - Data-driven background estimation
- What next?
  - New strategies and analysis techniques
- Summary

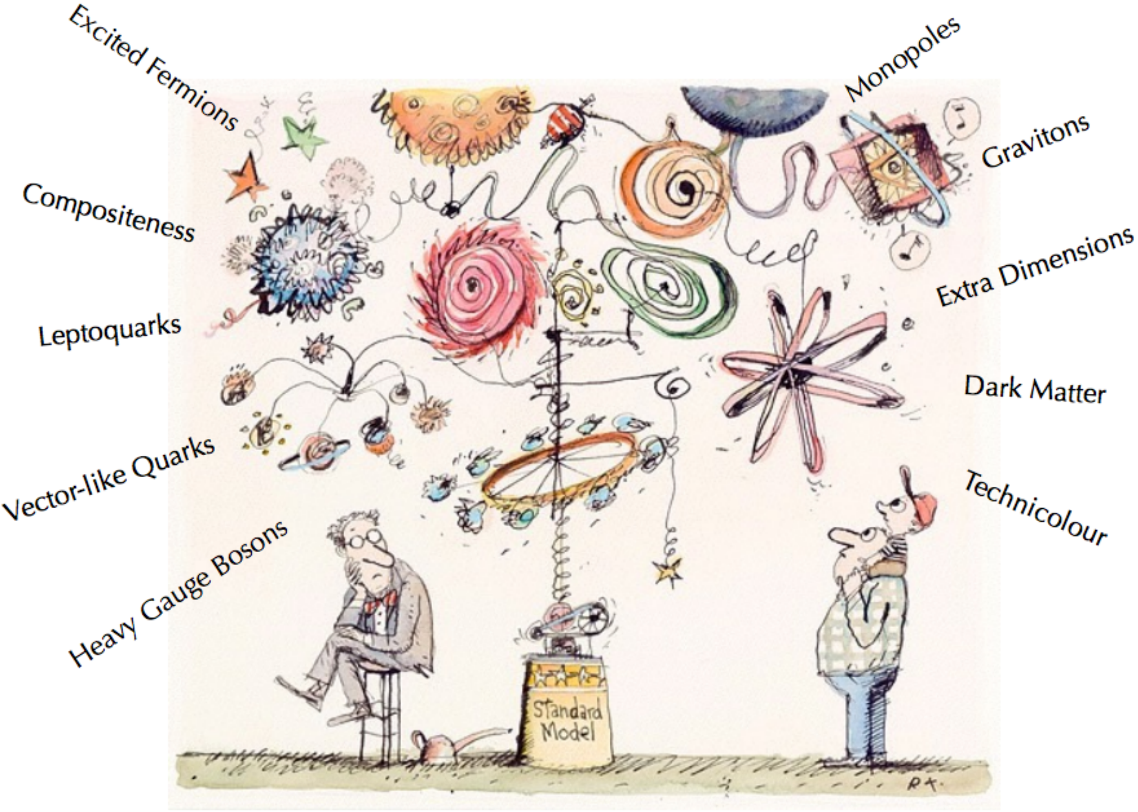
Motivation

The Standard Model is not the complete picture



- No known candidate for dark matter.
- Matter dominance over anti-matter.
- No explanation for masses of particles.
- No explanation for number of generations.
- Gravity not taken into account.
- …

How do we extend it?

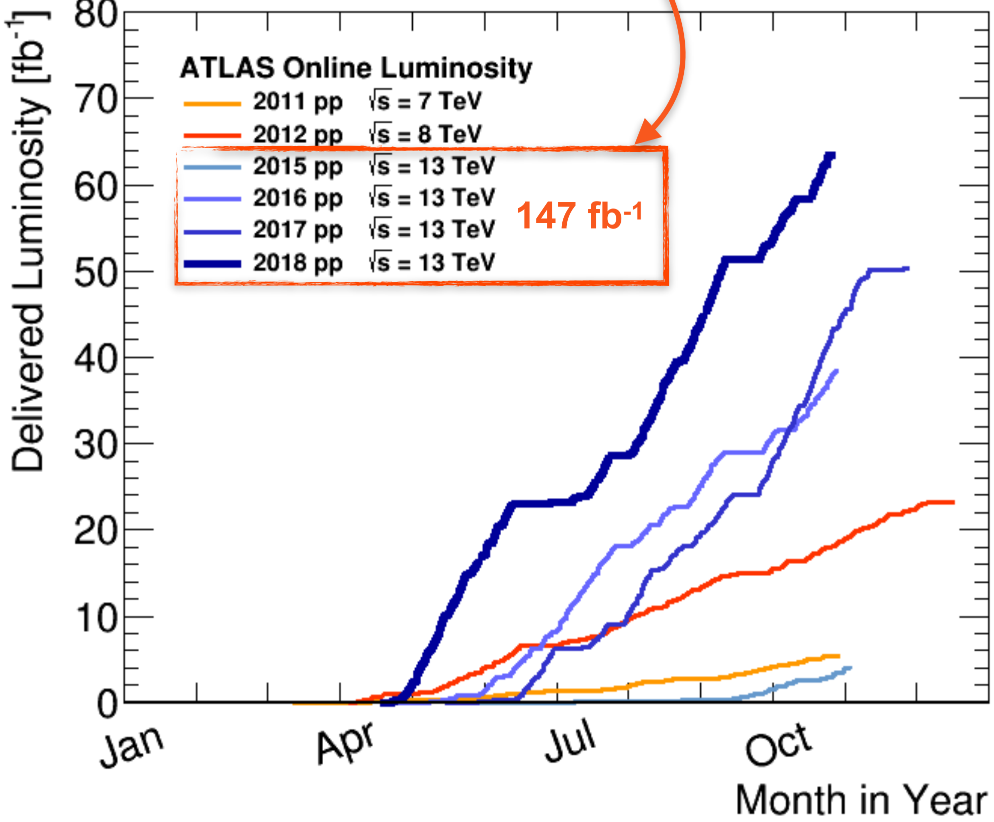


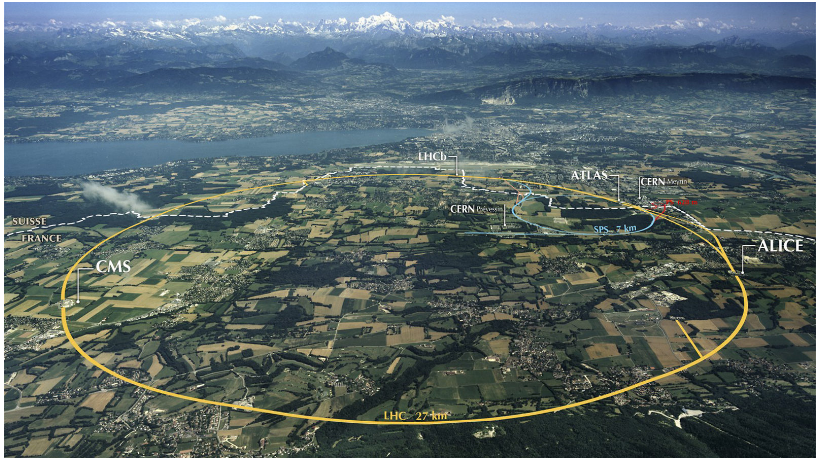
- There are many ways to extend the SM to address some / many of the missing pieces.
- …but no single well-motivated scenario until there is a discovery!

➡ **Need for unbiased and comprehensive approach to New Physics searches to take advantage of the vast and rich LHC dataset.**

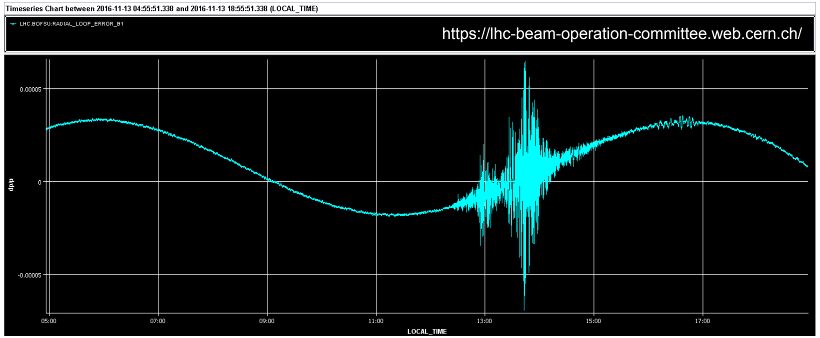
The LHC

Full Run 2 dataset





2/19 calibration

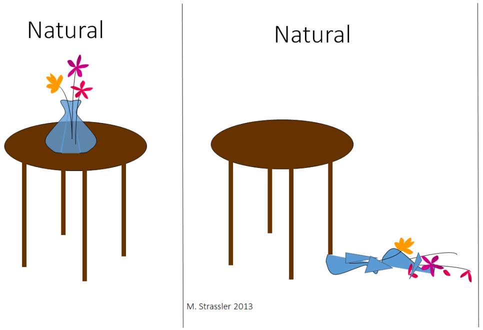


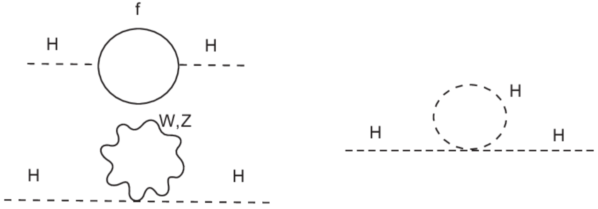
The New Zealand earthquake (2016).

LHC beam orbit displacement.

Higgs and gauge bosons as gates to the unknown

- Precision measurements SM processes so far in excellent agreement with predictions.
- **A new resonance would provide the most dramatic signal of New Physics, with minimal assumptions.**
- Many SM extensions predict new resonances that couple to the gauge and Higgs bosons:
  - Could alleviate **naturalness** problem of the Higgs boson mass…

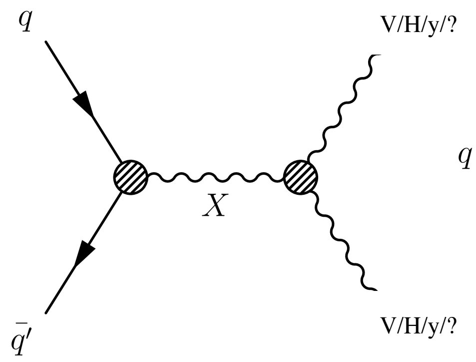


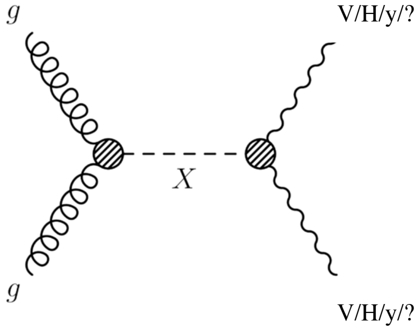


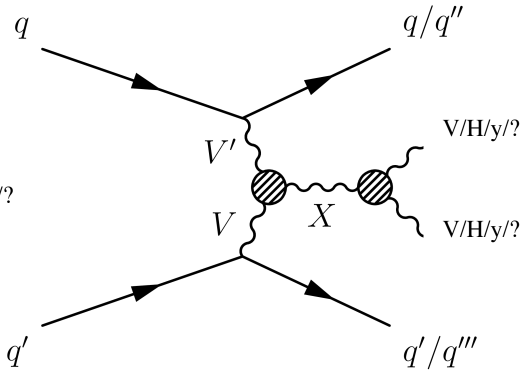
- Experimentally, Higgs and gauge bosons have well-defined signatures that can be targeted using state-of-the-art techniques.

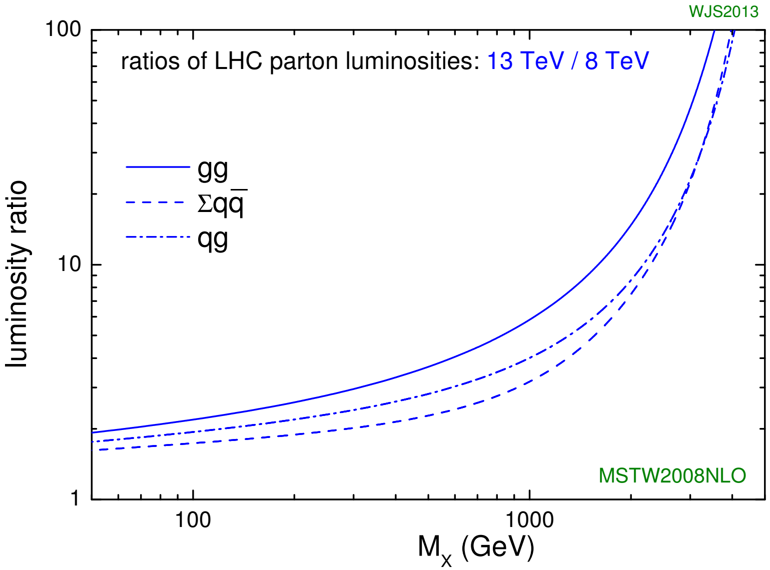
Searches for diboson resonances

X $\stackrel{?}{=}$ gravitons, radions, new vector bosons, new scalars, …





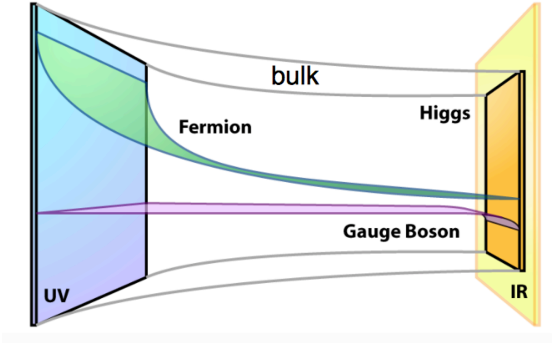


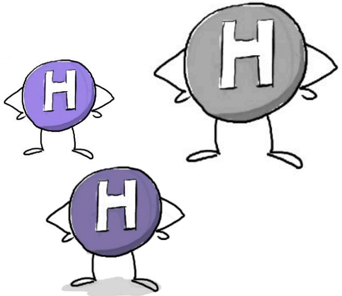


➡ Very high mass reach with 13 TeV pp collisions and access to different initial states.

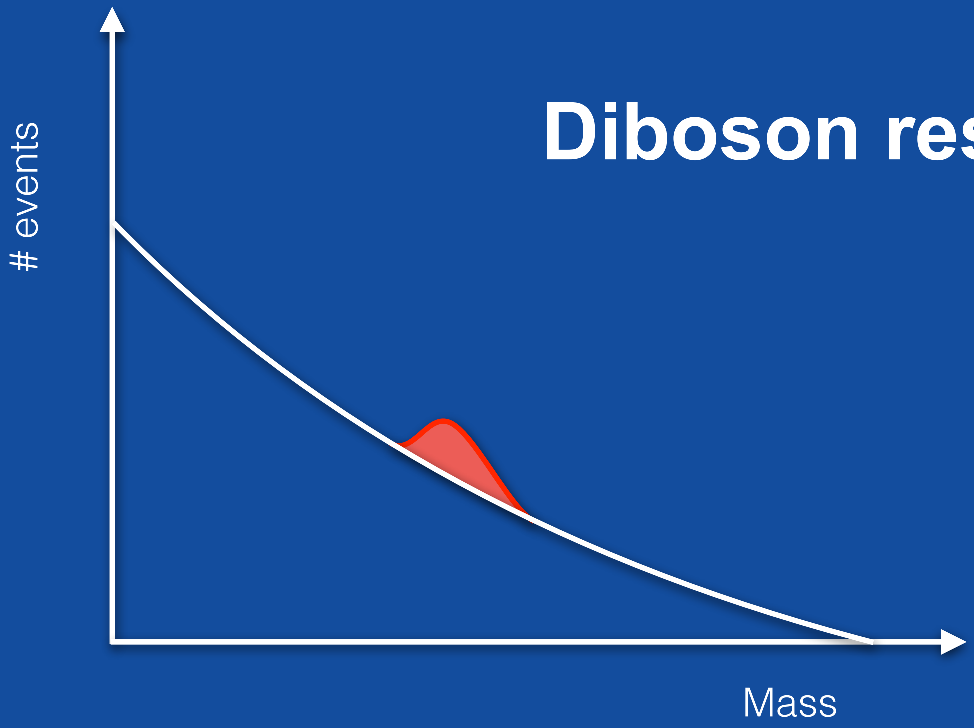
A rich phenomenology

- Looking for Beyond-the-SM physics means exploring a vast and multi-dimensional space.
  - *Simplified models* used to make generic predictions on specific processes.
- Examples of BSM scenarios predicting diboson resonances:
  - Spin-0: extended Higgs sector (e.g. 2HDM)
  - Spin-1: W’ and Z’ bosons from new gauge groups
  - Spin-2: gravitons in warped extra dimensions





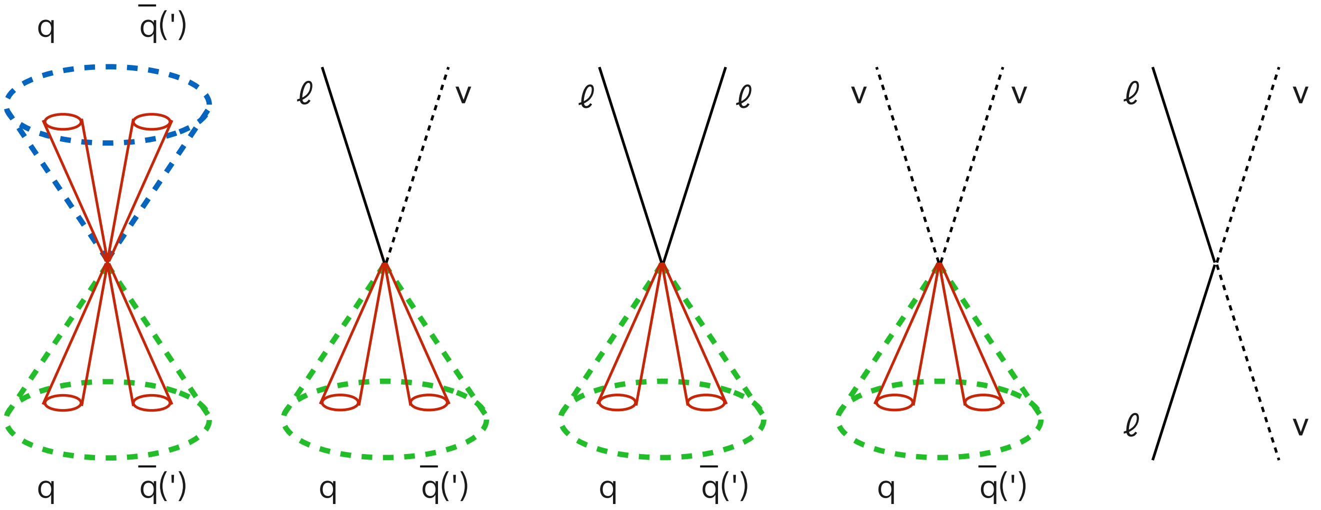
Diboson resonances



A vast collection of final states

BR(W/Z→qq)~70%

BR(H→bb)~60%



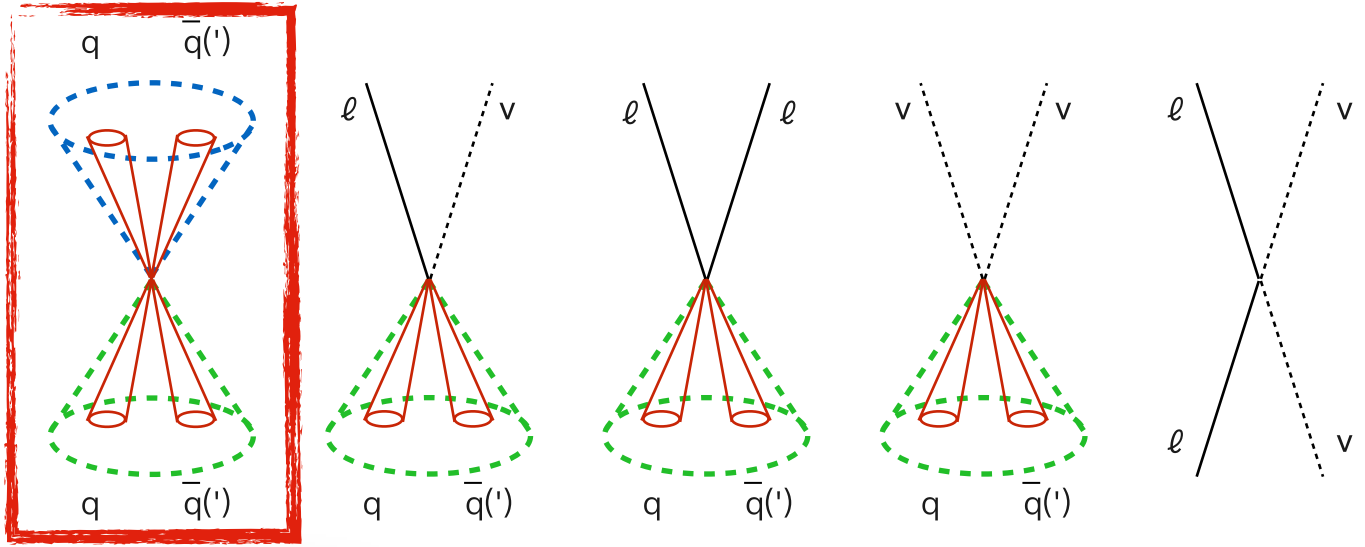
+ W/Z/H+ɣ combinations

- Different analysis target different combinations of SM bosons:
  ➡ W, Z, H, photons
- …and the W,Z leptonic or hadronic decay channels:
  ➡ Trade-off between signal purity and branching ratios.

