Search for pair production of Higgs bosons in the $b\bar{b}b\bar{b}$ final state using proton–proton collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector

A DISSERTATION PRESENTED
BY

Baojia Tong
TO
The Department of Physics

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF
Doctor of Philosophy
IN THE SUBJECT OF
Physics

Harvard University
Cambridge, Massachusetts
May 2017

Thesis advisor: Professor Melissa Franklin Baojia Tong

Search for pair production of Higgs bosons in the $b\bar{b}b\bar{b}$ final state using proton–proton collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector

Abstract

We present a search for Higgs boson pair production, with two $b\bar{b}$ pairs in the final state. This analysis uses the full 2015 and 2016 data collected collected by the ATLAS Collaboration at $\sqrt{s} = 13$ TeV, corresponding to $X \pm y$ fb^{-1} of 2015 and $A \pm B$ fb^{-1} of 2016 pp collision data. The data are interpreted in the context of the bulk Randall-Sundrum warped extra dimension model with a Kaluza-Klein graviton decaying to hh. Relative to the 2015 analysis, this analysis focuses on improvements in the boosted analysis in the highest resonance mass range (between 2000 GeV and 3000 GeV). The data is found to be compatible with the Standard model, and no signs of new physics have been observed.

Contents

Listing of figures

Listing of tables

Everything is meaningless. Buddha.

Acknowledgments

Thanks to everyone in the ATLAS collaboration, who have supported this remarkable program and have contributed to every bit of the result in my thesis. Sir Issac Newton said he was standing on the gaint's showlders, and I am standing on the ATLAS(member)'s showlders. It's eight stories high so I hope it doesn't shrug. Without their work on detector design, commisioning, operational works, reconstruction, data processing, performance studies and recommendations, software support, computing support, and analysis discussions and guidance, I could not have completed this work. I truely and sincerely thank everyone deeply for their contributions.

0
Introduction

Why do we look for $hh \to 4b$?

There are two types of analysis in particle physics. The first one is measurement, which yeilds a observable with an uncertainty. This could either imrpove our knowledge of the Standard Model, or show some inconsistency with the Standard Model. The other type is search, which generally assumes some new physics model and try to justfy in data wether the new model is justified in some observables. A successful search turns the subject into a measurement, yet a null result will set a new limit for a given physics model.

Knowledge knows no bounds.

Creator

1

Motivation for searches beyond the Standard Model

There's something out there that we don't know.

Run.

Random Person

2

Machine of discovery–The Large Hadron Collider

2.1 Design

2.2 Performance

DAMN.

God

3

Eyes of giant–The ATLAS Detector

We love ATLAS.

3.1 Trigger and Data Aquasition

To avoid too high accept rates for certain triggers, the triggers are often prescaled, which means the accepted events get rejected at the prescale. For example, a prescale of two means only every second event passing all trigger conditions gets accepted.

I hate my life.

Tony

4

Game of Jigsaw–Reconstruction

4.1 Introduction to objects

Leptons are rare in proton-proton collisions. Less than 1% of the total tracks are made from leptons, and these are often related to very interesting physics processes, especially invovling the electroweak forces.

Electrons are maximal inoizing in the tracking system. They get absorbed completely in ECAL nd leaves no signature in the HCAL.

Muons are special for leaving minimally ionizing signatures in the detector. They usually penetrate the calorimeter and form tracks in the Muon Spetrometer. This is useful for reducing the generate rate of events, hence for triggering. They are also very clean in reconstruction and have exceelnt resolutions up to 1 TeV. Material and station information, see note.

Jet mass, see note. Jet mass is one of the best tools for distinguishing massive particle decays from QCD background. Jets are trimmed by re-clustering the constituents fo the jet into subjets. k_t algorithm with $R_{sub} = 0.2$ is used, and if the subjets has p_T less than 5% of the original jet p_T, the constituent is removed. The calorimeter-based jet mass (m^{calo}), with clorimeter-cell cluster constituents i with energy E_i, momentum $\vec{p}_i(|\vec{p}_i| = E_i)$ is defined as:

$$m^{calo} = \sqrt{(\sum_i E_i)^2 - (\sum_i p_i)^2} \tag{4.1}$$

For a boosted massive particle, the angular spread in the decay products scales as $\frac{1}{p_T}$. For highly boosted cases, the spread could be comparable with the 0.1.1 calorimeter granularity. Tracking information can be used to maintain performance beyond this. The track-assited jet mass (m^{TA}) is defined as:

$$m^{TA} = \frac{p_T^{calo}}{p_T^{track}} \times m^{track} \tag{4.2}$$

where p_T^{calo} is the calorimeter measurement, p_T^{track} is the four-vector sum of tracks associated to the large-radius calorimetere jet, and m^{track} is the invariant mass of this four-vector sum(the track mass is set to m). This ratio corrects for charged-to-neutral fluctuations, thus improve the resolution with respect to track-only jet mass.

Need to insert a figure here.

Tracks reconstruction, see note. The first step is creating clusters based on pixel and SCT meausred energy deposits, which are space-points. Then three space-points form a seed, and tey are cobmined to build track candidates using a Kalman filter. After ambiguity solving, an artificial neural network is trained and used to identy merged clusters. The last step is a high resolution fit, which is CPU intensive. The min p_T is 400 MeV, and $|\eta| < 2.5$, and at least seven hits in the pixel or SCT. Total number of holes has to be less than two per track, and no more than one in the pixels. $Z_0^{BL} \sin\vartheta$ is requred to be less than 3 mm.

b-tagging. see 2016 run2 note.

