Study of the jet multiplicity of a new heavy diphoton resonance in the *ATLAS* experiment

Stage overview (LPNHE Biennale) - 04/10/2016

Ahmed Tarek

Supervised by Sandrine Laplace

ATLAS EXPERIMENT

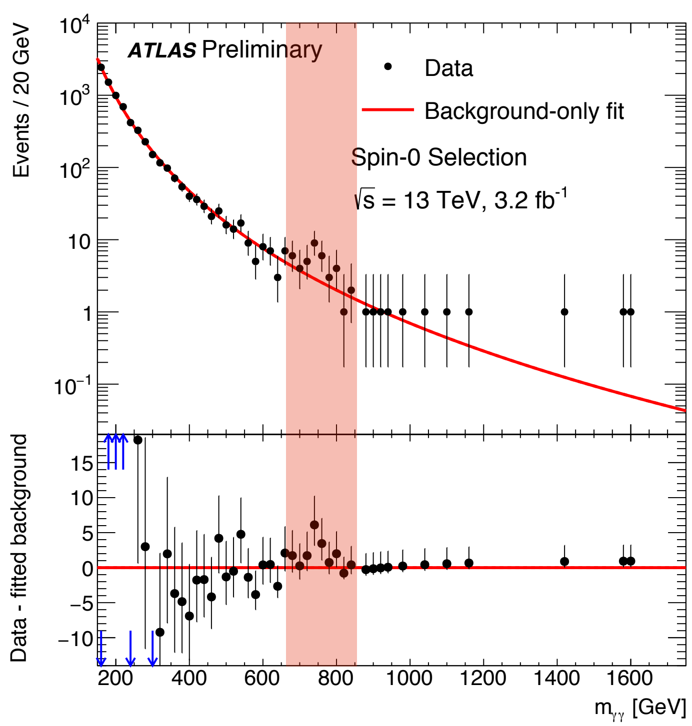
LPNHE PARIS

Introduction

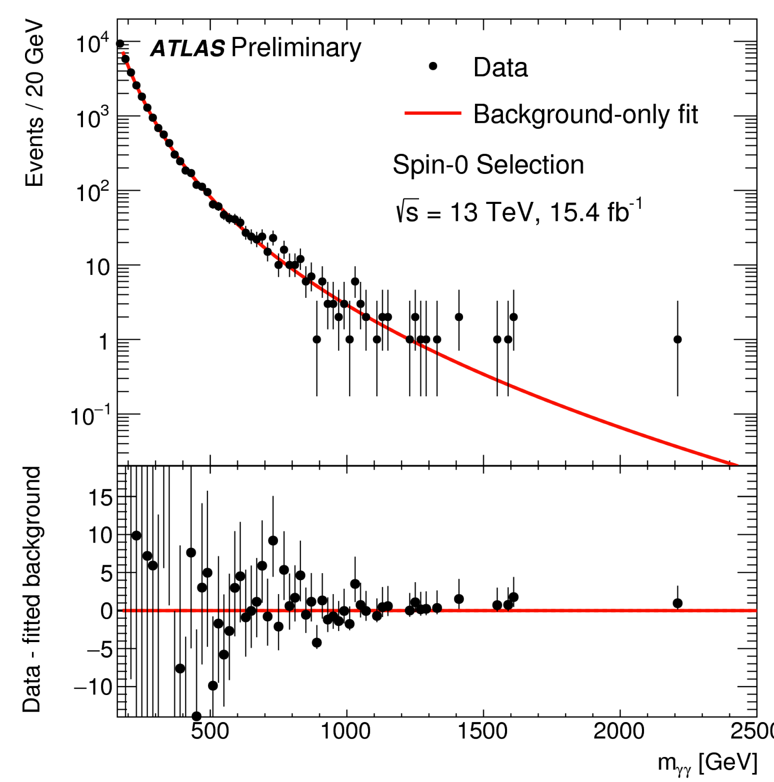
Motivation - diphoton channel

- Search for high mass resonances decaying to two photons is motivated by several extensions to the standard model (SM)
- Diphoton searches benefit from a **very clean experimental signature**
  - Two **high p_T photon candidates** reconstructed from energy deposit in the calorimeter
  - Resonance appears in the invariant mass spectrum as a **local excess of events** over **well-know background** of SM direct photon pair production
- Using 2015 run-2 data at $\sqrt{s}$ = 13 TeV, both ***ATLAS*** and **CMS** reported a moderate excess of events around mass of ~ 750 GeV

2015



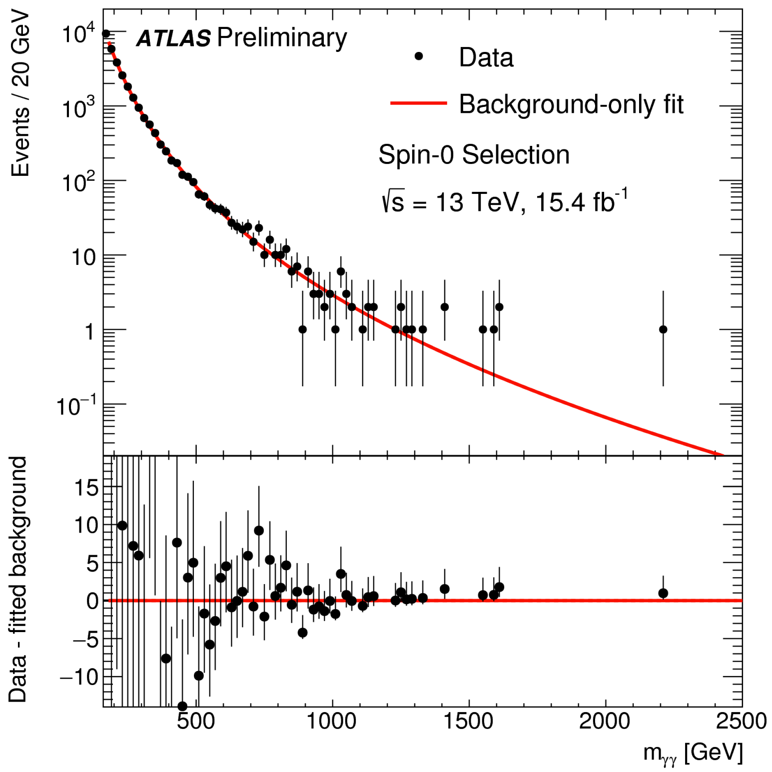
2016



Introduction

Motivation - diphoton channel

- Search for high mass resonances decaying to two photons is motivated by several extensions to the standard model (SM)
- Diphoton searches benefit from a **very clean experimental signature**
  - Two **high p_T photon candidates** reconstructed from energy deposit in the calorimeter
  - Resonance appears in the invariant mass spectrum as a **local excess of events** over **well-know background** of SM direct photon pair production
- Using 2015 run-2 data at $\sqrt{s}$ = 13 TeV, both ***ATLAS*** and **CMS** reported a moderate excess of events around mass of ~ 750 GeV



Categorization
Overview

Categorization is the study of data subsamples with specific characteristics and analyzing them independently

- **Enhance the significance**
- **Measure properties of the process**

Model independent *Based on detector characteristic*
e.g. η , photon conversion ..etc

In this case, signal fractions per category is determined a-priori

Model dependent *Based on resonance dynamics*
e.g. $\boldsymbol{N_{jets}}$*, Missing transverse energy ..etc*

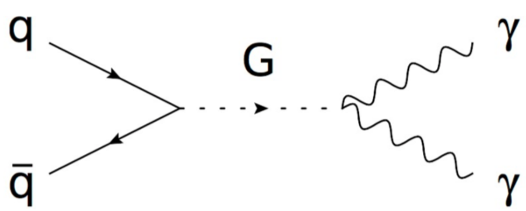
Here, signal fractions per category is determined from data

Data is divided into 4 categories chosen from **[0-1-2- ≥3] jets**

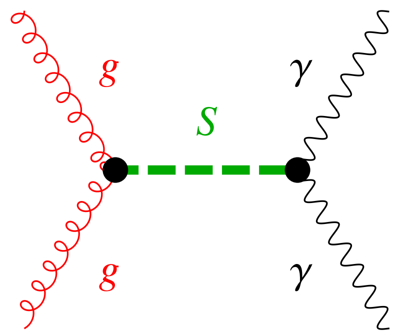
- **p_T > 25 GeV, |η| < 4.4, JVT cuts**

Study of Jet Multiplicity

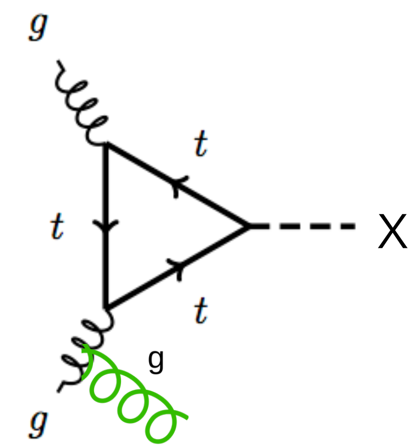
Production Mechanism



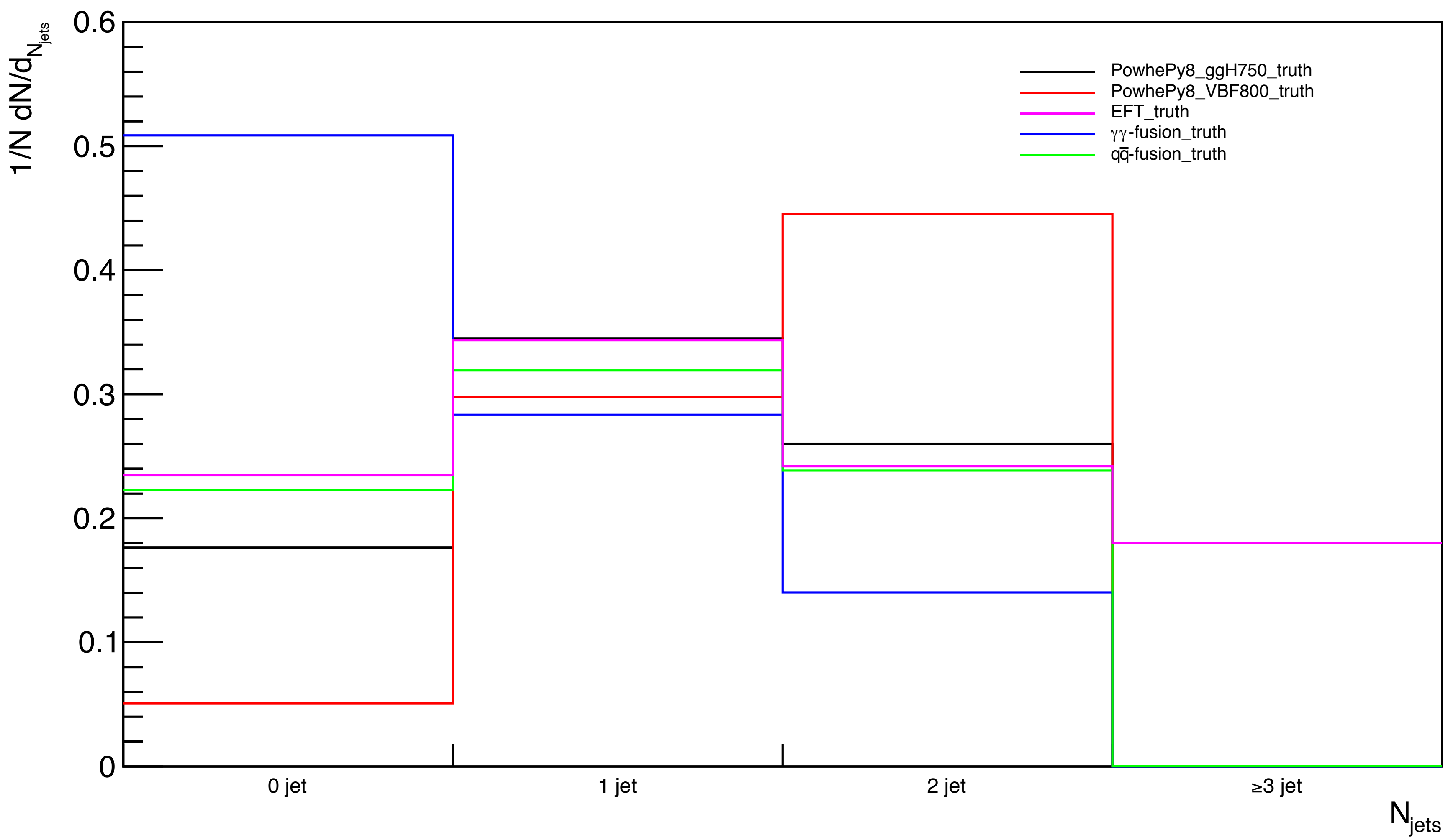
- Quark-antiquark annihilation

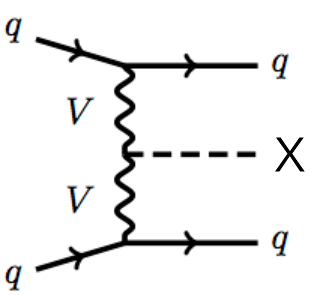


- Effective field theory EFT

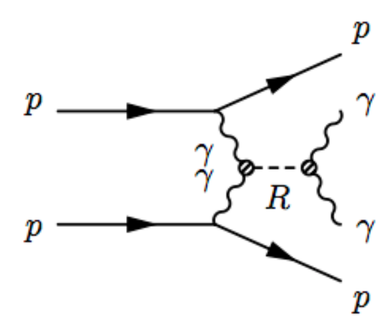


- SM-like Higgs : Gluon-gluon fusion ggH





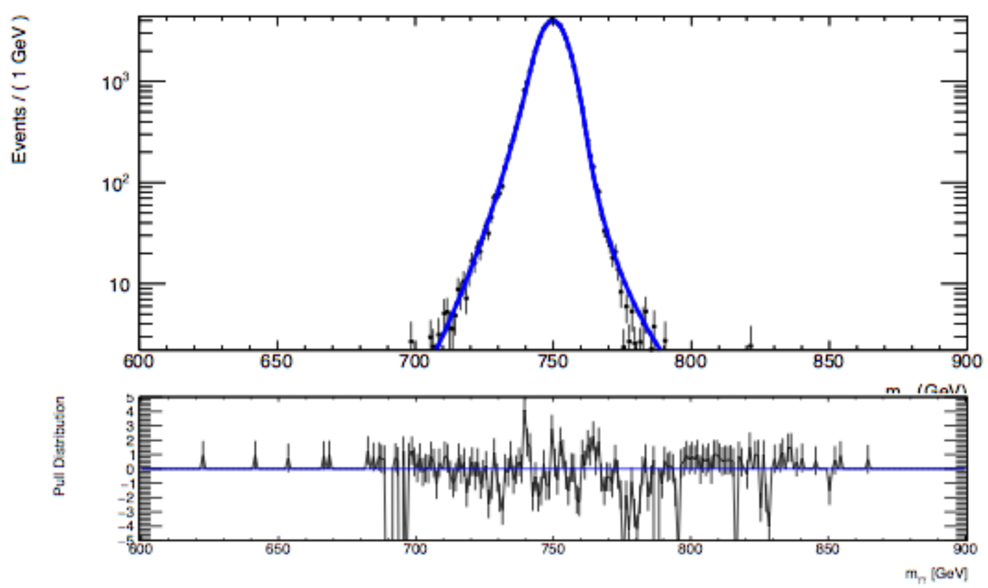
- SM-like Higgs : Vector boson fusion VBF

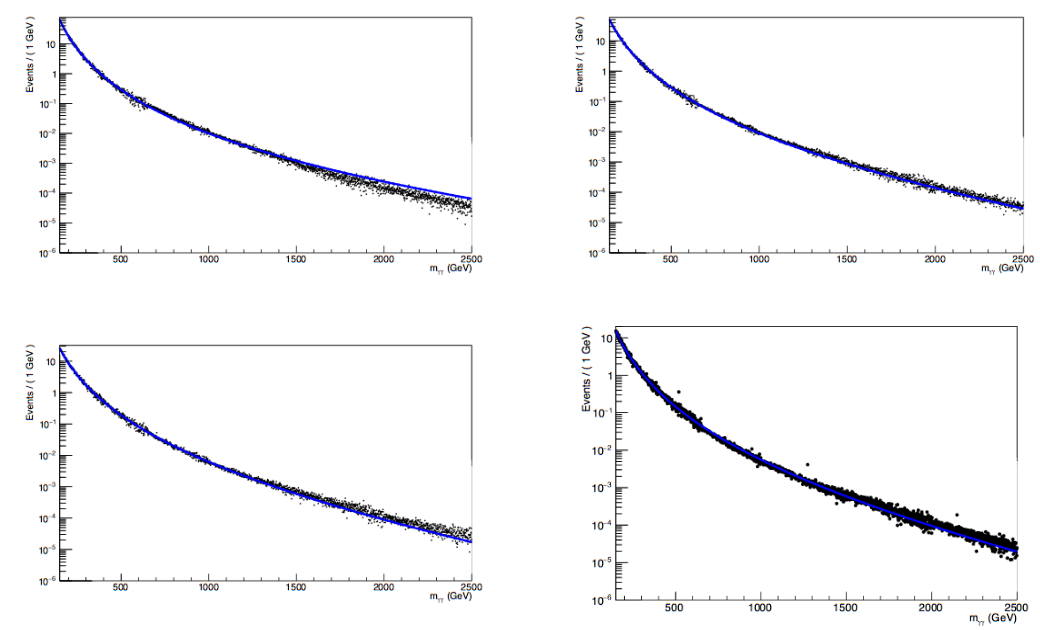


- Photon-photon fusion

Analysis Overview

- The number of signal and background events are extracted from **Likelihood fit** on the $m_{\gamma\gamma}$ distribution
  - Using models of the **signal** and SM **background**





- The Likelihood function per event is defined as follows

$$\begin{aligned}\mathcal{L}(m_{\gamma\gamma};\sigma_{fid},m_X,\alpha_X,N_{bkg},\boldsymbol{a},\boldsymbol{\theta}) &= \boxed{N_X(\sigma_{fid},m_X,\boldsymbol{\theta}_{N_X},\theta_{SS})f_X(m_{\gamma\gamma},\boldsymbol{x}_X(m_X,\alpha_X),\theta_{\sigma})}\\ &+ \boxed{N_{bkg}f_{bkg}(m_{\gamma\gamma},\boldsymbol{a})}\end{aligned}$$

