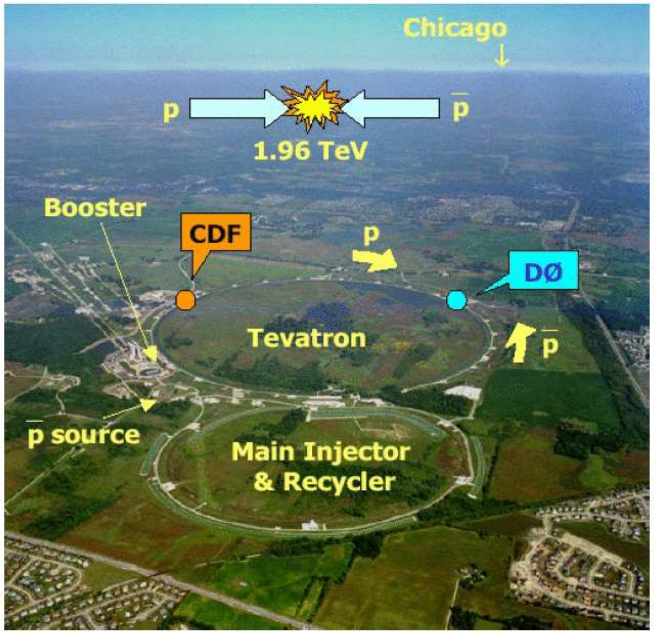
The Multivariate Template Method: A Matrix Element Top Mass Measurement in the ℓ + Jets Channel

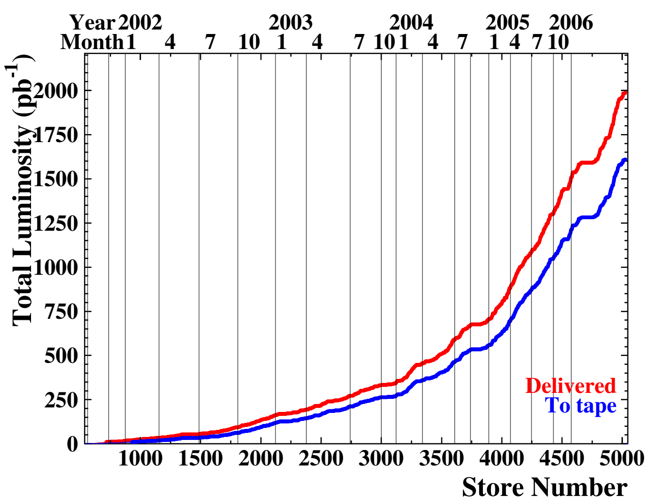
Paul Lujan
UC-Berkeley / Lawrence Berkeley National Laboratory
for the CDF II Collaboration

October 31, 2006

- Experimental setup
- Analysis technique
- Monte Carlo results
- Conclusion & future plans

Fermilab Tevatron





- The Fermilab Tevatron is a $p\bar{p}$ collider with a CM energy $\sqrt{s} = 1.96$ TeV.

- Two large, general-purpose detectors are located at interaction points along the ring: CDF and DØ.

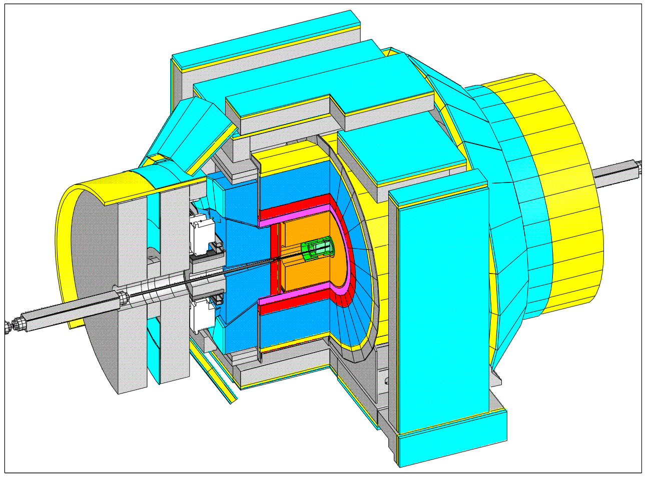
- The top was first discovered here in 1995.

- Both detectors underwent major upgrades for Run II beginning in 2001.

