

Search for Heavy Boson Resonances in Final States with a W or Z Boson and a Photon Using 139 fb⁻¹ of Proton-Proton Collision Data at $\sqrt{s} = 13$ TeV Collected with the ATLAS Detector

Nov. 12, 2020

Minyu Feng
Duke University





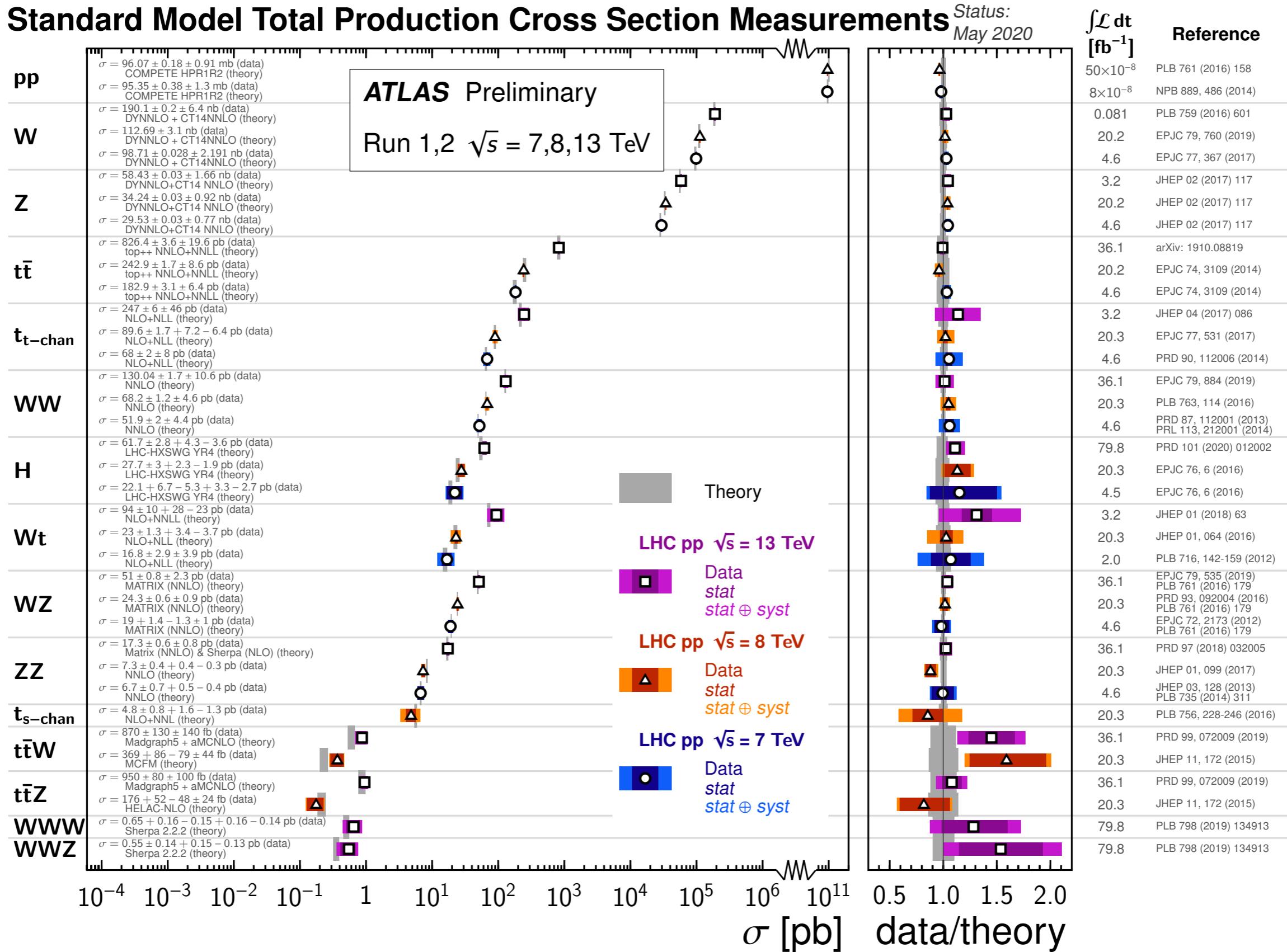
Theory Motivation: Standard Model

QUARKS		LEPTONS		GIGE BOSONS	
mass →	$\approx 2.3 \text{ MeV}/c^2$				
charge →	$2/3$				
spin →	$1/2$				
	u				
	up				
	$\approx 1.275 \text{ GeV}/c^2$				
	$2/3$				
	$1/2$				
	c				
	charm				
	$\approx 173.07 \text{ GeV}/c^2$				
	$2/3$				
	$1/2$				
	t				
	top				
	$\approx 0 \text{ GeV}/c^2$				
	0				
	0				
	g				
	gluon				
	$\approx 4.8 \text{ MeV}/c^2$				
	$-1/3$				
	$1/2$				
	d				
	down				
	$\approx 95 \text{ MeV}/c^2$				
	$-1/3$				
	$1/2$				
	s				
	strange				
	$\approx 4.18 \text{ GeV}/c^2$				
	$-1/3$				
	$1/2$				
	b				
	bottom				
	$0.511 \text{ MeV}/c^2$				
	-1				
	$1/2$				
	e				
	electron				
	$105.7 \text{ MeV}/c^2$				
	-1				
	$1/2$				
	μ				
	muon				
	$1.777 \text{ GeV}/c^2$				
	-1				
	$1/2$				
	τ				
	tau				
	$91.2 \text{ GeV}/c^2$				
	0				
	1				
	Z				
	Z boson				
	$<2.2 \text{ eV}/c^2$				
	0				
	$1/2$				
	ν_e				
	electron neutrino				
	$<0.17 \text{ MeV}/c^2$				
	0				
	$1/2$				
	ν_μ				
	muon neutrino				
	$<15.5 \text{ MeV}/c^2$				
	0				
	$1/2$				
	ν_τ				
	tau neutrino				
	$80.4 \text{ GeV}/c^2$				
	± 1				
	1				
	W				
	W boson				

- Standard Model unified all particles and interactions by a symmetry of:

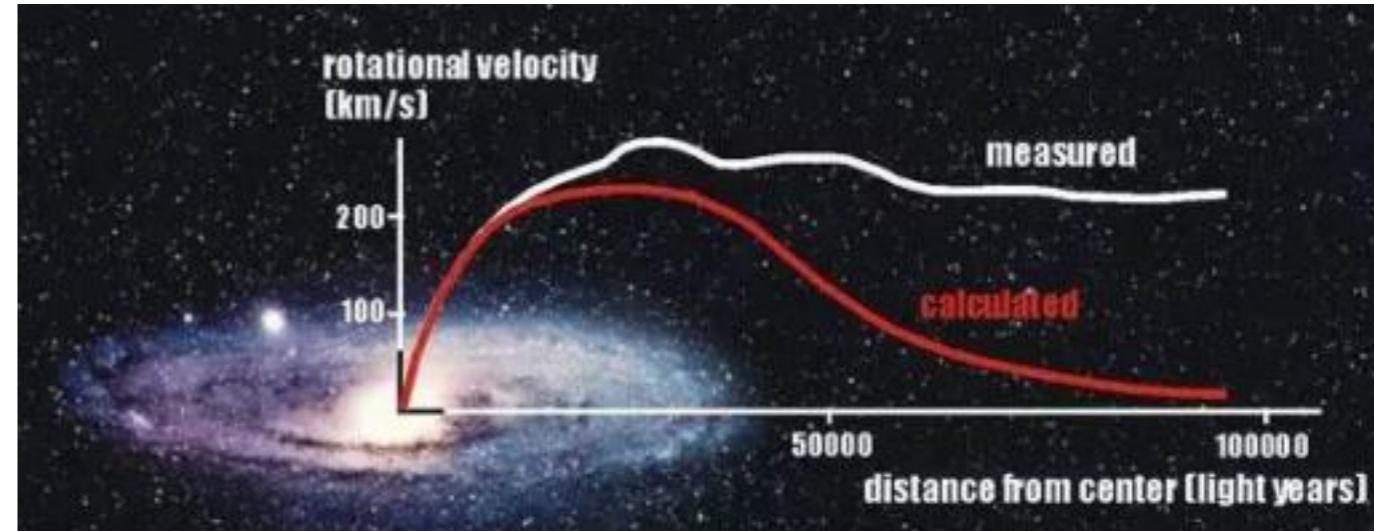
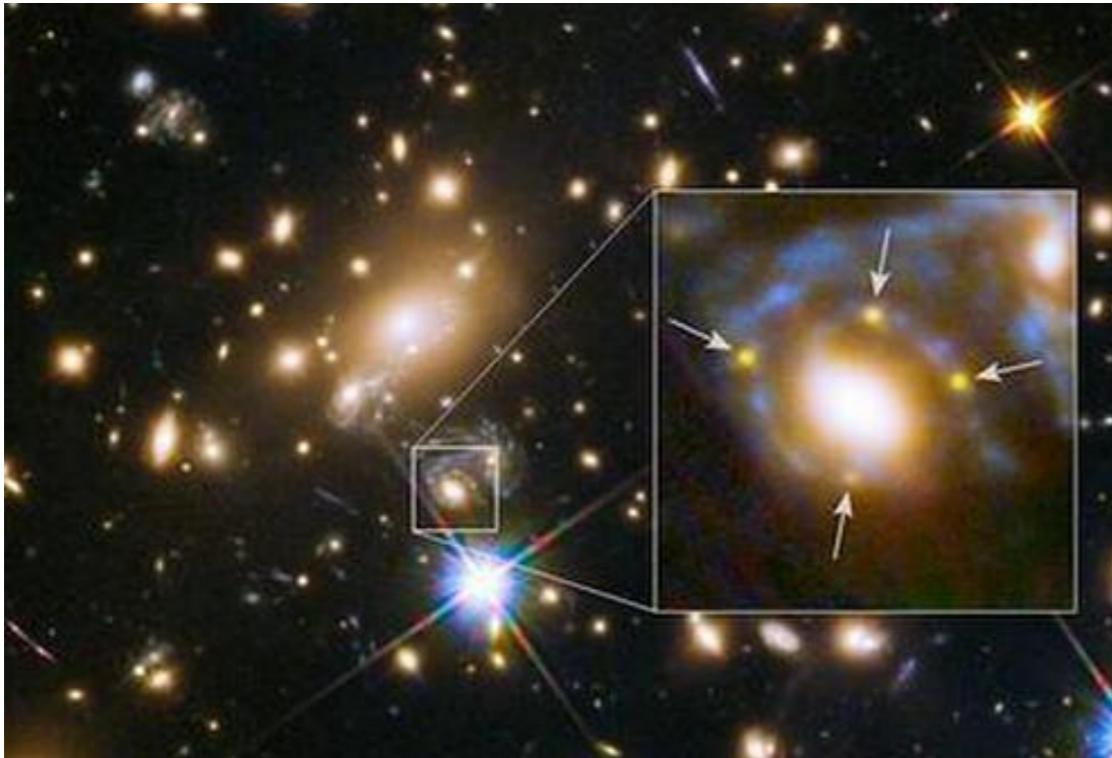
$$SU(3)_C \times SU(2)_L \times U(1)_Y$$

Theory Motivation: Measurements





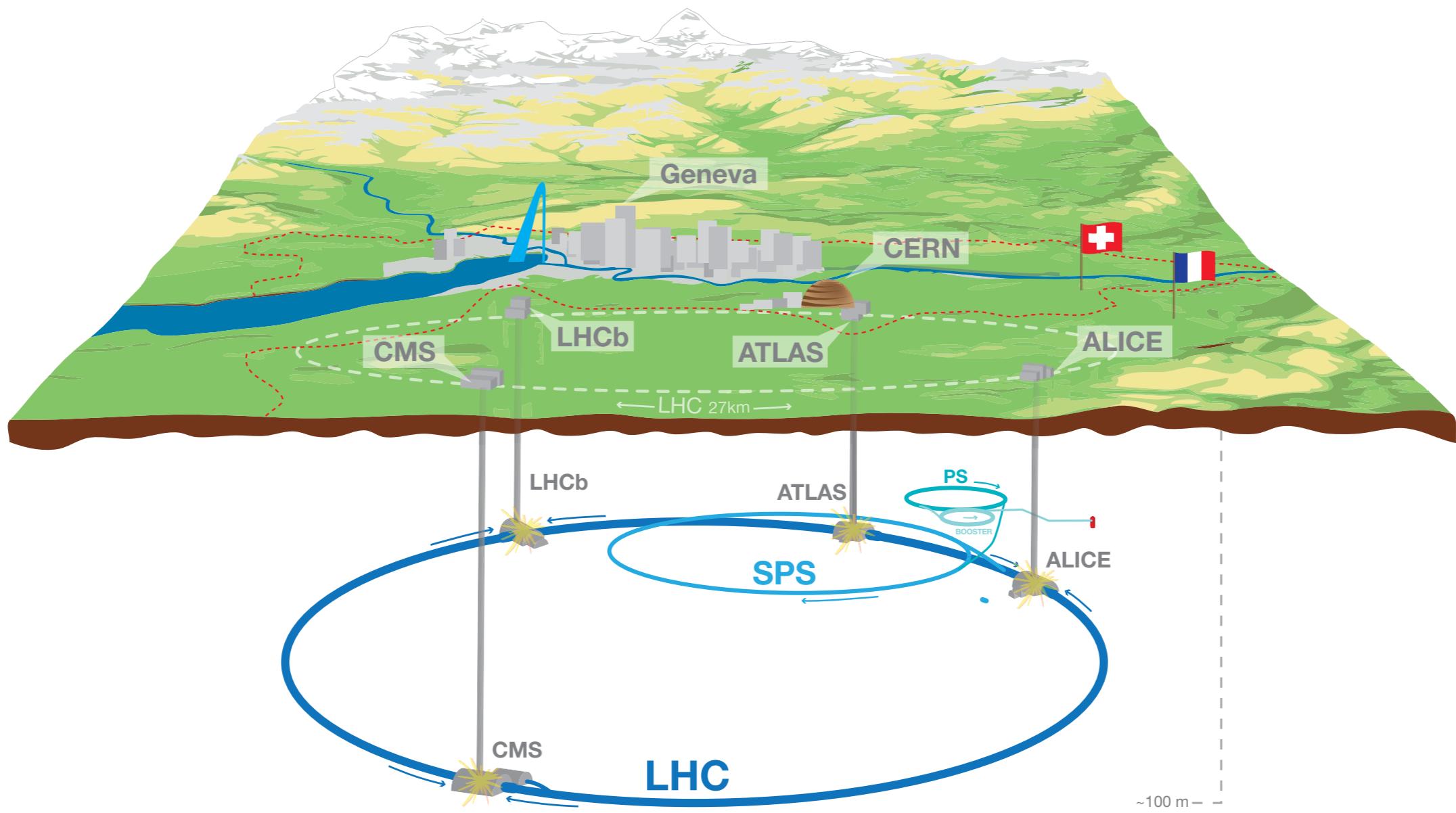
Theory Motivation: Beyond the Standard Model



- More and more evidences implying new physics beyond the SM:
 - ❖ Left gravity out of the model
 - ❖ Observation of dark matter
 - ❖ Neutrino non-zero mass
- Electroweak sector could be the connection between SM and BSM
- Many theories provide mechanisms extending SM from electroweak sector. We use some of the theories as our benchmark models for a generic search.

BSM Model	Boson Spin	Process
Singlet Scalar	0	$gg \rightarrow X \rightarrow \gamma Z (\rightarrow q\bar{q})$
Heavy Vector Triplet	1	$qq' \rightarrow X \rightarrow \gamma W (\rightarrow qq')$
Deconstruction of Extra Dimensions	2	$gg \rightarrow X \rightarrow \gamma Z (\rightarrow q\bar{q})$
		$q\bar{q} \rightarrow X \rightarrow \gamma Z (\rightarrow q\bar{q})$

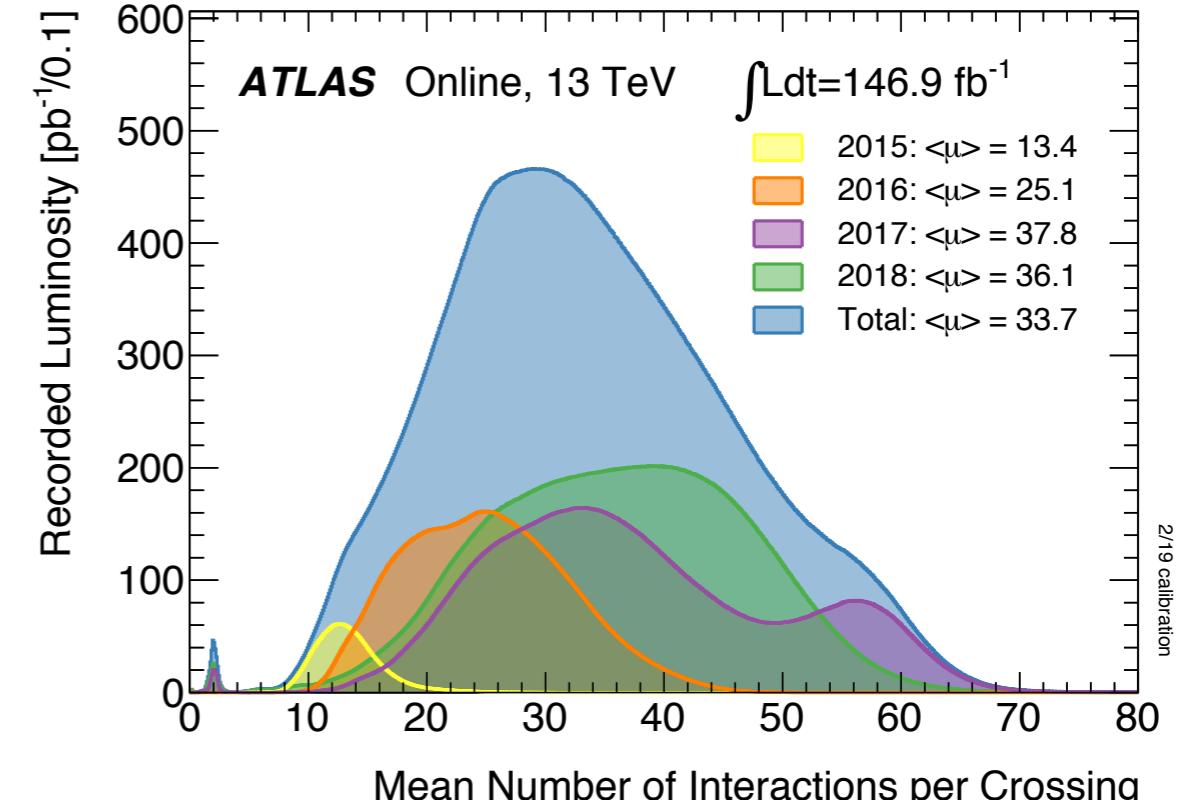
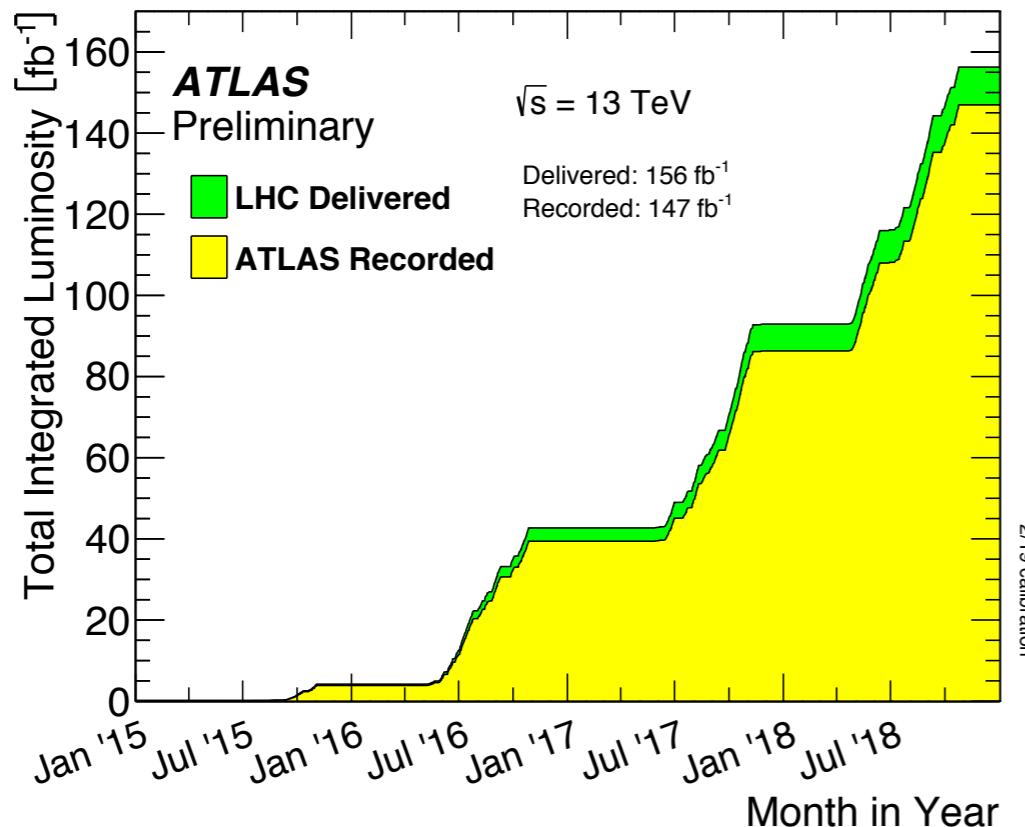
LHC and ALTAS: LHC Overview



- LHC boost protons to 6.5 TeV
- Protons are grouped in bunches ($\sim 1.1 \times 10^{11}$ each), separated by 25 ns
- The collision rate is ~ 40 MHz



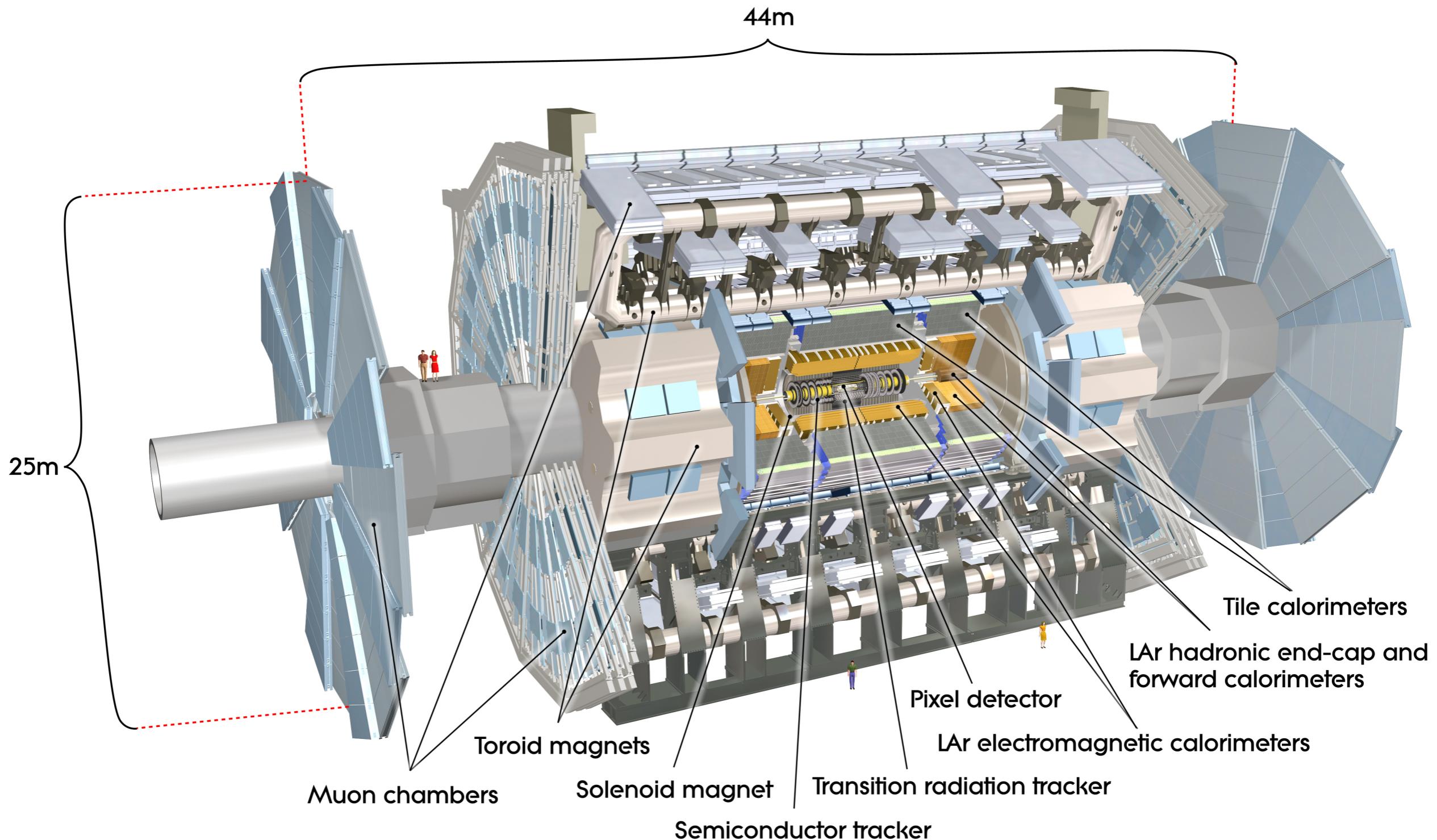
LHC and ATLAS: Luminosity



- LHC has delivered 156 fb^{-1} data, ATLAS has recorded 147 fb^{-1} , and 139 fb^{-1} is regarded good for physics usage
- The pileup μ is caused by multiple pp collisions happening simultaneously
- Vertices reconstructed to reduce the background noise from pileup