- An integer index is added to the likelihood to discriminate the different categories ⇒ multiplying the number of fit parameters

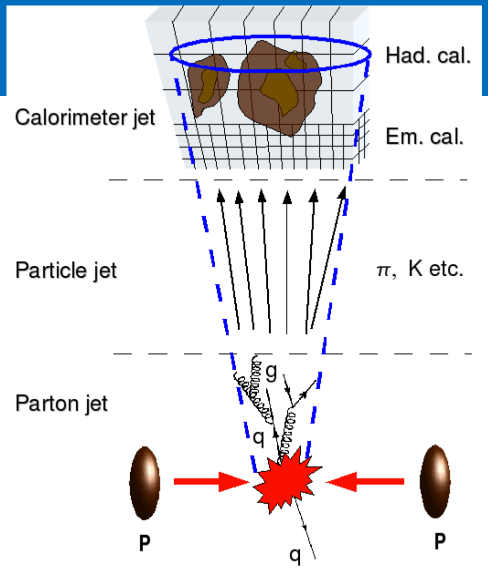
Impact of jet multiplicities on significance

- Compute the expected significance from an injected signal (2.2 fb)
- The expected significance is computed twice
  - **inclusive analysis** as done in ***ATLAS*** so far and the expected significance for the injected cross -section is **2.7 σ**
  - Adding **jet multiplicity** information (i.e. perform 4 fits instead of 1, for N_{jet}=0,1,2,≥3) the expected significance is

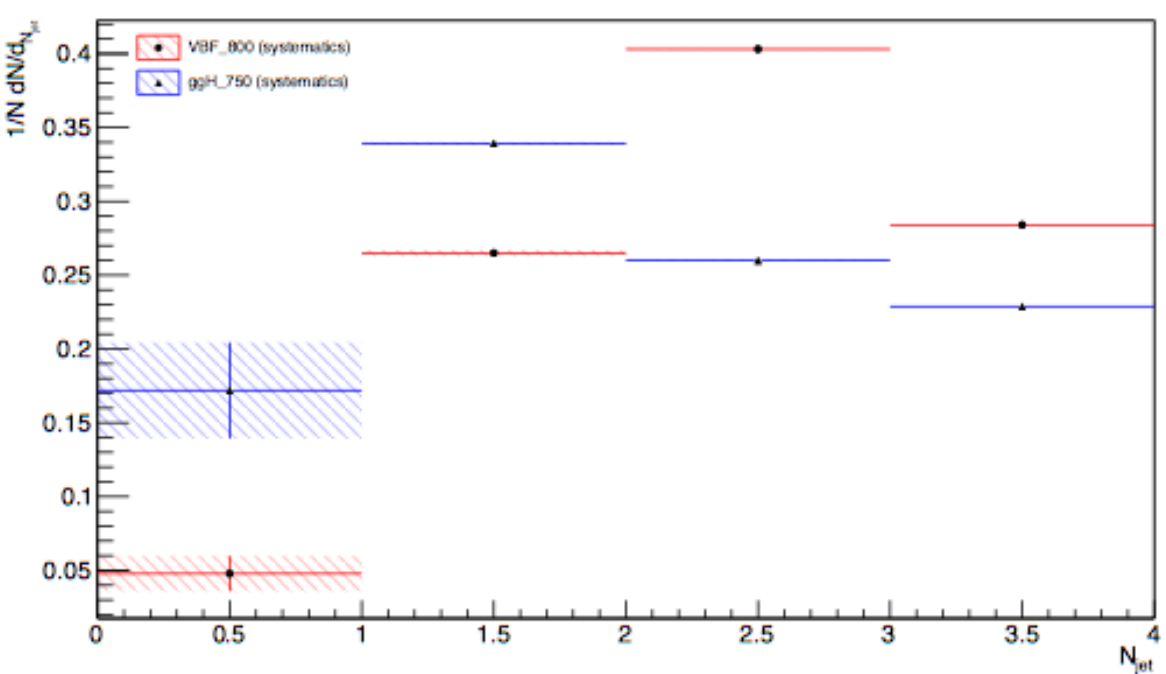
	ggH	qq	EFT	VBF	γγF
combined	2.8 σ	2.8 σ	2.7 σ	**3.0 σ**	**2.9 σ**

- **No major change in significance is observed**

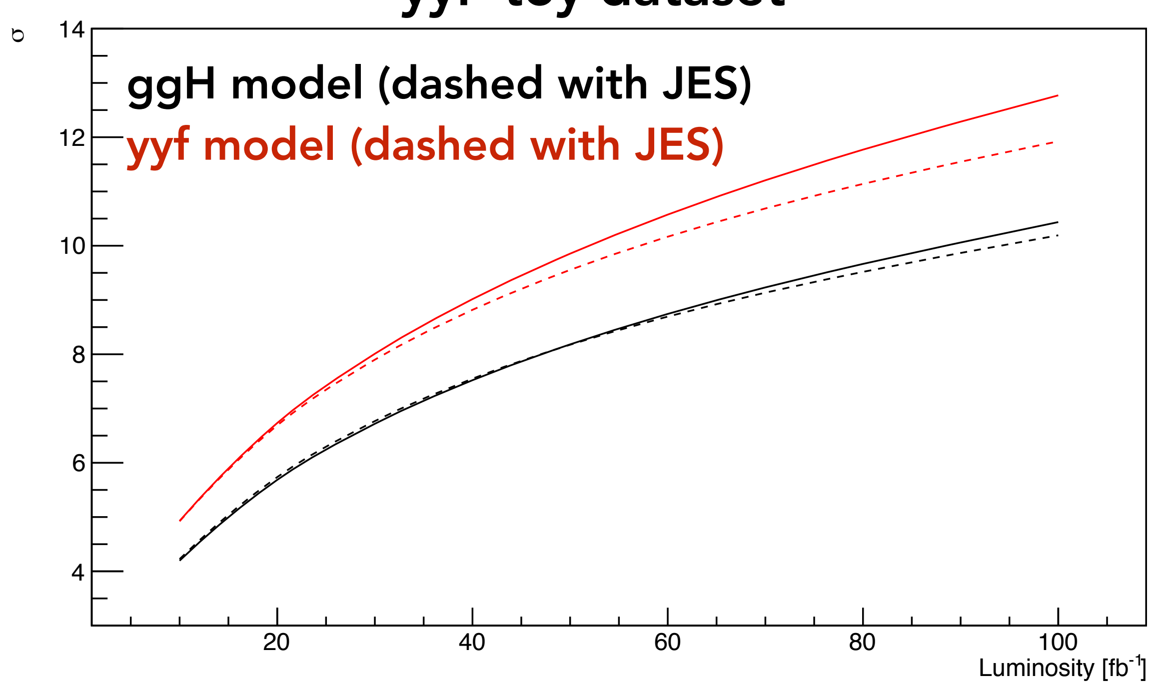
Jet Systematics



- **Jets are difficult to calibrate** and Jet Calibration systematics can affect jet p_T
- Jets with p_T around the threshold will be affected by systematic variations causing ***bin migration***
- Including the jet systematics to the fit requires **reduction to the full set of jet systematic uncertainties** (up to 17 parameter) without loosing correlations
- Biased fit is performed using fixed fractions to distinguish different models



yyF toy dataset



Summary and outlook

Jet multiplicities can be used to understand the production mechanism of the resonance

- **N_{jet} categories dont increase the discovery significance significantly, but allow to distinguish between different production modes**
- **Still true even with JES uncertainties.**

Outlook

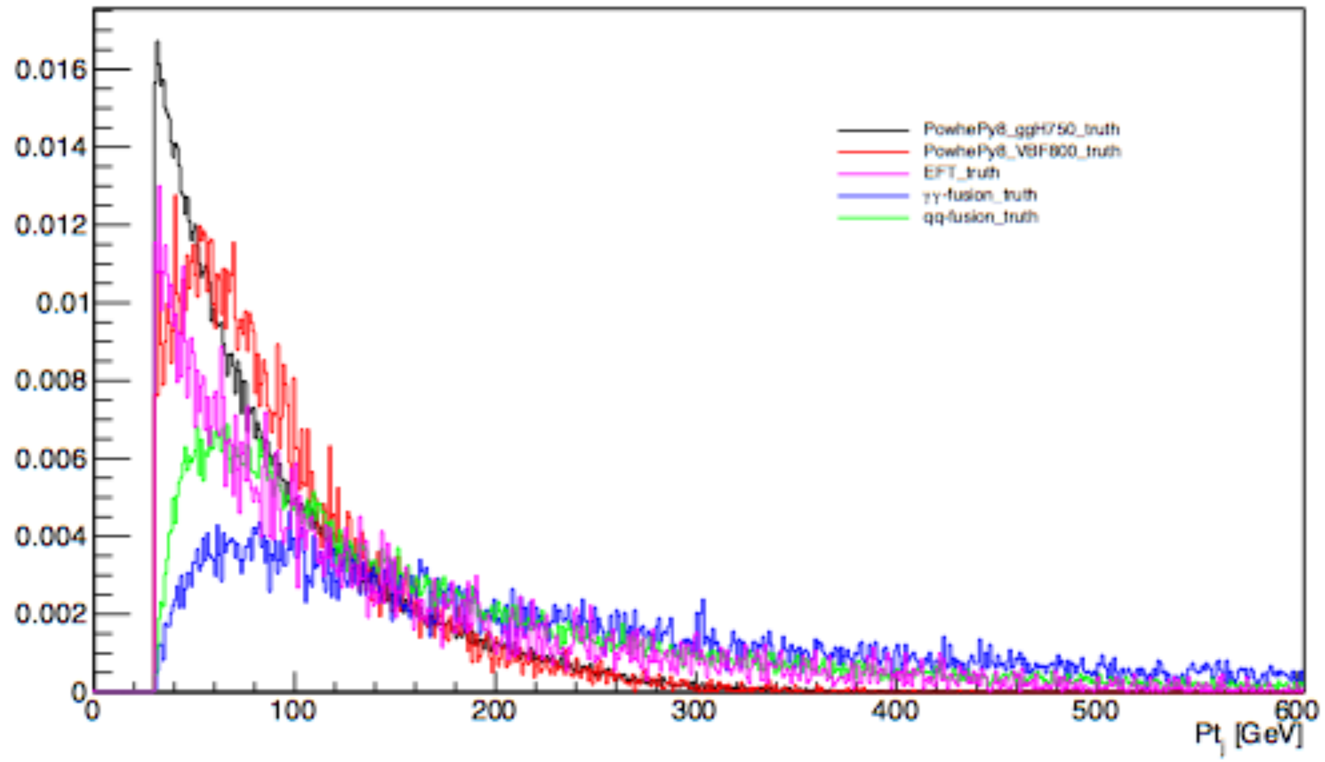
- **Started qualification task : E/Gamma Calibration : The pre-sampler scale**
  - Check the stability of the presampler energy scale as function of presampler HV changes
  - Determine the presampler scale for 2015-2016 data
- **Analysis : measurement of cross-sections in H(125)-> YY for constraints on EFTs**

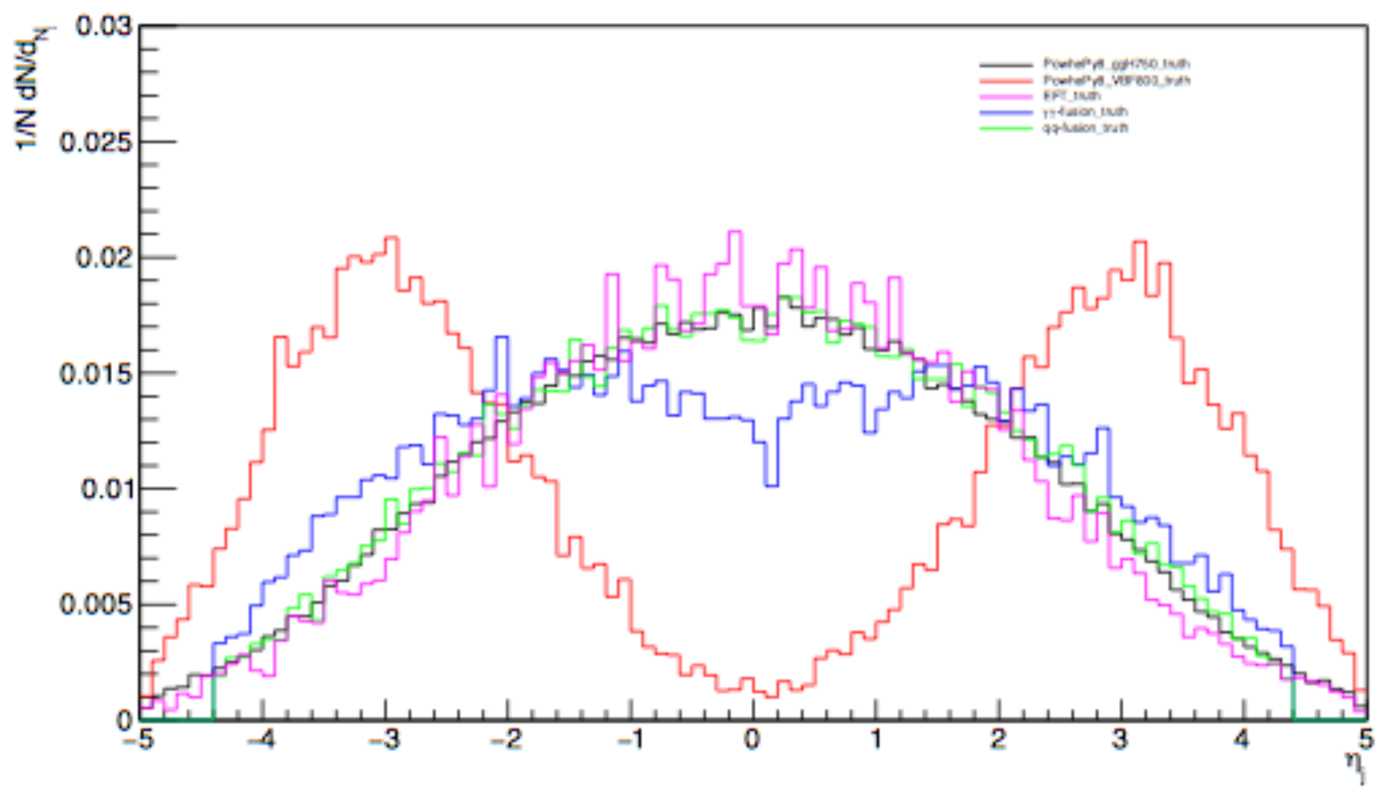
Thank you for your attention

Backup

Categorization

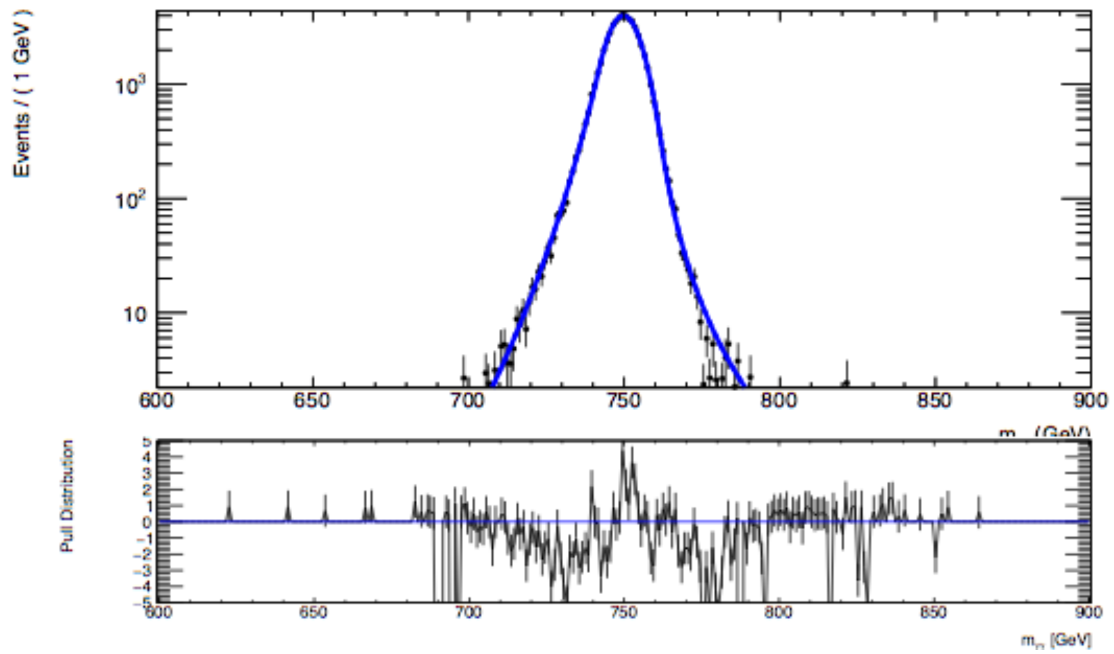
Other jet variables





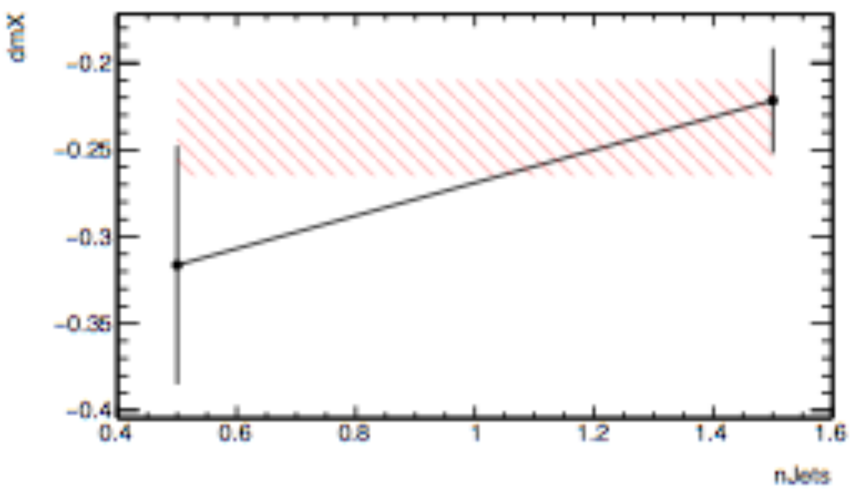
Analysis Overview

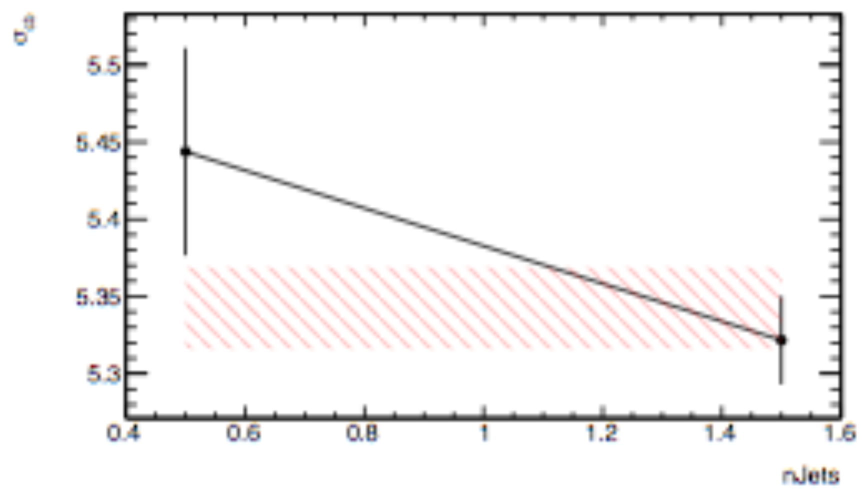
Signal Parametrization (SP) - 2jet cat.