A vast collection of final states

BR(W/Z→qq)~70%

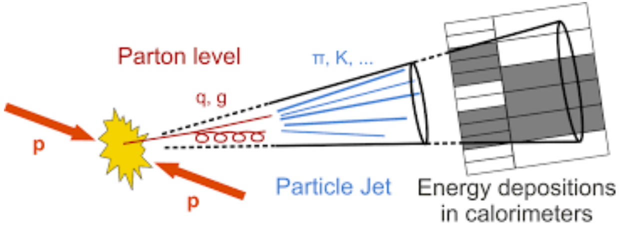
BR(H→bb)~60%

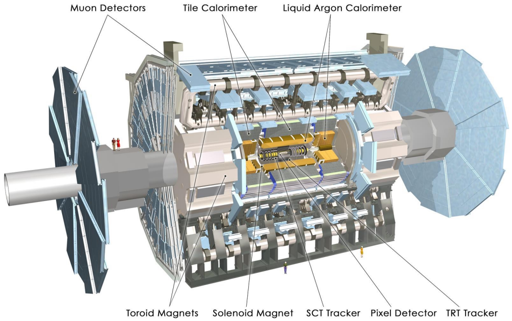


+ W/Z/H+γ combinations

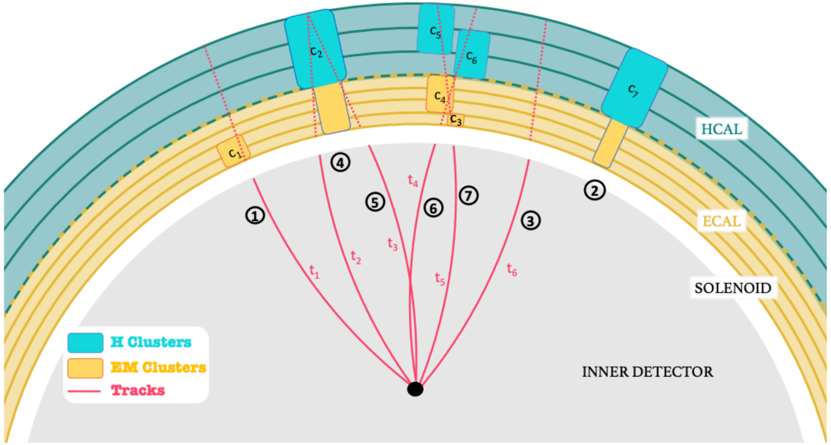
- Different analysis target different combinations of SM bosons:
  - ➡ W, Z, H, photons
- …and the W,Z leptonic or hadronic decay channels:
  - ➡ Trade-off between signal purity and branching ratios.

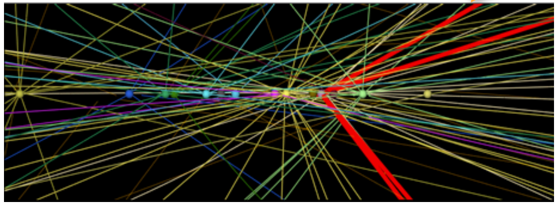
Aside: jets in the ATLAS detector





- A jet is a collection of particles produced by outgoing quarks or gluons:
  - Built from a combination of charged particle tracks and calorimeters deposits.
- Jets for W/Z/Higgs reconstruction:
  - ✓ Large radius to capture full decay products.
  - ✓ Removal of pile-up and underlying event contributions.





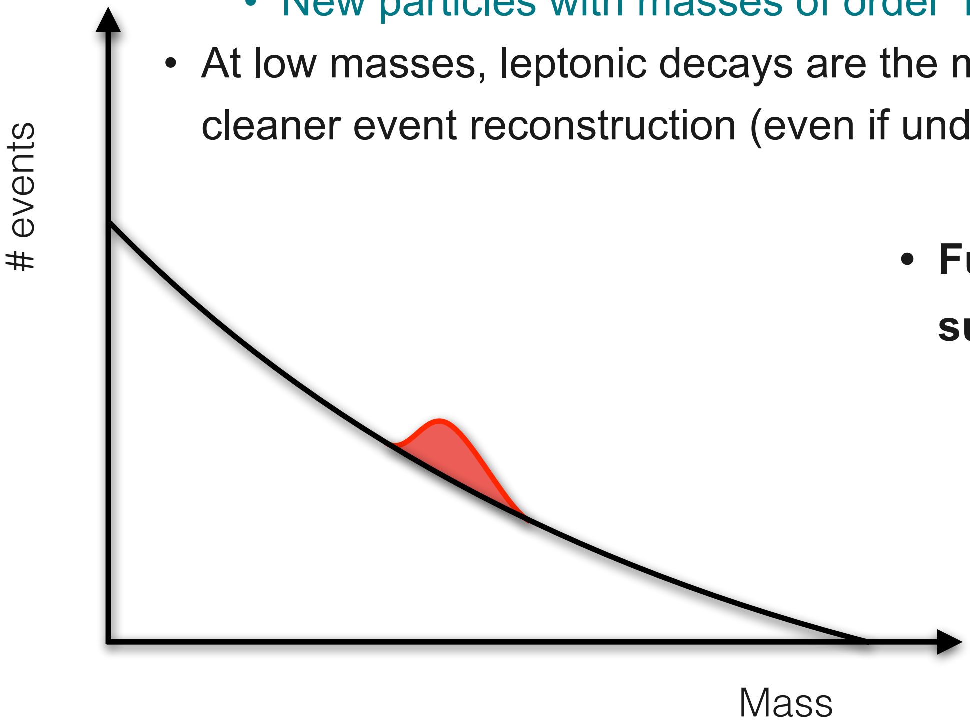
Average 33 collisions per bunch crossing in Run 2.

Hadronic final states

BR(W/Z→qq)~70%

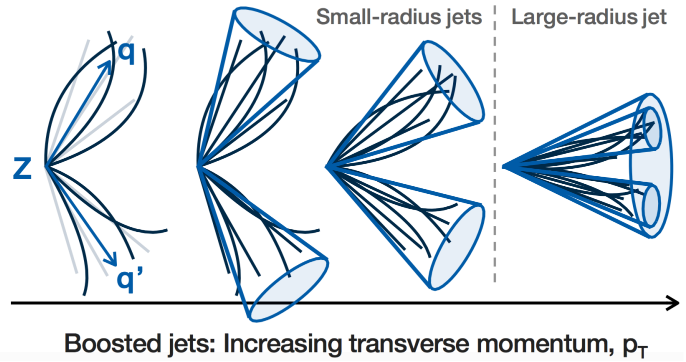
BR(H→bb)~60%

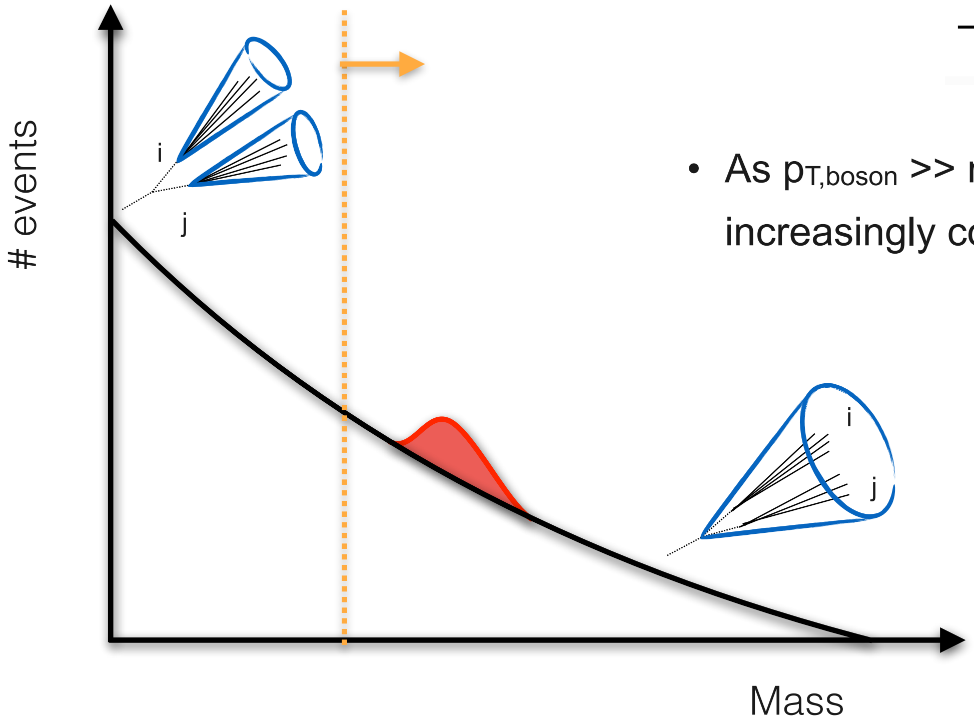
- Diboson resonance searches target a vast kinematic regime:
  - New particles with masses of order 100 GeV all the way up to several TeV!
- At low masses, leptonic decays are the most sensitive: easier trigger strategy, cleaner event reconstruction (even if under-constrained in case of neutrinos).



- **Fully-hadronic final states are ideally suited for the high mass region:**
  - ✓ Branching ratio advantage: dominant decay modes of W,Z,H bosons.
  - ✓ Smoothly falling Standard Model background (dominated by quark and gluon initiated jets - "QCD processes”).

The “boosted” regime



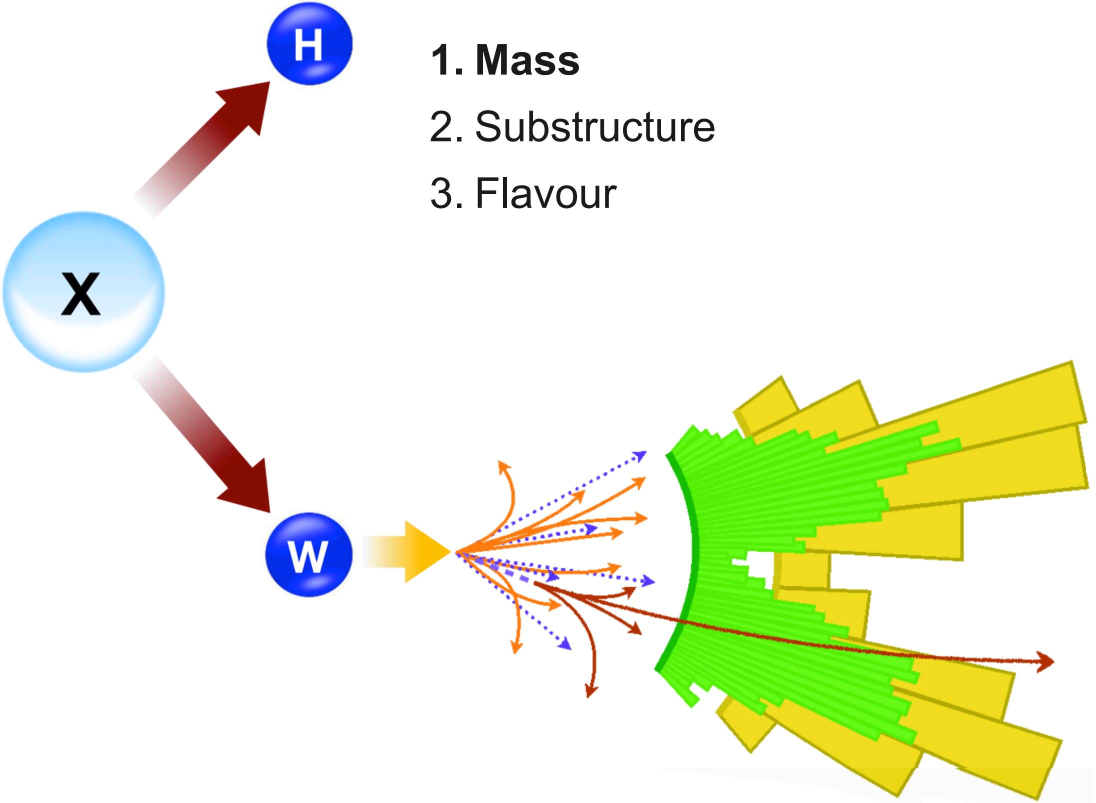


- As $p_{T,boson}$ >> m_{boson}, the boson decay products become increasingly collimated in the lab frame.

- We need *boosted boson tagging* techniques to identify jets from boson decays and suppress the QCD background.

Ingredients for boosted boson tagging

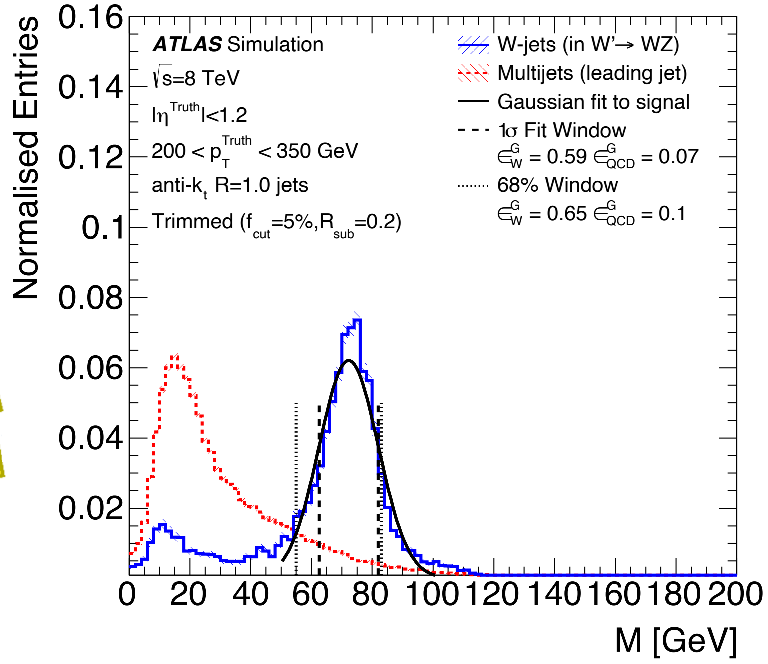
Is the jet mass compatible with a W, Z, H?



