Search for heavy top-philic resonances with the ATLAS detector

Elise Le Boulicaut Ennis (Duke University)

US LUA Annual Meeting, Fermilab

Lightning Round

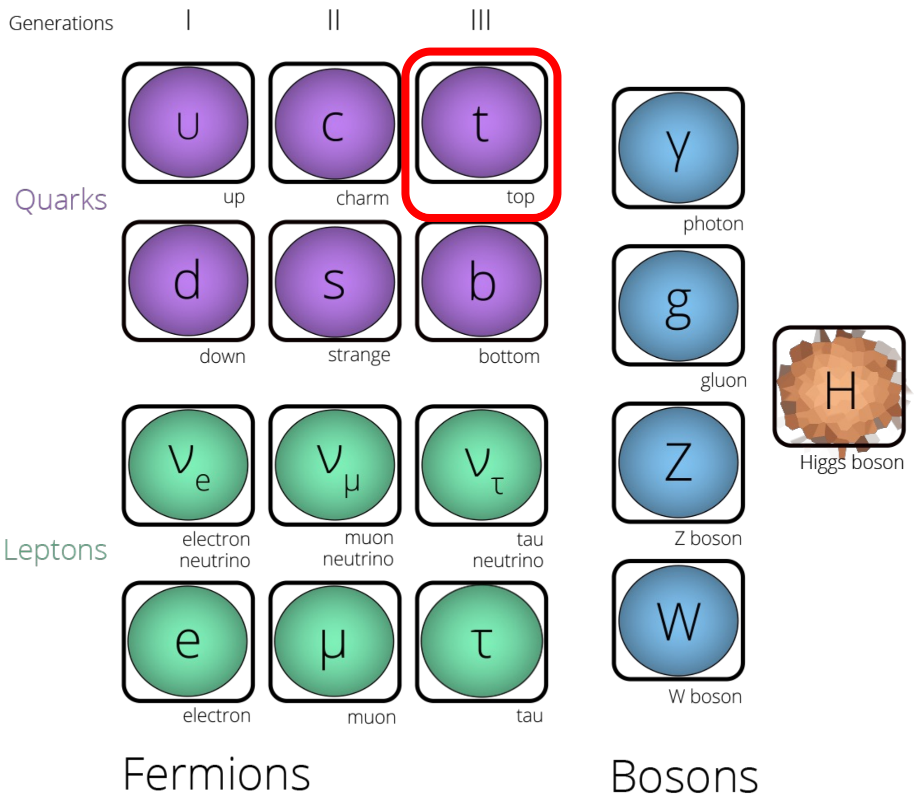
12/14/2023

arXiv:2304.01678

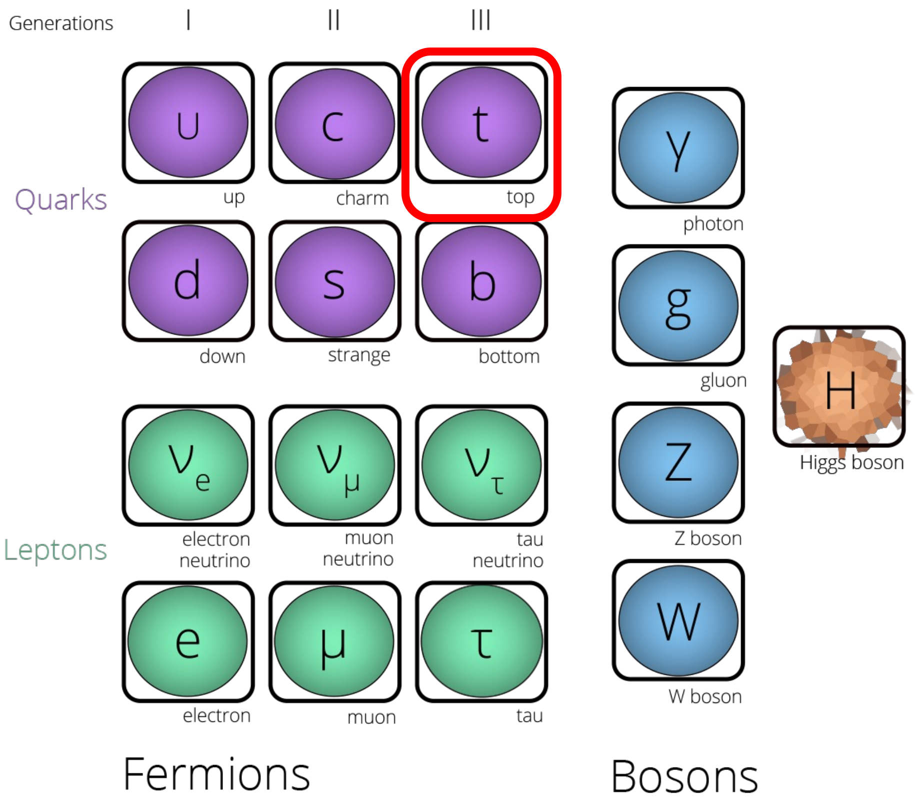
Accepted to EPJC

The top quark as a probe for new physics

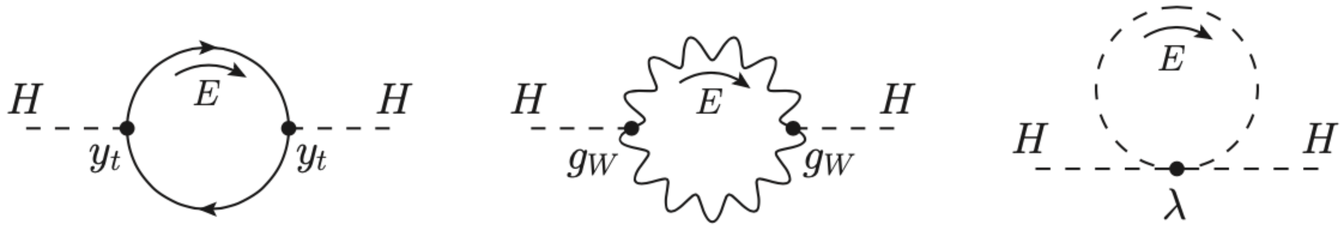
ATLAS EXPERIMENT | Duke UNIVERSITY



See [here](#) for a "cheat sheet" on the SM and other topics



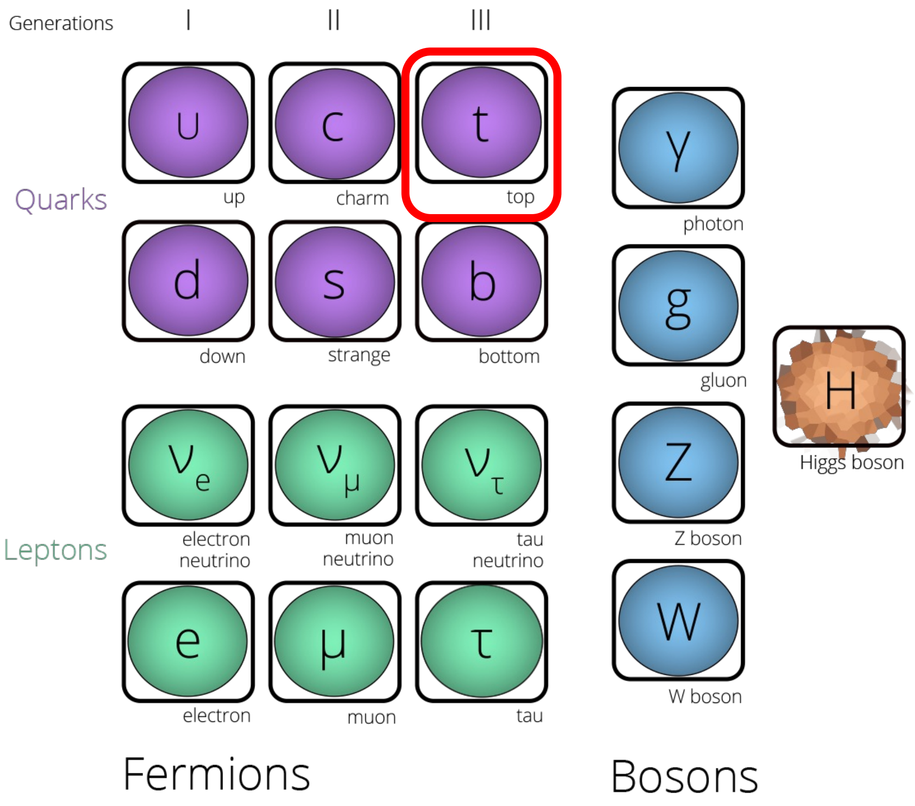
Naturalness problem: need to cancel loops that contribute to Higgs mass
→ new resonances could cancel them "naturally".



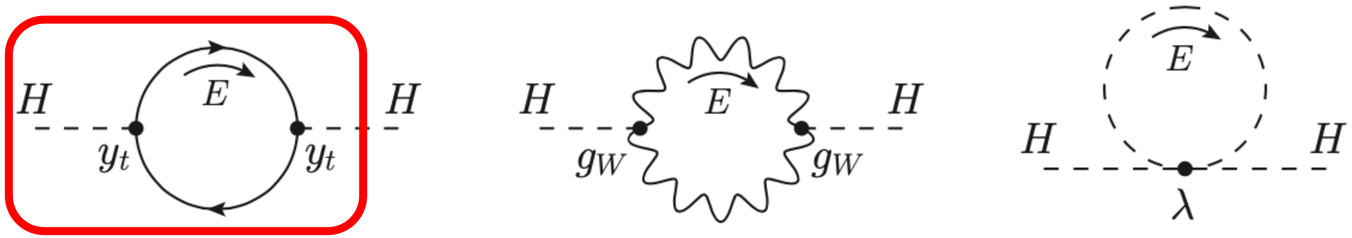
See here for a "cheat sheet" on the SM and other topics

Diagram from The Composite Nambu-Goldstone Higgs by G. Panico and A. Wulzer

The top quark as a probe for new physics



Naturalness problem: need to cancel loops that contribute to Higgs mass
→ new resonances could cancel them "naturally".



Top quarks contribute a lot to these loops
→ new resonances could be top-philic

See [here] for a "cheat sheet" on the SM and other topics

Diagram from [The Composite Nambu-Goldstone Higgs] by G. Panico and A. Wulzer

Signal model

Simplified model of color-singlet vector (spin-1) boson Z' coupling exclusively to top quarks:

$c_{L/R}$ = coupling to left/right-handed tops

$c_t = \sqrt{c_L^2 + c_R^2}$

$tan\theta = c_R/c_L$

$$\mathcal{L} = \bar{t}\gamma_\mu(c_L P_L + c_R P_R)tZ'^\mu$$

$$\mathcal{L} = c_t\bar{t}\gamma_\mu(\cos\theta\, P_L + \sin\theta\, P_R)tZ'^\mu$$

*Equation from J. High Energ. Phys. **2015**, 29 (2015)*

Signal model

Simplified model of color-singlet vector (spin-1) boson Z′ coupling exclusively to top quarks:

$c_{L/R}$ = coupling to left/right-handed tops

$c_t = \sqrt{c_L^2 + c_R^2}$

$tan\theta = c_R / c_L$

$$\mathcal{L} = \bar{t}\gamma_\mu (c_L P_L + c_R P_R) t Z'^\mu$$

$$\mathcal{L} = c_t \bar{t}\gamma_\mu (\cos\theta\, P_L + \sin\theta\, P_R) t Z'^\mu$$

Free parameters:

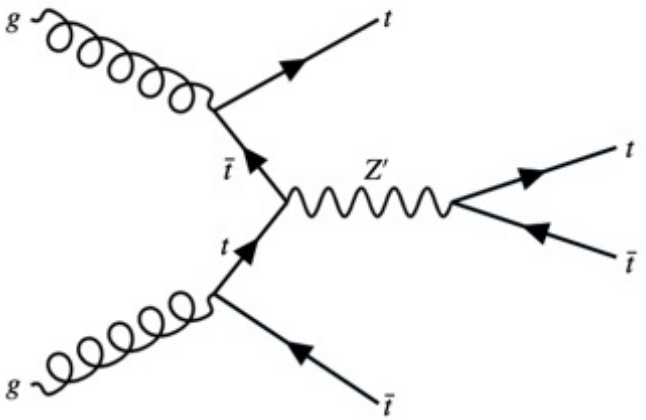
- **Mass** of the resonance $m_{Z'}$
- **Coupling** c_t: related to the width by $\frac{\Gamma}{m_{Z'}} \approx \frac{c_t^2}{8\pi}$
- **Chirality** angle θ

