Monte Carlo study of the impact of event selection on top-quark mass extraction in the dileptonic $t\bar{t}$ decay channel at $\sqrt{s} = 100$ TeV

Vinayak Kapoor

August 2022

Acknowledgements

Thank you for the guidance of Andrew James Wildridge from Purdue University in the development of this research paper.

Abstract

Measurements of the top-quark mass (m_t) are presented using events containing two oppositely charged leptons (e^+e^-,$e^\pm\mu^\mp$ or $\mu^+\mu^-$) produced in proton-proton collisions at a center-of-mass energy of 100 TeV. The data used for these measurements are produced by a Monte Carlo (MC) event generator for the FCC-hh, an energy-frontier hadron collider. By fitting the distribution of top-quark masses at the generator level with a relativistic Breit-Wigner continuous probability distribution, the top-quark mass without any event selection is measured to be $m_t = 252.48 \pm 86.00$ GeV. An event selection, which involves cuts on the number of leptons, jets, and b-jets, is then carried out to separate signal events with top quarks from background events originating from other physics processes. A Breit-Wigner distribution is then fit to the masses of top quarks that pass the event selection, and the top-quark mass after event selection is measured to be $m_t = 281.24 \pm 103.10$ GeV.

1 Introduction

The top quark is the most massive of all observed fundamental particles. Its mass, m_t, is an important parameter of the standard model (SM) of particle physics, particularly because of how strongly the top-quark interacts with the Higgs field [5]. Because the top quark is roughly 50 times more massive than the next massive fermion, the bottom quark, it is the most important fermion when describing electroweak (EW) theory - which can be described by only the W-boson mass, the Z-boson mass, and the top-quark mass. Current central values of the Higgs boson mass and the top quark mass, however, suggests that the effective Higgs potential develops an instability about the scale $\Lambda_I \approx 10^{11}$ GeV [9]. Hence, New Physics (NP) is required to stabilize the Higgs field and ensure that the electroweak vacuum is stable, and that vacuum decay through quantum tunneling does not occur. This NP is likely to occur at energy scales of 10^{19} GeV as this is where EW theory does not hold up. Hence, measurement of m_t at energy scales greater than 13 TeV – the center-of-mass energy for run 2 of the LHC – (such as at 100 TeV which is the subject of the paper) may strongly disagree with SM predictions, as was seen in the recent W-boson mass measurement [2], hinting toward NP. Indeed even more precise measurements can lead to NP since deviations from SM/EW predictions occurring in higher-order diagrams from virtual particle exchanges of these non-SM interactions may be seen.

At the Large Hadron Collider (LHC) and proposed hadron Future Circular Collider (FCC-hh), top-quarks are produced primarily in quark-antiquark pairs ($t\bar{t}$) through quantum chromodynamics (QCD) interaction that is dominated by gluon-induced processes. These $t\bar{t}$ pairs decay predominantly into a bottom-quark and W-bosons – the decays of which are often used to classify $t\bar{t}$ events as belonging to the all-jets channel, lepton+jets channel, and di-lepton channel [8]. In this paper, the di-lepton channel is analyzed. The final state, therefore, consists of two oppositely-charged leptons.

As a result of several measurements of m_t at the Tevatron and the LHC, the top-quark mass has been found with excellent precision, making it the most precisely known quark mass [7]. In the di-lepton channel, the top quark mass has been measured by the CMS Collaboration with p-p data at $\sqrt{s} = 13$ TeV to be 172.33 ± 0.70 GeV. An indirect mass measurement method was employed using the measured top quark production cross section and theoretical predictions to determine m_t. In this paper, we use a simple mass extraction method on data generated using MC simulators at $\sqrt{s}$= 100 TeV - which is the energy scale of the proposed FCC-hh collider - to find a measurement of m_t. The paper uses a maximum likelihood expectation of a Breit Wigner distribution to model the observed mass distribution as closely as possible, extracting a measurement for the top quark mass. This method is used to find m_t of generator level top quarks before and after event selection to observe the effect of event selection on top quark mass measurements at the FCC-hh.

2 Background

Top Quark Decay Channels

The top quark decay channel chosen to study is responsible for determining both the number of $t\bar{t}$ events available for analysis as well as the number of background events present.