1. Mass

2. Substructure

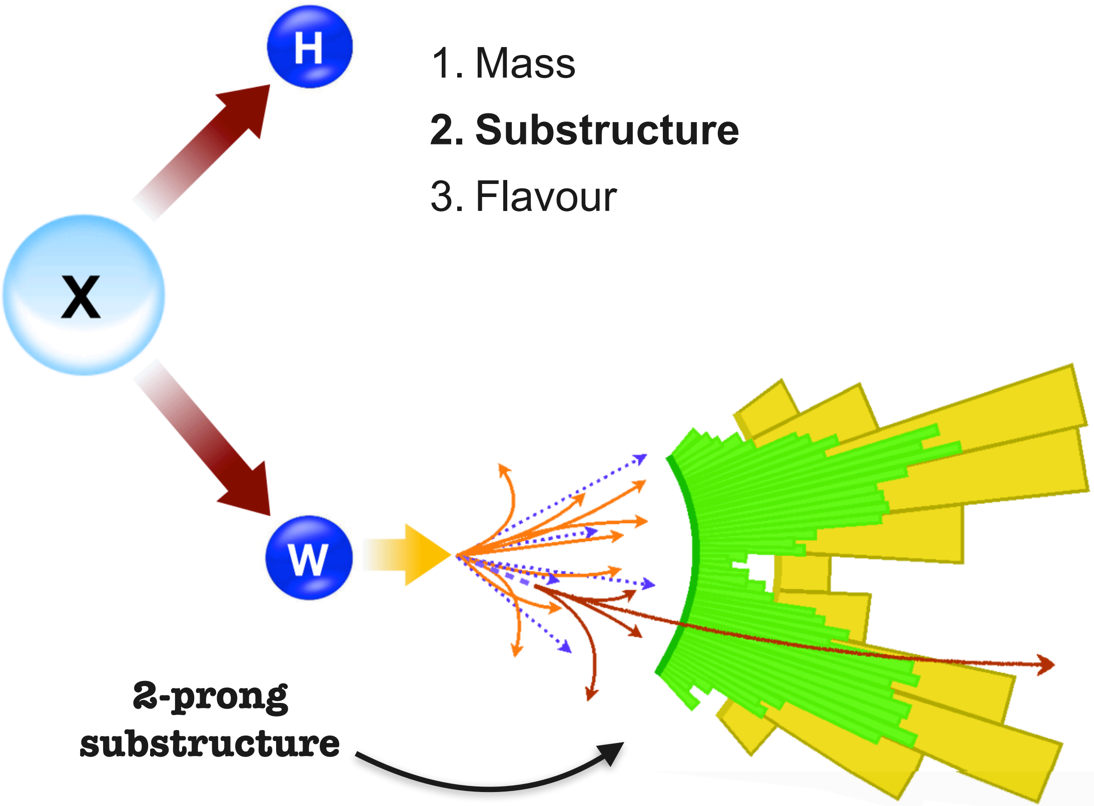
3. Flavour

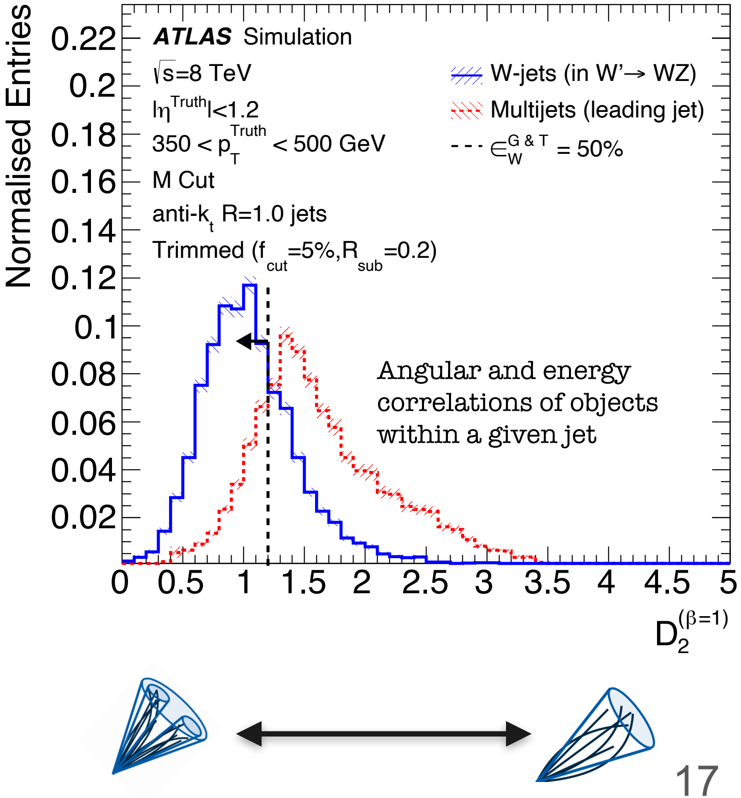


Ingredients for boosted boson tagging

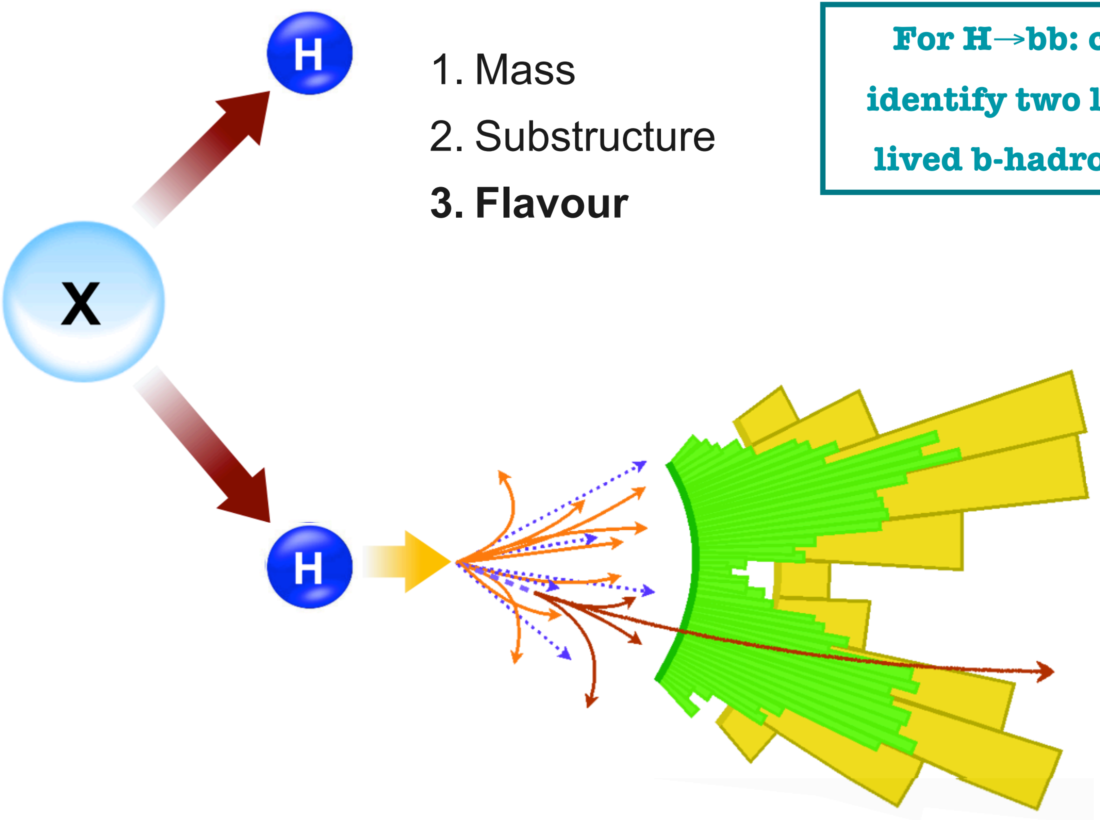
Does the jet have a 2-prong structure?

1. Mass
2. **Substructure**
3. Flavour



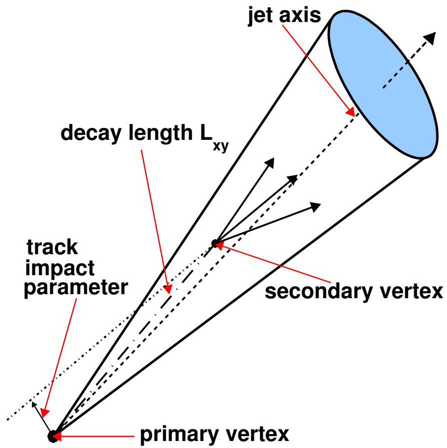


Ingredients for boosted boson tagging



1. Mass
2. Substructure
3. **Flavour**

For H→bb: can identify two long-lived b-hadrons?



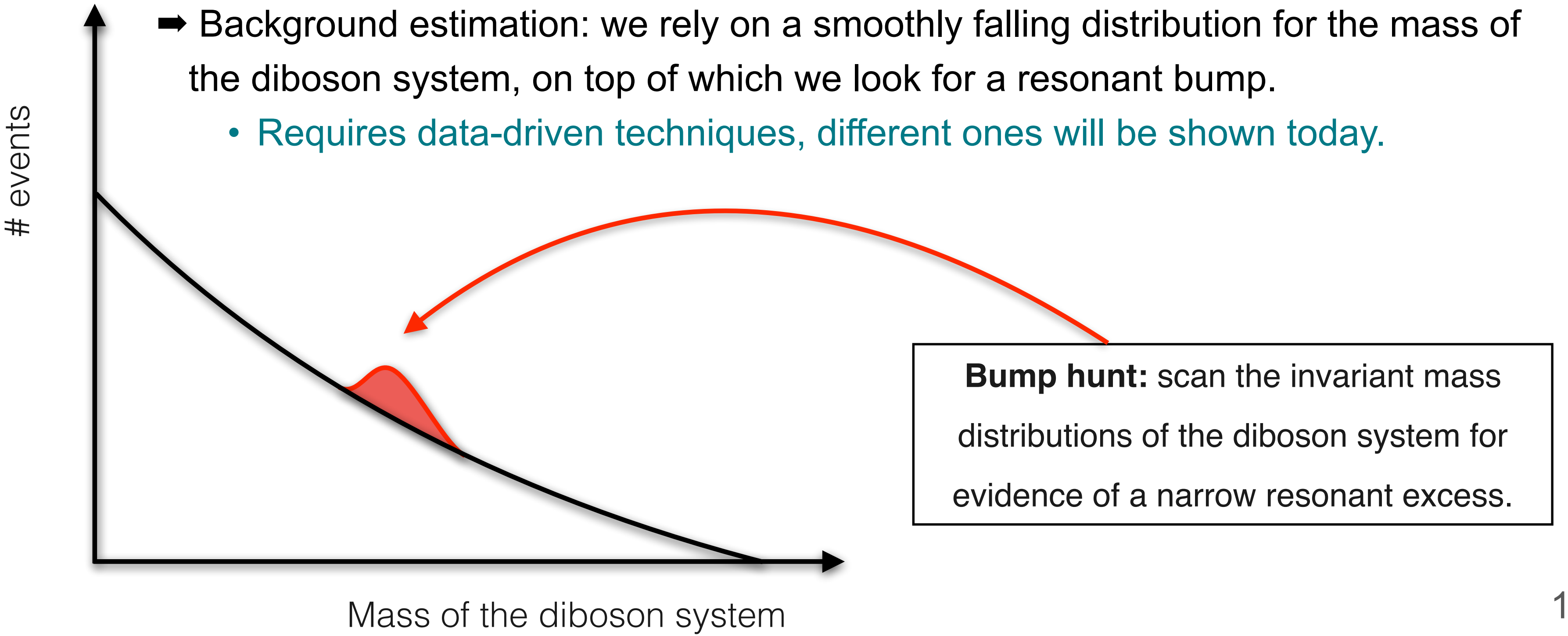
- A b-hadron decay provides a measurable displaced **secondary vertex** in the detector.
- **B-tagging** algorithms for identifying b-jets and suppressing light-jets.

Search strategy

✓ Boosted boson tagging techniques

➡ Background estimation: we rely on a smoothly falling distribution for the mass of the diboson system, on top of which we look for a resonant bump.

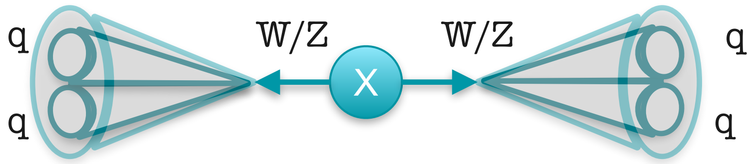
- Requires data-driven techniques, different ones will be shown today.

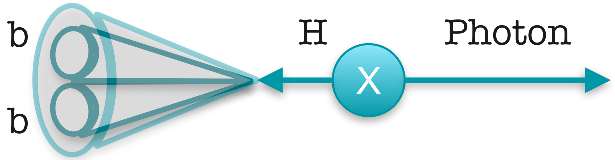


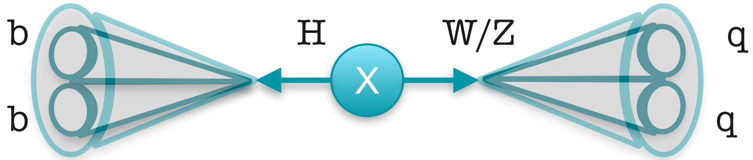
In this talk

- **Will focus on three results from the ATLAS Collaboration in fully-hadronic decay channels:**
  - Y→VV→qqqq: JHEP09(2019)091
  - Y→VH→qqbb: 2007.05293 (recently submitted to PRD)
  - Y→Hɣ→bbɣ: 2008.05928 (just accepted to PRL)

V = W, Z

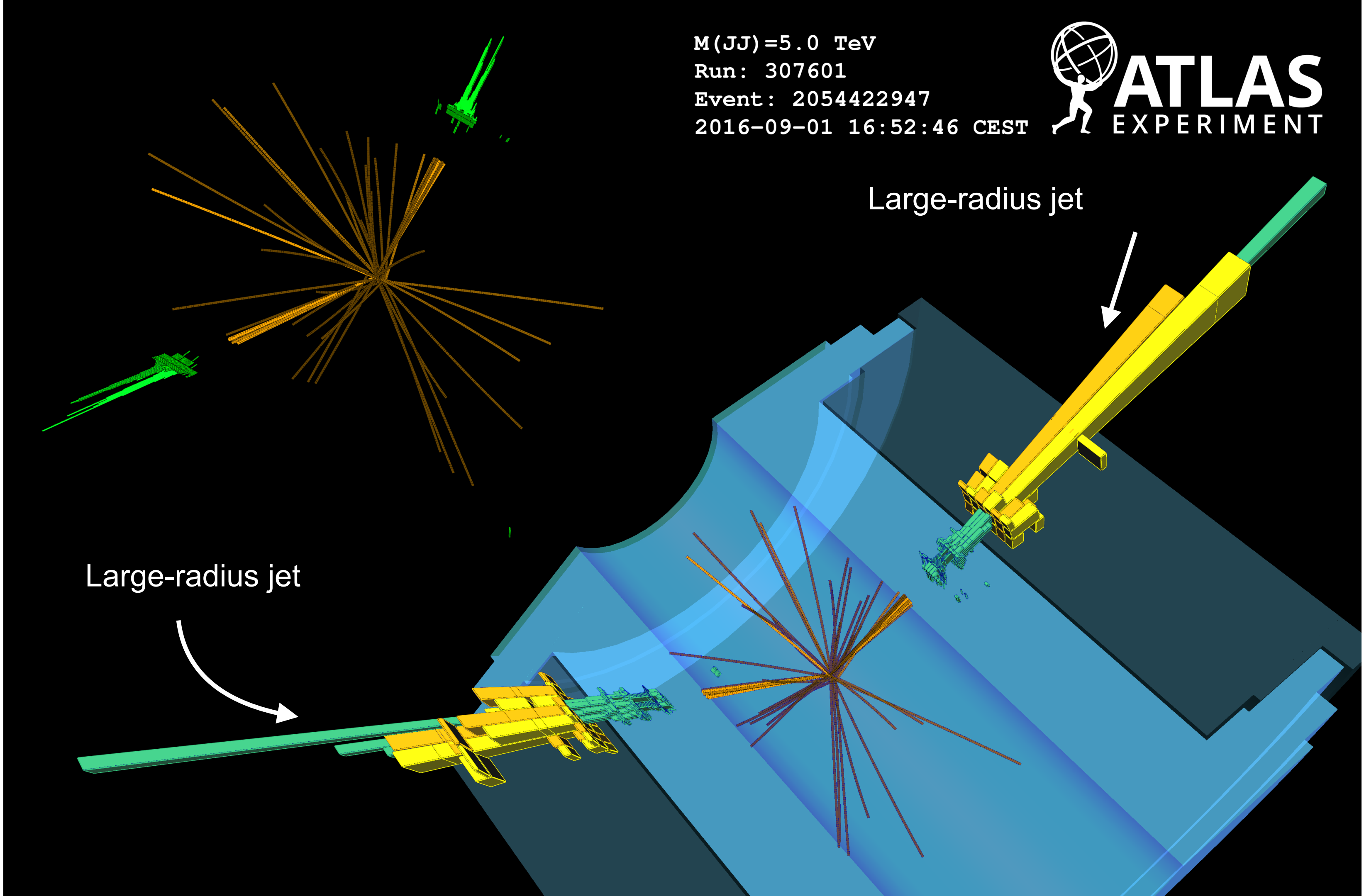






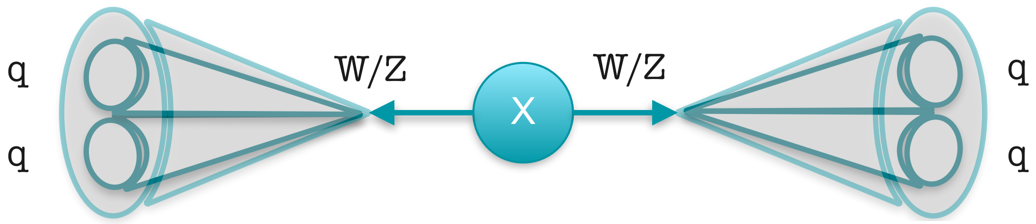
Latest results

$VV \rightarrow qqqq$



JHEP09(2019)091

VV→qqqq
W/Z taggers



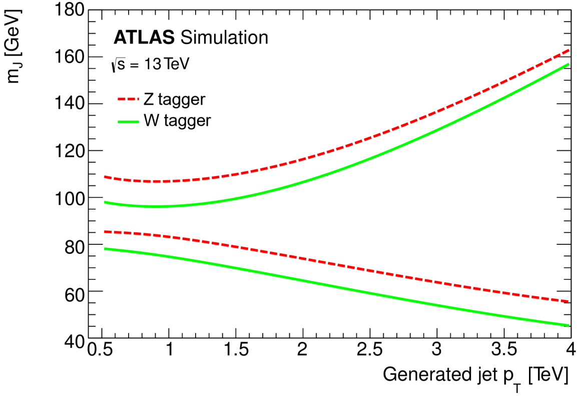
Topology and boson tag:

$|\Delta y| = |y_1 - y_2| < 1.2$

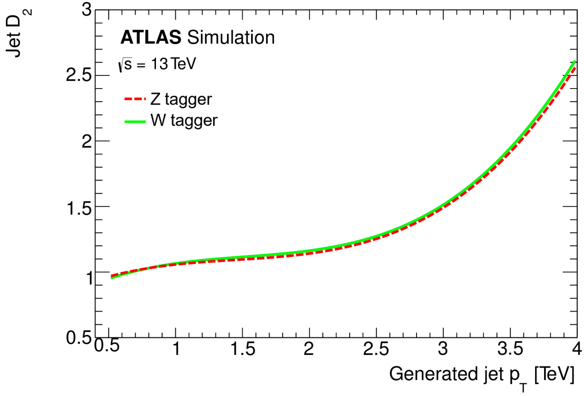
$A = (p_{T1} - p_{T2}) / (p_{T1} + p_{T2}) < 0.15$

Boson tag with D_2 variable, n_{trk} variable, and W or Z mass window

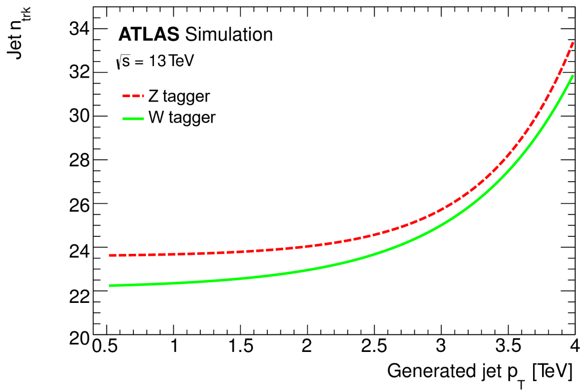
Mass



D_2



Number of tracks



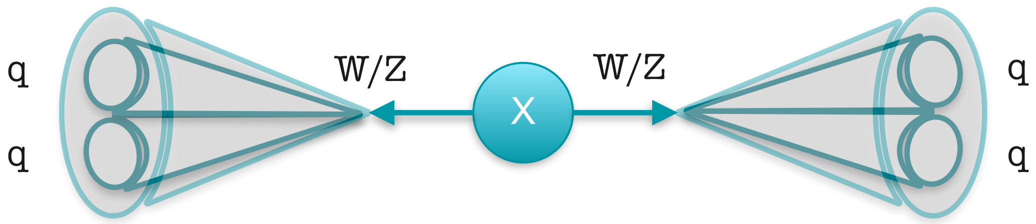
Cuts optimized to maximize signal significance

- Larger windows at high masses where background is lower

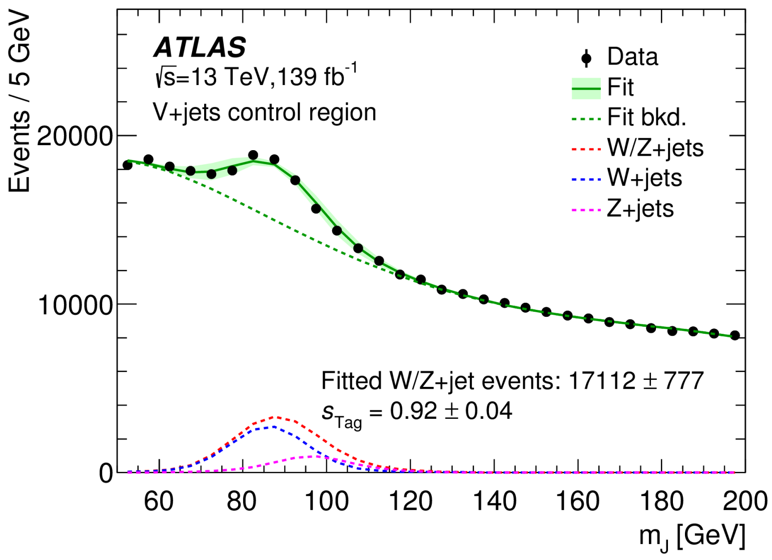
Signal efficiencies of ~20% at low p_T and 40% at high p_T

JHEP09(2019)091

VV→qqqq
Background estimation



V+jets:

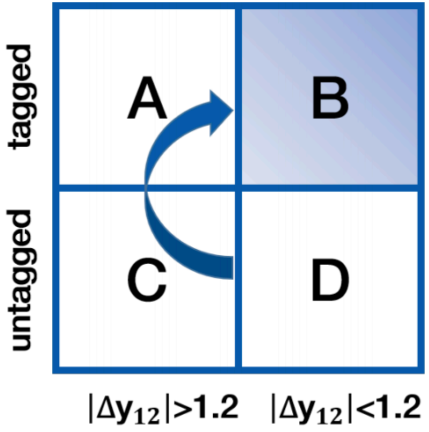


Used to extract W/Z tagger efficiency corrections between data and simulation.

QCD:

Direct fit to observed m_{VV} spectrum in the signal region B.

Validated in regions A, C, D.

