Using ePump to Understand Future Top Quark Mass Measurements

Jason Gombas
Reinhard Schwienhorst
Sara Sawford
Jarrett Fein

https://arxiv.org/abs/2203.08064

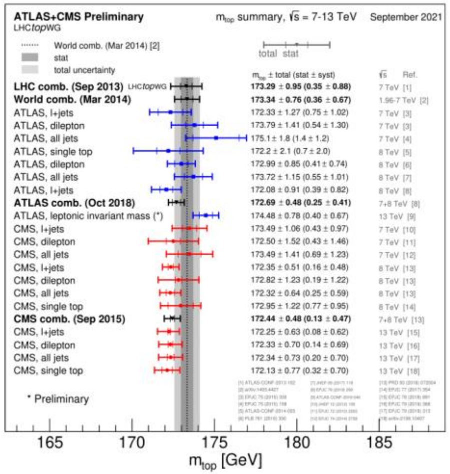
Importance of the top quark mass measurement

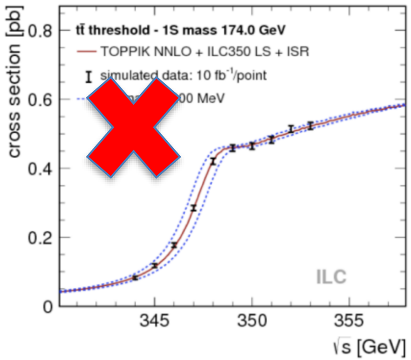
- This is important to understand if vacuum is stable or unstable
  - If there is no new physics up to very high scales, then the vacuum itself might not be stable
  - Determines the fate of the universe
- We can constrain Standard Model parameters by comparing top, W, and Higgs boson mass measurements.
- Perform precision electroweak fits to probe electroweak symmetry breaking

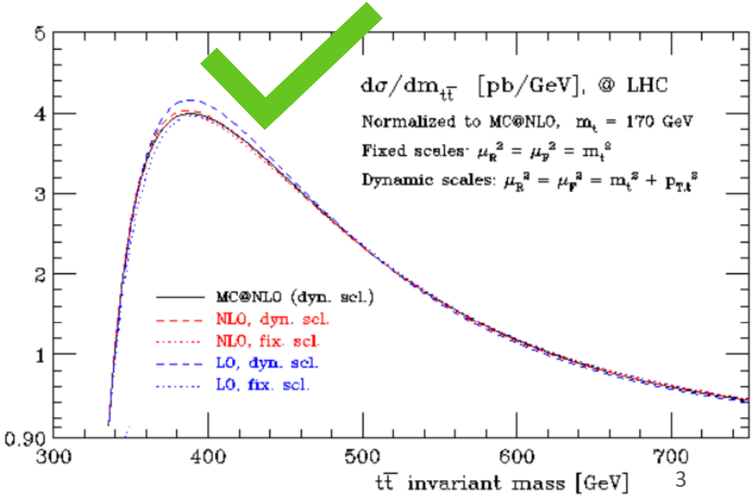
How to measure top quark mass

- Direct measurement of the decay products of the top quark (not well understood)
- Scanning through beam energies is not possible with proton-proton beam, so reconstruct top-pair invariant mass (very well understood theoretically)

Direct top mass measurement





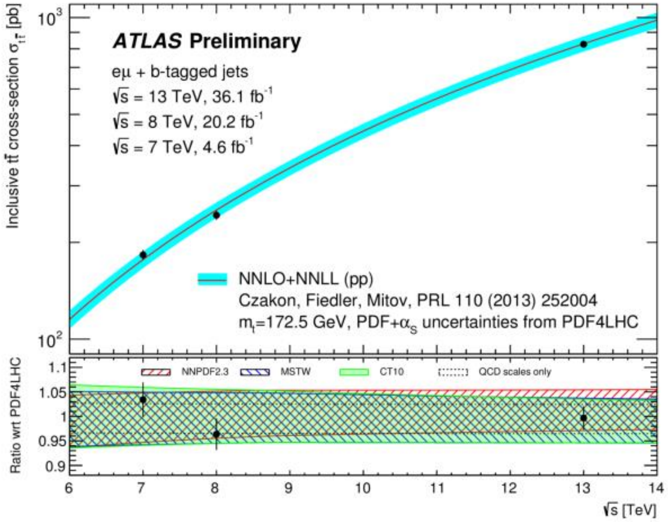


An introduction to our study

- Generate ttbar events at NLO using Madgraph
  - Obtain PDF weights for CT18NLO PDF set from Madgraph
  - Don’t decay the top, and look at the best-case scenario
- Calculate χ^2 of $m^{t\bar{t}}$ between nominal and different mass points using PDF uncertainties from CT18 as the only uncertainty
- Perform an update to the global PDF fit using $p_z^{t\bar{t}}$ and use the resulting $m^{t\bar{t}}$ PDF uncertainties to calculate χ^2 of $m^{t\bar{t}}$

Why incorporate Parton Distribution Functions?