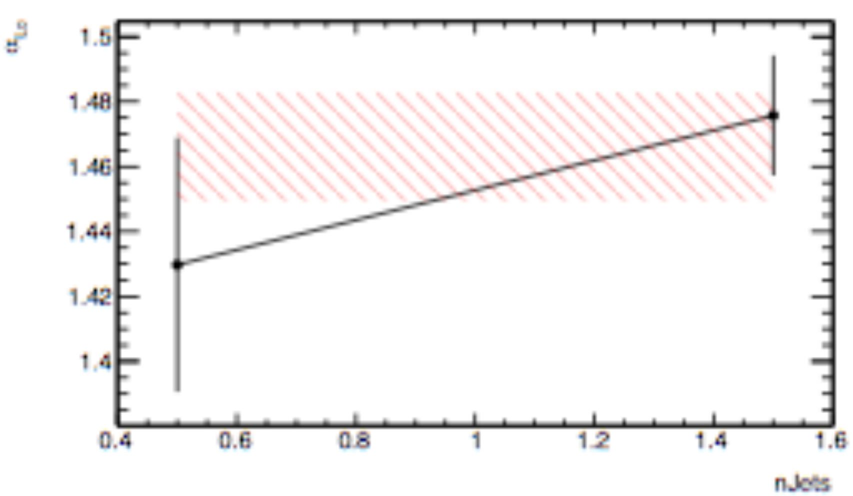


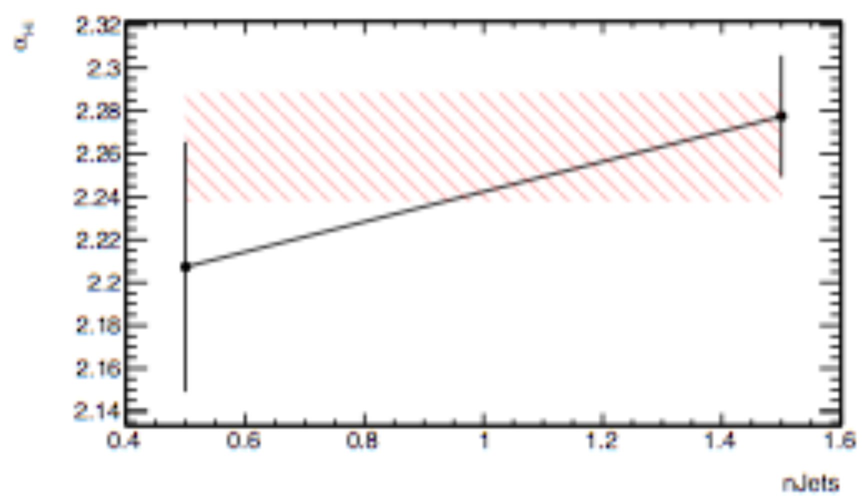
Signal Fractions / category :

- 0j : 0.17
- ≥1j : 0.83





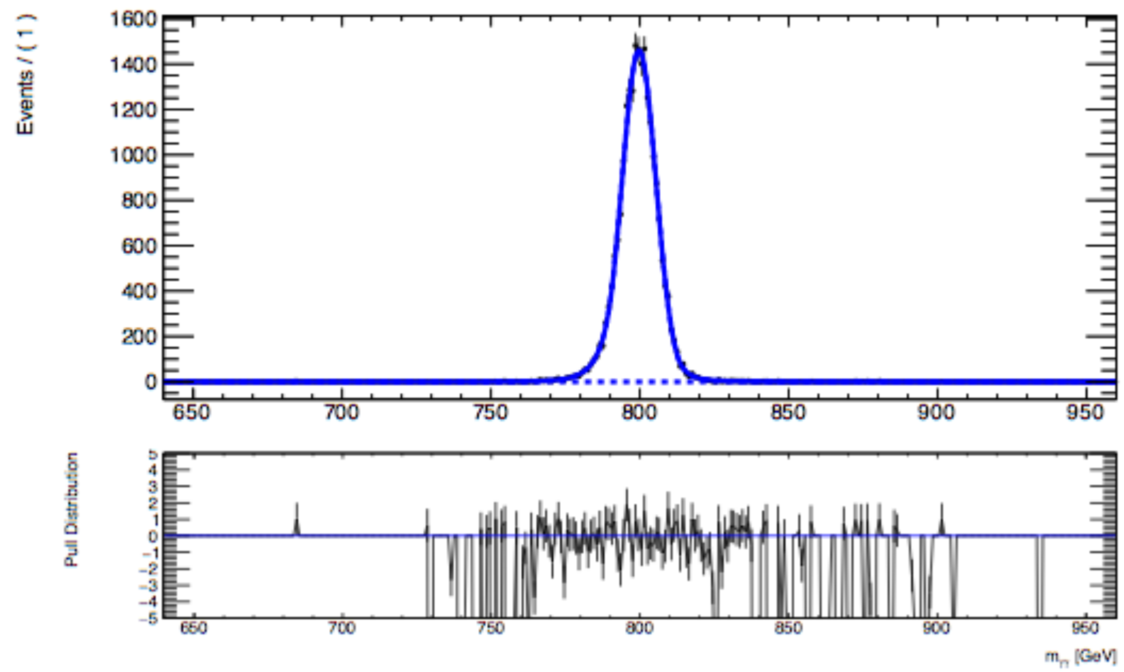




nLo=8.9 and nHi =3.3

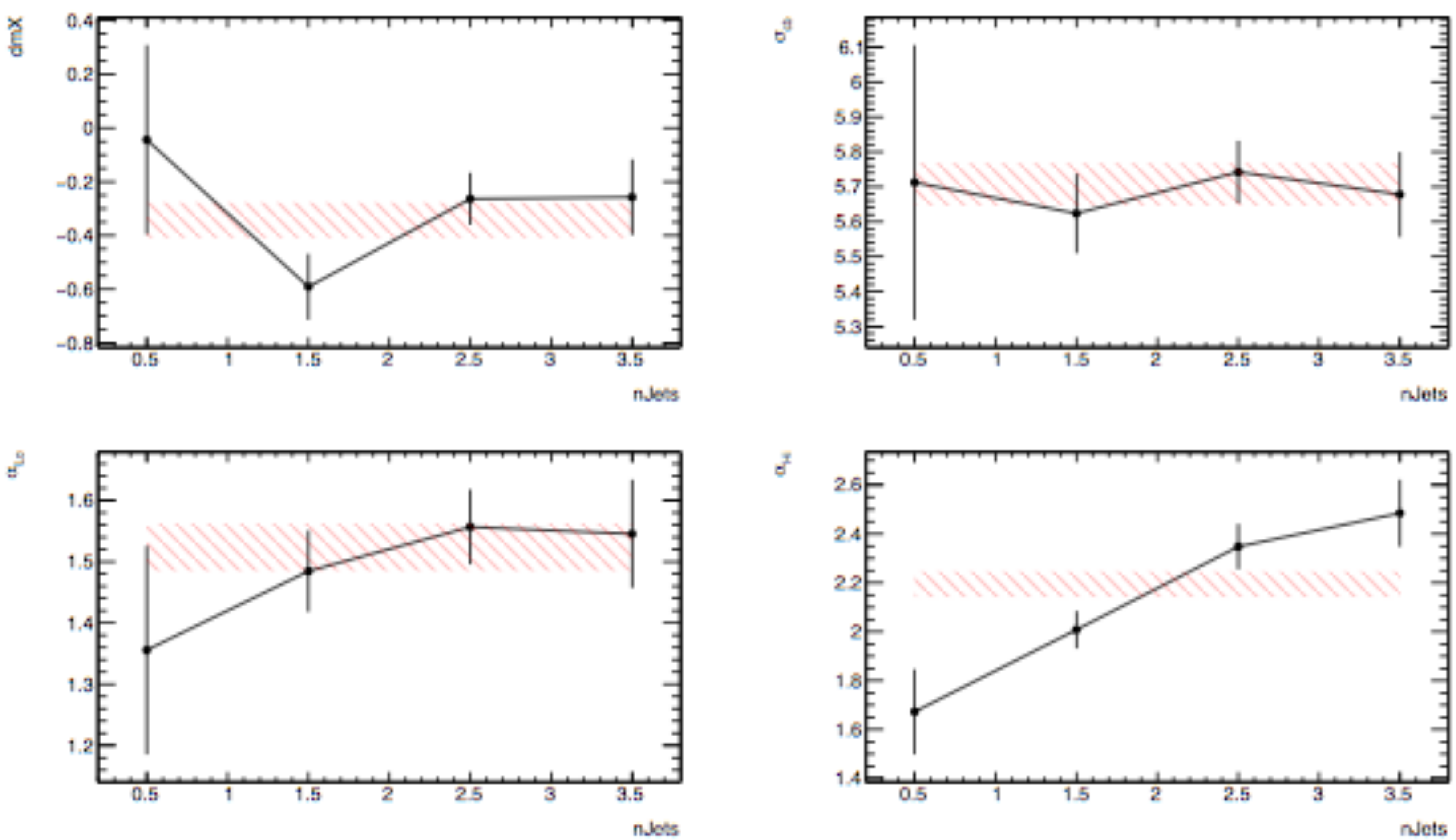
Analysis Overview

Signal Parametrization (SP) - VBF



Signal Fractions / category :

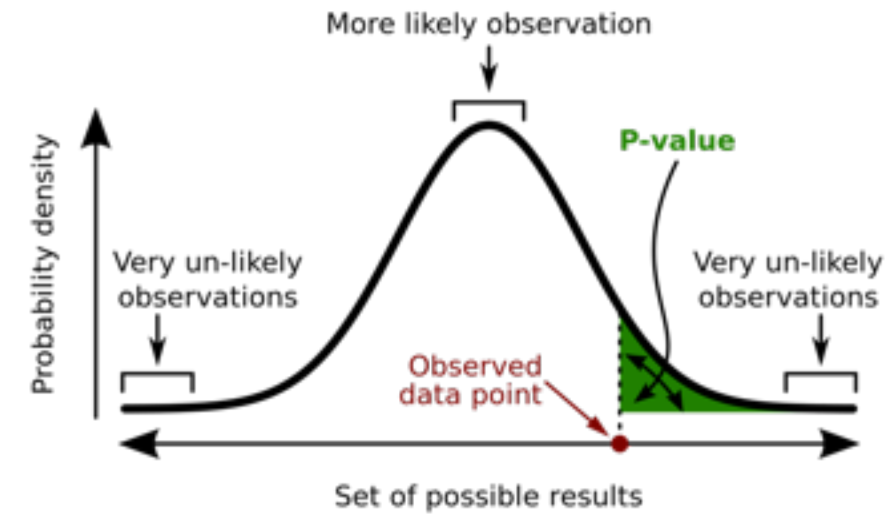
- 0j : 0.05
- 1j : 0.30
- 2j : 0.45
- ≥3j : 0.20



nLo=8.9 and nHi =3.3

Analysis Overview
Statistical Modeling

- The **p-value test** gives the probability that result is as or less compatible with the background only hypothesis
  - p-value is calculated from **profile likelihood ratio**



- The Likelihood function per event is defined as follows

$$\begin{aligned}\mathcal{L}(m_{\gamma\gamma};\sigma_{fid},m_X,\alpha_X,N_{bkg},\boldsymbol{a},\boldsymbol{\theta}) &= N_X(\sigma_{fid},m_X,\boldsymbol{\theta}_{N_X},\theta_{SS})f_X(m_{\gamma\gamma},\boldsymbol{x}_X(m_X,\alpha_X),\theta_\sigma)\\ &+ N_{bkg}f_{bkg}(m_{\gamma\gamma},\boldsymbol{a})\end{aligned}$$

- N_X and N_{bkg} are the signal and background number of events, f_X and f_{bkg} are the signal and background Probability density functions

- θ is nuisance parameters to describe systematic uncertainties

- An integer index is added to the likelihood to discriminate the different categories hence multiplying the number of fit parameters

Analysis Overview
Statistical modeling

- The full likelihood is shown below

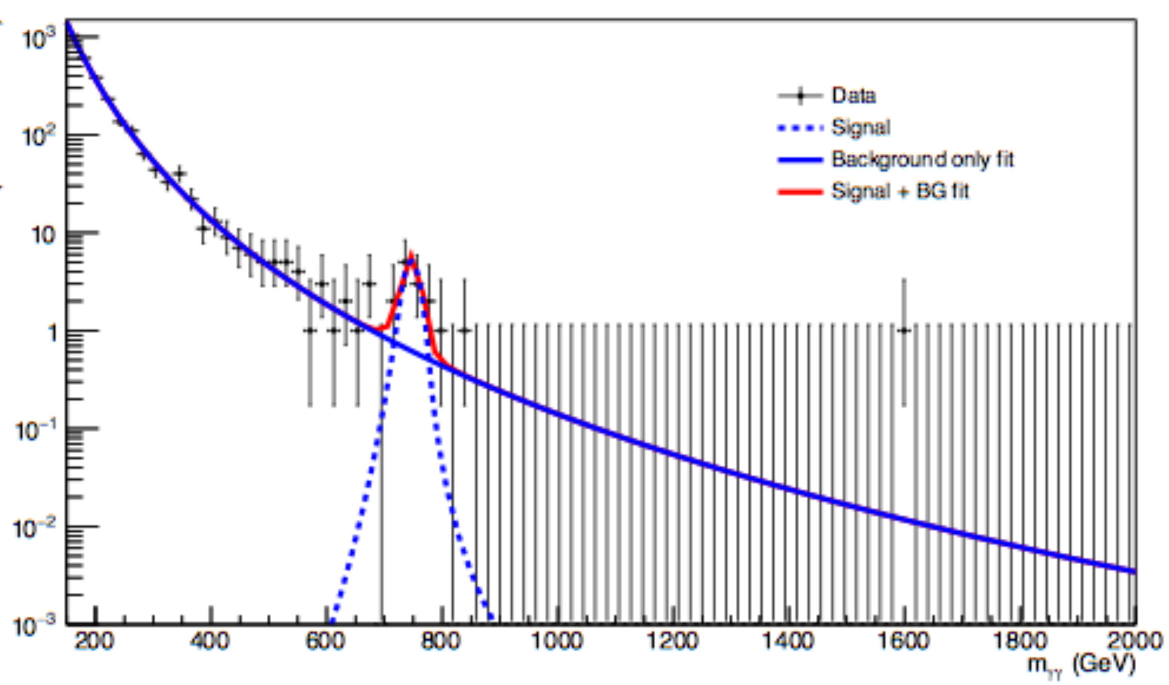
$$\mathcal{L}(\sigma_{fid}, m_X, \alpha_X, N_{bkg}, \boldsymbol{a}, \boldsymbol{\theta}) =$$
$$e^{-(N_X+N_{bkg})} \left[\prod_{i=1}^{n} \mathcal{L}(m_{\gamma\gamma_i}; \sigma_{fid}, m_X, \alpha_X, N_{bkg}, \boldsymbol{a}, \boldsymbol{\theta})\right] \left[\prod_{k=1}^{\dim \boldsymbol{\theta}} \exp\left(-\frac{1}{2}\left(\theta_i - \theta_i^{\mathrm{aux}}\right)^2\right)\right]$$

