Using ePump to Understand Future Top Quark Mass Measurements

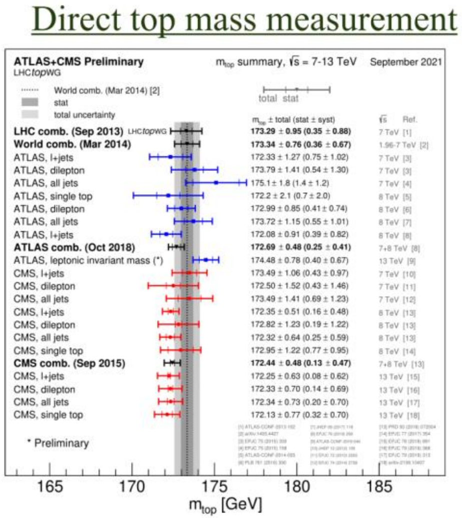
Jason Gombas
Reinhard Schwienhorst
Sara Sawford
Jarrett Fein

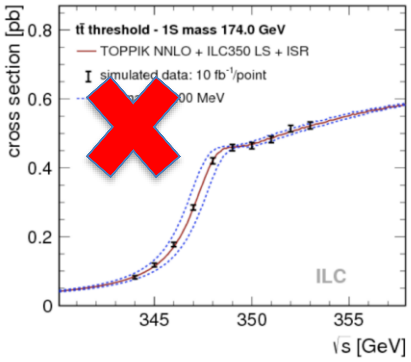
Importance of the top quark mass measurement

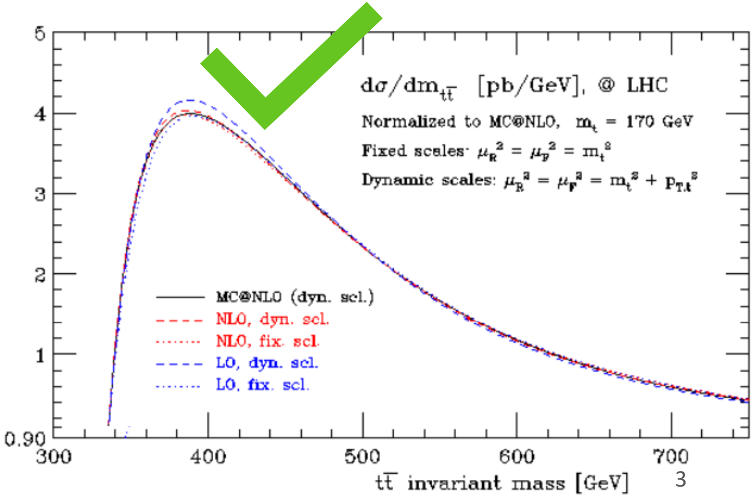
- This is important to understand if vacuum is stable or unstable
  - If there is no new physics up to very high scales, then the vacuum itself might not be stable
  - Determines the fate of the universe
- We can constrain Standard Model parameters by comparing top, W, and Higgs boson mass measurements.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





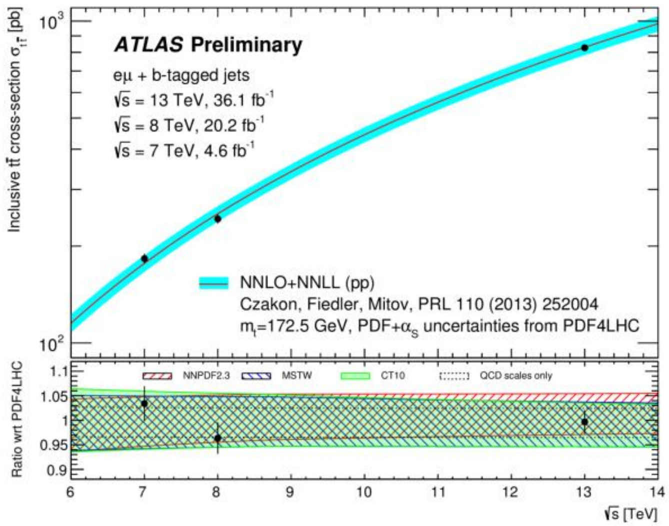


An introduction to our study

- 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
  - We will ‘smear’ the top distributions to at least approximately reproduce ATLAS and CMS extractions from their differential cross-section measurements
- Run calculations and obtain total χ^2 from global PDF data and Madgraph pseudo data
- Then adjust smearing to study potential detector and reconstruction improvements