- Most recent top pole mass studies have the highest contribution to their overall uncertainty being the PDF uncertainty.
  - About 5% uncertainty on the total cross-section
  - Gluon PDF at large x and large scale μ



A brief introduction to ePump

- ePump is a tool that allows the user to see the impact that new data will have on PDF sets without performing the large global fit
  - ePump runs within seconds compared to a global fit which takes much more time (several hours at least)
  - ePump assumes eigen directions don’t change, just their amplitudes
- To update the PDFs, you need data files and theory files
- For a particular observable, you need the theory file that contains the calculated observable from the best fit and each error PDF (for us this comes from Madgraph calculations)

Madgraph event generation

- We change the top mass in Madgraph and generate 10 million events for each top mass
- Data sets were generated with top masses from 171 GeV to 174 GeV in 0.25 GeV increments
- We then do this for proton-proton beam energies of:
  - 8 TeV – To compare with previous studies
  - 13 TeV – To benchmark what is possible with Run 2 data
  - 13.6 TeV – To see what is possible with Run 3 data
  - 14 TeV – To see what is possible with the high luminosity LHC
  - 100 TeV – To see what is possible with the FCC-hh

ePump input

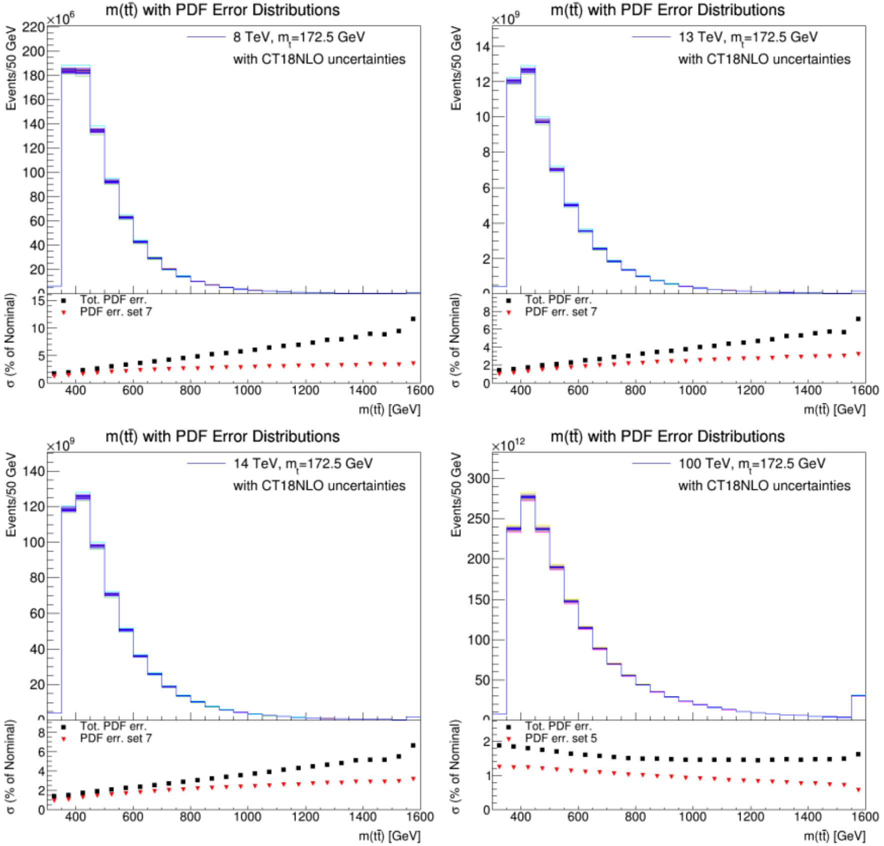
- To run ePump, we must reformat generated plots from Madgraph into ePump .theory and .data files
- Theory is set to nominal 172.5 GeV top mass distribution
  - “Nominal” corresponds to the PDF best-fit
- Data (in our case pseudo-data) is set differently for each mass point
  - Each pseudo-data distribution is the nominal theory distribution of that mass point
  - Here we set statistical error and correlated systematic error to 0
  - Correlated systematic error is set to 1%; future studies will change this to 5%, 10% and 15%

ePump output

- ePump will output many things including the updated PDFs
- We are interested in the updated PDF error of the $m^{t\bar{t}}$ distribution
- For each top mass, we run ePump and extract the updated PDF errors
- These updated uncertainties then increase the calculated χ^2 values which constrain the top quark mass