- If both W-bosons decay hadronically the final state is referred to as the all-jets channel:

$t\bar{t} \rightarrow W^+bW^-\bar{b} \rightarrow q\bar{q}^{'}bq^{''}\bar{q}^{'''}\bar{b}$

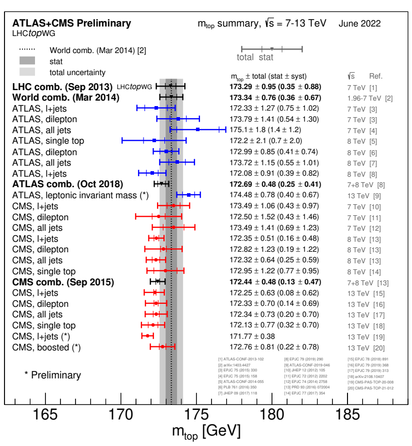
- If one W-boson decays hadronically, and the other leptonically the final state is referred to as the lepton+jets channel:
  $t\bar{t} \rightarrow W^+bW^-\bar{b} \rightarrow l^+v_lbq\bar{q}^{'}\bar{b}$ and
  $t\bar{t} \rightarrow W^+bW^-\bar{b} \rightarrow q\bar{q}^{'}bl^-\bar{v}_l\bar{b}$

- If both W bosons decay leptonically final state is referred to as the di-lepton channel:
  $t\bar{t} \rightarrow W^+bW^-\bar{b} \rightarrow l^+v_lbl^{'-}\bar{v_{l'}}\bar{b}$

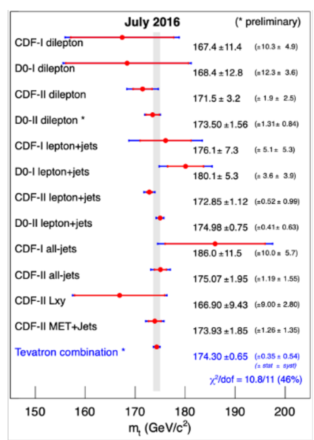
The three decay channels have a different branching fraction - which refers to the fraction of particles that decay by an individual decay channel with respect to the total number of particles that decay. The all-jets channel suffers from the largest background yet has the largest branching fraction. The lepton+jets channel has a moderate branching fraction as well as moderate backgrounds. The di-lepton channel has the smallest branching fraction and smallest backgrounds. The lepton+jets channel provides a fully constrained system and is often studied since it provides the most accurate mass measurement for the top-quark among all final states. Moreover, the moderate branching fraction does not pose a challenge due to the large number of top quarks produced at colliders such as the LHC.

Recent Top Quark Mass Results

Figure 1a and Figure 1b depict the most recent measurements of the top-quark mass from the CMS and ATLAS experiments at the LHC, and the CDF and D0 experiments at the Tevatron, respectively.



(a) Compilation of recent top-quark mass measurements at the LHC [6]



(b) Compilation of recent top-quark mass measurements at the Tevatron [12]

Figure 1: Summary plots of recent top-quark mass measurements

The measurements by the Tevatron collaborations lead to a combined value of $m_t = 174.30 \pm 0.65$ GeV. The most precise measurement for the top-quark mass at the LHC of $m_t = 171.77 \pm 0.38$ GeV [1] was made by the CMS collaboration by studying $t\bar{t}$ events in the lepton+jets channel and using a profile likelihood method to extract the top-quark mass. That being said, there is a tension of around 3.35σ between the Tevatron combined value and the most precise top-quark mass measurement by the CMS collaboration. This amount of tension is fairly common in particle physics and is often eliminated by better analysis methods. In this case, however, since this amount of tension has persisted for nearly a decade and there is a trend of increasing tension between the combined Tevatron measurement and newer top-quark mass measurements at the LHC, there may be the possibility of new physics. Hence making top-quark mass extraction all the more worthwhile, due to its role in discovering new physics.

Top Quark Mass Measurement Methods

The top-quark mass is measured at the CMS, ATLAS, CDF, and D0 experiments in different decay channels and by using several different techniques. These techniques can be broadly classified as direct m_t measurements, which involve using information from the kinematic reconstruction of top-quark decay products, and indirect m_t measurements, which involve comparing the inclusive or differential $t\bar{t}$ production cross-section to the corresponding theory calculations [3].

Direct mass measurements of m_t can further be categorized into the following groups:

- *Template method*: involves comparing the top-quark mass distribution reconstructed from data with Monte Carlo (MC) simulated top-quark mass distributions, which are known as templates, with different values for the top-quark mass parameter in the MC simulation.
- *Matrix element method*: based on the likelihood of observing a set of selected events in the detector. Each event is assigned an m_t-dependent likelihood.
- *Ideogram method*: combines features of both the template and matrix element method (MEM). It can be thought of as a less computing-intensive alternative to the MEM.

Indirect mass measurements of m_t can further be categorized into the following groups:

- *Top-quark mass from inclusive $t\bar{t}$ production cross-section*: exploits the theoretical relation between m_t and the $t\bar{t}$ production cross-section, by comparing the measured cross-section to the theory calculation.

- *Top-quark mass from differential $t\bar{t}$ production cross-section*: uses the differential cross-sections to extract m_t and overcomes the problem of limited precision associated with using the inclusive $t\bar{t}$ production cross-section to extract m_t.