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 a few million events for each top mass
- We then do this for proton-proton beam energies of:
  - 4000 GeV – To verify with previous studies
  - 6500 GeV – To benchmark what is possible with Run 2 data
  - 6800 GeV – To see what is possible with Run 3 data
  - 7000 GeV – To see what is possible with the high luminosity LHC
  - 50000 GeV – To see what is possible with the FCC-hh

ePump input

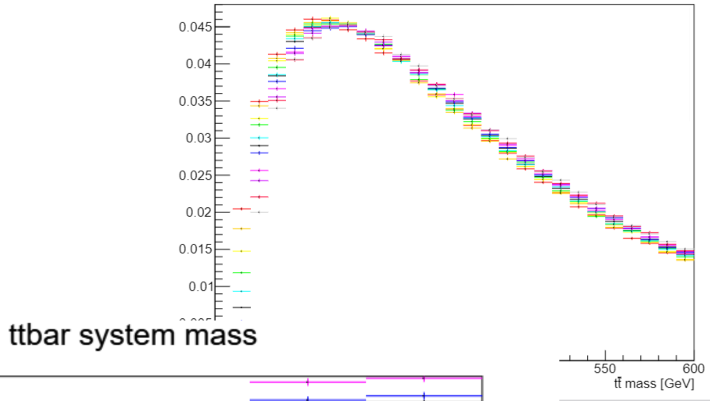
- To run ePump, we must reformat generated plots from Madgraph into ePump .theory and .data files
- .theory files set to nominal 172.5 GeV top mass
- .data files set differently for each mass point
  - Here we set statistical error and correlated systematic error to 0
  - We then vary uncorrelated systematic error to try and match results from 8 TeV study and approximate experimental uncertainties in the differential distributions
    - [171 GeV, 172 GeV, 172.5 GeV, 173 GeV, 174 GeV]