θ^{aux} *are set of auxiliary measurement to constrain systematic uncertainties*

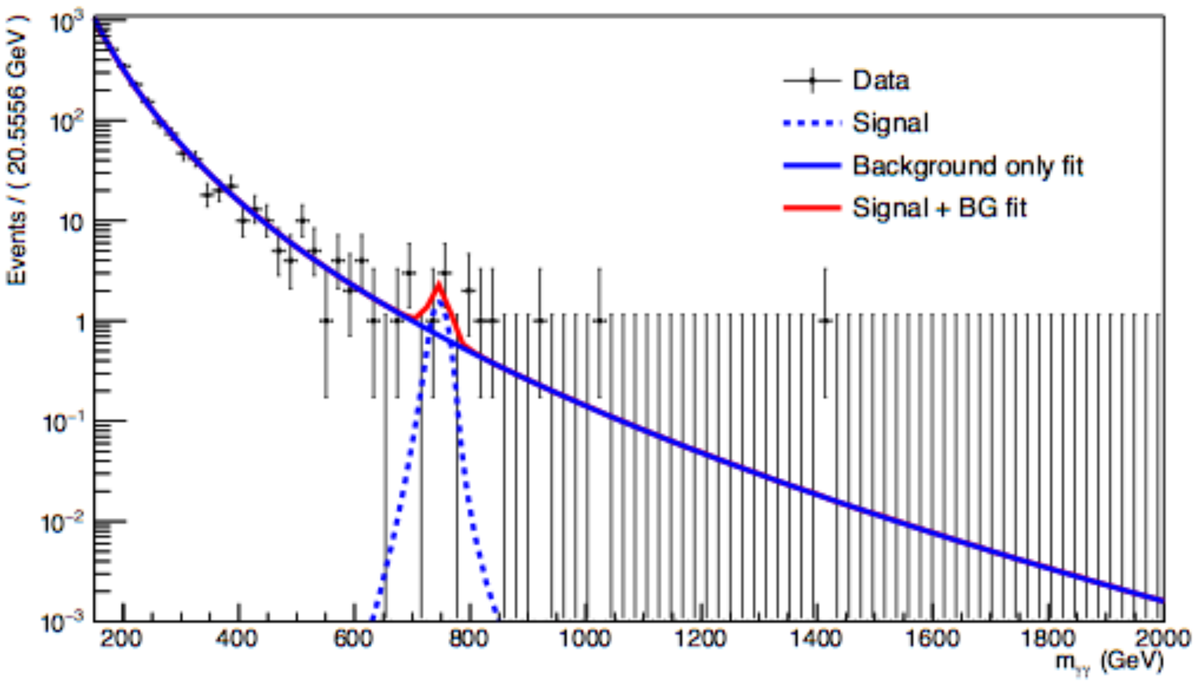
- The local significance Z^{local} is computed from the asymptotic approximation as $Z=\sqrt{q_0}$
- Expected significance is computed from **Asimov dataset** which is a generated single dataset corresponding to the exact model PDF
- The analysis was done using `HFitter + RooFit` packages assuming NW and fixed nuisance parameters