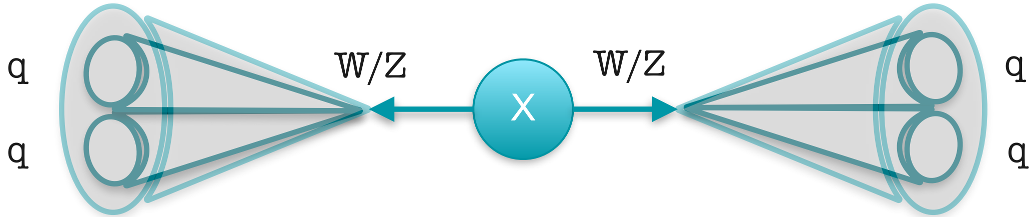


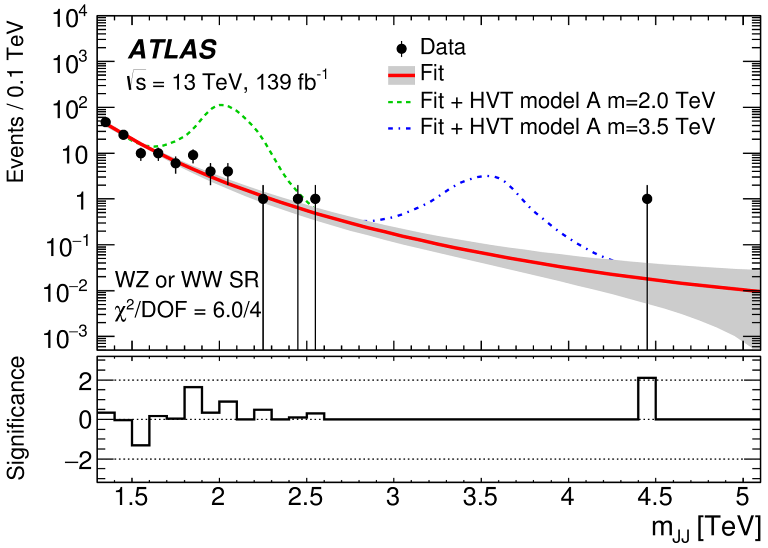
$$\frac{dn}{dx} = p_1(1-x)^{p_2-\xi p_3}x^{-p_3}$$

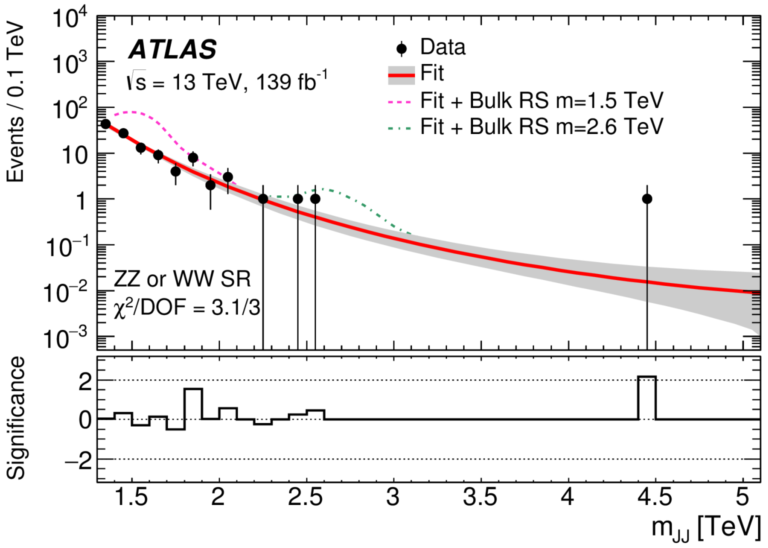
$$x = m_{JJ}/\sqrt{s}$$

JHEP09(2019)091

VV→qqqq
Results





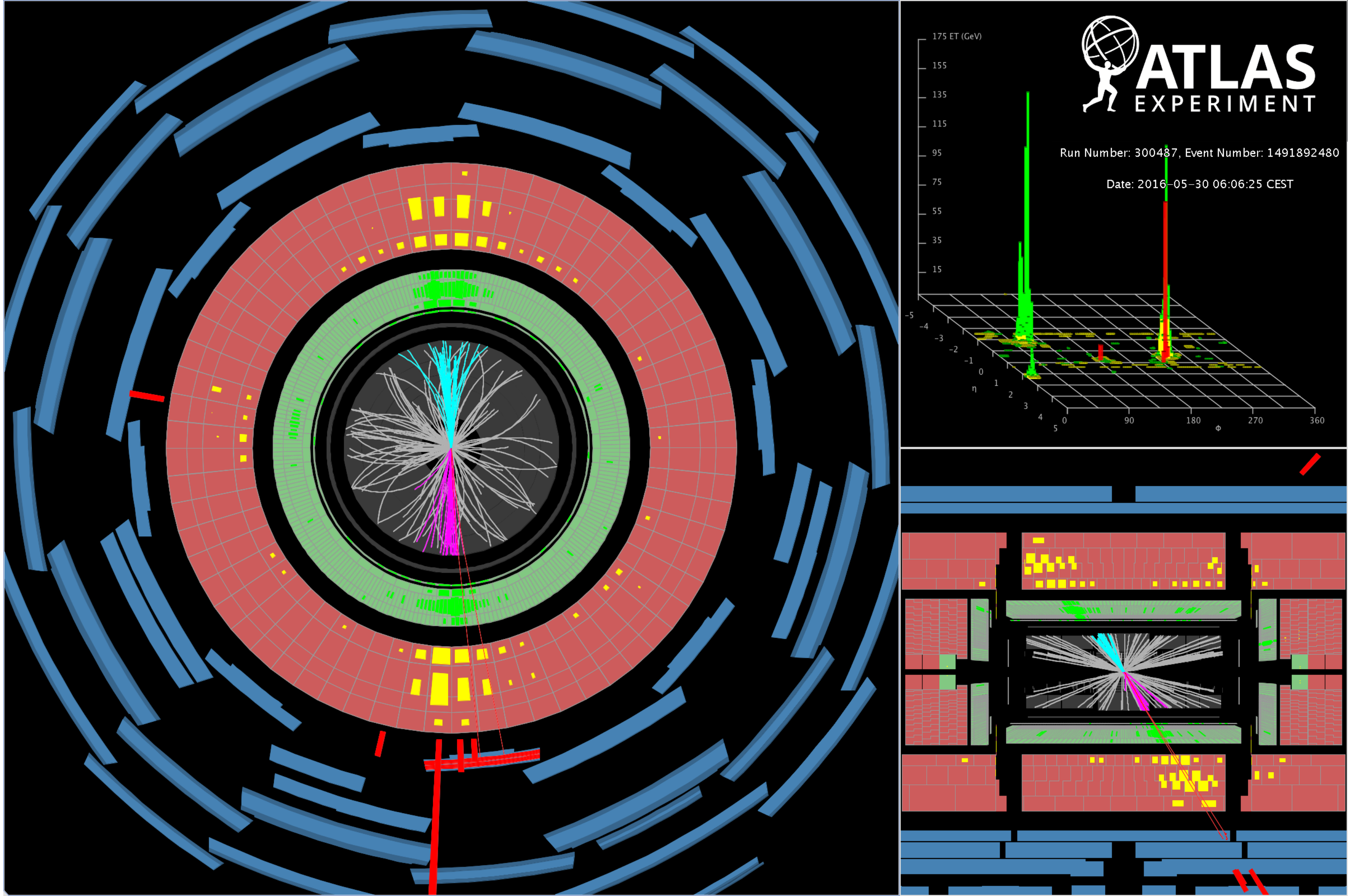


Done separately for three overlapping selections: WW, WZ and ZZ.

Can combine results into WW + WZ and WW + ZZ according to signal interpretation.

JHEP09(2019)091

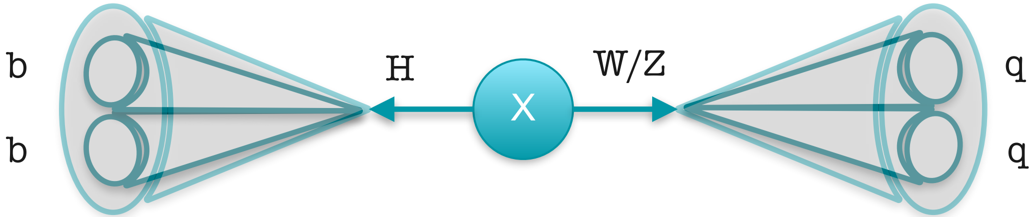
VH→qqbb



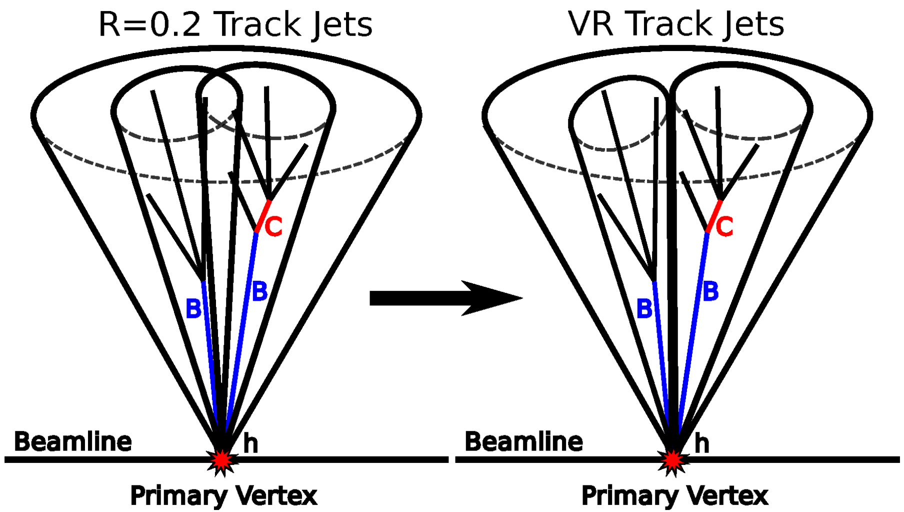
2007.05293

VH→qqbb

H→bb identification



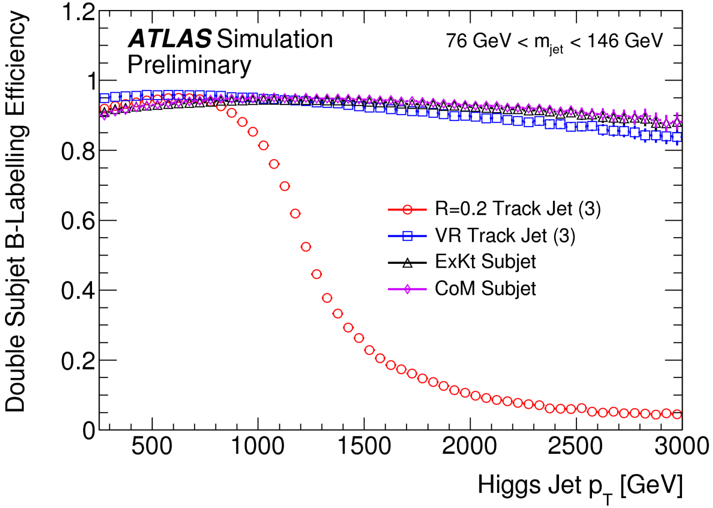
Most powerful handle on Higgs boson identification: b-tagging of **two jets** inside large-radius jet.



Previous searches used jets with R=0.2:

Can do much better e.g. with variable-radius jets

Large acceptance gain at high transverse momentum

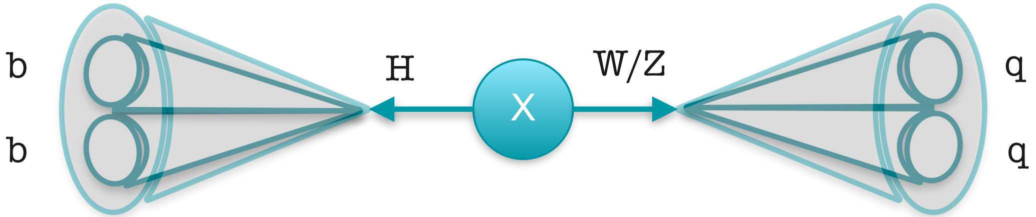


ATL-PHYS-PUB-2017-010

"VR" = variable-radius

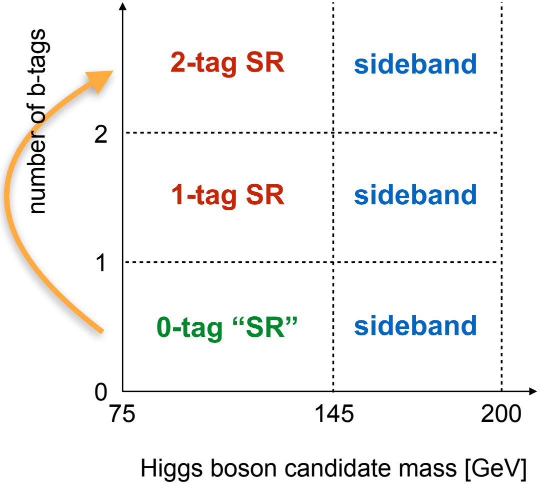
VH→qqbb

Background estimation



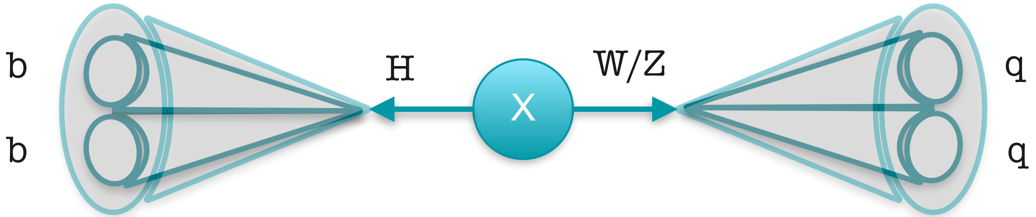
Background dominated by QCD:

1. Extract template from data events where no b's are found ("0-tag" region)



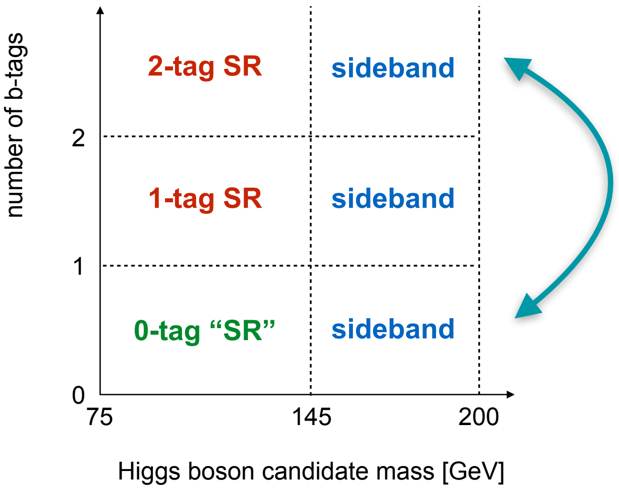
2007.05293

VH→qqbb
Background estimation



Background dominated by QCD:

1. Extract template from data events where no b's are found ("0-tag" region)
2. Using a sideband, extract multi-dimensional **ratio** between 0-tag and 2-tag with a boosted decision tree (BDT)



$$r(x) = \frac{h(x)}{1-h(x)}$$

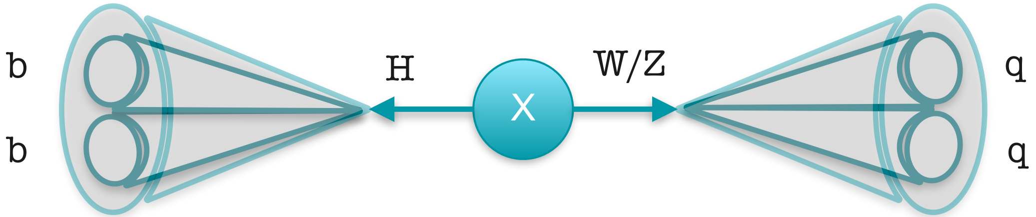
h(x) is the classifier output

x = four-vectors of small radius jets, W/Z and Higgs p_T, ...

2007.05293

VH→qqbb

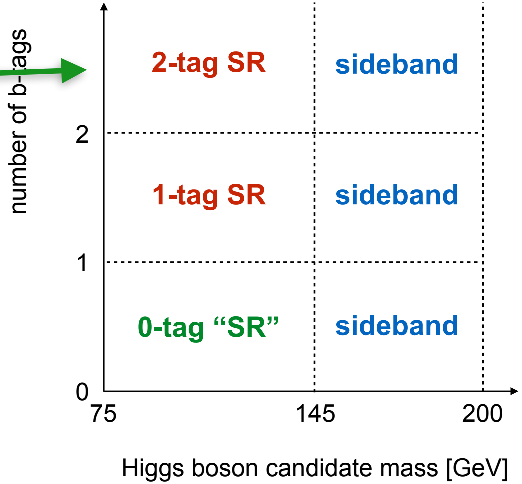
Background estimation



Background dominated by QCD:

1. Extract template from data events where no b's are found ("0-tag" region)
2. Using a sideband, extract multi-dimensional ratio between 0-tag and 2-tag with a boosted decision tree (BDT)
3. "Correct" 0-tag template to obtain a background prediction in 2-tag

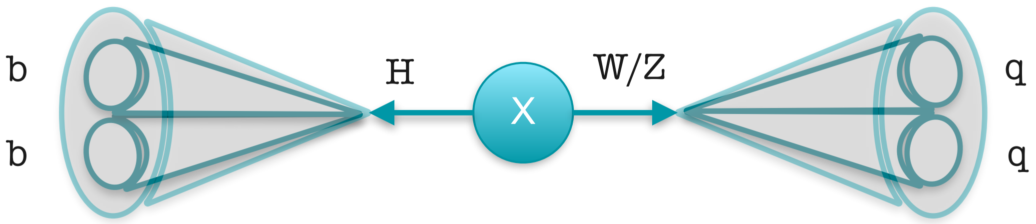
$$\boxed{f_{2-tag}(x) = r(x) f_{0-tag}(x)}$$



2007.05293

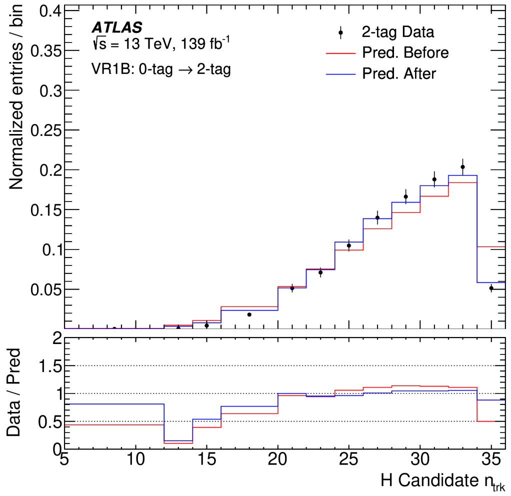
VH→qqbb

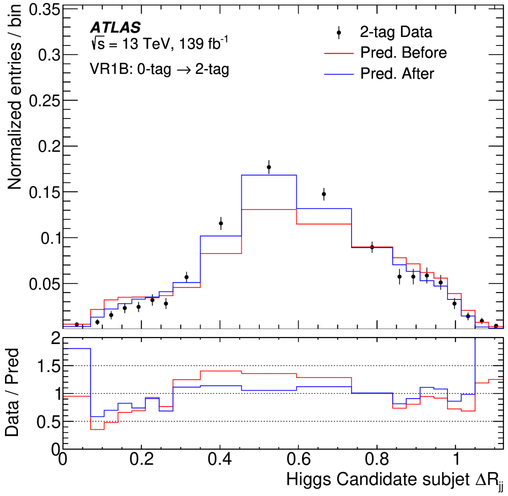
Background estimation

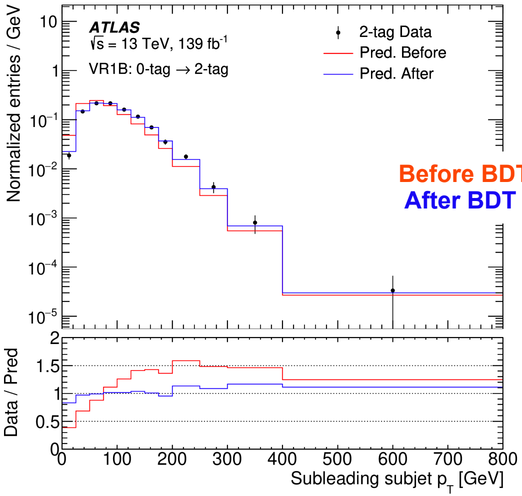


Background dominated by QCD:

✓ Use validation regions in data to confirm quality of background description





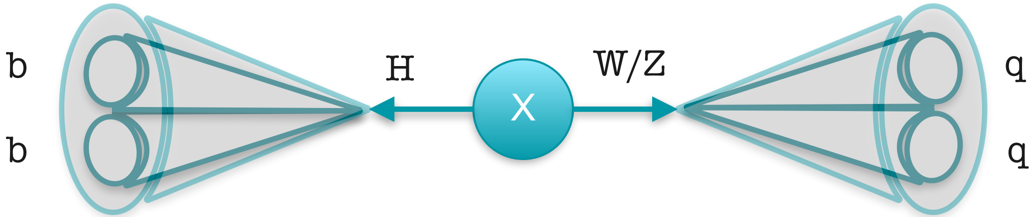


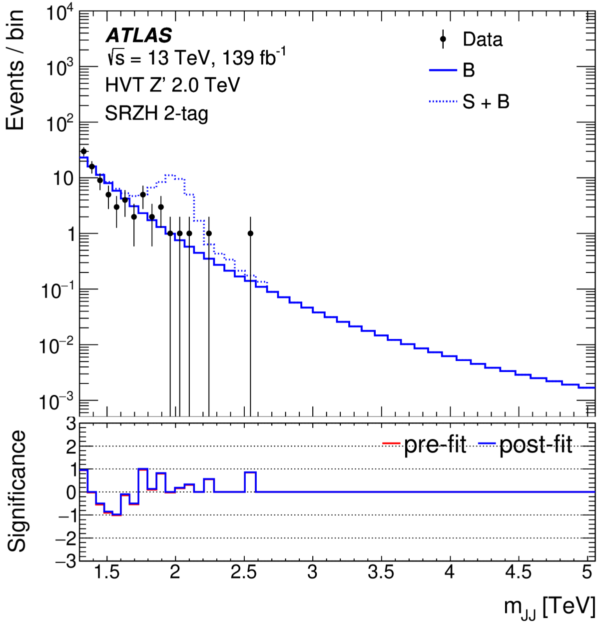
Before BDT reweighting
After BDT reweighting

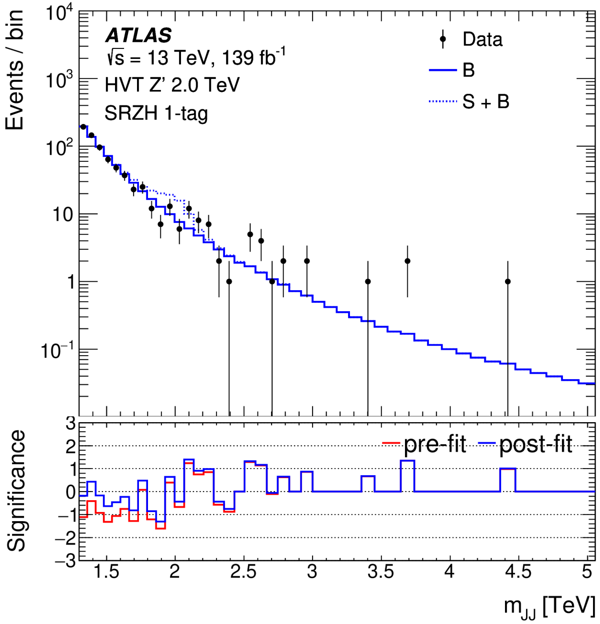
2007.05293

VH→qqbb

Results





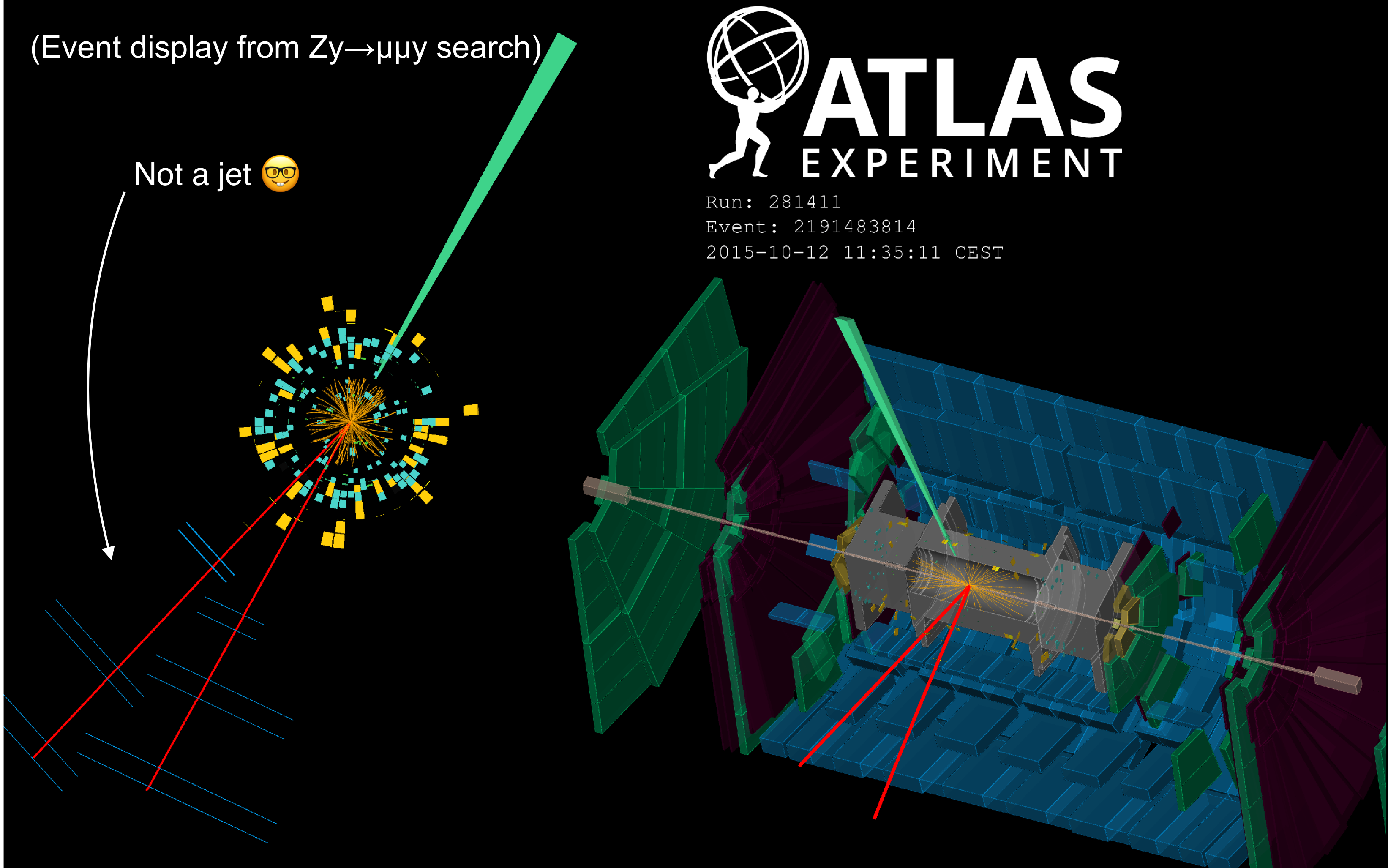


➡ 1-tag events contribute at very high masses

Results produced separately for WH and ZH search (W and Z regions overlap)

ZH results shown here: events with 2 b-jets on the left, events with 1 b-jet on the right

$H\gamma \to bb\gamma$

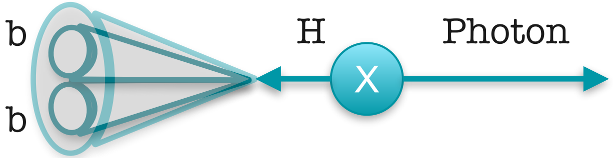


Phys. Lett. B 764 (2017) 11

2008.05928

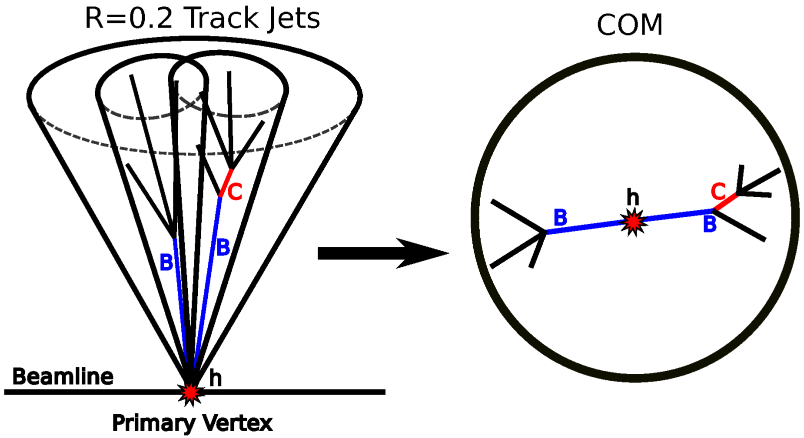
$H\gamma \rightarrow bb\gamma$

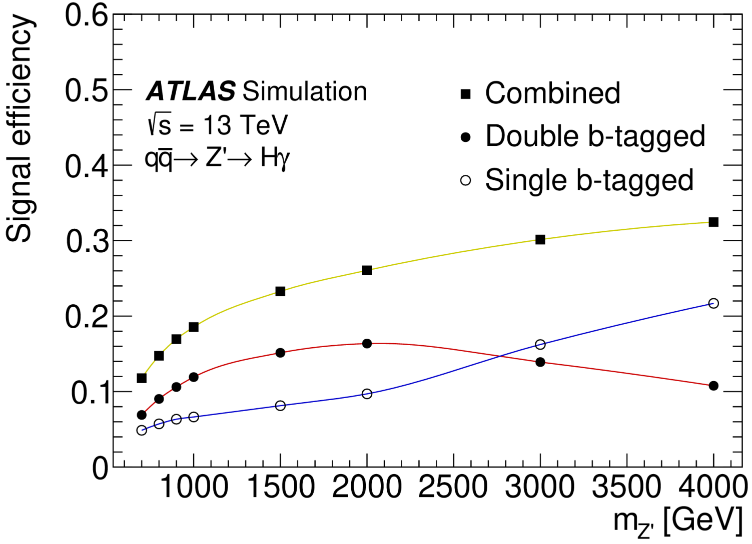
$H \rightarrow bb$ identification



Another method for identifying boosted Higgs bosons decaying to pairs of b-quarks

- Tracks are associated to each candidate "b-jet" in the rest frame of the Higgs candidate
  - Also large acceptance gain at high transverse momentum



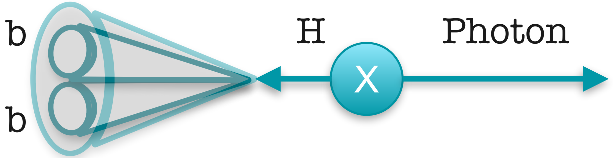


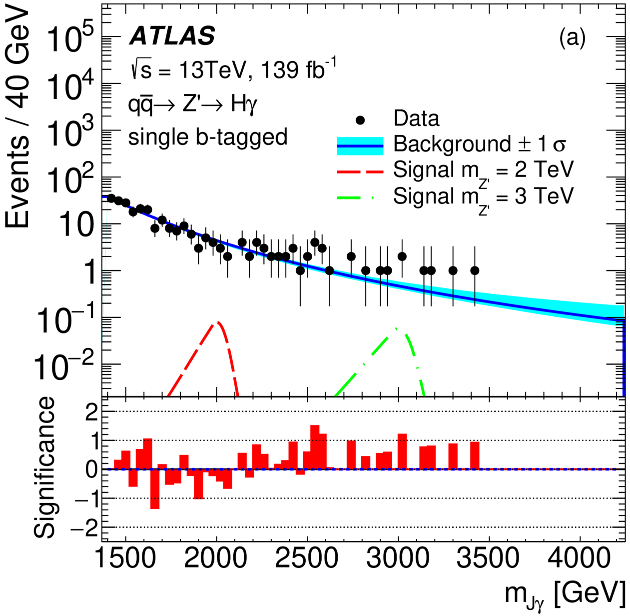
ATL-PHYS-PUB-2017-010

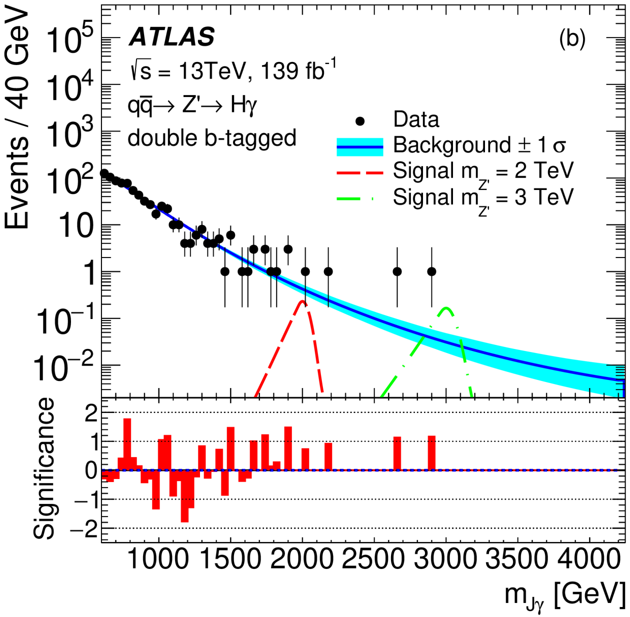
Hγ→bbγ

Results

HOT OF THE PRESS!







Dominant background from γ+jets

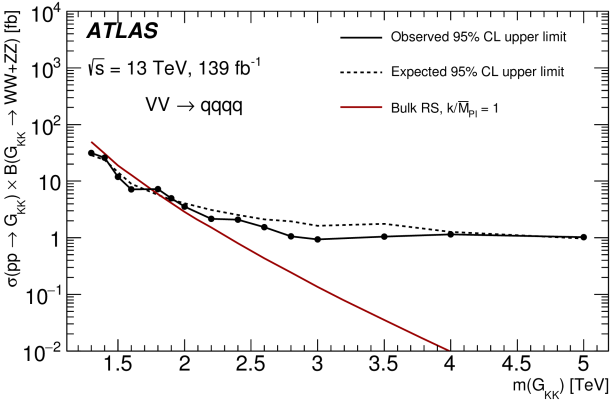
- Fit function validated in control data samples (sidebands and 0-tags)

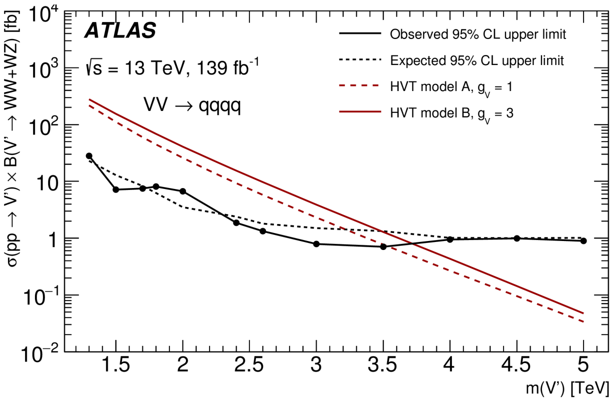
$$B(m_{J\gamma}) = (1-x)^{p_1} x^{p_2+p_3\log(x)}$$