Top Quark Mass Measurement at the FCC-hh

The FCC-hh is a proposed hadron (pp) collider and hence in this study, only hadron-hadron collisions are simulated, not electron-positron ones. It would operate at seven times the center-of-mass energy of the LHC, at 100 TeV, and collect about 10 times more data [10].

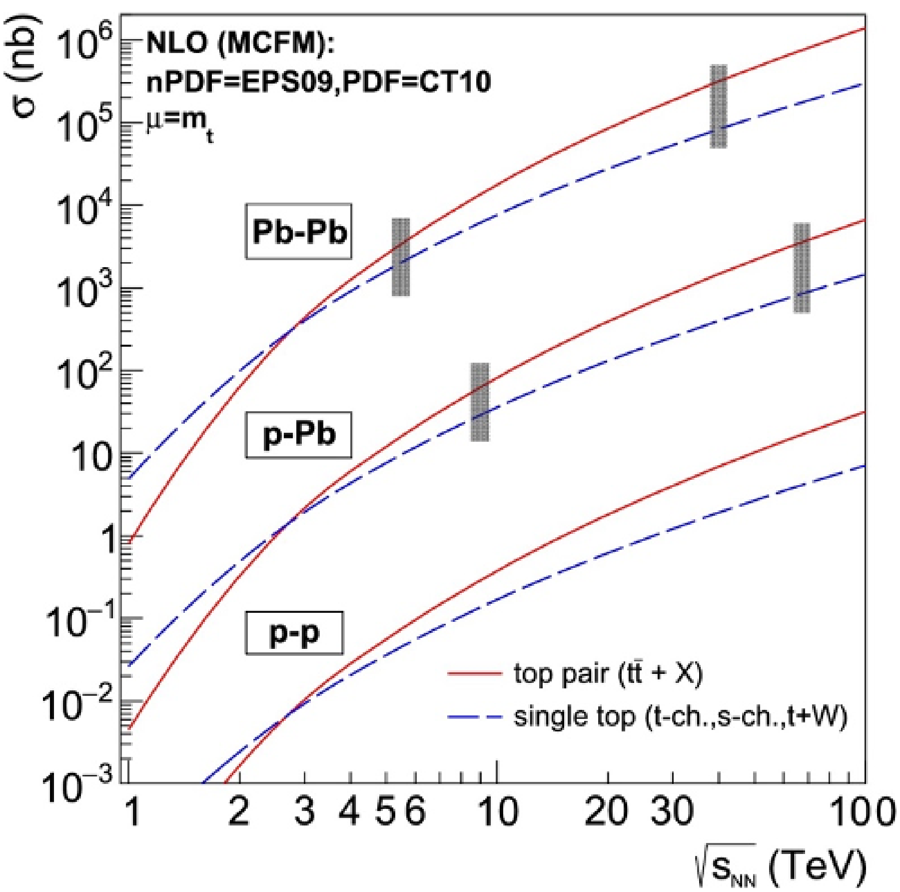


Figure 2: Total cross sections for top pair and single-top pair production in Pb-Pb, p-Pb, and p-p collisions as a function of center-of-mass energy. [4]

Figure 2 shows how at a hadron collider with $\sqrt{s} = 100$ TeV the $t\bar{t}$ production cross-section increases by a factor of approximately 40 when compared to the

LHC, which has center-of-mass energy of 13 TeV. Since at the higher center-of-mass energy of $\sqrt{s} = 100$ TeV, the total cross-section for $t\bar{t}$ events is larger, and more top quarks are produced, hence, leading to a more precise measurement of m_t. That being said top quark physics is constrained by systematic errors. As seen in the most recent measurements of m_t by the CMS collaboration, the majority of uncertainty arises from systematic uncertainty and not statistical uncertainty which only contributes 0.04 GeV to the total uncertainty of 0.48 GeV [1]. Hence, the benefits of a 40 times greater production cross-section will not translate to improvements in sensitivity by this factor.

At higher center-of-mass energies there is a greater probability to produce high-energy top-quark pairs. If new physics, such as super-symmetric top-quarks, exists, it is likely to do so at these higher energy scales. We would see more events than predicted at these higher energies on m_t distribution which is likely to lead to disagreements in the m_t measurement at the FCC-hh and the LHC, due to new physics.

3 Event generation

Data of hadron-hadron collisions – which are responsible for $t\bar{t}$ production - is generated using software tools based upon the Monte Carlo (MC) method. These tools generate simulated $t\bar{t}$ collision events as well as some background contributions that resemble experimental data at the center-of-mass energy of 100 TeV. The LO and NLO Monte Carlo event generator MadGraph5 aMC@NLO is used with top quark mass values m_t^{gen} of 173 GeV. The parton distribution functions (PDFs) set and the strong coupling constant value used for data generation is unknown. Pythia8 is used to handle parton showering and hadronization. Finally, all simulated events were processed through Delphes - a fast multipurpose detector response simulation.