Analysis Overview

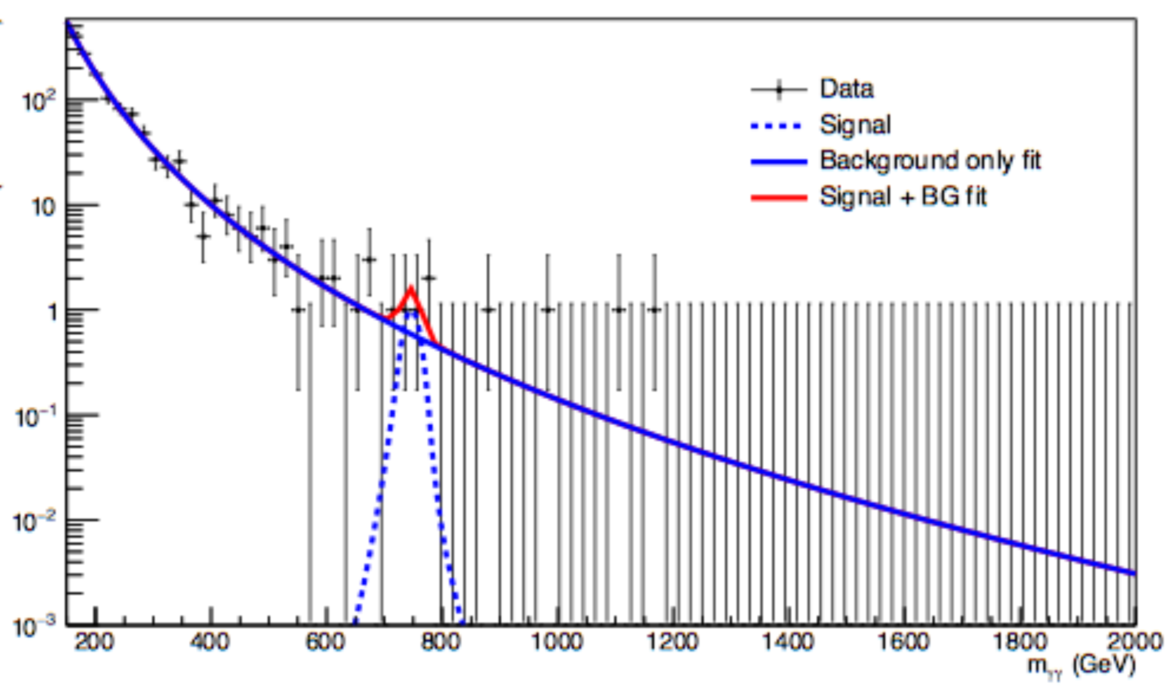
S+B fit on data



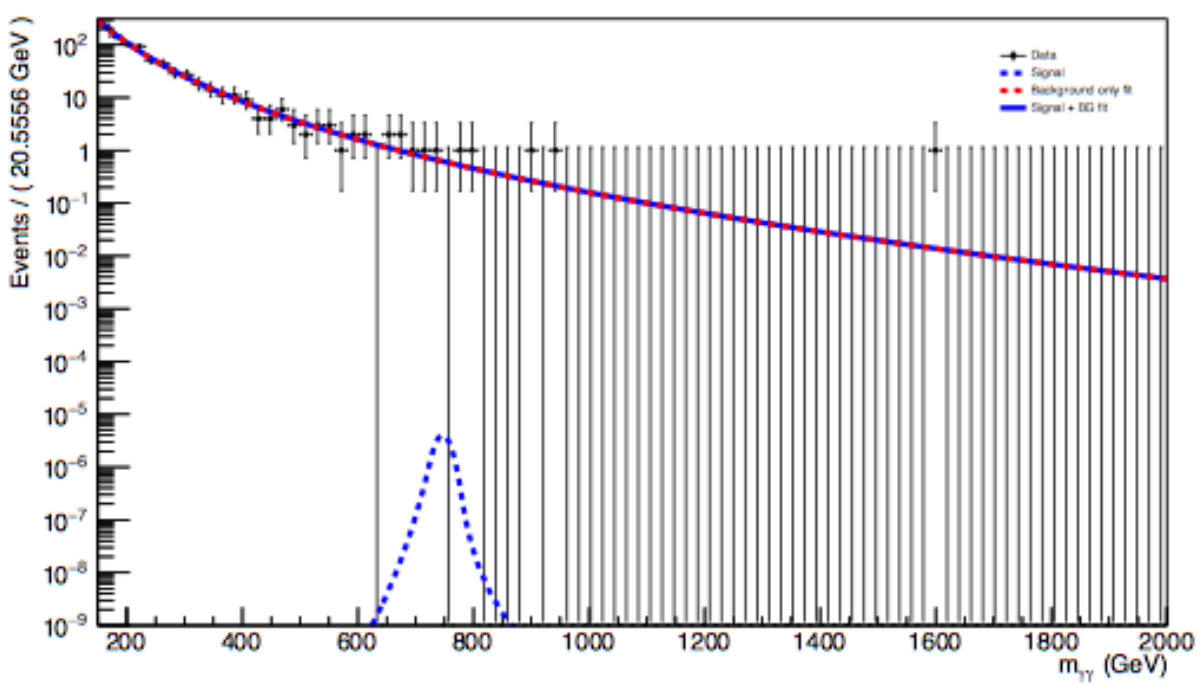
0-jet



1-jet



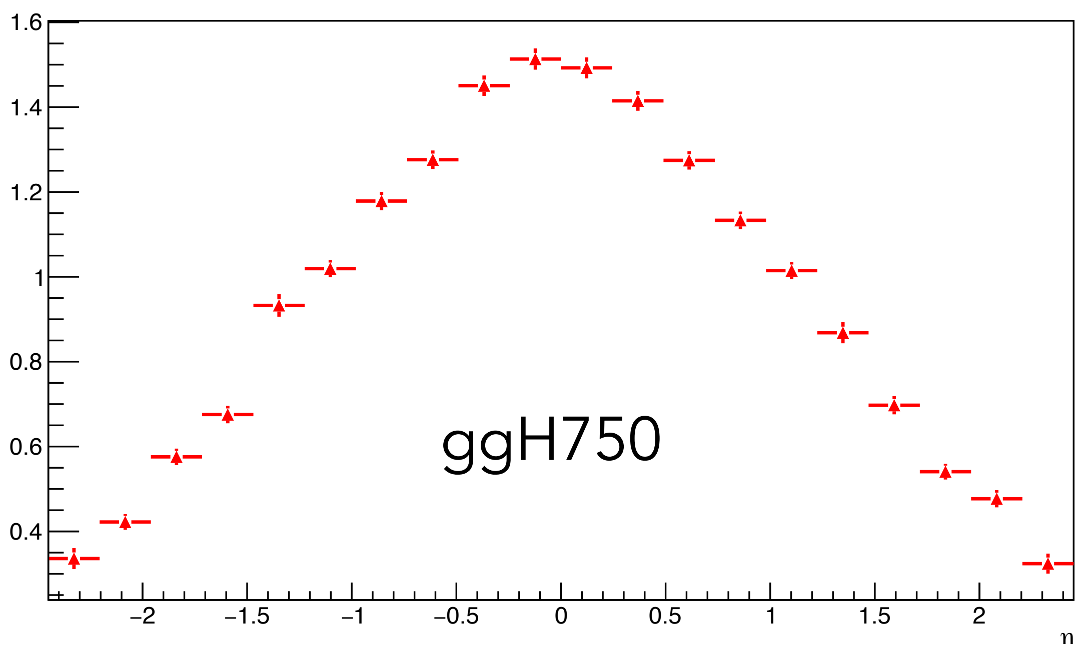
2-jet

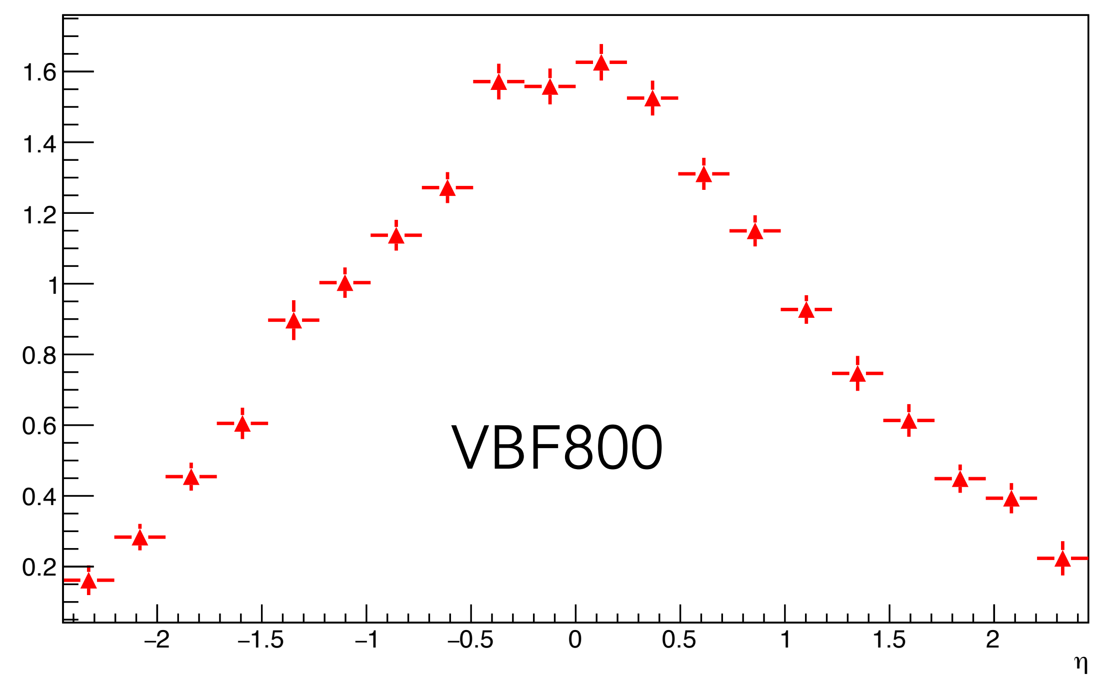


≥3-jet

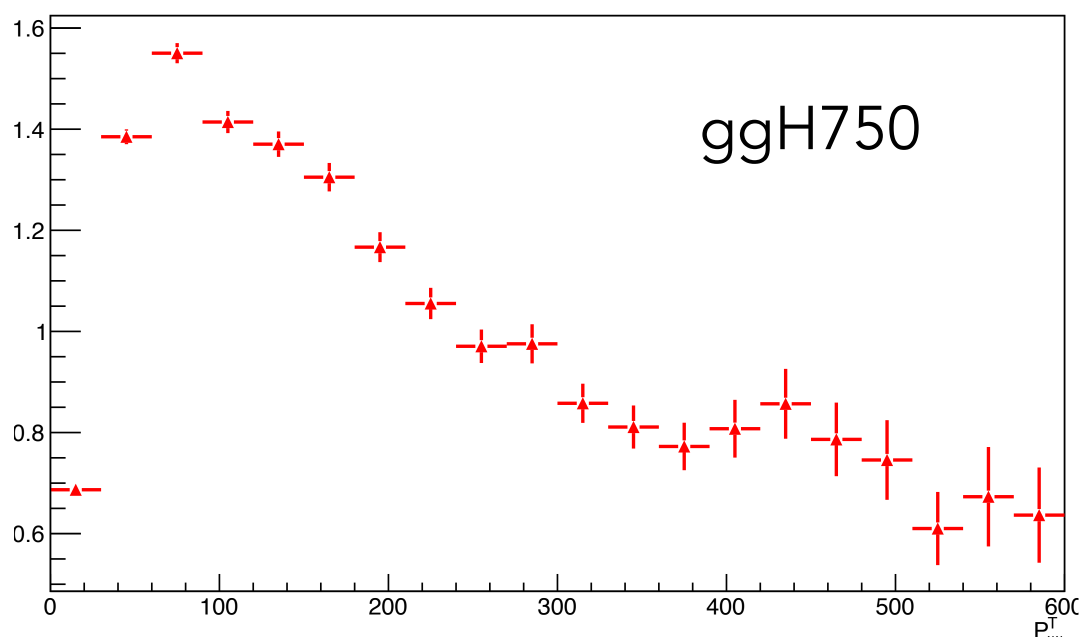
Categories
Other variables

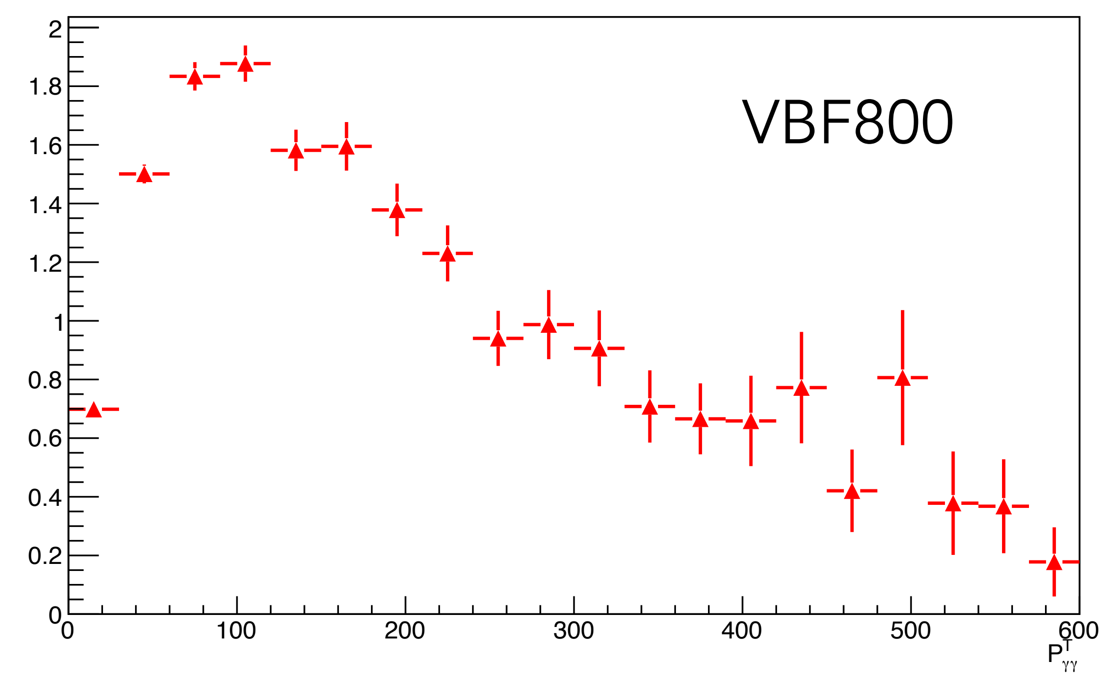
η leading photon





$p^T_{\gamma\gamma}$



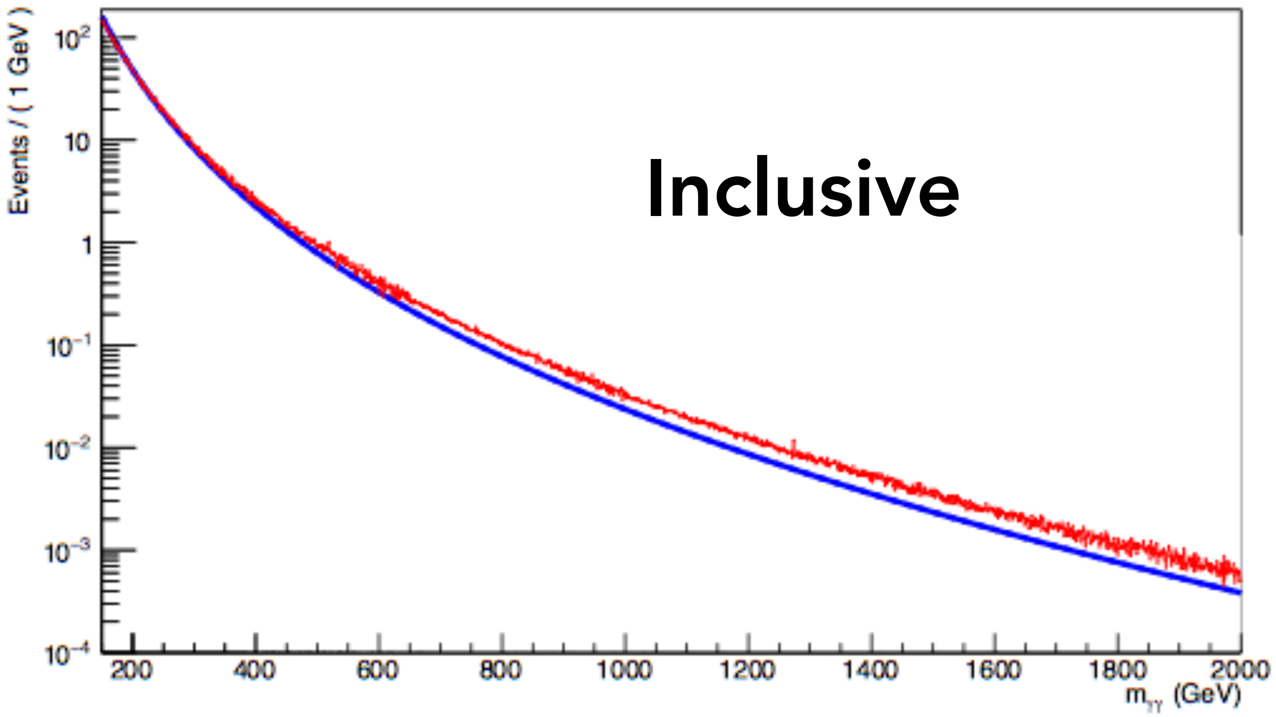


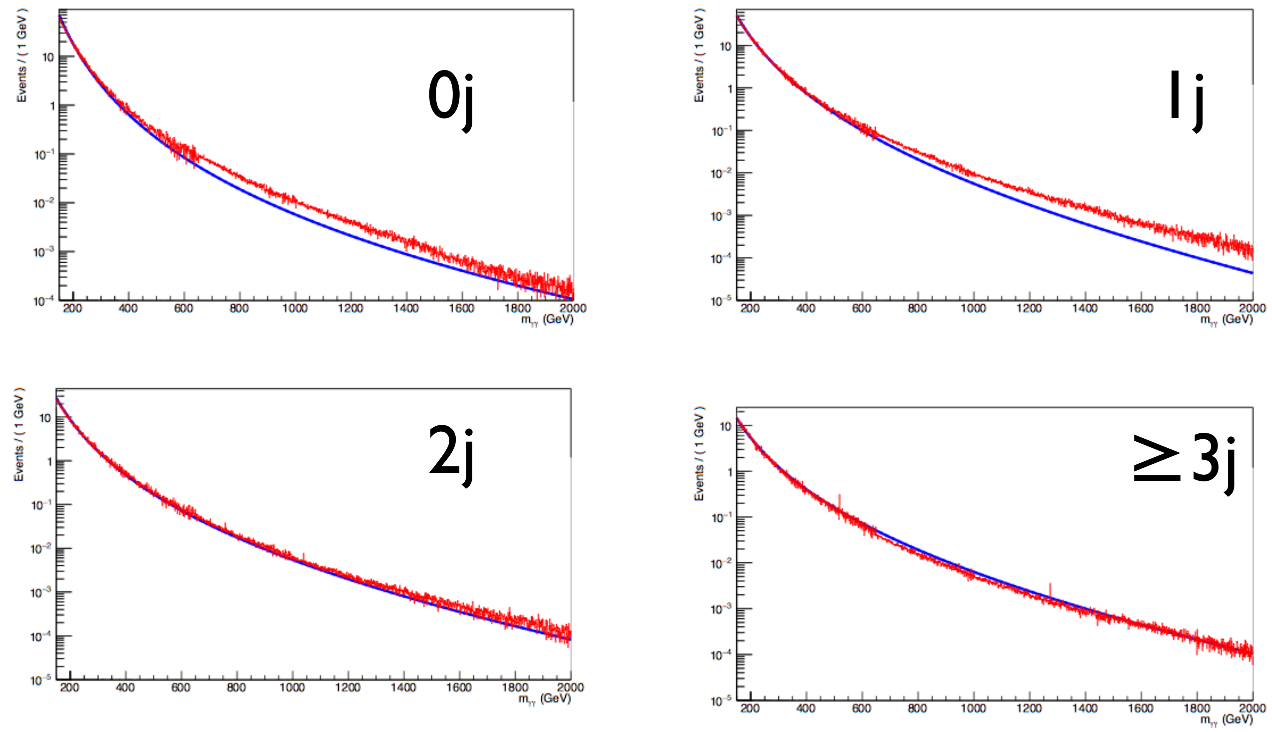
Analysis Overview