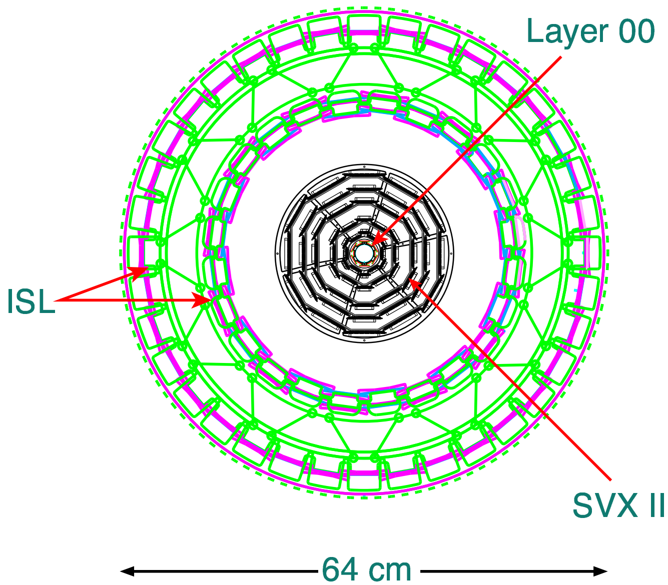
- The Tevatron maximum instantaneous luminosity is $\sim$ 2.2 x 10^{32} cm^{-2} s^{-1}.

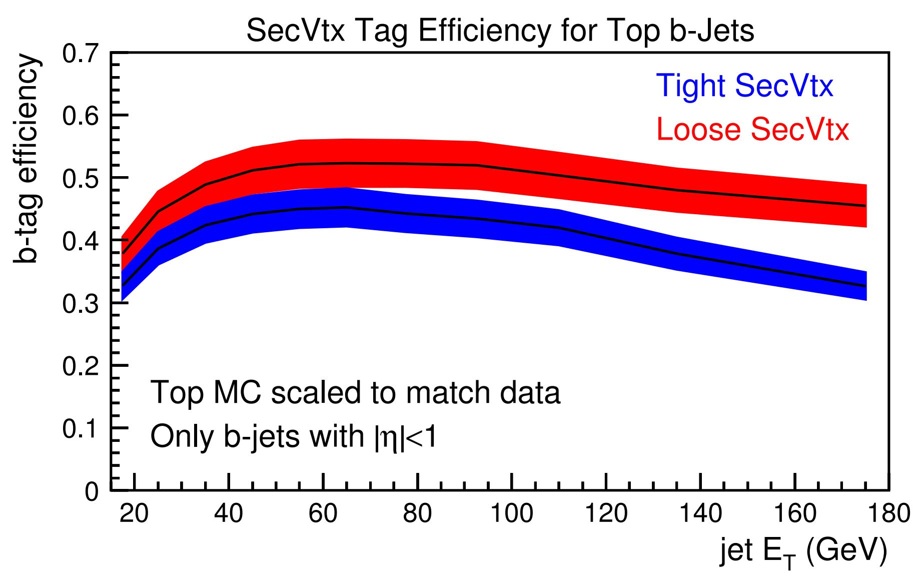
- In Run II, the Tevatron has delivered almost 2 fb^{-1} integrated luminosity; our analysis is based on 1 fb^{-1} of good data taken at CDF.

CDF II Detector



- The CDF II detector contains a silicon vertex detector, a central tracking system, electromagnetic and hadronic calorimeters, and muon chambers.
- The silicon vertex detector (bottom left) is particularly important for high-efficiency b-tagging.