*Equation from J. High Energ. Phys. **2015**, 29 (2015)*

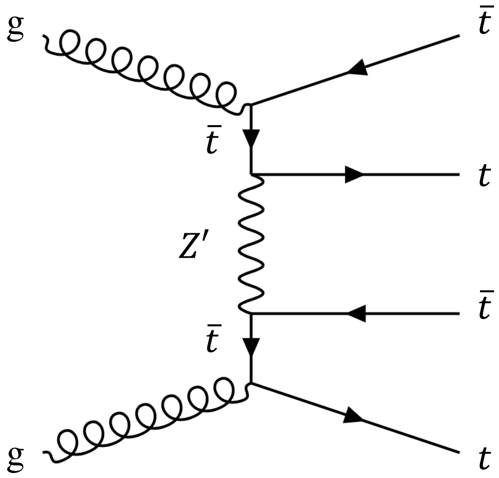
Production modes

$t\bar{t}Z'$

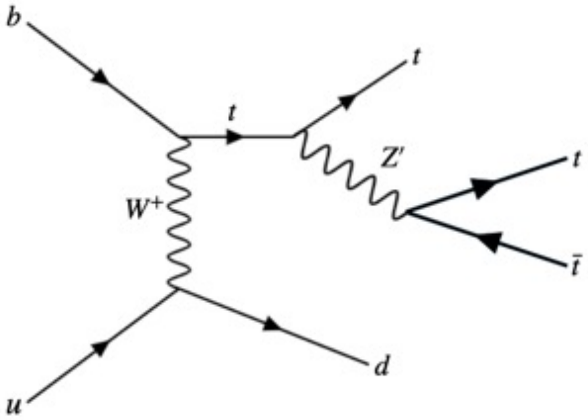
s-channel



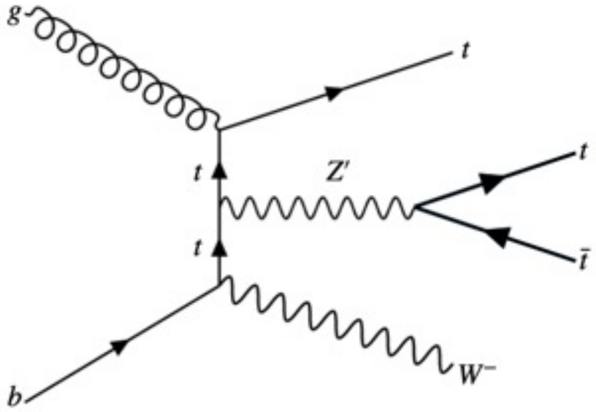
t-channel



tjZ'



tWZ'

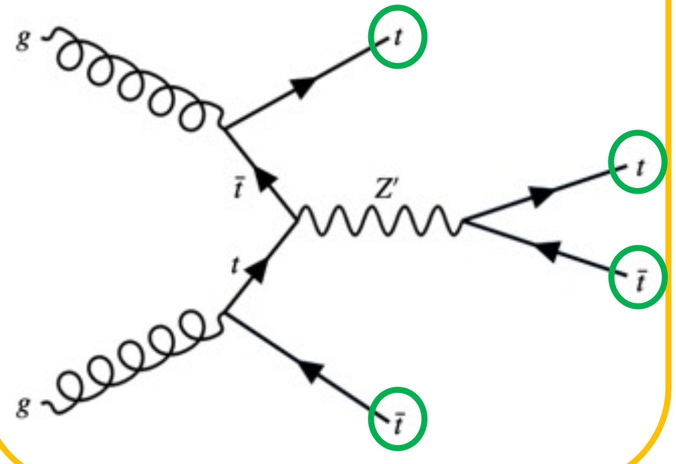


Production modes

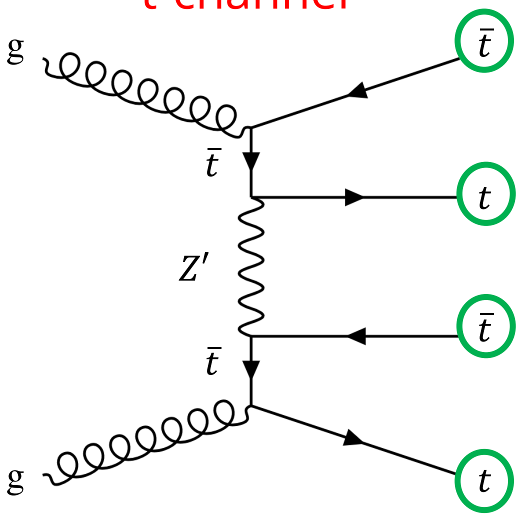
Main target

$t\bar{t}Z'$

s-channel

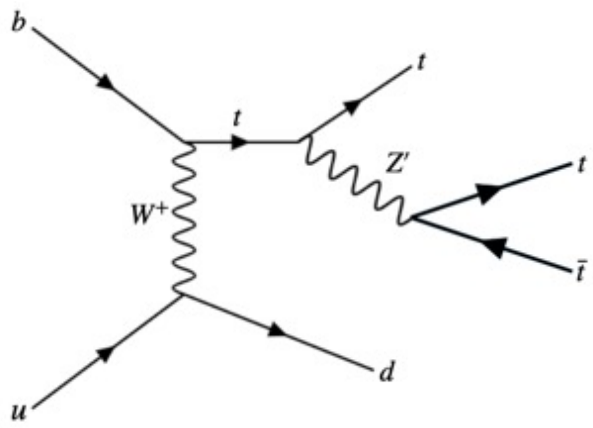


t-channel

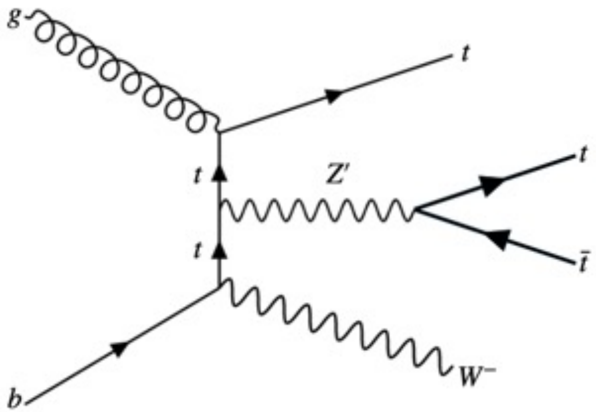


4 top final state

tjZ'



tWZ'

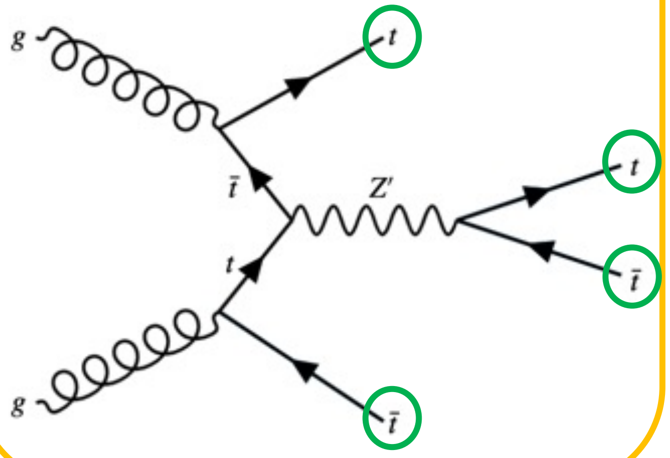


Production modes

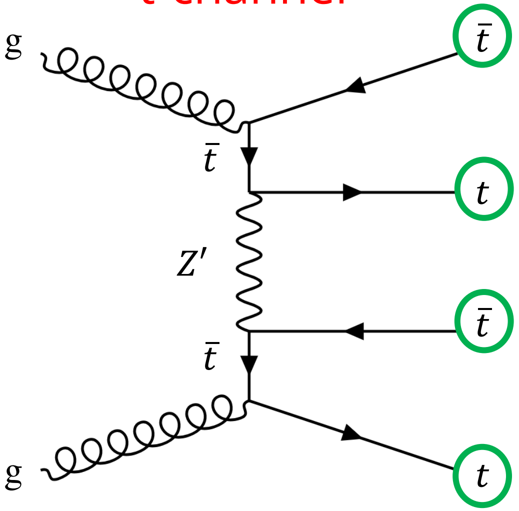
Main target

$t\bar{t}Z'$

s-channel

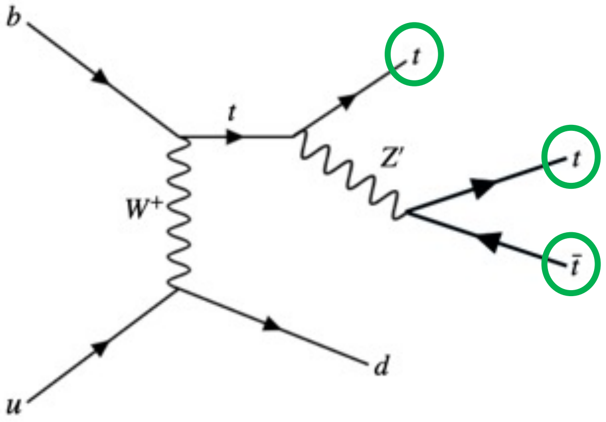


t-channel

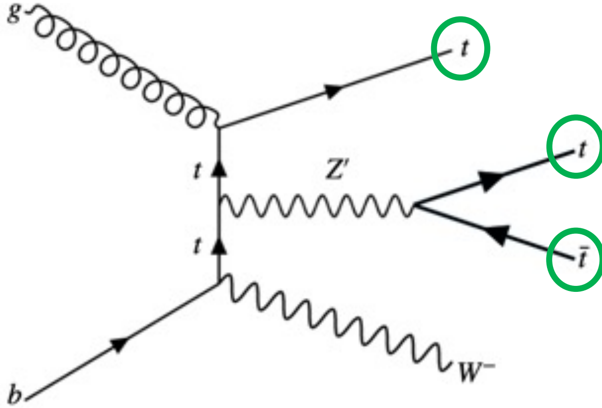


4 top final state

tjZ'



tWZ'



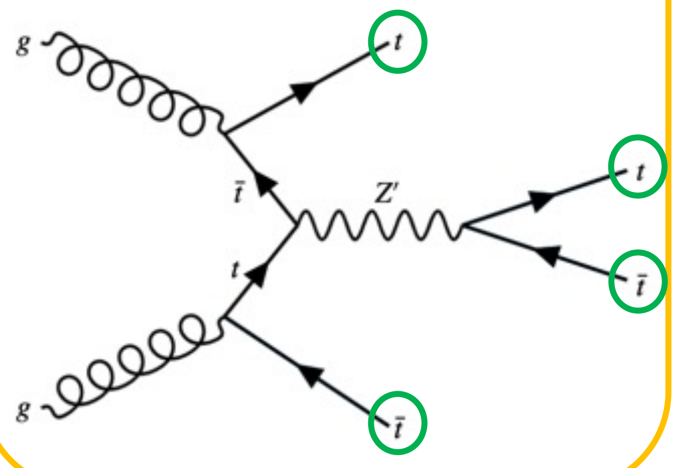
3 top final state

Production modes

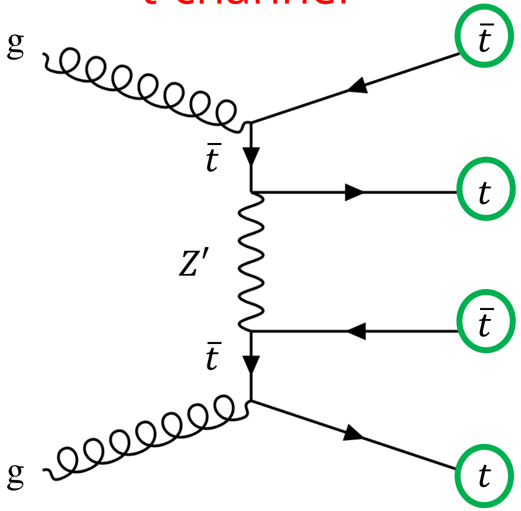
$t\bar{t}Z'$ | tjZ' | tWZ'