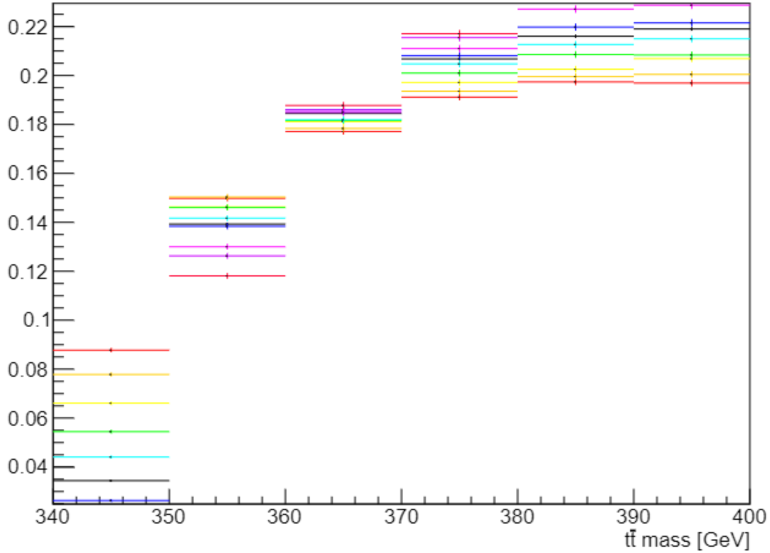
ePump output

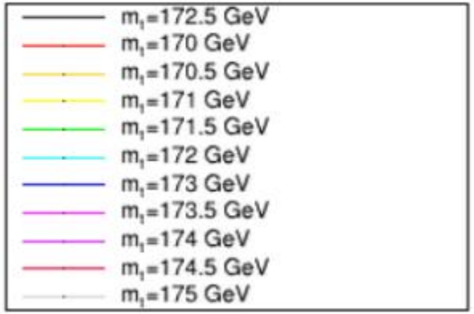
- ePump will output many things including the updated PDFs
- We are interested in the total χ^2 for the original global analysis plus our pseudo data set
- As we scan through top pole masses, the χ^2 will form a parabolic curve and we can estimate the top mass uncertainty by looking at when this value increases by T^2
- In ePump, χ^2 is calculated by: $\chi^2 = T^2 \sum_{i=1}^{N} z_i$, where z_i are the N parameters of the PDF set
  - z_i=0 corresponds to the global best fit
- $T = 1$ for data errors that are precisely Gaussian and internally consistent, $T > 1$ to accommodate experimental inconsistencies. CTEQ-TEA has used $T = 10$ in their analyses to date

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

- Masses of 170.0 GeV to 175 GeV in steps of 0.5 GeV are plotted
- The variations are found in the early few bins, after the peak region, the fluctuations are dominated by statistics



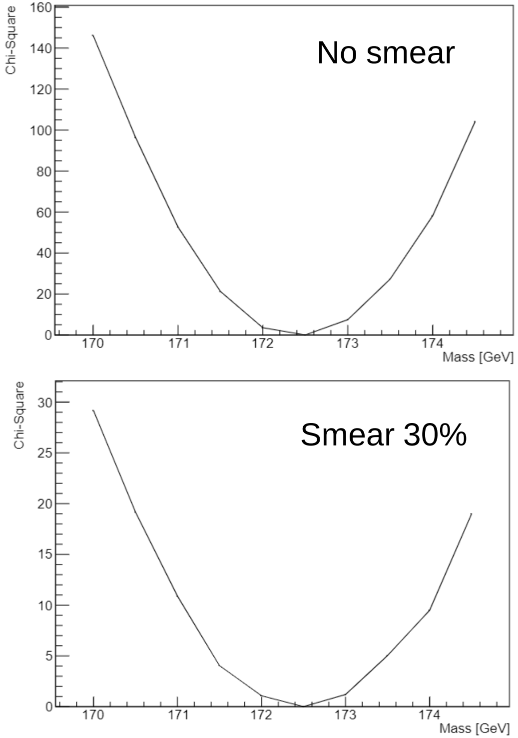




Alternative χ^2 calculation

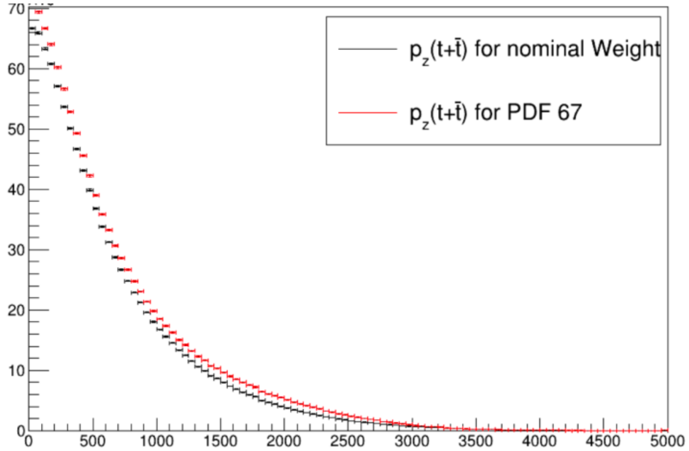
- Take the nominal 172.5 GeV $m^{t\bar{t}}$ distribution and calculate χ^2 from the nominal $m^{t\bar{t}}$ distribution of the other masses with: $\chi^2 = \sum \frac{(x_{theory} - x_{obs})^2}{\sigma_i^2}$
- Does not factor in updated PDF fit
- Smear by multiplying mass by gaussian random number centered at 1 with variance of 0.3
- Should reproduce ePump's original χ^2 calculation and will be used to verify ePump's output