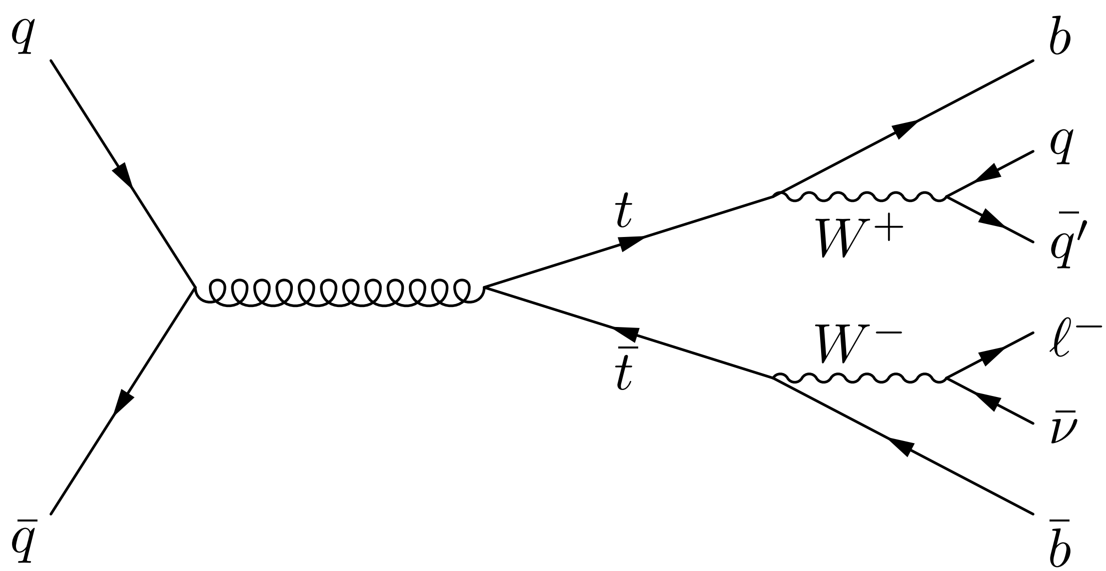
Lepton + Jets

- We search in the "lepton + jets" channel, where one of the W quarks produced decays hadronically and the other leptonically: $t\bar{t} \rightarrow WWb\bar{b} \rightarrow b\bar{b}q\bar{q}'\ell\nu$

- This channel produces the best combination of statistics ($\sim$ 30% of all $t\bar{t}$ events) and sample purity.

- So far, the single best top mass measurements have all come from this channel.

- We look for events with 4 jets, a lepton, and missing energy from the neutrino.



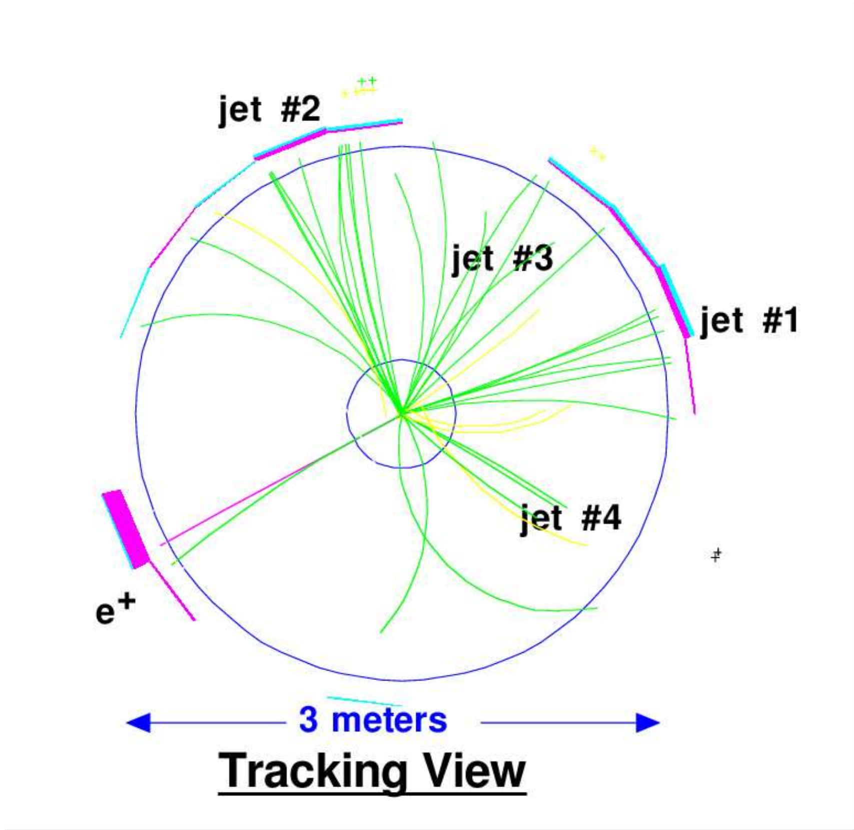
Principal backgrounds:

- W + heavy flavor ($b\bar{b}, c\bar{c}$)
- W + light jets (mistag)
- non-W QCD

Event Selection

Our selection requirements are as follows:

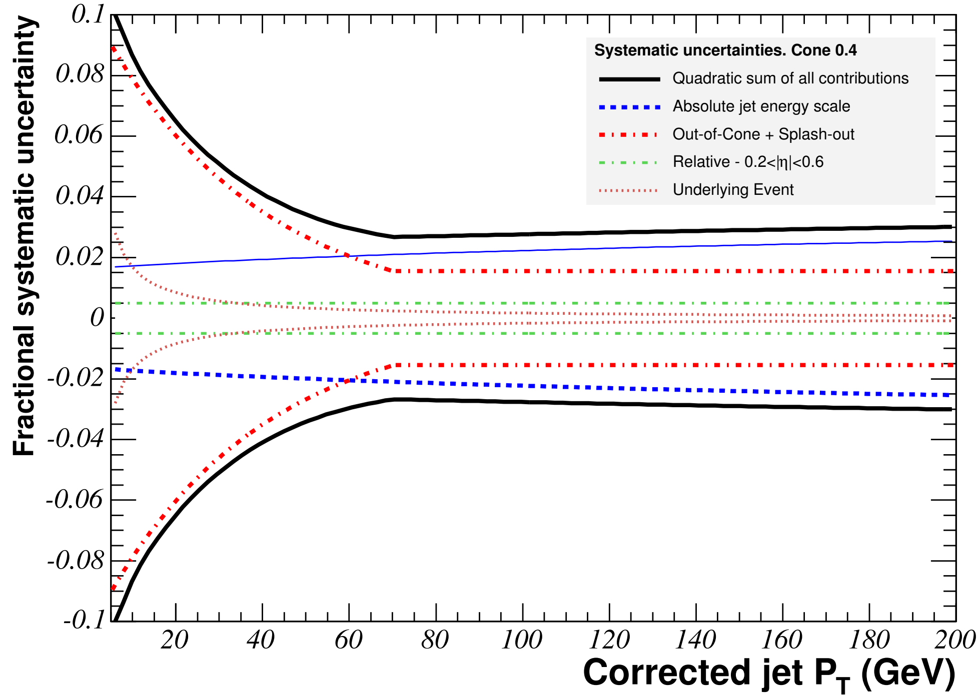
- Exactly one tight lepton with $P_T > 20$ GeV in the central region ($|\eta| < 1$), separated from all jets
- Exactly 4 tight jets with $E_T > 15$ GeV in the central region ($|\eta| < 2$)
- No additional jets with $E_T > 8$ GeV in the central region (to remove initial-state and final-state radiation)
- Missing $E_T > 20$ GeV
- At least one jet tagged as being from a b quark



Sample $t\bar{t} \rightarrow b\bar{b}q\bar{q}'\ell\nu$ event

Jet Energy Systematics

- The uncertainty in measuring the jet energies is the single largest systematic in making a top mass measurement.

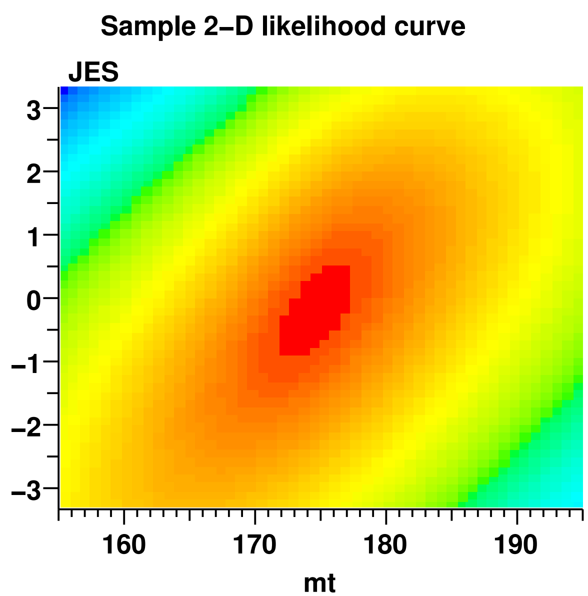


- We introduce the jet energy scale JES, a scale factor to the measured jet energies, as an additional parameter to our likelihood. This allows us to use the information in the W decay to determine the JES in an event.