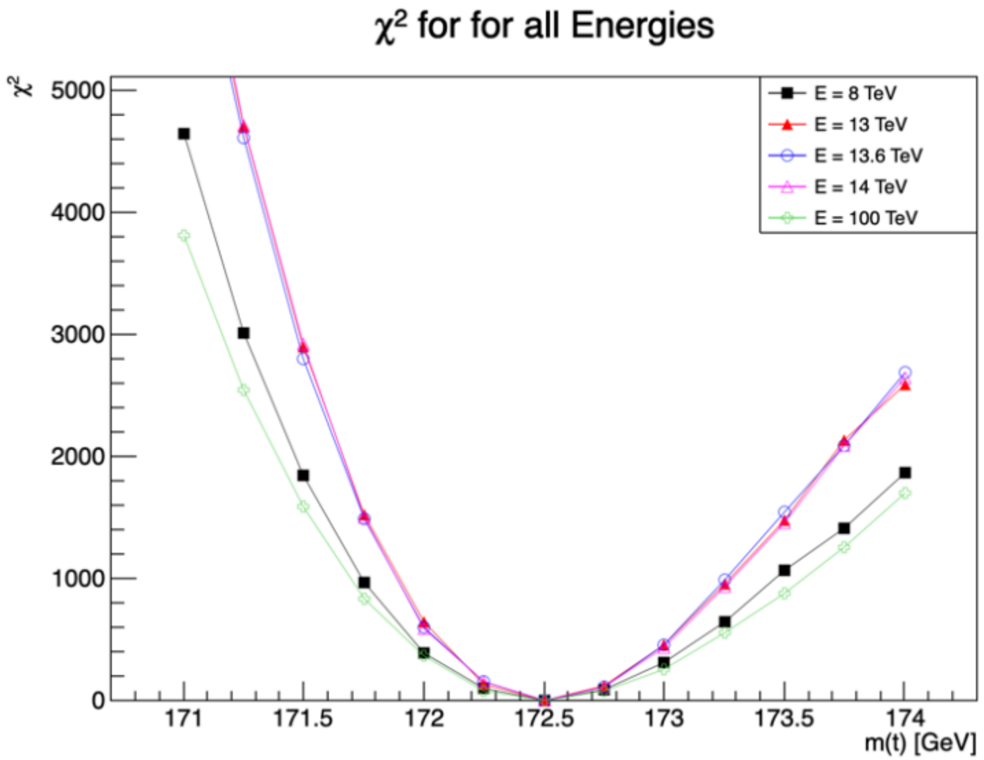
Pseudo Data - $m^{t\bar{t}}$

- Shown here are the $m^{t\bar{t}}$ for CM energy of 8 TeV, 13 TeV, 14 TeV, and 100 TeV
- The PDF sensitivity is mainly at high $m^{t\bar{t}}$, whereas top mass sensitivity is mainly at the turn-on region at about ~350 GeV
- The 100 TeV distribution is quite different than the other lower CM beam energies



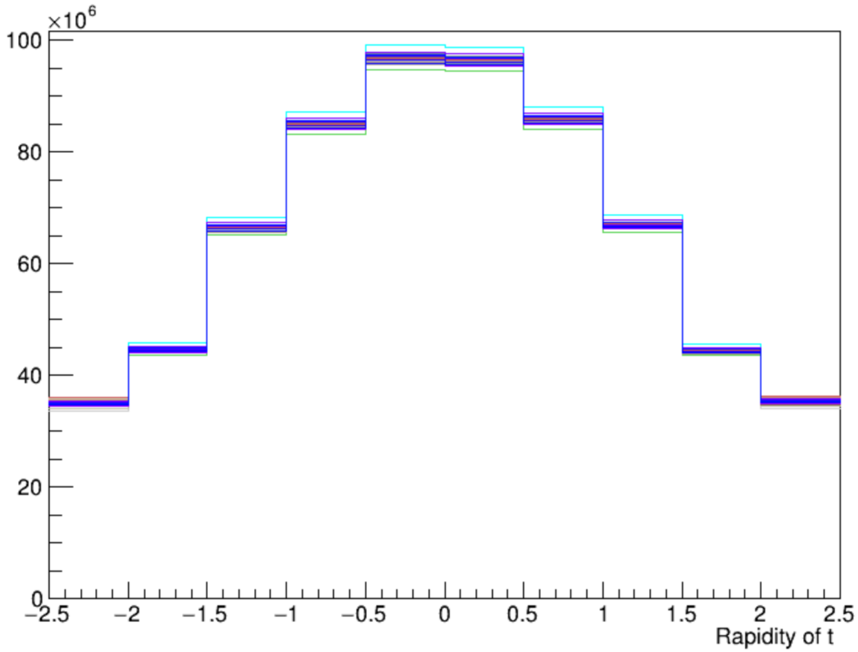
χ^2 Calculations

- χ^2 is calculated by taking $m^t = 172.5$ GeV to be 'expected' and every other mass to be 'observed'
- The uncertainty in each bin is set to the PDF uncertainty
- Most of the χ^2 comes from the first bins