The increase of tracks from fragmentation in the high jet p_T region is the main reason for the performance degradation. As the jet p_T increases, the number of fake vertices is increasing, while the secondary vertex reconstruction effieincy for b and c jets slighly decreases with jet p_T.

Maybe. Just maybe.

Who

5

Censor of work–Data Quality

5.1 Data Flow in ATLAS

5.2 Online Monitoring

5.3 Offline Monitoring

Pick the berry.

Bear

6

Event Selection

6.1 Optimization Strategy

To improve the analysis, a quantity which defines the sensitivity of analysis is maximized. This is usually described as $\frac{S}{\sqrt{B}}$, when no knowledge of the model cross section is available, or $\frac{S}{\sqrt{S+B}}$, if the signal cross section is known. These parametrizations have limitations, particularly when the signal yield and the number of estiamted background are both small. A better parametrization for low signal strengh is $\frac{S}{\sqrt{1+B}}$, where the extra $\sqrt{1+B}$ accounts for poisson fluctuations.

For this analysis, two methods are used: one is calculate the number of signal and backgrounds within 68% of the signal m_{hh}mass window, the other is to implement the full signal and background predicions after smoothing and compare the asymptotic expected exclusion limits. Both methods yield comparable results.

To avoid bias during the optimzation process, data's signal regions are blinded.

6.2 *b* tagging

b-tagging, which is the identification of the *b*hadron, [1] is the core and main limiting factor of this analysis. Because of the relatively long lifetime, it is possible to tag the *b*hadron using the inner detector informations. A higher *b*-tagging efficiency will increase the signal selection efficiency, while a lower *b*-tagging fake rate will reduce the background like $gg \to c\bar{c}$ in the signal regions.

6.3 Signal Efficiency

Acceptance referes to purely geometric fiducial volume of the detector. Efficiency refers to purely detector effectivenss in finding objects. The final value in the study is Acceptance $\times$ Efficiency, where both effects are considered.

Tough part.

Revolutionaries.

7

Background Estimation

7.1 Background composition

Backgrounds of this analysis are mostly QCD multijet events and $t\bar{t}$events. $t\bar{t}$cross section increased by a factor of 3.4.

7.2 Sideband Region Distributions

7.3 Control Region Distributions

This section shows comparisons of data with the prediction of QCD multi-jets and $t\bar{t}$in the control region (CR), which is identical to the signal region (SR) except the large-R jets are required to have masses close but not too close to the Higgs mass. The definition can be seen in Section ??. The predicted and observed event yields are summarized in Tables ?? and ??.

Figures 7.1, 7.2, 7.3, and 7.4 show predictions of various kinematics of the large-R jets and their associated track jets in the 4b selection. The shapes and normalization are a feature of the prediction, where the normalization is derived in the SB. The quality of the prediction is generally good, and no clear systematic biases are observed.

Figures 7.5, 7.6, 7.7, and 7.8 show predictions of various kinematics of the large-R jets and their associated track jets in the 3b selection. The shapes and normalization are a feature of the prediction, where the normalization is derived in the SB. The quality of the prediction is generally good, and no clear systematic biases are observed.

Figures 7.9, 7.10, 7.11, and 7.12 show predictions of various kinematics of the large-R jets and their associated track jets in the 2bs selection. The shapes and normalization are a feature of the prediction, where the normalization is derived in the SB. The quality of the prediction is generally good, and no clear systematic biases are observed.

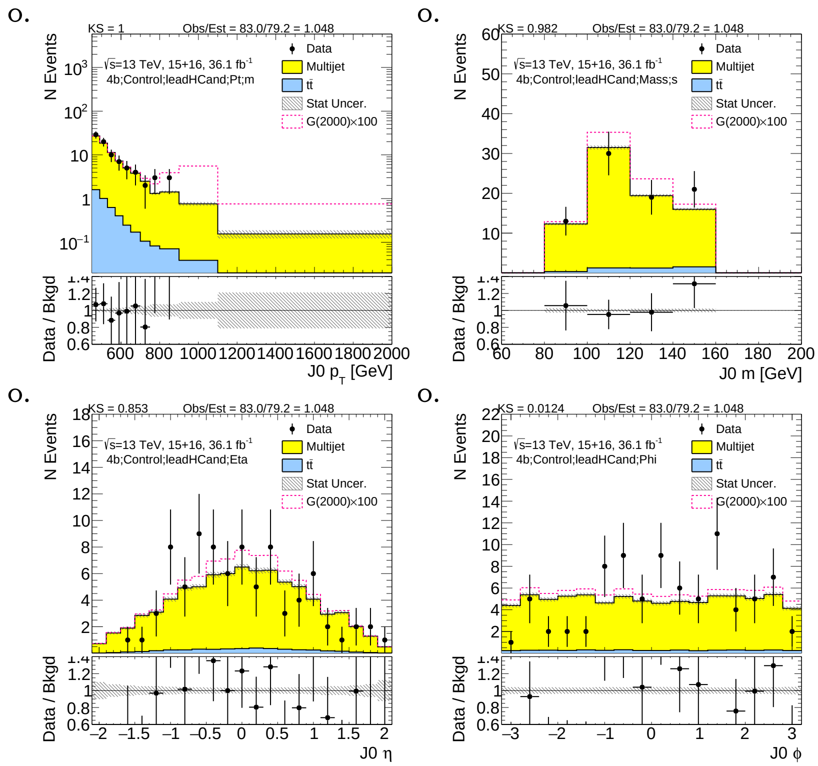

Figure 7.1: Kinematics of the lead large-R jet in data and prediction in the sideband region after requiring 4 b-tags. The normalization agrees by construction, and the shapes are a feature of the prediction.