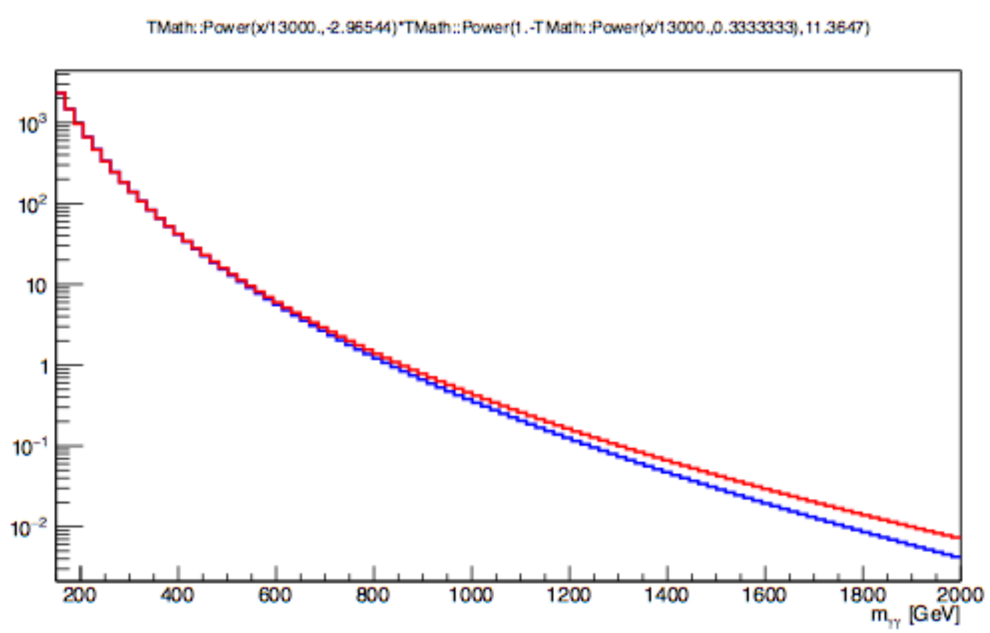
Background - Data/MC comparison

- *right* Comparison between fit on data (without signal region) and MC (Sherpa yy) normalized to nBackground from fit

- *down* BG Fit on data (without signal region) using full range and to 800 GeV

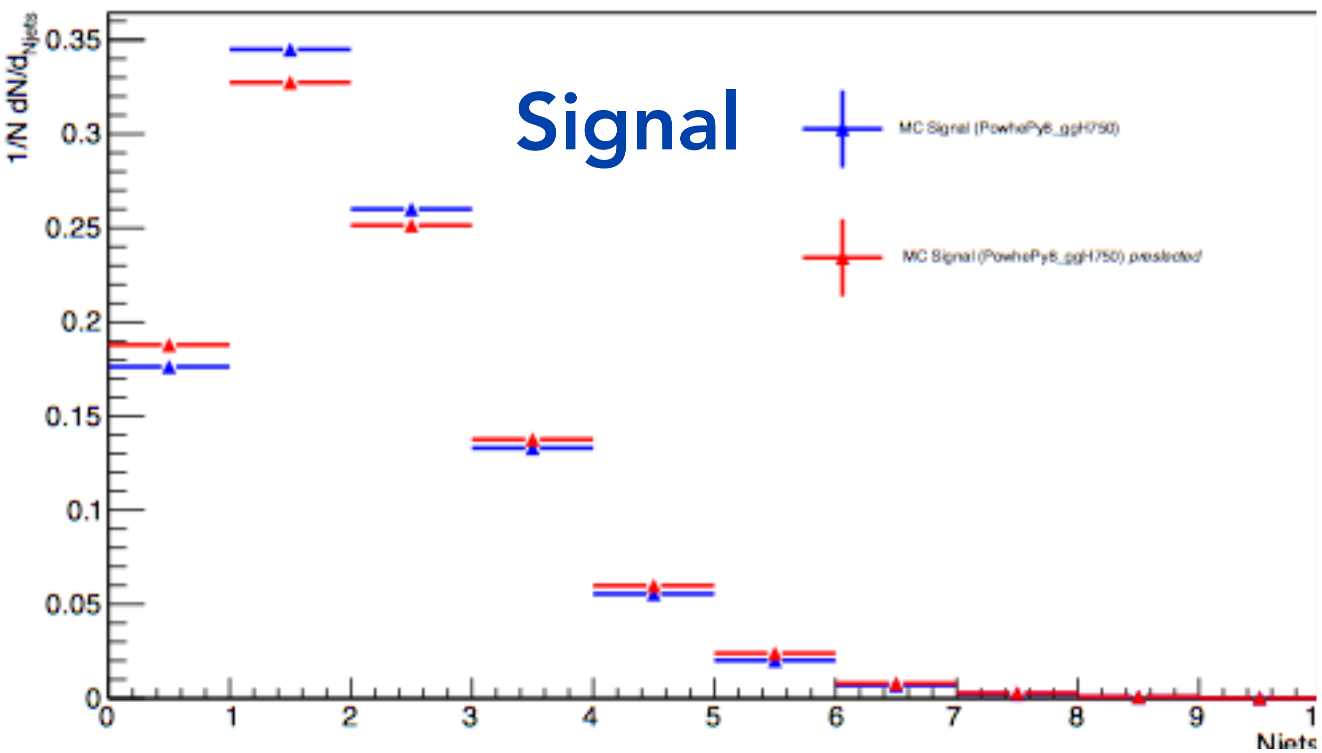


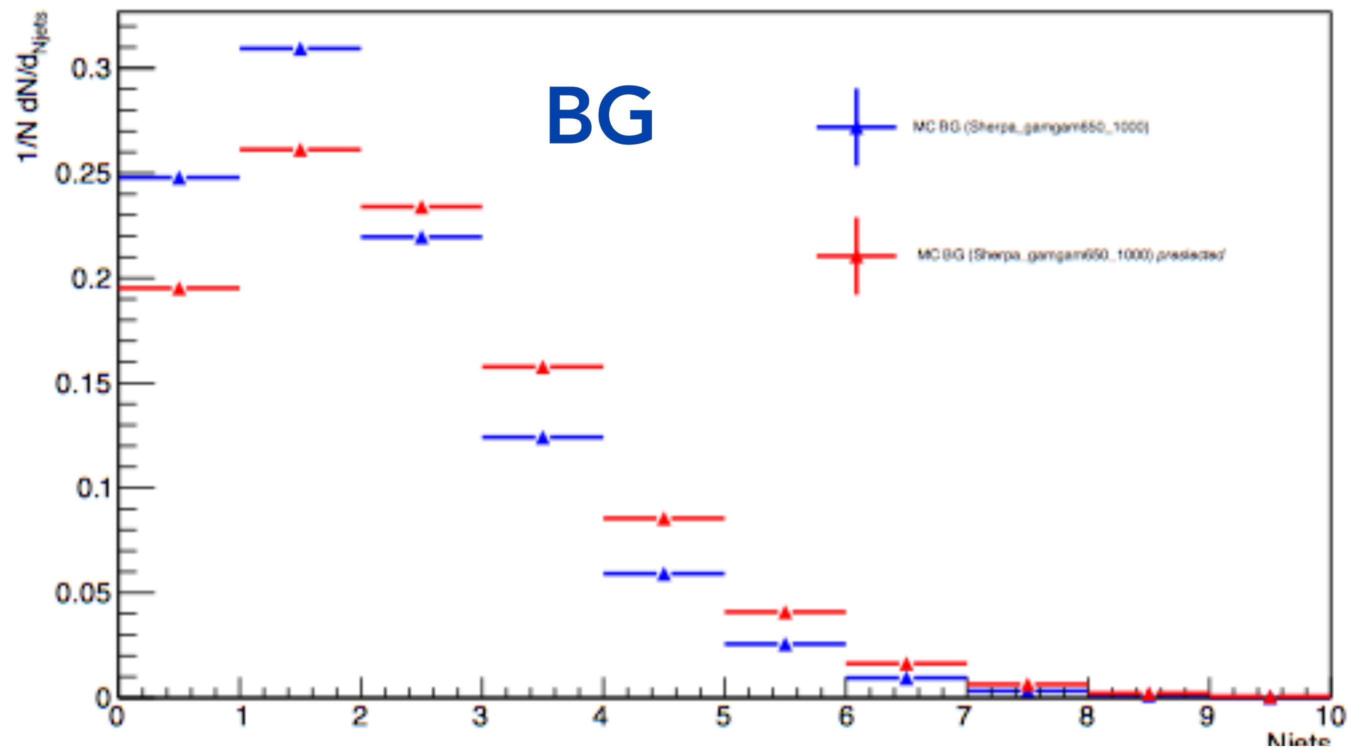


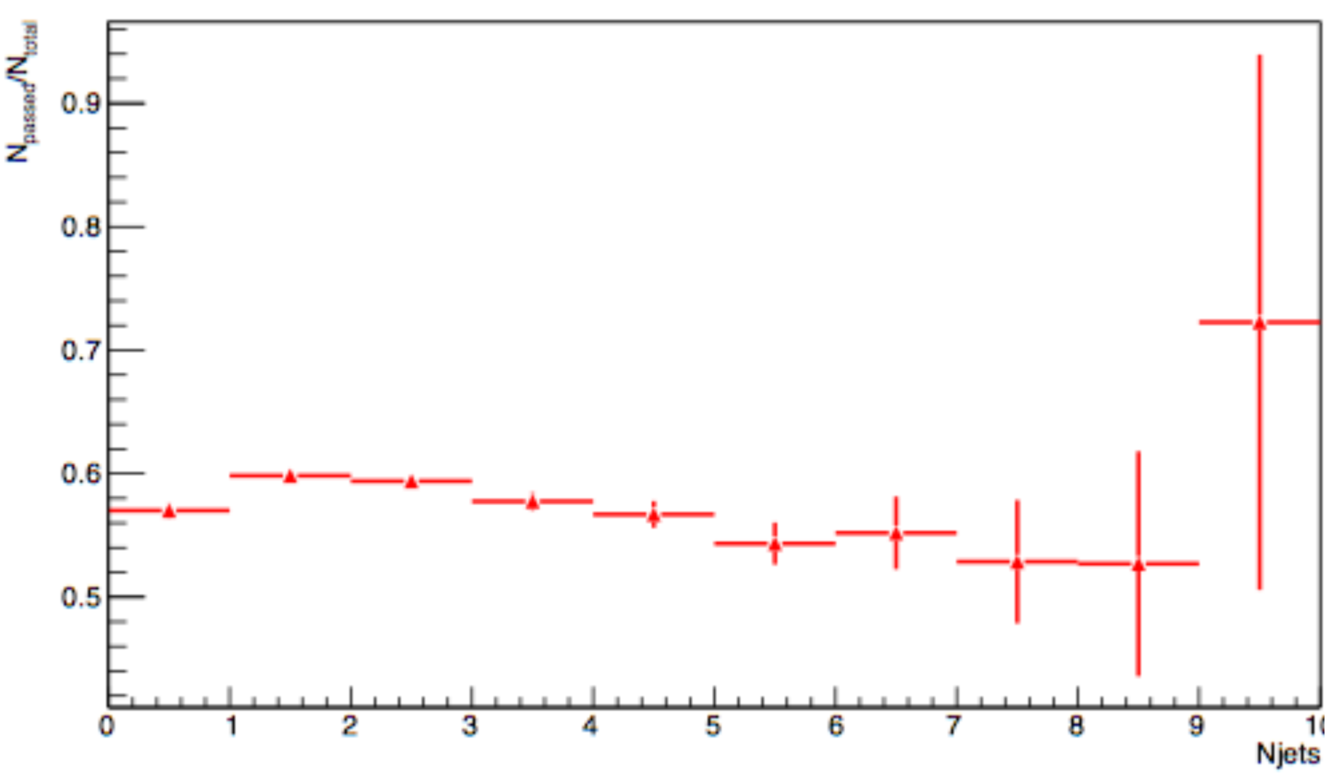


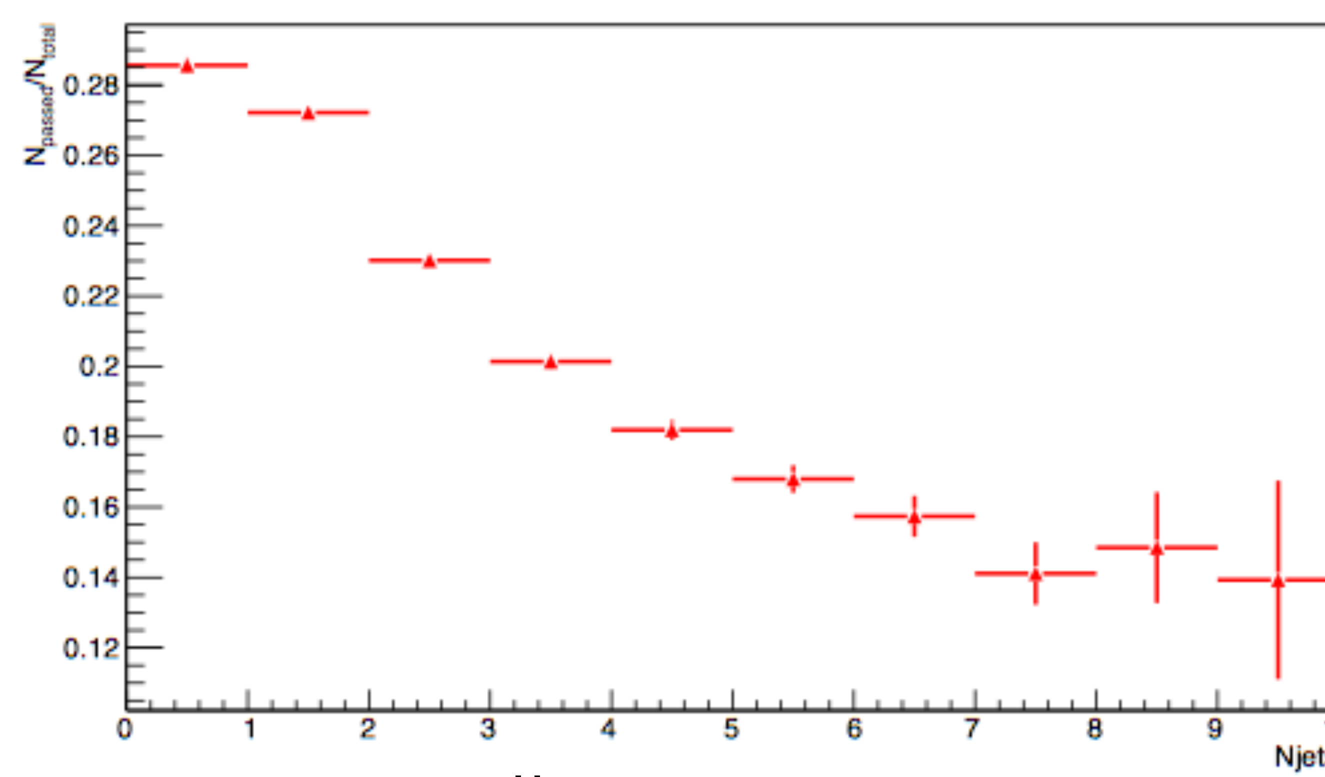
Analysis Overview

Selection effect on N_{jet} cat.