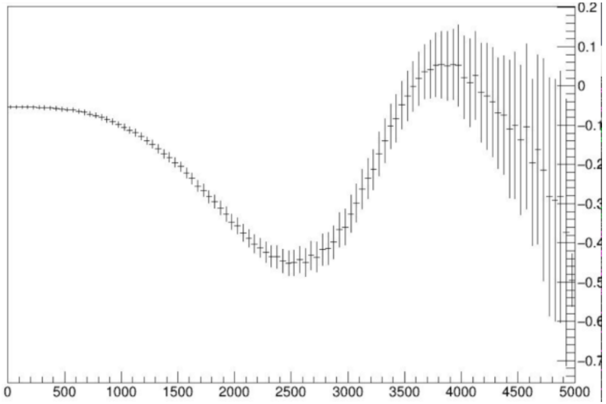
Preliminary Results



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

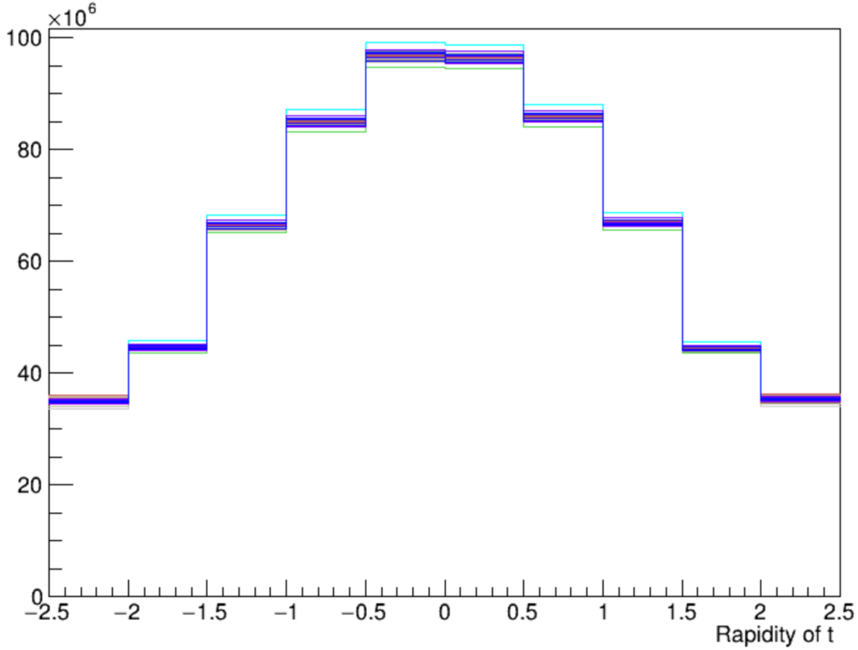
- Here is the nominal $p_z^{t\bar{t}}$ distribution for beam energy of 13 TeV with an example error PDF distribution
- This distributions change when the top mass and beam energy is changed
- Below is the relative difference plot between the nominal and PDF error $p_z^{t\bar{t}}$ distribution