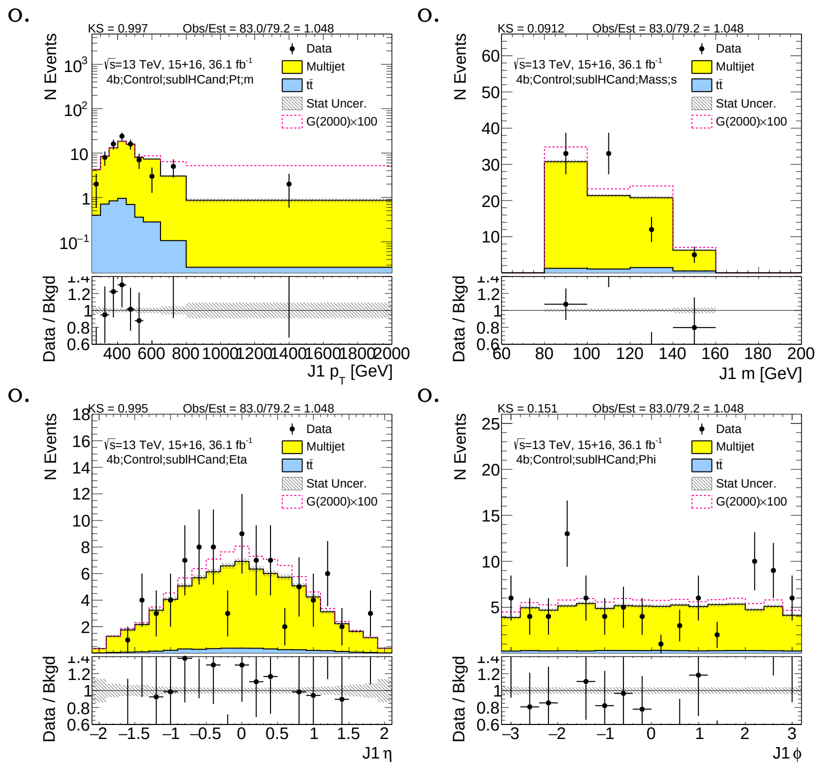

Figure 7.2: Kinematics of the sub-lead large-R jet in data and prediction in the sideband region after requiring 4 b-tags. The normalization agrees by construction, and the shapes are a feature of the prediction.