N_{total}: all events in MxAOD

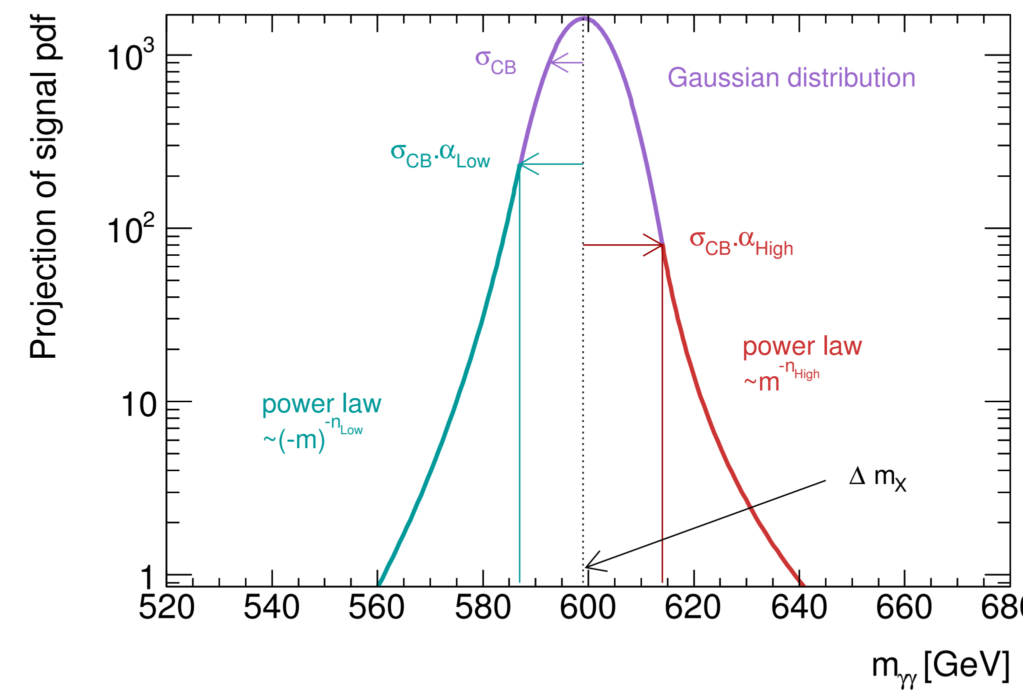
Analysis Overview

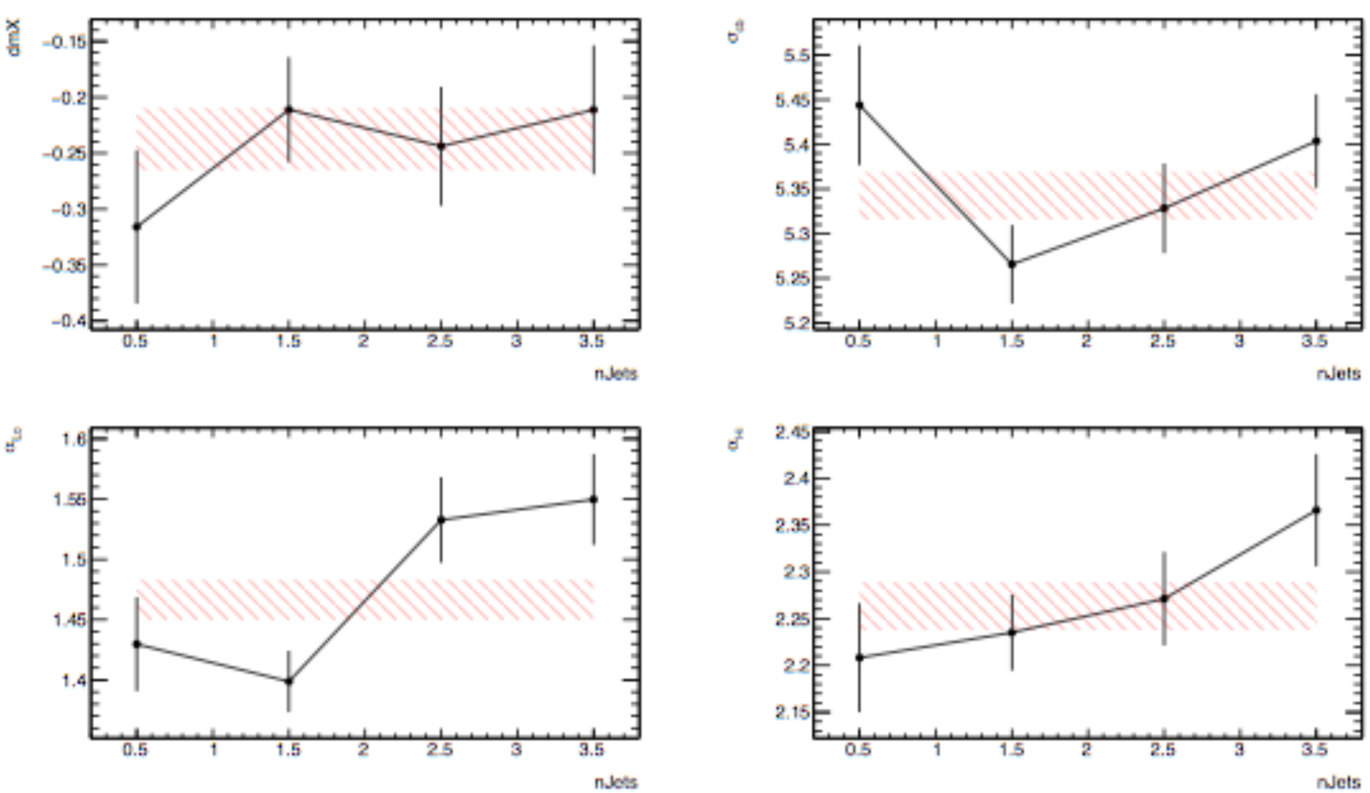
Signal Parametrization (SP)

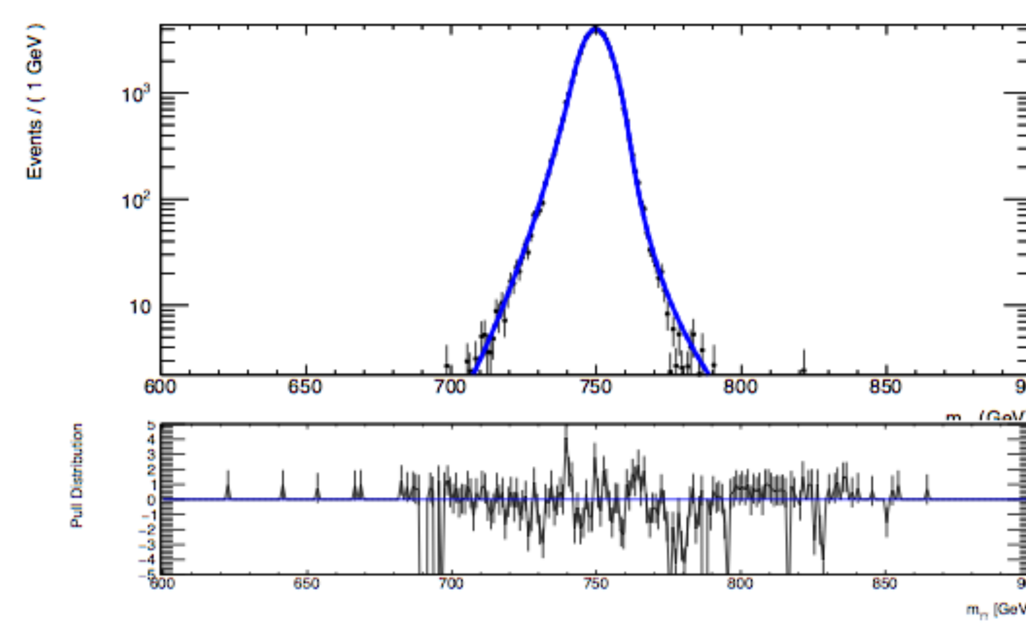
- **Heavy Higgs-like signal** : Theoretical line shape (Narrow width of 4 MeV) dominated by detector response
  - **Modeled for each category** with double sided crystal ball function (Gaussian core extended by power law tail)

No observed change in SP from the inclusive case (expected)

- Signal fraction per category is extracted



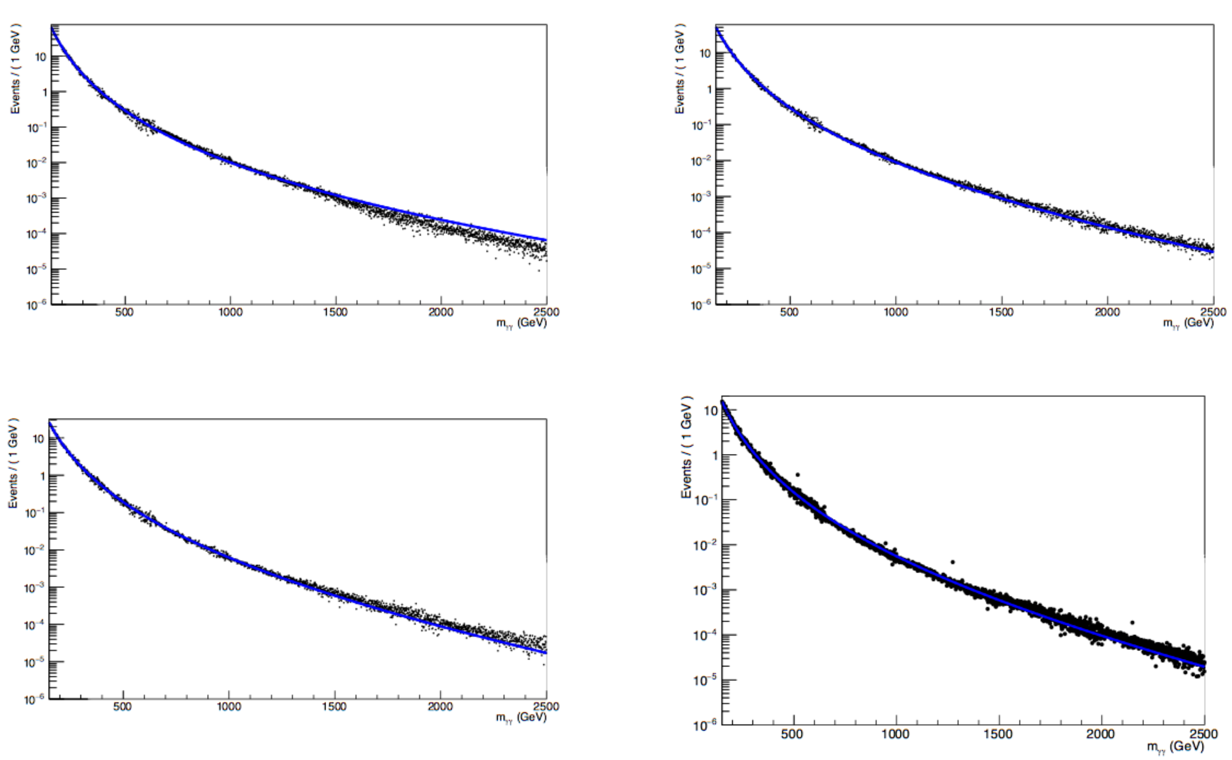


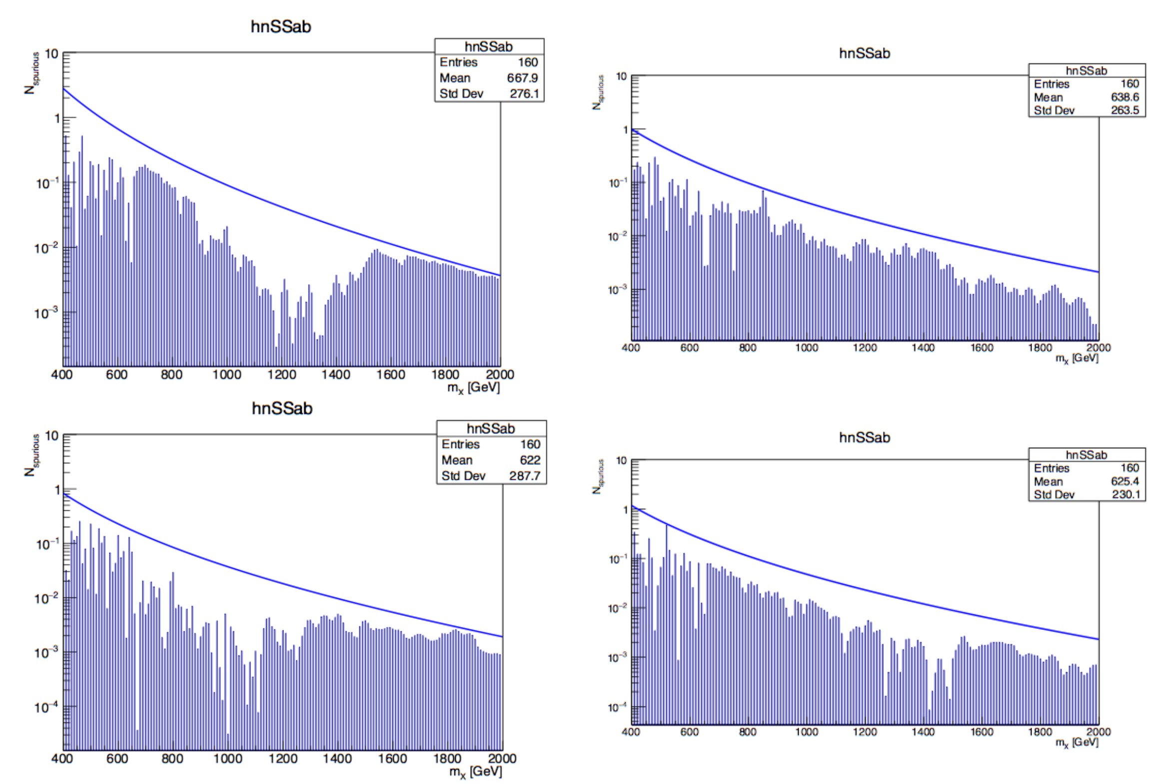


Analysis Overview

Background Modeling

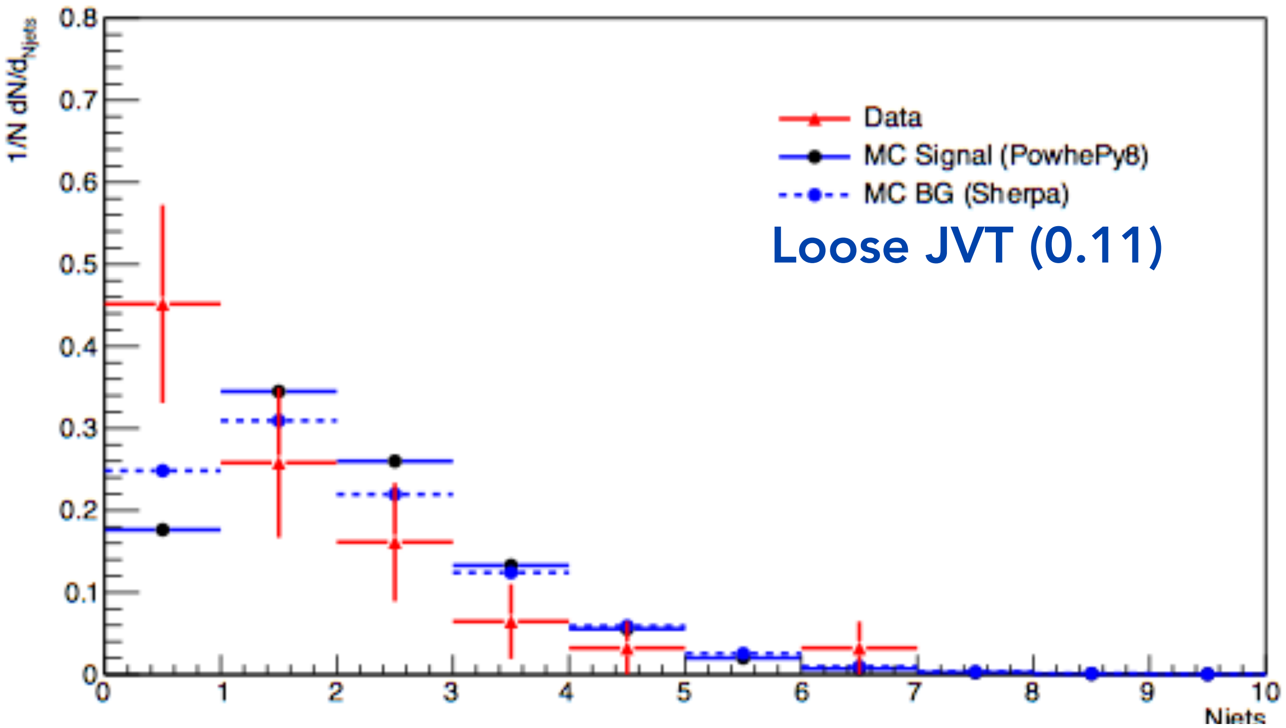
- The continuous BG can be modeled from MC γγ or using data-driven method ranging from [200-2000] GeV
- The BG function is modeled with $x^{-p_2}(1-x^{1/3})^{p_1}; x = m_{\gamma\gamma}/13000$ with all parameters free.
- A dedicated fit for each category is done
- Bias from choice of function estimated from **spurious signal (SS)** test
  - S+B fit on background-only for various m_X signal hypothesis