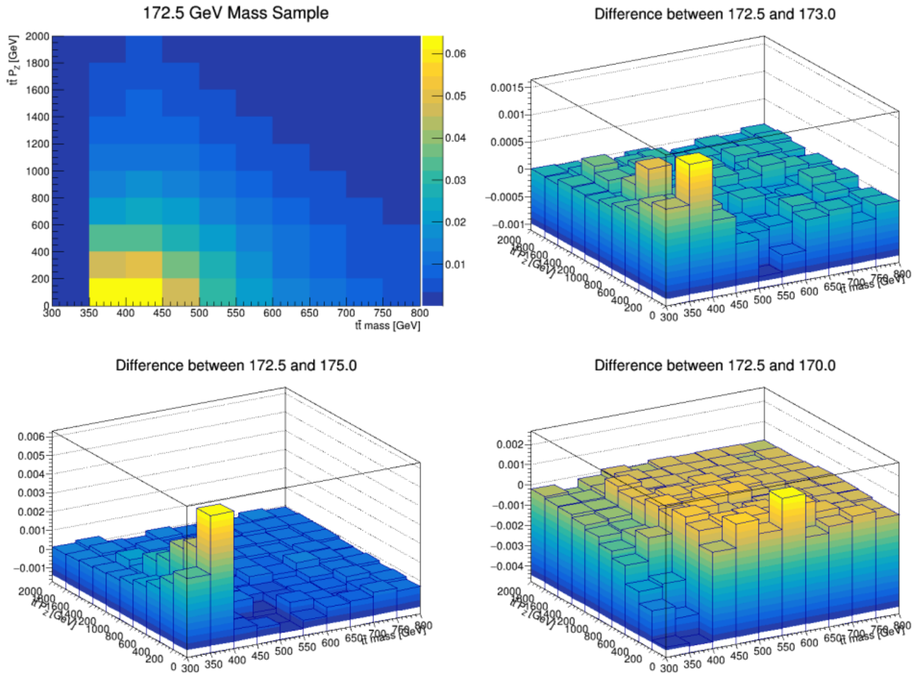
Pseudo Data - η^t

- Here are the η^t distributions for beam energy of 13 TeV with nominal 172.5 GeV top mass
  - There are 59 histograms overlayed here, 58 error histograms and 1 nominal (best fit) histogram
- This distributions change when the top mass and beam energy is changed



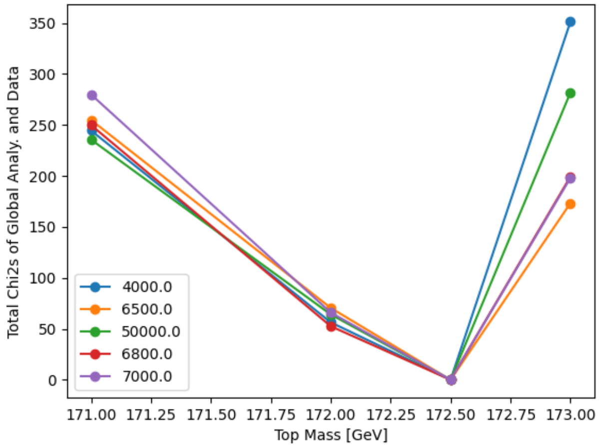
Pseudo Data – 2D Histograms of $p_z^{t\bar{t}}$ and $m^{t\bar{t}}$

- Most differences, as expected, come from the mass 'turn-on' region at about 350-360 GeV
- Motivation for this 2D fit is that $p_z^{t\bar{t}}$ is sensitive to PDF uncertainties and $m^{t\bar{t}}$ is sensitive to top mass



Preliminary Results

- For the 2D plot of $m^{t\bar{t}}$ and $p_z^{t\bar{t}}$ we have calculated the total χ^2 from global PDF data and Madgraph pseudo data for each top pole mass and each beam energy
- We expect to produce this curve for $m^{t\bar{t}}$ as well as rapidity of t
- We are working on understanding the skewness of the χ^2 and we expect that lower energies should have a narrower χ^2 compared to higher energies



Outlook

- We hope to implement smearing to reproduce ATLAS and CMS resolutions at 8 TeV and then scale from there to project to the future
- To this point I have been normalizing the pseudo-data plots, but we think this is removing some PDF information
- All scripts to this point are pretty much automated and so small adjustments can be easily made within the next two weeks
- We are hoping to provide top mass uncertainty projections, extracted from the χ^2 distributions
- We aim to complete note by March 15

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