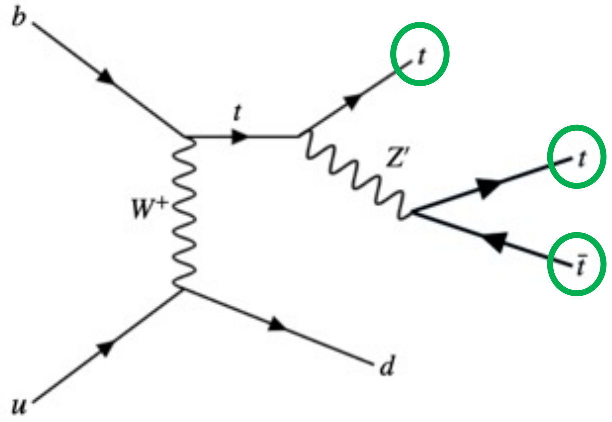
Main target

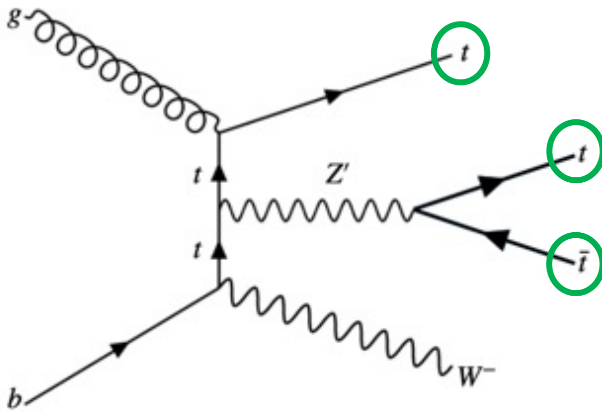
s-channel



t-channel







4 top final state

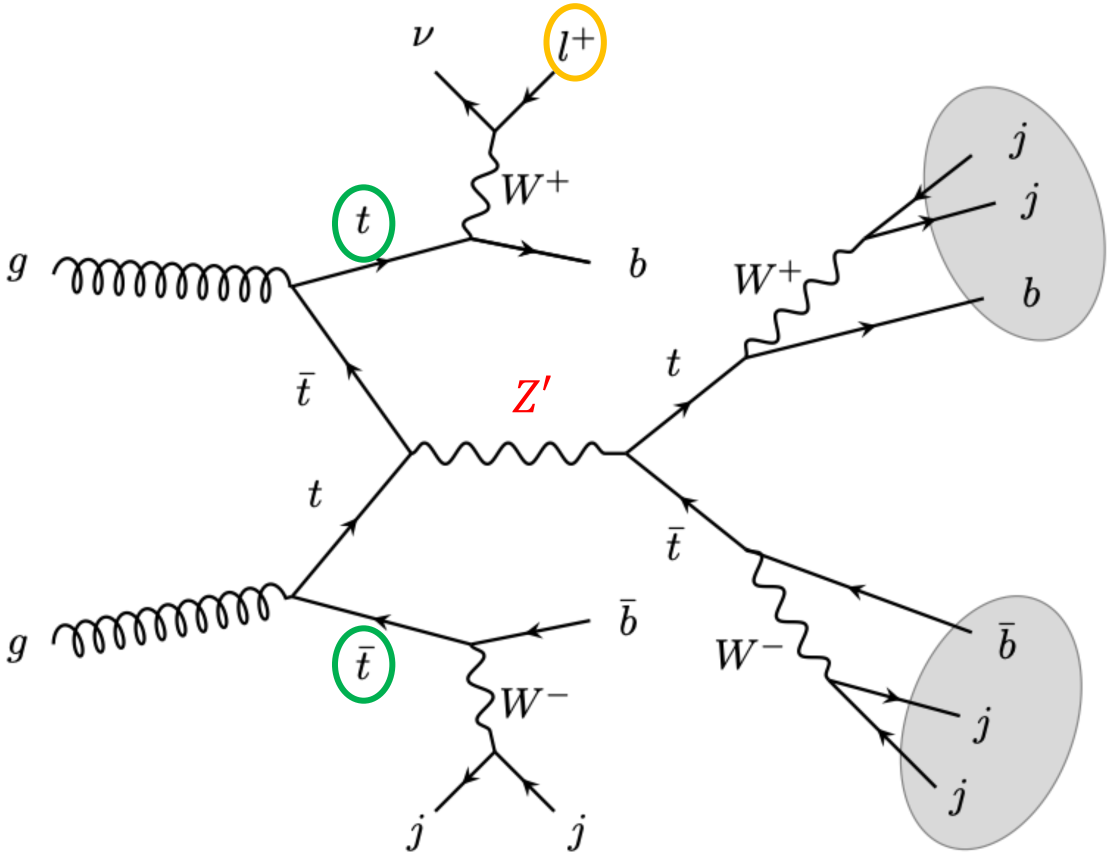
3 top final state

Cross section independent of θ

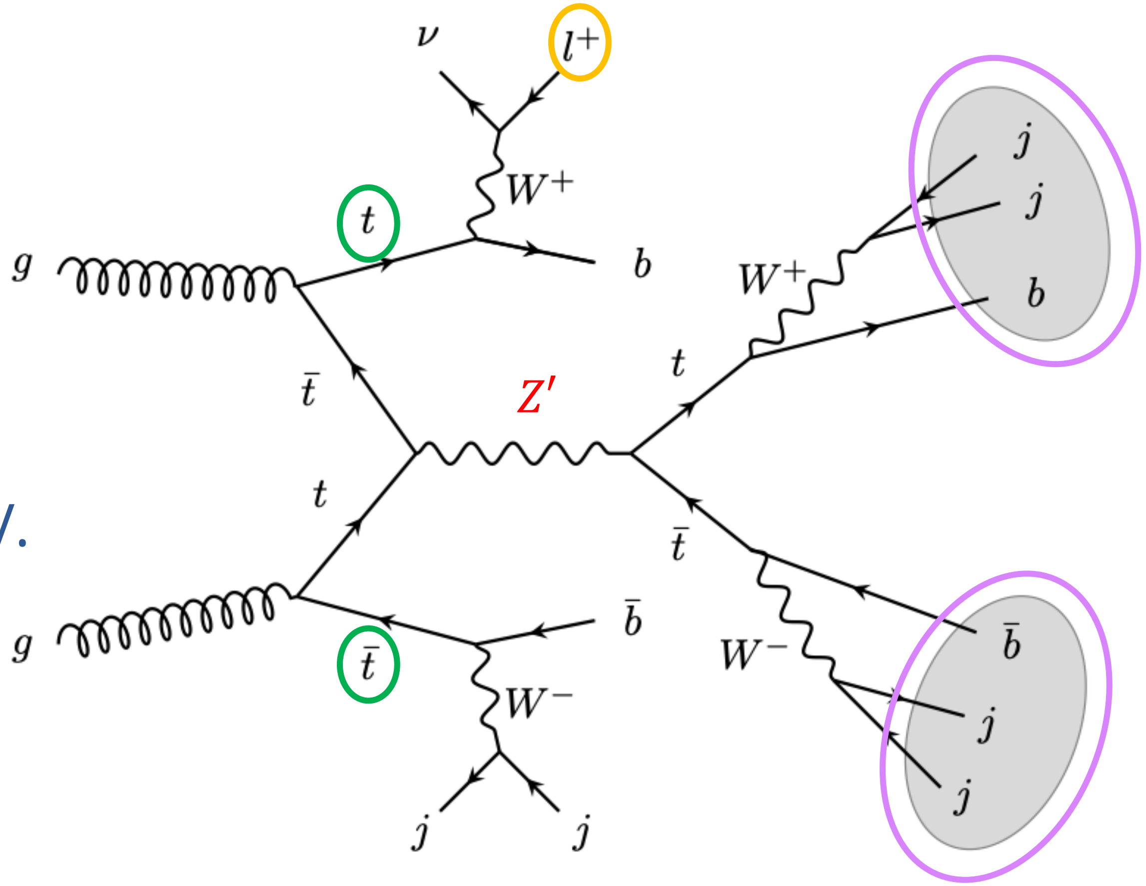
Cross section min when $\theta = \frac{\pi}{2}$, max when $\theta = 0$

Reconstruction & event selection

- Target events with **exactly one lepton**, preferably coming from a "spectator" top.

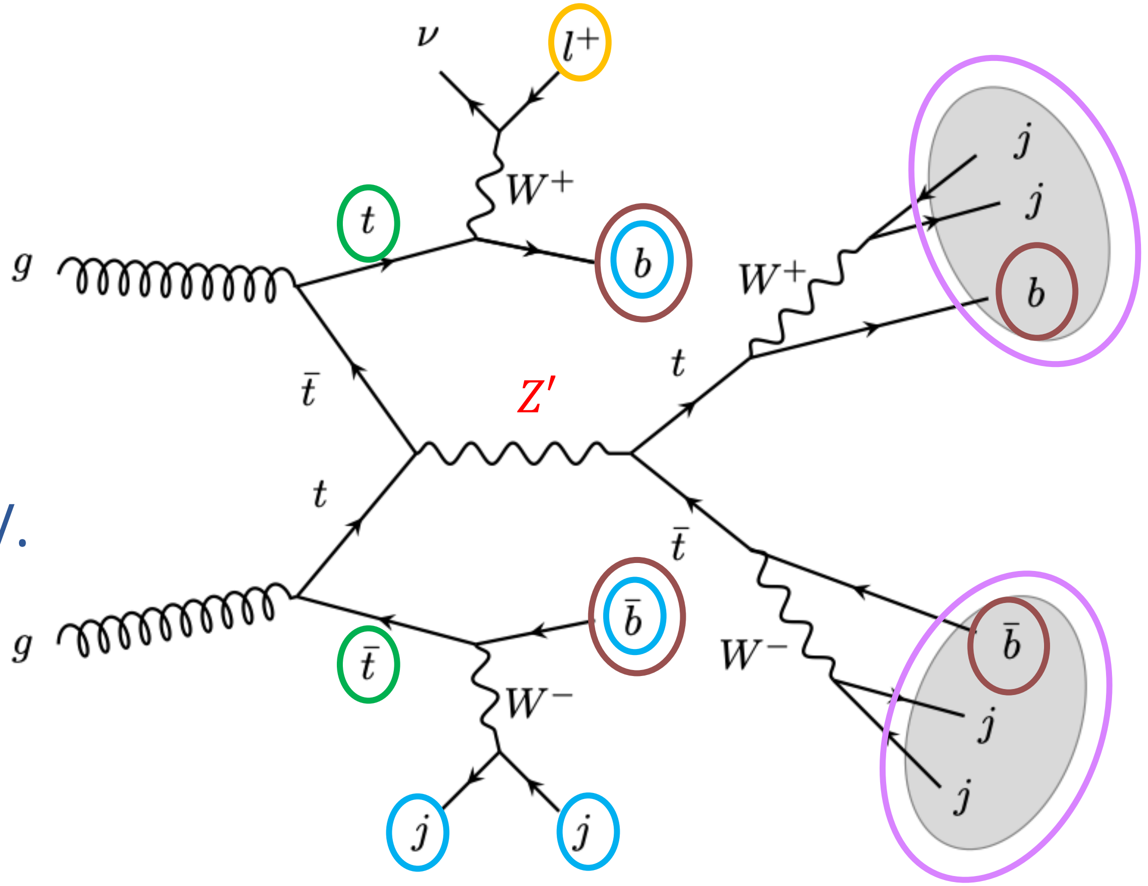


- Target events with **exactly one lepton**, preferably coming from a "spectator" top.
- Heavy Z′ → 2 **re-clustered (RC) jets** with R = 1.0, $p_T \geq 300$ GeV and $m \geq 100$ GeV.



Reconstruction & event selection

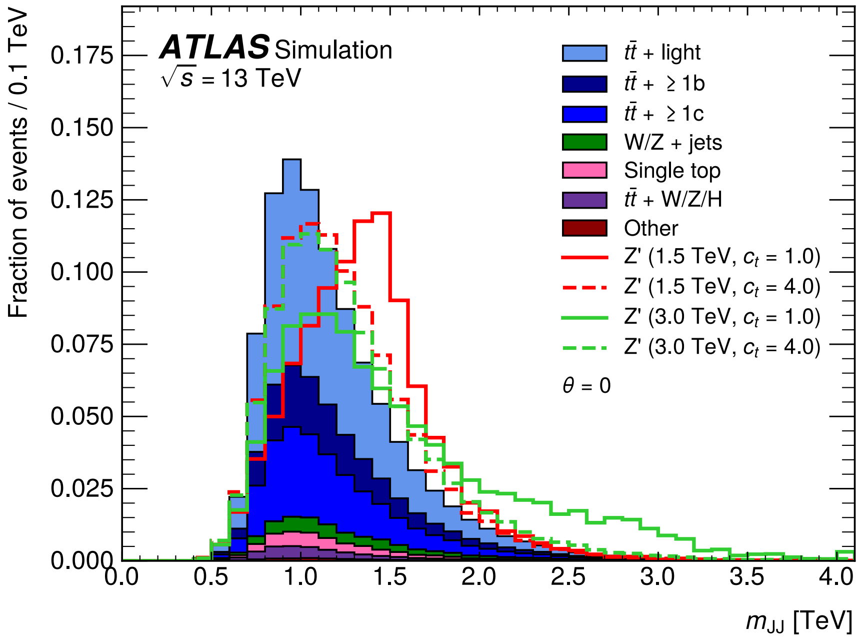
- Target events with **exactly one lepton**, preferably coming from a "spectator" top.
- Heavy Z′ → 2 **re-clustered (RC) jets** with R = 1.0, $p_T \geq 300$ GeV and $m \geq 100$ GeV.
- Number of **b-jets** and **additional jets** used to define regions.



Ideal event = 1 lepton + 2 RC jets + 4 b-jets + 4 additional jets

Analysis strategy

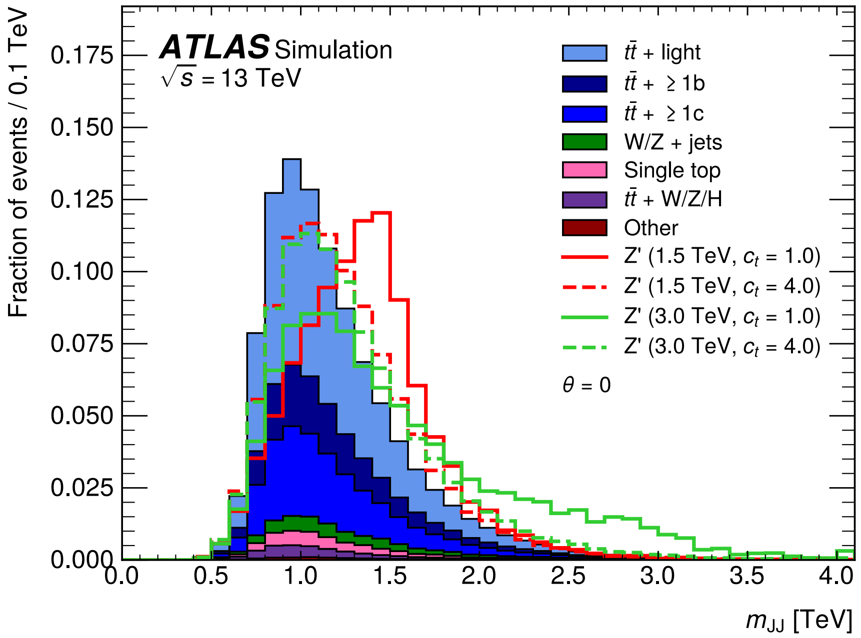
1. **Reconstruct** the mass of the Z' as the invariant mass of the 2 RC jets: m_{JJ}