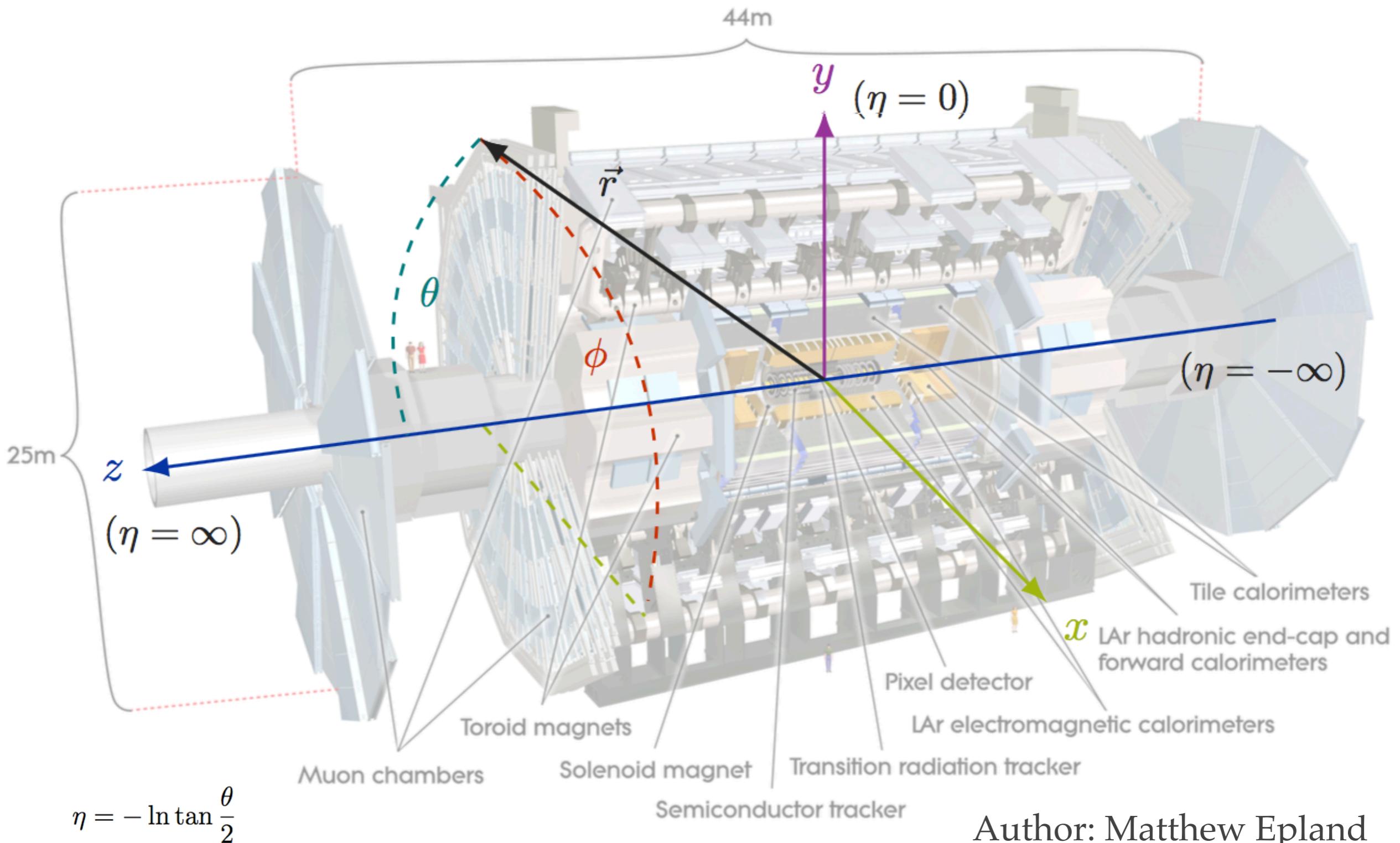


LHC and ALTAS: ATLAS Overview





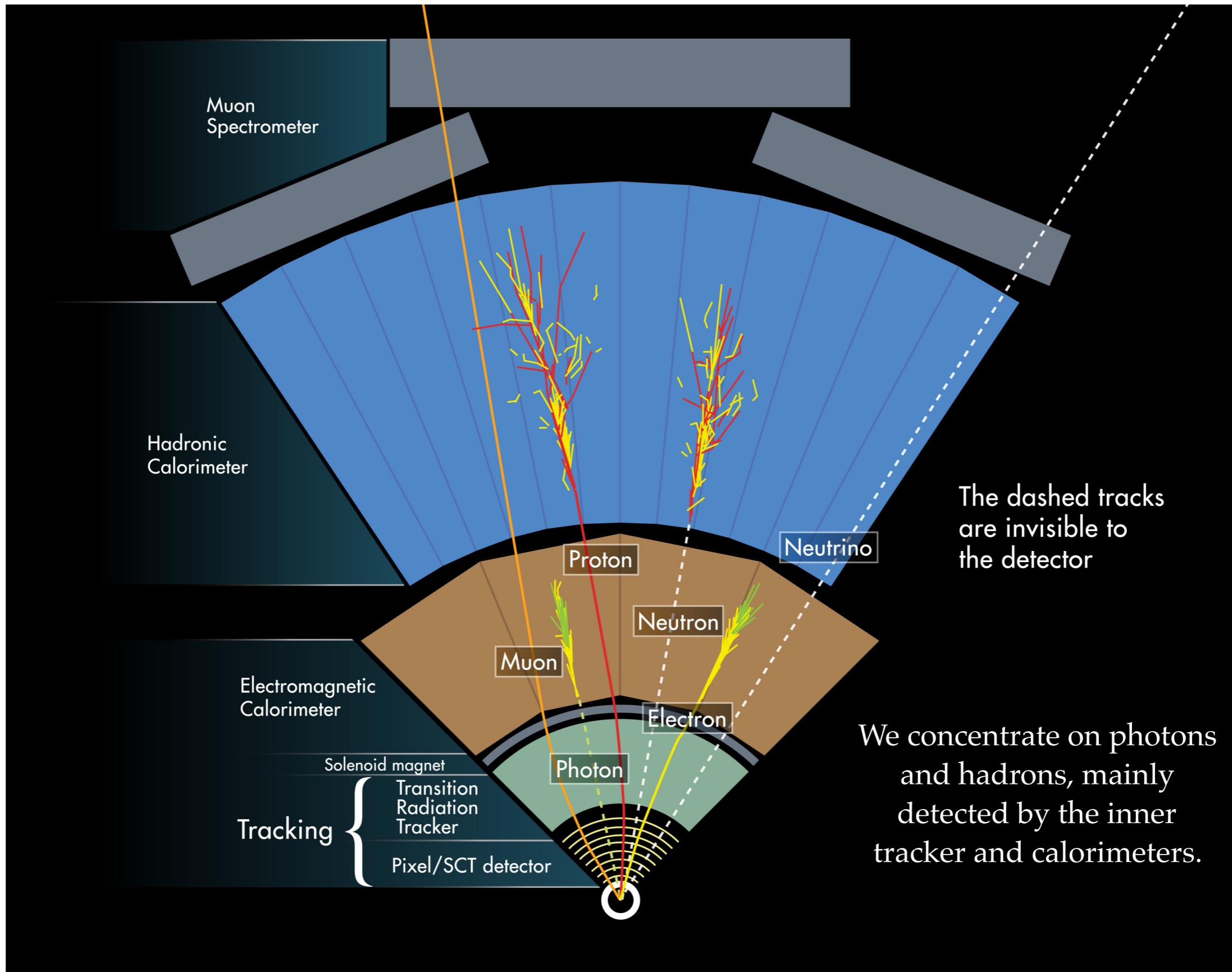
LHC and ALTAS: Coordinate System



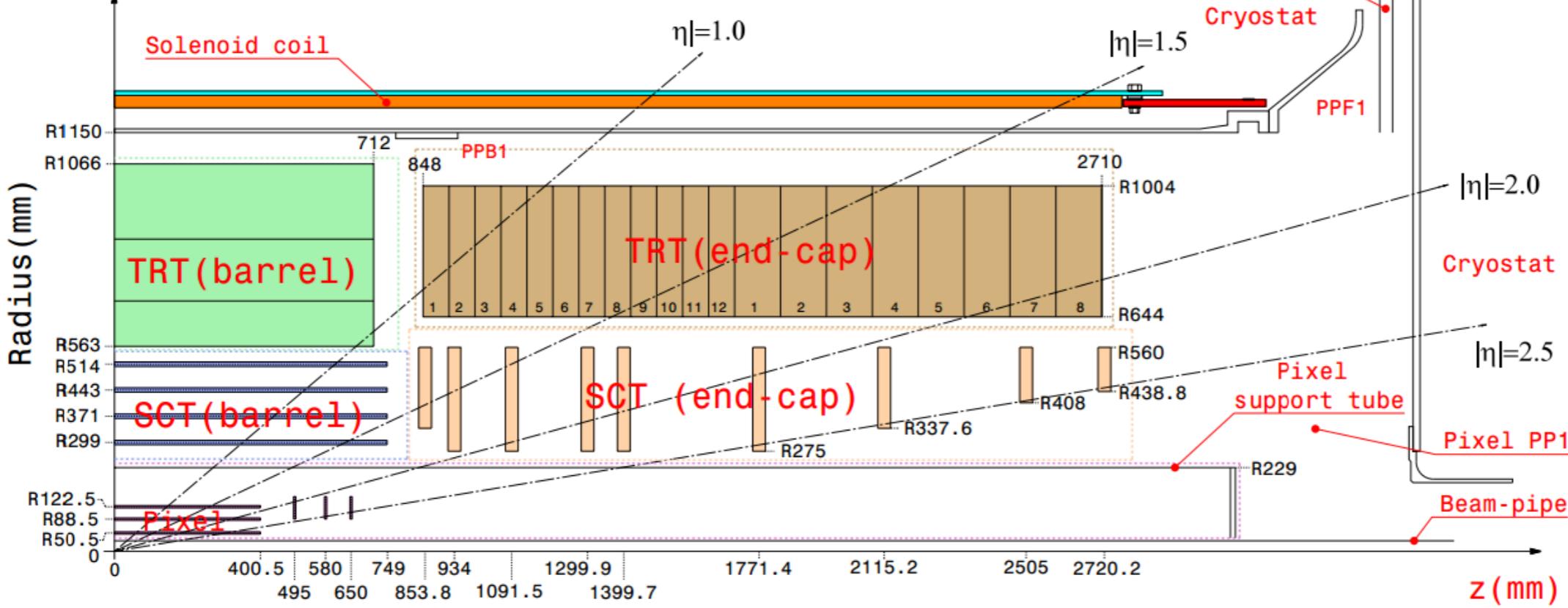
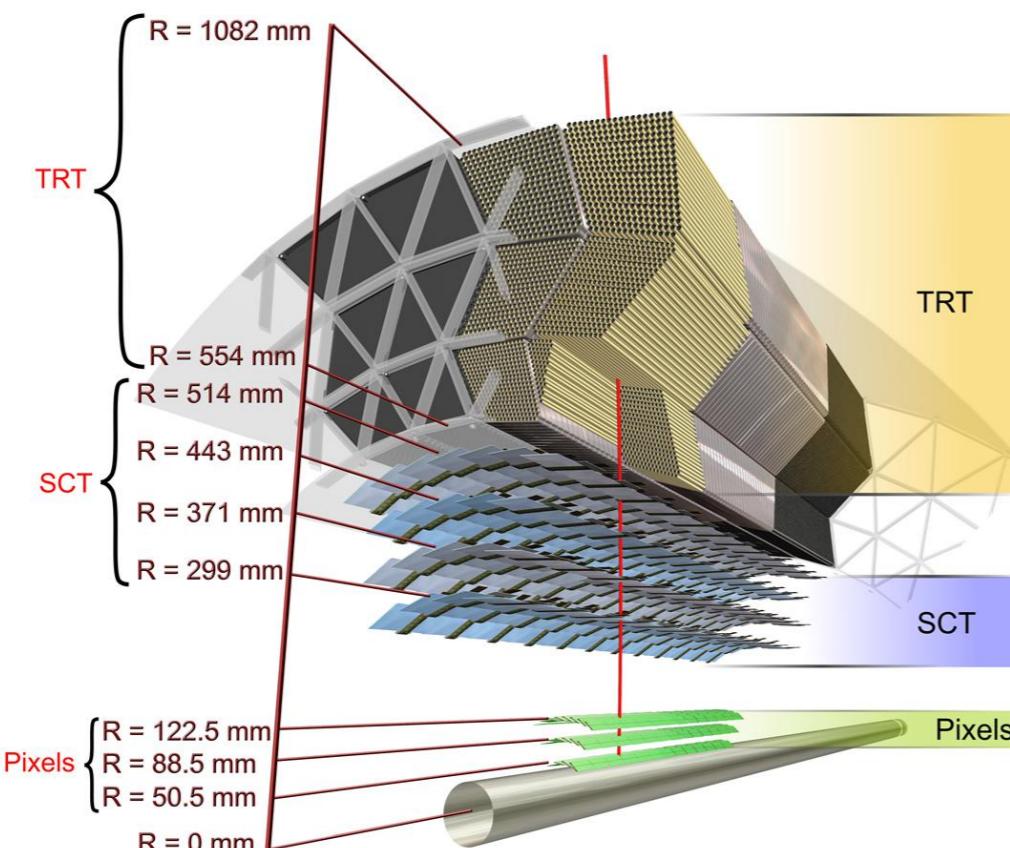
Author: Matthew Epland



LHC and ALTAS: Detecting Particles

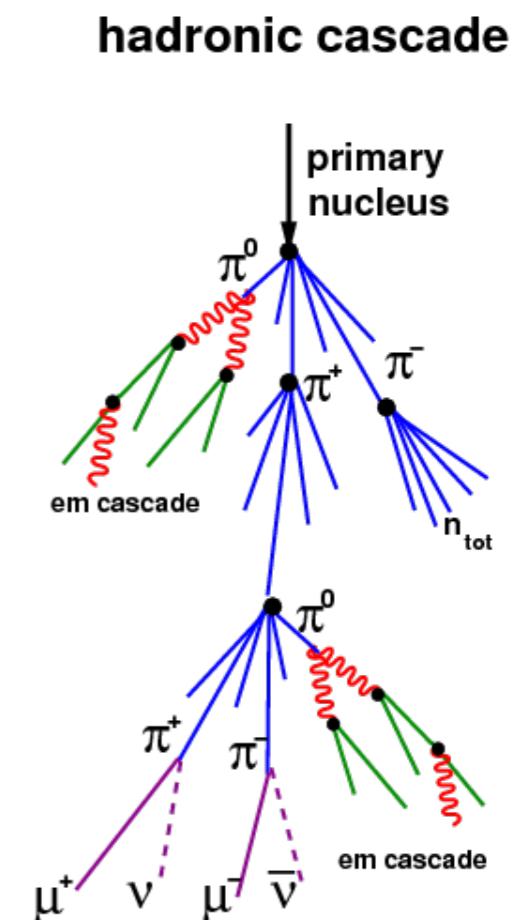
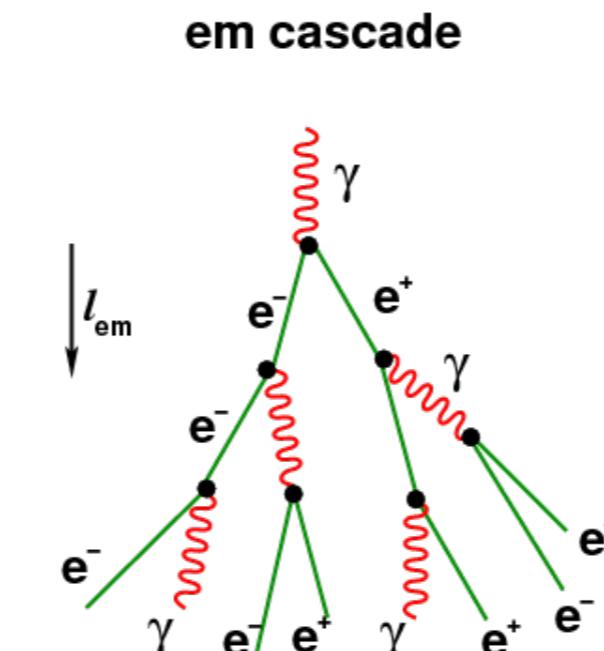
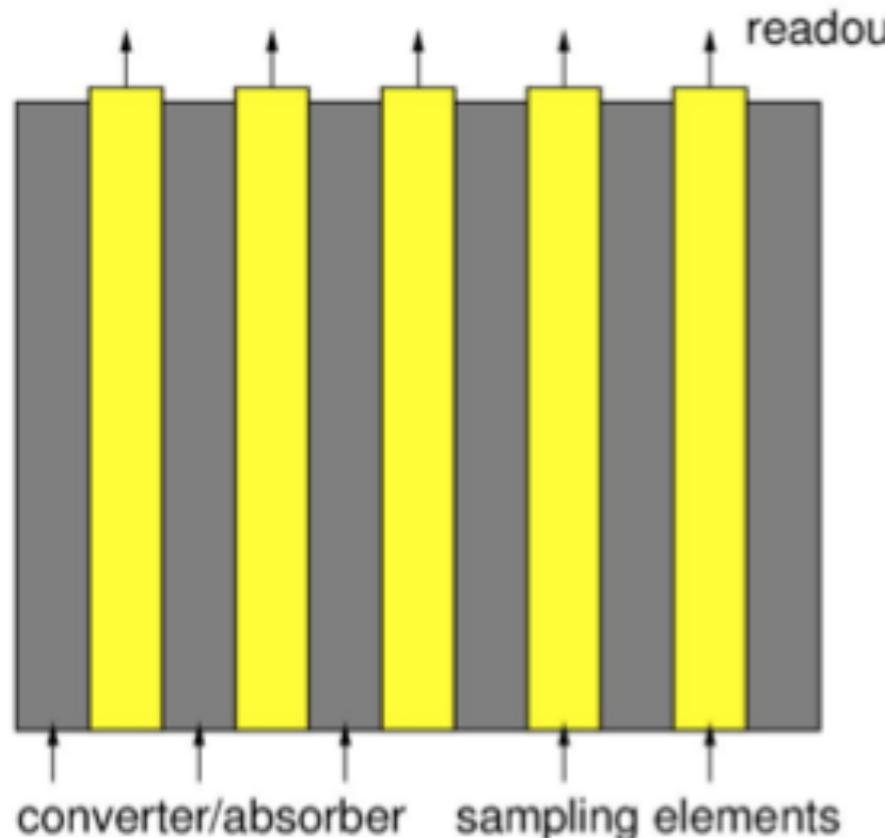


LHC and ALTAIR: Inner Detector



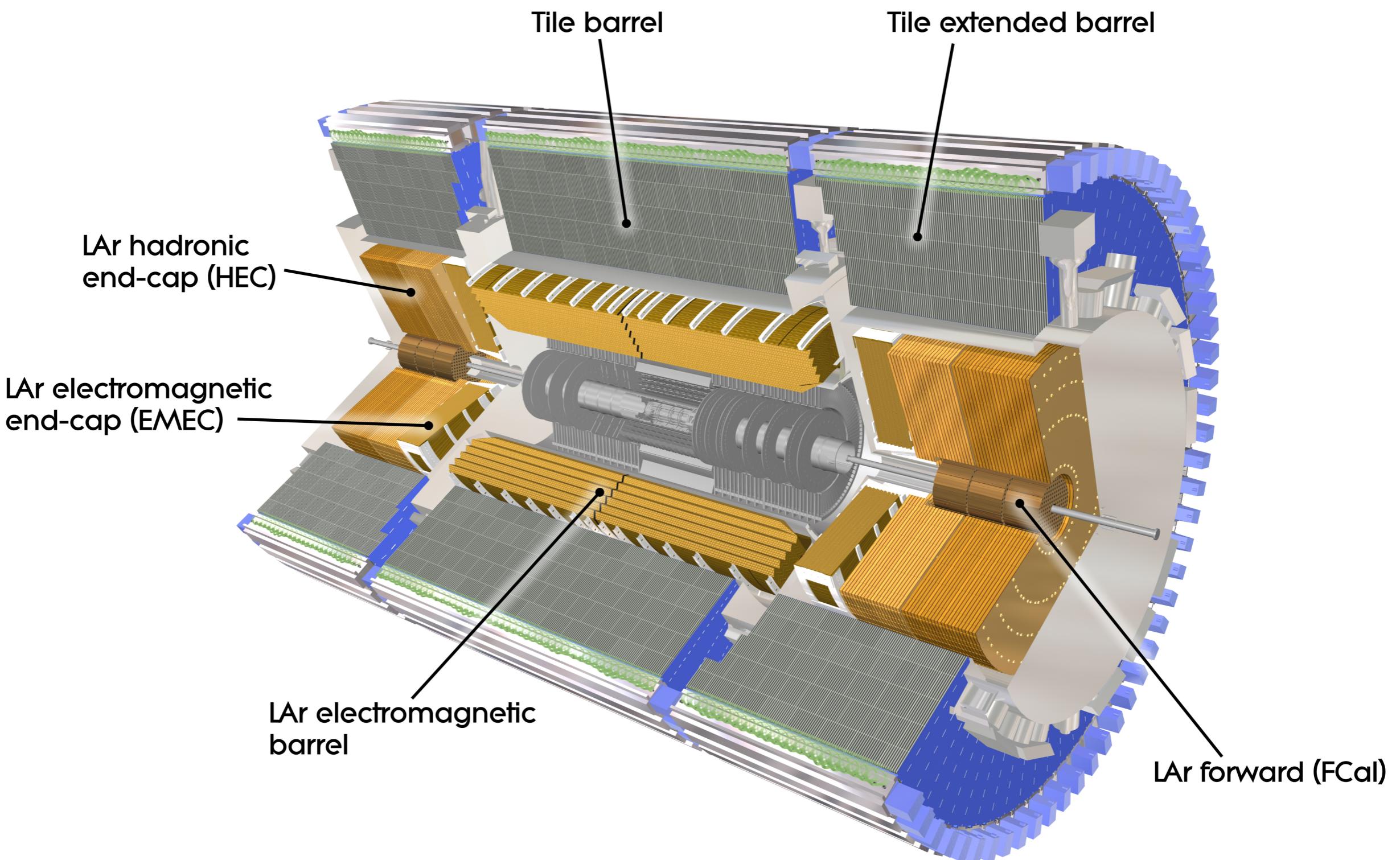
- Inner detector measures the path of charged particles.
- TRT helps discriminating electrons and pions.
- An insertable B-layer is added since 2014 to improve tracking and b-tagging performance.

LHC and ALTA^S: Calorimeters

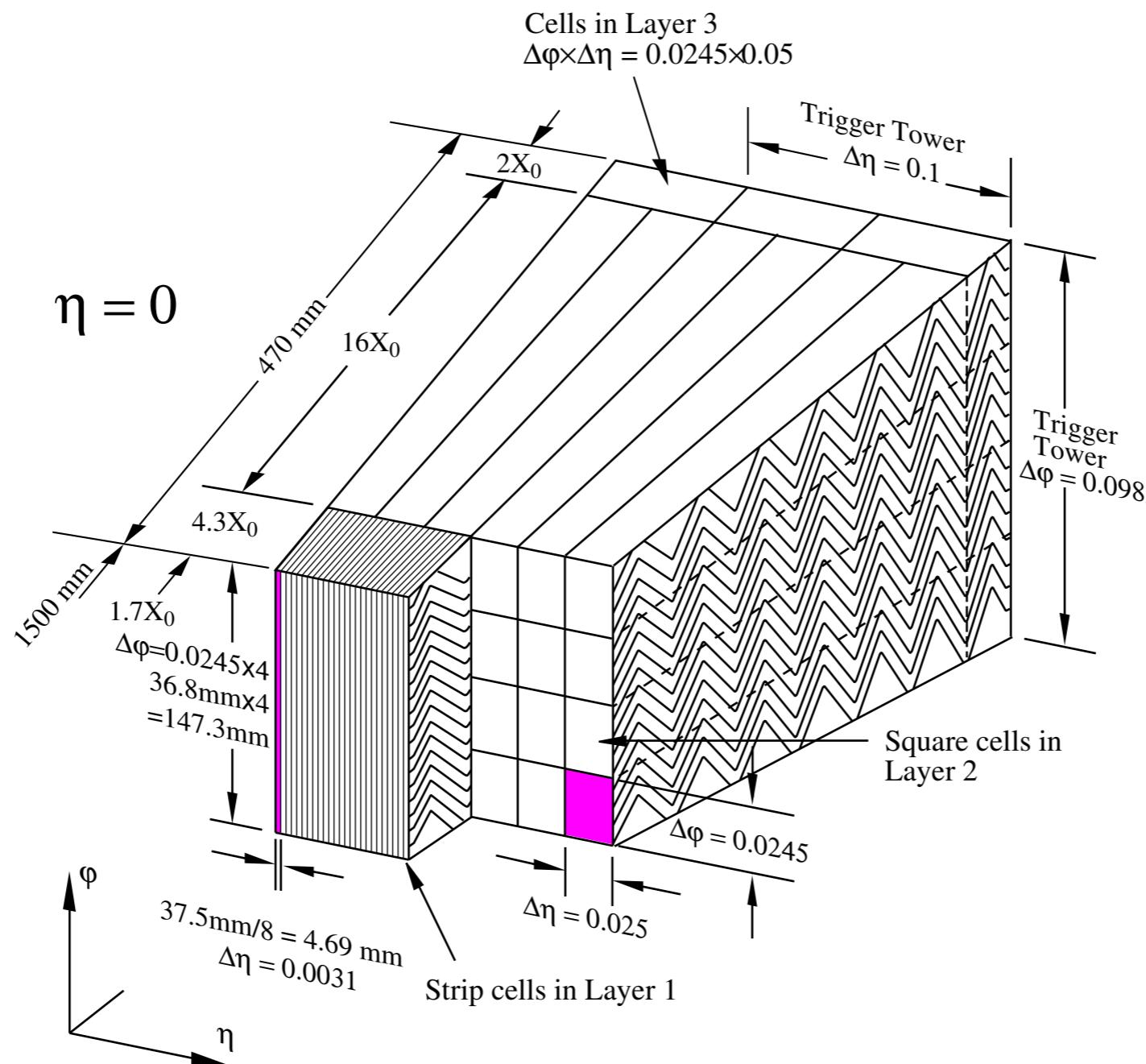


- Calorimeters measure particle energy by absorbing it and producing corresponding electronic signals.
- Usually built by alternating layers of absorber and sampling elements
- EM cascade and hadronic cascade have different features

LHC and ALTAS: Calorimeters



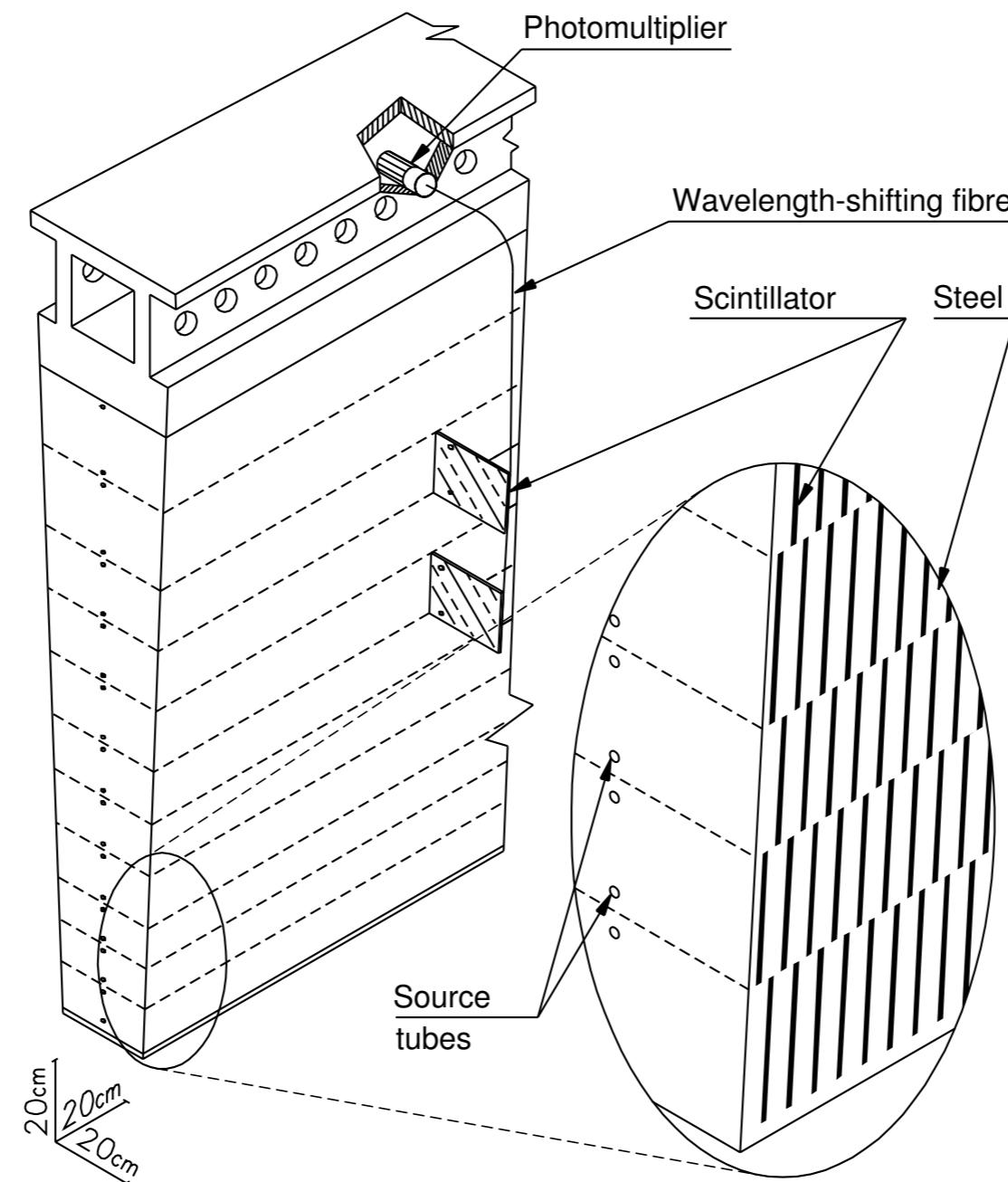
LHC and ALTAS: EM Calorimeter



$$\frac{\sigma(E)}{E} = \frac{10\%}{\sqrt{E(\text{GeV})}} \oplus 0.17\%$$



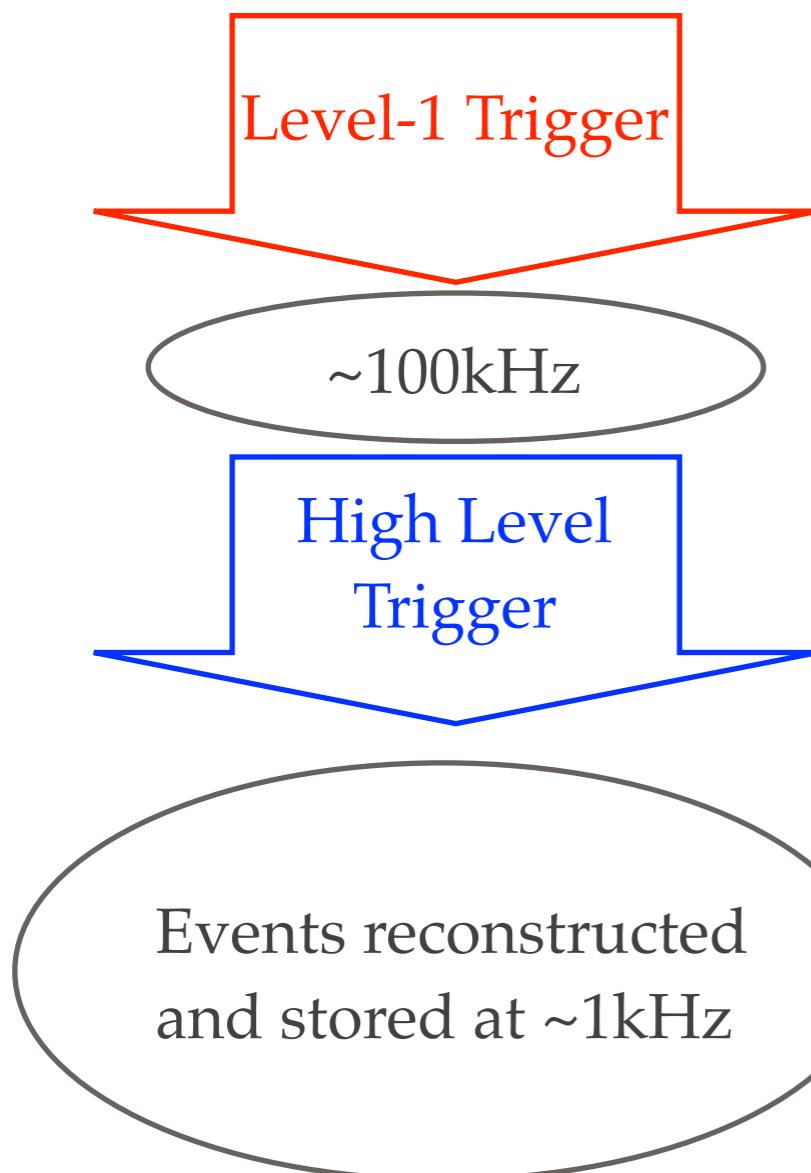
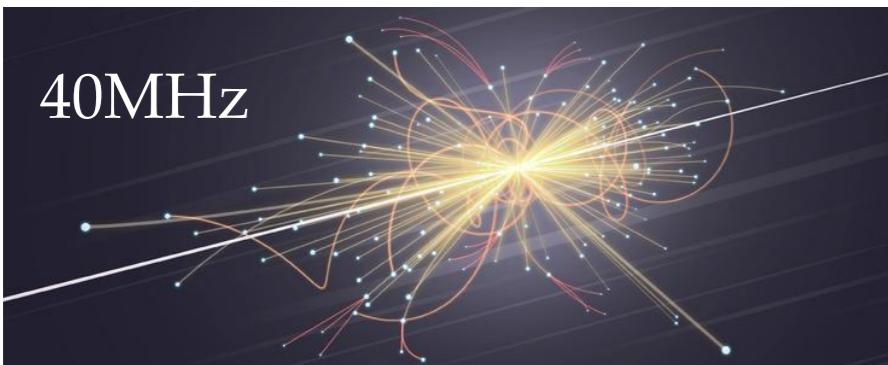
LHC and ALTAS: Hadronic Calorimeter



$$\frac{\sigma(E)}{E} = \frac{56.4\%}{\sqrt{E(\text{GeV})}} \oplus 5.5\%$$



LHC and ALTA^S: Triggers



With 1-2 MB for each event, the data need to be stored in a rate of **40-80 TB/s**.

Hardware based trigger. No complicated algorithms.
(Need to make decision in $2.5 \mu\text{s}$.)
Look for events with interesting features like high pT particles or large MET...

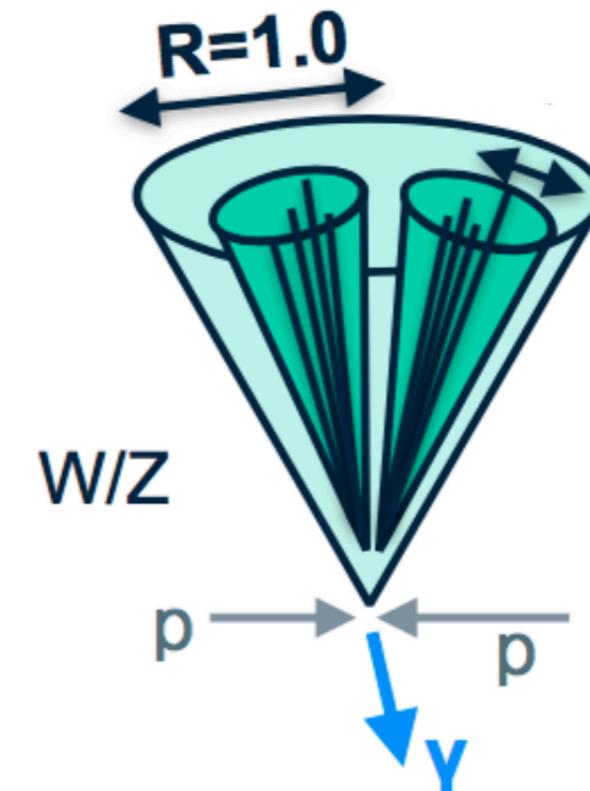
Mostly software based triggers. The average event processing time for HLT is about 40 ms.
Complex algorithms like particle identifications can be applied with much longer processing time.