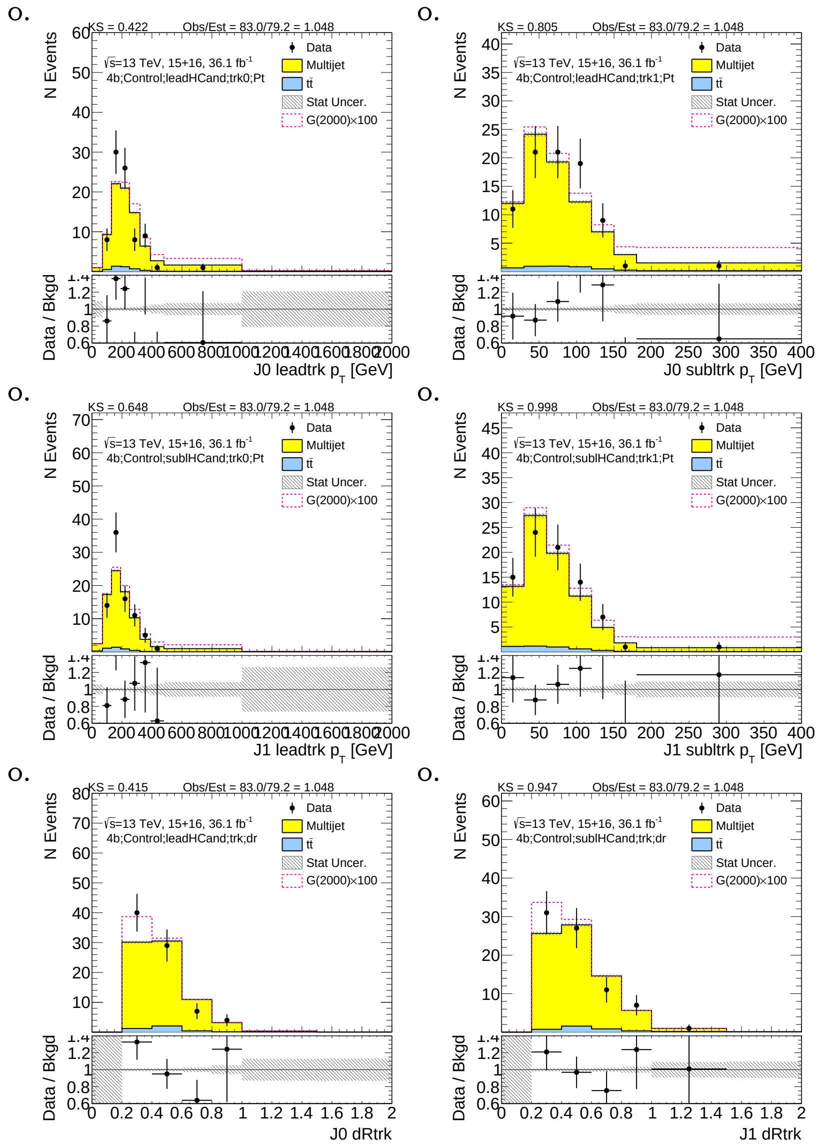

Figure 7.3: First two rows show the kinematics of the lead (left) and sub-lead (right) small-R track jets associated to the lead (first-row) and sub-lead (second-row) large-R jet in data and prediction in the sideband region after requiring 4 b-tags. Third row shows the ΔR between two leading small-R track-jets associated to the leading (left) and sub-leading (right) large-R jet. The normalization agrees by construction, and the shapes are a feature of the prediction.

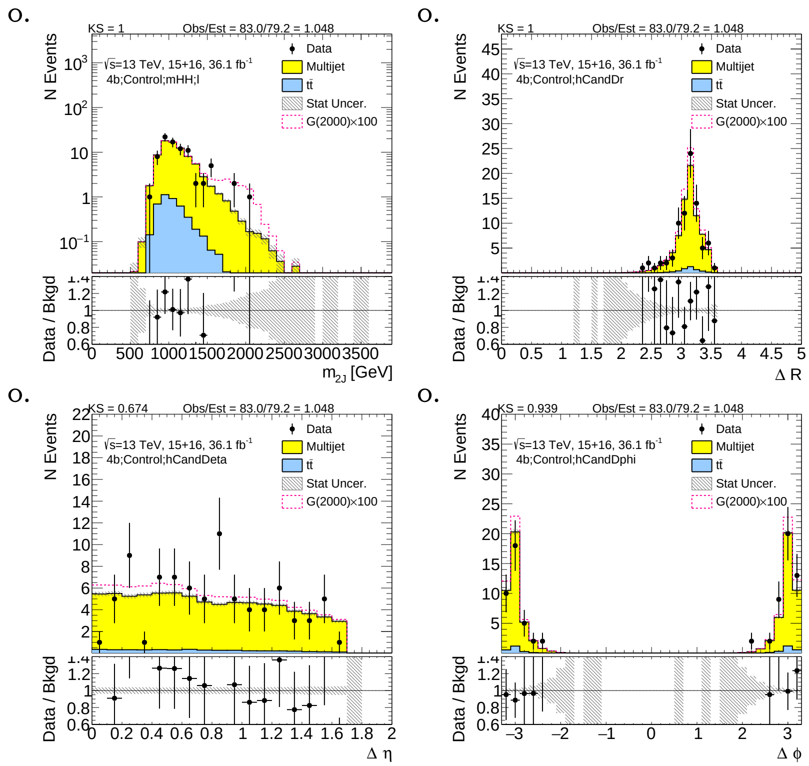

Figure 7.4: Kinematics of the large-R jet system in data and prediction in the sideband region after requiring 4 b-tags. The normalization agrees by construction, and the shapes are a feature of the prediction.

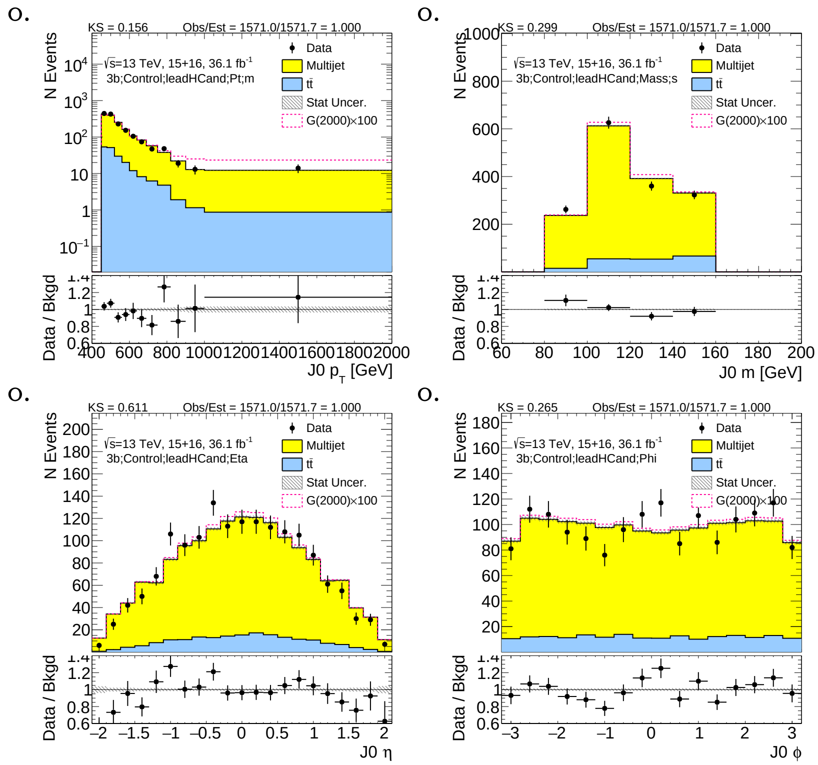

Figure 7.5: Kinematics of the lead large-R jet in data and prediction in the sideband region after requiring 3 b-tags. The normalization agrees by construction, and the shapes are a feature of the prediction.