Method Overview

- Our method is built on calculating a likelihood for seeing the observed event by integrating over the differential cross-section for $t\bar{t}$ events.

- For each event, we build a 2-D likelihood curve $L(\vec{y}|m_t, \mathrm{JES})$ representing the probability of seeing the observed detector-level quantities ($\vec{y}$) as a function of the pole ("true") top mass (m_t) and the jet energy scale JES.



- Then, we combine these curves by multiplying their likelihoods, reduce the likelihood to a function of top mass by using the profile likelihood, and use the peak of the curve to obtain our top mass result and error.

- We do not have results on data yet, so this talk will focus on the method with some Monte Carlo results.

Integration Formula

We build our likelihood by integrating over the unknown quantities:

$$L(\vec{y}|m_t, \mathrm{JES}) = \frac{1}{N(m_t)A(m_t, \mathrm{JES})} \sum_{\mathrm{perms}} \int f(z_1)f(z_2)w(\vec{y}\cdot\mathrm{JES}|\vec{x})|M(m_t, \vec{x})|^2 d\Phi(\vec{x})$$

- $\vec{x}$ are the particle-level momenta, and $\vec{y}$ are the detector-level momenta.
- N is the normalization factor as a function of top mass, and A is the acceptance factor (or "efficiency") as a function of top mass and JES.
- $f(z)$ are the quark/gluon parton distribution functions.
- $w(\vec{y}|\vec{x})$ are the transfer functions connecting the partons and jets, obtained from Monte Carlo.
- M is the matrix element for $t\bar{t}$ production and decay. Our matrix element includes both $q\bar{q}$ and gg as well as full spin correlations.
- $\Phi(\vec{x})$ is the particle-level phase space being integrated over.

Integration Assumptions

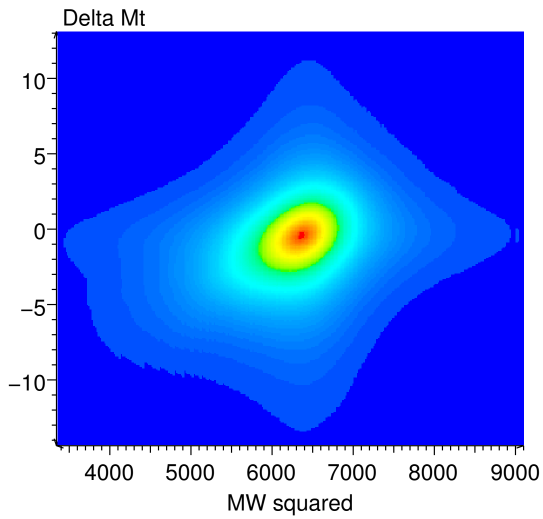
- The full phase space $\Phi(\vec{x})$ has a total of 22 variables, far too many to practically integrate over. Consequently, we make some assumptions: quark angles and lepton momentum are perfectly measured, and all quarks are on mass shell, except the leptonic b, which is massless.

- This leaves us with seven integration variables: M_W^2 and M_t^2 on the hadronic side, M_W^2 and M_t^2 on the leptonic side, $\beta = \log \frac{p_q}{p_{\bar{q}}}$ (the logarithm of the ratio of the momenta of the two hadronic W decay products), and the P_T of the $t\bar{t}$ system (two variables).

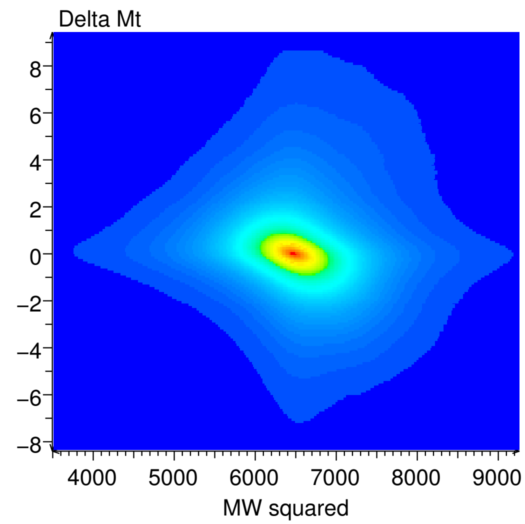
- However, these assumptions require changes to the distributions of M_W^2 and M_t^2 that we integrate over – they are no longer simply Breit-Wigners.