Preselection region:
1 lepton
≥2 RC jets
≥2 b-jets
≥2 additional jets

Analysis strategy

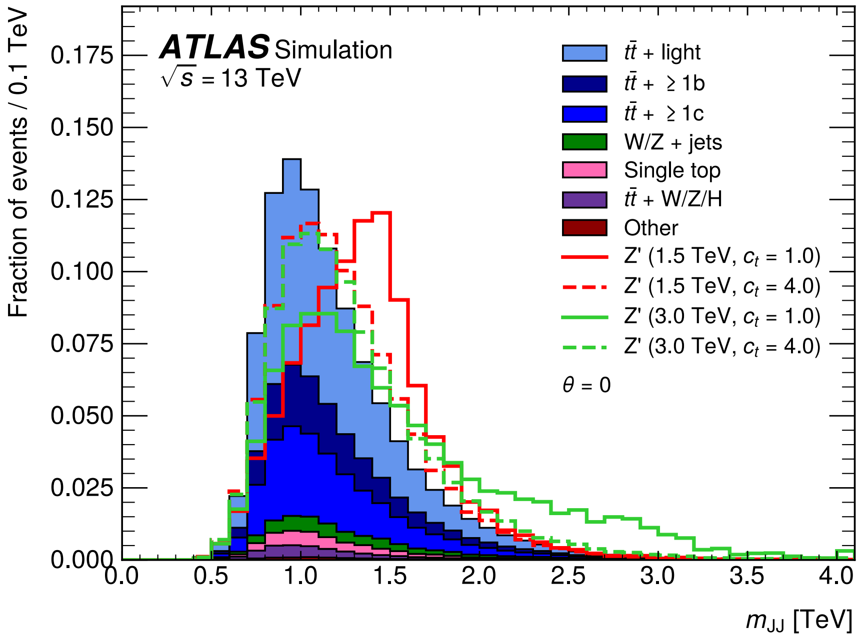
1. **Reconstruct** the mass of the Z' as the invariant mass of the 2 RC jets: m_{JJ}
2. **Estimate** the background



Preselection region:
1 lepton
≥2 RC jets
≥2 b-jets
≥2 additional jets

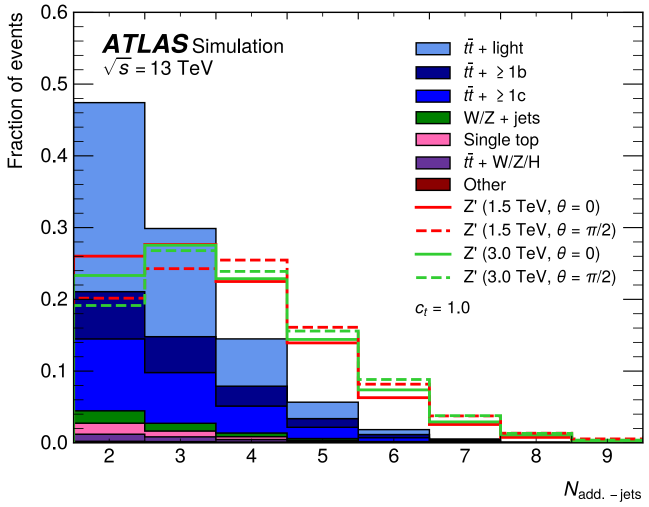
Analysis strategy

1. **Reconstruct** the mass of the Z' as the invariant mass of the 2 RC jets: m_{JJ}
2. **Estimate** the background
3. **Interpret** the results
   A. Model independent using BumpHunter
   B. Model dependent: significance or limit

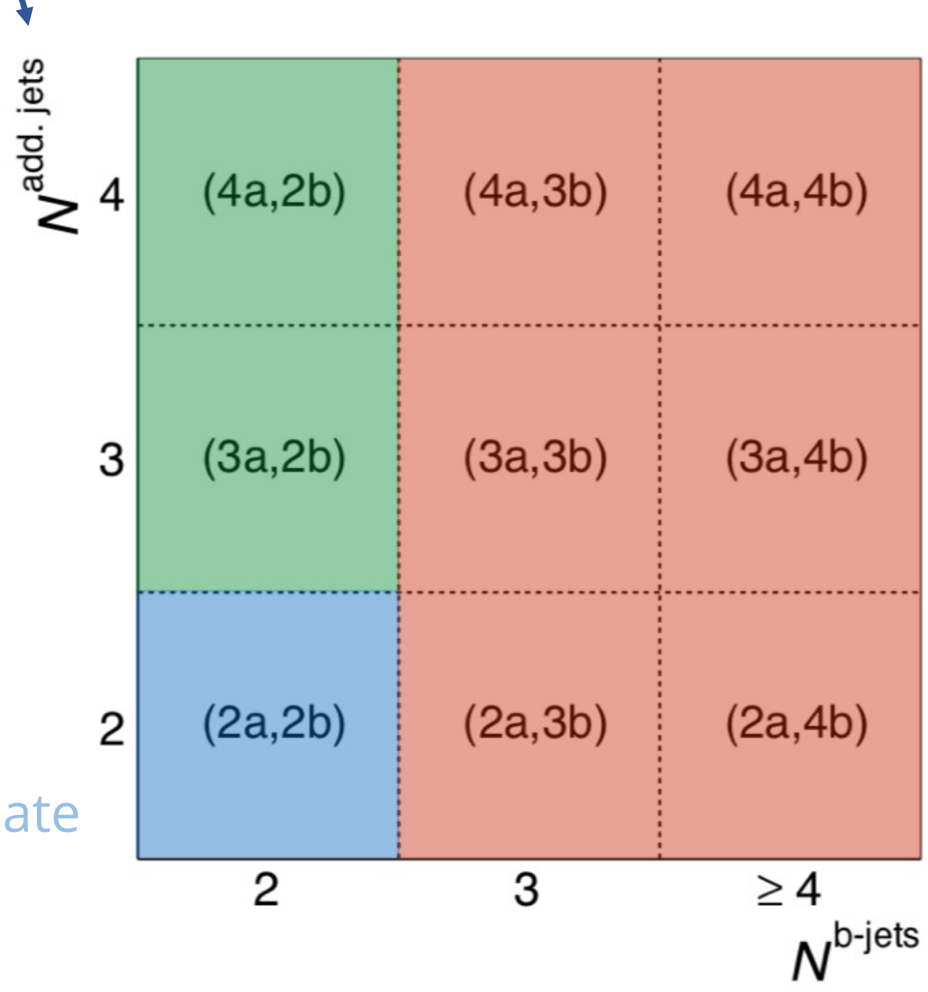


Preselection region:
1 lepton
≥2 RC jets
≥2 b-jets
≥2 additional jets

Region definition

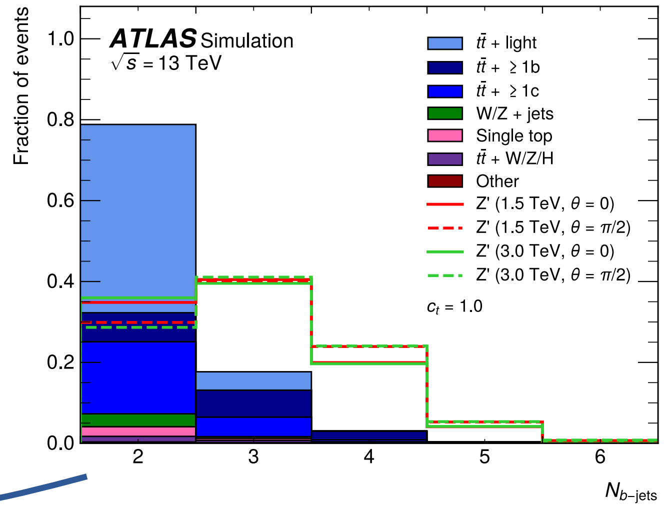


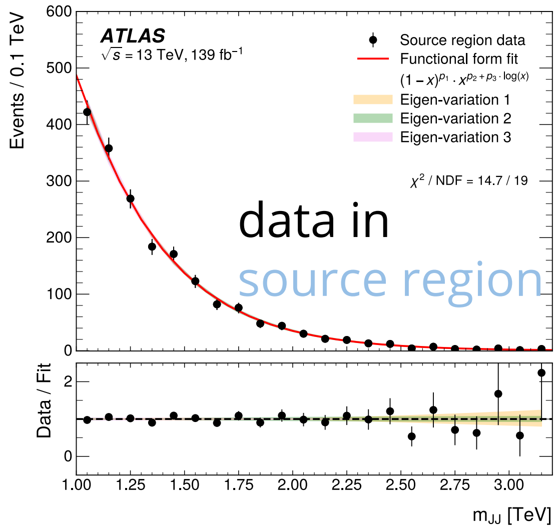
Validation regions: for checking

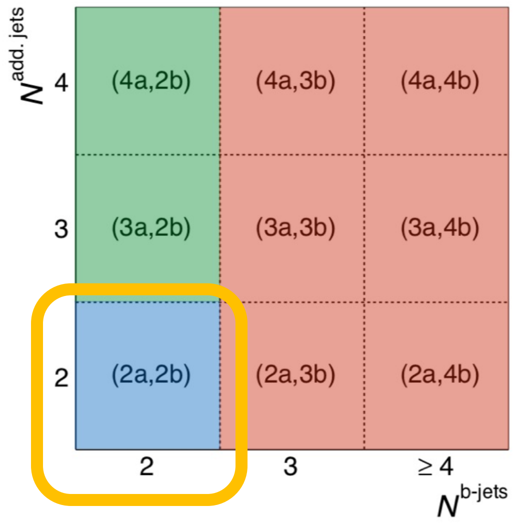


Signal regions:
for final interpretation

Source region:
for background estimate





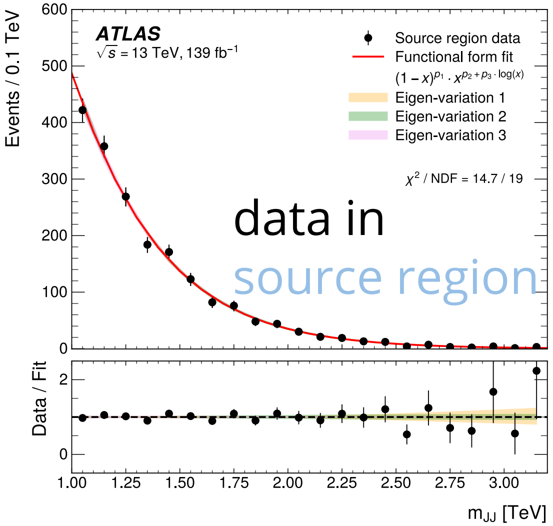


Dijet function:

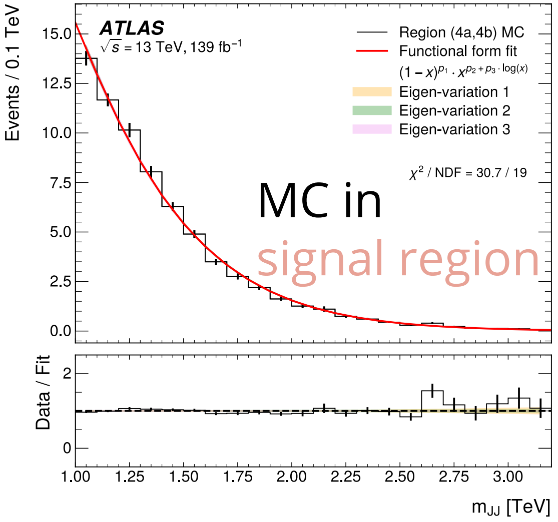
$$f(x) = (1-x)^{p_1} \times x^{p_2+p_3\log(x)}$$