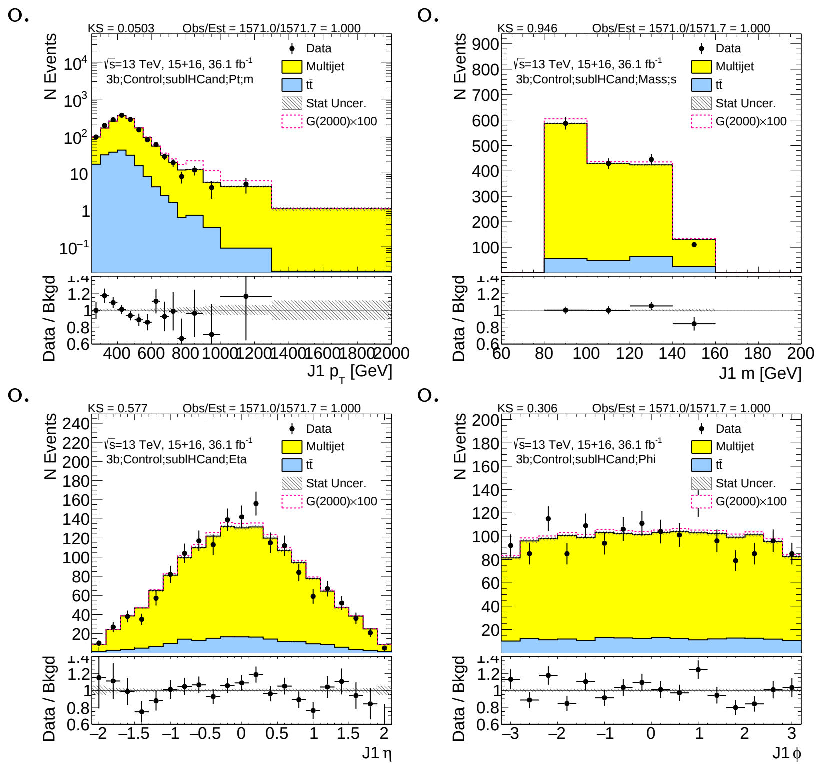

Figure 7.6: Kinematics of the sub-lead large-R jet in data and prediction in the sideband region after requiring 3 b-tags. The normalization agrees by construction, and the shapes are a feature of the prediction.

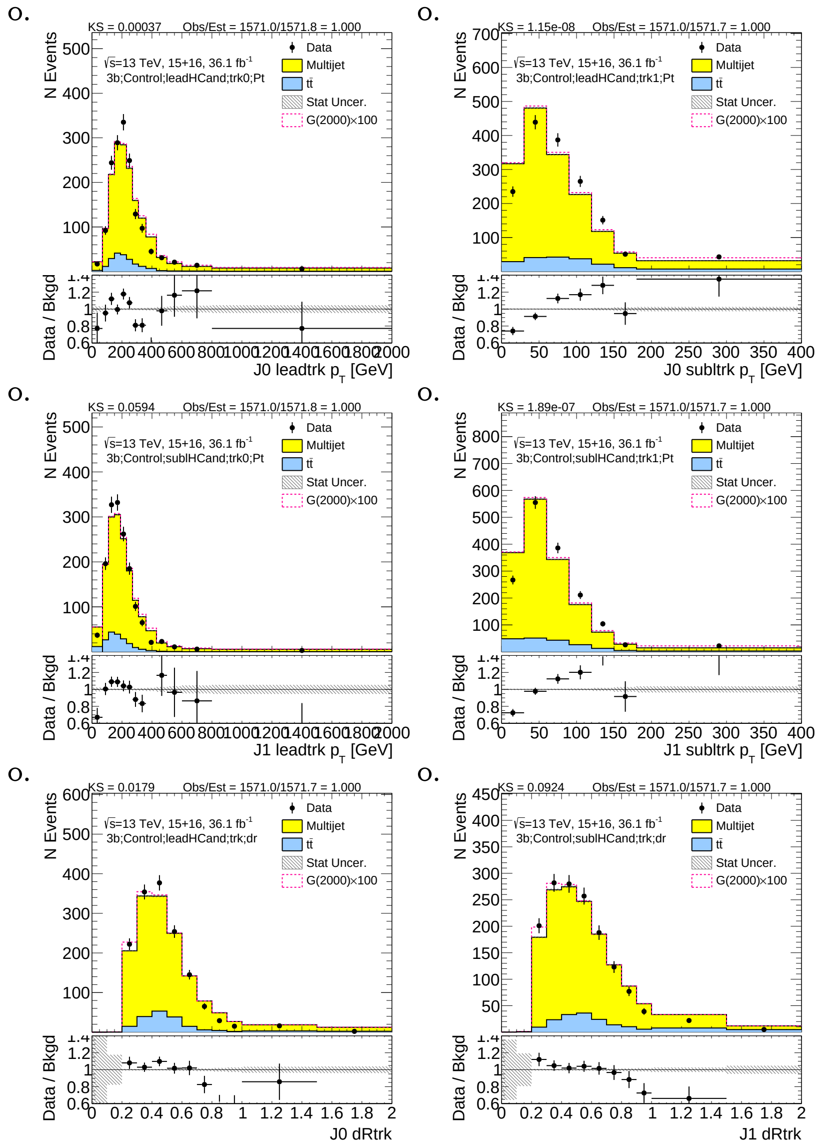

Figure 7.7: First two rows show the kinematics of the lead (left) and sub-lead (right) small-R track jets associated to the lead (first-row) and sub-lead (second-row) large-R jet in data and prediction in the sideband region after requiring 3 b-tags. Third row shows the ΔR between two leading small-R track-jets associated to the leading (left) and sub-leading (right) large-R jet. The normalization agrees by construction, and the shapes are a feature of the prediction.

