# Di-Higgs Production Limits from Top Quark Pairs ($t\bar{t}HH$)

*Japnidh Thakral, Franny Setti, Hualin Mei, Nick Amin, Claudio Campagnari*
*Dept. of Physics at University of California, Santa Barbara*

College of Creative Studies

## Introduction

The discovery of the **Higgs Boson** as the fundamental scalar particle in the Standard Model has been an important impetus of recent particle physics efforts. One of the sought-after properties of the Higgs boson is the ability to self interact and "decay" into two of itself in the Feynman diagram level (di-Higgs), which while confirmed in theory, is considered a very rare process experimentally. In the landscape of di-Higgs decays, the top-antitop quark channels ($t\bar{t}HH$, see Fig 1.) have limited investigation and the signatures for this process are similar to events from a previous analyses conducted at UCSB ($t\bar{t}t\bar{t}$). Here, we aim to leverage the tools of these analyses to set bounds for the **cross section** for this process, which is intimately related to the probability of this di-Higgs event occurring in quantum field theory.

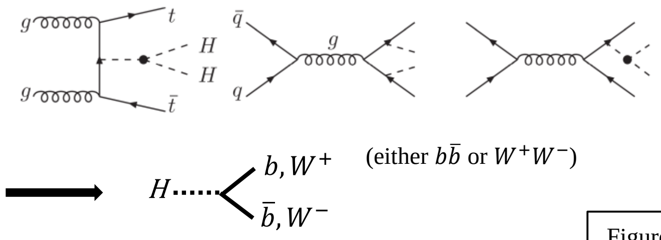


Figure 1

## Setup

To establish our data, we use Monte Carlo's (MC) or simulations of data from the Compact Muon Solenoid (CMS) detector at the Center of European Nuclear Research (CERN), hosting the Large Hadron Collider. Through the miniAOD data structure, we have access to simulations that trace the lifetime of particles, from 13 TeV proton-proton collisions, that go through various scattering, decay, and oscillation processes that leave signatures that eventually reach detectors, and in our case, a final state containing b quarks and W bosons (not counting the ones from the top quark decay, which have opposite charge). This final state has interference with other processes that originate from the top quark production, so it is imperative to **isolate our signal**, $t\bar{t}HH$, from this background to assess the cross section. After cutting processes below a certain energy threshold and angles within the detector, the main backgrounds that we account for are: $t\bar{t}H$, $t\bar{t}W$, $t\bar{t}Z$, $t\bar{t}$ with a pair of bosons, nonprompt and charge misidentified lepton processes, single photon processes, rare diboson/triboson along with single top quark processes, and finally $t\bar{t}t\bar{t}$.

## Data Acquisition

To facilitate the isolation of this data, we make use of ROOT, a C++ based software language that is adept for high level data structures tailored for particle physics, and the use of **boosted decision trees** (BDT's), a data algorithm that isolates a particular signal from a background multivariately by machine training (CERN TMVA tools).

Since the signatures for $t\bar{t}t\bar{t}$ and $t\bar{t}HH$ are very similar, more so that the remainder of the background, through comparable final states, we train a BDT for discerning $t\bar{t}HH$ from a combined $t\bar{t}HH + t\bar{t}t\bar{t}$ sample, and another for discerning it from the rest of the background (labeled "$ttV$"). Below (Fig. 2) is an example graph of input variables added on top of UCSB's $t\bar{t}t\bar{t}$ analysis, where the red fill indicates **background** ($t\bar{t}t\bar{t}$) and blue being **signal** ($t\bar{t}HH$) for our first BDT. The top three graphs are invariant mass combinations of b-quarks and leptons with the highest b-jet discriminants, or likelihoods of being b-quarks. The bottom three represent their respective angular separations.

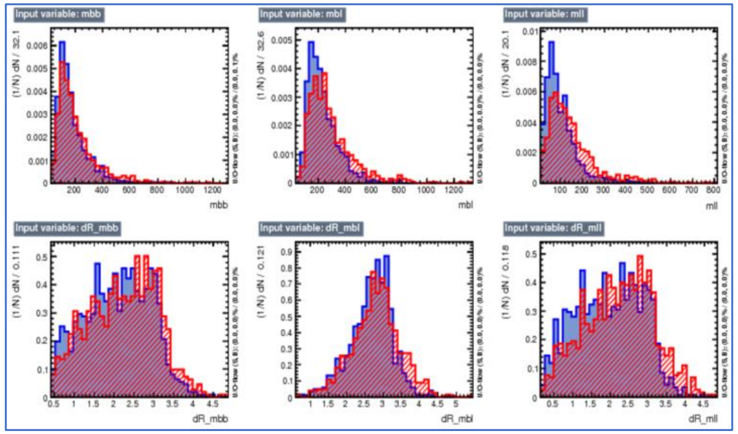


Figure 2

Then, with both BDT's, they will output a score for each data point. We can analyze the 2D score histogram where the training outputs higher ratios of signal in a bin, or in other words, the quantitative **statistical significance**, defined as cumulative signal data over the square root of cumulative total data ($S/\sqrt{S+B}$) (see Fig. 3 for significance vs BDT scores). These will be the regions with the highest sensitivity, experimentally speaking. From the isolated regions, we decompose the signal and background into their respective yields.

*2D Significance for BDT Scores ($t\bar{t}t\bar{t}$ vs $t\bar{t}V$)*

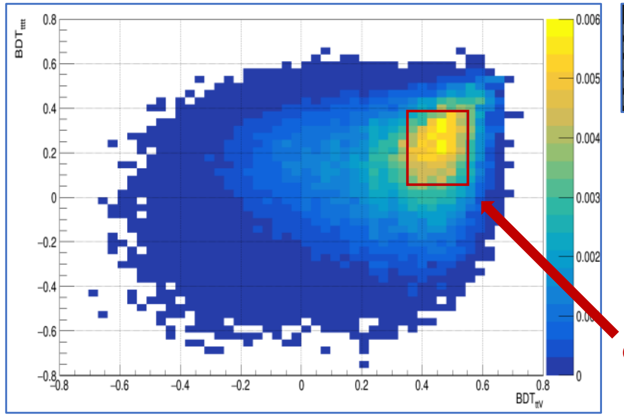


Figure 3

```
-- AsymptoticLimits ( CLs ) --
Expected  2.5%: r < 36.7251
Expected 16.0%: r < 49.1860
Expected 50.0%: r < 68.6250
Expected 84.0%: r < 96.5263
Expected 97.5%: r < 130.2596
```

Figure 4

## Analysis and Conclusion

Inputting the yields of these sensitive bins and looking at the asymptotic limits (confidence level of CERN's HiggsCombine tool, see Fig. 4), setting the upper limit of the cross section of this process to be around **100 times the standard model cross section at 95% confidence level**. This channel lends itself to an inherently small cross section than the more conventional gluon-gluon to di-Higgs into b quark and photon final states, and this channel also presents a complex topology, with multiplicities and combinatorics, as well as employing the use of machine learning. However, this report would serve as a great compliment to existing experimental di-Higgs searches and has potential for competitive results as it takes account of the electroweak interaction.

## Acknowledgements

My CCS Summer Project was possible due to The Create Fund, thanks to the generosity of John and Susan Axline. I would like to thank Nick Amin (former graduate student, now postdoc) and Prof. Claudio Campagnari for their expertise on the subject, having been part of the $t\bar{t}t\bar{t}$ and same-sign dilepton analyses and building much of the groundwork for this project to be possible. I would also like to thank Franny Setti and Hualin Mei for working closely with me and acquainting me with the complex software environment.