The final data is recorded at a rate of **1-2 GB/s**



Performance: Physics Objects of Interests

BSM Model	Boson Spin	Process
Singlet Scalar	0	$gg \rightarrow X \rightarrow \gamma Z (\rightarrow q\bar{q})$
Heavy Vector Triplet	1	$qq' \rightarrow X \rightarrow \gamma W (\rightarrow qq')$
Deconstruction of Extra Dimensions	2	$gg \rightarrow X \rightarrow \gamma Z (\rightarrow q\bar{q})$
	2	$q\bar{q} \rightarrow X \rightarrow \gamma Z (\rightarrow q\bar{q})$

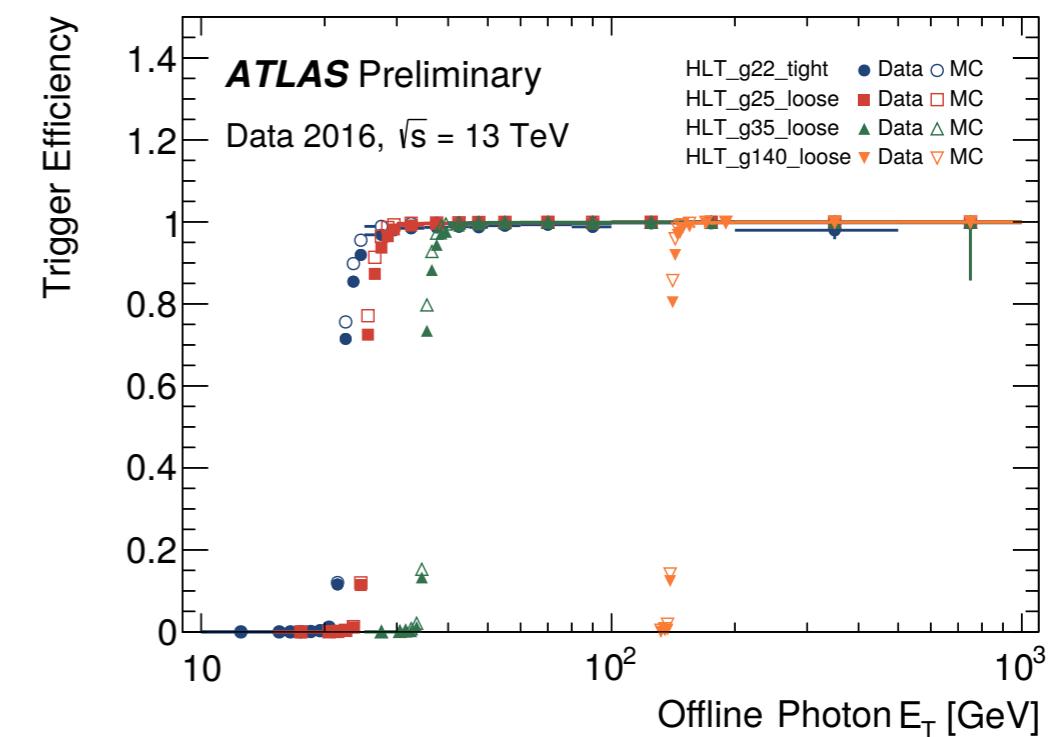
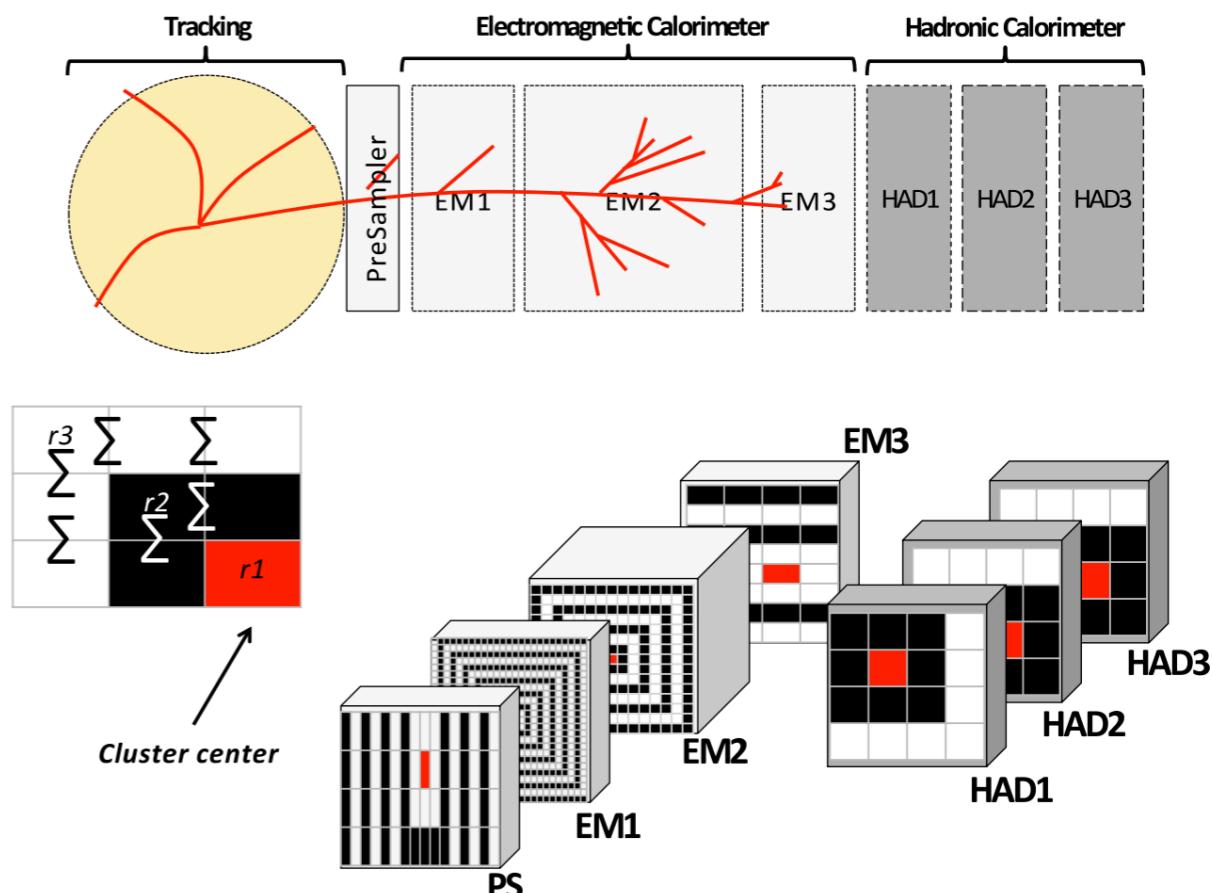


- **Photon** with high energy and passing the identification
- **W/Z boson** with high energy, and decaying hadronically

Performance: Photon Measurement



- Inner tracker information is used for identification of particles
- Calorimeter information is used to determine the energy of photon
- An algorithm called “ringer” is used to improve the identification and reconstruction performance
 - ❖ A BDT is trained with the inputs of the total energy deposit for each of the rings around the central cell
 - ❖ It saves processing time and improve rejection of background with similar efficiency
- The trigger used in this search (HLT_g140_loose) is almost 100% efficient, and the turn on is much below our selection cut.





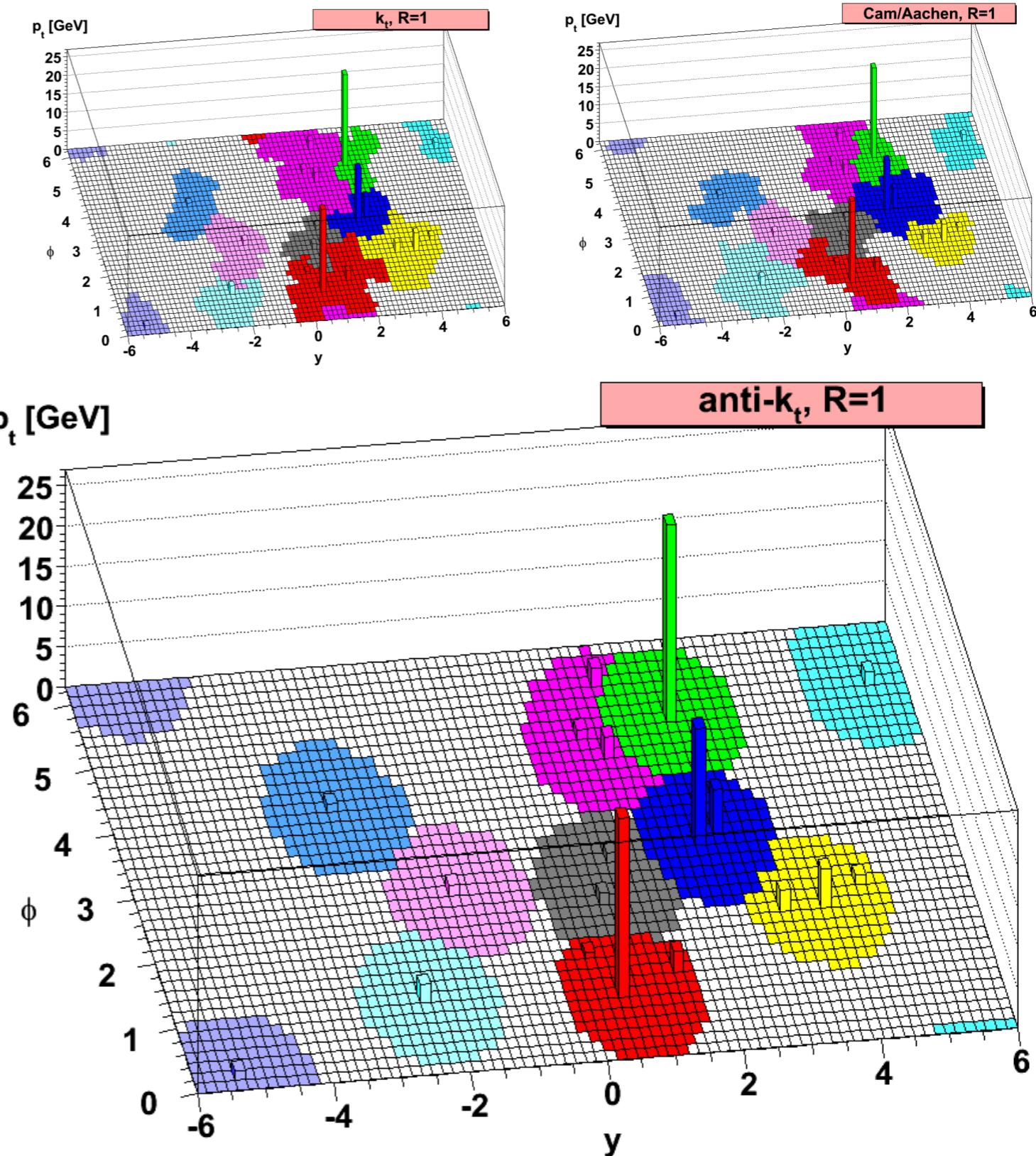
Performance: Boosted Jet Reconstruction

The clustering algorithm is conducted by repetitively grouping the two closest cells into a cluster, with the distance d between i-th and j-th cell defined as:

$$d_{ij} = \min(k_{ti}^{2p}, k_{tj}^{2p}) \frac{\Delta R_{ij}^2}{R_{jet}^2}$$

$$d_{iB} = k_{ti}^{2p}$$

$p = -1$ is taken for the anti- k_t algorithm, k_{ti} means transverse momentum for i-th cell, R_{jet} is the defined radius of jet, $\Delta R^2 = \Delta\phi^2 + \Delta\eta^2$ and d_{iB} is a baseline which implies the i-th jet should be regarded as a separate jet.





Performance: Identification of W/Z jets

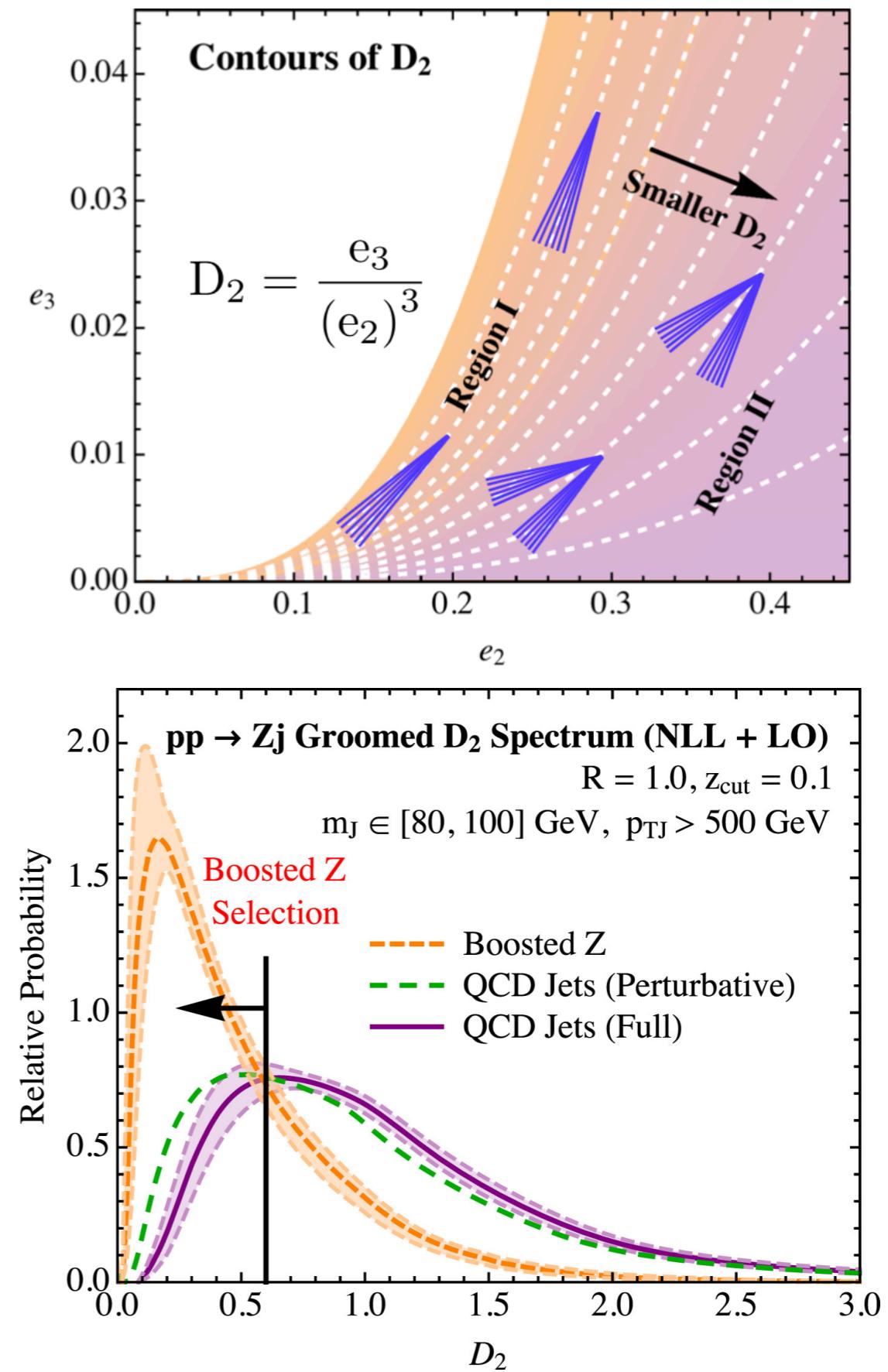
The high level variable D2 used for boson tagging is defined as:

$$\epsilon_2^{(\alpha)} = \frac{1}{E_J^2} \sum_{i < j}^N E_i E_j \left(\frac{2p_i \cdot p_j}{E_i E_j} \right)^{\alpha/2}$$

$$\epsilon_3^{(\alpha)} = \frac{1}{E_J^3} \sum_{i < j < k}^N E_i E_j E_k \left(\frac{2p_i \cdot p_j}{E_i E_j} \frac{2p_i \cdot p_k}{E_i E_k} \frac{2p_j \cdot p_k}{E_j E_k} \right)^{\alpha/2}$$

$$D_2^{(\alpha)} = \frac{\epsilon_3^{(\alpha)}}{(\epsilon_2^{(\alpha)})^3}$$

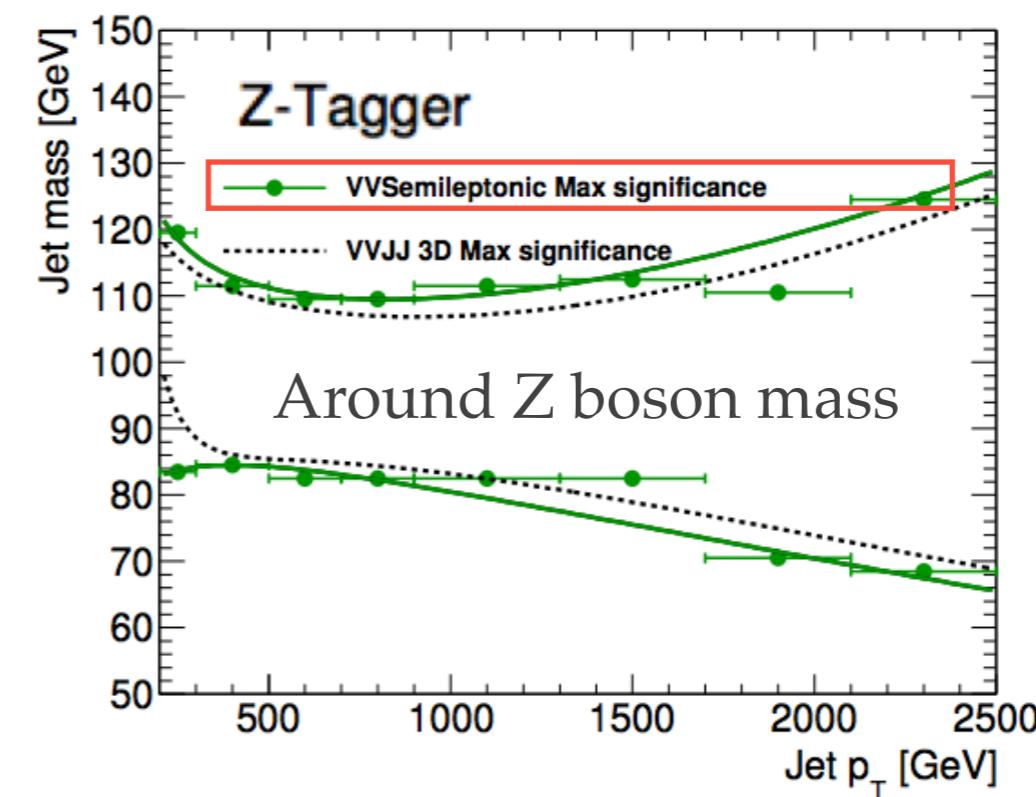
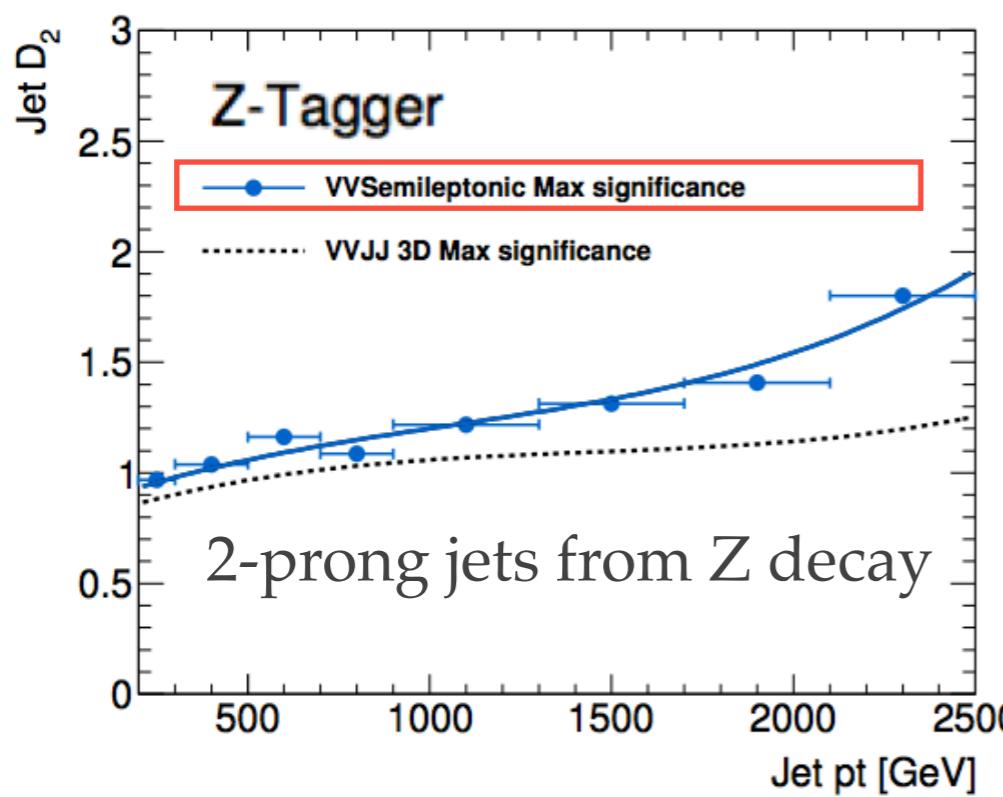
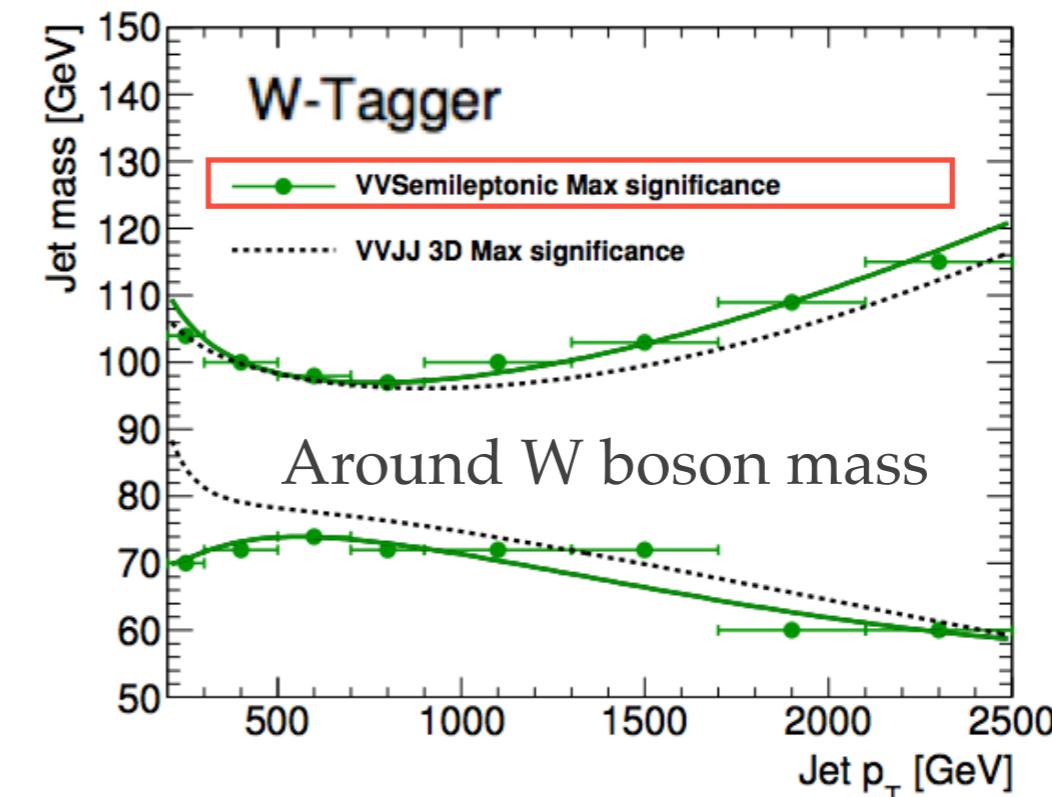
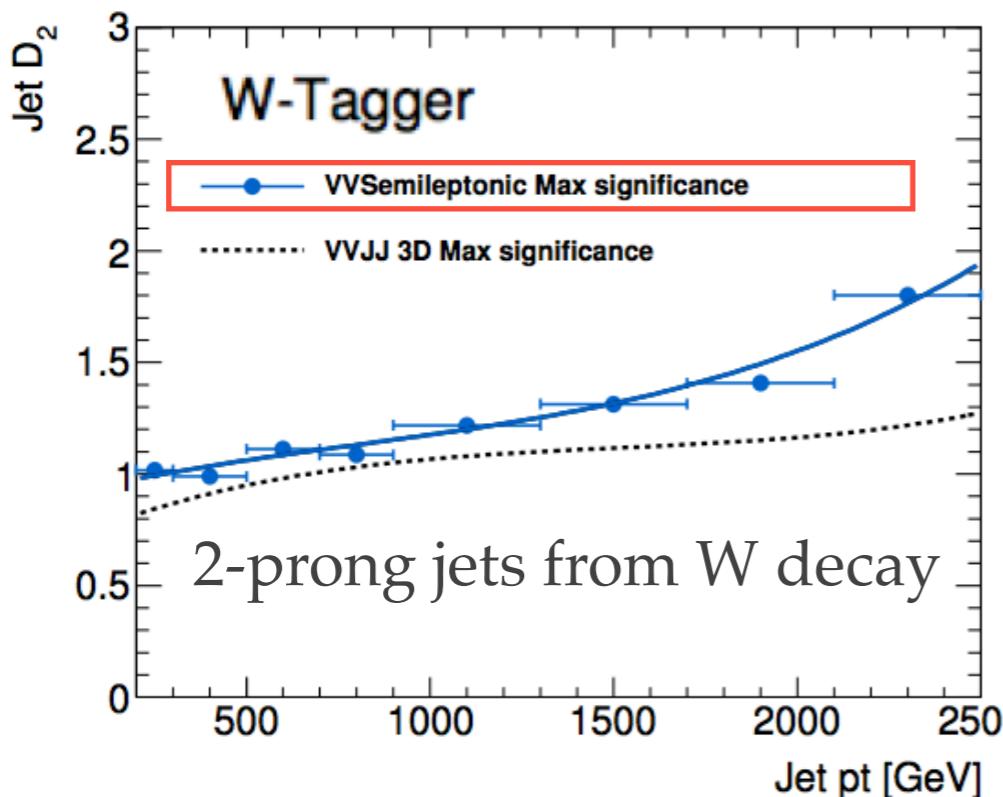
The D₂ is used as convention for abbreviation of D₂ = D₂^(α = 1)





Performance: Identification of W/Z jets

The D₂ variable and the mass of jet are combined to identify jets from boosted bosons.

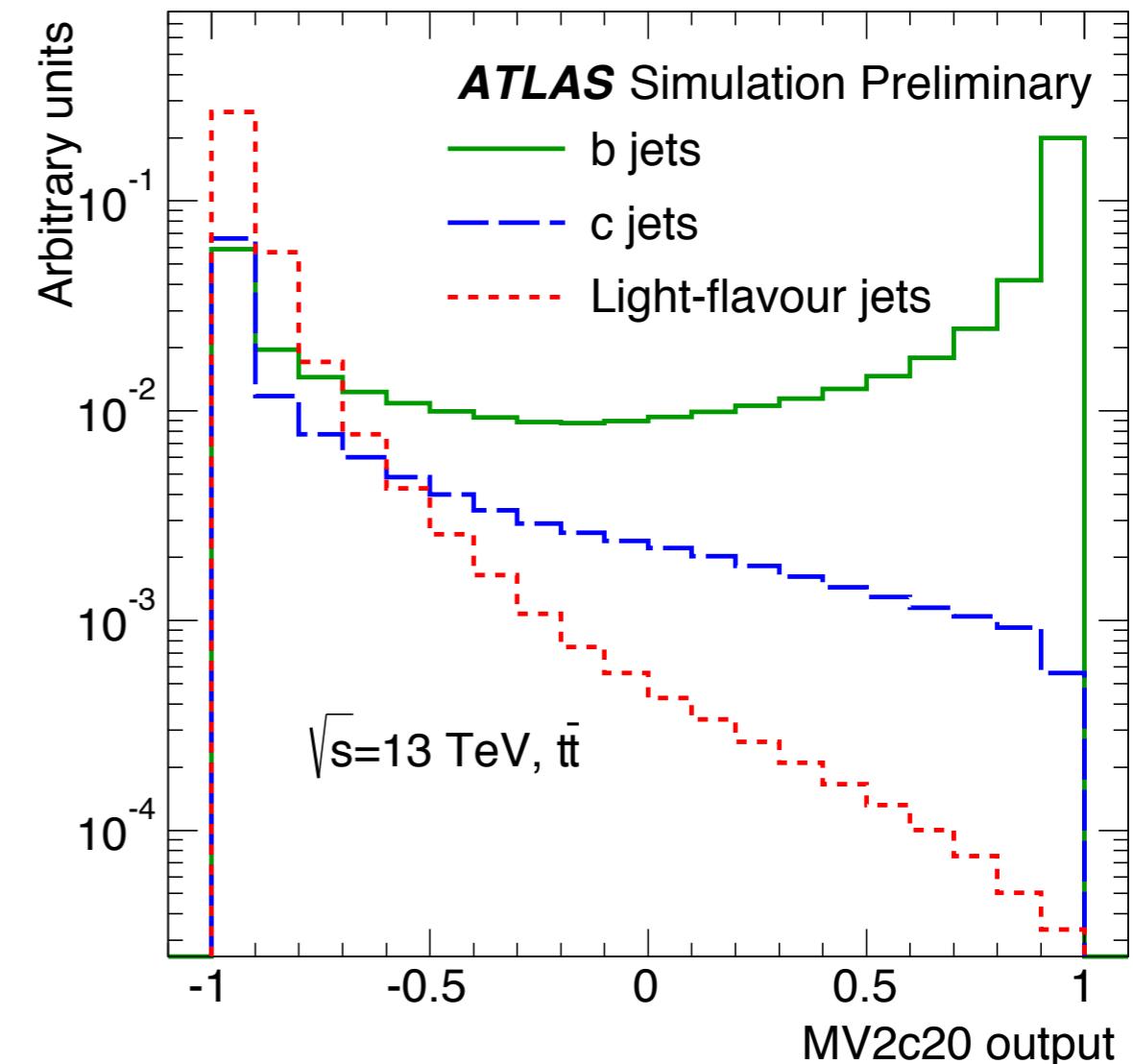




Performance: Identification of b-hadron jets

A boosted decision tree (BDT) is trained to optimize the identification of b-hadron jets, with inputs from multiple algorithms.

This BDT has different working points with different efficiencies. The most loose one with 60% efficiency is used in this search.