2008.05928

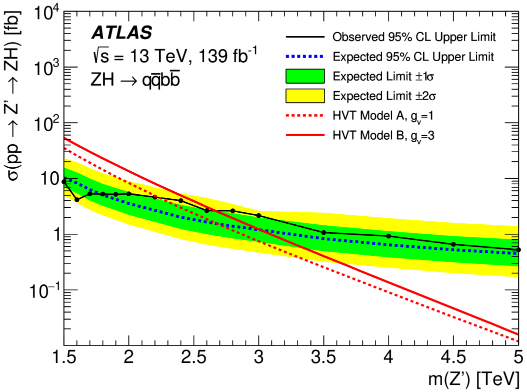
Limits and exclusions

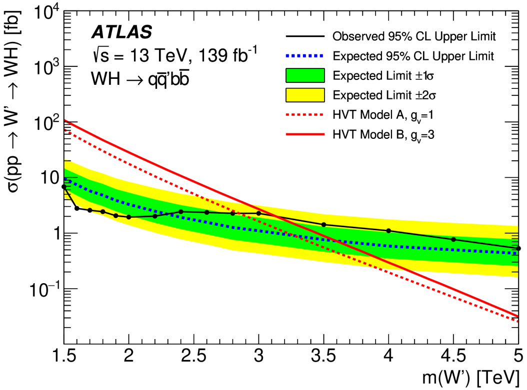
VV



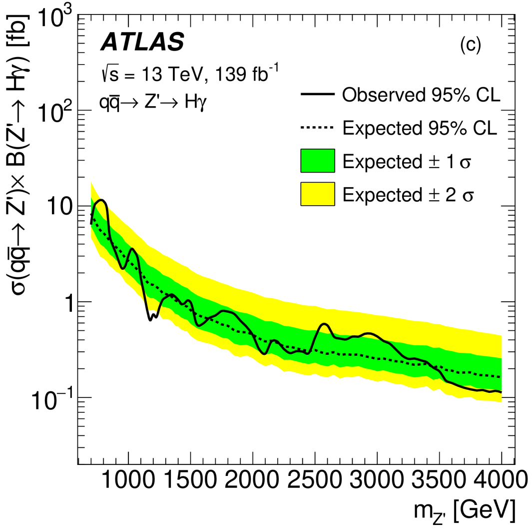


VH

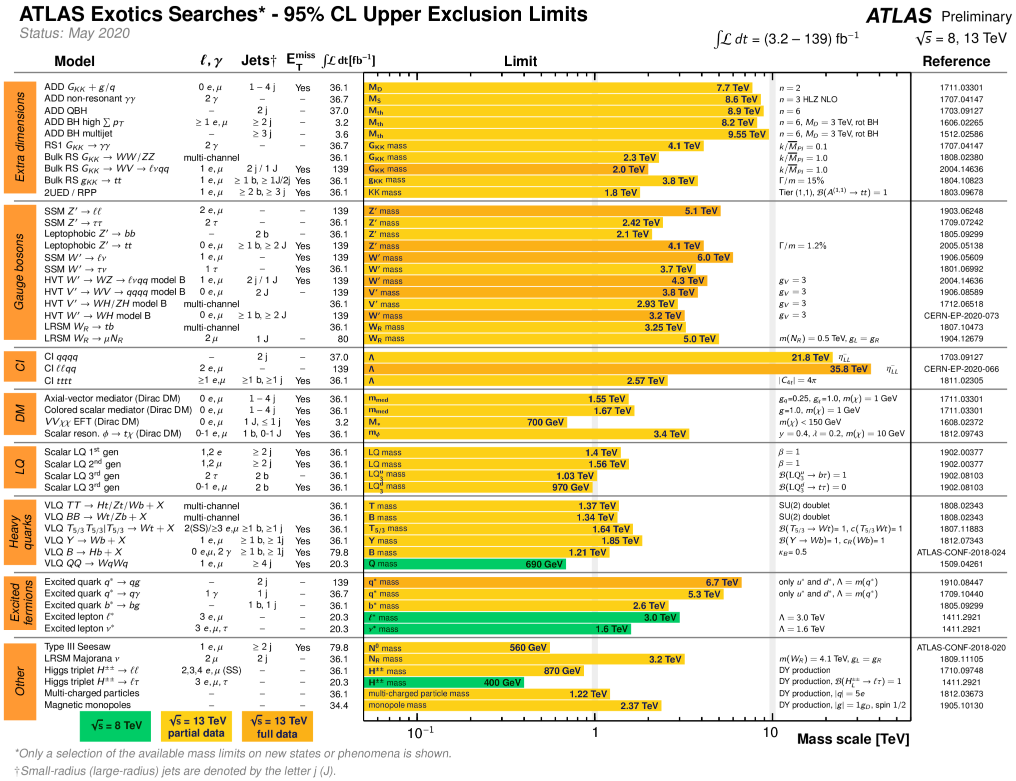


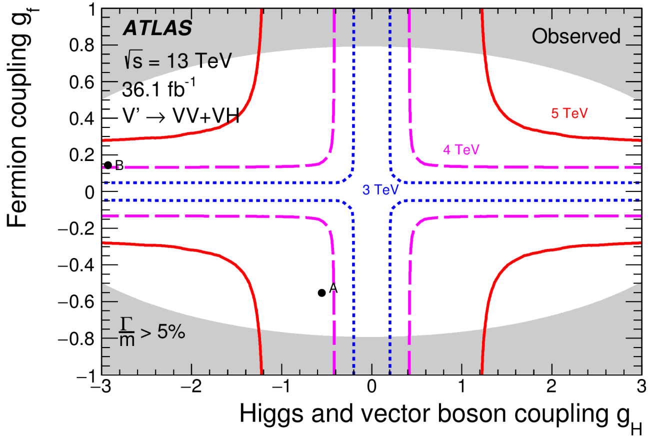


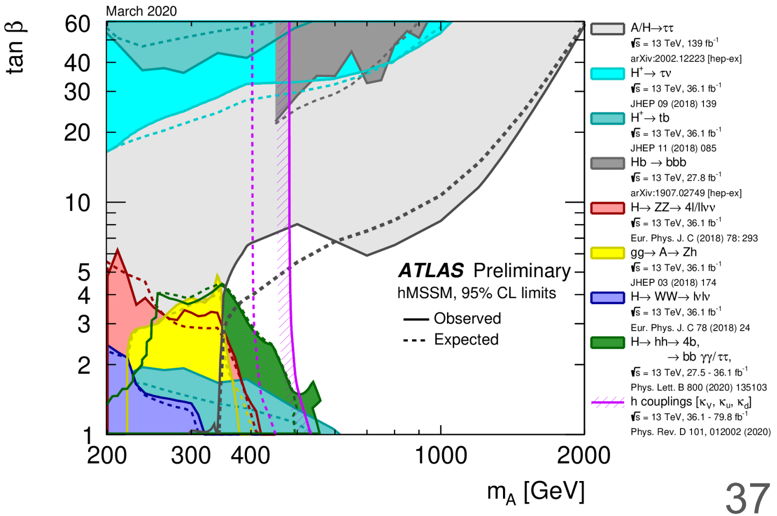
Hy



Towards a big picture

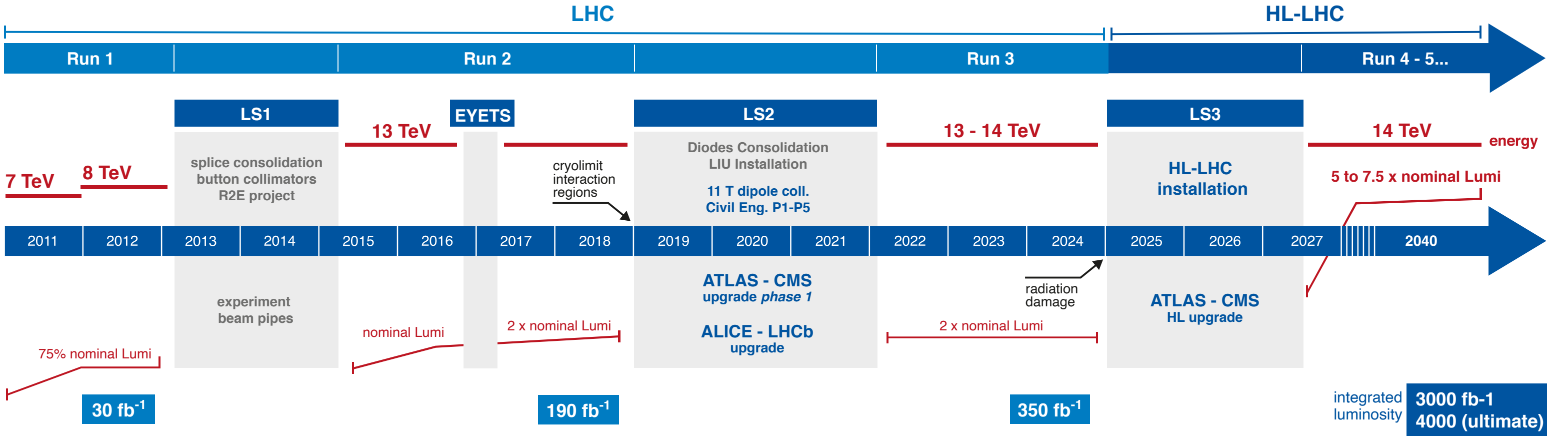






What next?

The LHC timeline



- The experiments are currently undergoing Phase-I of the planned upgrades in order to improve and/or maintain trigger rates and data taking capabilities.
- These are relatively early days: HL-LHC integrated luminosity goal of 3000-4000 fb^{-1}.
- But the energy reach won't increase significantly and it will take some time until we double the current integrated luminosity:
  - It is critical to keep developing new analysis ideas and methods to fully explore the Run 2 data.

New techniques

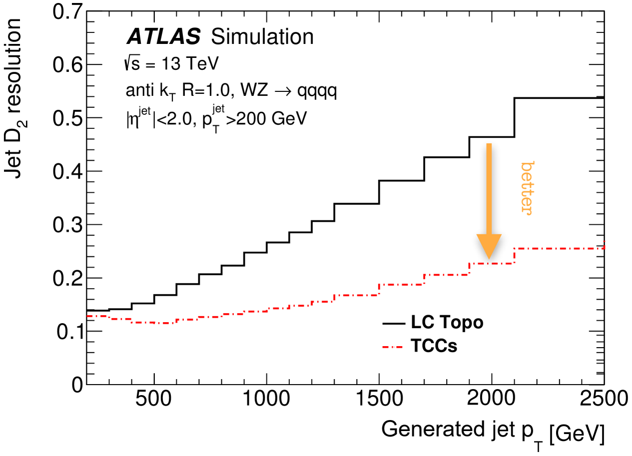
ATL-PHYS-PUB-2017-015

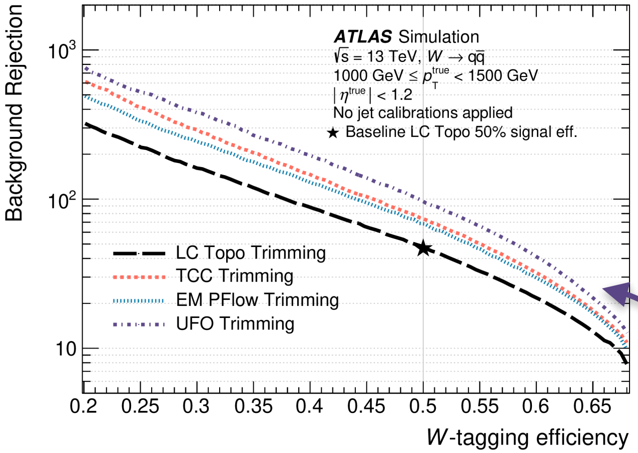
JETM-2018-06

Inputs to jet reconstruction

- Tracking information can be incorporated into jet substructure to benefit from better spatial resolution of the tracker in addition the excellent energy resolution of the calorimeter.

(TCCs: already used in VH,VV results presented)





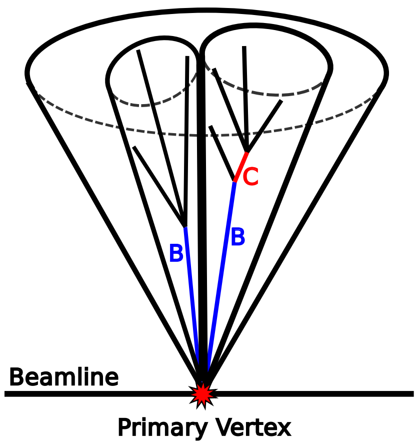
- New "Unified Flow Objects" will provide optimal performance across a wide kinematic range and in dense environments typical of high p_T jets.

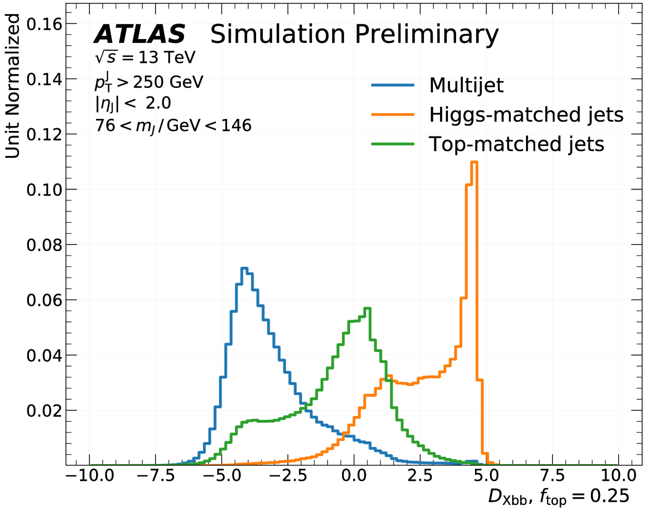
ATL-PHYS-PUB-2020-019

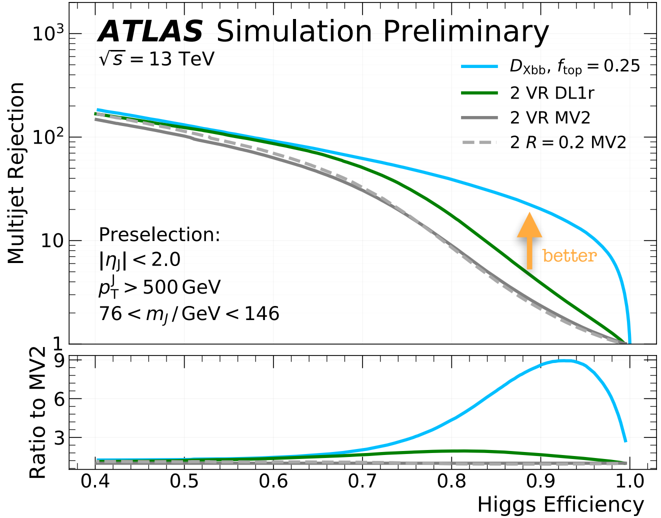
Dedicated Higgs boson taggers

H→bb

- Mature techniques in place for identifying individual b-jets using deep neural networks, optimised for vast kinematic range.
- Can take on step further and train classifier neural networks directly on boosted H→bb jets against main expected backgrounds (coming from top and QCD processes):
  - Exploiting flavour information correlations within Higgs candidate.





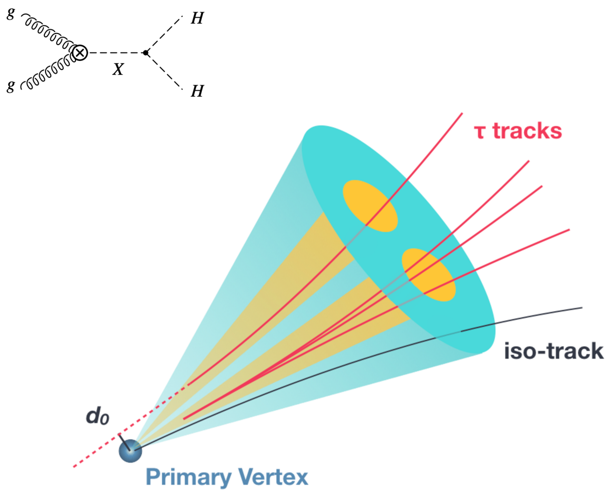


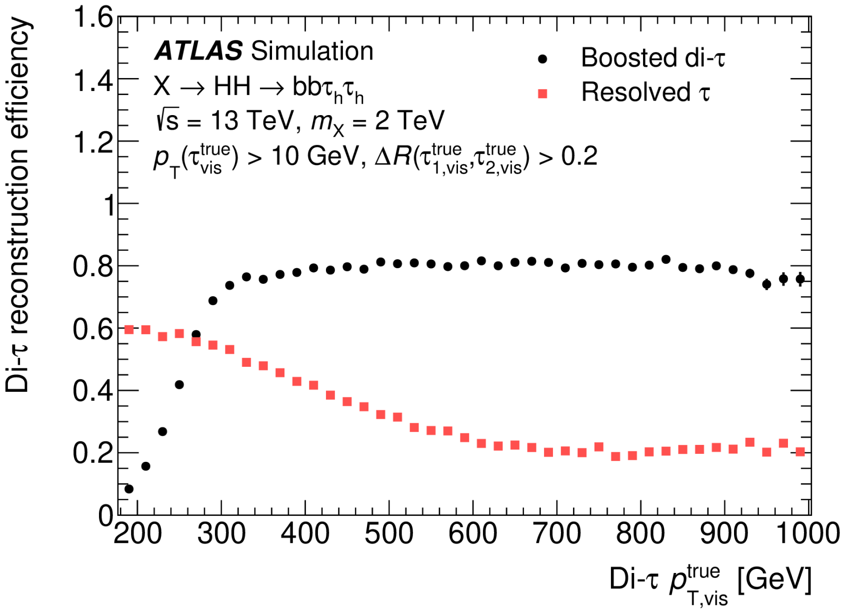
Dedicated Higgs boson taggers

$H \rightarrow \tau\tau$

2007.14811

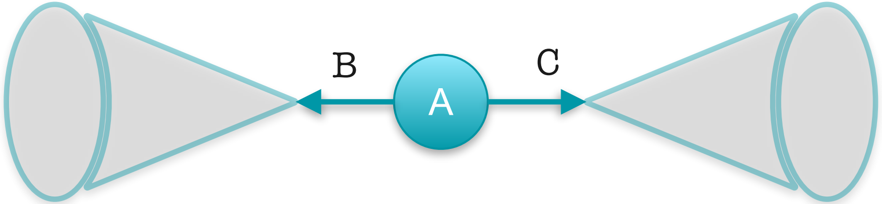
- Exploring calorimetric shower shapes and tracking information to distinguish between ττ pairs (when both τ-leptons decay hadronically) and QCD background.
  - BDT trained to classify Higgs to ττ jets against background dominated data.





Broadening the scope

Broadening the scope



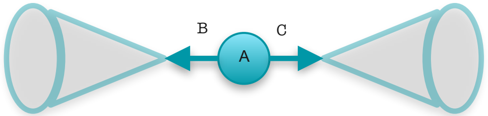
- What if we extend the search phase-space by not assuming Standard Model bosons?
- First exploration by ATLAS with 2015+2016 data:
  - $Y \rightarrow XH$ where X is unknown: assumption is that it decays to jets and has a mass in the range 50 GeV to O(1 TeV), Phys. Lett. B 779 (2018) 24.

- Can be taken further by searching for $A \rightarrow BC$ events with no assumption on either particle:
  - **A task for anomaly detection techniques!**

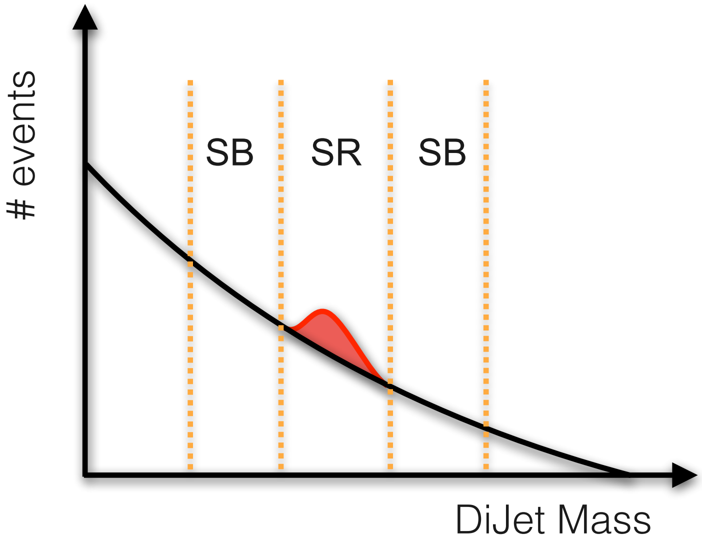
Generic search A→BC

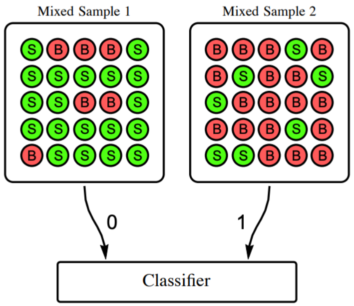
Weak supervision with CWoLa

[PhysRevLett.125.131801]



- Generic search for new resonances via anomaly detection procedure:
  - Suited for massive resonance decay with di-jet topology, using with large-radius jets.
- CWoLa method: Classification WithOut LAbels ([PhysRevLett.121.241803]).
  - Train neural networks to distinguish between signal region and sidebands in data.

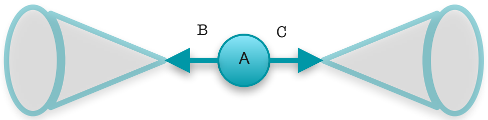




Generic search A→BC

Weak supervision with CWoLa

PhysRevLett.125.131801



- Generic search for new resonances via anomaly detection procedure:
  - Suited for massive resonance decay with di-jet topology, using with large-radius jets.
- CWoLa method: Classification WithOut LAbels (PhysRevLett.121.241803).
  - Train neural networks to distinguish between signal region and sidebands in data.

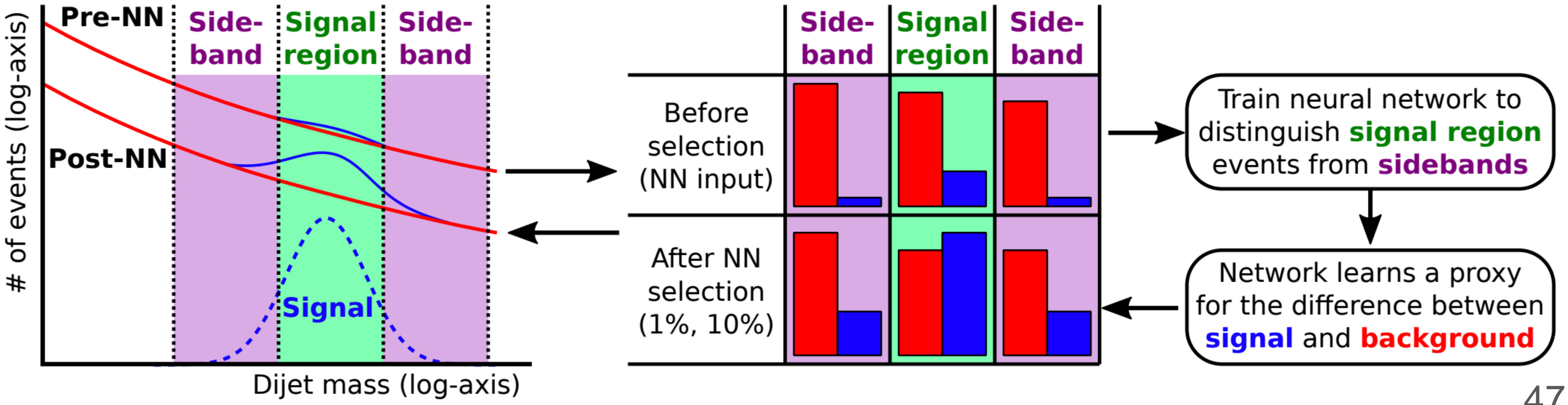
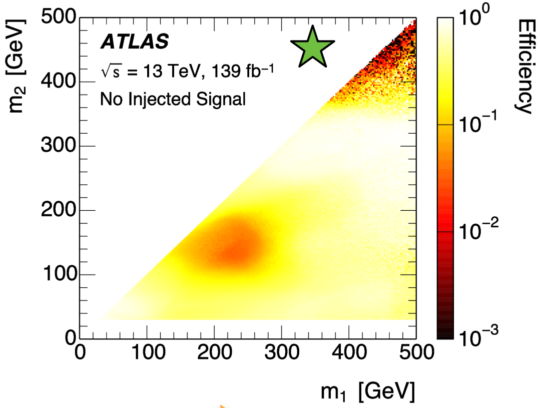
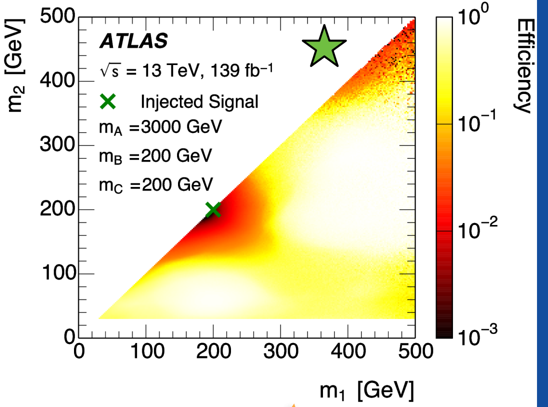


Diagram by S. Schramm, ICHEP2020

Generic search A→BC

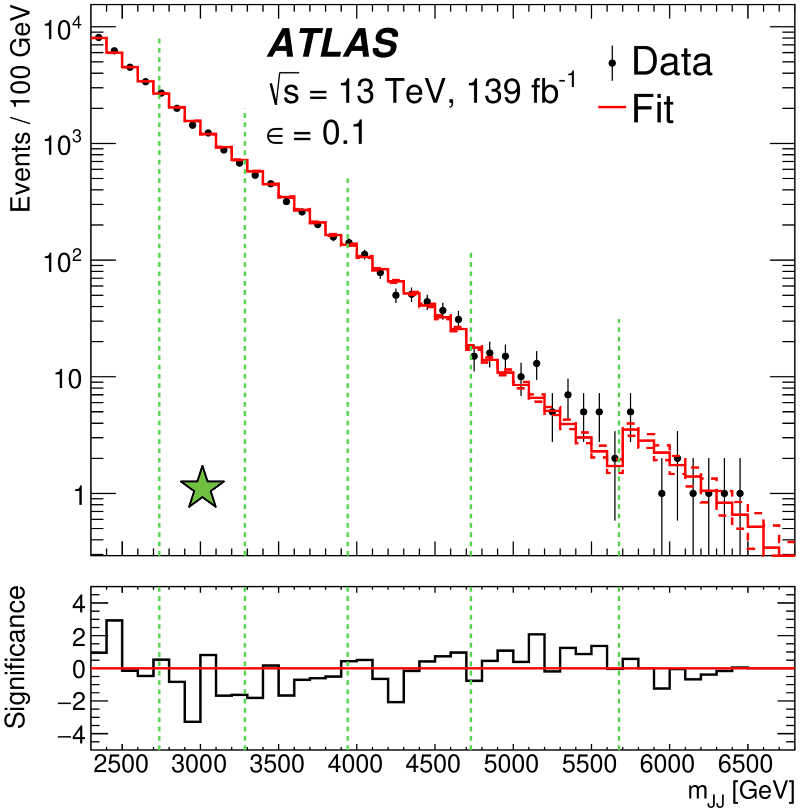
Method & Results





NN output training directly on data

NN output with injected signal at x



- NNs trained with data for each m_{JJ} signal region.
- Fits performed to a subset of the m_{JJ} distribution after cut on NN output at efficiencies of 1%, 10%.
- Model-independent: no evidence of excess.
- Model-dependent: exclusions covering new regions of phase-space.