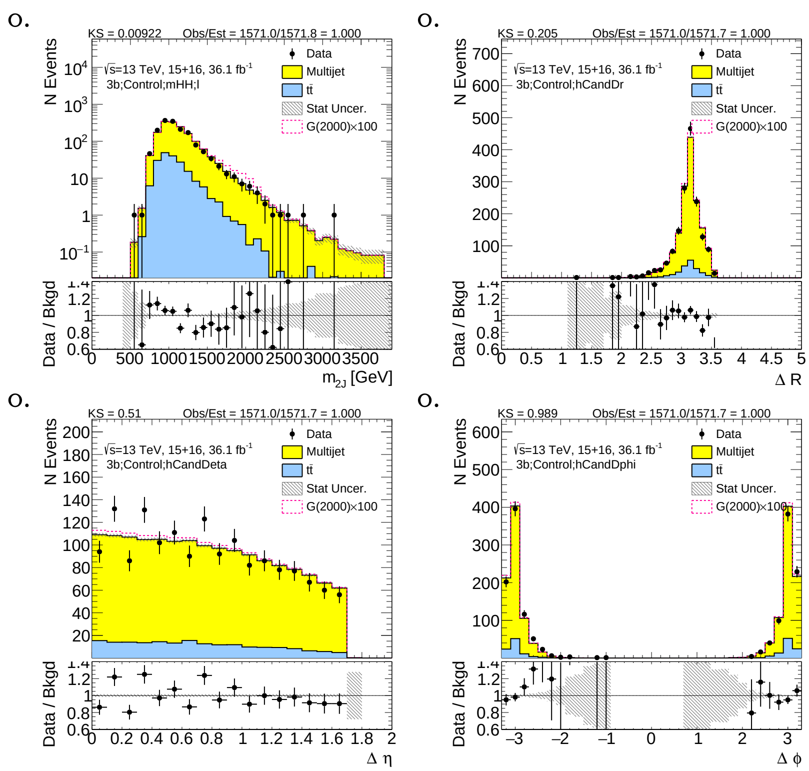

Figure 7.8: Kinematics of the large-R jet system in data and prediction in the sideband region after requiring 3 b-tags. The normalization agrees by construction, and the shapes are a feature of the prediction.

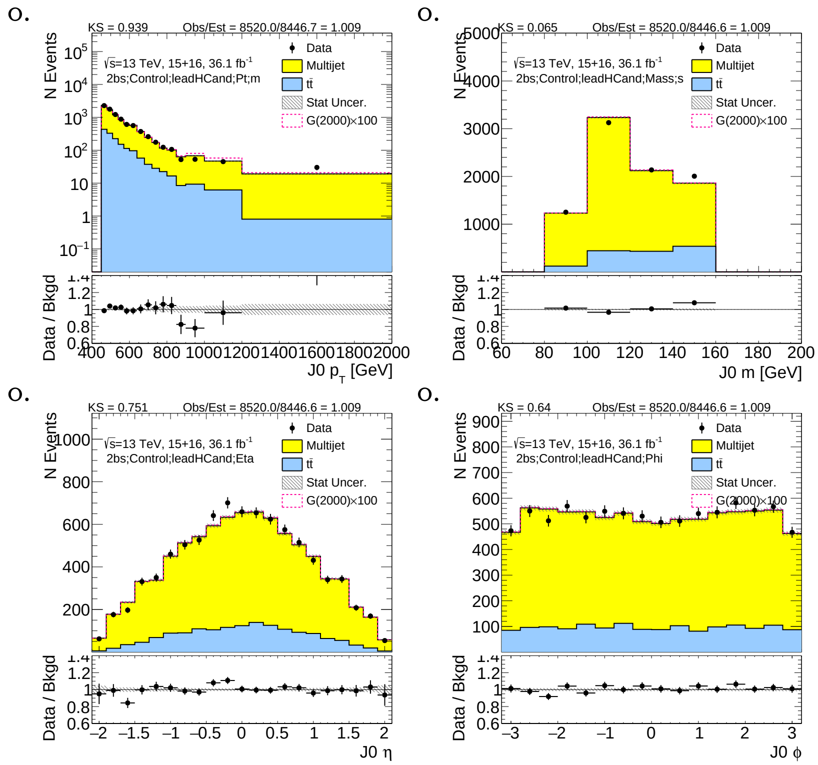

Figure 7.9: Kinematics of the lead large-R jet in data and prediction in the sideband region after requiring 2 b-tags split. The normalization agrees by construction, and the shapes are a feature of the prediction.