4 Data selection

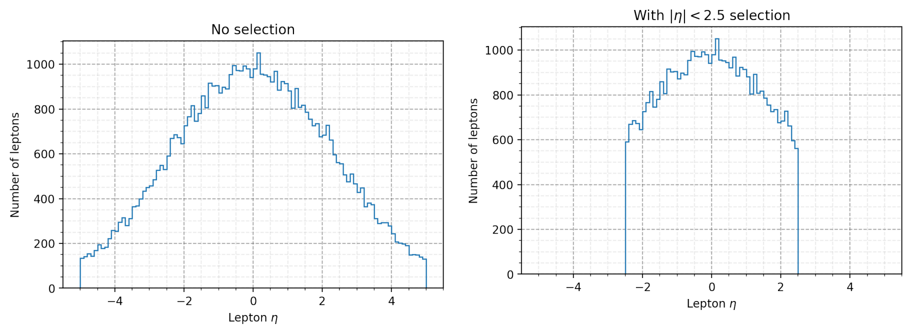
Top Quark selection at the generator level

At the generator level, only top quarks produced by beam-remnant treatment are selected since it is these top quarks that decay into a bottom quark and a W-boson.

Object selection

The leptons used for analysis should be well isolated from other particles and have large transverse momenta. Hence, only leptons that have p_T>20 GeV and $|\eta|$<2.5, and which are isolated from jet candidates within $\Delta R = \sqrt{\Delta\eta^2 + \Delta\phi^2} >$

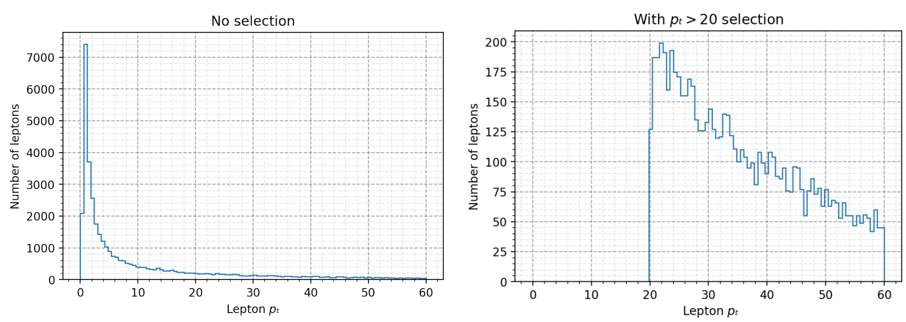
0.02, where $\Delta\eta$ and $\Delta\phi$ are the differences in pseudorapidity and azimuthal angles between the jet and lepton candidate, are kept. The latter is done to reduce contamination as a result of leptons from hadronic weak decays in the jets.



(a) Lepton pseudorapidity before cut

(b) Lepton pseudorapidity after cut of $|\eta|<2.5$

Figure 3: Lepton η cut

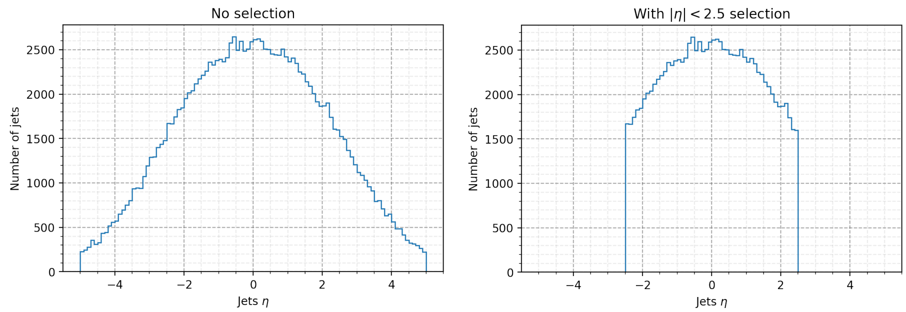


(a) Lepton transverse momentum before cut

(b) Lepton transverse momentum after cut of p_T>20 GeV

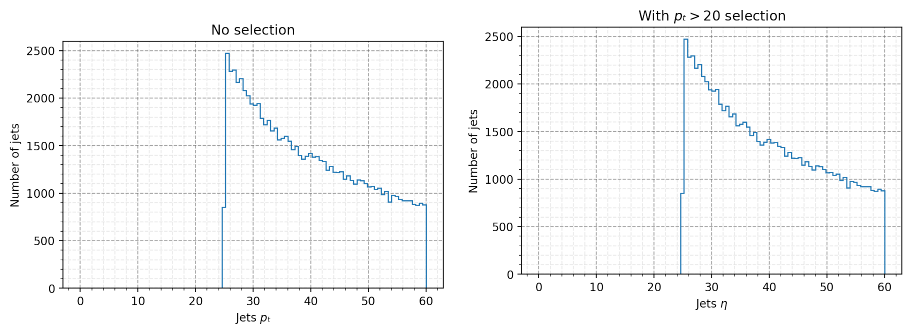
Figure 4: Lepton p_T cut

Moreover, only jets with p_T>20 GeV and within $|\eta|<2.5$ are used for analysis. Top quark decays almost always produce a bottom quark, which hadronizes into a B meson. b-jets which are jets originating from these B mesons are identified using b-tagging algorithms. b-tagging algorithms are based on the distinct properties of the b hadron. These include its long lifetime, high mass, and its semileptonic decay. As a result of their long lifetimes experimentally large displacement of secondary vertices from the primary collision vertex, which can be used to identify b-jets [8].