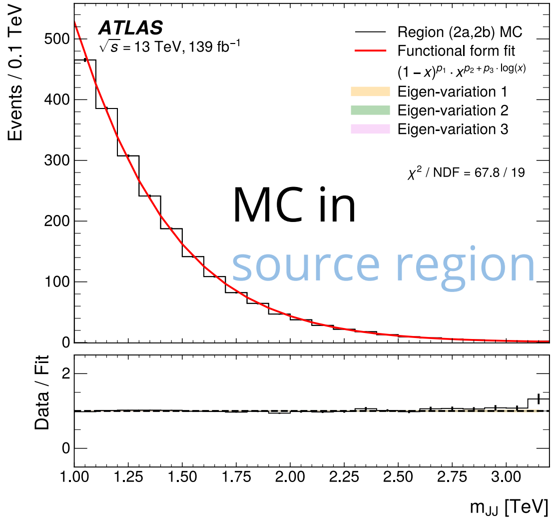
Background estimate

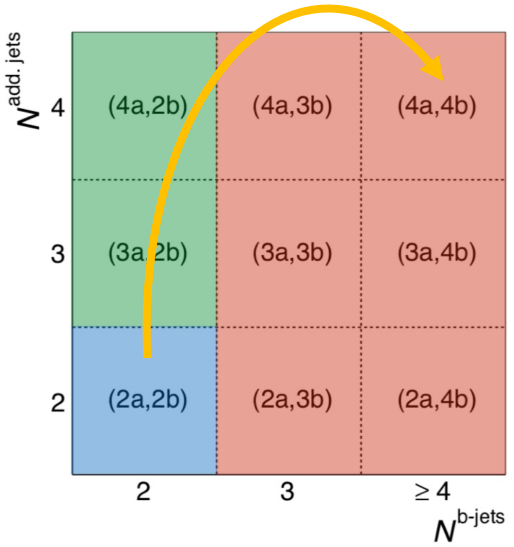
Extrapolation function



×







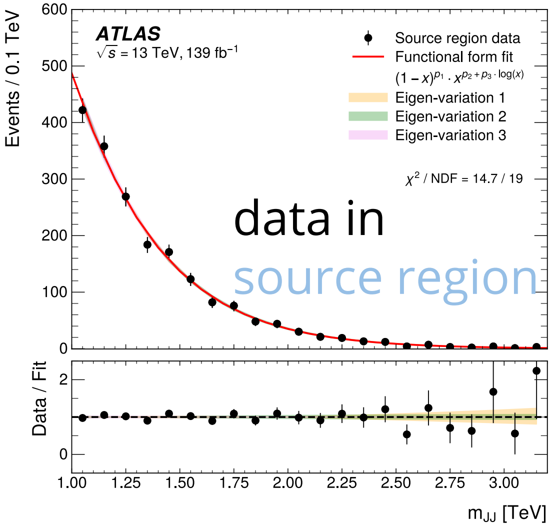
= Estimated background in signal region

Dijet function:

$$f(x) = (1-x)^{p_1} \times x^{p_2+p_3 \log(x)}$$

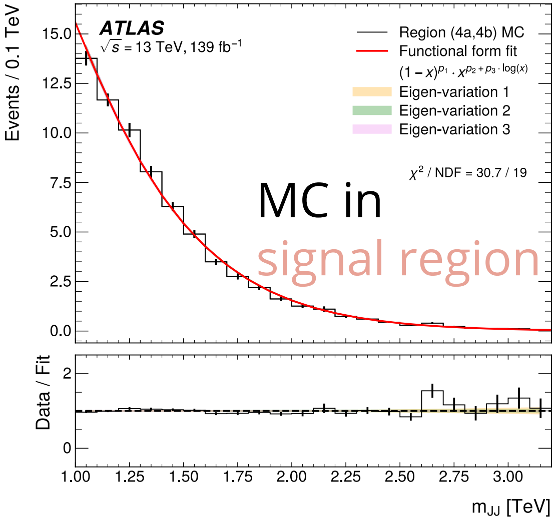
Background estimate

Extrapolation function

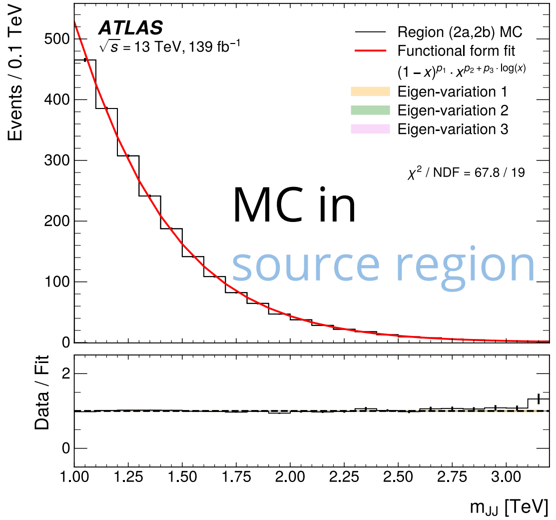


data in source region

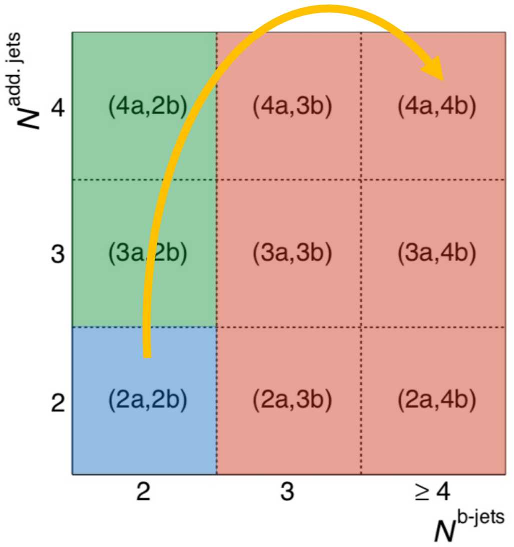
×



MC in signal region



MC in source region



= Estimated background in signal region

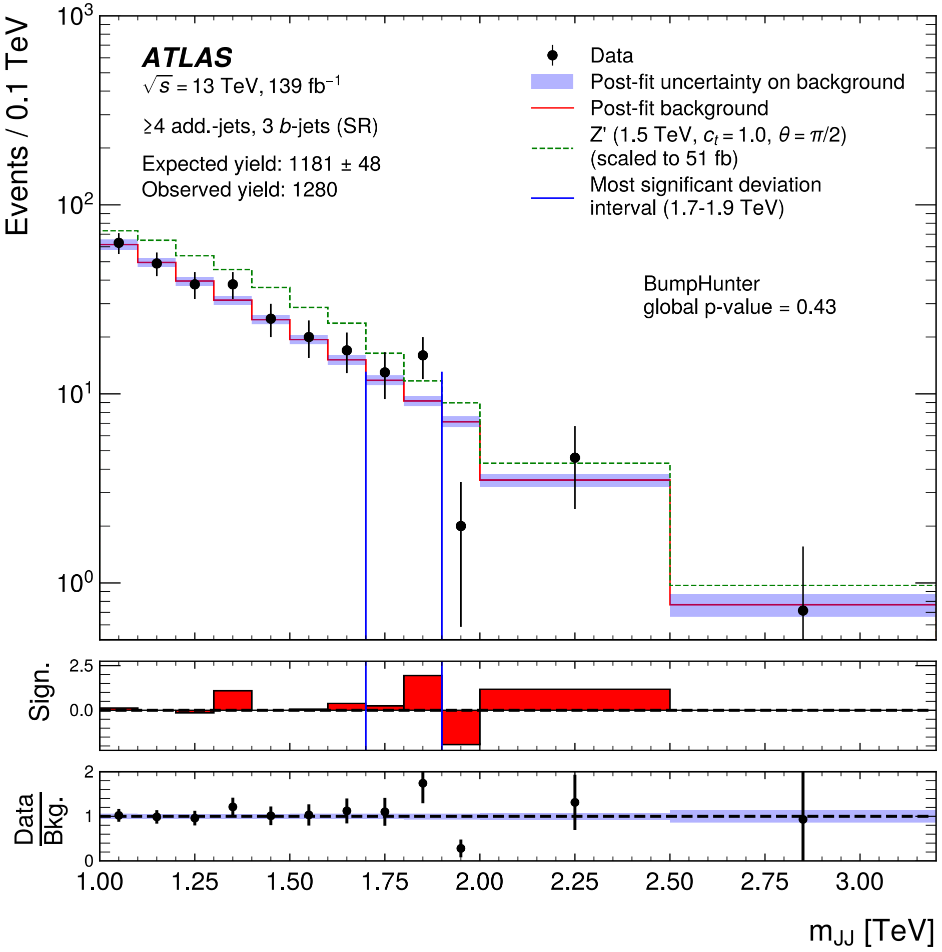
Dijet function:

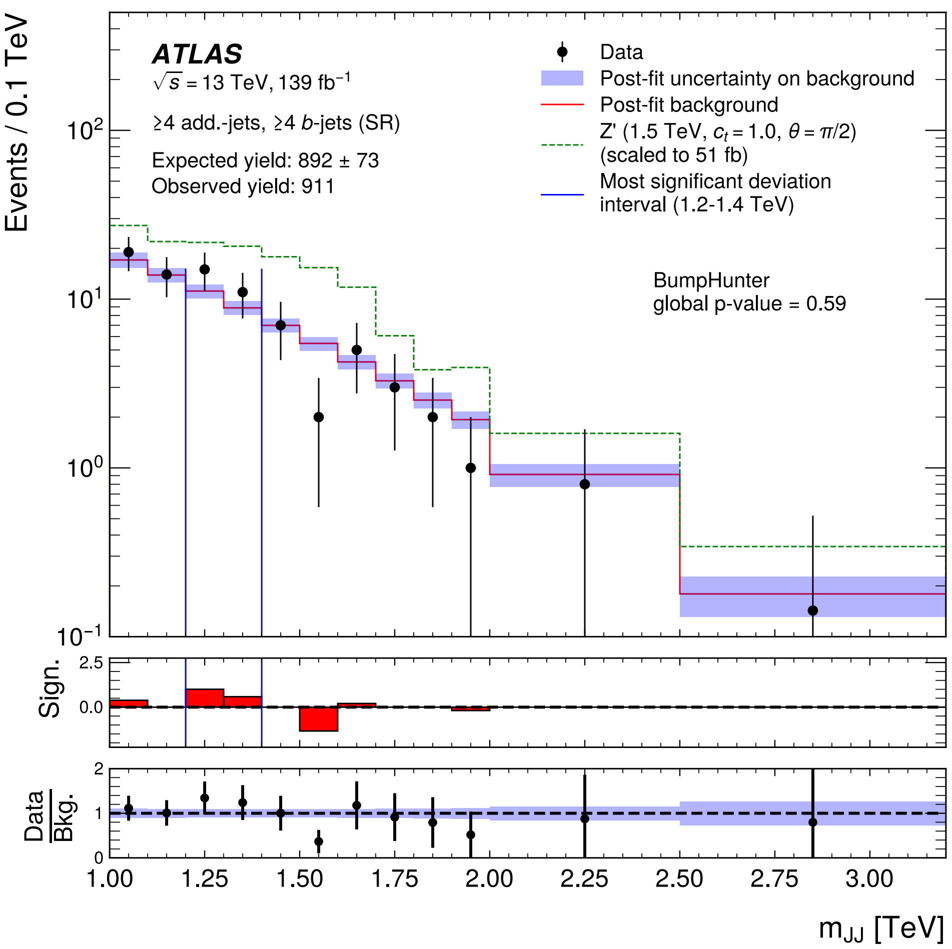
$$f(x) = (1 - x)^{p_1} \times x^{p_2 + p_3 \log(x)}$$

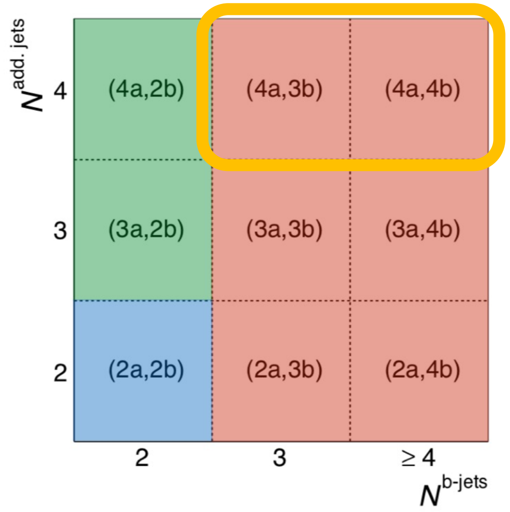
Systematic uncertainties are propagated by re-computing the extrapolation functions from varied MC.