- To compensate, we build "effective propagators" by reconstructing Monte Carlo events so that they do adhere to our assumptions and using those to build distributions of M_W^2 and M_t^2.

Effective Propagators



Hadronic-side propagator



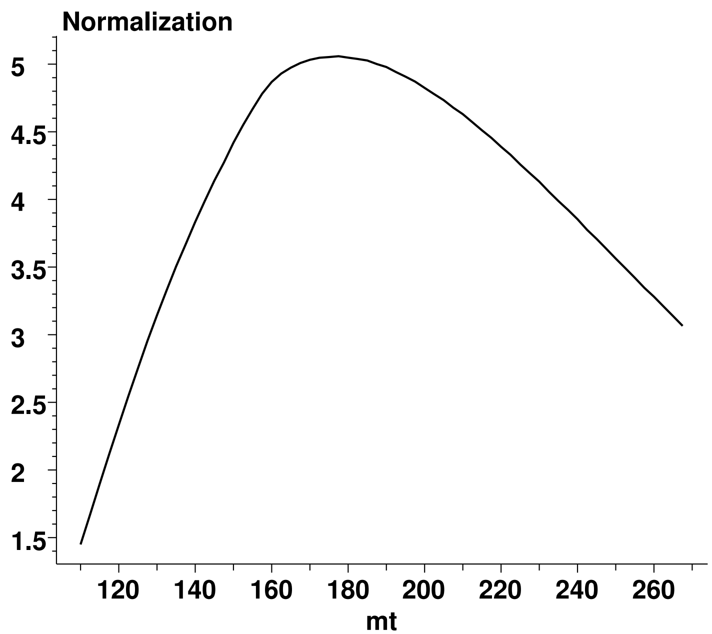
Leptonic-side propagator

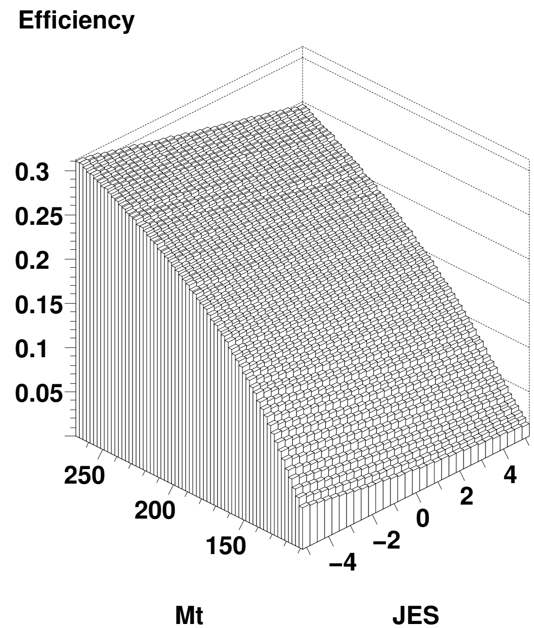
- The propagators are currently built in M_W^2 and ΔM_t (the difference between the pole top mass and the event top mass). ΔM_t is on the y-axis and M_W^2 on the x-axis.

- Note that the propagators are much broader than they would be if they were merely Breit-Wigners.

Normalization and Acceptance

In order to obtain sensible results, we must ensure that the likelihood is normalized, and ensure that the effect of our selection cuts is taken into account.



