Exploring the $Z' \to t\bar{t}$ heavy resonance at FCC-hh

Rachel Emma Clarke Smith
University of Illinois at Urbana-Champaign

Advisors:
Michele Selvaggi, Clement Helsens, & Filip Moortgat

August 11, 2017

1 Introduction

The future circular proton-proton collider (FCC-hh) is a proposed post-LHC particle accelerator. It will reach energies up to 100 TeV and be capable of producing exotic particles of masses up to 40 TeV. The top (t) quark, the heaviest particle discovered so far, is likely coupled to these new massive particles. Therefore, it is important to explore the feasibility of detecting final states containing top quarks at 100 TeV. Because they are produced with significant Lorentz boost, the top quark decay products will be highly collimated and will overlap. It is therefore also important to understand how the FCC detector resolution and granularity affects the detection of these new heavy particles.

This project focused on the $Z' \to t\bar{t}$ heavy resonance, where Z' is a BSM, heavy, Z-like gauge boson. In this process, Z' decays to two top-jets. These top-jets can either be hadronic ($t \to bW \to bq\bar{q}'$) or leptonic ($t \to bW \to b\ell\nu$). The full process is then either all-hadronic, semi-leptonic, or di-leptonic. This project focused only on hadronic decays (no cuts were made on the existence of a lepton in the event). In this project I determined the exclusion cross-section of $Z' \to t\bar{t}$ and the integrated luminosity required to make a discovery at the baseline FCC detector at 95% confidence level.

2 Analysis

2.1 Workflow

All events used in this analysis were produced by MadGraph MG5 gridpacks and Pythia8 or Pythia8 alone. Delphes, a fast multipurpose detector response simulation framework, was then used to simulate the response of the FCC baseline detector to these events. Finally, heppy, a modular python framework for the analysis of collision events, was used to run over the FCCSW samples and produce flat trees with relevant variables. These flat trees were then used to produce the results shown.

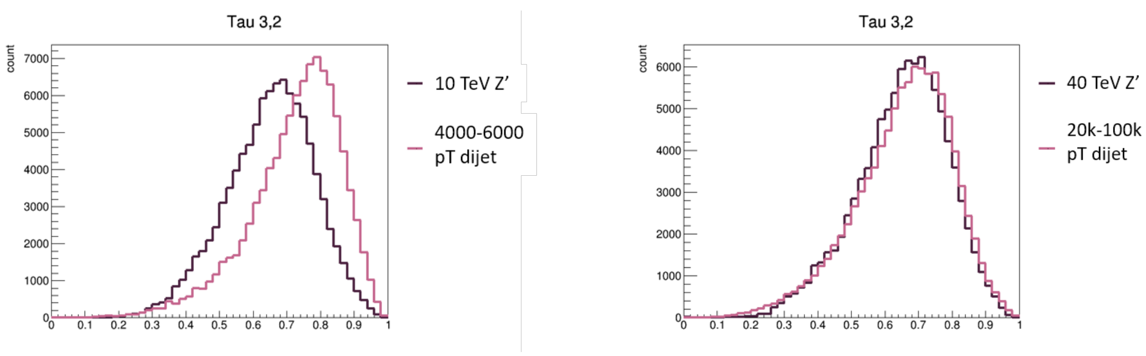
2.2 Background reduction

The major backgrounds for $Z' \to t\bar{t}$ are QCD dijets, irreducible $t\bar{t}$, and VV standard model processes ($V = W/Z$). It was necessary to identify variables and develop techniques to reduce these backgrounds since, for example, the irreducible $t\bar{t}$ background increases approximately $\times 40$ from 14 TeV to 100 TeV.

2.2.1 Jet substructure

One way to distinguish between jets produced from boosted tops and jets produced from light quarks or gluons (QCD jets) is to look at jet substructure. In short, a variable dubbed "N-subjettiness" or τ_N is used to quantify the likelihood of the jet being N-prong. The ratio τ_3/τ_2 is used to distinguish 3-prong objects and the ratio τ_2/τ_1 is used to distinguish 2-prong objects.

As the mass of the Z' increases and the tops are more boosted, the limited detector resolution makes it increasingly difficult to distinguish the individual jet components. As an estimate, consider a 10 TeV Z'. The distance between outgoing partons would be approximately $\Delta R \sim 0.07$. For a 40 TeV Z', $\Delta R \sim 0.02$. The distance between actual jet components would be even closer. The typical ECAL/HCAL resolutions at the LHC (FCC) are 0.02/0.1 (0.01/0.05).