Model independent results

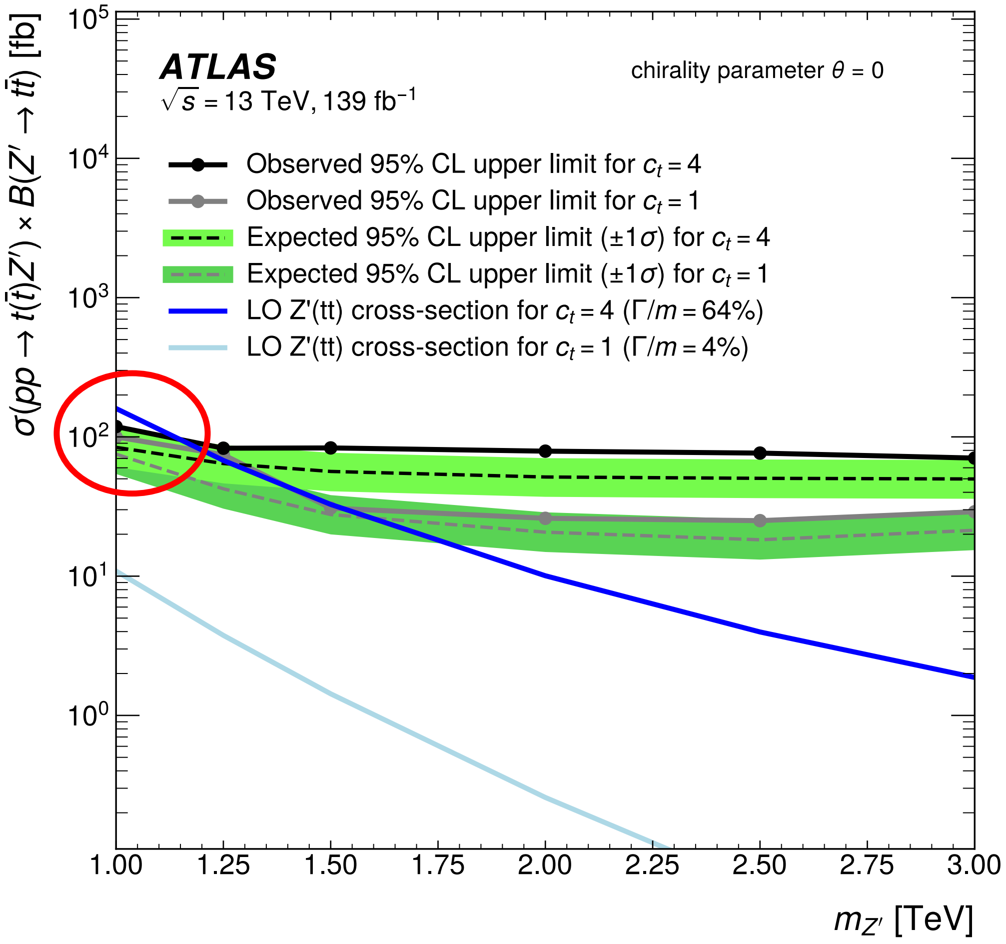
Most signal sensitive regions

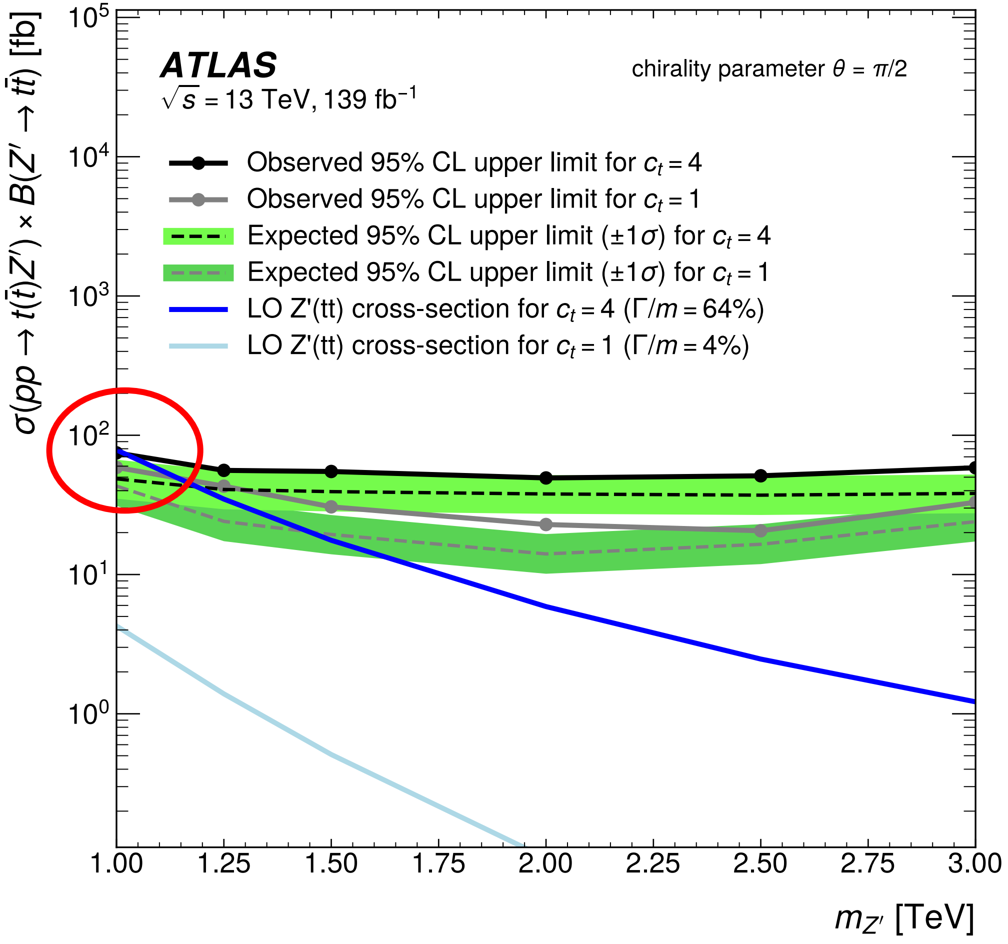






Model dependent results





⇒ Can exclude 1 TeV mass point for $c_t = 4$.

Possible improvements

Using machine learning:

Inputs (particle variables)

$(E, p_x, p_y, p_z)_1$

$(E, p_x, p_y, p_z)_2$

…

$(E, p_x, p_y, p_z)_N$

Classifier

Outputs

$(P_{t,spec}, P_{\bar{t},spec}, P_{t,res}, P_{\bar{t},res})_1$

$(P_{t,spec}, P_{\bar{t},spec}, P_{t,res}, P_{\bar{t},res})_2$

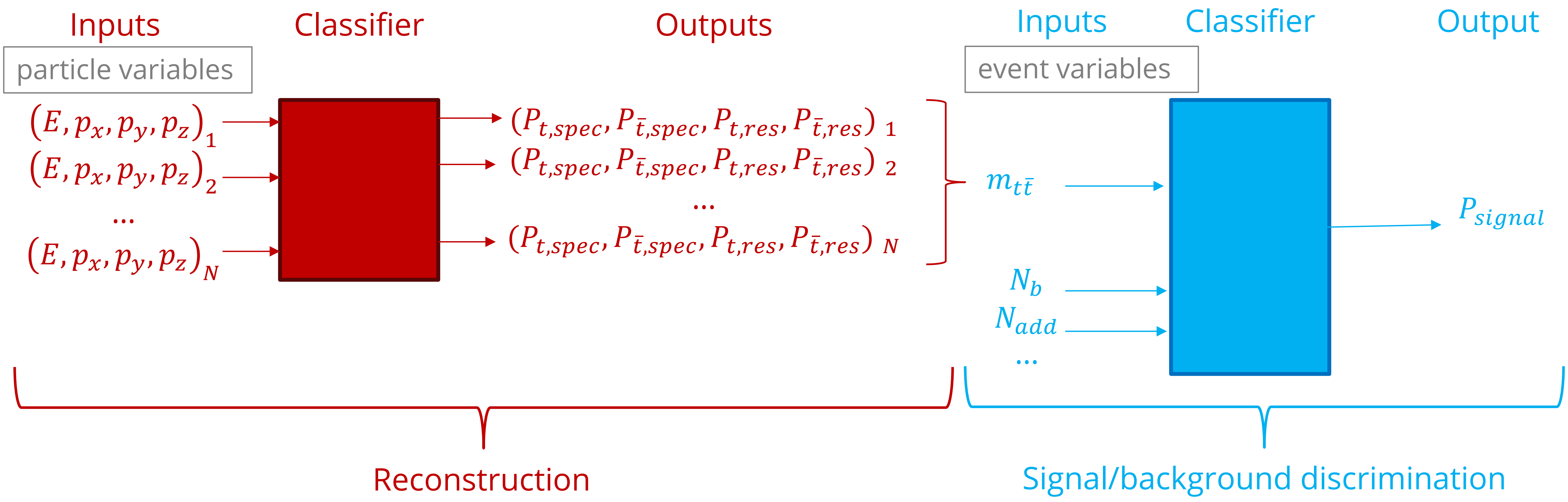
…

$(P_{t,spec}, P_{\bar{t},spec}, P_{t,res}, P_{\bar{t},res})_N$

Reconstruction

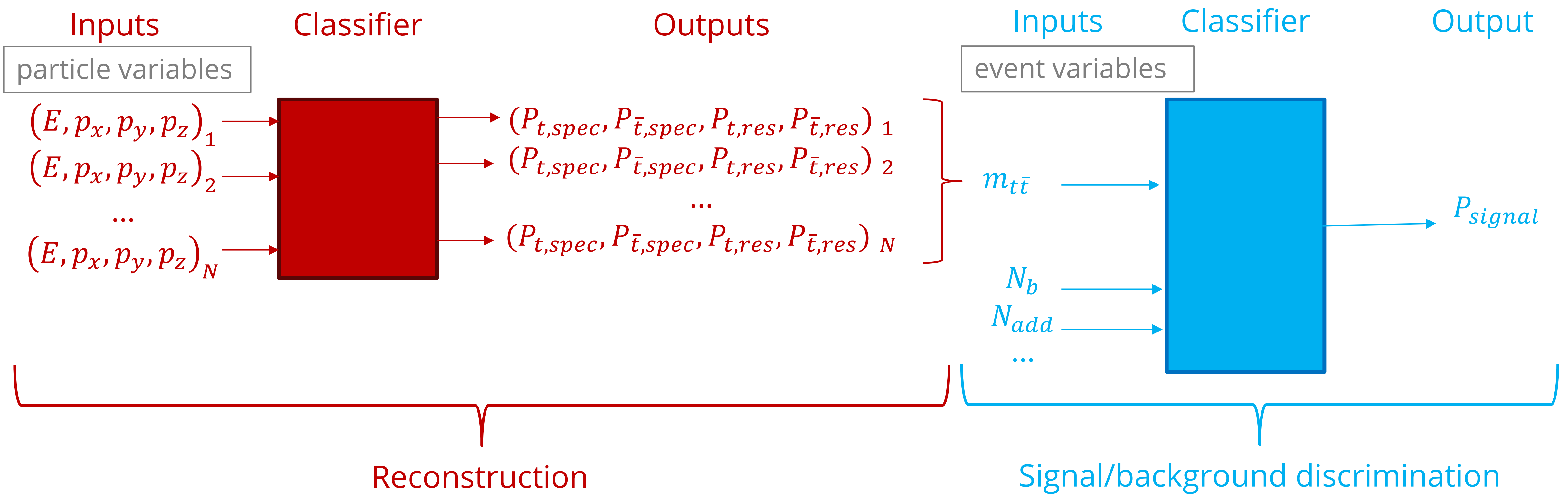
Possible improvements

Using machine learning:



Possible improvements

Using machine learning:



Some promising avenues:

- SPANet [SciPost Phys **12**, 178 (2022)]
- Topograph [Phys. Rev. D **107**, 116019 (2023)]
- $\nu^{(2)}$-flows [SciPost Phys. **14**, 159 (2023) and arxiv:2307.02405]

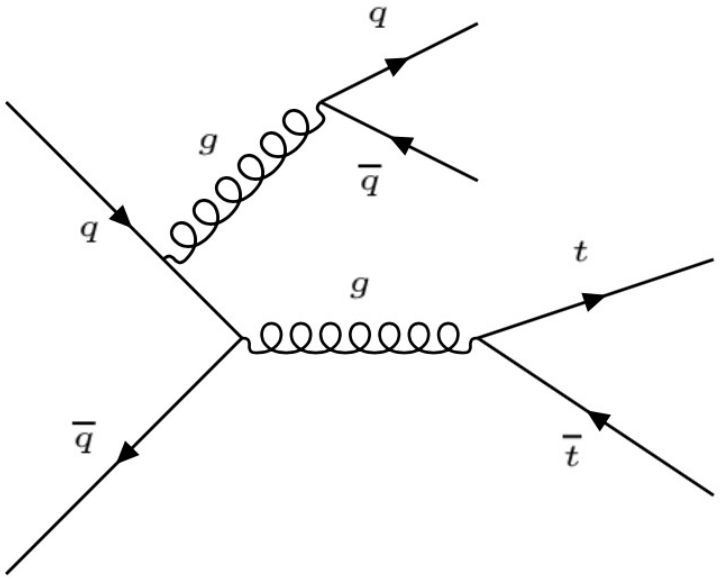
Thank you for your attention!

…

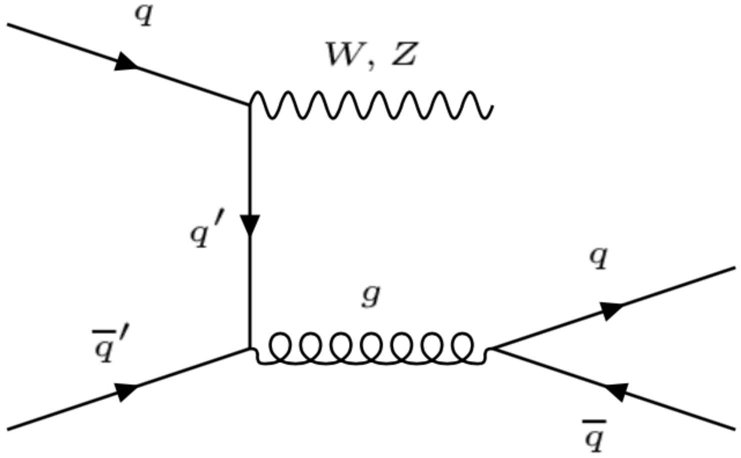
Backup

Main backgrounds

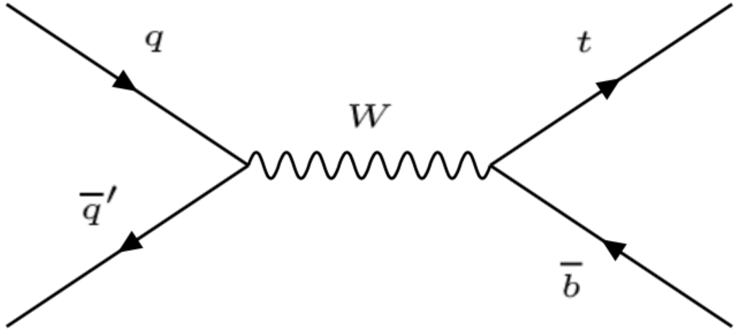
$t\bar{t}$ + jets
88%



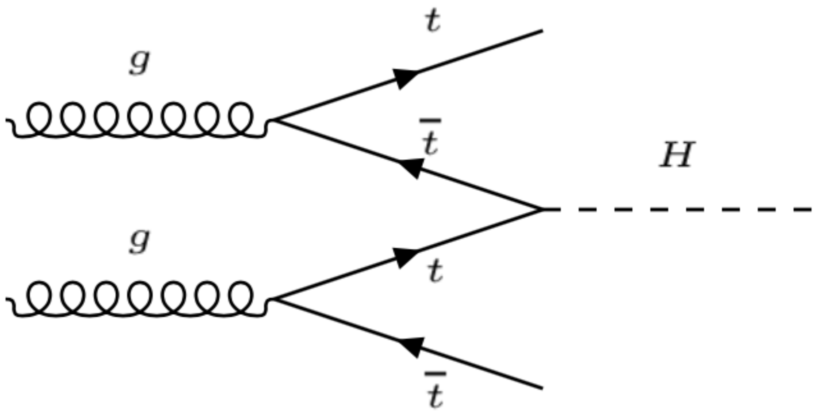
W, Z + jets
5%



Single top
4%



$t\bar{t}V$
$(\mathrm{V = H, W, Z})$
3%



Percentages correspond to pre-selection level

Uncertainties on dijet fits

- For fits to **nominal MC**, estimate uncertainties using the **bootstrap method**:
  1. Obtain a large number (1000) of toys based on original MC histogram.
  2. Apply dijet fit to each toy and record central values of parameters.
  3. Compute covariance matrix from set of pseudo-experiments.
  4. Compute eigen-decomposition using covariance matrix from previous step.
- Use result of single fit for nominal background estimation and **"stitch"** eigen-variations on top of nominal.
- For fits to data, use single fit for central value and uncertainties.
- Resulting nuisance parameters:
  - ➢ 3 NPs from fit to data in source region, correlated across signal regions
  - ➢ 3 NPs from fit to MC in each region (signal and source)

  ⇒ **24 NPs total**.