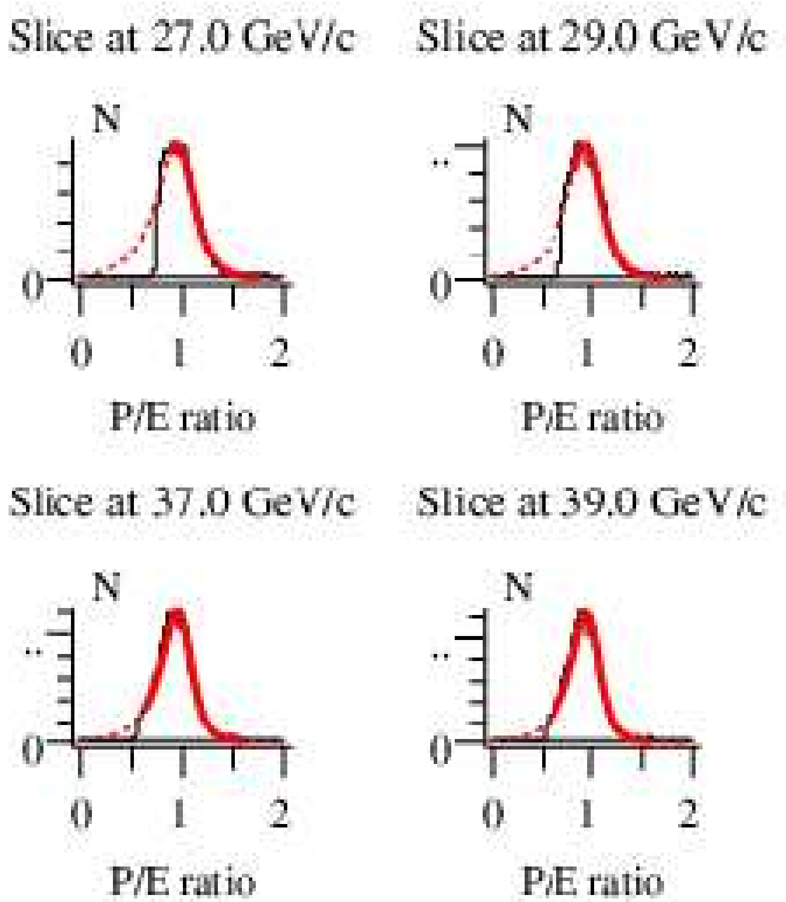


- Left: Normalization, obtained by imposing the condition $\int P(\vec{y}, m_t) d\vec{y} = 1$ for all m_t. Analytically, the normalization comes out proportional to $\sigma_{t\bar{t}} \cdot \Gamma_t^2 / m_t^2$.

- Right: Acceptance, obtained by computing the number of events from Monte Carlo samples passing our selection cuts as a function of m_t and JES.

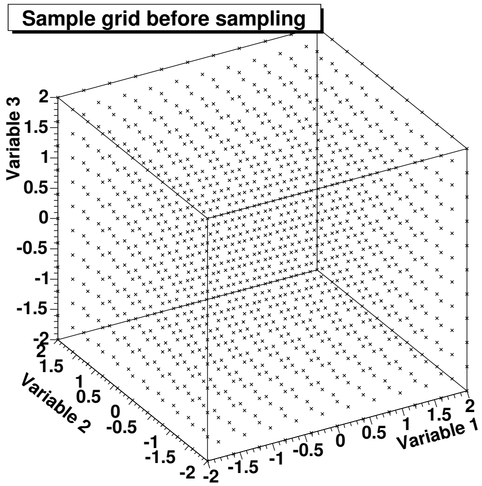
Transfer Functions

- The transfer functions $w(\vec{y}|\vec{x})$ are a crucial component of any matrix-element based analysis. They give the probability of seeing a reconstructed jet with momentum $\vec{y}$ given a parton with momentum $\vec{x}$.

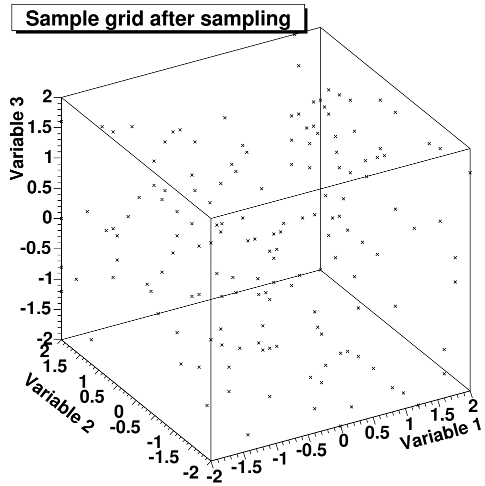


- The transfer functions are built by matching jets to partons in Monte Carlo events as a function of the ratio of parton energy to jet momentum, fitted, and binned in 4 different η bins separately for b and light quarks.
- At left: Sample fitted transfer functions for central light quarks in various P_T bins.
- Dashed lines indicate fit extrapolation below our momentum cutoff from event selection.