(a) Jets pseudorapidity before cut (b) Jets pseudorapidity after cut of $|\eta|$<2.5

Figure 5: Jets η cut



(a) Jets transverse momentum before cut (b) Jets transverse momentum after cut of p_T>20 GeV

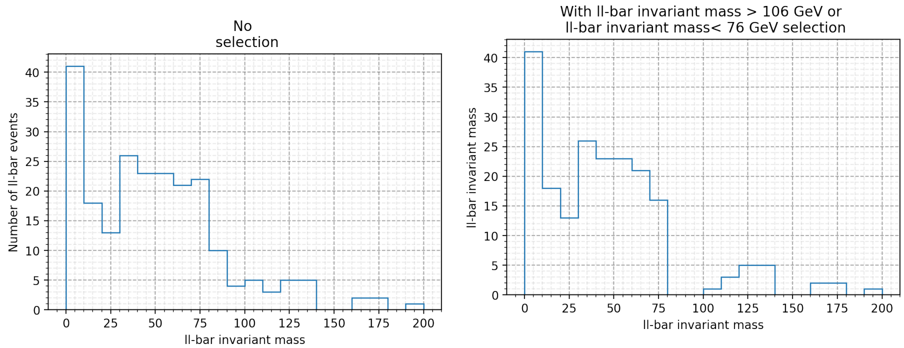
Figure 6: Jets p_T cut

Event Selection

The following paper finds a measurement for m_t by studying the di-lepton channel. This channel is characterized by the presence of two isolated and oppositely charged leptons (e^+e^-,$e^\pm\mu^\mp$ or $\mu^+\mu^-$), E_T^{miss} arising from the neutrinos from W boson decay, and two b-jets.

Hence, only events that contain two oppositely charged leptons (e, μ) are selected. In addition, only events containing two b-jets are accepted. Moreover, backgrounds from Z+jets processes are further suppressed by vetoing events with 76 $<$ $m_{l\bar{l}}$ <106 GeV.

A total of 96 events pass the selection criteria out of 23556 candidate events.



(a) $m_{l\bar{l}}$ invariant mass before cut

(b) $l\bar{l}$ invariant mass after cut of $76 > m_{l\bar{l}}$ or $m_{l\bar{l}} > 106$ GeV

Figure 7: Invariant mass of the dilepton system cut

5 Mass determination

The relativistic Breit-Wigner is a continuous probability distribution and its probability density function is [11],

$$f(E) = \frac{k}{(E^2 - M^2)^2 + M^2\Gamma^2} \tag{1}$$

where k is a constant of proportionality,

$$k = \frac{2\sqrt{2}M\Gamma\gamma}{\pi\sqrt{M^2 + \gamma}} \tag{2}$$

with

$$\gamma = \sqrt{M^2(M^2 + \Gamma^2)} \tag{3}$$

The relativistic Breit-Wigner is used to model unstable particles, such as the top quark, in high-energy physics. In this case, E represents the center-of-mass energy of the particle, i.e. its mass for a non-composite particle. M is the mass of the particle, and Γ is the decay width. For non-composite particles since in the center-of-mass frame of reference, p in $E^2 = m^2 + p^2$ is always equal to 0, the center-of-mass energy is equal to mass.

This paper employs a maximum likelihood expectation of a relativistic Breit-Wigner distribution to model the top quark mass distribution, and consequently arrive at a mass measurement for the top quark.

The Python package SciPy is used to find the relativistic Breit-Wigner distribution that models the observed mass distribution of the top quark as closely as possible. A curve fitting algorithm, CURVE_ FIT, is used to fit the relativistic Breit-Wigner function to the data by taking an independent variable as the first argument, which in this case is the generator level top quark masses, along with other parameters, including the mass of the resonance, decay width of the resonance along with a normalization factor. The algorithm returns the array POPT which contains the fit results for the parameters, and a $2d$ array PCOV which is a covariance matrix and can be used to represent the variance of the fitted parameters. The former includes the measurement for m_t while the latter is essentially the uncertainty found for the measurement.

6 Results and analysis

Figure 8 and Figure 9 show the distribution of the observed top quark mass without and with event selection, respectively. The relativistic Breit-Wigner is fit on these distributions of m_t.