Efficiency	Cut Value	c-jet Rejection	τ -jet Rejection	Light-jet Rejection
60%	0.4496	21	93	1900
70%	-0.0436	8.1	26	449
77%	-0.4434	4.5	10	140
85%	-0.7887	2.6	3.8	28



Analysis: MC Samples

MC Sample	Process	Generator	Boson spin	W/Z Boson Polarization
Z γ Spin 0	$gg \rightarrow X \rightarrow \gamma Z (\rightarrow q\bar{q})$	Powheg+Pythia8	0	Transverse
W γ Spin 1	$q\bar{q}' \rightarrow X^\pm \rightarrow \gamma W^\pm (\rightarrow q\bar{q}')$		1	Longitudinal
qqZ γ Spin 2	$q\bar{q} \rightarrow X \rightarrow \gamma Z (\rightarrow q\bar{q})$	MadGraph + Pythia8	2	Transverse
ggZ γ Spin 2	$gg \rightarrow X \rightarrow \gamma Z (\rightarrow q\bar{q})$		2	Transverse
SM γ +Jet (Background)		Sherpa NLO	-	-



Analysis: Event Selection

Baseline Selection:

Trigger: HLT_g140_loose (the highest pT photon trigger)

Photon pass the tight identification and tight isolation selection

$P_T(\gamma) > 200 \text{ GeV}$ and $|\eta(\gamma)| < 1.37$

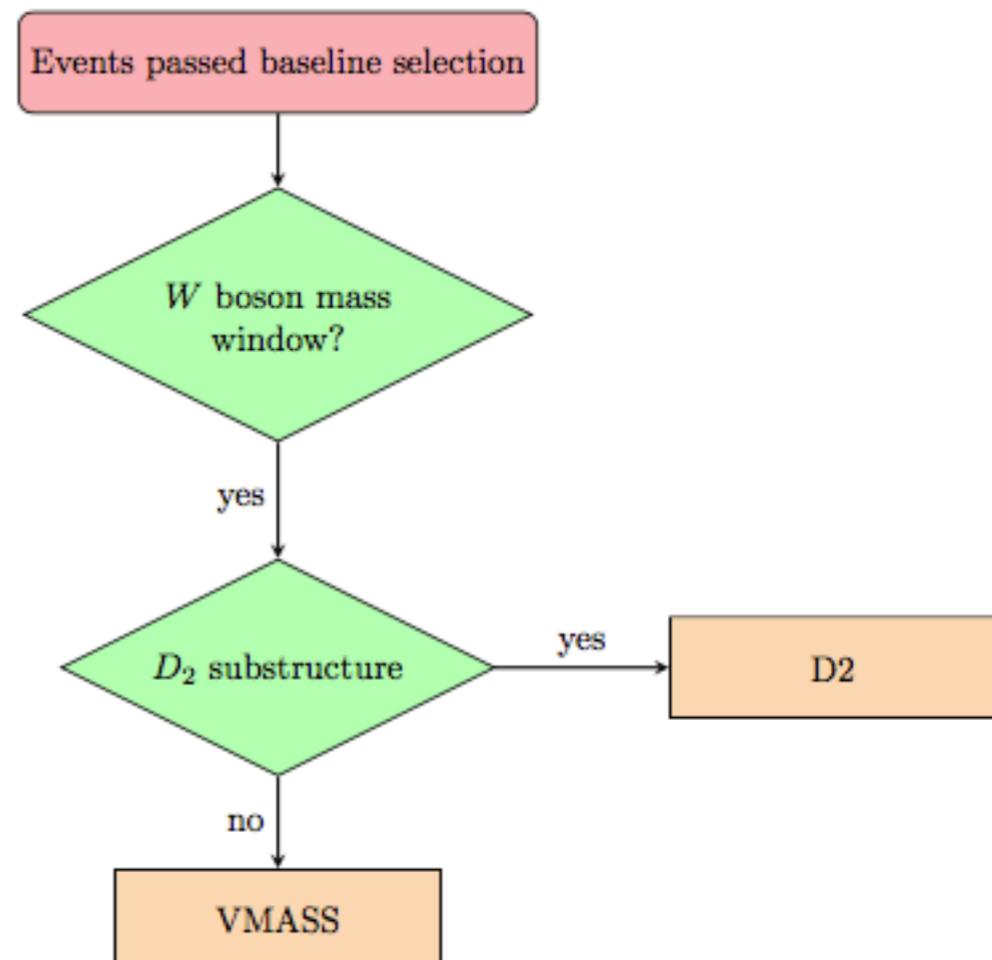
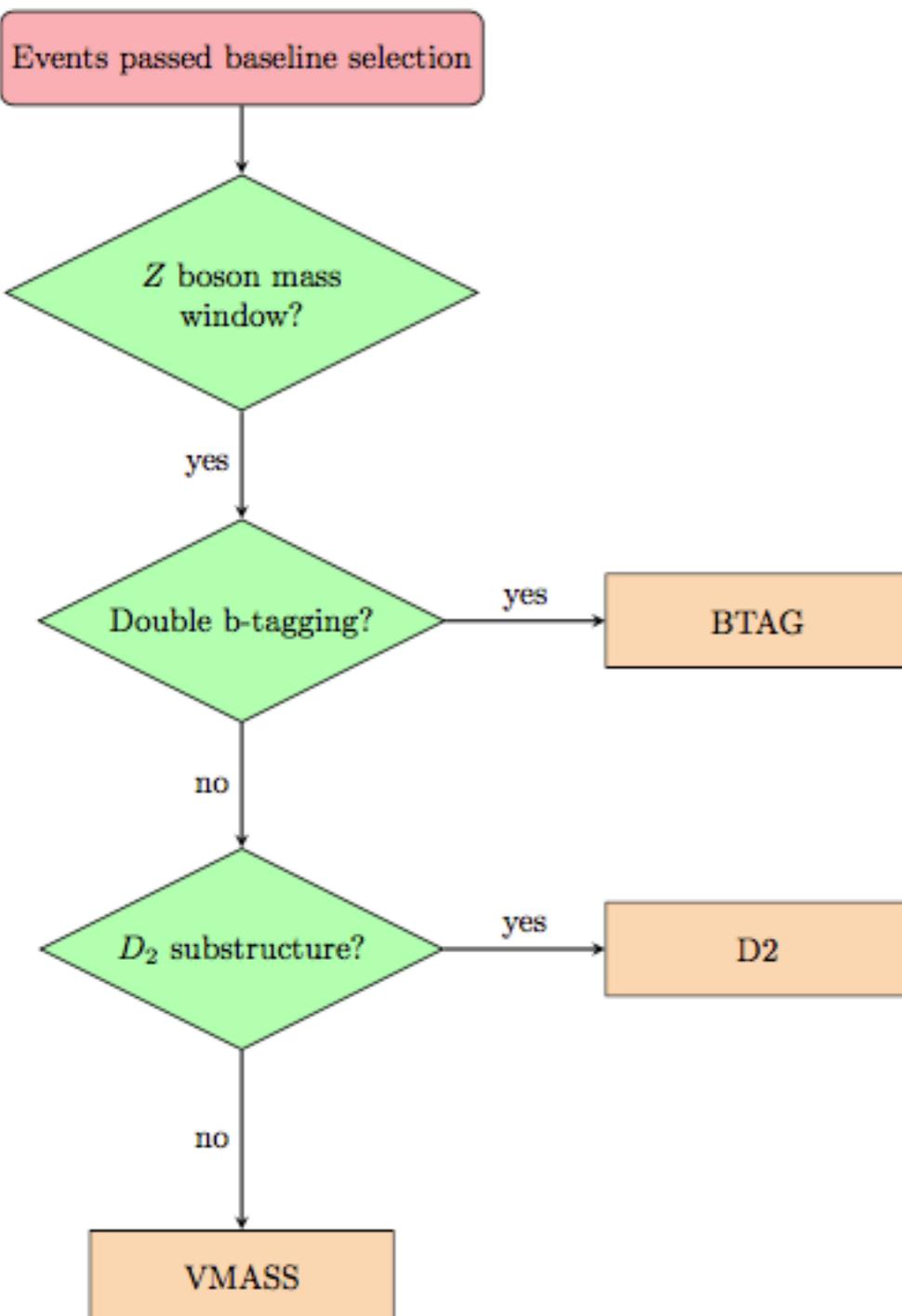
The boosted jet is not overlap with selected photon

$P_T(\text{jet}) > 200 \text{ GeV}$ and $|\eta(\text{jet})| < 2.0$

An additional photon/jet pT cut is applied after categorization

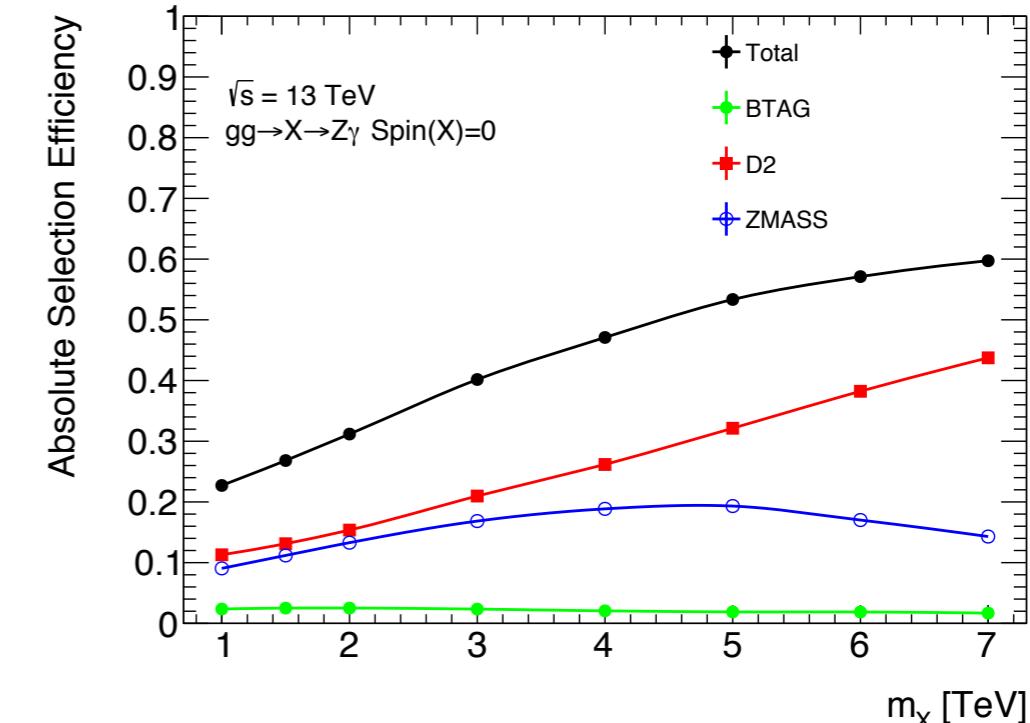
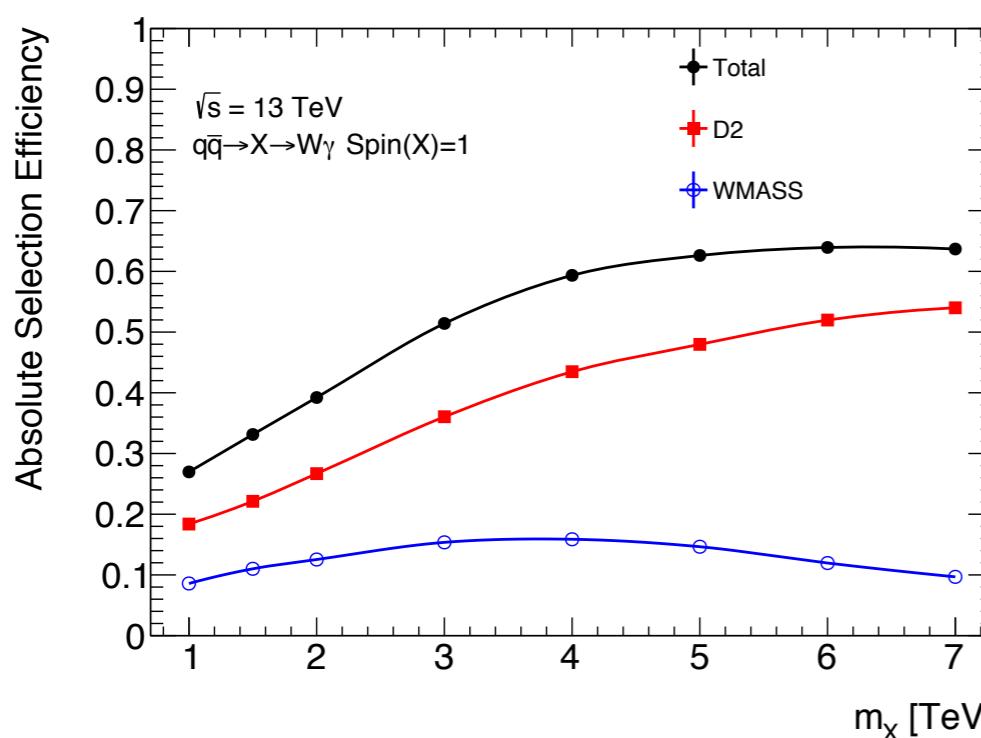
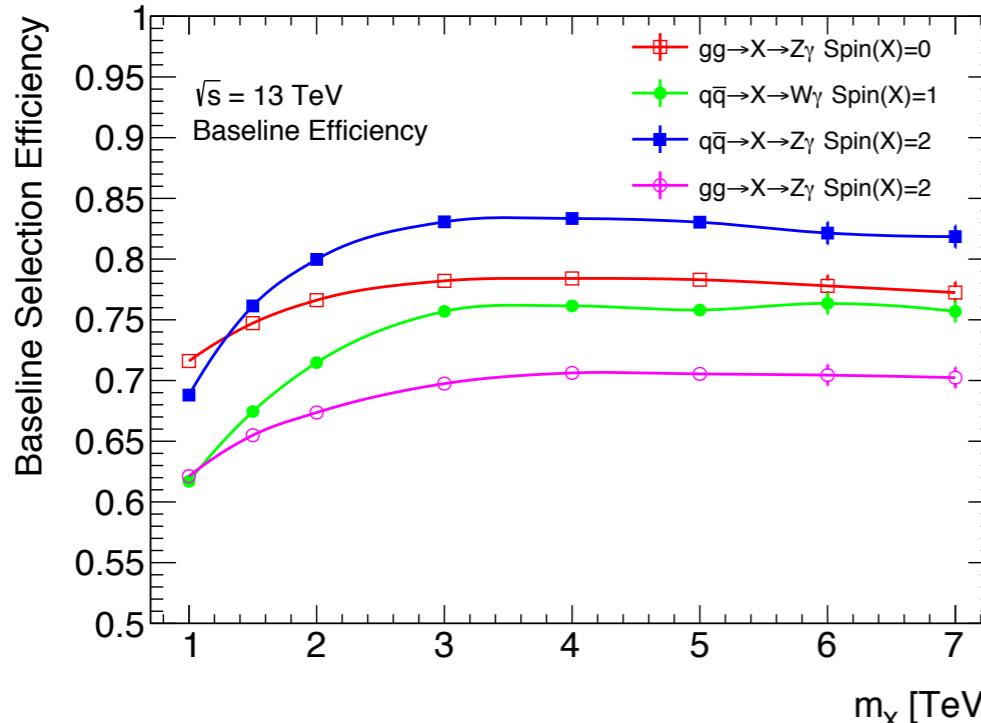


Analysis: Categorization





Analysis: Signal Efficiency



The signal categorization efficiency here is after applying the additional P_T cut

Due to the boson polarization, the boson tagging on W boson is more efficient than Z boson

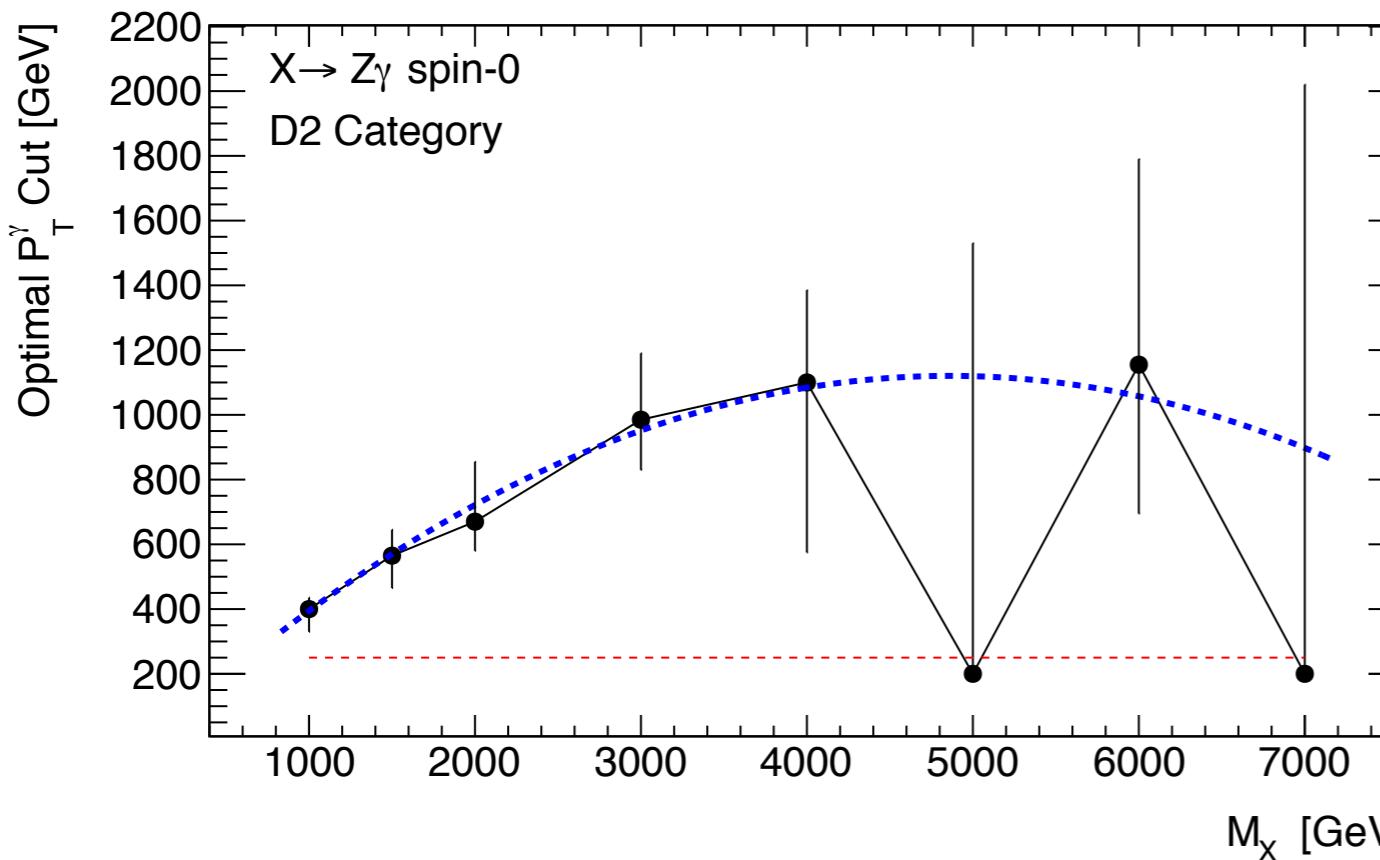
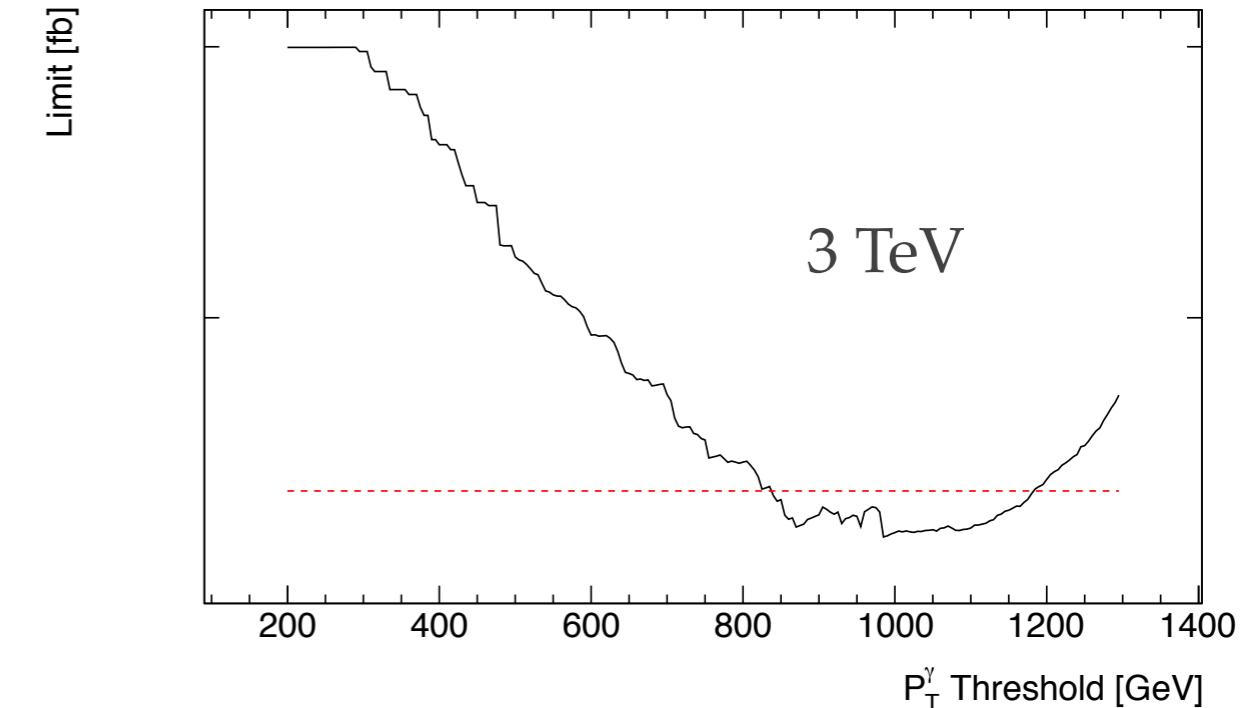
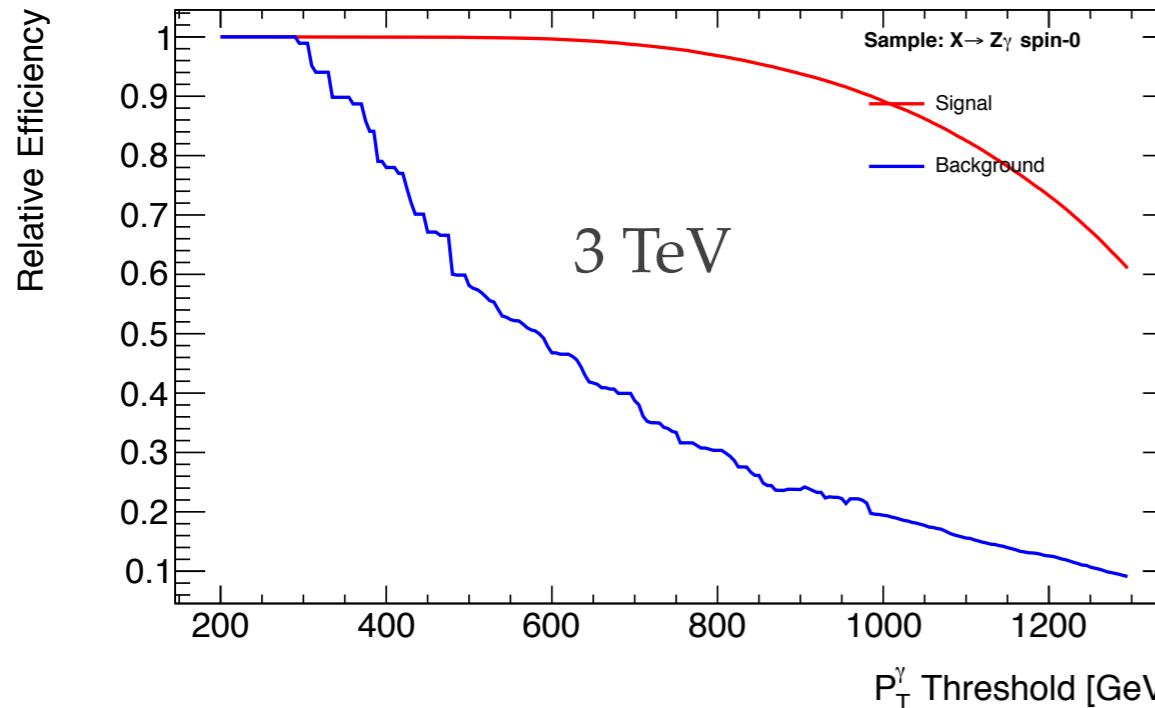
Variable-radius jet b-tagging making b-tag efficiency smooth after 3 TeV

The total number excluded the events filtered by the additional pT cut on non-btag categories



Analysis: Further Optimization

P_T cut is optimized in D2 and Vmass categories



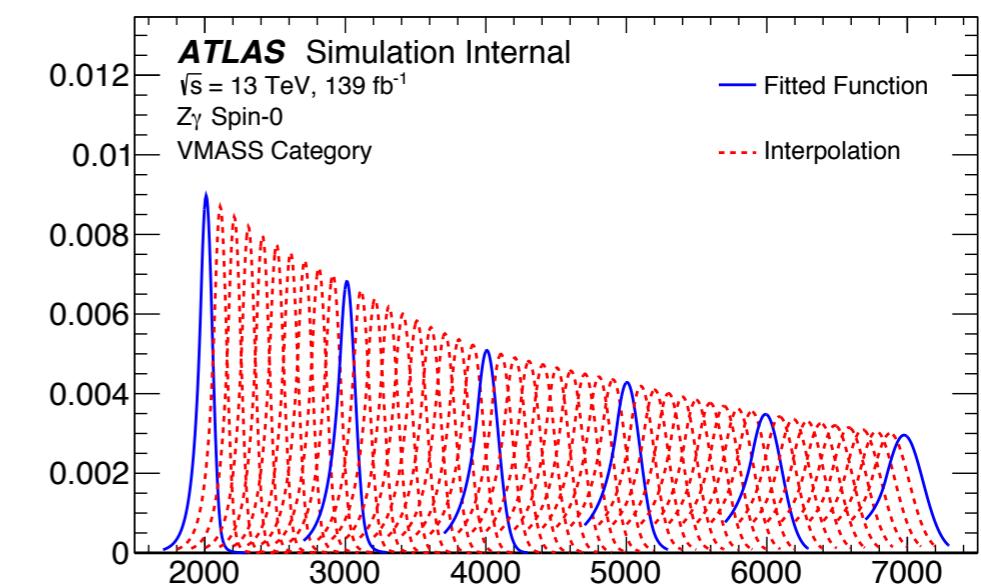
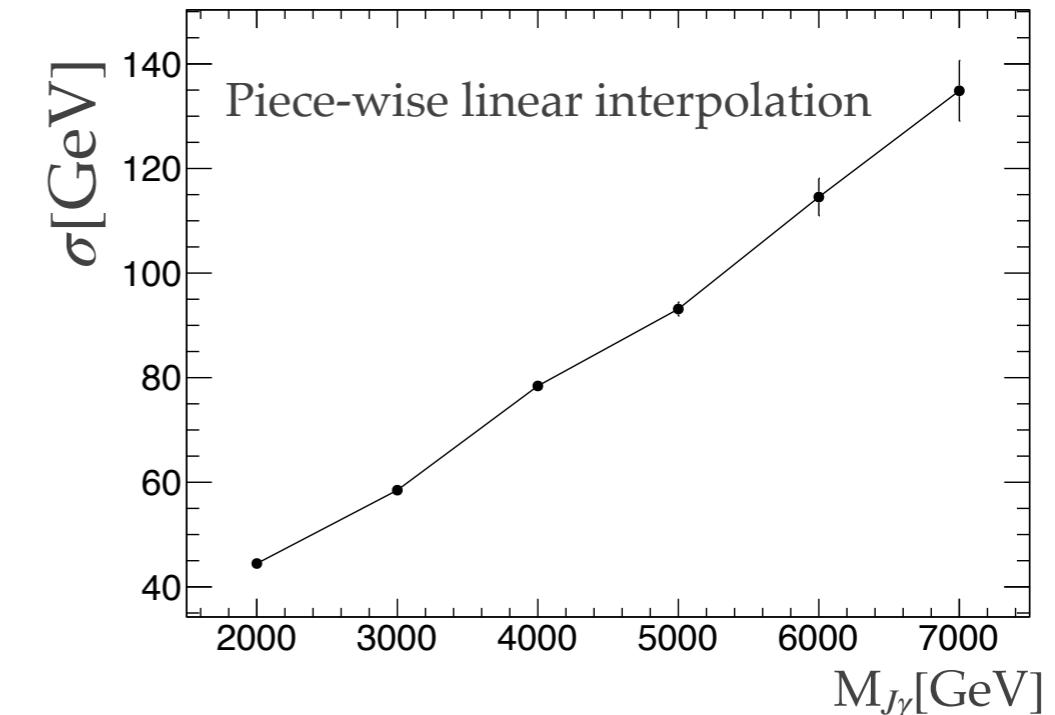
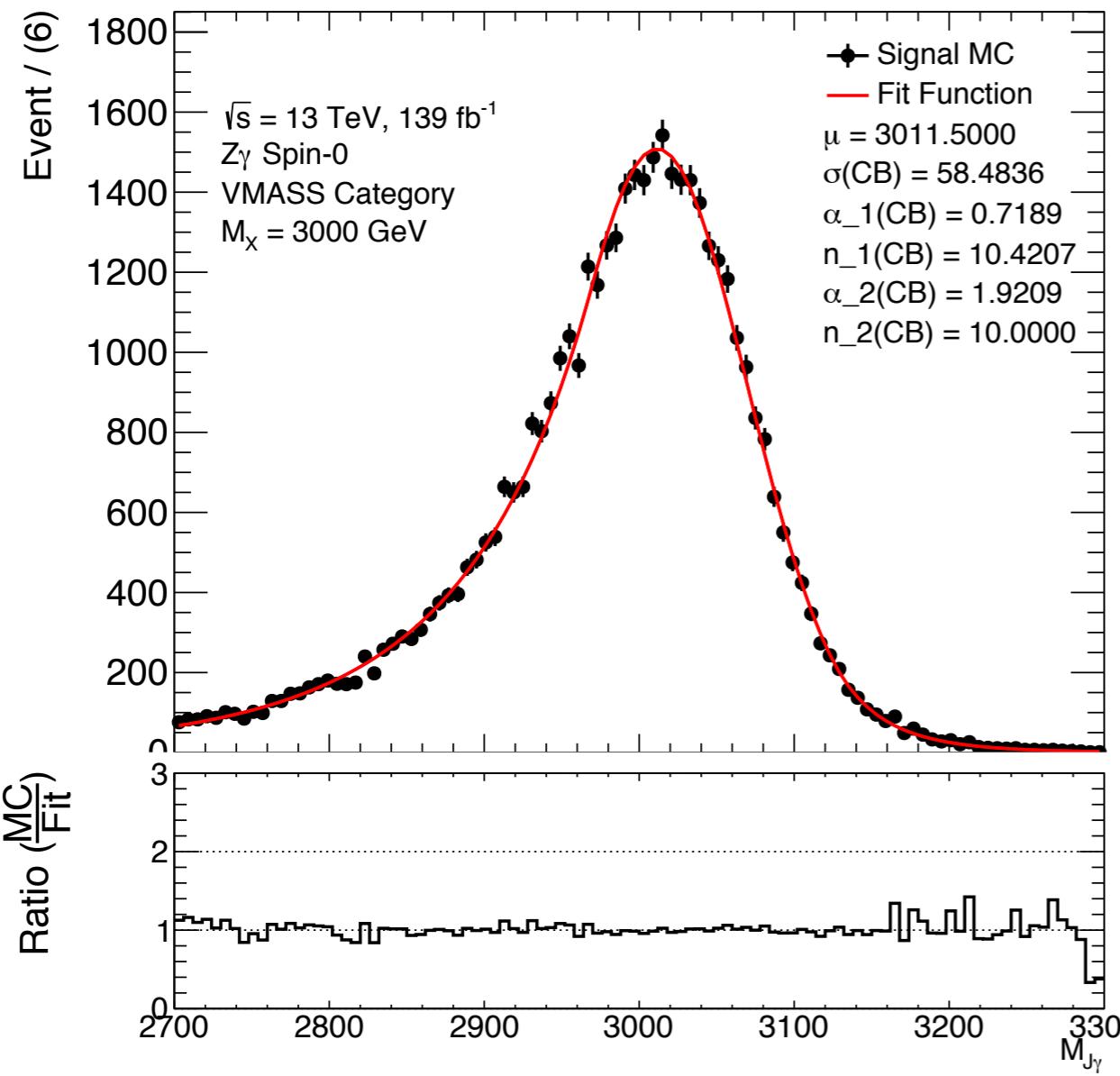
- The motivation of this cut is on the difference of pT distribution between signal and bkg
- For each category and mass point, scan through all possible pT cuts with 5 GeV step
- The optimal cut is taken based on the statistical only limit (asymptotic calculation)
- The error on each point in right bottom plot is set by the cut value window with less than 1.05 optimal limit
- A parabola is fitted for cut values, and after the maximum of parabola, cut value become constant



Analysis: Signal Modelling

$$DSCB(m; N, \mu, \sigma, \alpha_1, n_1, \alpha_2, n_2) =$$

$$N \cdot \begin{cases} \left(\frac{n_1}{|\alpha_1|} \right)^{n_1} \exp \left(-\frac{|\alpha_1|^2}{2} \right) \left(\frac{n_1}{|\alpha_1|} - |\alpha_1| - \frac{m-\mu}{\sigma} \right)^{-n_1} & \frac{m-\mu}{\sigma} \leq -\alpha_1 \\ \exp \left(-\frac{(m-\mu)^2}{2\sigma^2} \right) & -\alpha_1 < \frac{m-\mu}{\sigma} \leq \alpha_2 \\ \left(\frac{n_2}{|\alpha_2|} \right)^{n_2} \exp \left(-\frac{|\alpha_2|^2}{2} \right) \left(\frac{n_2}{|\alpha_2|} - |\alpha_2| + \frac{m-\mu}{\sigma} \right)^{-n_2} & \alpha_2 < \frac{m-\mu}{\sigma} \end{cases}$$