Summary

Summary

- The LHC experiments are in a great position to tackle some of the questions left unanswered by the Standard Model.
- There is a plethora of experimental results placing constraints on SM extensions that predict diboson resonances.
  - Focus here was on fully-hadronic searches at very high masses, part of a broad and comprehensive search programme.
- Progress depends on the development of new analysis ideas and methods to keep exploring the LHC data for New Physics:
  - Improved “tagging” techniques to extract more sensitivity from the Run 2 dataset.
  - Anomaly detection for broadening the scope of analyses.
  - …a lot more that couldn’t fit in this talk.

Thank you for your attention!

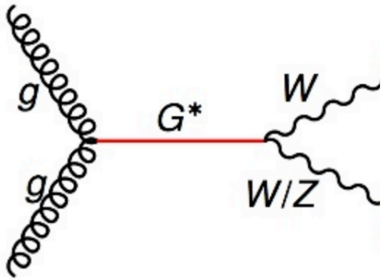
Backup

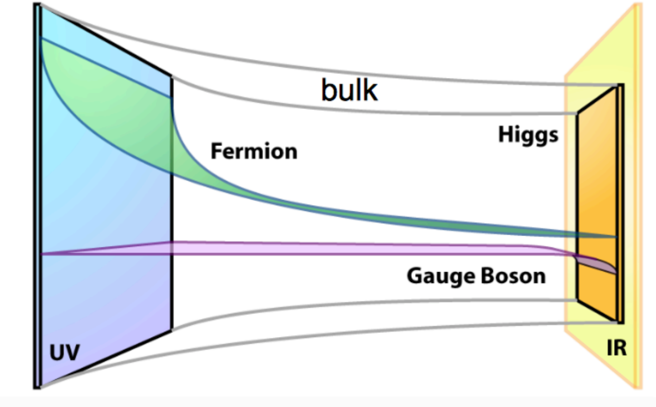
Theoretical scenarios (I)

PhysRevD.76.036006

Warped Extra Dimensions / bulk “RS" model

- Extension of Randall Sundrum models: gravity propagates in warped extra dimension.
  - The original RS model confines SM particles to a 4D brane, in the bulk RS model the SM particles extend into the “bulk”.
- The most distinctive feature of this scenario is the existence of spin-2 Kaluza-Klein (KK) gravitons whose masses and couplings to the SM are set by the TeV scale.
- Couplings to light fermions suppressed.
- Gluon-gluon fusion dominant production channel.





Theoretical scenarios (II)

JHEP09(2014)060

Heavy Vector Triplets

- Benchmark models are defined according to different parameter values.
- Model A: gV = 1: Extended gauge symmetry, with comparable branching ratios into bosons and fermions.
- Model B: gV = 3: Strongly coupled scenarios, suppressed branching ratios into fermions
- VBF model: Couplings to fermions set to zero, couplings to boson similar to Model A.
- For Models A and B, intrinsic width assumed much narrower than detector resolution: ~2.5%

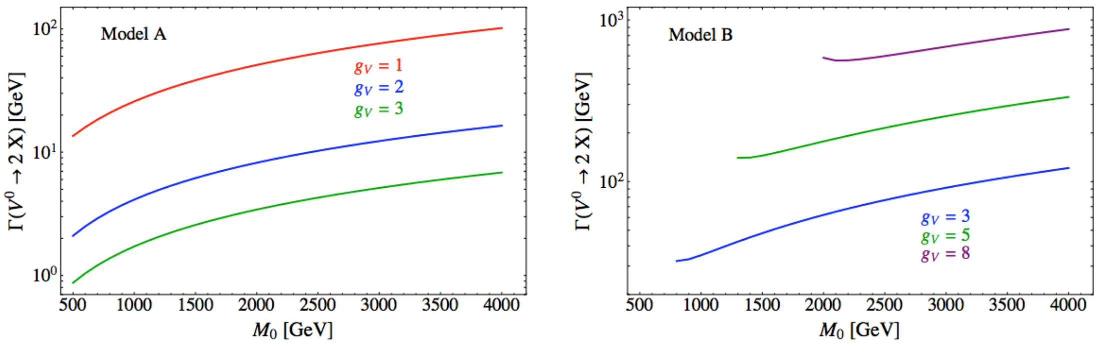


Figure 2.1: Upper panel: Branching Ratios for the two body decays of the neutral vector V^0 for the benchmarks $A_{g_V=1}$ (left) and $B_{g_V=3}$ (right). Lower panel: Total widths corresponding to different values of the coupling g_V in the models A (left) and B (right).

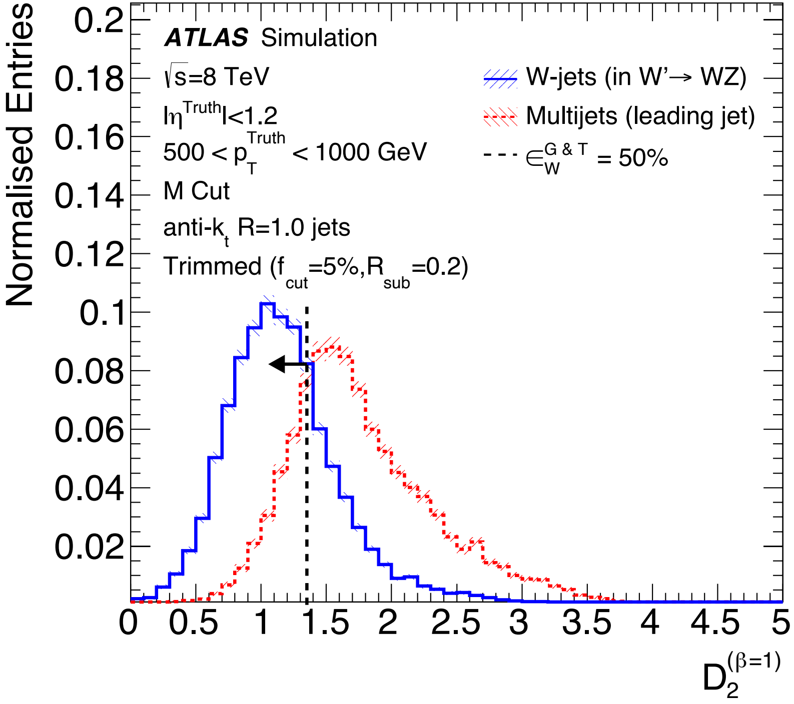
Jet substructure

- The $D_2^{\beta=1}$ variable is useful in identifying jets with two-prong substructures.
- Defined from n-point energy correlation functions:

$$E_{CF1}(\beta) = \sum_{i \in J} p_{T_i},$$

$$E_{CF2}(\beta) = \sum_{i<j \in J} p_{T_i} p_{T_j} \left(\Delta R_{ij}\right)^{\beta},$$

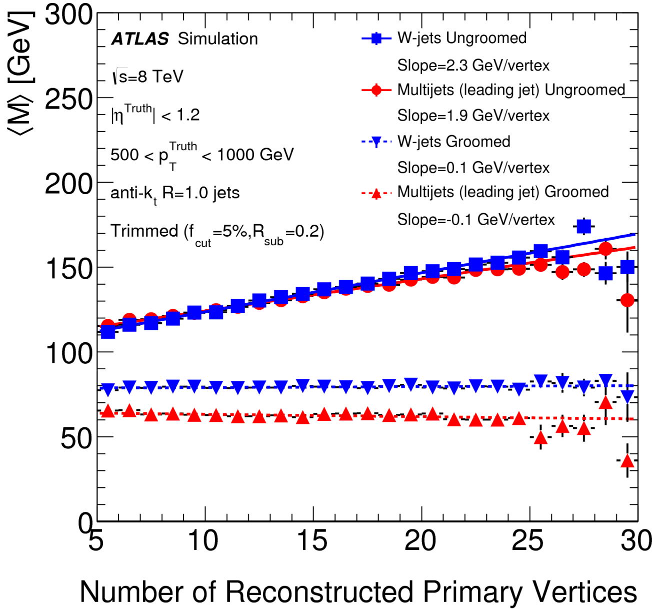
$$E_{CF3}(\beta) = \sum_{i<j<k \in J} p_{T_i} p_{T_j} p_{T_k} \left(\Delta R_{ij} \Delta R_{ik} \Delta R_{jk}\right)^{\beta}$$

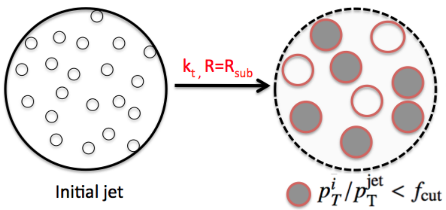


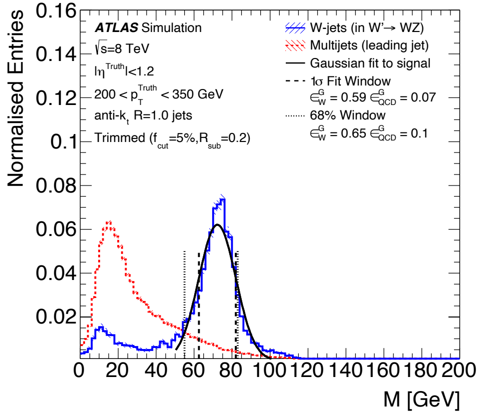
$$D_2^{\beta=1} = E_{CF3}\left(\frac{E_{CF1}}{E_{CF2}}\right)^3$$

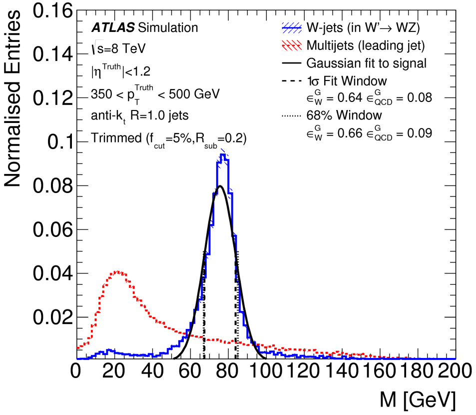
Jet mass

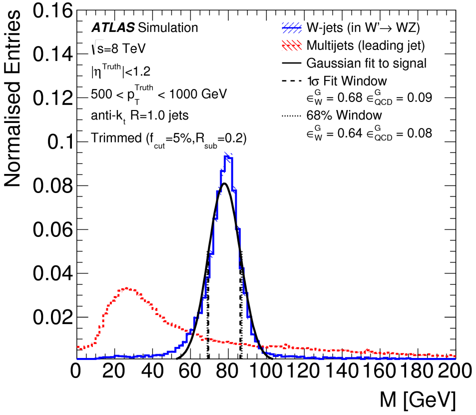
PERF-2015-03







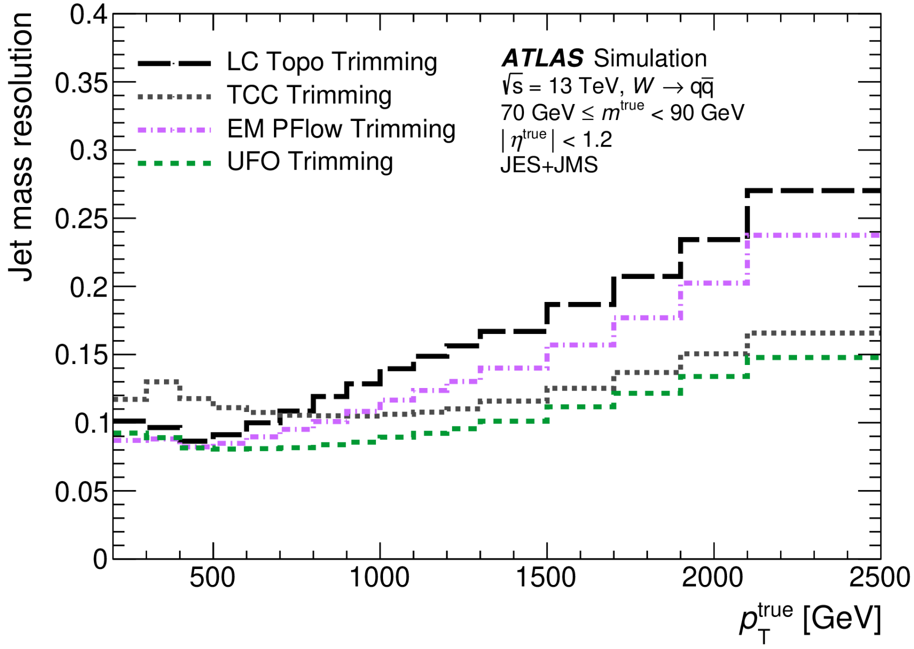




Jet mass

ATL-PHYS-PUB-2017-015

JETM-2018-06



Sensitivity based optimisation

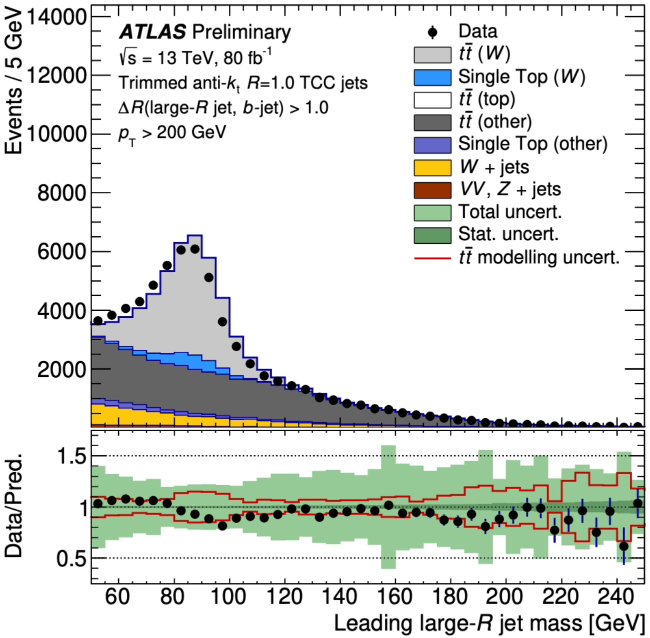
https://arxiv.org/pdf/physics/0308063.pdf

- Significance measure independent of cross-section of new processes:

$$\epsilon/(a/2 + \sqrt{B}),$$

ε = per signal jet / event efficiency
a = number of standard-deviations corresponding to a one-sided Gaussian distribution
B = number of background jets / events after selection

VV



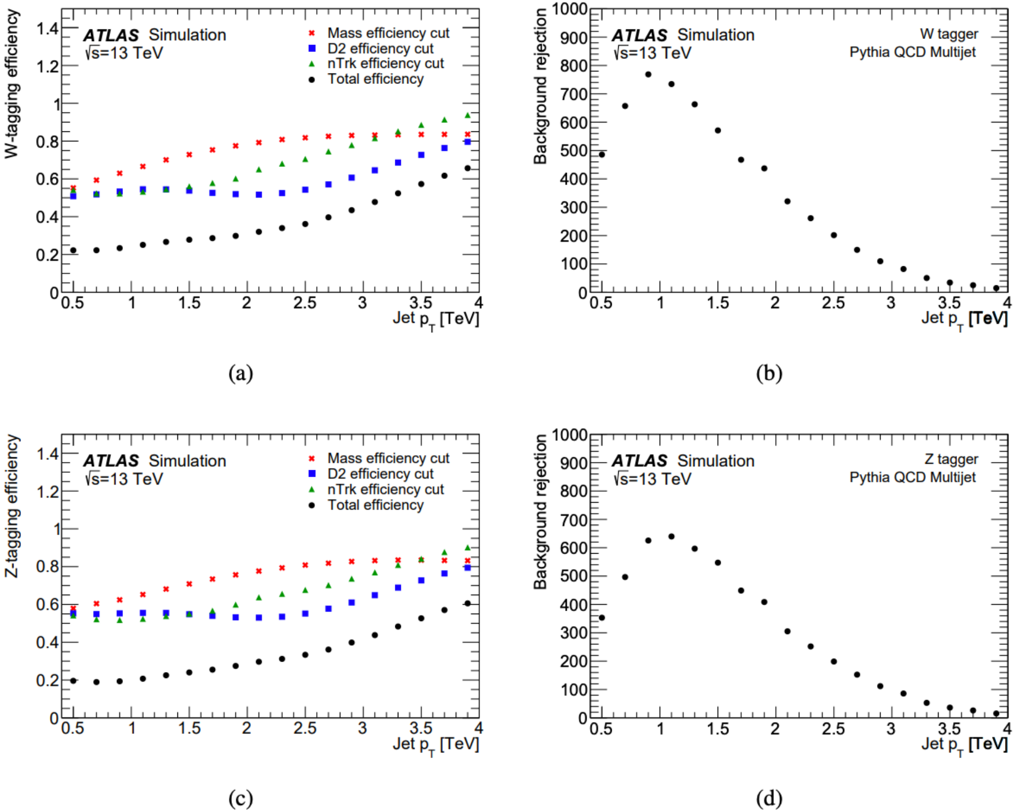
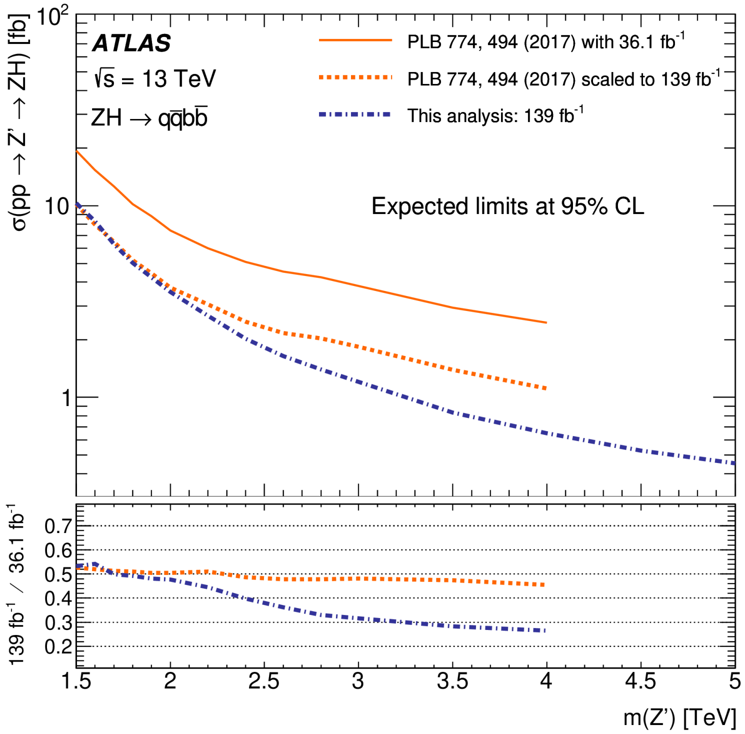
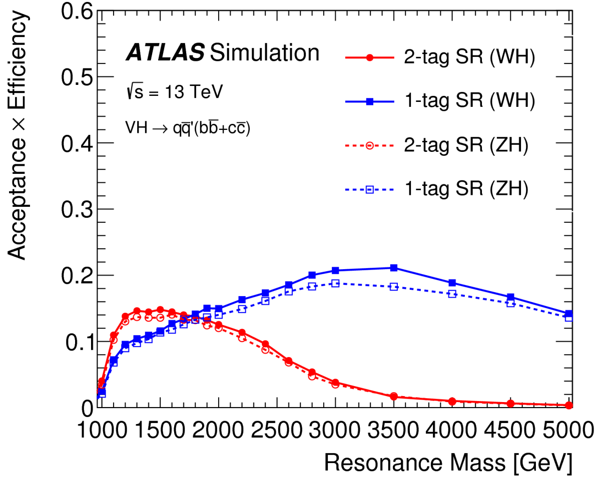


Figure 3: The (a) per-boson signal efficiency for the jet mass, D_2, and n_{trk} selections, as well as the combined efficiency and (b) background rejection (1/efficiency) of the W tagger for HVT $W' \to WZ \to qqqq$ and MC simulated multijets as a function of the jet p_T. Corresponding values for the Z tagger are shown in (c) and (d).