7.4 Signal Region Predictions

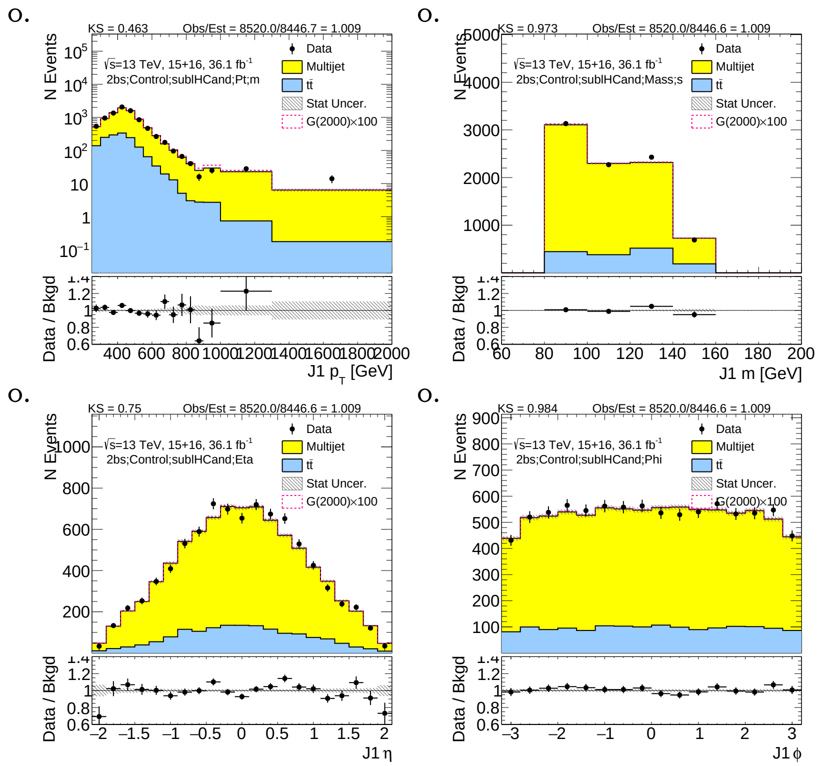

Figure 7.10: Kinematics of the sub-lead large-R jet in data and prediction in the sideband region after requiring 2 b-tags split. The normalization agrees by construction, and the shapes are a feature of the prediction.

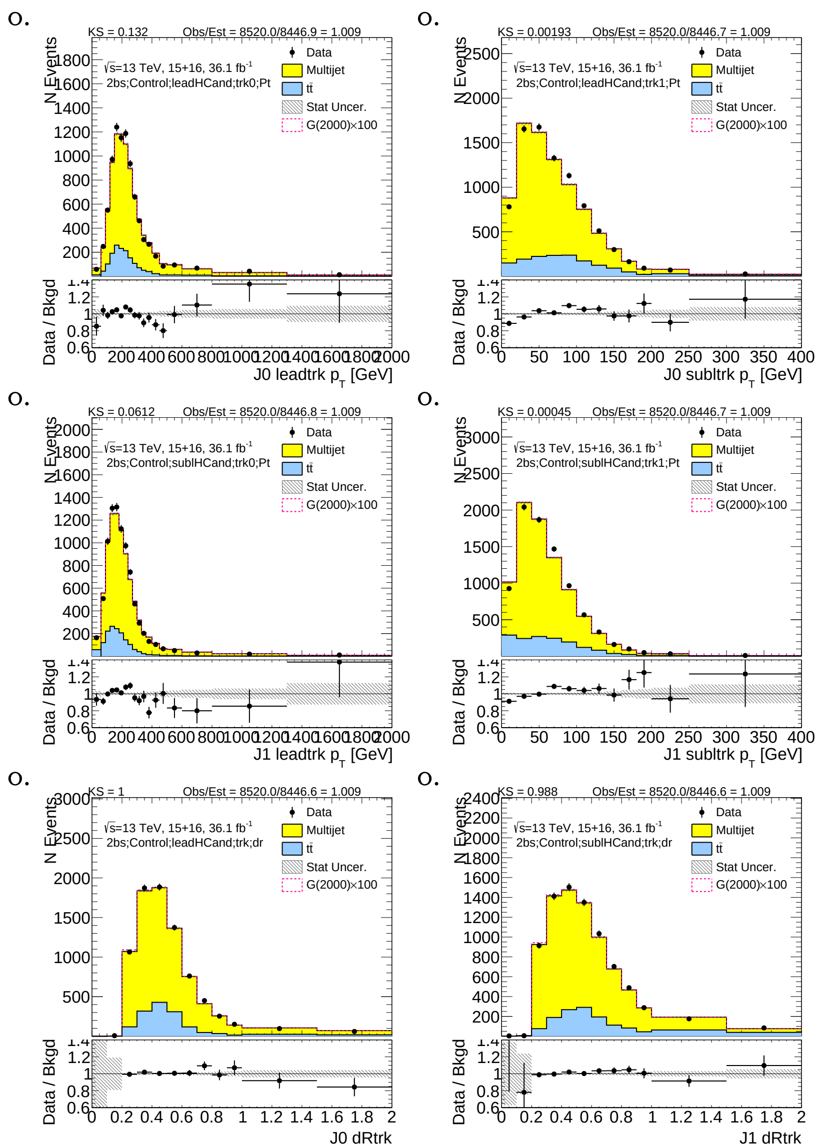

Figure 7.11: First two rows show the kinematics of the lead (left) and sub-lead (right) small-R track jets associated to the lead (first-row) and sub-lead (second-row) large-R jet in data and prediction in the sideband region after requiring 2 b-tags split. Third row shows the ΔR between two leading small-R track-jets associated to the leading (left) and sub-leading (right) large-R jet. The normalization agrees by construction, and the shapes are a feature of the prediction.