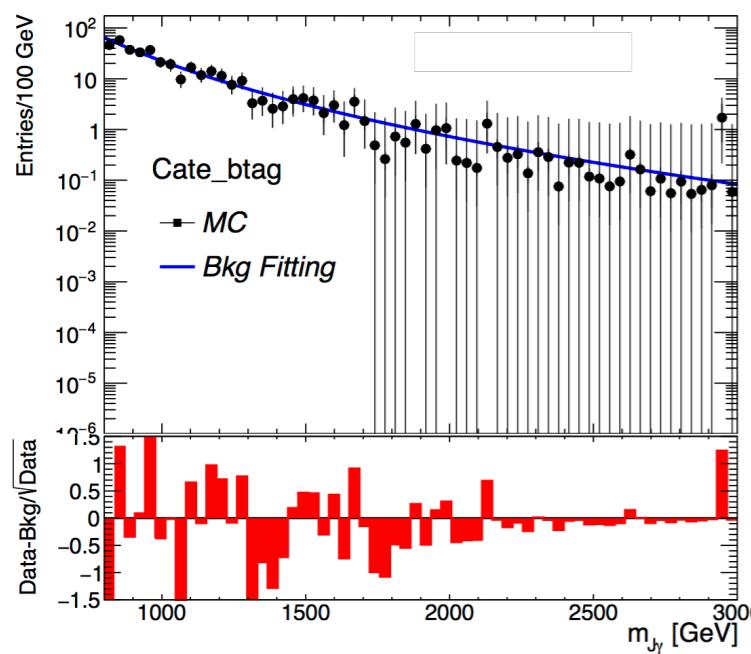


Analysis: Background Estimation

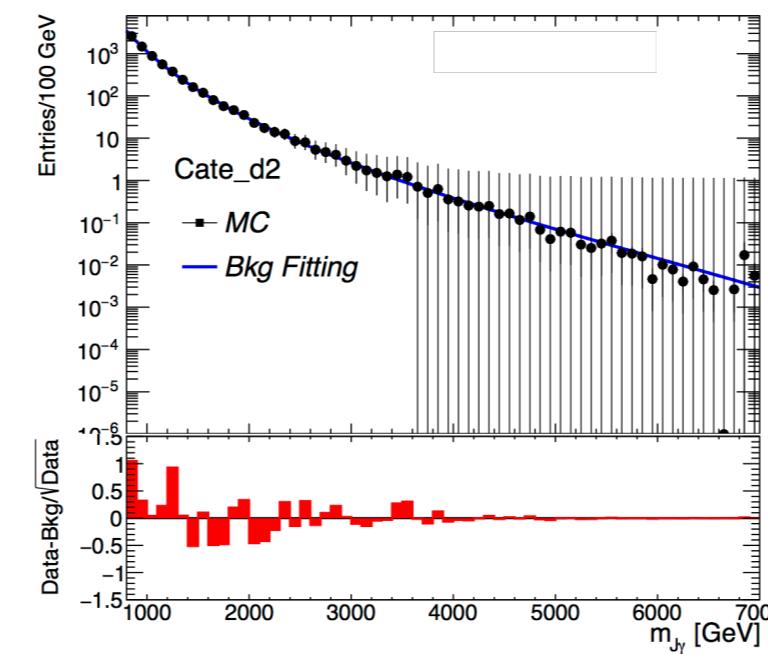
- The background estimation is data-driven, but the form of fit function is determined by spurious signal test on MC
- Test on the MC to determine number of parameters for di-jet function:

$$B(m_{\gamma J}; p_i) = (1 - x)^{p_1} x^{p_2 + p_3 \log(x + p_4 \log^2(x) + p_5 \log^3(x))}$$
- Use the spurious signal test result to estimate background modeling uncertainty
- Cross checked with F-test to validate our function form choice

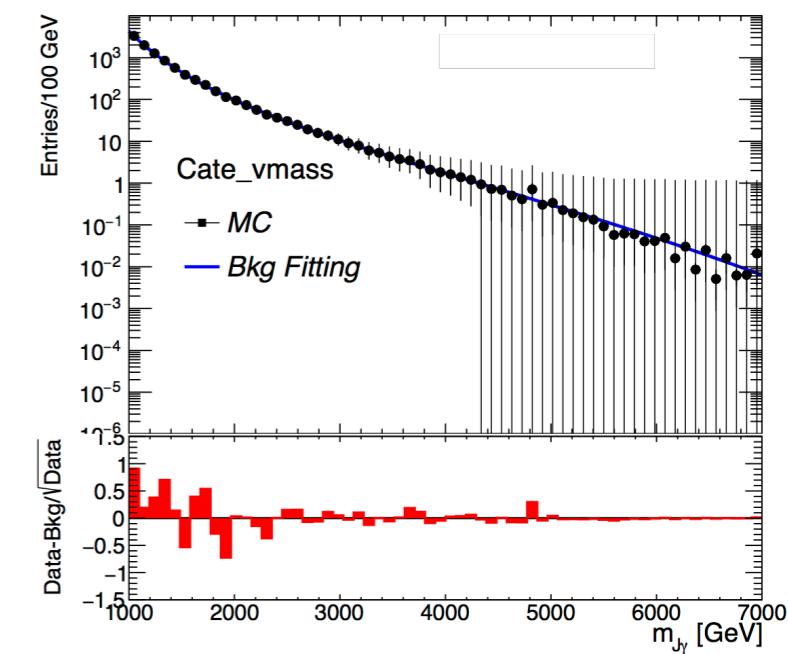
$Z\gamma$ Spin-0 Btag Category



$Z\gamma$ Spin-0 D2 Category



$Z\gamma$ Spin-0 Vmass Category

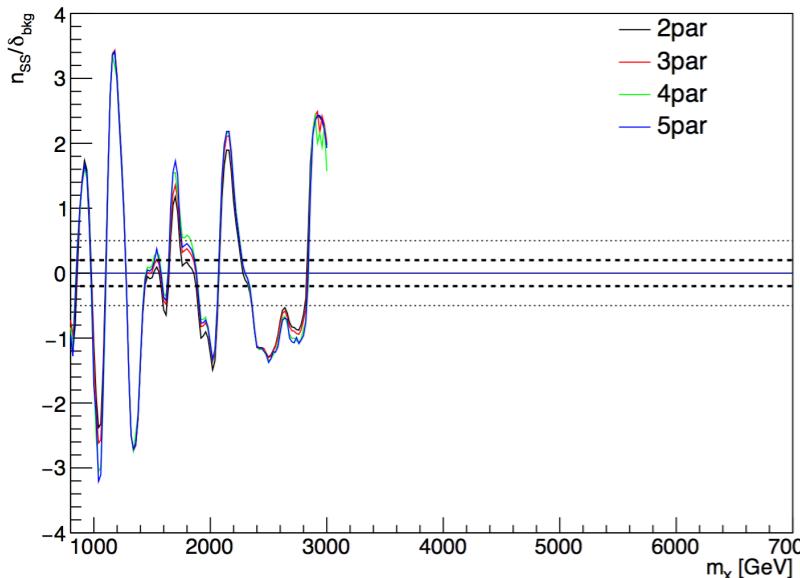




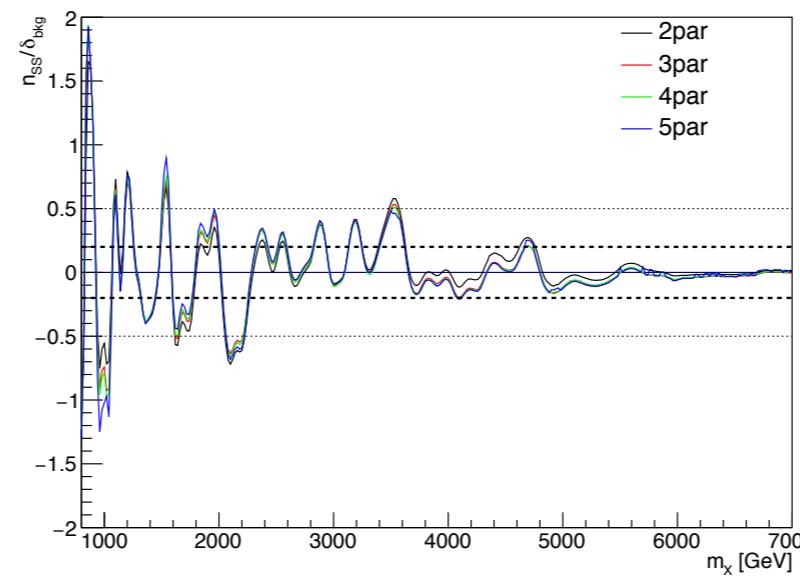
Analysis: Spurious Signal Test

- Fit the distribution with signal+background fit, the number of signal fitted is N_{SS}
- Estimate the statistical uncertainty assuming the distribution is background only, which is called δ_B . If the background is large enough, it will be about $\sqrt{N_B}$
- The ratio N_{SS}/δ_B is taken as the criteria of goodness of fit.
- If no significant improvement, use the lowest number of parameters

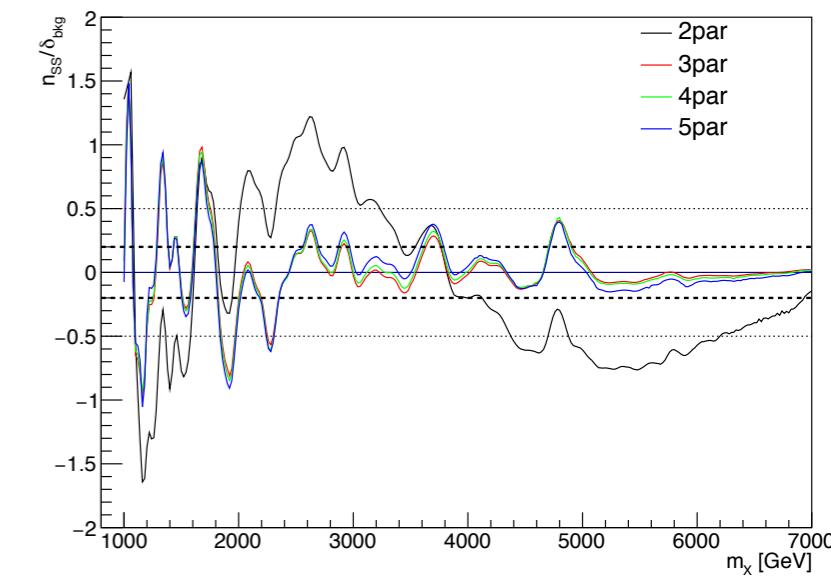
$Z\gamma$ Spin-0 Btag Category



$Z\gamma$ Spin-0 D2 Category



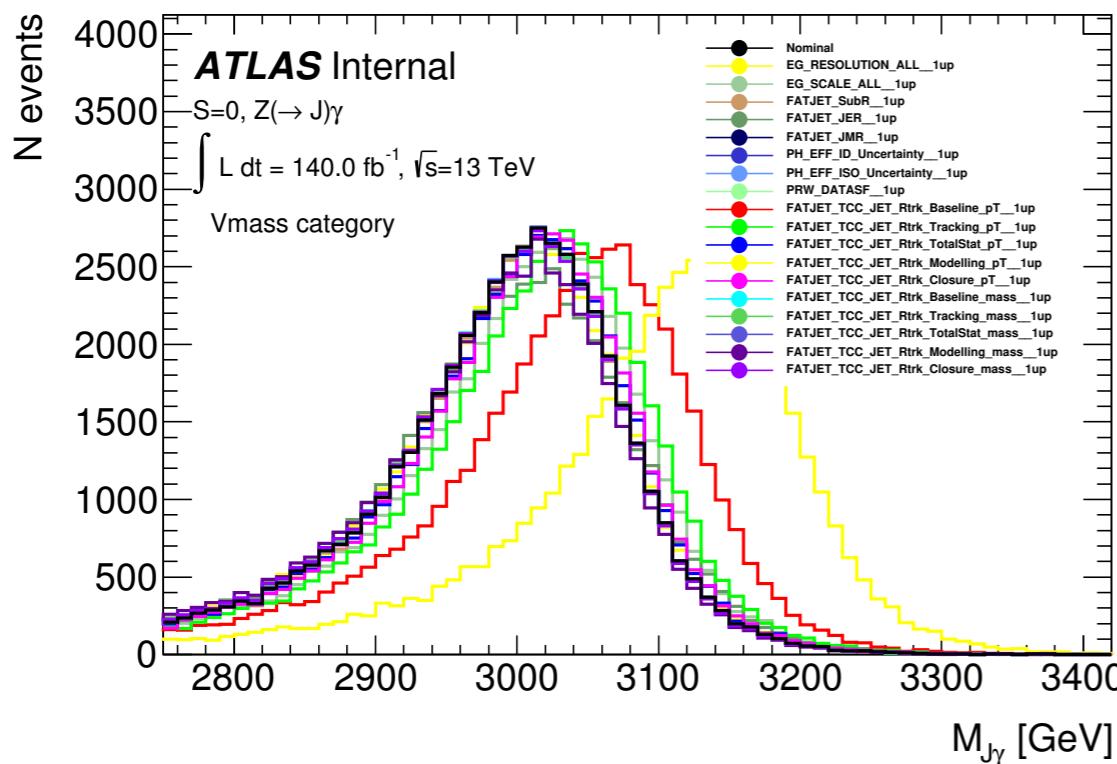
$Z\gamma$ Spin-0 Vmass Category



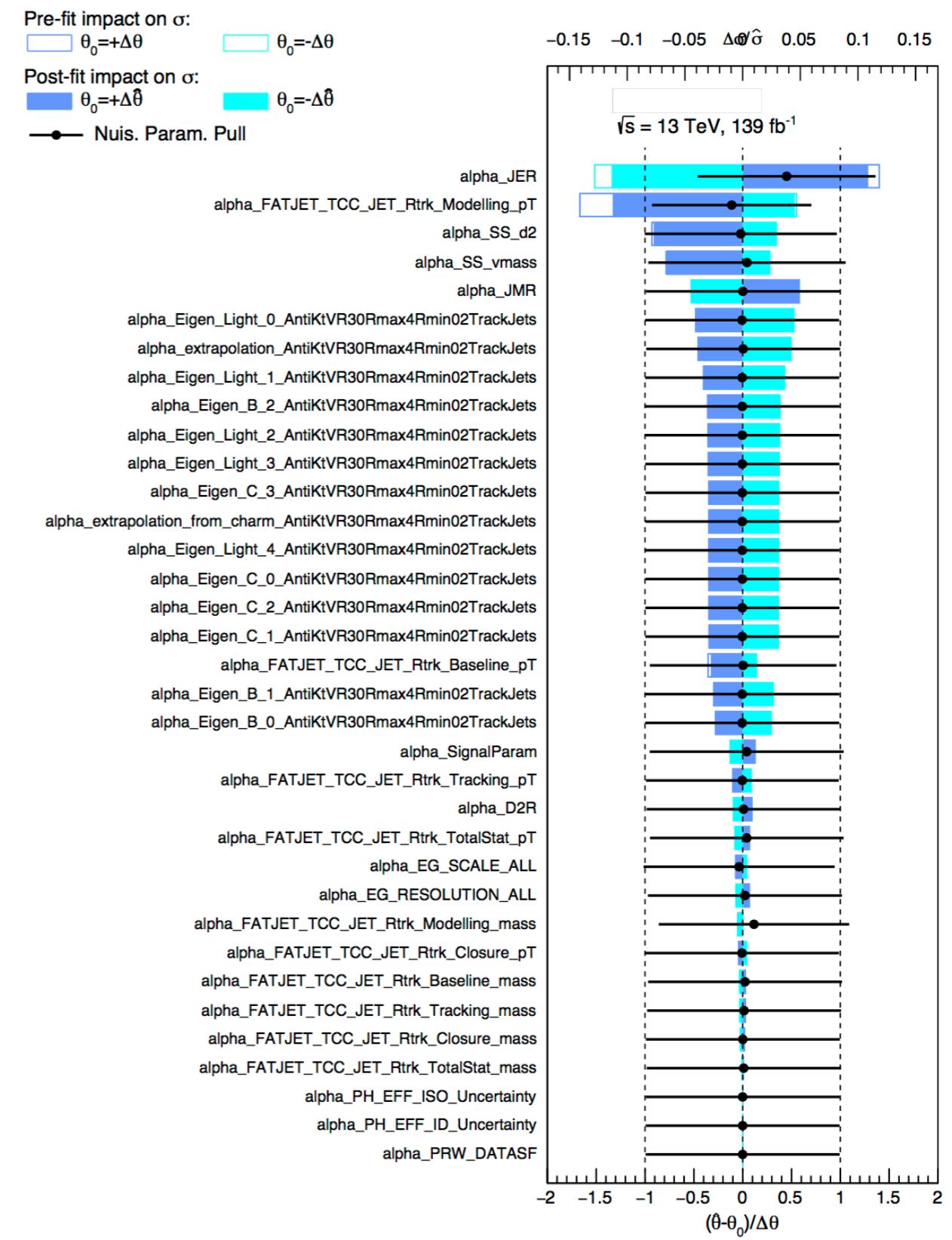
Parameter Number Choice

Channel	btag	d2	vmass
$gg \rightarrow X(spin = 0) \rightarrow \gamma Z(\rightarrow q\bar{q})$	2	2	3
$qq \rightarrow X^\pm(spin = 1) \rightarrow \gamma W^\pm(\rightarrow qq')$	-	2	3
$qq \rightarrow X(spin = 2) \rightarrow \gamma Z(\rightarrow q\bar{q})$	2	3	3
$gg \rightarrow X(spin = 2) \rightarrow \gamma Z(\rightarrow q\bar{q})$	2	2	3

Analysis: Systematical Uncertainties

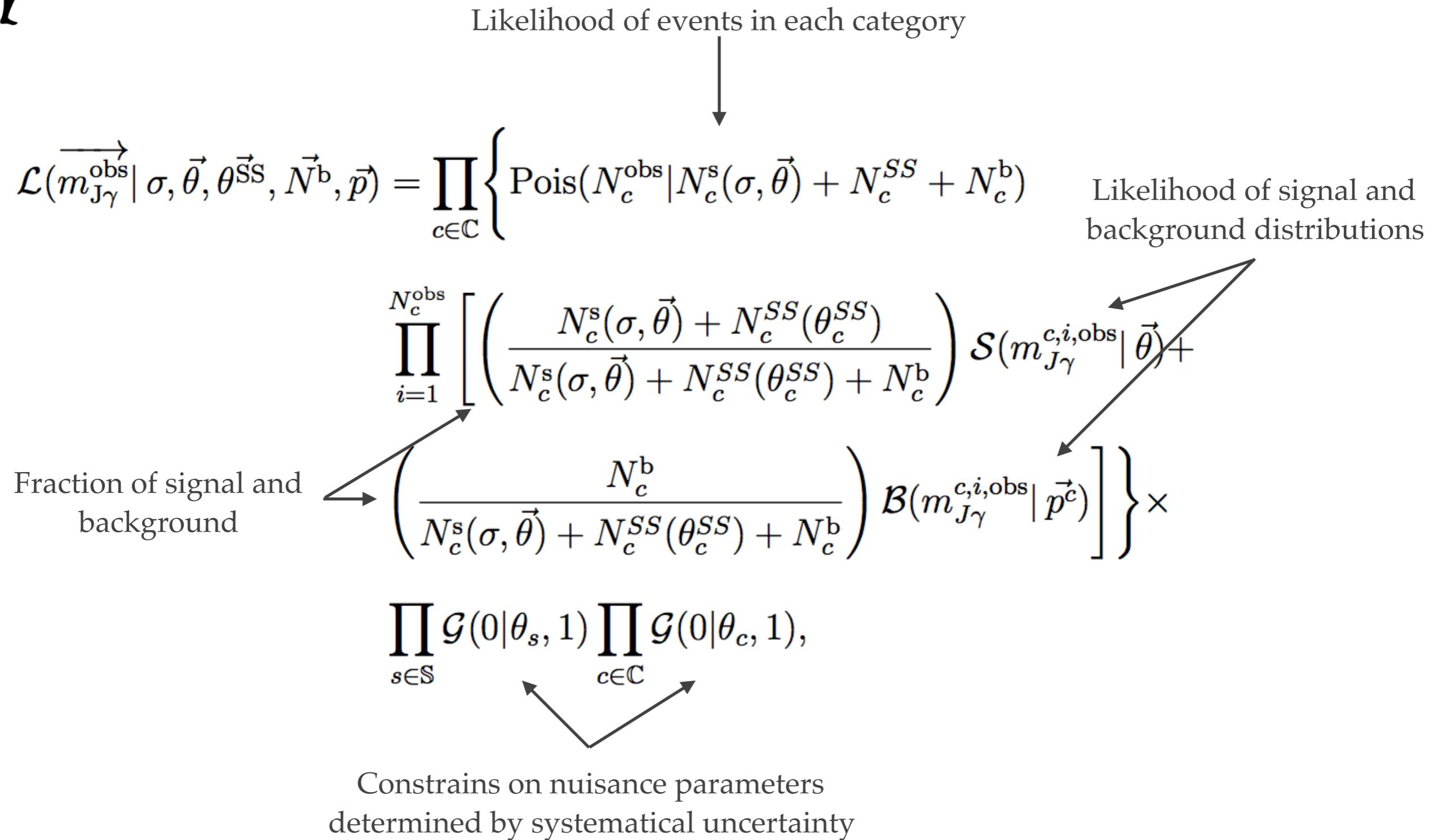


- Luminosity ($\sim 1.7\%$)
- Photon energy**
 - Dominant at very high mass
- Photon identification and isolation
 - Negligible
- Jet energy and mass**
 - Major contribution in full range
- Jet tagging
 - Negligible





Analysis: Statistical Implementation

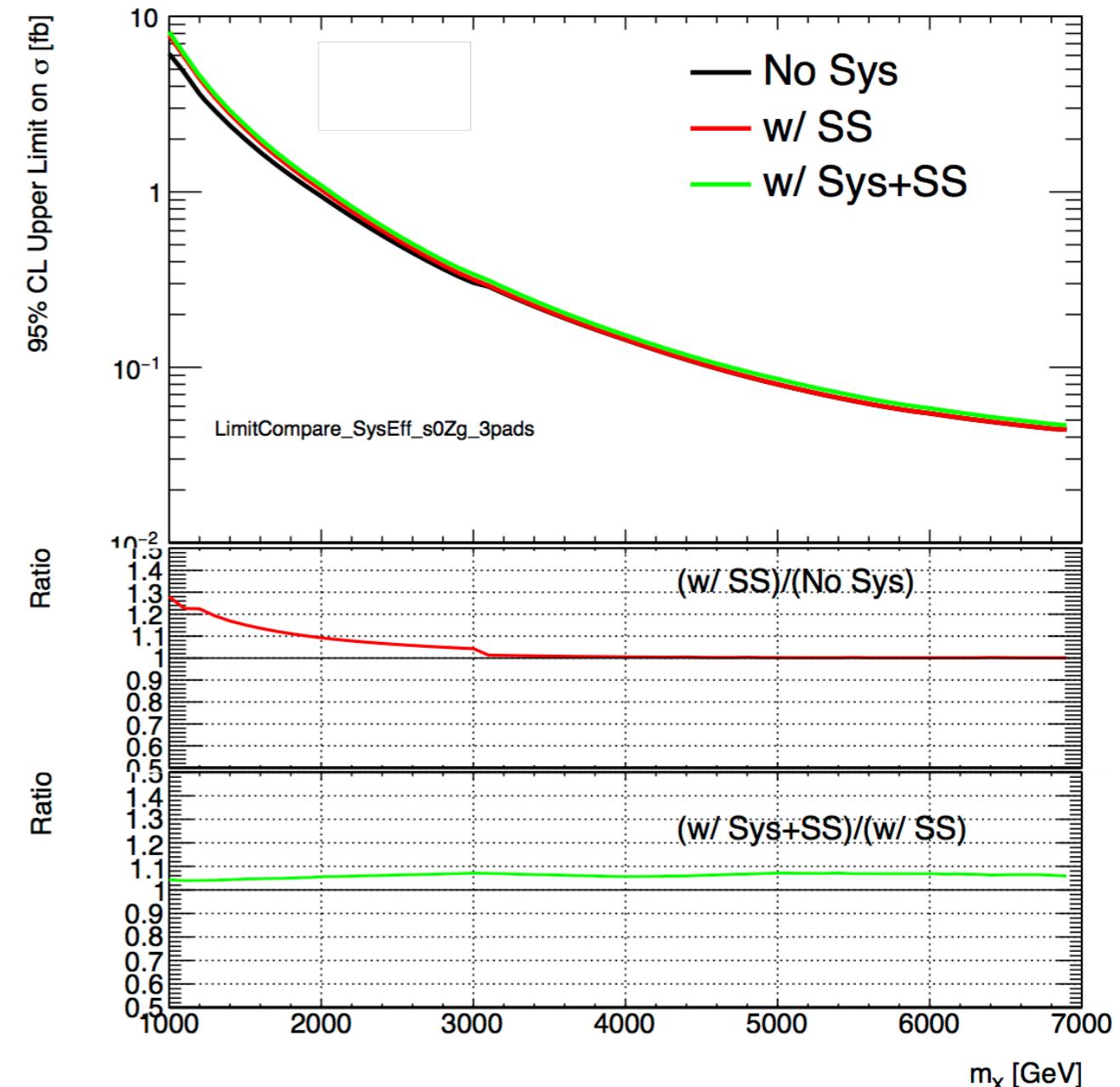




Analysis: Expected Limits

- Spurious signal contribution is only in low mass range
- Systematical uncertainty is below 10% in the full range
- The major limitation of this search is from the statistical uncertainty
- Similar results for other signal channels

$$gg \rightarrow X \rightarrow \gamma Z (\rightarrow q\bar{q})$$



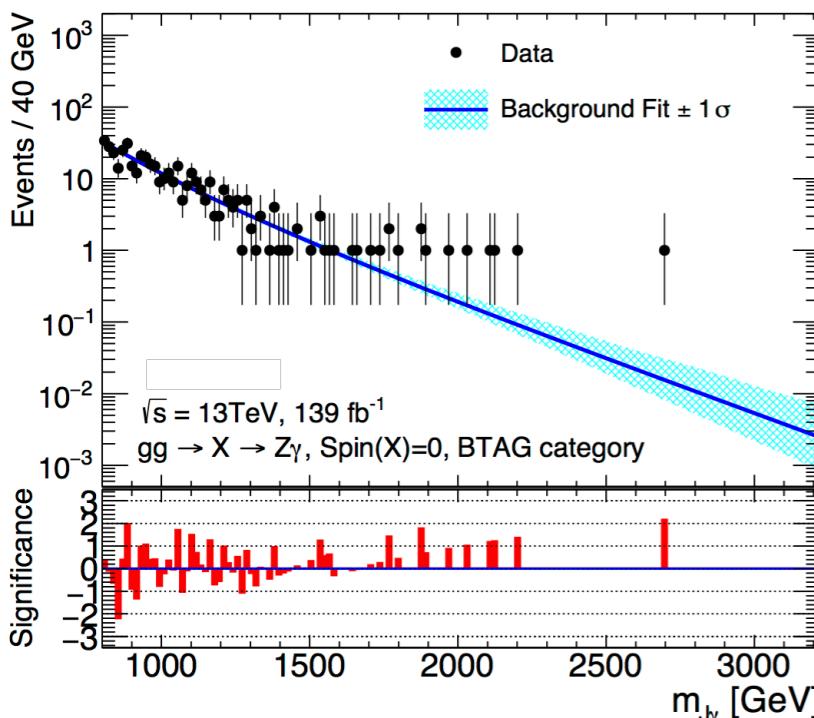


Analysis: Fit to the Data

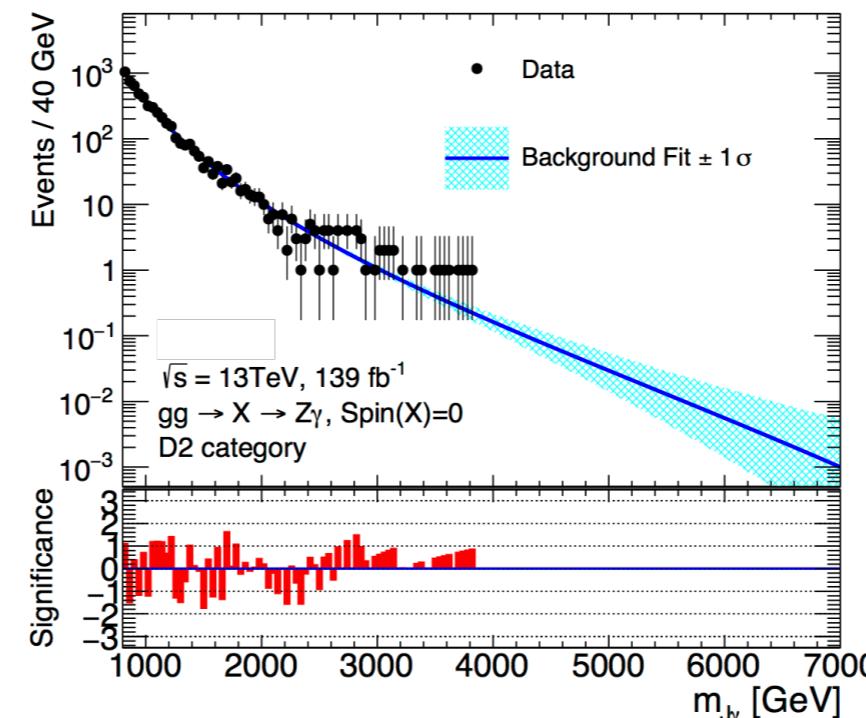
- After the search strategy completely defined, we are permitted by ATLAS to use the collected Run-2 data of 139 fb^{-1}
- The fitting function is:

$$B(m_{\gamma J}; p_i) = (1 - x)^{p_1} x^{p_2 + p_3 \log(x + p_4 \log^2(x) + p_5 \log^3(x))}$$