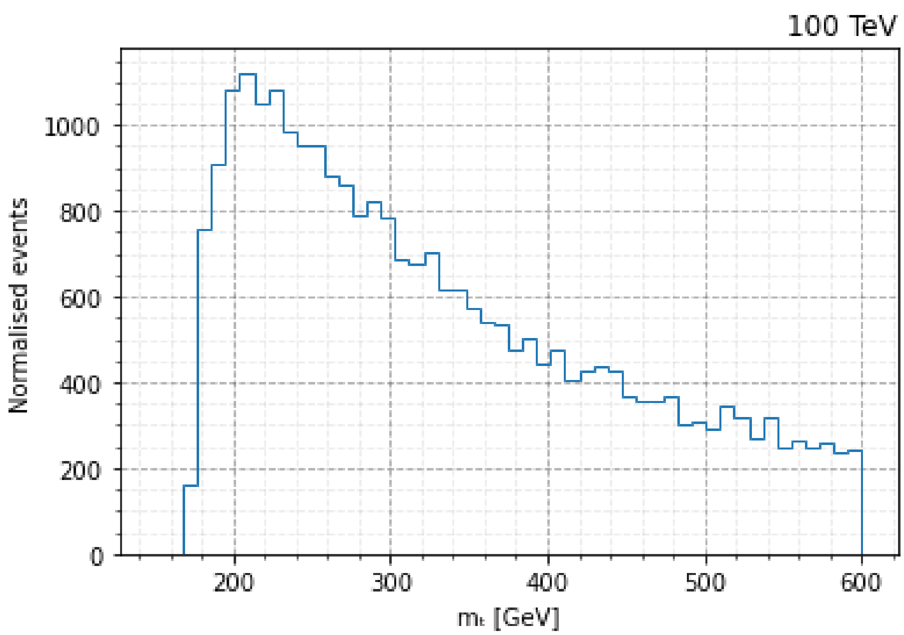


Figure 8: Distribution of m_t with no event selection

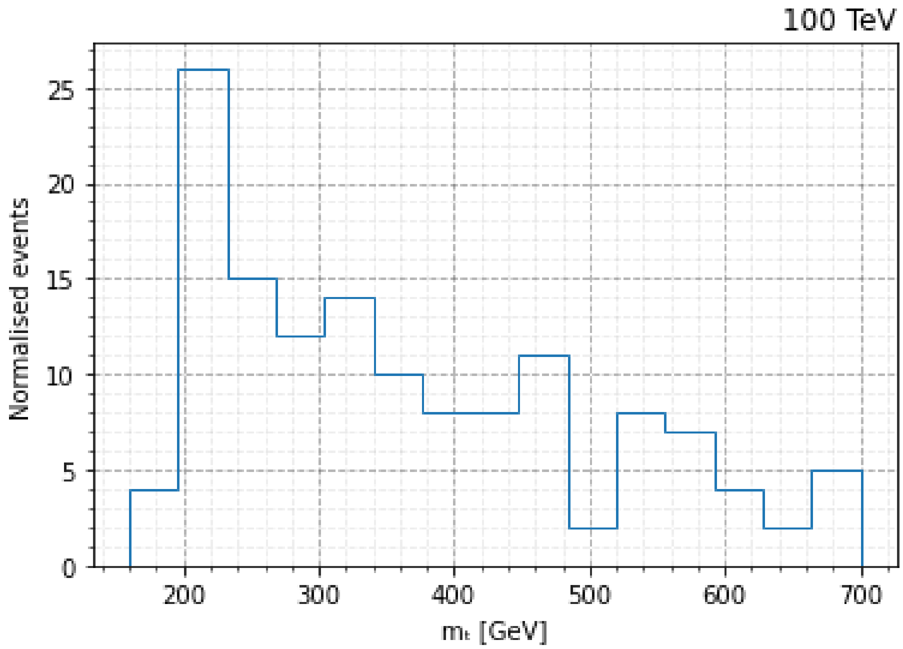


Figure 9: Distribution of m_t with event selection

The results of the curve fitting are shown in Figure 10. The mass of the top quark without any event selection is measured to be 252.48 ± 86.00 GeV, while the mass of the top quark with event selection is found to be 281.24 ± 103.10 GeV.