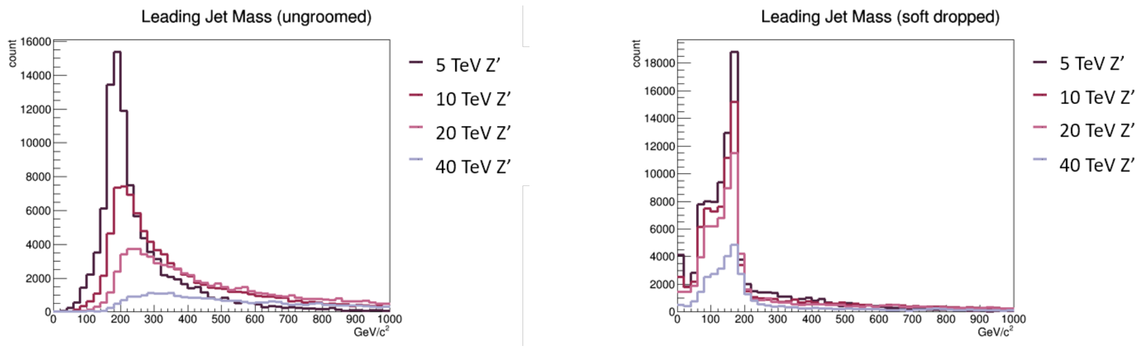


(a) (b)

Figure 1: τ_3/τ_2 of Z' and QCD background for 10 TeV and 40 TeV mass points.

2.2.2 Jet grooming

Jet grooming is a method of removing soft radiation (initial/final state radiation, pile-up, the underlying event, etc.) from a jet. I tested three different grooming algorithms: "soft drop," "trimming," and "pruning." Jet grooming improves the resolution of the reconstructed jet mass. The mass can then be used as a cut variable to reduce background.



(a) (b)

Figure 2: (a) Final state radiation adds unnecessary radiation to the jet cone and smears the reconstructed jet mass. (b) Effects of soft drop jet grooming algorithm on the reconstructed jet mass.

2.2.3 Top tagging

As part of my project, I explored the possibility of using boosted decision trees (BDTs) to develop a "top tagger" capable of distinguishing between boosted top jets and QCD jets. I used the TMVA framework within ROOT. I supplied the framework with a set of "training" events and discriminating variables. The framework then produced a set of boosted decision trees to optimally separate the "signal" and "background" based on the discriminating variables. Finally, the framework ran over a separate set of "testing" events to test for over-training. In the end, I did not have time to implement this top tagger in the analysis, so the cuts used in the Results section are all rectangular cuts on single variables.

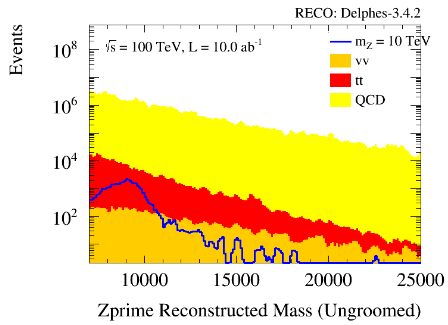
3 Results

The following cuts were used:

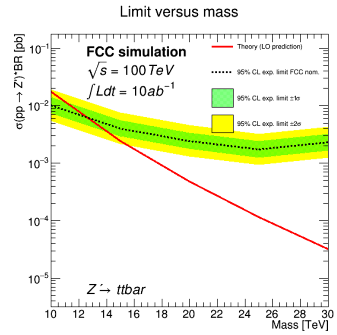
<table>
<tr><td colspan="2">At least 2 "fat" jets in the event</td></tr>
<tr><td colspan="2">Jet[1] p_T > 2.8 TeV and Jet[2] p_T > 2.8 TeV</td></tr>
<tr><td colspan="2">abs(Jet[1] eta) < 3.0 and abs(Jet[2] eta) < 3.0</td></tr>
<tr><td colspan="2">and either/or:</td></tr>
<tr><td>Jet[1] soft drop inv mass > 140 GeV</td><td>Jet[2] soft drop inv mass > 140 GeV</td></tr>
<tr><td>Jet[1] $Tau_{32} < 0.7$</td><td>Jet[2] $Tau_{32} < 0.7$</td></tr>
<tr><td>0.3 < Jet[1] $Tau_{21} < 0.8$</td><td>0.3 < Jet[2] $Tau_{21} < 0.8$</td></tr>
</table>

Table 1: Cuts used in this analysis.