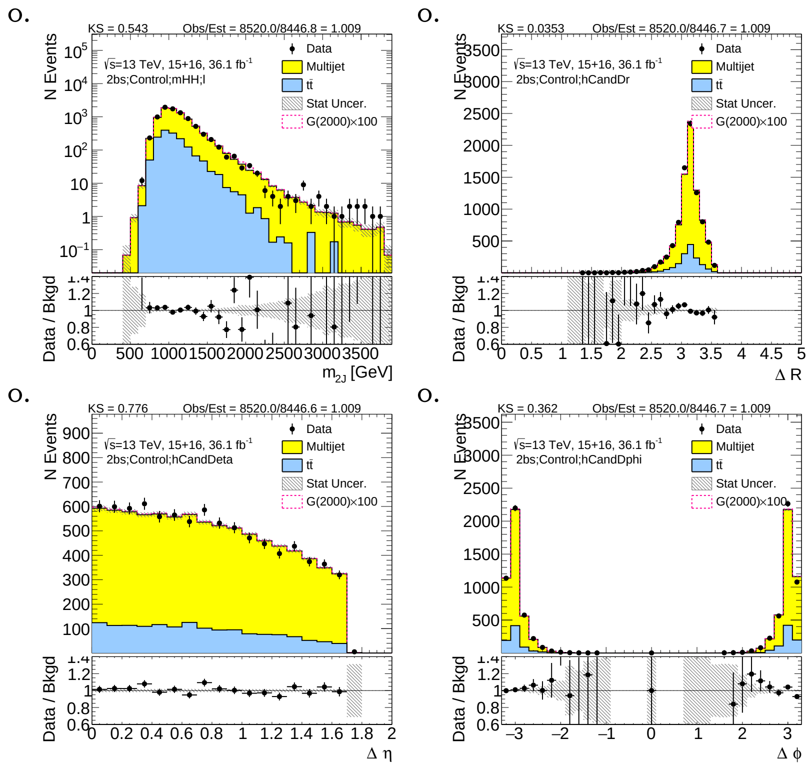

Figure 7.12: Kinematics of the large-R jet system in data and prediction in the sideband region after requiring 2 b-tags split. The normalization agrees by construction, and the shapes are a feature of the prediction.

Since we do physics.

Tony Tong

8 Systematics

Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Morbi commodo, ipsum sed pharetra gravida, orci magna rhoncus neque, id pulvinar odio lorem non turpis. Nullam sit amet enim. Suspendisse id velit vitae ligula volutpat condimentum. Aliquam erat volutpat. Sed quis velit. Nulla facilisi. Nulla libero. Vivamus pharetra posuere sapien. Nam consectetuer. Sed aliquam, nunc eget euismod ullamcorper, lectus nunc ullamcorper orci, fermentum bibendum enim nibh eget ipsum. Donec porttitor ligula eu dolor. Maecenas vitae nulla consequat libero cursus venenatis. Nam magna enim, accumsan eu, blandit sed, blandit a, eros.

Enfin.

French Person

9 Result

Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Morbi commodo, ipsum sed pharetra gravida, orci magna rhoncus neque, id pulvinar odio lorem non turpis. Nullam sit amet enim. Suspendisse id velit vitae ligula volutpat condimentum. Aliquam erat volutpat. Sed quis velit. Nulla facilisi. Nulla libero. Vivamus pharetra posuere sapien. Nam consectetuer. Sed aliquam, nunc eget euismod ullamcorper, lectus nunc ullamcorper orci, fermentum bibendum enim nibh eget ipsum. Donec porttitor ligula eu dolor. Maecenas vitae nulla consequat libero cursus venenatis. Nam magna enim, accumsan eu, blandit sed, blandit a, eros.

Tell the world.

Marathon

10

Intepretation

In order to avoid the Oops-Leon cases, commonly accepted standard for announcing the discovery of a particle is that the number of observed events is 5 standard deviations (σ) above the expected level of the background.

With no excess observed, a limit needs to be set on the cross section of the signal [2].

An estimator should satisfy three criterias:

- Consistency: the value of the estimator should converge to the truth value if the sample size goes to infinity
- Efficiency: theory limits the variance of the true value, given a sample size N (Minimum Variance bound, or MVB). If the variance of the estimator is equal to the MVB the estimator is

called efficient.

- Unbiased: the estiator should have no difference from the true value, otherwise it is biased.

Maximum likelihood estimator is a good estimator, based on these criterias.

For measurements, most of the time https://arxiv.org/abs/1611.01927 is done to compare with generator level distributions. This accounts for detector effects, statistical fluctuations and background mis-identification. Since the analysis is a search, this $hh \rightarrow b\bar{b}b\bar{b}$is a search, unfolding is less applicable in this case.

11

Conclusion

We will find something new one day, because there is always soemthing new.

A

Some extra stuff

Some appendix.

References

[1] Georges Aad et al. Performance of *b*-Jet Identification in the ATLAS Experiment. *JINST*, 11 (04):P04008, 2016. doi: 10.1088/1748-0221/11/04/P04008.

[2] Glen Cowan, Kyle Cranmer, Eilam Gross, and Ofer Vitells. Asymptotic formulae for likelihood-based tests of new physics. *Eur. Phys. J.*, C71:1554, 2011. doi: 10.1140/epjc/s10052-011-1554-0,10.1140/epjc/s10052-013-2501-z. [Erratum: Eur. Phys. J.C73,2501(2013)].