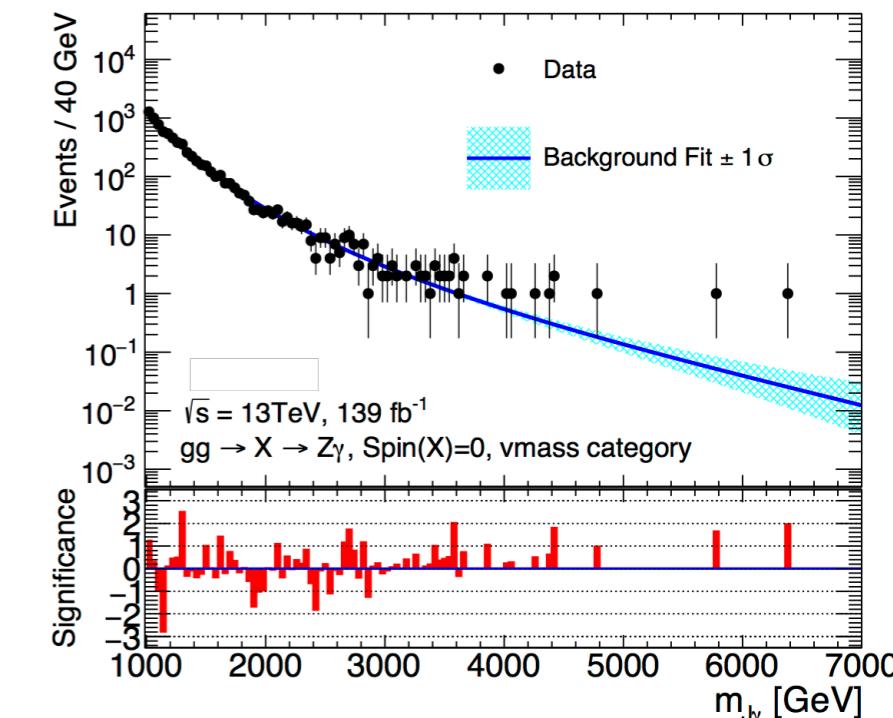
$Z\gamma$ Spin-0 Btag Category



$Z\gamma$ Spin-0 D2 Category

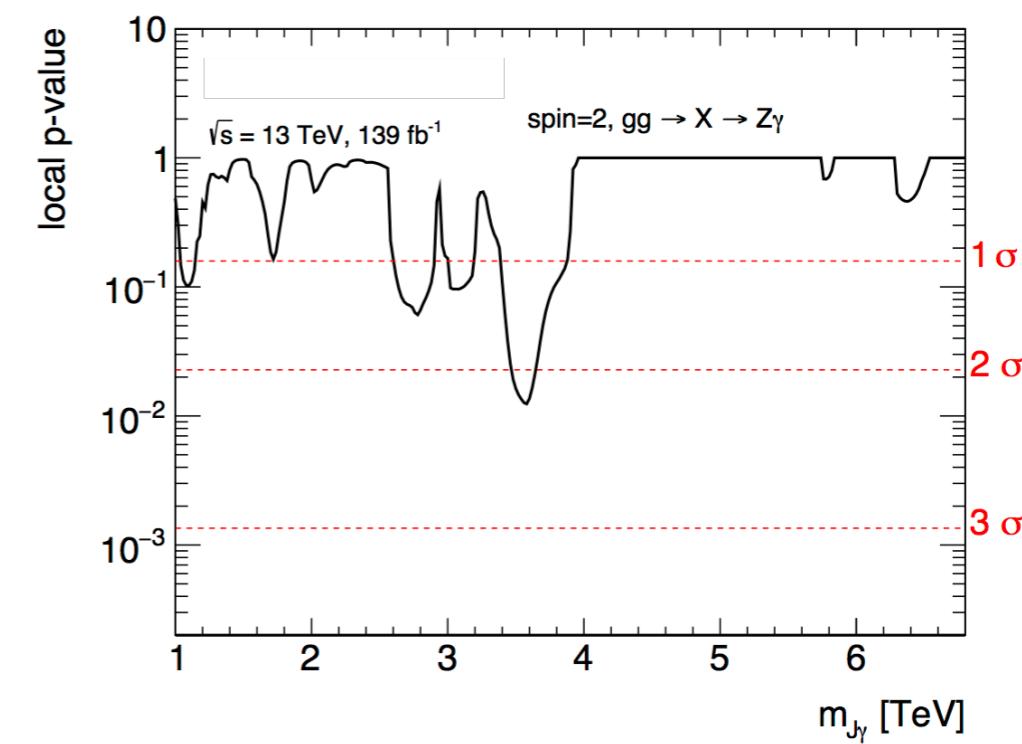
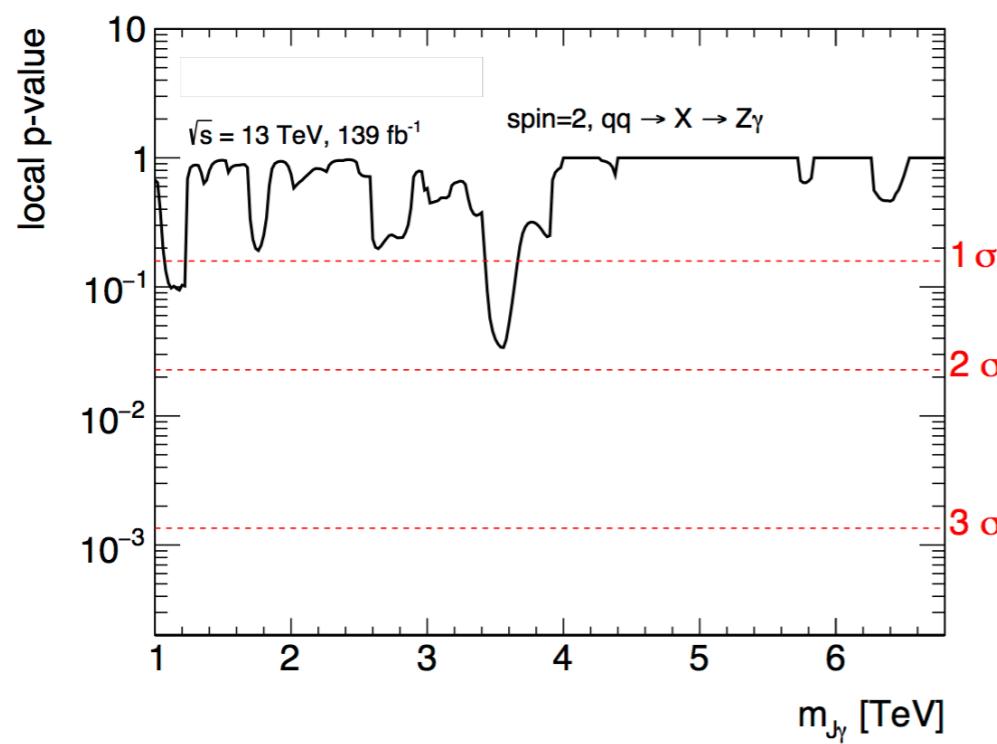
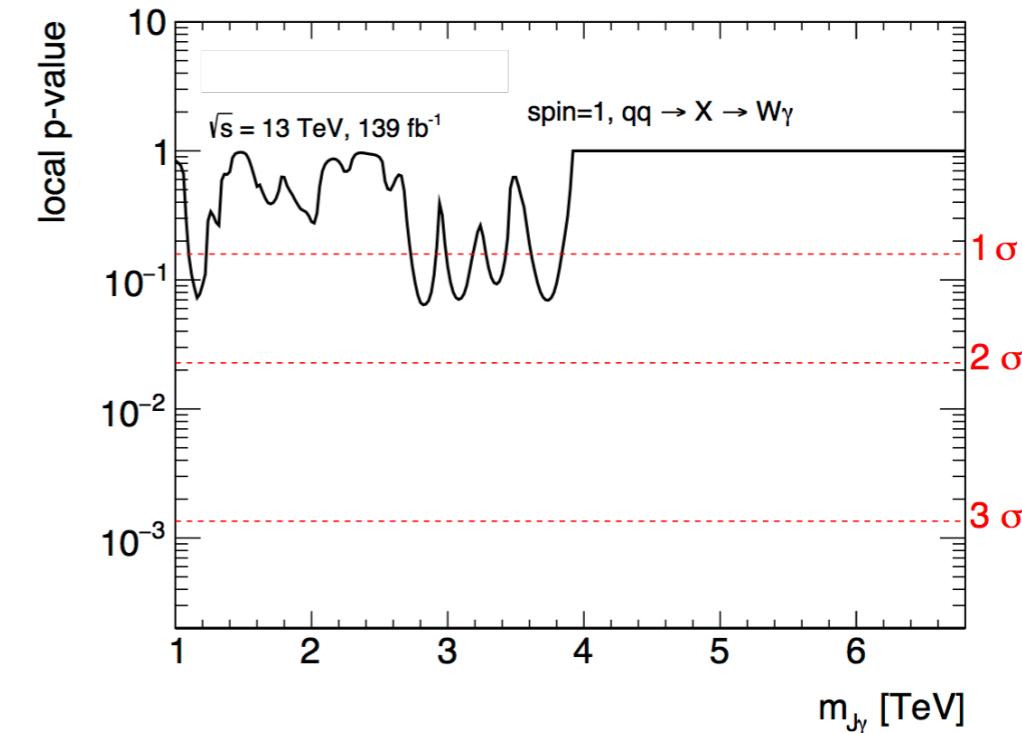
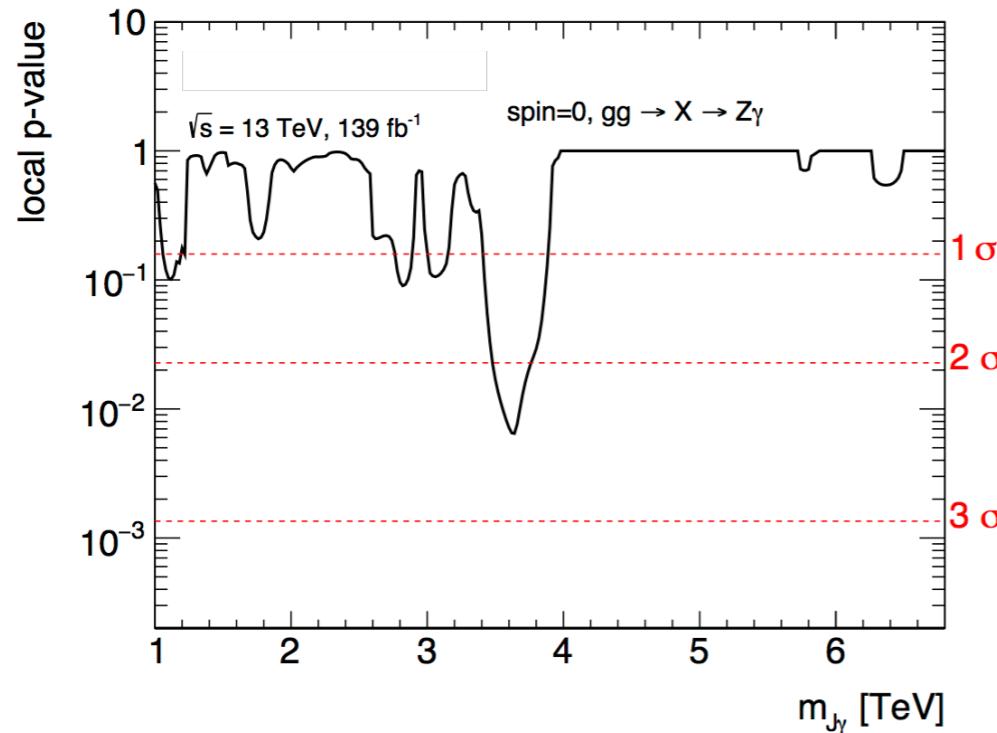


$Z\gamma$ Spin-0 Vmass Category



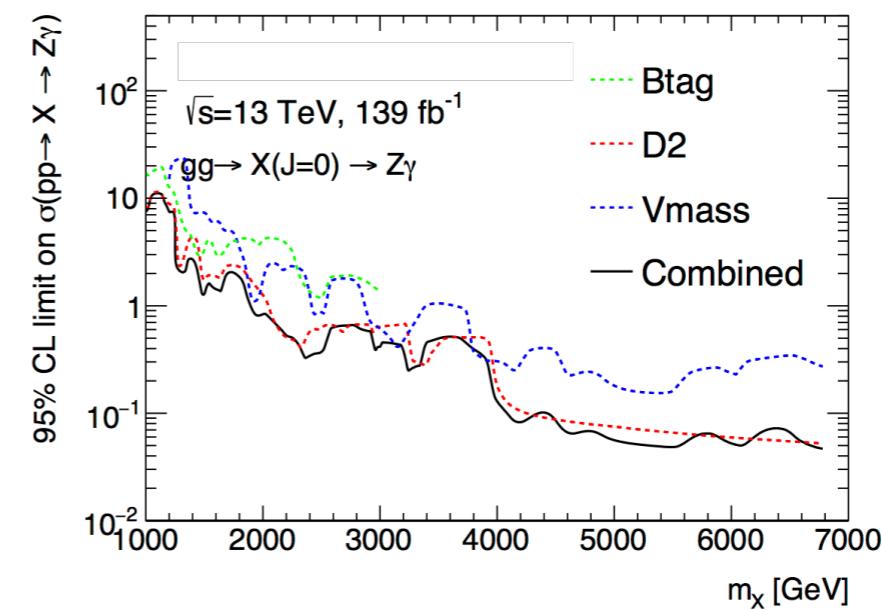
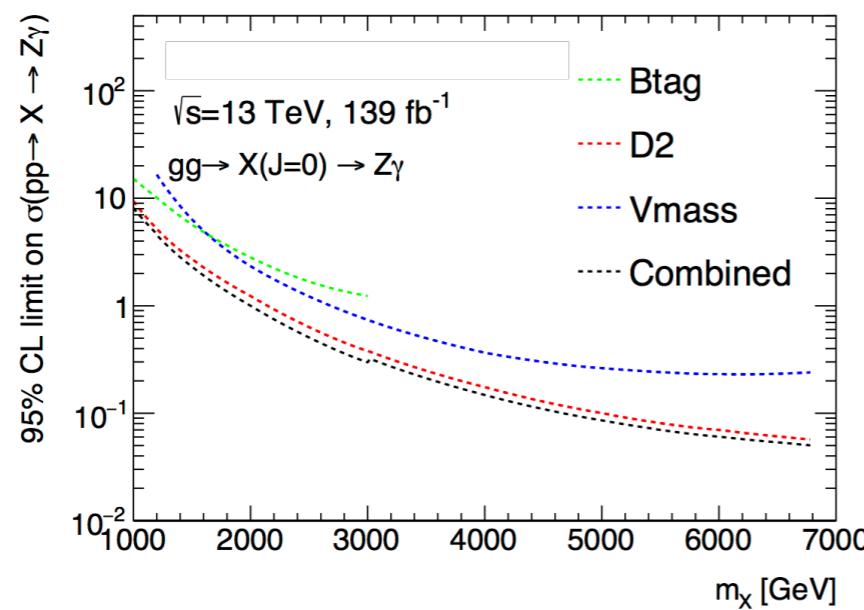
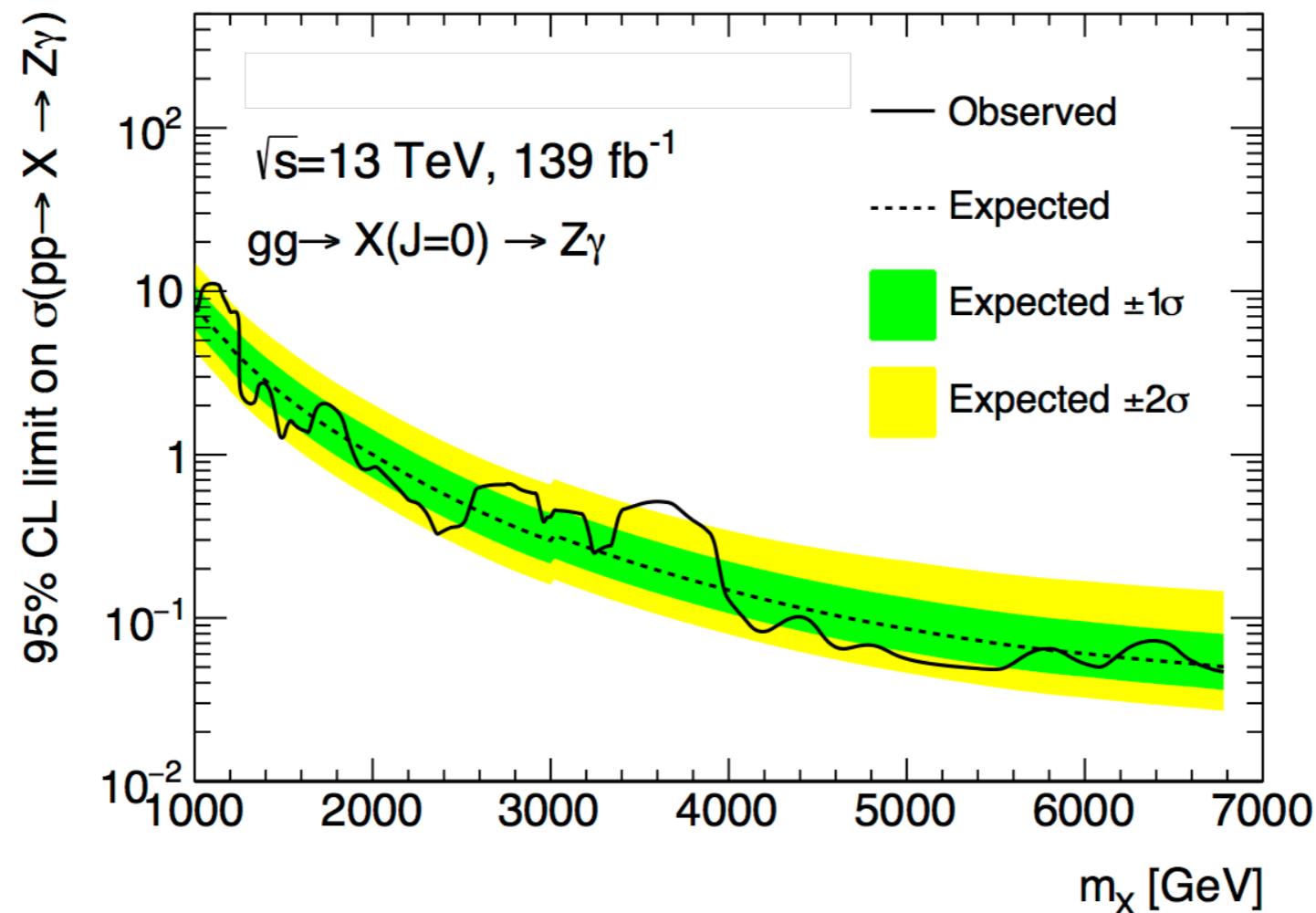


Analysis: Observed Significance



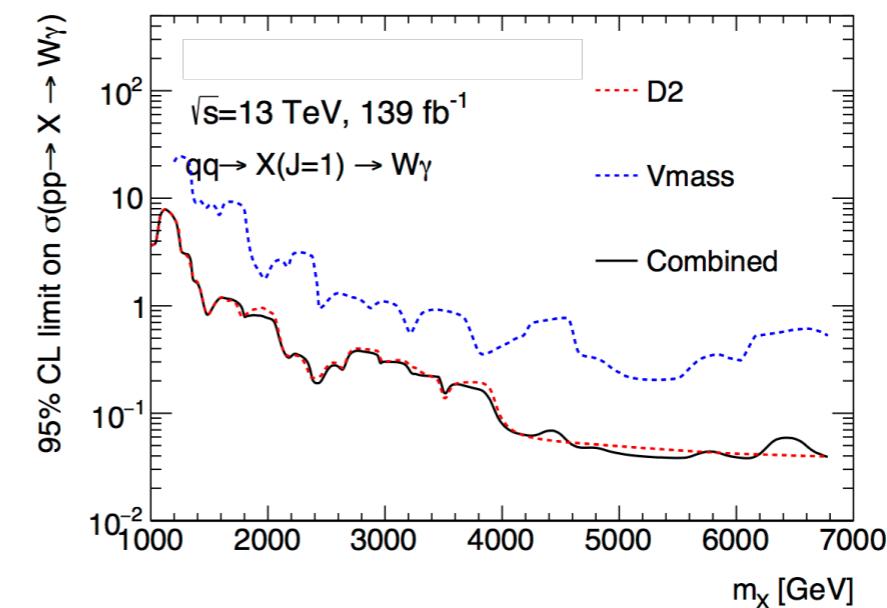
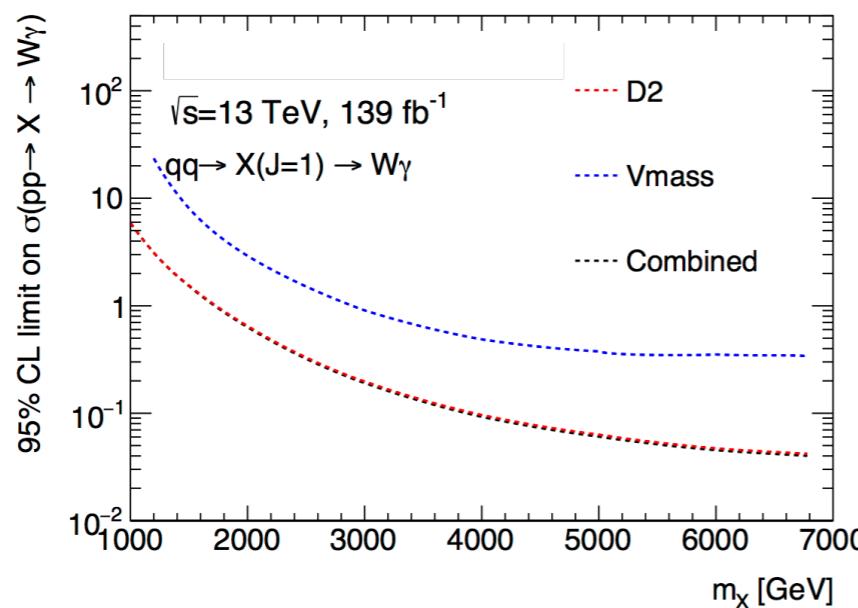
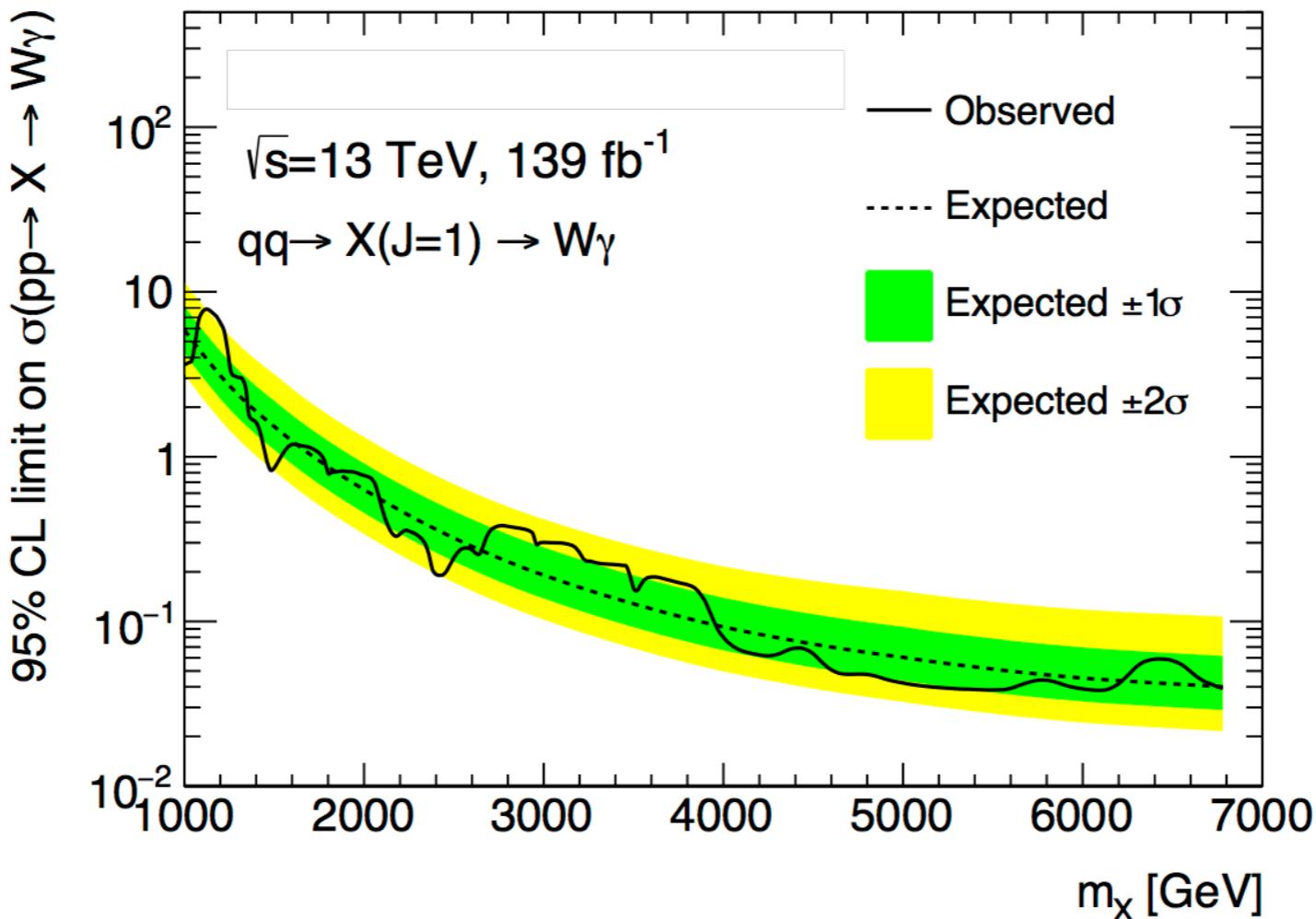


Analysis: Observed Limits



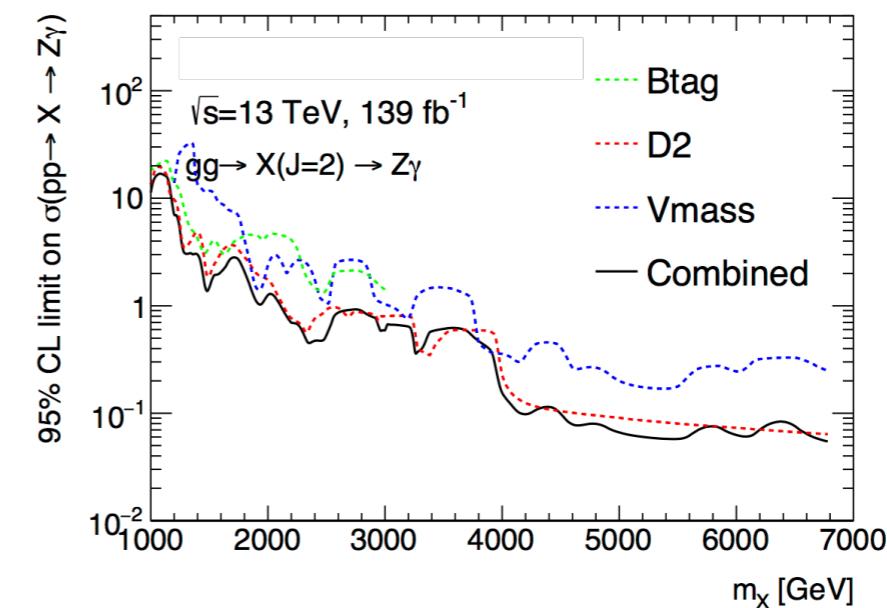
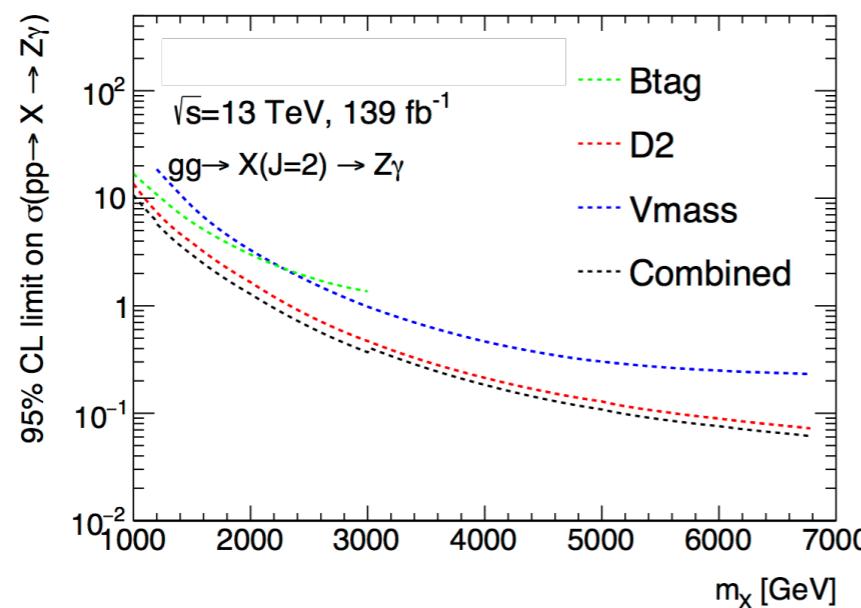
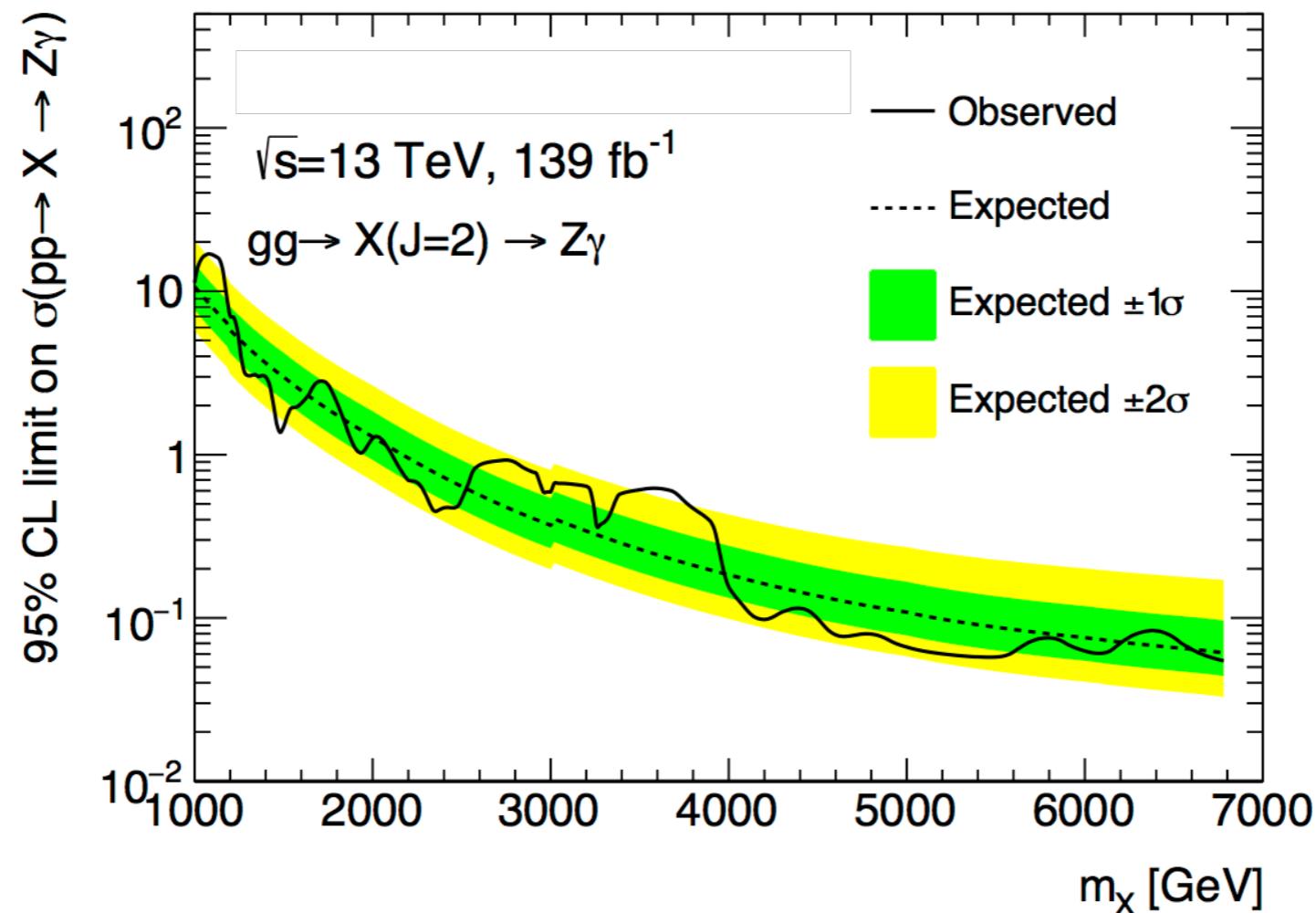


