

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

©2014 – BAOJIA TONG
ALL RIGHTS RESERVED.

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 by the ATLAS Collaboration at $\sqrt{s} = 13$ TeV, corresponding to $X \pm \gamma \text{ fb}^{-1}$ of 2015 and $\mathcal{A} \pm B \text{ fb}^{-1}$ 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

0	INTRODUCTION	I
1	MOTIVATION FOR SEARCHES BEYOND THE STANDARD MODEL	2
1.1	the Standard Model	2
1.2	Problems with the Standard Model	2
1.3	Paths Beyond the Standard Model	2
2	MACHINE OF DISCOVERY—THE LARGE HADRON COLLIDER	3
2.1	Design	3
2.2	Performance	3
3	EYES OF GIANT—THE ATLAS DETECTOR	4
3.1	Trigger and Data Acquisition	4
4	GAME OF JIGSAW—RECONSTRUCTION	5
4.1	Introduction to objects	5
5	CENSOR OF WORK—DATA QUALITY	8
5.1	Data Flow in ATLAS	8
5.2	Online Monitoring	8
5.3	Offline Monitoring	8
6	EVENT SELECTION	9
6.1	Optimization Strategy	9
6.2	b tagging	10
6.3	Signal Efficiency	10
7	BACKGROUND ESTIMATION	11
7.1	Background composition	11
8	SYSTEMATICS	12

9	RESULT	13
10	INTEPRETATION	14
11	CONCLUSION	16
	APPENDIX A SOME EXTRA STUFF	17
	REFERENCES	18

Listing of figures

Listing of tables

THIS IS THE DEDICATION.

Acknowledgments

THANK YOU MELISSA FRANKLIN, the best advisor ever. It was five years ago when I made the decision, that of all the schools I visited, Melissa is the only one that will make my life "suffer" the most for the next five years. And I think it is safe to say that I was still right. This is probably one of the hardest paths I could foresee, yet at every corner, Melissa is there to guide me and show me the lights at the end of the tunnel. She is always pushing me to understand the details I ignore, and think about every action I make in analysis or operations. Her encouragements, or complains about my progress, have always been inspirational and keep me moving ahead. There have been multiple nights when I was worried about work, and after talking to her, I always felt much more confident. Thank you, Melissa. Thank you, Melissa. Thank you, Melissa.

THANKS TO MY FRIEND, STEPHEN—I DON'T KNOW HOW TO PRONAUNCE HIS MIDDLE NAME—CHAN, for his great accompaniment and help. Dr. Chan has been extremely awesome-o.

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 giant's shoulders, and I am standing on the ATLAS(member)'s shoulders. Without

their work on detector design, commissioning, 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 truly and sincerely thank everyone deeply for their contributions.

THANK YOU MY FATHER WEIDONG TONG, who has always been the greatest inspiration for me. He has always been the smartest person I know, and he will always be.

THANK YOU MY MOTHER LIPING SUN, who has always been the greatest understanding for me. She has always been the wisest person I know, and she will always be.



Introduction

Why do we look for $b\bar{b} \rightarrow 4b$?

There are two types of analysis in particle physics. The first one is measurement, which yields a observable with an uncertainty. This could either improve 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 justify in data whether 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.

Random People

1

Motivation for searches beyond the Standard Model

THERE'S SOMETHING OUT THERE THAT WE DON'T KNOW.

1.1 THE STANDARD MODEL

1.2 PROBLEMS WITH THE STANDARD MODEL

1.3 PATHS BEYOND THE STANDARD MODEL

Run.

Random Person

2

Machine of discovery–The Large Hadron Collider

2.1 DESIGN

2.2 PERFORMANCE

DAMN.

Quoteauthor Lastname

3

Eyes of giant–The ATLAS Detector

We love ATLAS.

3.1 TRIGGER AND DATA ACQUISITION

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.

Quoteauthor Lastname

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 involving the electroweak forces.

Electrons are maximal ionizing in the tracking system. They get absorbed completely in ECAL and 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 Spectrometer. This is useful for reducing the gen-

erate rate of events, hence for triggering. They are also very clean in reconstruction and have excellent resolutions up to 1 TeV.

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 of the jet into subjets. k_t algorithm with $R_{sub} = 0.2$ is used, and if the subjets have p_T less than 5% of the original jet p_T , the constituent is removed. The calorimeter-based jet mass (m^{calo}), with calorimeter-cell cluster constituents i with energy E_i , momentum \vec{p}_i ($|\vec{p}_i| = E_i$) is defined as:

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

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-assisted jet mass (m^{TA}) is defined as:

$$m^{TA} = \frac{p_T^{calo}}{p_T^{track}} \times m^{track} \quad (4.2)$$

where p_T^{calo} is the calorimeter measurement, p_T^{track} is the four-vector sum of tracks associated to the large-radius calorimeter 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 improves 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 measured energy deposits, which are space-points. Then three space-points form a seed, and they are combined to build track candidates using a Kalman filter. After ambiguity solving, an artificial neural network

is trained and used to identify 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. $Z_o^{BL} \sin \vartheta$ is required 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 efficiency for b and c jets slightly decreases with jet p_T .

Nulla facilisi. In vel sem. Morbi id urna in diam dignissim feugiat. Proin molestie tortor eu velit. Aliquam erat volutpat. Nullam ultrices, diam tempus vulputate egestas, eros pede varius leo.

Quoteauthor Lastname

5

Censor of work–Data Quality

5.1 DATA FLOW IN ATLAS

5.2 ONLINE MONITORING

5.3 OFFLINE MONITORING

Nulla facilisi. In vel sem. Morbi id urna in diam dignissim feugiat. Proin molestie tortor eu velit. Aliquam erat volutpat. Nullam ultrices, diam tempus vulputate egestas, eros pede varius leo.

Quoteauthor Lastname

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 estimated background are both small. A better parametrization for low signal strength 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 predictions after smoothing and compare the asymptotic expected exclusion limits. Both methods

yield comparable results.

To avoid bias during the optimization 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 $g\bar{g} \rightarrow c\bar{c}$ in the signal regions.

6.3 SIGNAL EFFICIENCY

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

Nulla facilisi. In vel sem. Morbi id urna in diam dignissim feugiat. Proin molestie tortor eu velit. Aliquam erat volutpat. Nullam ultrices, diam tempus vulputate egestas, eros pede varius leo.

Quoteauthor Lastname

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.

Nulla facilisi. In vel sem. Morbi id urna in diam dignissim feugiat. Proin molestie tortor eu velit. Aliquam erat volutpat. Nullam ultrices, diam tempus vulputate egestas, eros pede varius leo.

Quoteauthor Lastname

8

Systematics

LOREM IPSUM DOLOR SIT AMET, consectetur 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 consectetur. 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.

Nulla facilisi. In vel sem. Morbi id urna in diam dignissim feugiat. Proin molestie tortor eu velit. Aliquam erat volutpat. Nullam ultrices, diam tempus vulputate egestas, eros pede varius leo.

Quoteauthor Lastname

9

Result

LOREM IPSUM DOLOR SIT AMET, consectetur 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 consectetur. 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.

Nulla facilisi. In vel sem. Morbi id urna in diam dignissim feugiat. Proin molestie tortor eu velit. Aliquam erat volutpat. Nullam ultrices, diam tempus vulputate egestas, eros pede varius leo.

Quoteauthor Lastname

10

Interpretation

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 estimator should have no difference from the true value, otherwise it is biased.

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

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 $b\bar{b} \rightarrow b\bar{b}b\bar{b}$ is a search, unfolding is less applicable in this case.

11

Conclusion

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



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)].