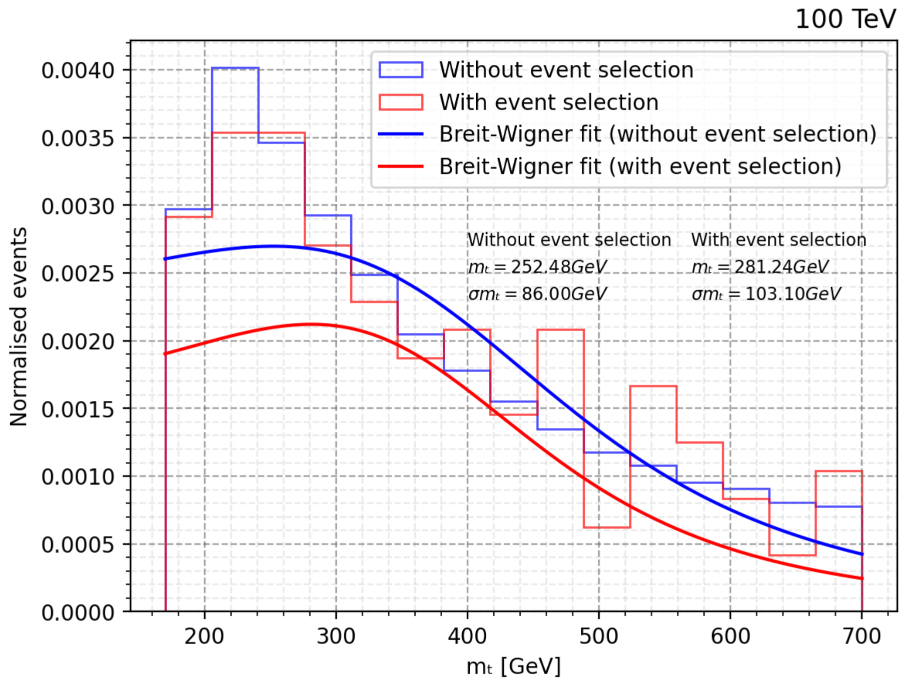


Figure 10: Plot showing the observed mass distributions with the relativistic Breit-Wigner

Figure 11 is a summary plot of the results of this paper and clearly shows that the measurement for m_t are indistinguishable within uncertainty, and hence, event selection does not play a significant role in top quark mass measurements at the FCC-hh. The large relative uncertainty may be attributed to the fact that only a small number of events are studied.

Indeed, comparing the results of the paper, with the most precise top quark mass measurement by the CMS collaboration of 172.33 ± 0.70 GeV [1], we see that the two are indistinguishable within uncertainty (and vary by only 1.1σ).

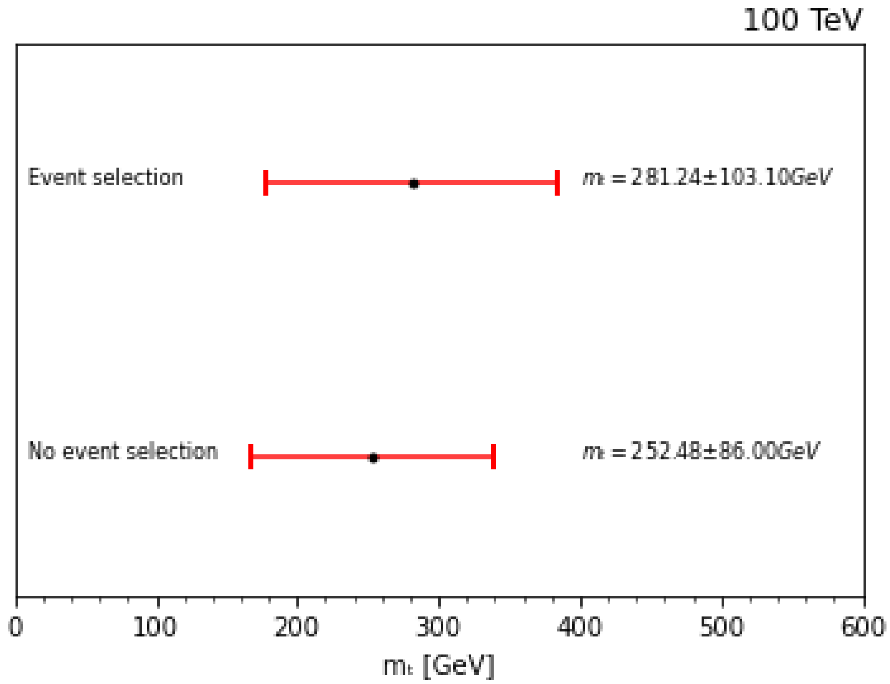


Figure 11: Summary plot showing the extracted value of m_t with and without event selection

7 Limitations and discussion

There exist numerous limitations in this measurement of m_t. Perhaps the most prominent is that in this paper we have plotted the energy of the top quark rather than the mass. This is why we have such a large relative uncertainty and a measurement of m_t that varies significantly from mass measurements at the Tevatron and the LHC.