Analysis: Observed Limits



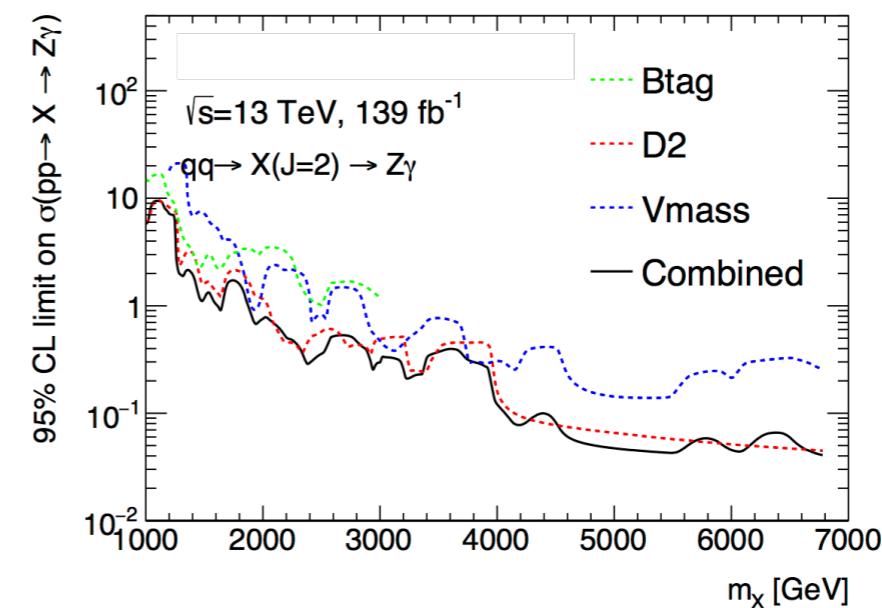
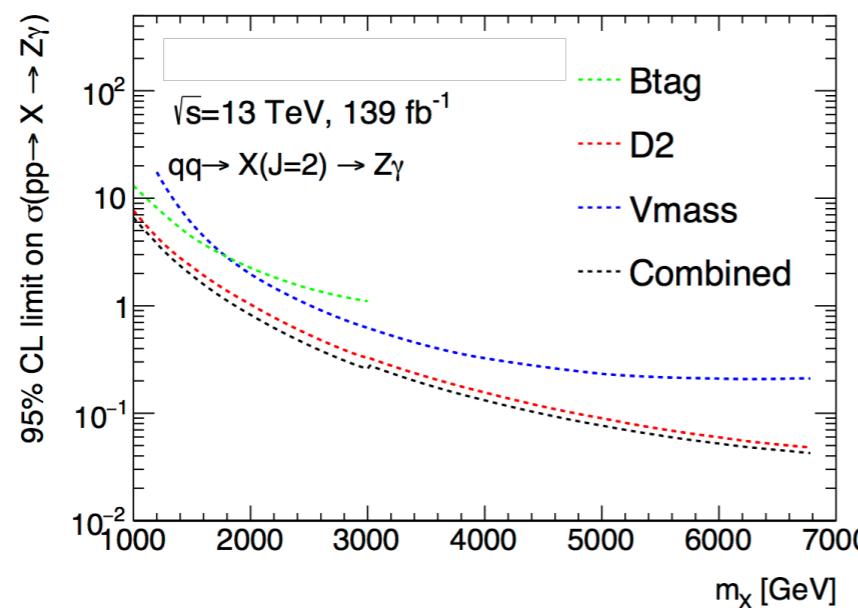
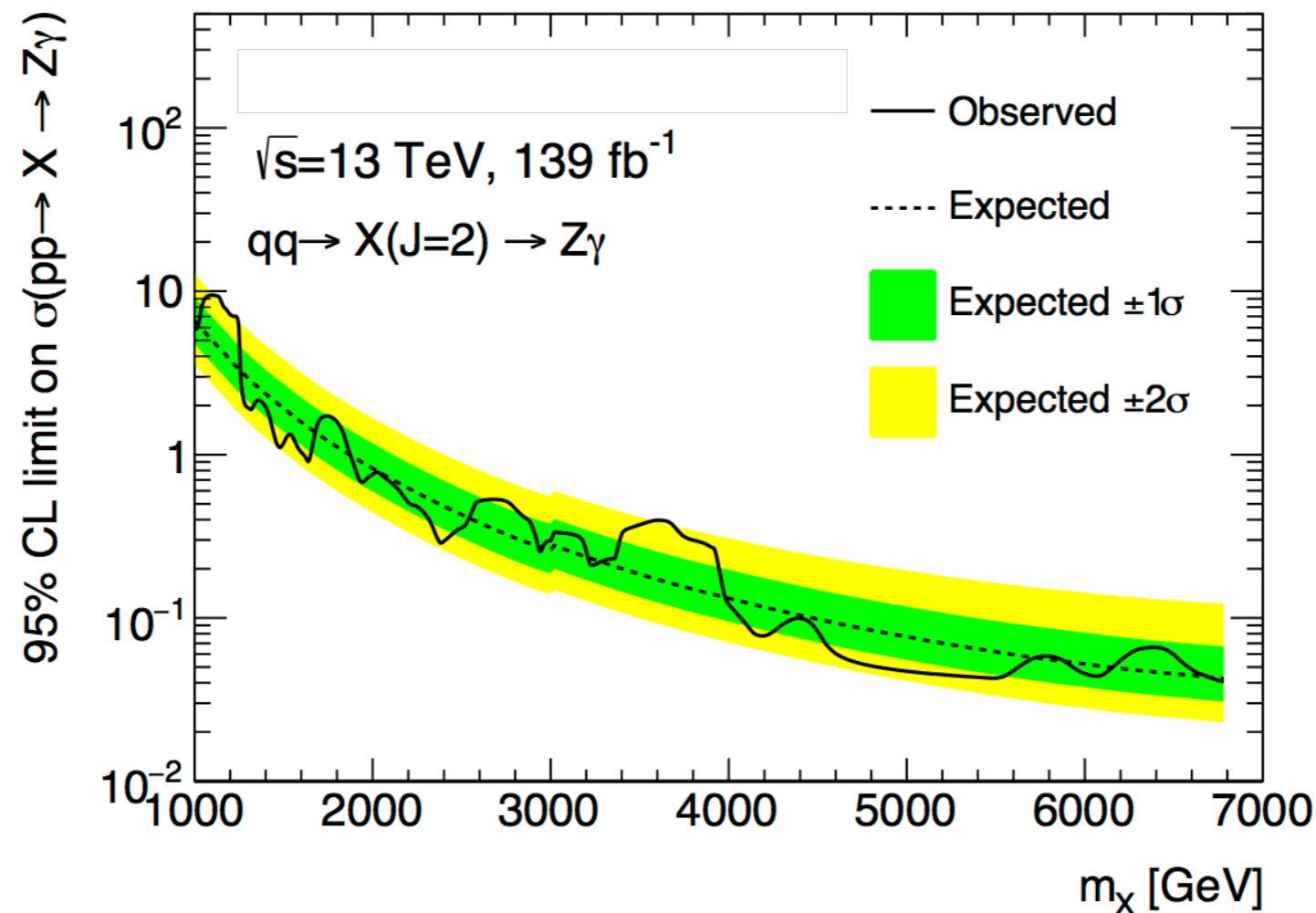


Analysis: Observed Limits





Analysis: Observed Limits

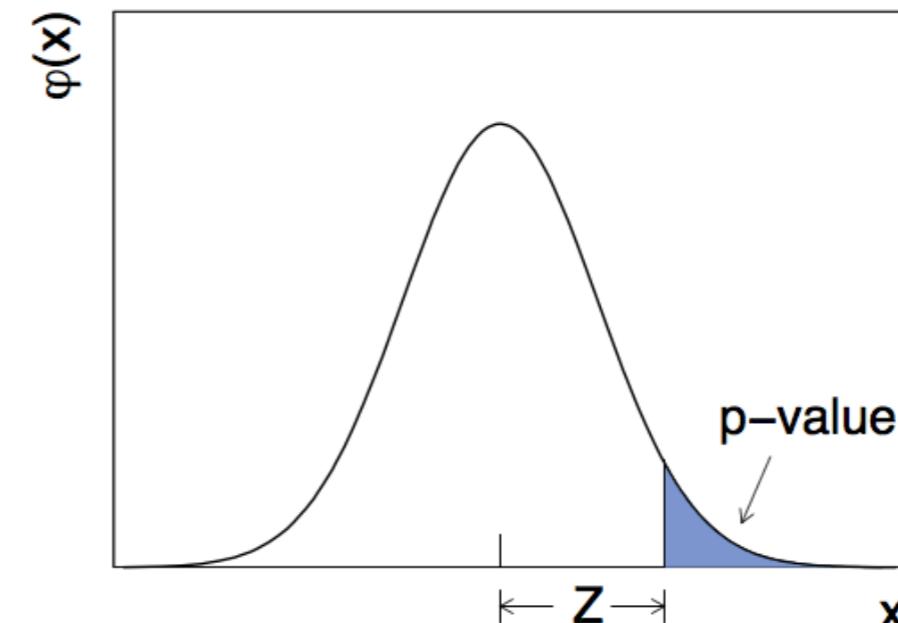
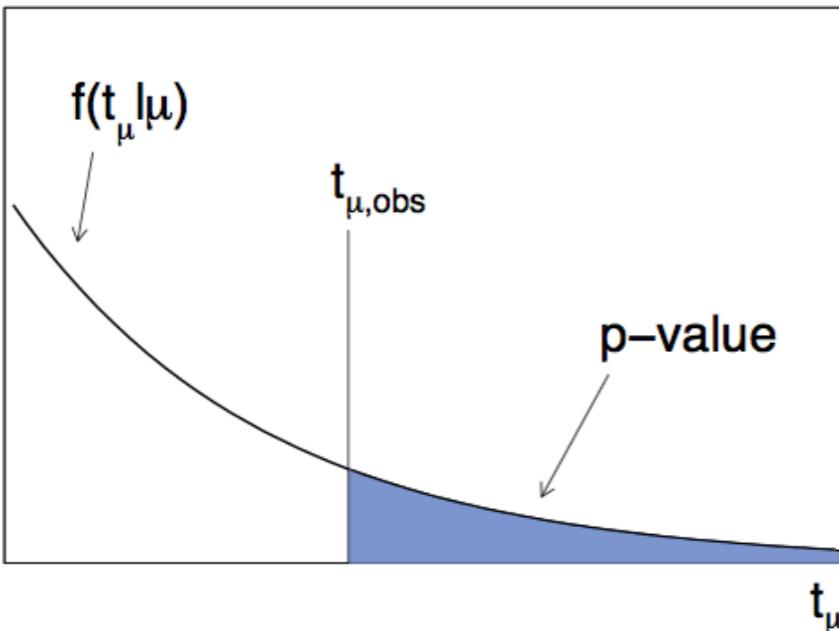


Backup





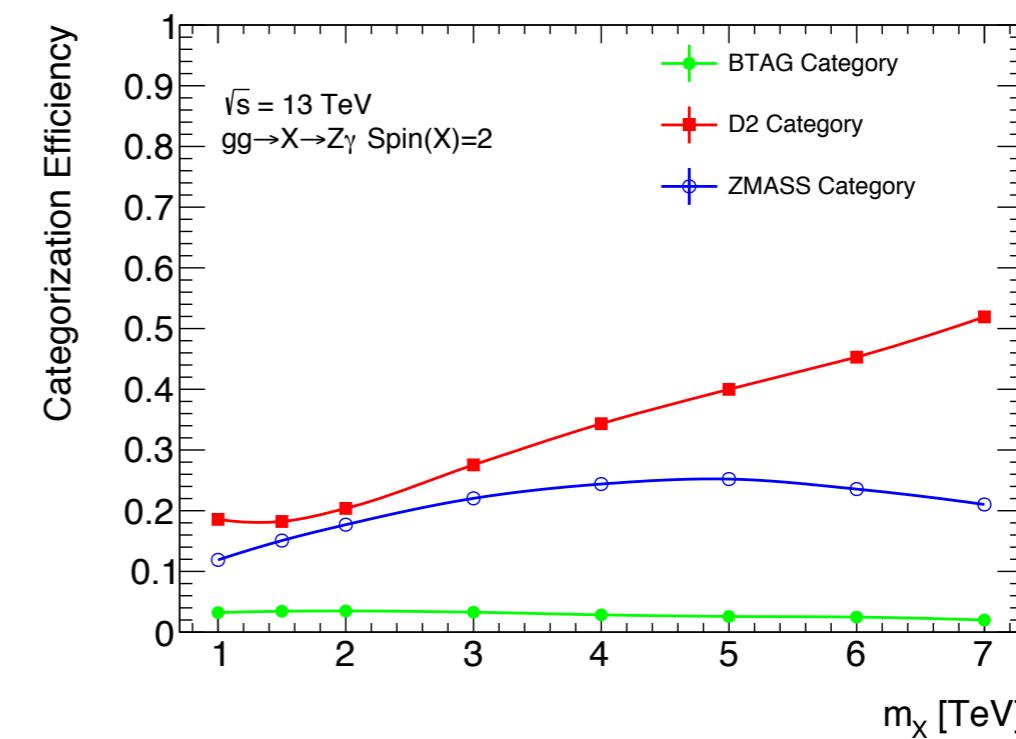
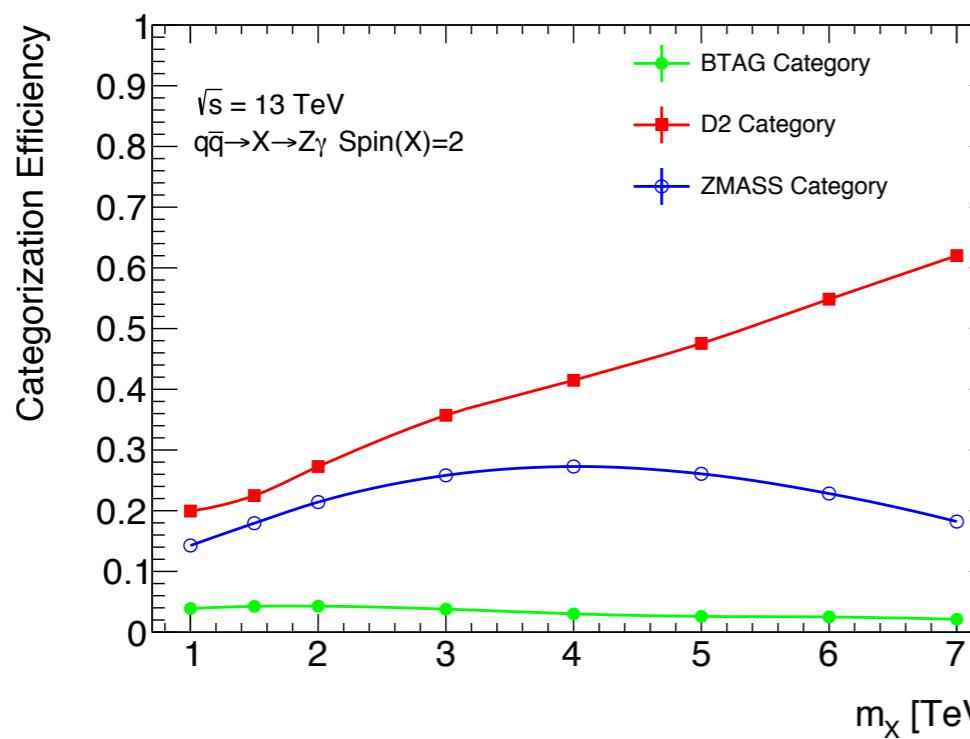
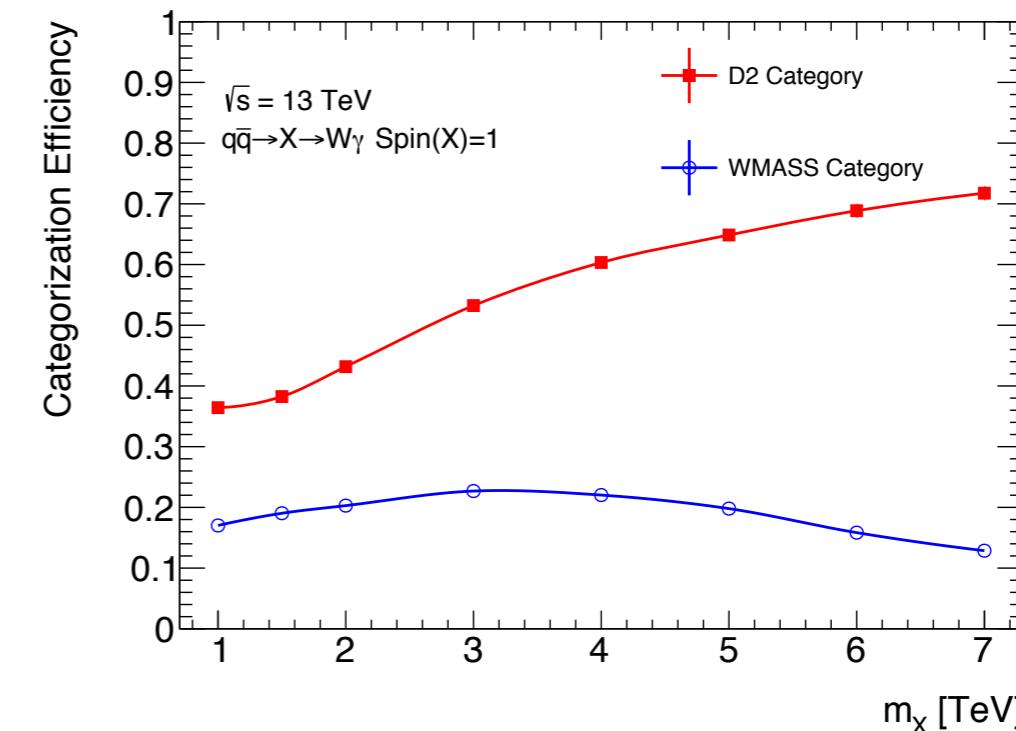
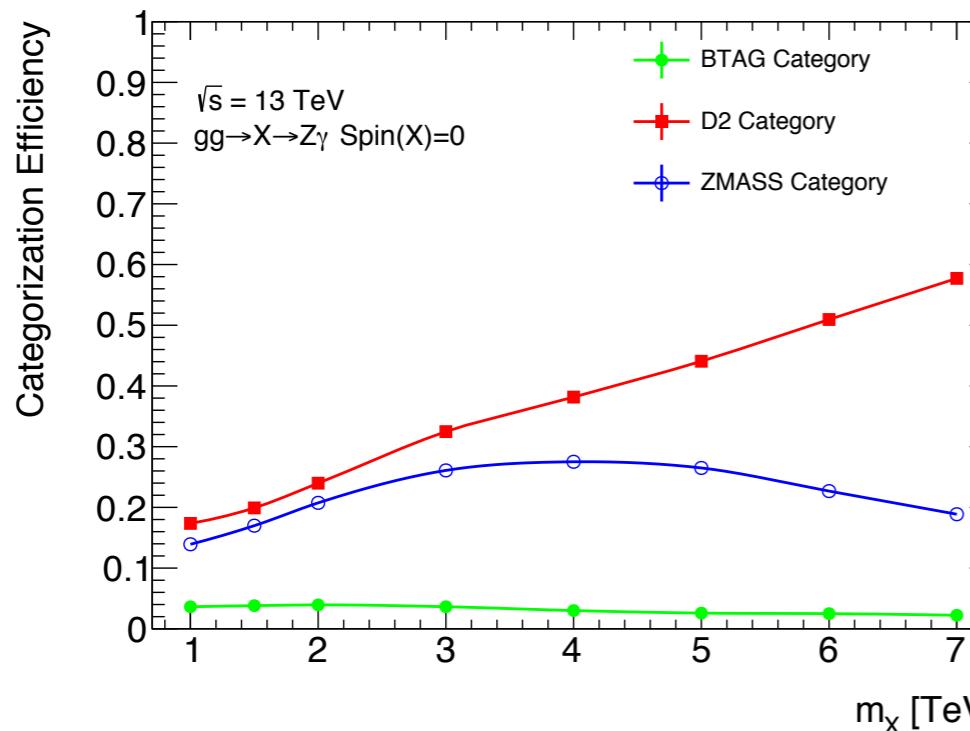
Significance and P-value



$$Z = \Phi^{-1}(1 - p)$$

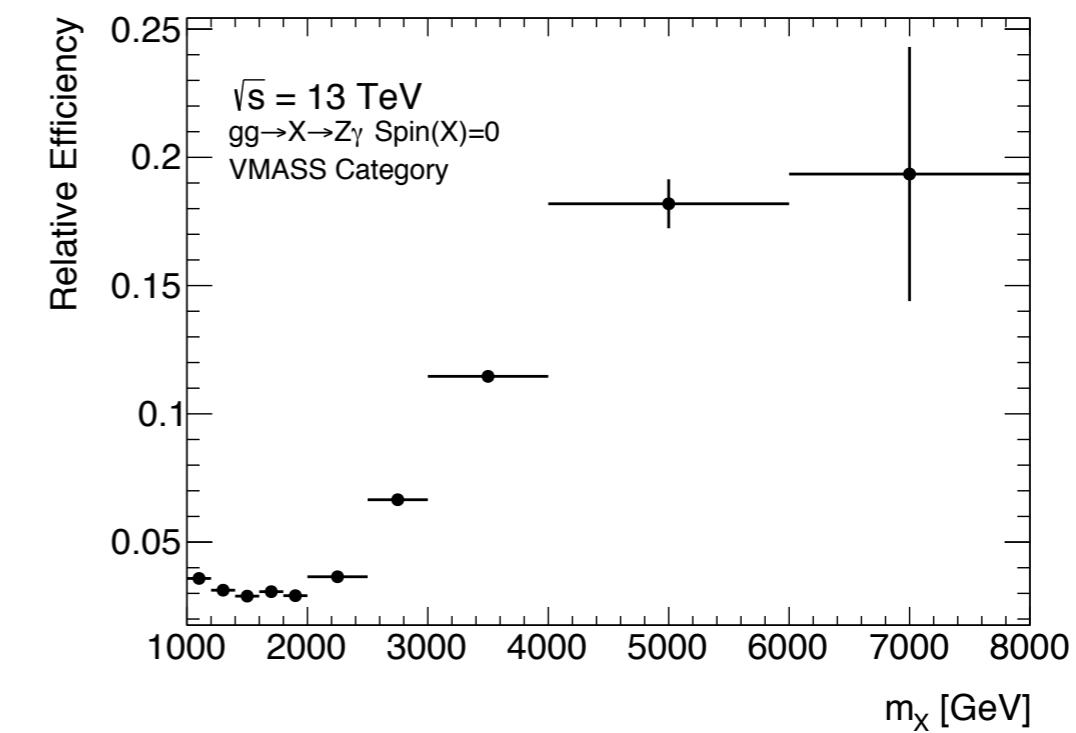
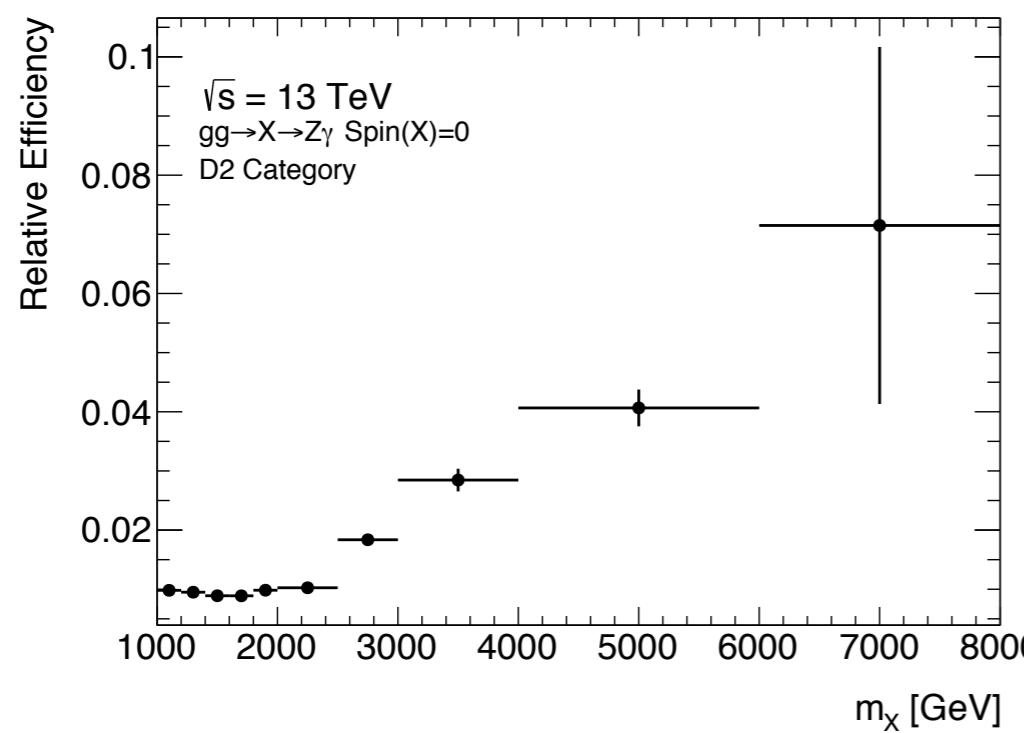
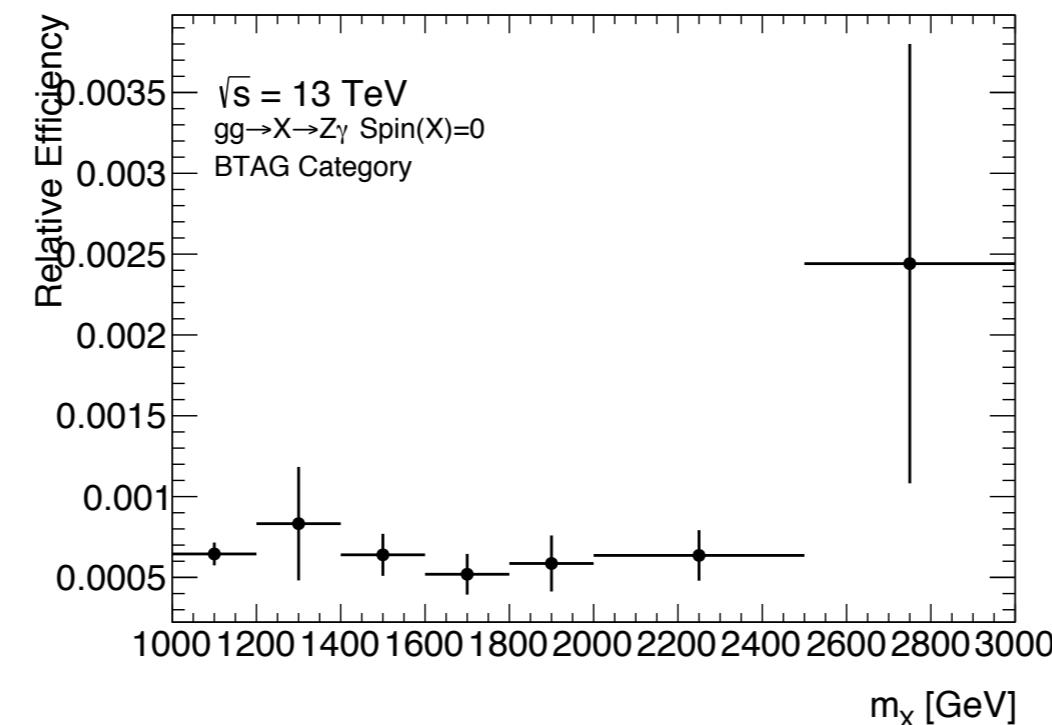


Signal Categorization Efficiencies



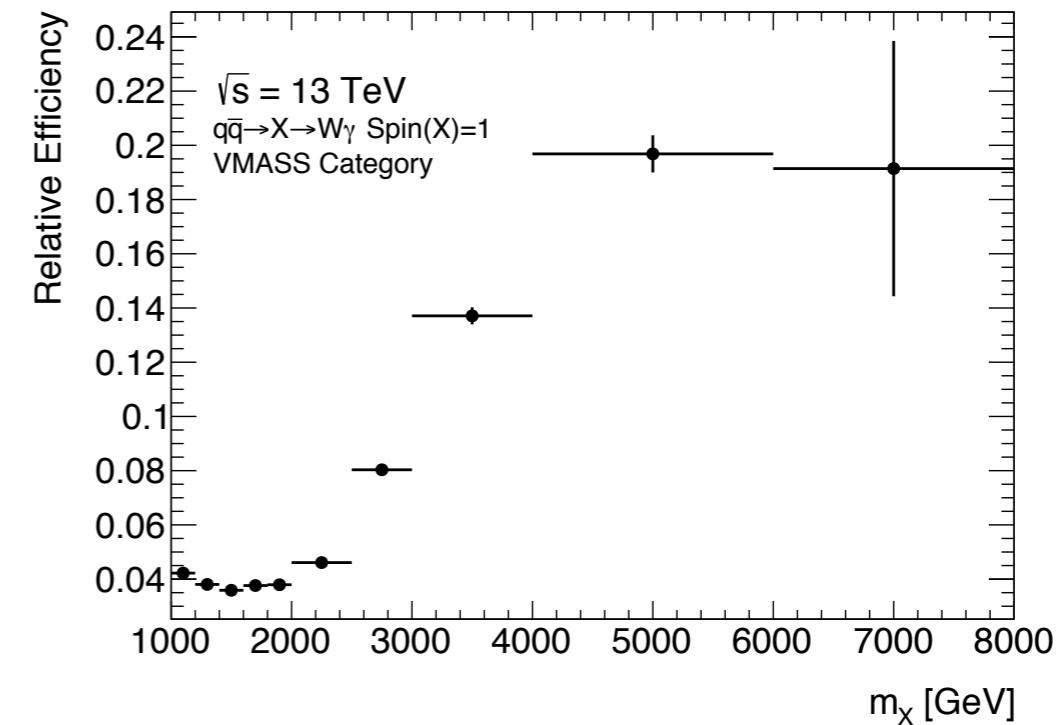
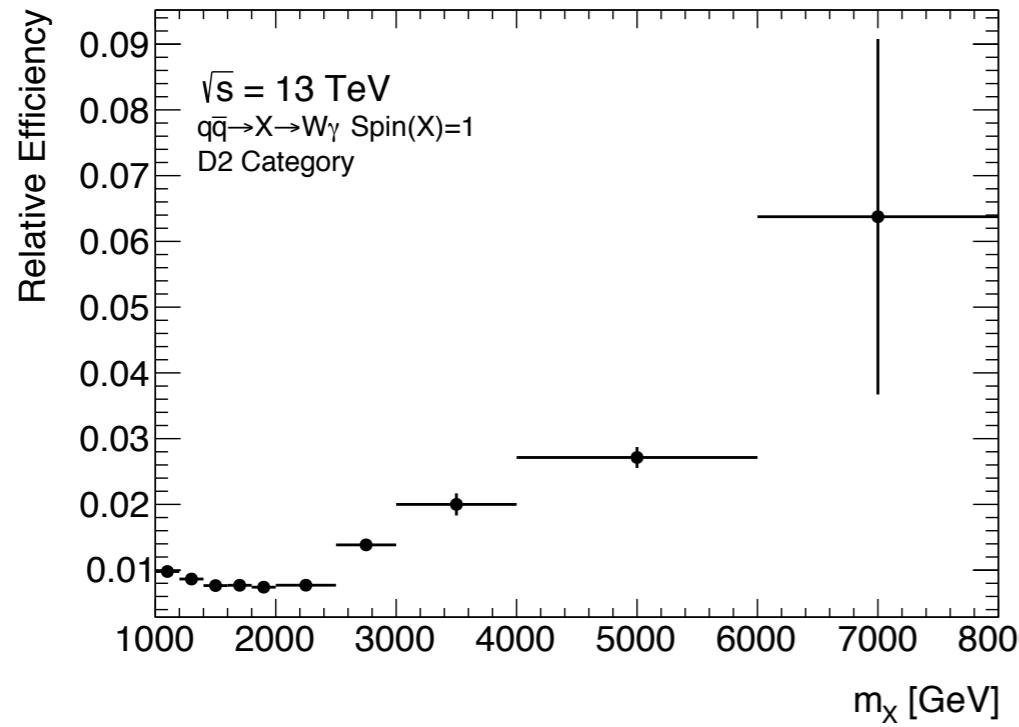


Background Categorization Efficiencies





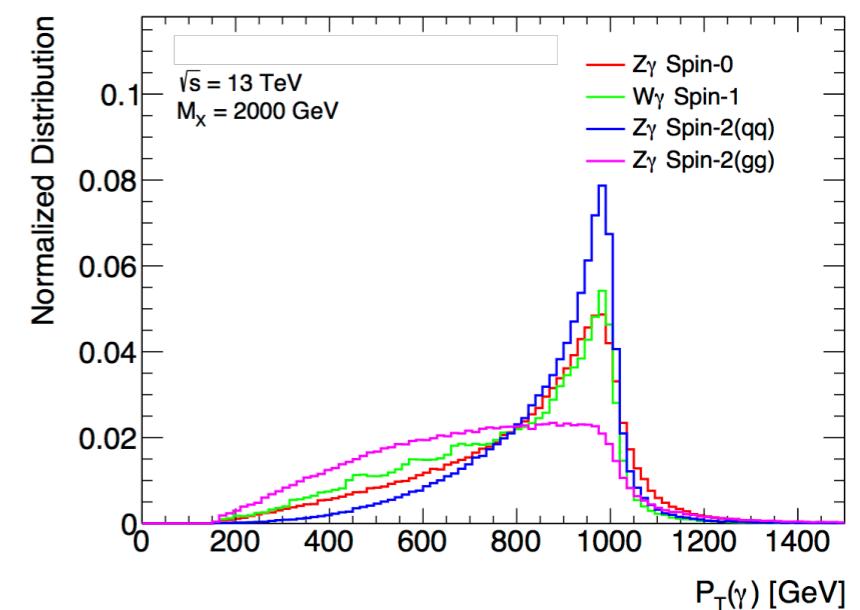
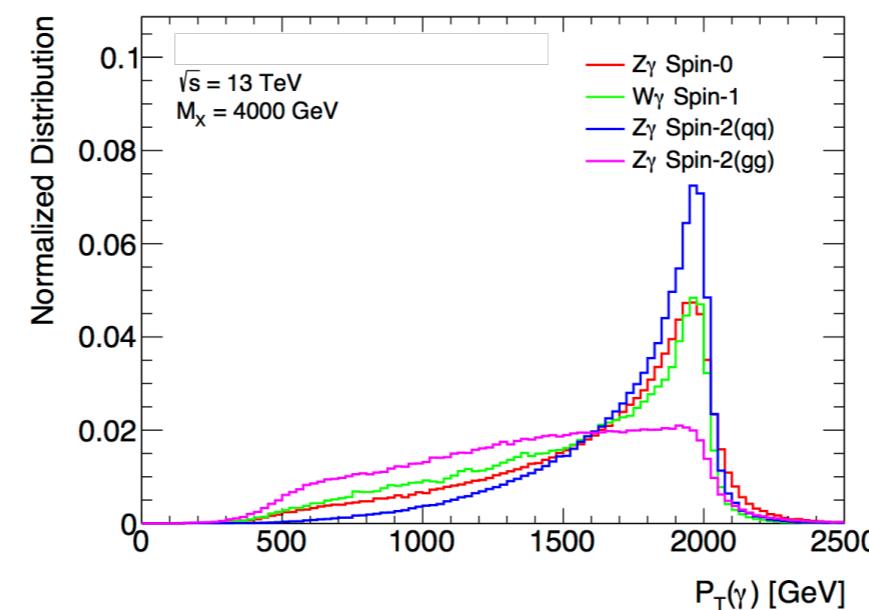
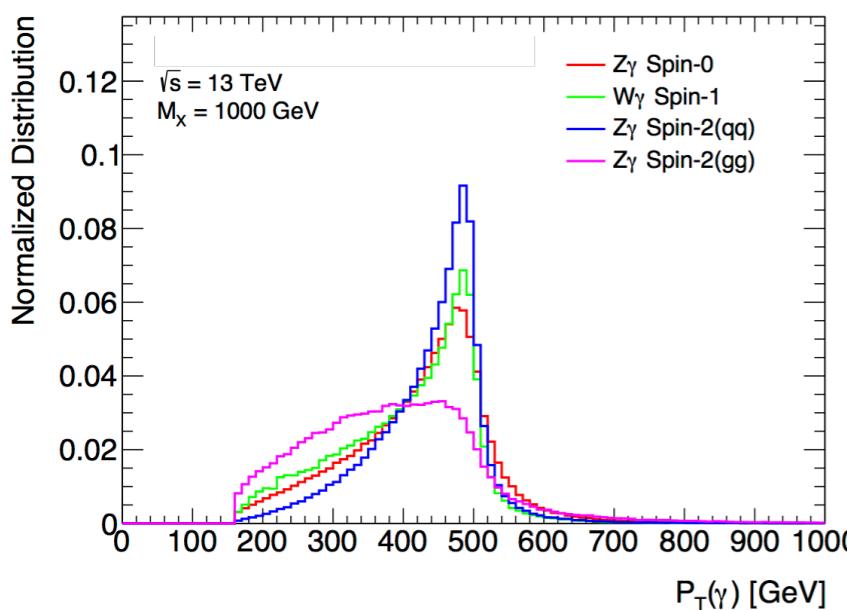
Background Categorization Efficiencies