Tests of background estimate and fit

- **Model-agnostic tests**:

1. Goodness-of-fit test
   - conditional $\mu = 0$ profile-likelihood fit and associated checks in Asimov and pseudo-data
   - χ^2 test with saturated model in ensemble of 500 background-only pseudo-data
2. Type I error probability test
   - run BumpHunter on ensemble of background-only pseudo-data sets, require less than 10% fraction of pseudo- experiments which result in BumpHunter p-values of less than 0.05
3. Signal injection studies
   - study signal extraction with ensemble of signal+background pseudo-data sets, expect small BumpHunter p-values for most pseudo-experiments

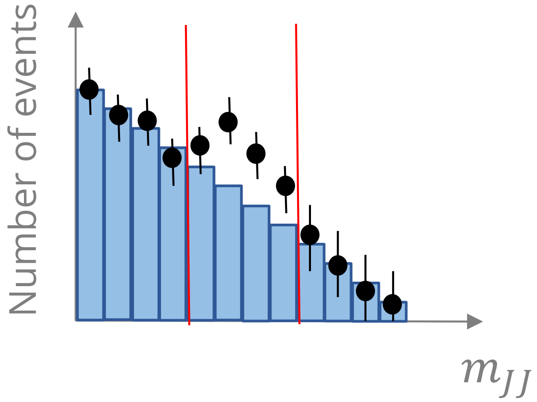
- **Model-dependent tests:**

1. Fit cross-checks for signal+background fits
   - unconditional profile-likelihood fits + checks in Asimov and pseudo-data for various signal hypotheses
2. Spurious signal test
   - run unconditional profile-likelihood fit for ensemble of background-only pseudo-data sets for various signal hypotheses, extracted μ expected to follow Gaussian distribution centered around 0, resulting $\langle\mu\rangle$ required to be compatible with zero within standard deviation
3. Signal extraction test
   - run unconditional profile-likelihood fit for ensemble of signal+background pseudo-data sets for various signal hypotheses with various injected μ_{inj}, expect linear dependence between μ_{inj} and $\langle\mu\rangle$

Statistical analysis

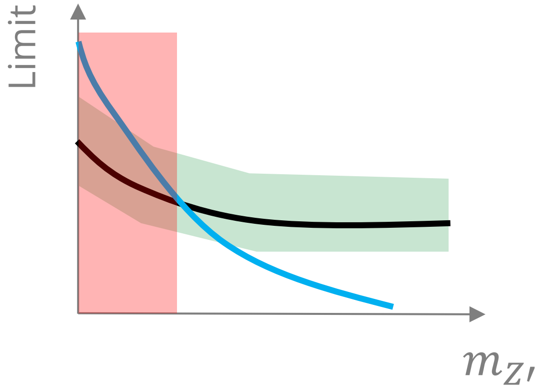
Model independent interpretation:

1. Run background-only profile likelihood fit.
2. Input the post-fit m_{JJ} distributions in the signal regions into BumpHunter.
3. Find the "most significant interval", and significance.



Model dependent interpretation:

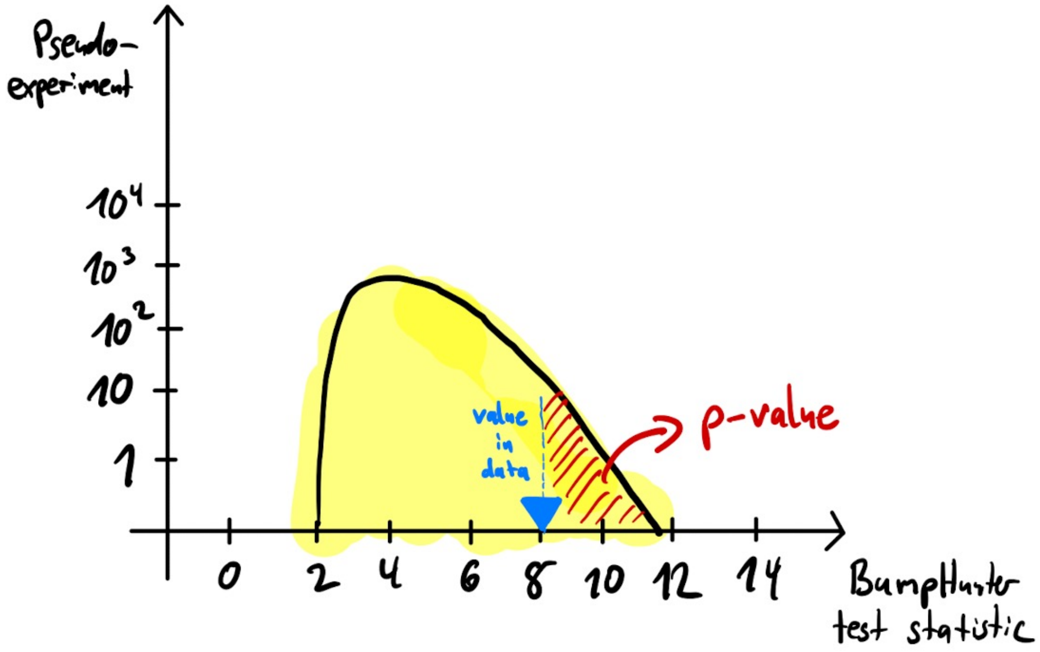
1. Run signal + background profile likelihood fit.
2. Calculate the significance (if a signal) or limit (if no signal).
3. Compare limit with theory prediction to determine which mass points are excluded.



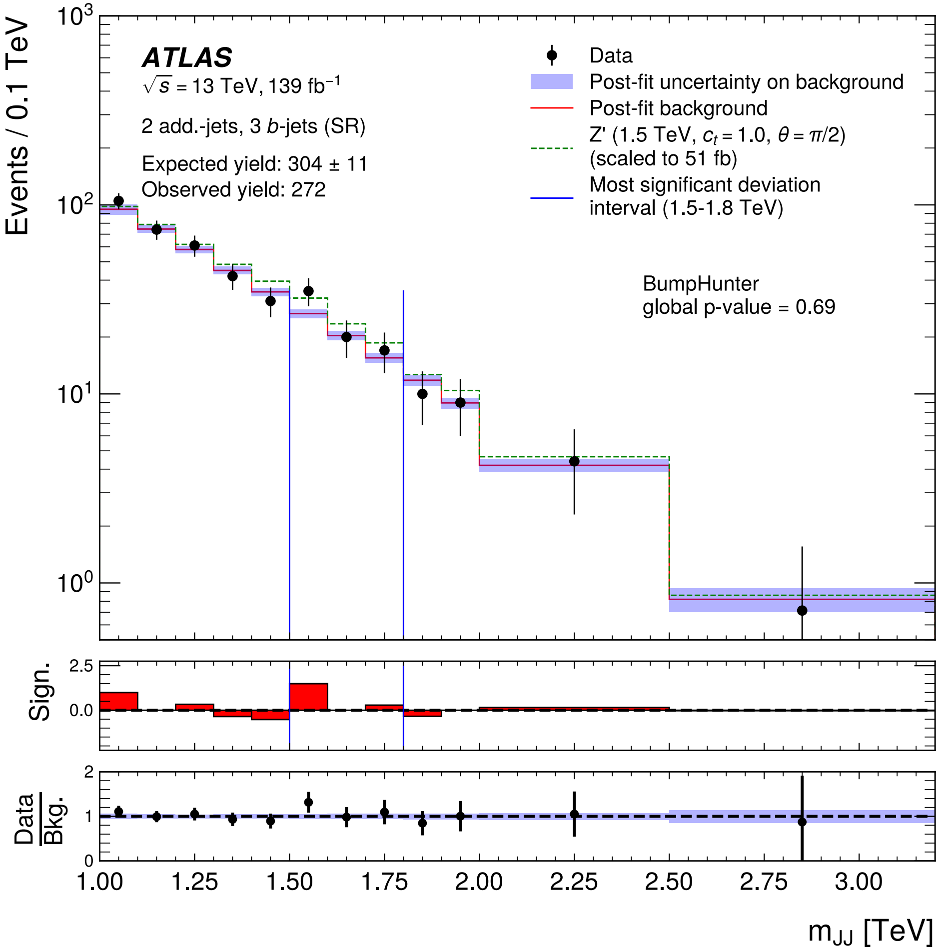
BumpHunter

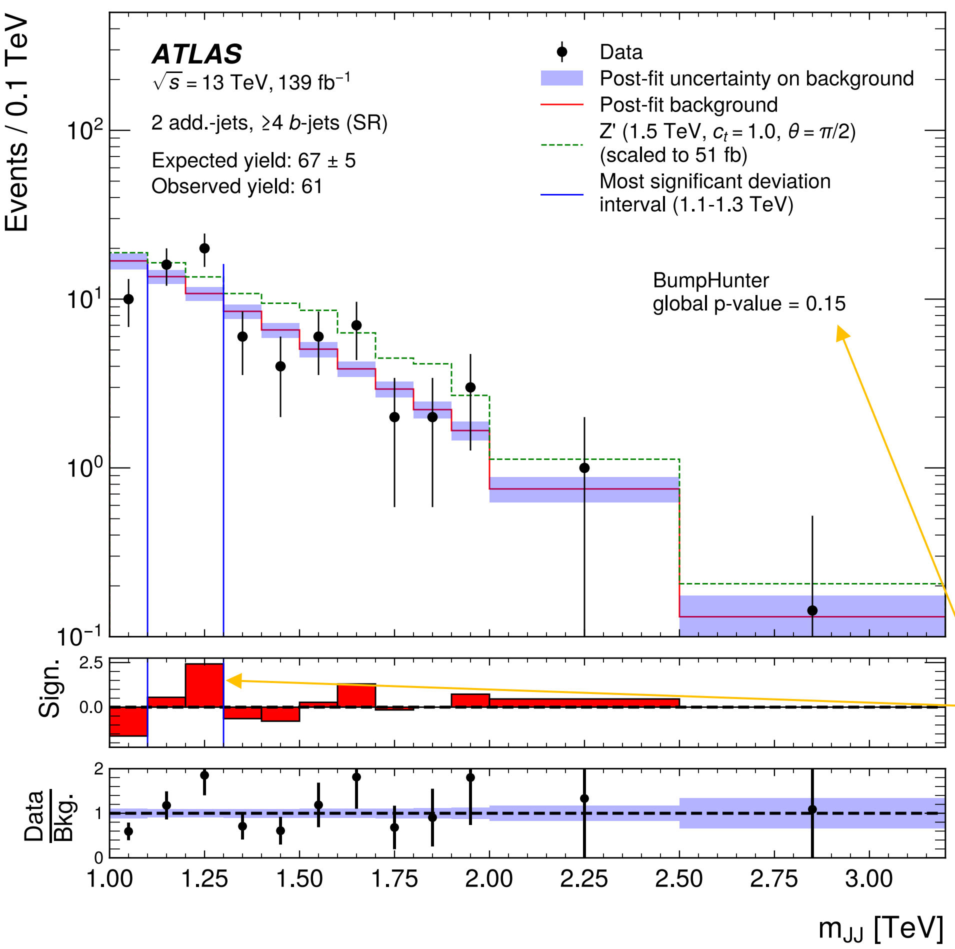
BumpHunter: locates the most significant local deviations from the bkg-only null hypothesis H_0, provides a p-value accounting for trials factor which corresponds to the Type I error probability.

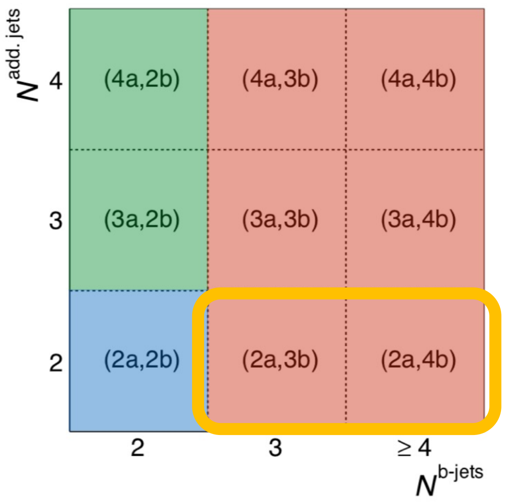
1. Generate large number of pseudo-data following H_0.
2. For each dataset, compute BumpHunter test statistic:
   - scan with sliding window and report test statistic $t = -\log(p\text{−}value_{min})$ based on window with smallest p-value.
3. Calculate p-value of the test based on observed data and t-distribution of pseudo-experiment.