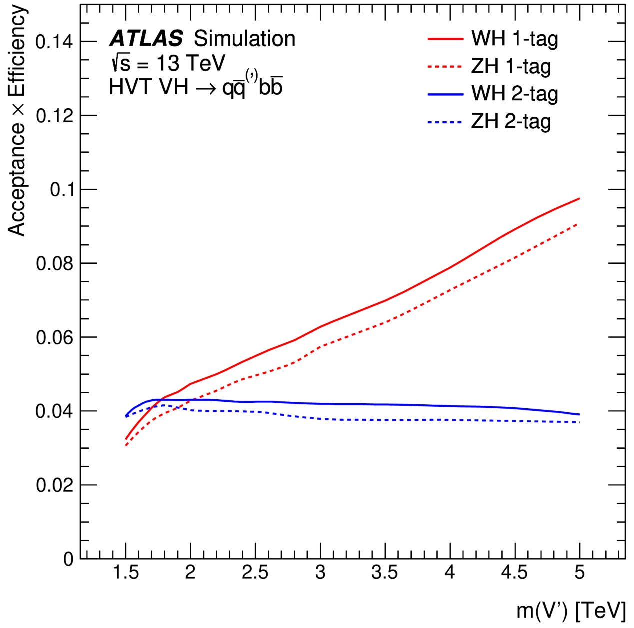
VHqqbb (I)

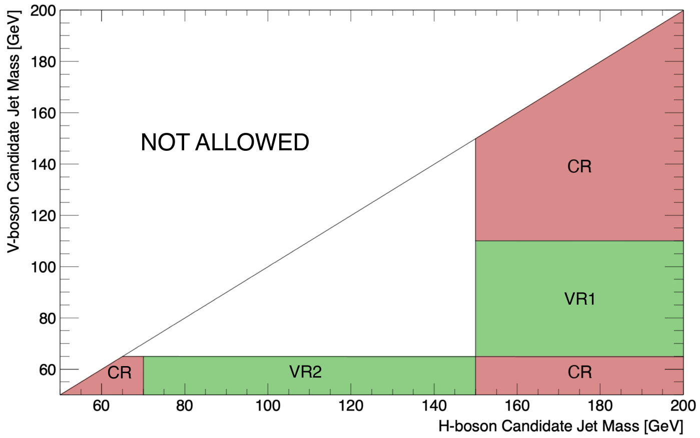


Previous publication with 36.1/fb

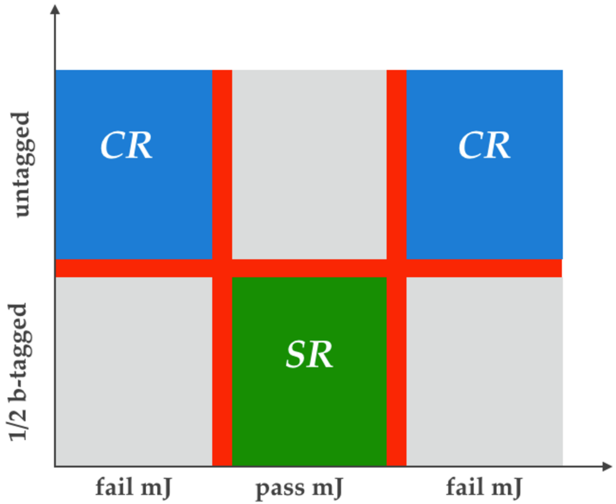


VHqqbb (II)

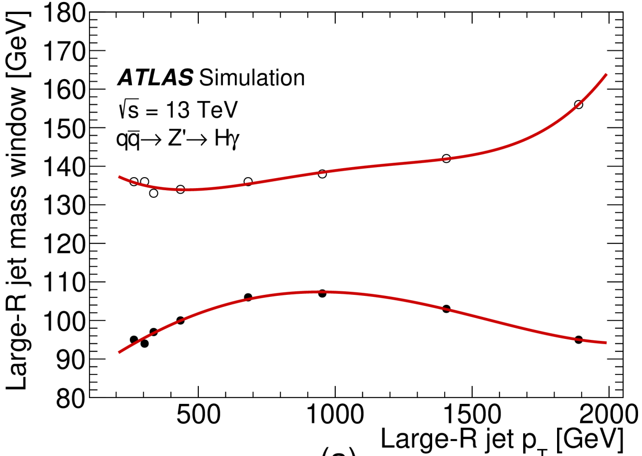

ATLAS Simulation
√s = 13 TeV
HVT VH → qq̄(')bb̄
WH 1-tag
ZH 1-tag
WH 2-tag
ZH 2-tag
Acceptance × Efficiency
m(V') [TeV]


NOT ALLOWED
CR
VR1
CR
CR
VR2
V-boson Candidate Jet Mass [GeV]
H-boson Candidate Jet Mass [GeV]

Hy

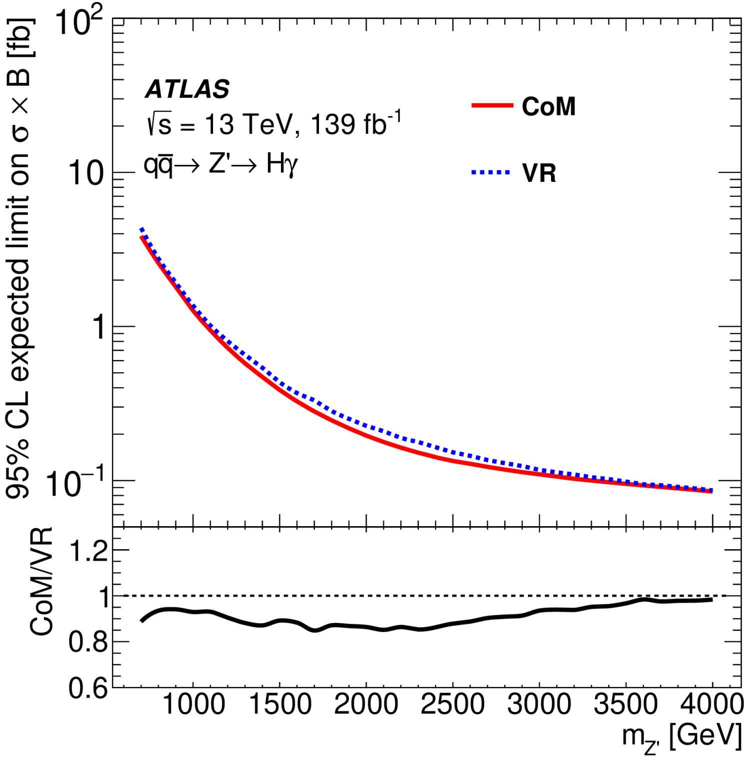


CR
CR
SR
untagged
1/2 b-tagged
fail mJ
pass mJ
fail mJ



ATLAS Simulation
√s = 13 TeV
qq̄→ Z'→ Hγ
Large-R jet mass window [GeV]
Large-R jet p_T [GeV]

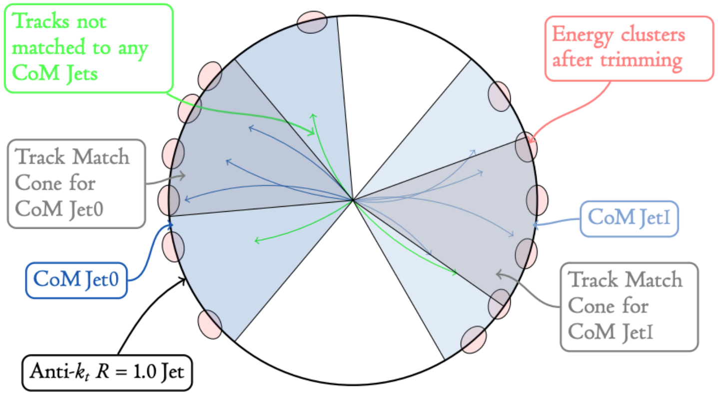
(a)

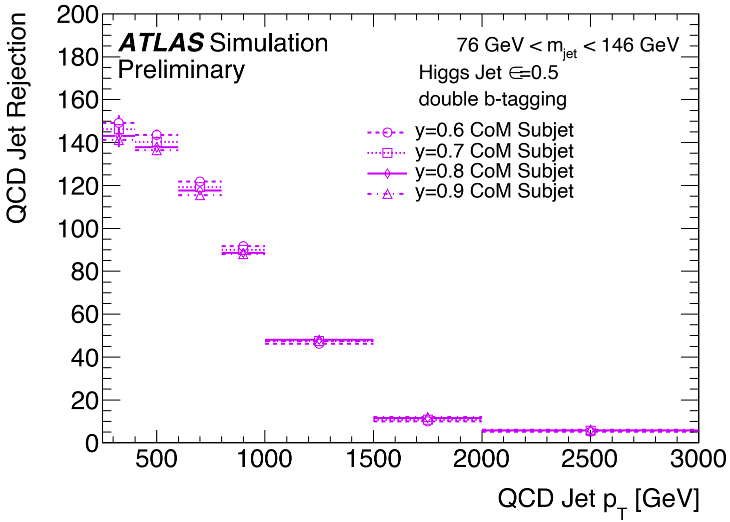


ATLAS
√s = 13 TeV, 139 fb-1
qq̄→ Z'→ Hγ
CoM
VR
95% CL expected limit on σ × B [fb]
CoM/VR
m_Z' [GeV]

Exclusive centre-of-mass (CoM)

- EECambridge algorithm on calorimeter constituents after boost to large-R jet rest frame, run in exclusive mode with number of subjets N_{sj}=2.
    - Based on angular separation $y_{ij}=2(1-\cos\theta_{ij})$.
- Track-to-subjet association is also based on their angular separation in the CoM frame (contrast with dR association used for other algorithms)
    - Background rejection studied for different values of y_{cut}.





Dedicated Higgs boson taggers

H→ττ

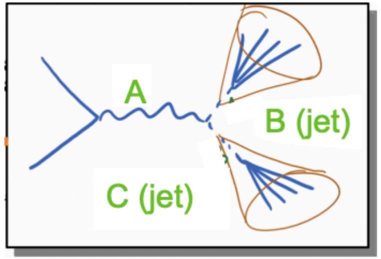
2007.14811

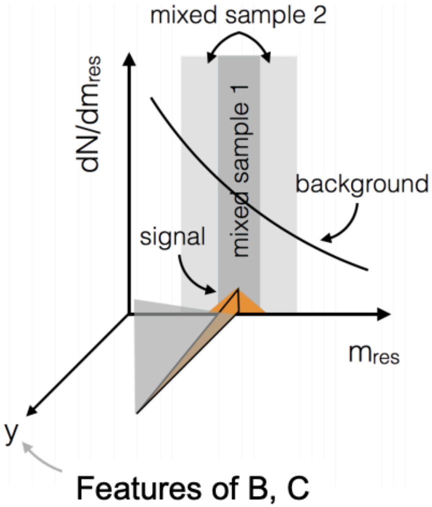
Table 1: Discriminating variables used in the di-τ identification BDT, aimed at rejecting the background from quark- and gluon-initiated jets. Here, LRJ refers to the seeding large-radius jet of the di-τ object, sj_1 and sj_2 stand for the first and second sub-jets ordered in p_T, respectively, and tracks refer to those matched to a sub-jet (τ track), unless specified otherwise.

Variable	Definition
$E^{sj_1}_{\Delta R<0.1}/E^{sj_1}_{\Delta R<0.2}$ and $E^{sj_2}_{\Delta R<0.1}/E^{sj_2}_{\Delta R<0.2}$	Ratios of the energy deposited in the core to that in the full cone, for the sub-jets sj_1 and sj_2, respectively
$p_T^{sj_2}/p_T^{LRJ}$ and $(p_T^{sj_1}+p_T^{sj_2})/p_T^{LRJ}$	Ratio of the p_T of sj_2 to the di-τ seeding large-radius jet p_T and ratio of the scalar p_T sum of the two leading sub-jets to the di-τ seeding large-radius jet p_T, respectively
$\log(\sum p_T^{\text{iso-tracks}}/p_T^{LRJ})$	Logarithm of the ratio of the scalar p_T sum of the iso-tracks to the di-τ seeding large-radius jet p_T
$\Delta R_{max}(\text{track}, sj_1)$ and $\Delta R_{max}(\text{track}, sj_2)$	Largest separation of a track from its associated sub-jet axis, for the sub-jets sj_1 and sj_2, respectively
$\sum[p_T^{\text{track}}\Delta R(\text{track}, sj_2)]/\sum p_T^{\text{track}}$	p_T-weighted ΔR of the tracks matched to sj_2 with respect to its axis
$\sum[p_T^{\text{iso-track}}\Delta R(\text{iso-track}, sj)]/\sum p_T^{\text{iso-track}}$	p_T-weighted sum of ΔR between iso-tracks and the nearest sub-jet axis
$\log(m^{\text{tracks}, sj_1}_{\Delta R<0.1})$ and $\log(m^{\text{tracks}, sj_2}_{\Delta R<0.1})$	Logarithms of the invariant mass of the tracks in the core of sj_1 and sj_2, respectively
$\log(m^{\text{tracks}, sj_1}_{\Delta R<0.2})$ and $\log(m^{\text{tracks}, sj_2}_{\Delta R<0.2})$	Logarithms of the invariant mass of the tracks with $\Delta R < 0.2$ from the axis of sj_1 and sj_2, respectively
$\log(\lvert d^{sj_1}_{0,\text{lead-track}}\rvert)$ and $\log(\lvert d^{sj_2}_{0,\text{lead-track}}\rvert)$	Logarithms of the closest distance in the transverse plane between the primary vertex and the leading track of sj_1 and sj_2, respectively
$n^{sj_1}_{\text{tracks}}$ and $n^{\text{sub-jets}}_{\text{tracks}}$	Number of tracks matched to sj_1 and to all sub-jets, respectively

A→BC with weak supervision

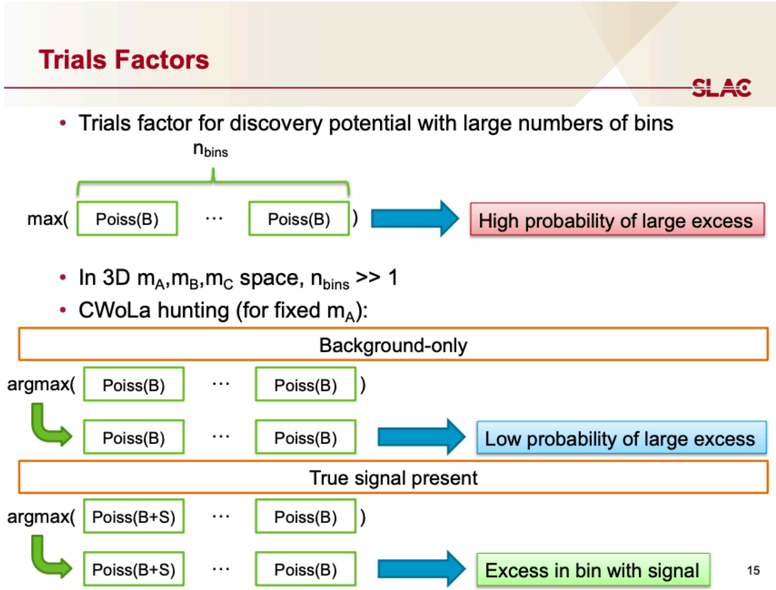
Non-ATLAS figures taken from CERN seminar by Aviv Cukierman.

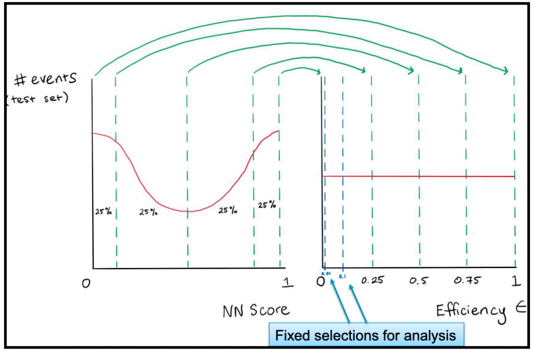


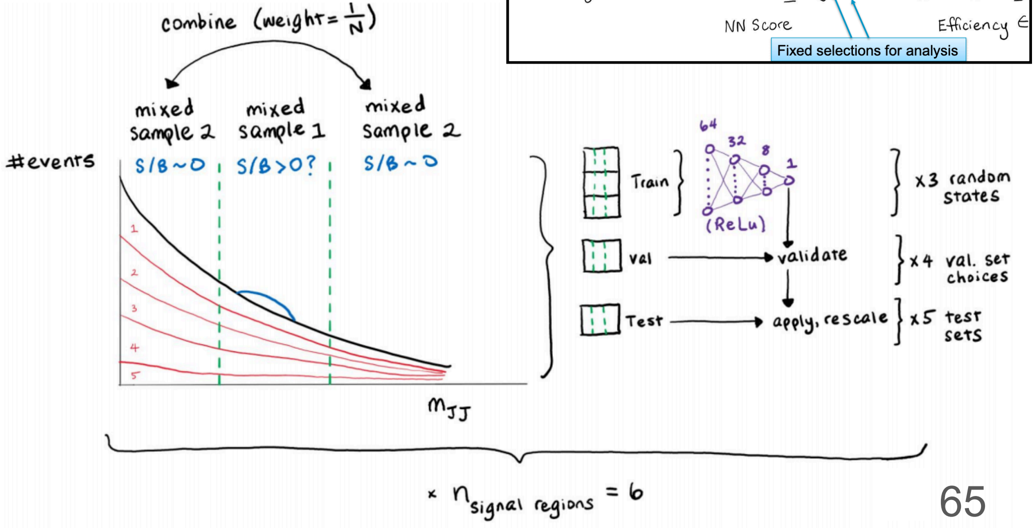


A→BC with weak supervision

Non-ATLAS figures taken from CERN seminar by Aviv Cukierman.







A→BC with weak supervision