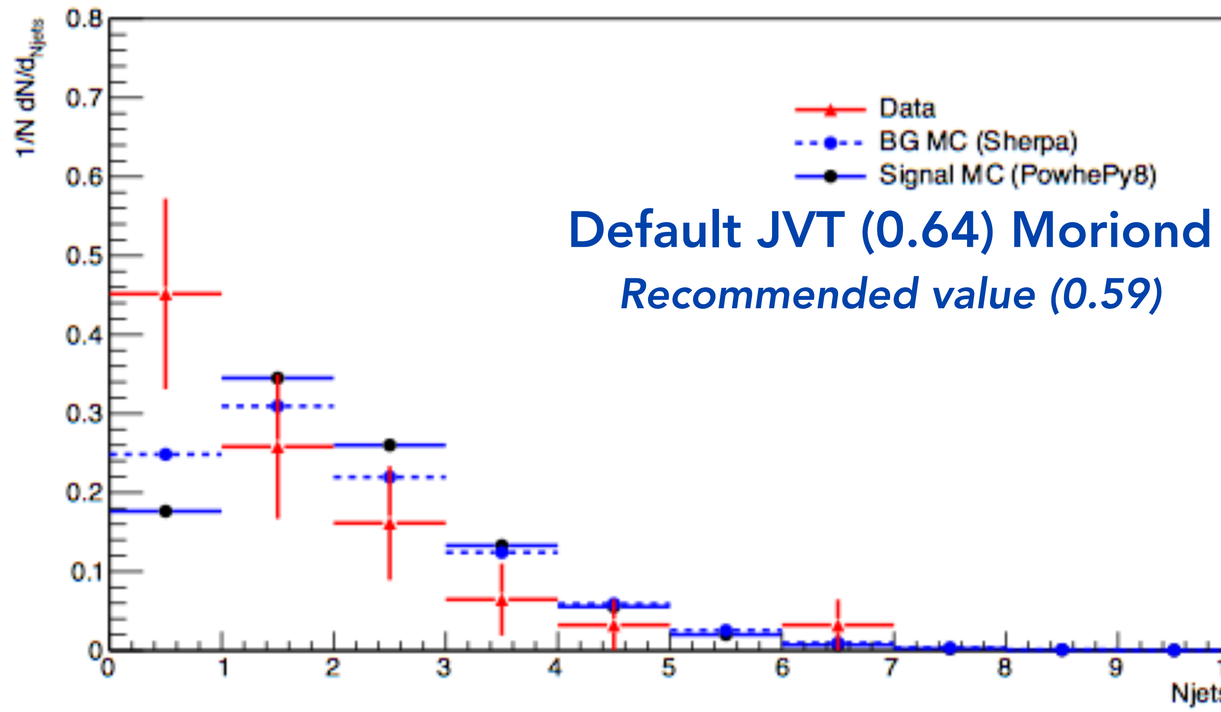


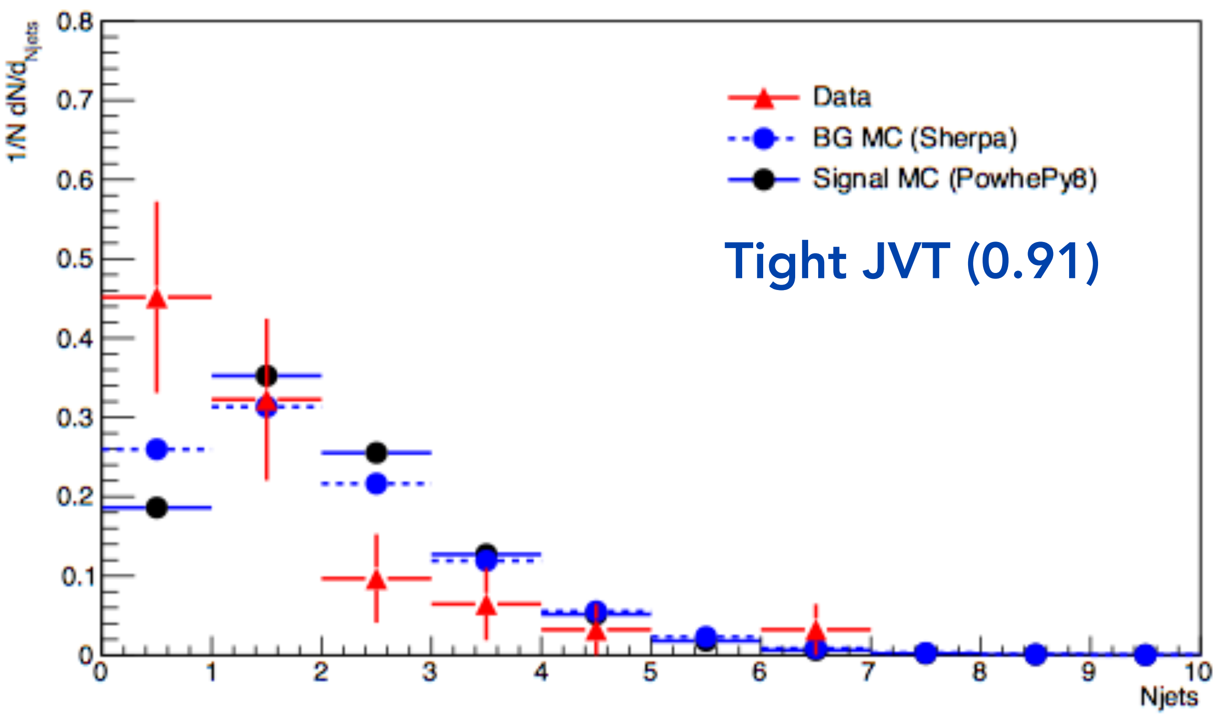


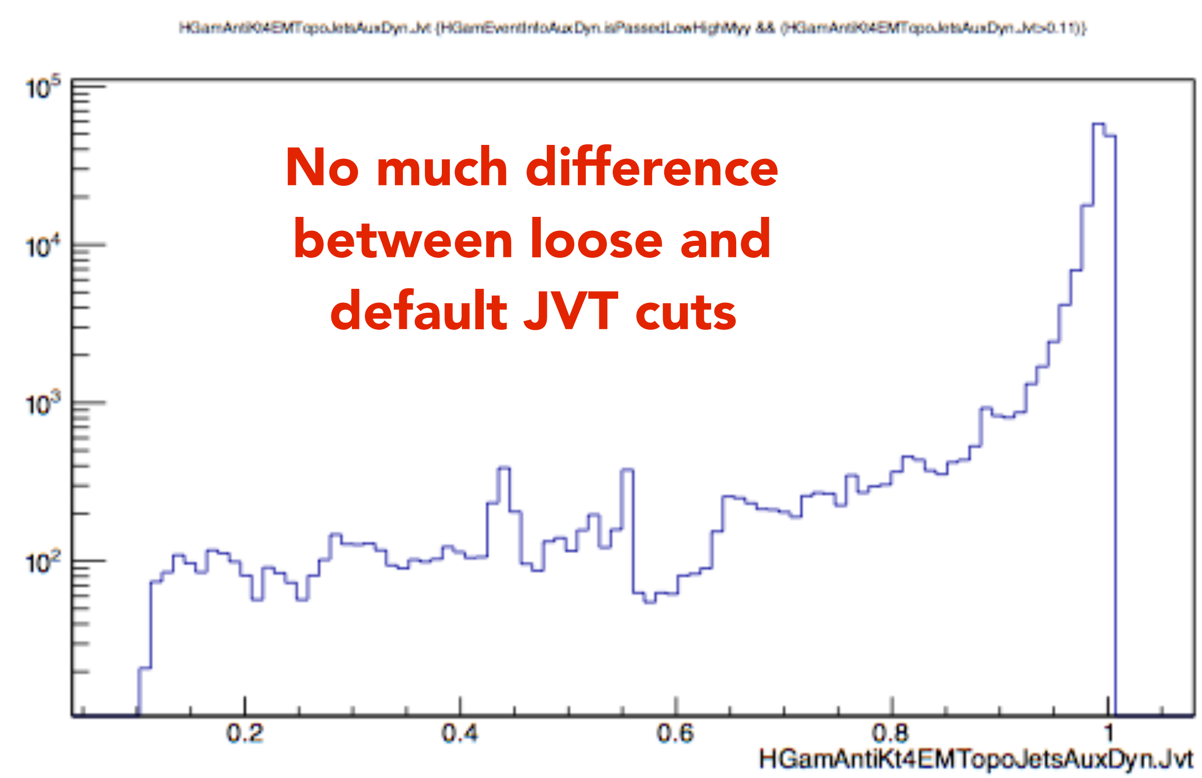
Analysis Overview

Different JVT cuts

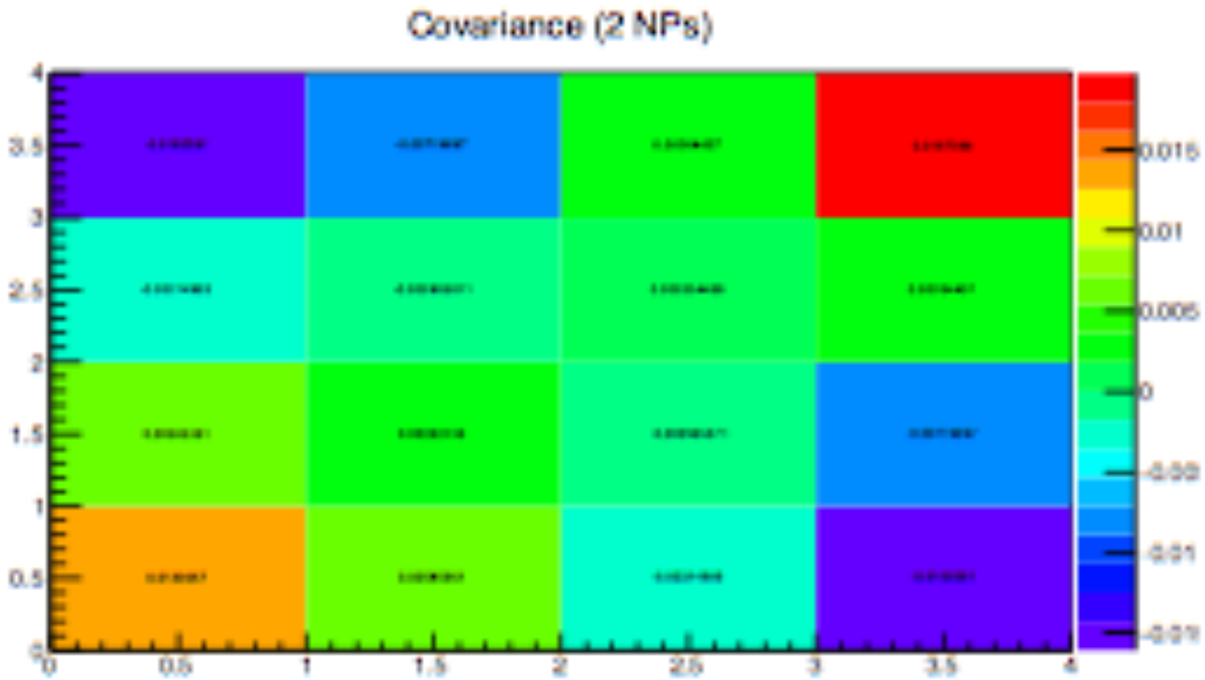


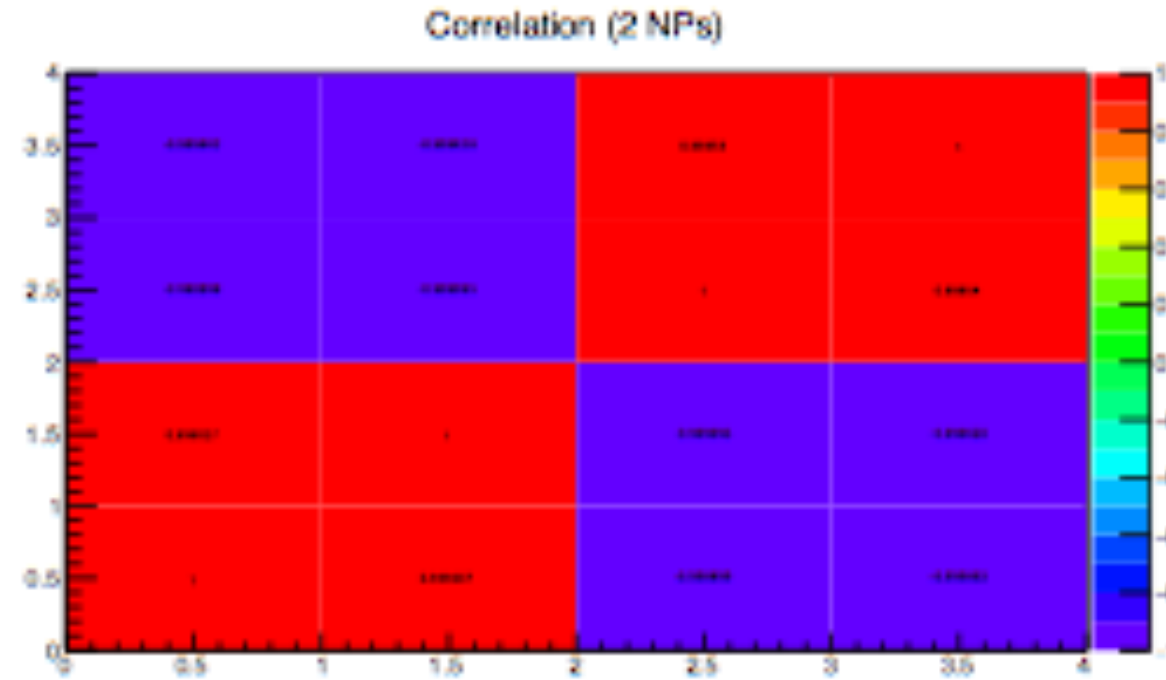


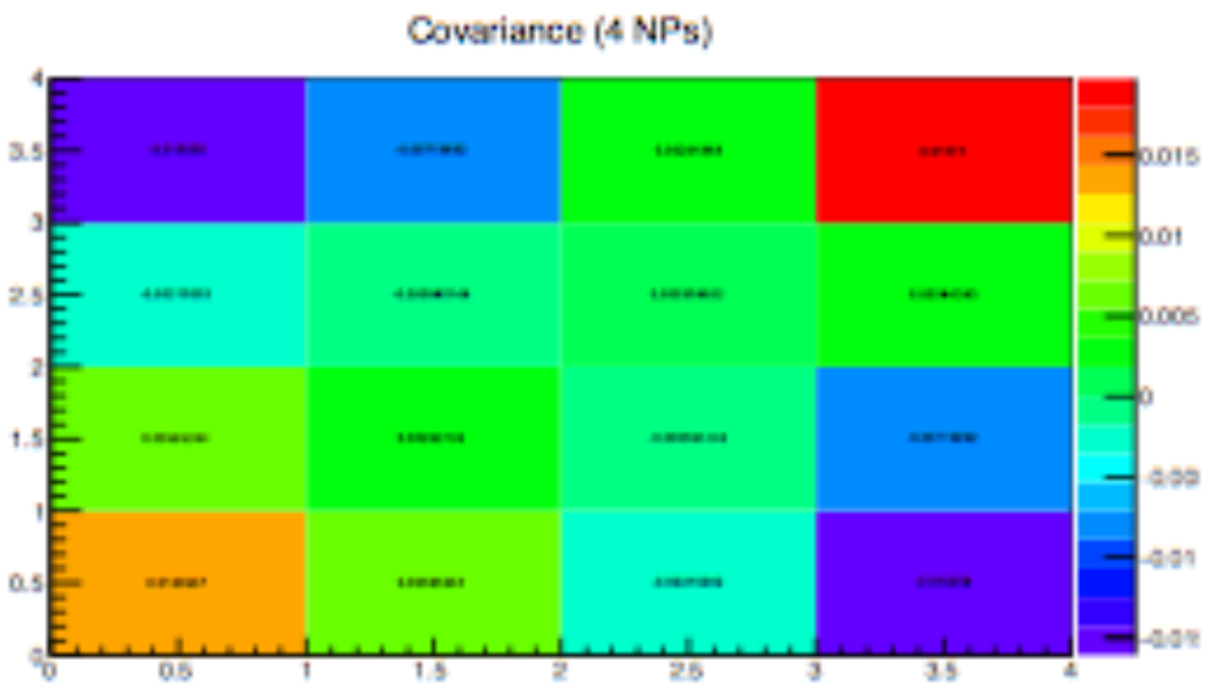


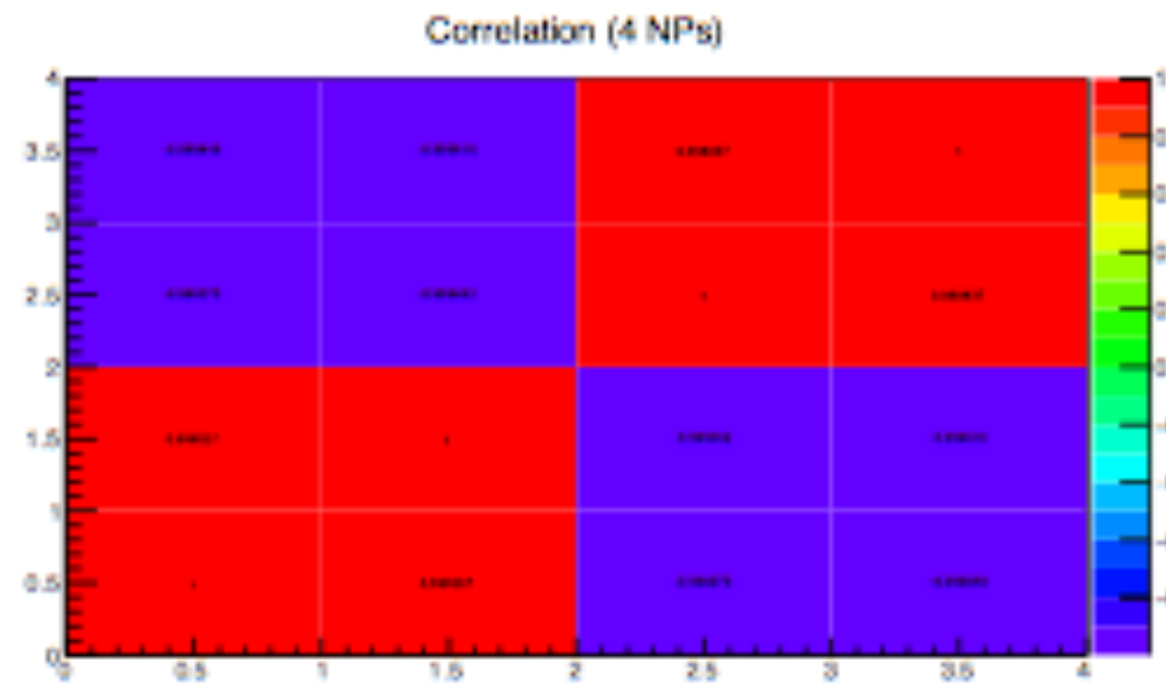


Results



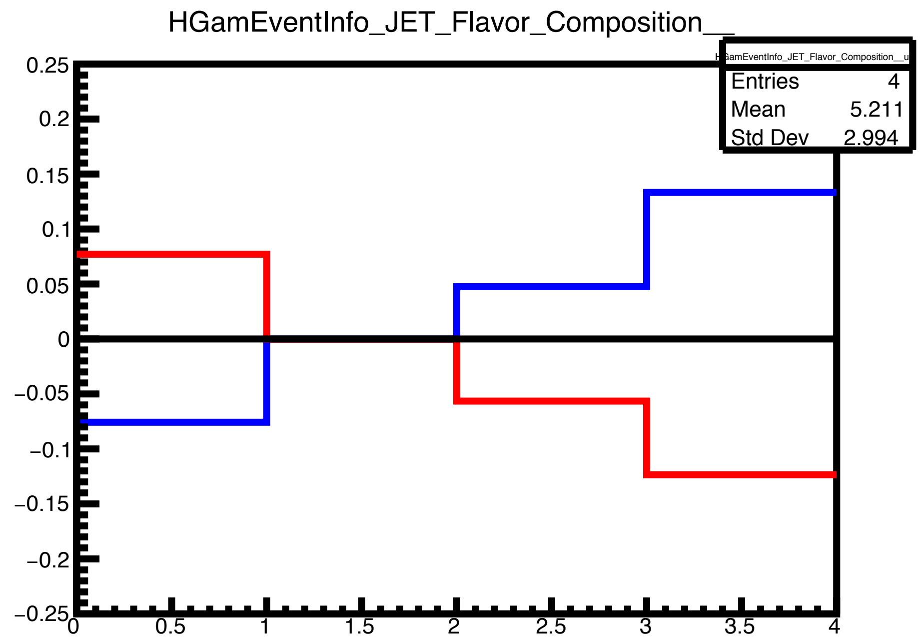


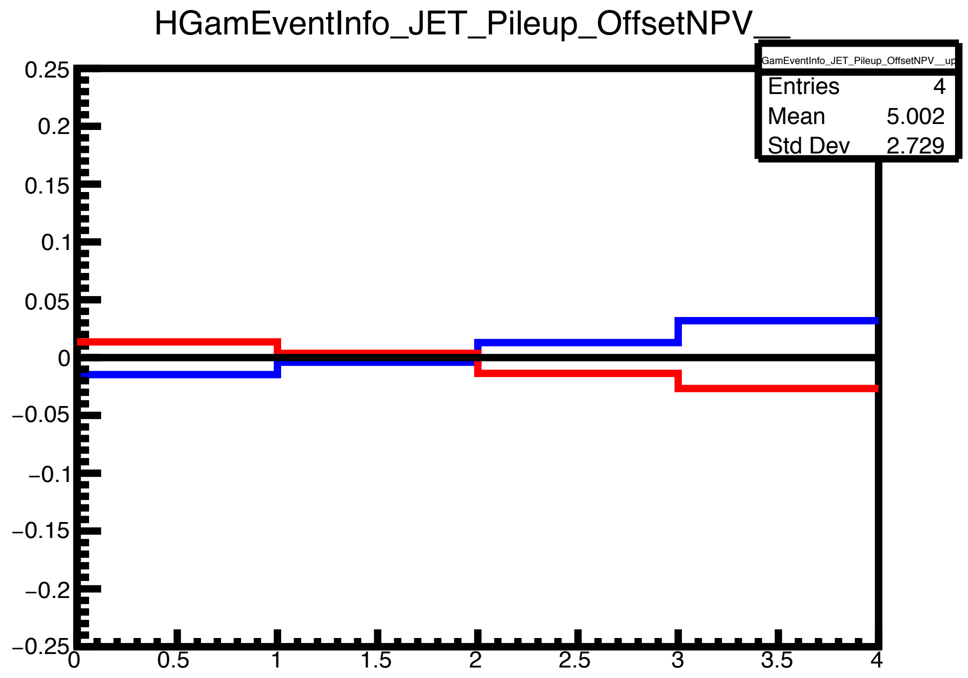


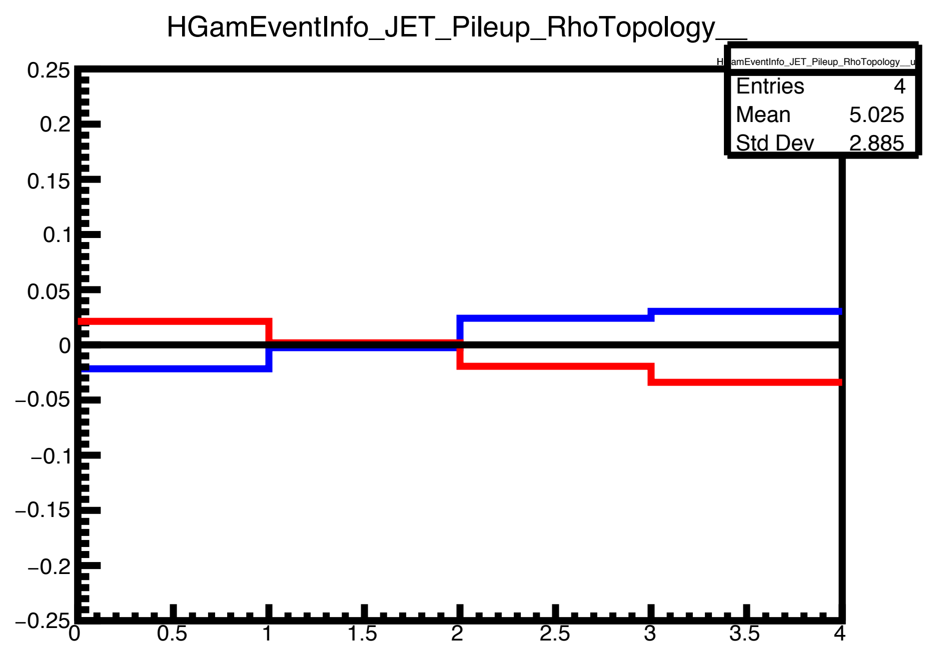


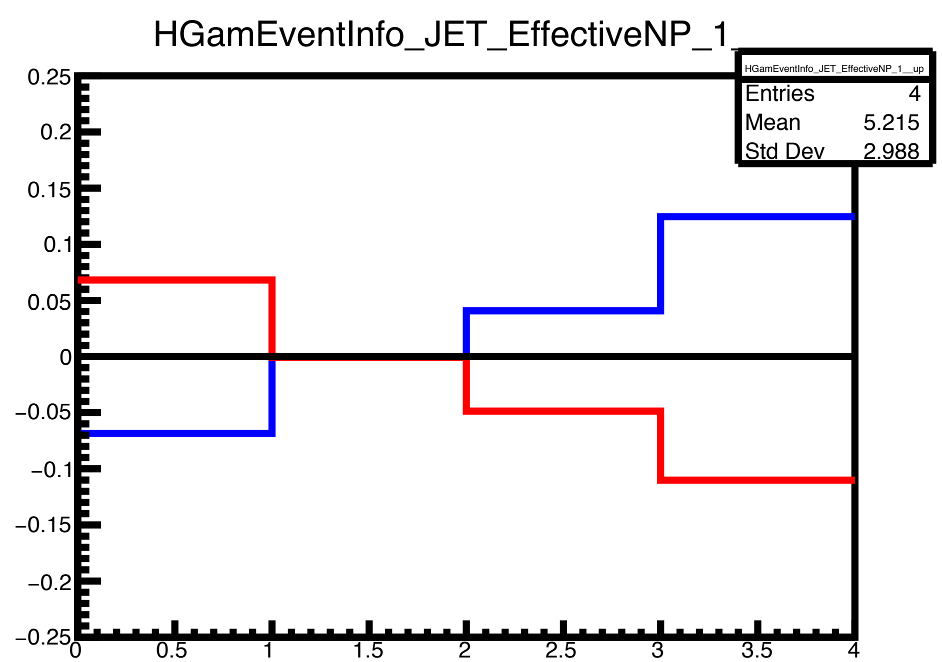
ggH750

Jet NPs *examples*





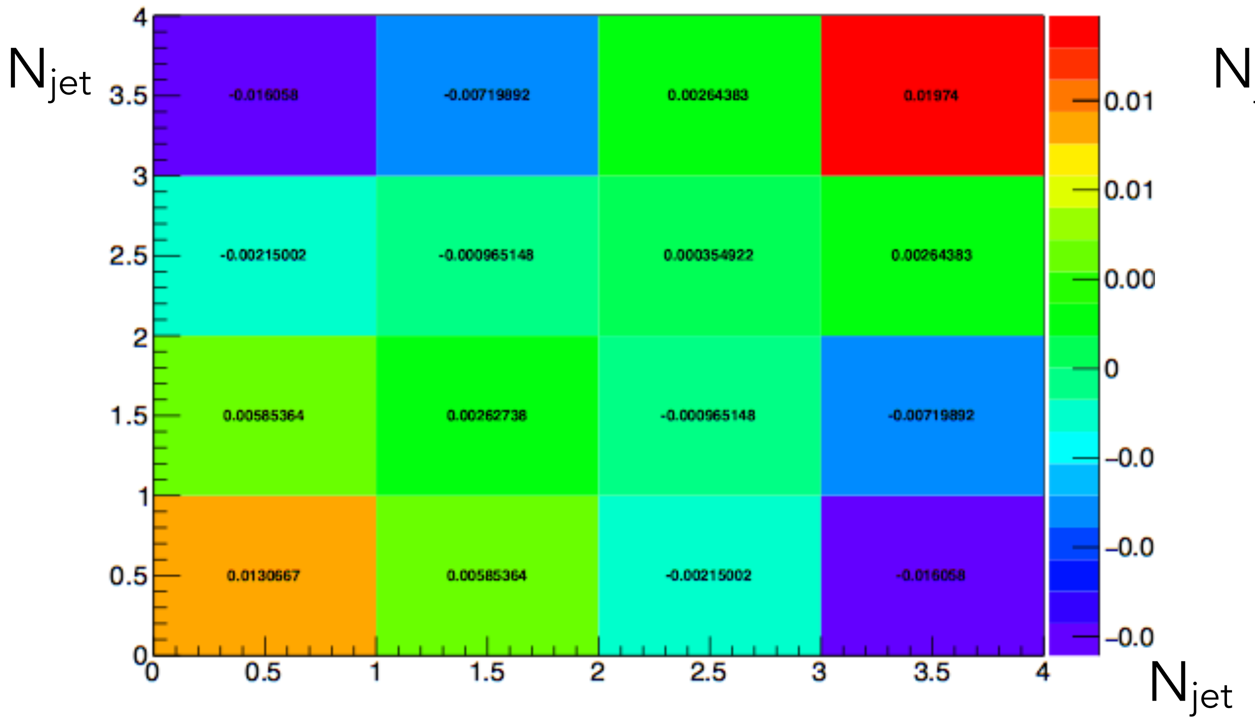


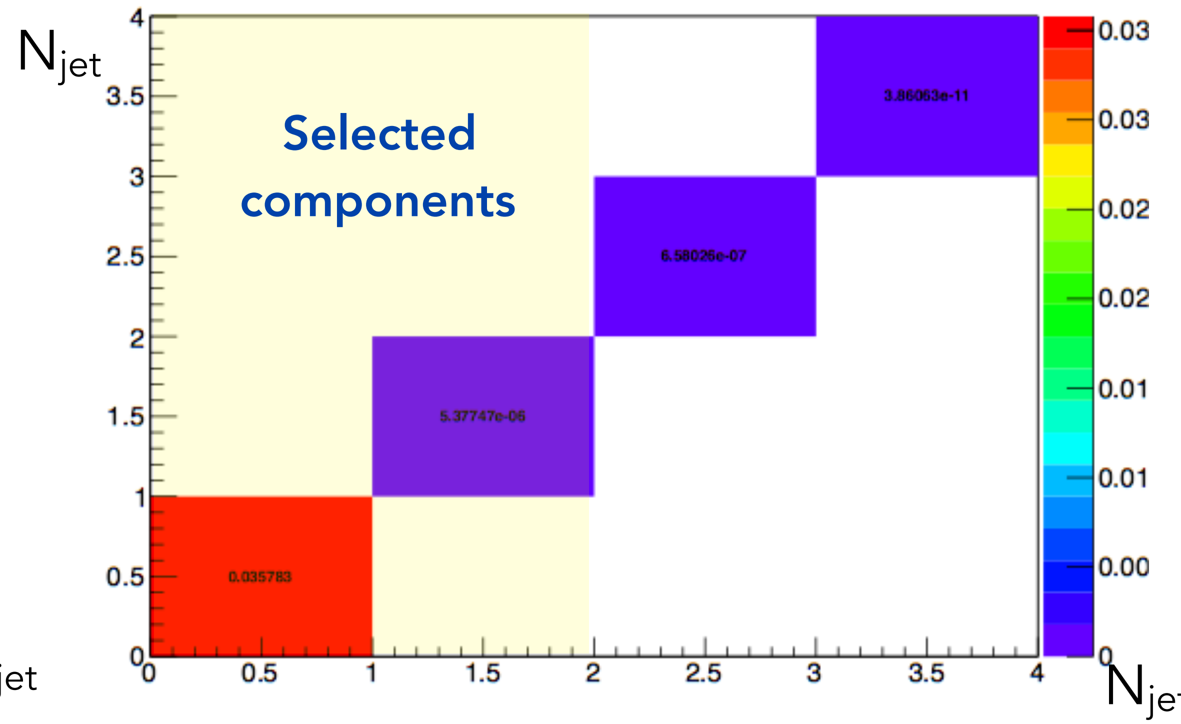


Jet Systematics

- The full set of jet systematic uncertainties can be up to 17 parameter *pile up, quark/gluon jet differences, energy calibration corrections*
- **Reduction of the number of systematic component** is needed to include jet systematics in the S+B fit

$$C^{\text{tot}} = \sum_{k=1}^{N_{\text{sources}}} C^k, \quad C^k_{ij} = s^k_i s^k_j \xrightarrow{\text{diagonalization}} C^{\text{tot}}_{ij} = \sum_{k=1}^{N_{\text{bins}}} \sigma_k^2 \, V^k_i \, V^k_j \xrightarrow{\text{Reduction}} C^{\text{tot}}_{ij} \approx \sum_{k=1}^{N_{\text{eff}}} \sigma_k^2 \, V^k_i \, V^k_j + C'$$





- Comparing correlations before and after reduction shows **no information loss reducing from 17 to 2 systematic component**

(details in backup)

Results

Samples with dedicated SP

	0-jet	1-jet	2-jet	≥ 3 jet	combined	inclusive
ggH BG (MC)	0.93 σ fSignal = 0.17	1.69 σ fSignal = 0.34	1.55 σ fSignal = 0.26	1.37 σ fSignal = 0.21	2.80 σ	2.77 σ
ggH BG (S+B fit on data)	0.99 σ fSignal = 0.17	1.78 σ fSignal = 0.34	1.52 σ fSignal = 0.26	1.27 σ fSignal = 0.21	2.83 σ	2.82 σ
ggH BG (MC)	0.93 σ fSignal = 0.17	2.67 σ fSignal = 0.83			2.73 σ	2.77 σ
ggH BG (S+B fit on data)	1.00 σ fSignal = 0.17	2.67 σ fSignal = 0.83			2.85 σ	2.82 σ
VBF BG (MC)	0.31 σ fSignal = 0.05	1.57 σ fSignal = 0.30	2.28 σ fSignal = 0.45	1.20 σ fSignal = 0.20	3.03 σ	2.77 σ
VBF BG (MC)	0.29 σ fSignal = 0.05	2.92 σ fSignal = 0.95			2.94 σ	2.77 σ

using inclusive SP

	0-jet	1-jet	2-jet	≥ 3 jet	combined	inclusive
γ γ fusion	2.31 σ fSignal = 0.50	1.45 σ fSignal = 0.29	0.92 σ fSignal = 0.14	0.53 σ fSignal = 0.07	2.92 σ	2.77 σ
qq fusion	1.22 σ fSignal = 0.23	1.60 σ fSignal = 0.32	1.44 σ fSignal = 0.24	1.37 σ fSignal = 0.21	2.81 σ	2.77 σ
EFT	1.19 σ fSignal = 0.24	1.65 σ fSignal = 0.34	1.42 σ fSignal = 0.24	1.19 σ fSignal = 0.18	2.73 σ	2.77 σ