Model independent results

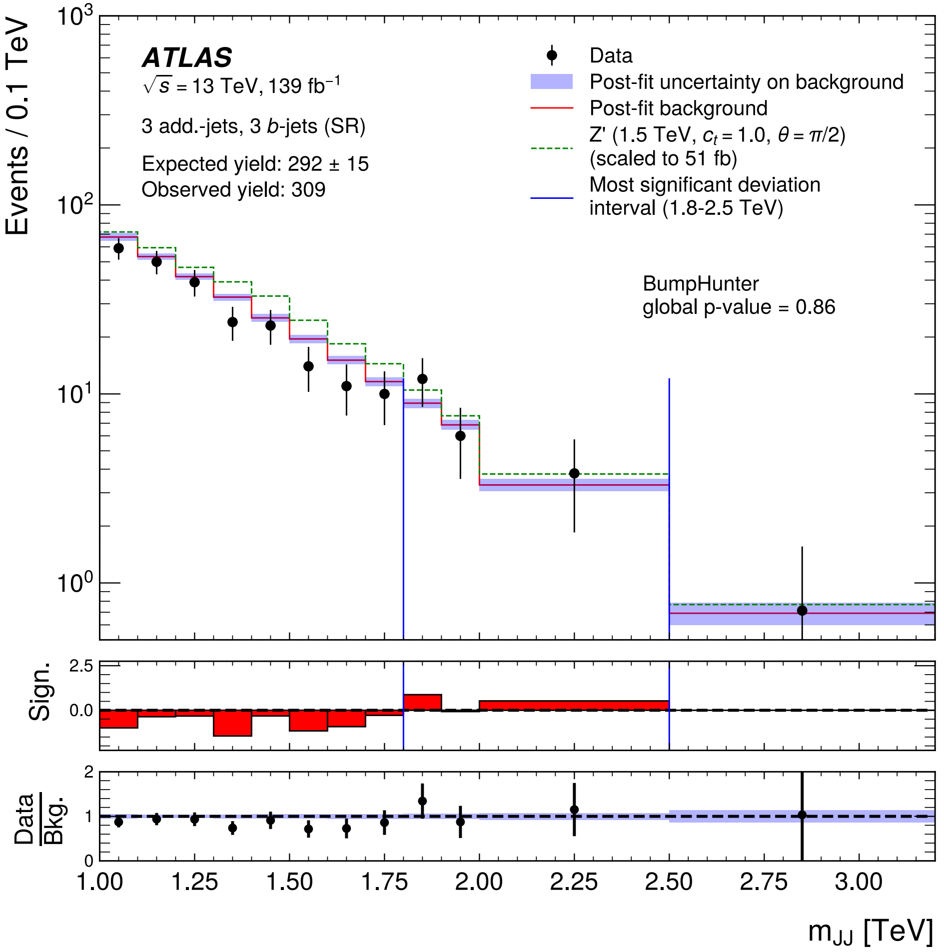


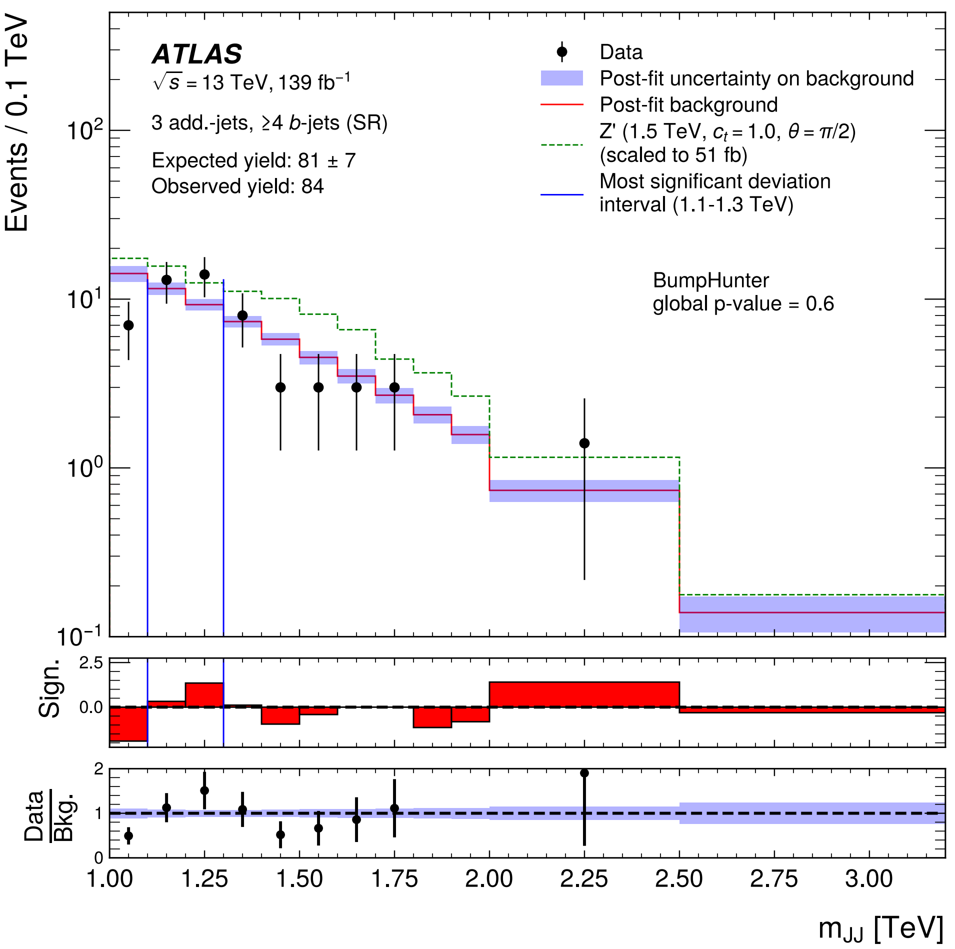


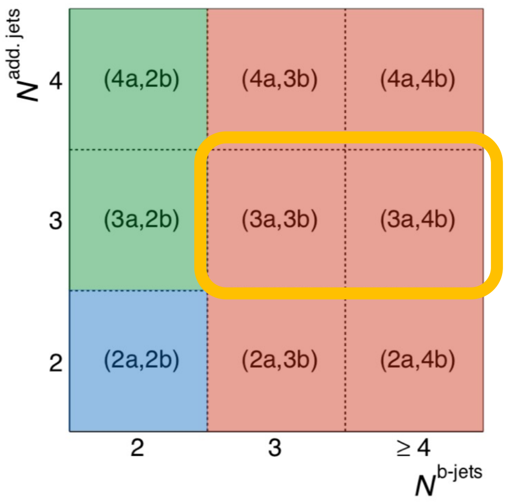


Largest local significance and smallest global p-value

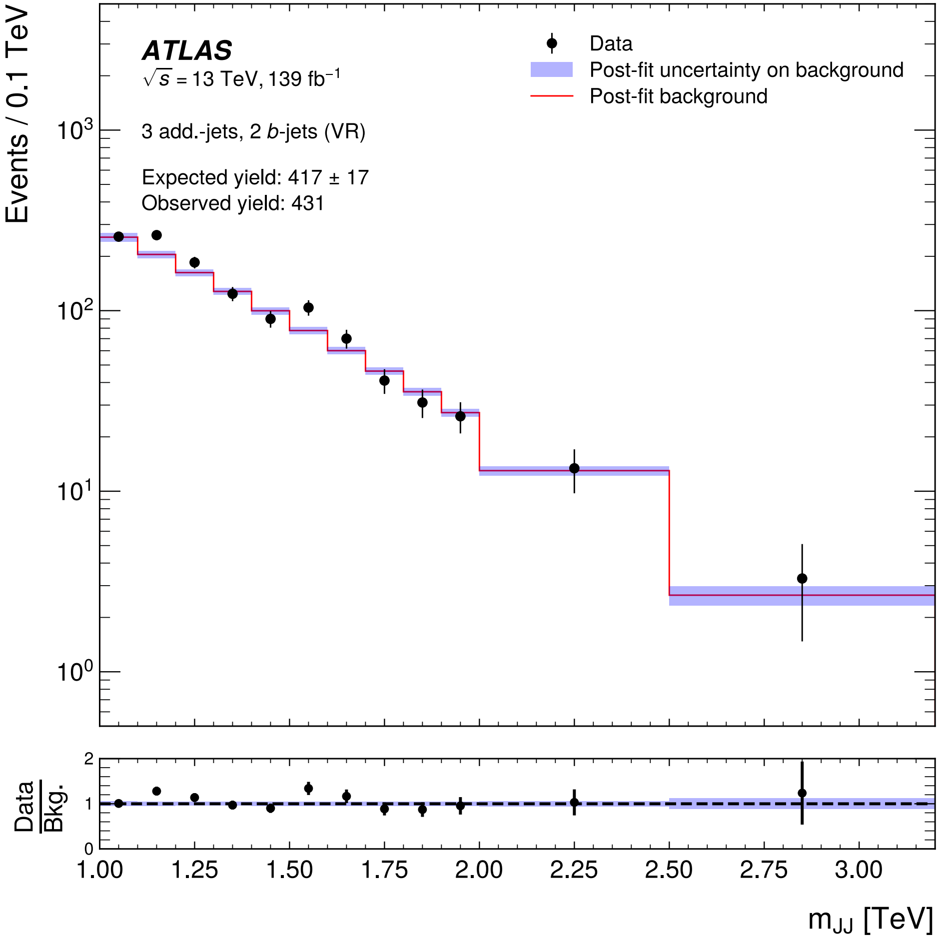
Model independent results

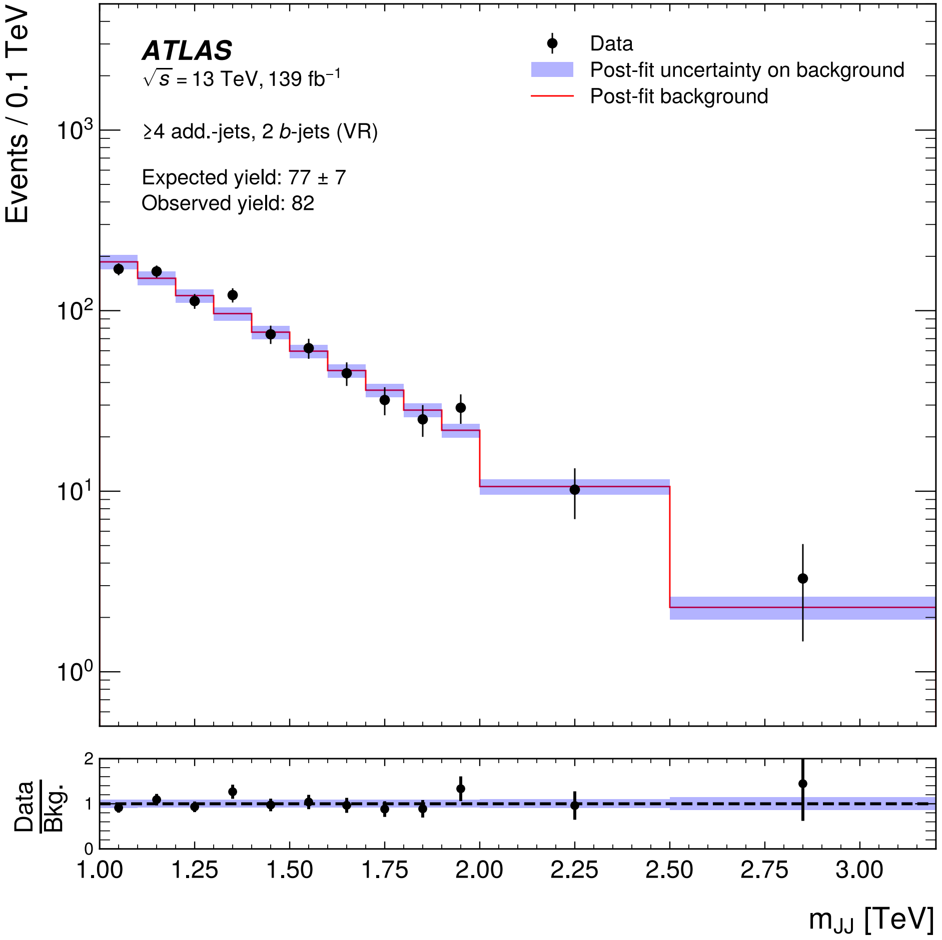


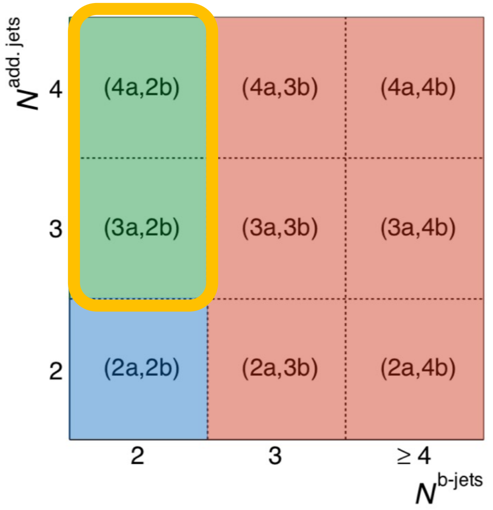




Model independent results – validation regions







Model dependent results – 2D limits