That being said, plotting the mass of the top quark instead results in a mass distribution that has a width of 0. This is because the MC uses the zero-width approximation to generate simulated events at the energy scale of the FCC-hh. Hence, a relativistic Breit-Wigner cannot be fit to the observed m_t distributions. Although such an approximation may be considered worthwhile, since it is mainly event reconstruction and not event selection that will possibly influence the mass of the measured top quark, a zero-width approximation does bias results. Moreover, since the precision of m_t at the LHC, and possibly at the FCC-hh too, is considerably smaller than the width of the top quark, a non-zero width may be needed in order to see the most precise possible measurement of

m_t that could be carried out. This problem could be worked around by generating top quark masses through the sampling of a relativistic Breit-Wigner while setting the resonant mass at 173 GeV and choosing a width of 1.4 GeV.

The method described to extract the mass in this paper is quite simple, and hence for more accurate results, alternative and more advanced methods such as the template method could be used. In this paper, Delphes is used to carry out detector simulation. However, using GEANT4 FastSim or FullSim for detector simulation may provide better results for this particular projection study. Indeed, better Monte Carlo is needed for a more complete and realistic study of the FCC-hh.

References

[1] "A profile likelihood approach to measure the top quark mass in the lepton+jets channel at $\sqrt{s} = 13$ TeV". In: (2022).

[2] T. Aaltonen et al. "High-precision measurement of the W boson mass with the CDF II detector". In: *Science* 376.6589 (2022), pp. 170–176. DOI: 10.1126/science.abk1781. eprint: https://www.science.org/doi/pdf/10.1126/science.abk1781. URL: https://www.science.org/doi/abs/10.1126/science.abk1781.

[3] Giorgio Cortiana. "Top-quark mass measurements: Review and perspectives". In: *Reviews in Physics* 1 (2016), pp. 60–76. ISSN: 2405-4283. DOI: https://doi.org/10.1016/j.revip.2016.04.001. URL: https://www.sciencedirect.com/science/article/pii/S2405428316300028.

[4] David d'Enterria, Krisztián Krajczár, and Hannu Paukkunen. "Top-quark production in proton–nucleus and nucleus–nucleus collisions at LHC energies and beyond". In: *Physics Letters B* 746 (2015), pp. 64–72. ISSN: 0370-2693. DOI: https://doi.org/10.1016/j.physletb.2015.04.044. URL: https://www.sciencedirect.com/science/article/pii/S0370269315002877.

[5] Michele Gallinaro. "Top quark physics: A tool for discoveries". In: *Journal of Physics: Conference Series* 447 (July 2013), p. 012012. DOI: 10.1088/1742-6596/447/1/012012. URL: https://doi.org/10.1088/1742-6596/447/1/012012.

[6] LHC Top Physics Working Group. *ATLAS and CMS Top Quark Mass Measurement Summary Plots.* 2022 2022. URL: https://twiki.cern.ch/twiki/pub/LHCPhysics/TopMassHistory/LHC_topmass_jun22.pdf.

[7] André H. Hoang. "What Is the Top Quark Mass?" In: *Annual Review of Nuclear and Particle Science* 70.1 (2020), pp. 225–255. DOI: 10.1146/annurev-nucl-101918-023530. eprint: https://doi.org/10.1146/annurev-nucl-101918-023530. URL: https://doi.org/10.1146/annurev-nucl-101918-023530.

[8] Ulrich Husemann. "Top-quark physics: Status and prospects". In: *Progress in Particle and Nuclear Physics* 95 (2017), pp. 48–97. ISSN: 0146-6410. DOI: https://doi.org/10.1016/j.ppnp.2017.03.002. URL: https://www.sciencedirect.com/science/article/pii/S0146641017300339.

[9] Kazunori Kohri and Hiroki Matsui. "Electroweak vacuum instability and renormalized vacuum field fluctuations in Friedmann-Lemaitre-Robertson-Walker background". In: *Phys. Rev. D* 98 (10 Nov. 2018), p. 103521. DOI: 10.1103/PhysRevD.98.103521. URL: https://link.aps.org/doi/10.1103/PhysRevD.98.103521.

[10] Michelangelo Mangano et al. *FCC Physics Opportunities: Future Circular Collider Conceptual Design Report Volume 1. Future Circular Collider*. Tech. rep. Geneva: CERN, Dec. 2018. DOI: 10.1140/epjc/s10052-019-6904-3. URL: https://cds.cern.ch/record/2651294.

[11] Torbjörn Sjöstrand, Stephen Mrenna, and Peter Skands. "PYTHIA 6.4 physics and manual". In: *Journal of High Energy Physics* 2006.05 (May 2006), pp. 026–026. DOI: 10.1088/1126-6708/2006/05/026. URL: https://doi.org/10.1088%2F1126-6708%2F2006%2F05%2F026.

[12] Tevatron Electroweak Working Group. *Combination of CDF and D0 results on the mass of the top quark using up* $9.7\,\mathrm{fb}^{-1}$ *at the Tevatron*. 2016. DOI: 10.48550/ARXIV.1608.01881. URL: https://arxiv.org/abs/1608.01881.