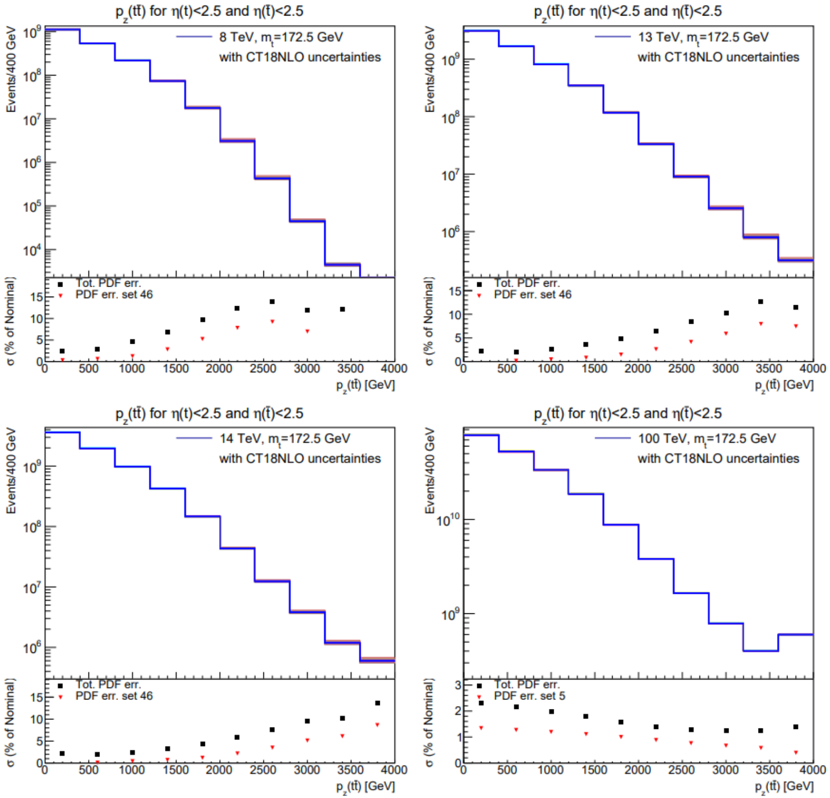
Rapidity distribution - η^t

- Here are the η^t distributions for beam energy of 13 TeV with 172.5 GeV top mass
  - There are 59 histograms overlayed here, 58 error histograms and 1 nominal (best fit) histogram
- Rapidity can typically only be experimentally measured up to ~2.5



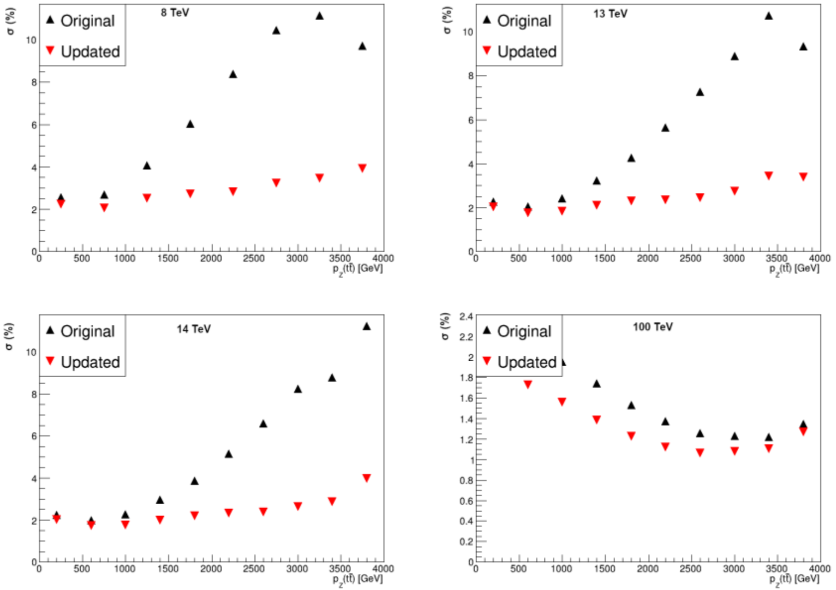
Pseudo Data – $p_z^{t\bar{t}}$ with small rapidity