Signal Comparison

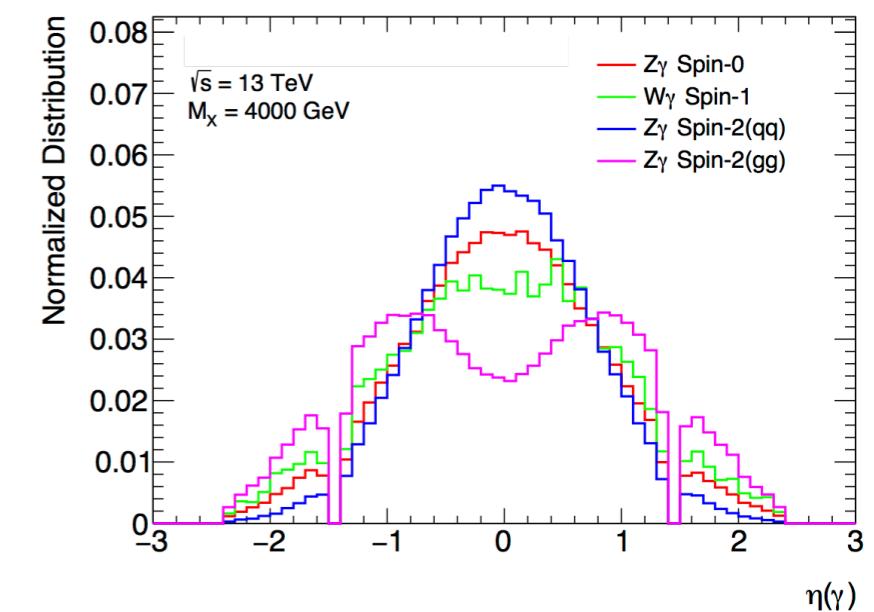
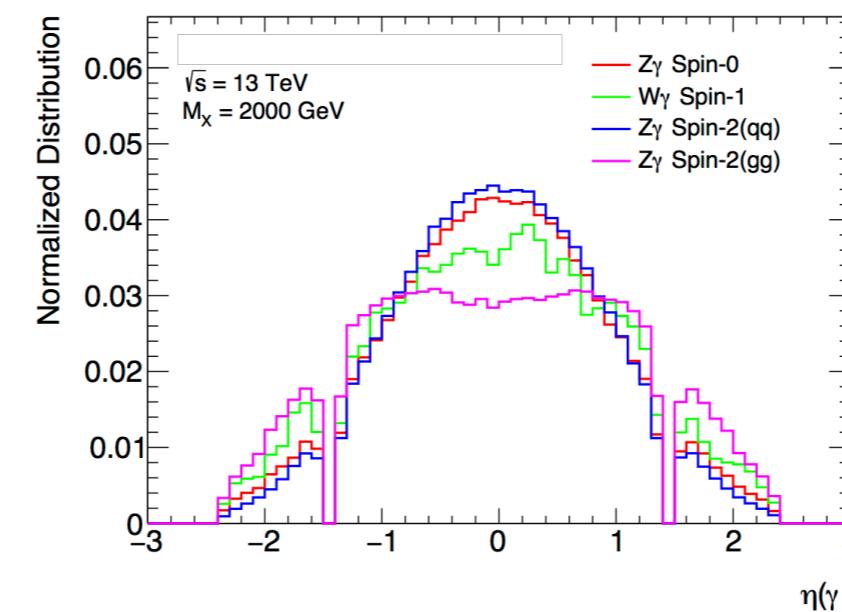
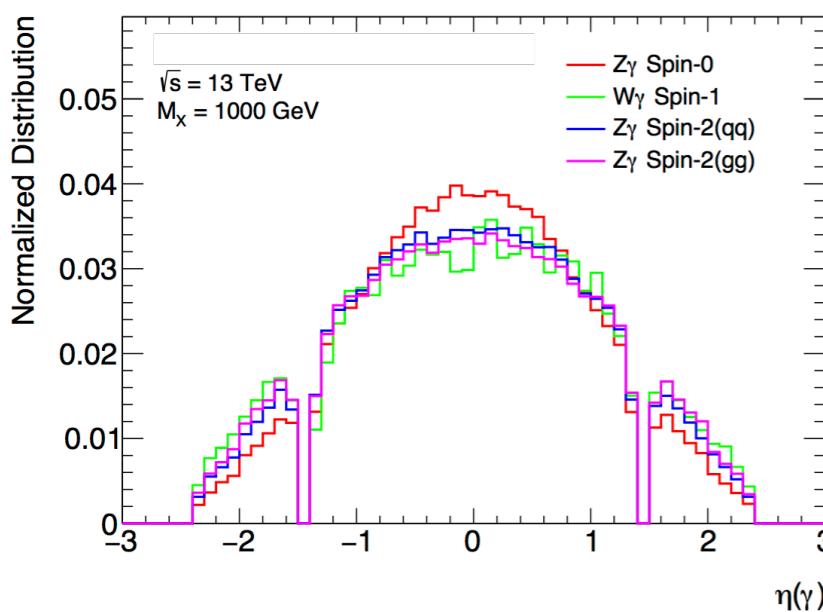
- The kinematics for each signal channels determines the different performance shown in signal efficiency plots.
- The pT distribution dominates the efficiency behavior in low mass region, and the angular distribution dominates high mass region.





Signal Comparison

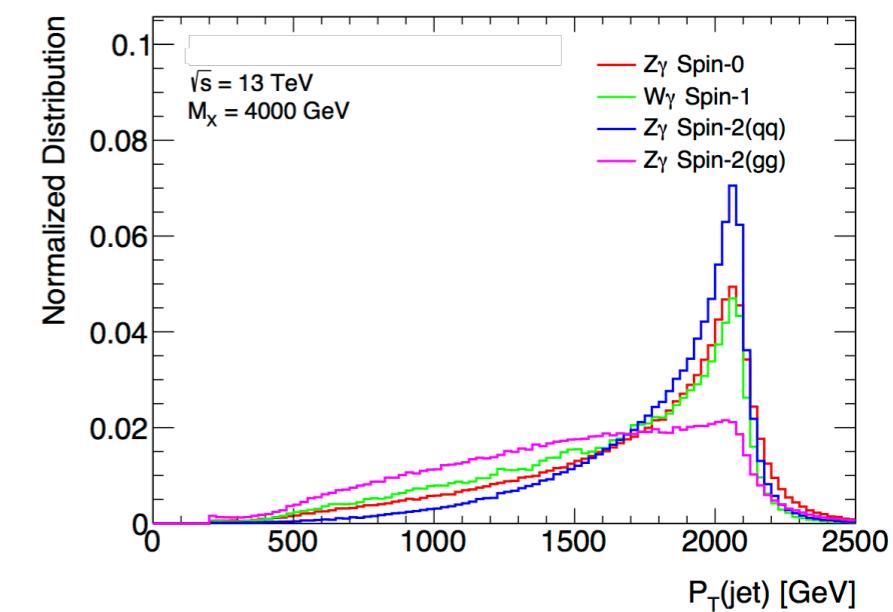
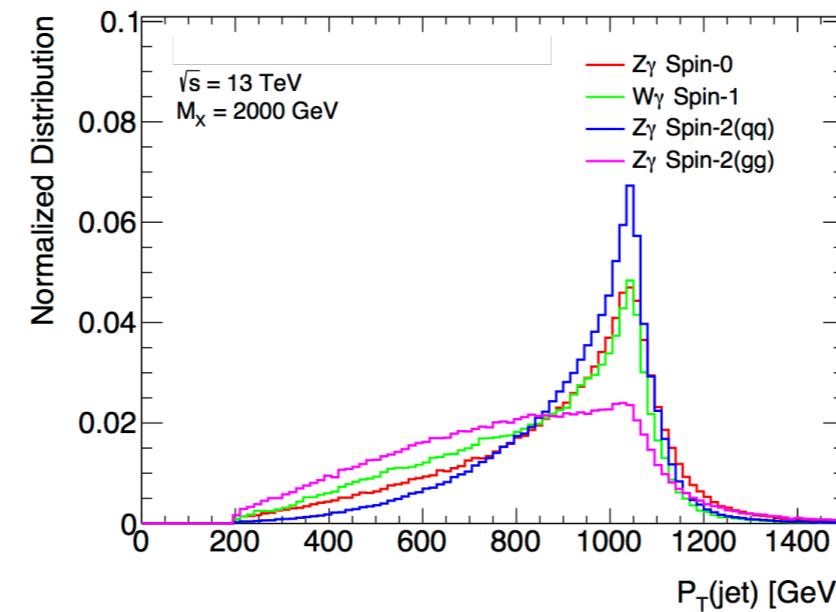
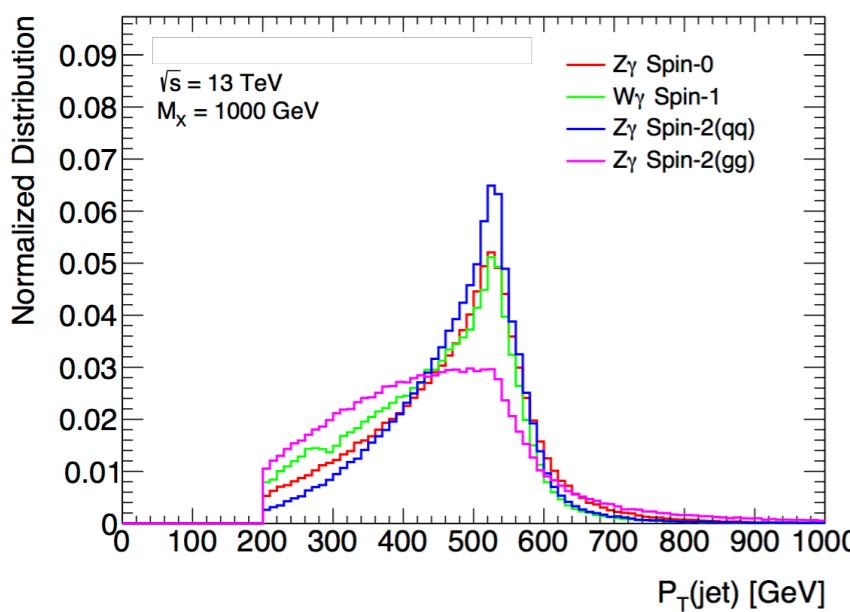
- The kinematics for each signal channels determines the different performance shown in signal efficiency plots.
- The pT distribution dominates the efficiency behavior in low mass region, and the angular distribution dominates high mass region.





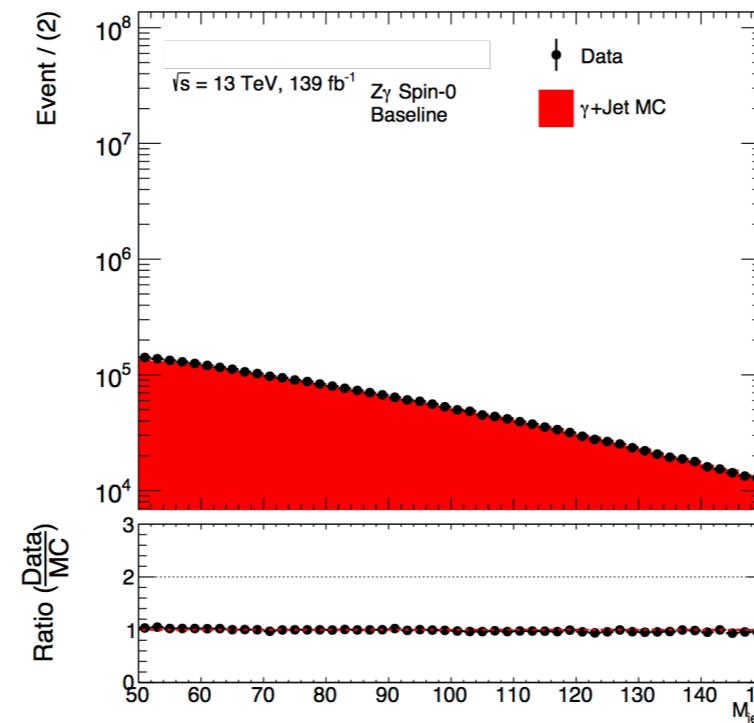
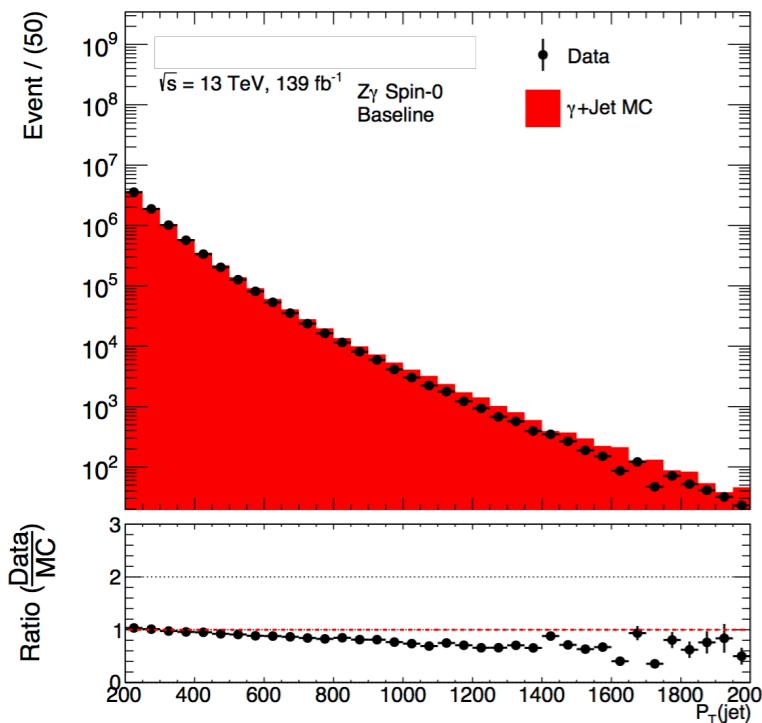
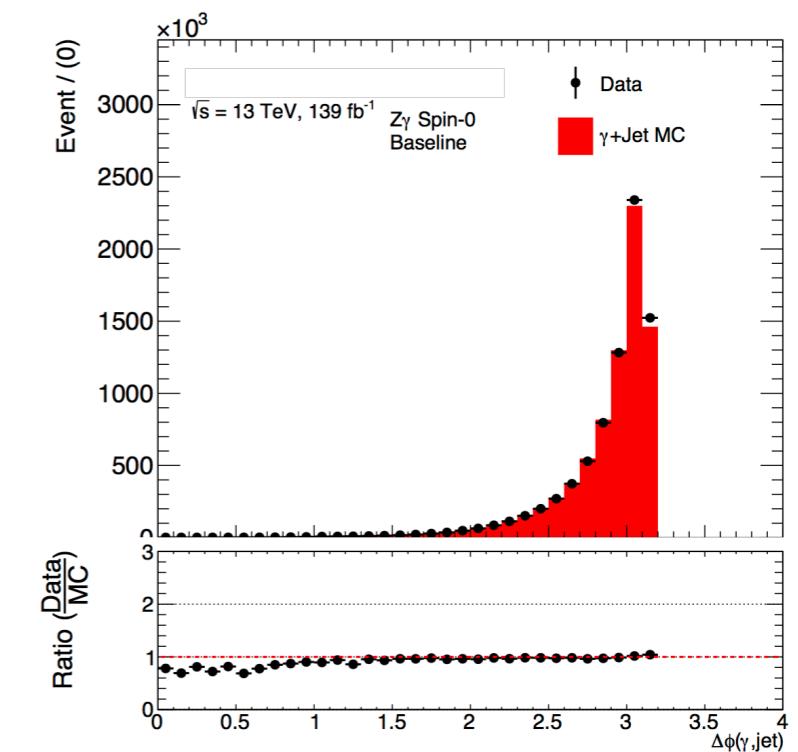
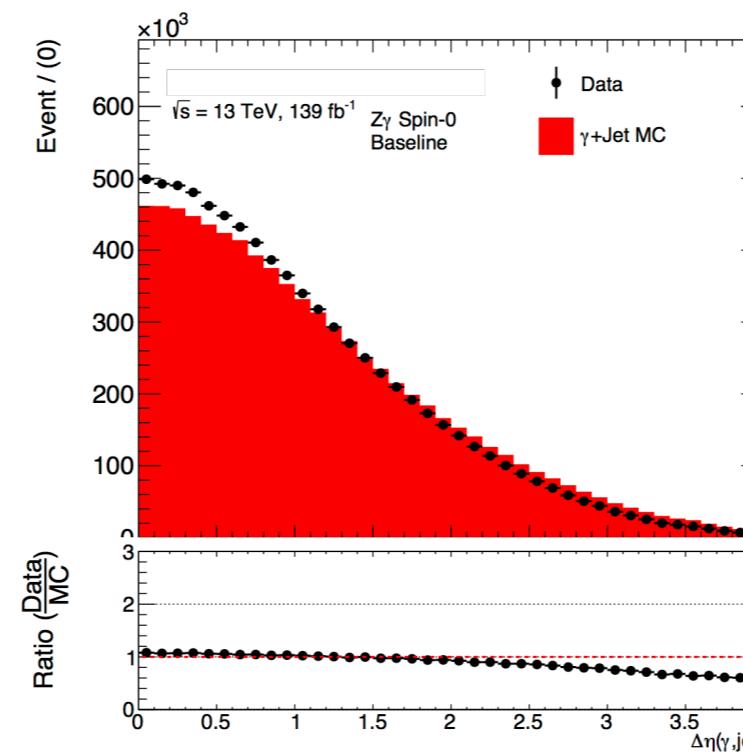
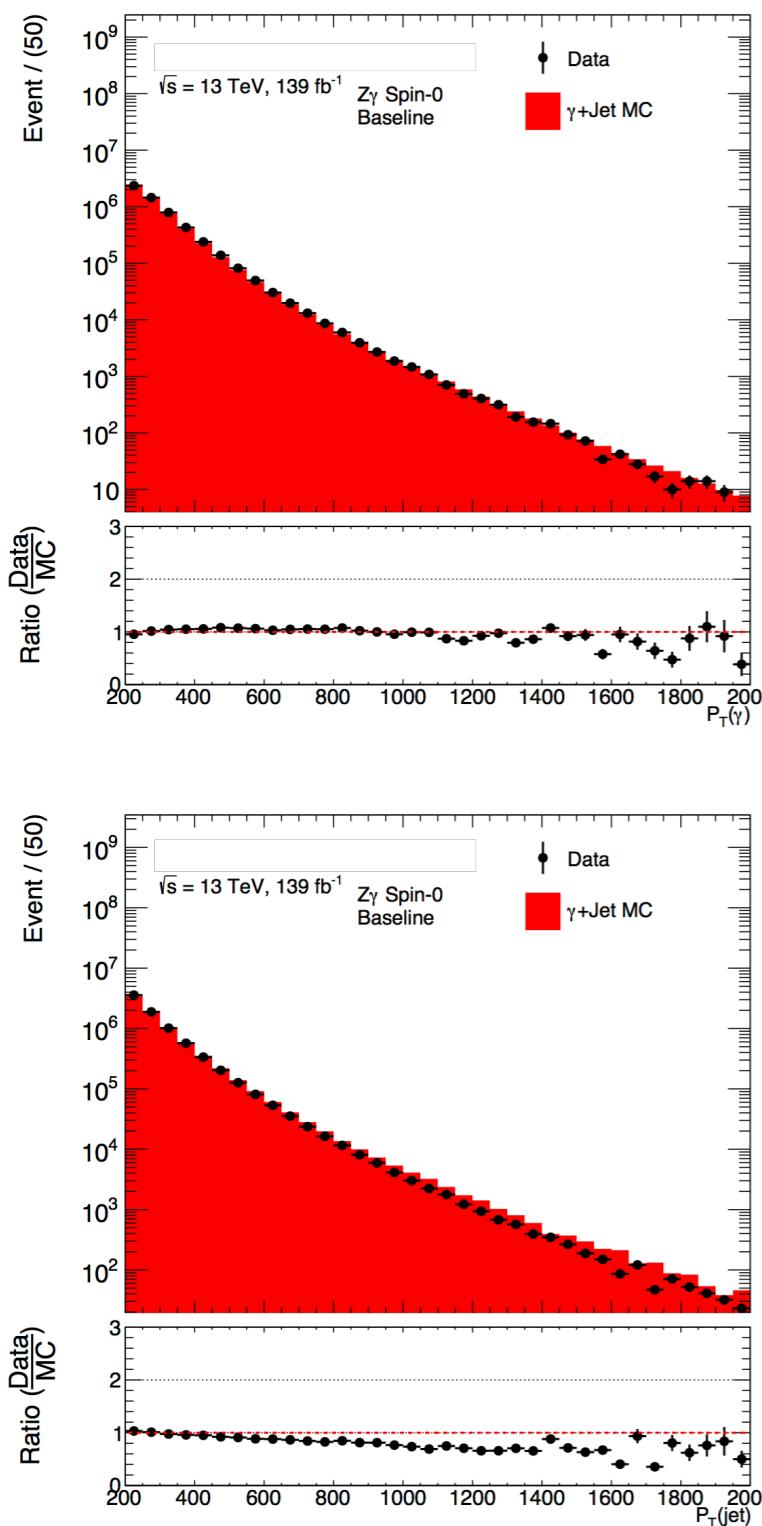
Signal Comparison

- The kinematics for each signal channels determines the different performance shown in signal efficiency plots.
- The pT distribution dominates the efficiency behavior in low mass region, and the angular distribution dominates high mass region.





Data/MC comparison

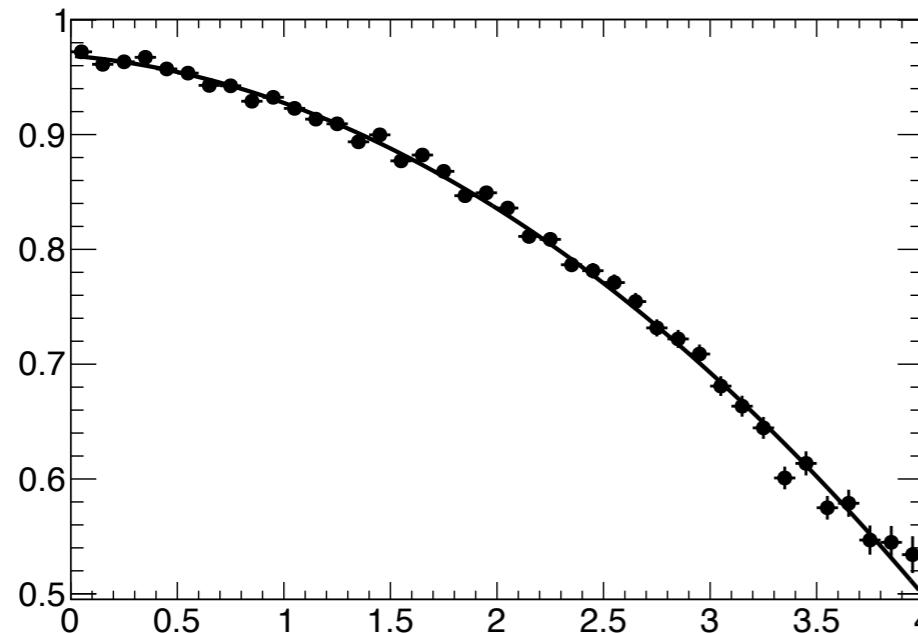


- No scale factors applied
- The ramp for jet pT is consistent with the sherpa NLO performance (<https://cds.cern.ch/record/2655001>)
- The study on re-weighted samples are done, no impact on our analysis strategy

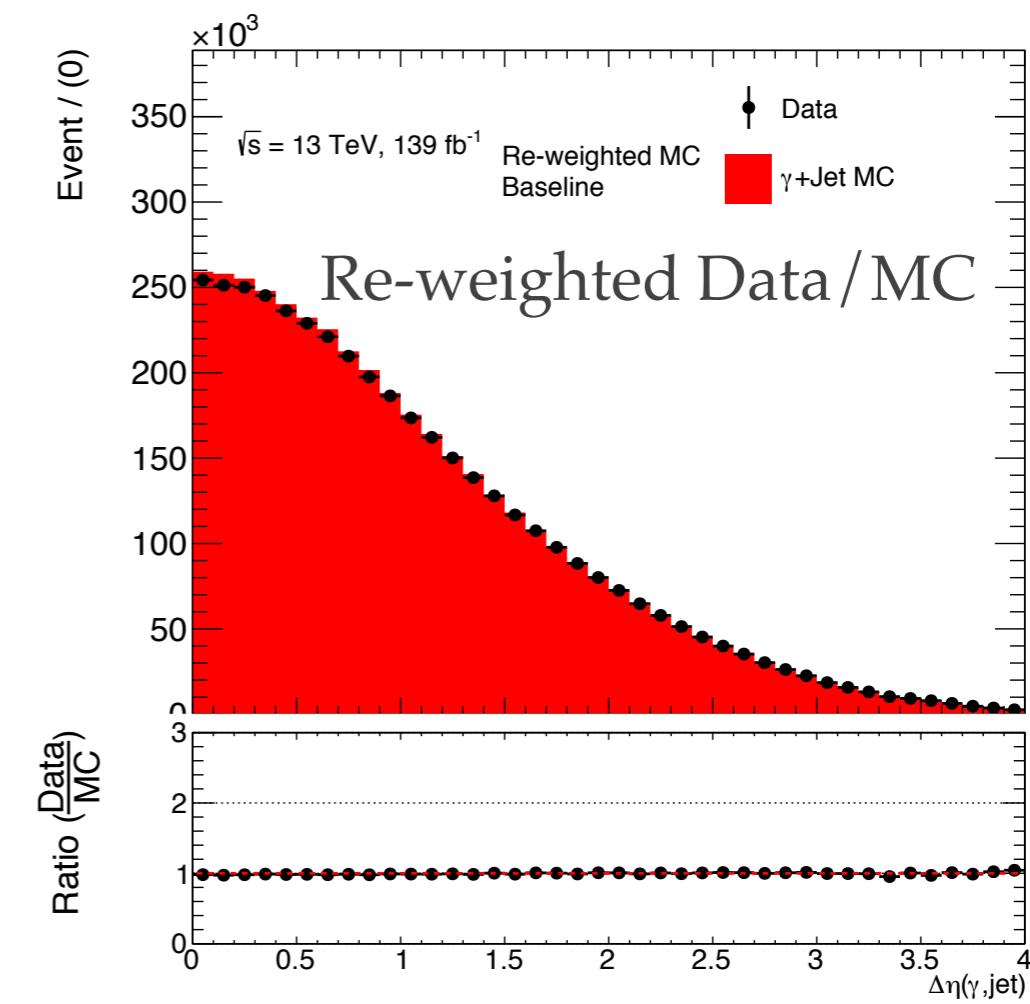


Re-weighted MC study

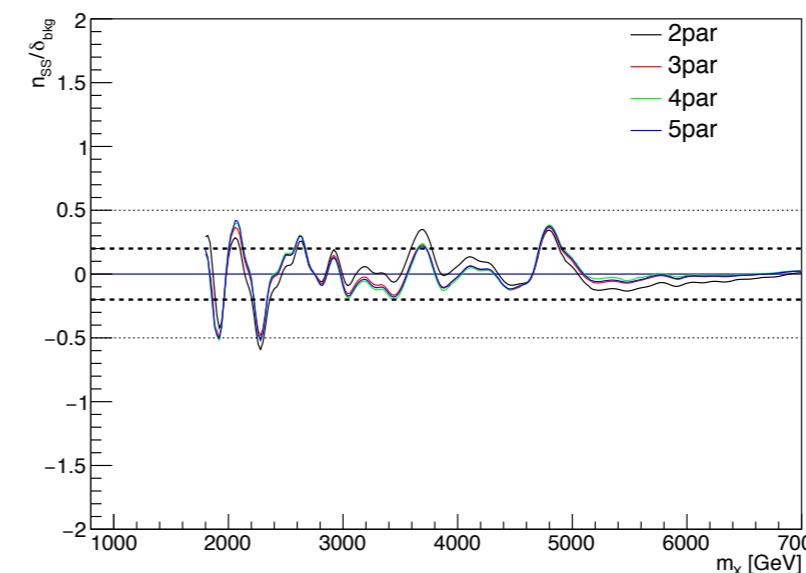
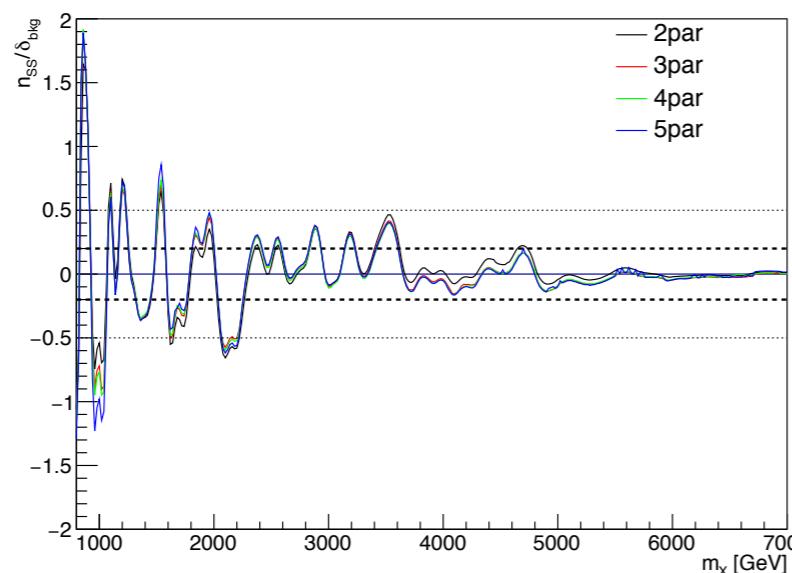
Ratio and Re-weighting function



$Z\gamma$ Spin-0 D2 Category



$Z\gamma$ Spin-0 Vmass Category





Fractional Spurious Signal Test

In order to identify the source of spikes in spurious signal test, the MC is evenly and randomly split into 2 or 3 subsets. Comparing the spurious signal test results for each of the subset, the correlation appears to be very weak. Therefore, the source of the spikes are probably from MC fluctuation limited by the size of the MC samples.