I first ran the analysis with these cuts over all signal and background:



(a) Reconstructed m_{jj}.



(b) Excluded cross-section.

Figure 3: Results with QCD background included.

Figure 3a shows the reconstructed mass of 10 TeV Z' vs all backgrounds. The mass is calculated from the combined invariant mass of the two jets with the highest p_T in the event. With these cuts, we exclude up to 12 TeV, as shown in Figure 3b.

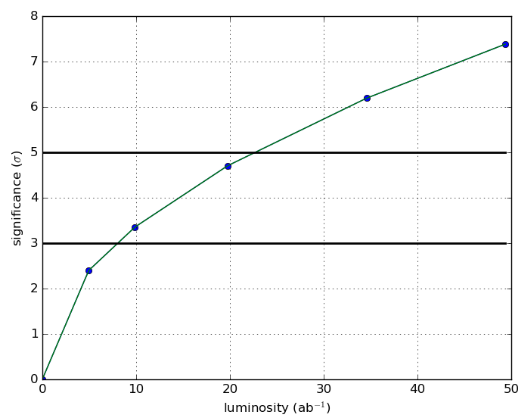
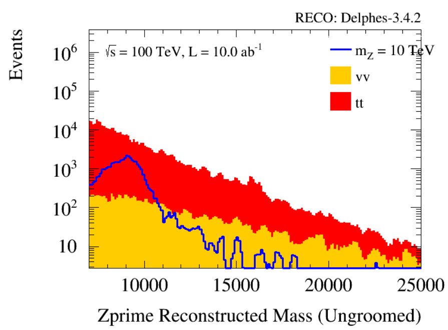
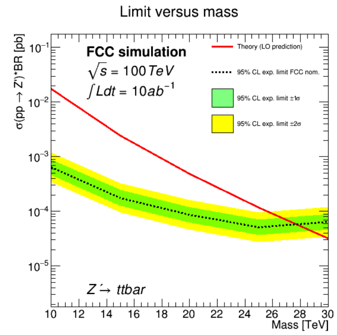


Figure 4: The integrated luminosity required to make a 5σ discovery of a 10 TeV Z' at the baseline FCC detector with the QCD background included.

I then tested how these results would change if the large QCD background was excluded. The QCD background can be suppressed, by b-tagging for instance, but I did not have enough time to implement such selections in the analysis.



(a) Reconstructed m_{jj}.



(b) Excluded cross-section.

Figure 5: Results with QCD background excluded.

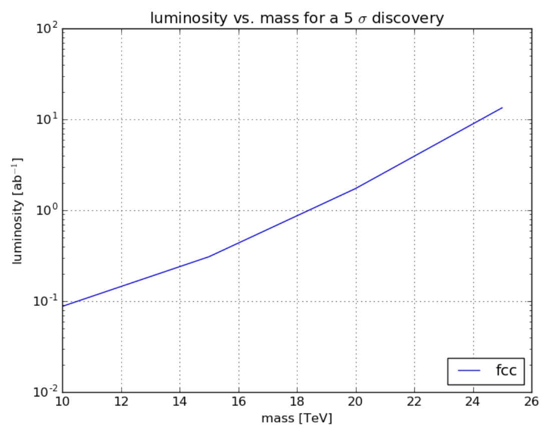


Figure 6: Integrated luminosity required to make a 5σ discovery at the FCC baseline detector for each Z' mass point at with the QCD background excluded.

4 Future work

This analysis can be further optimized to reduce backgrounds. Boosted decision trees and other multi-variable analysis methods can be used to reduce the QCD background with b-tagging and top-tagging. Applying cuts on the existence of leptons (semi-leptonic/di-leptonic decays) would also serve to kill the QCD background. With the analysis framework now in place, it should be relatively easy to perform the analysis on more selections and more detector configurations. Finally, it is necessary to ensure that all parts of this analysis are general and reusable so that these results are reproducible.

5 Acknowledgments

I'd like to sincerely thank Michele, Clement, Filip, Ine, Steve, Cari, Junjie, the CERN summer student programme, and the National Science Foundation for supporting my work at CERN.