- Create a cut on the rapidity and use the resulting $p_z^{t\bar{t}}$
- $p_z^{t\bar{t}}$ distributions were included in the global fit update using ePump

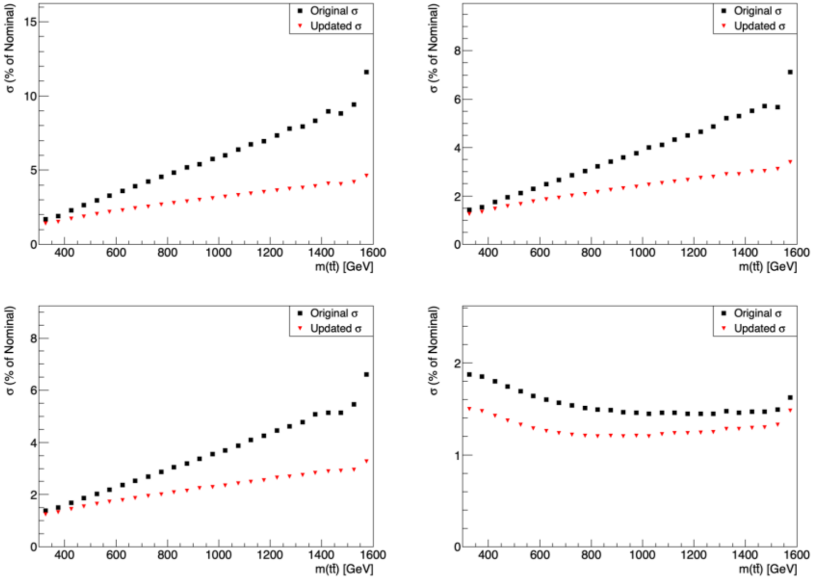


Updated PDF errors from ePump

$p_Z^{t\bar{t}}$ Relative Errors

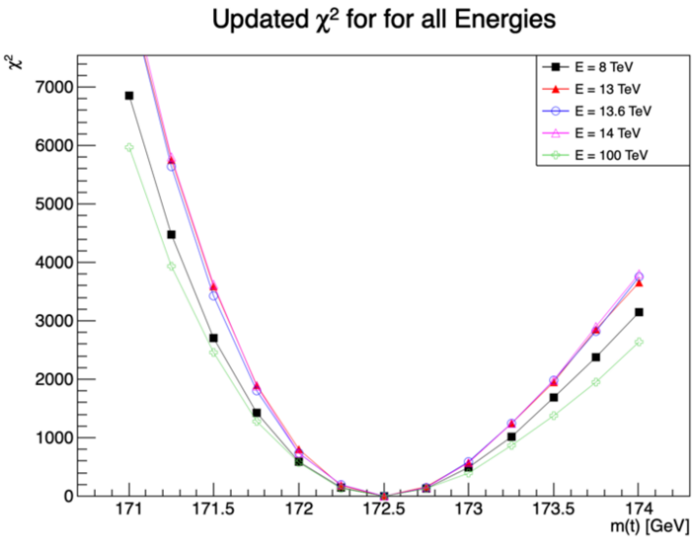


$m^{t\bar{t}}$ Relative Errors



Updated χ^2 Calculations

- The updated χ^2 calculations have the same general shape, but have increased leading to a smaller top mass uncertainty
- The result of including the $p_z^{t\bar{t}}$ distribution in the global PDF fit leads to a decrease in the mass uncertainty by 20%



Outlook

- A few things remain under consideration:
  - How do the results change if we increase the pseudo-data systematic error to be 5%, 10% or 15%?
  - Smear the data to replicate previous studies at 8 TeV, and then use the same smearing parameters to extrapolate to higher CM energies
    - Currently in progress

Executive Summary

This whitepaper proposes a way to reduce the top quark mass with upcoming improvements to the LHC and the future FCC-hh collider. Since the largest uncertainty of the top quark mass comes from the PDF uncertainty, top quark mass measurements can be improved by simultaneously updating the PDF best fit while fitting the top quark mass. This uncertainty of the top quark can be reduced by up to 20% by doing this.

Backup

Variables that can be changed

- Observable in the fit:
  - $m^{t\bar{t}}$, $p_z^{t\bar{t}}$, η^t, 2D fit
- Number of bins
- Bin placement
- Uncertainty in each bin
- Variance of the smearing