Integration Procedure



before



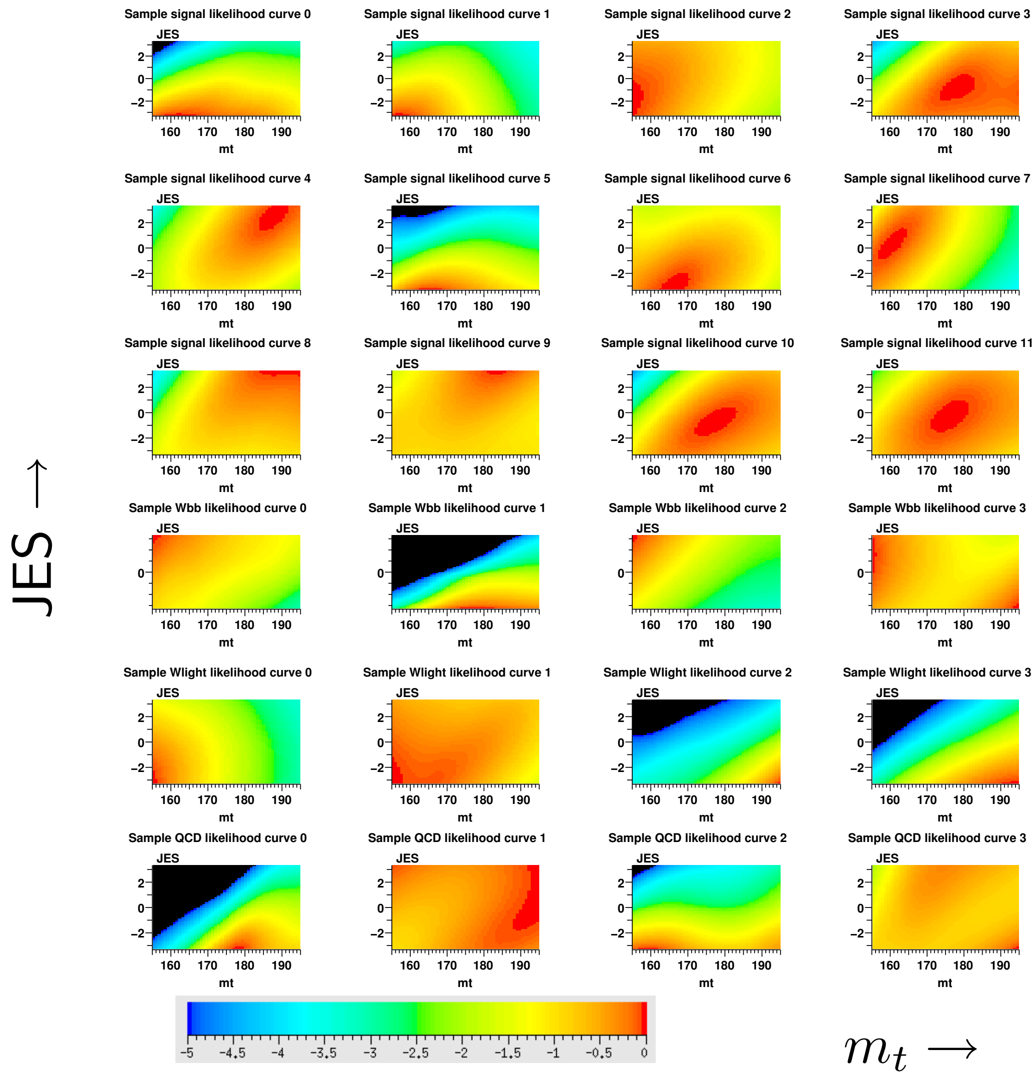
after

- To integrate, we create a 1-D grid equidistant in probability for each integration variable, combine the grids, and then quasi-randomly sample the resulting grid (similar to Monte Carlo integration).

- We sum over the possible jet-parton permutations with each permutation weighted by the appropriate tagging probabilities. For example, if a tagged jet is matched to a b parton, a weight of P(tag) is used. Conversely, if a b parton is matched to an untagged jet, a weight of $1 - P$(tag) is used. This is done for each jet in the event.

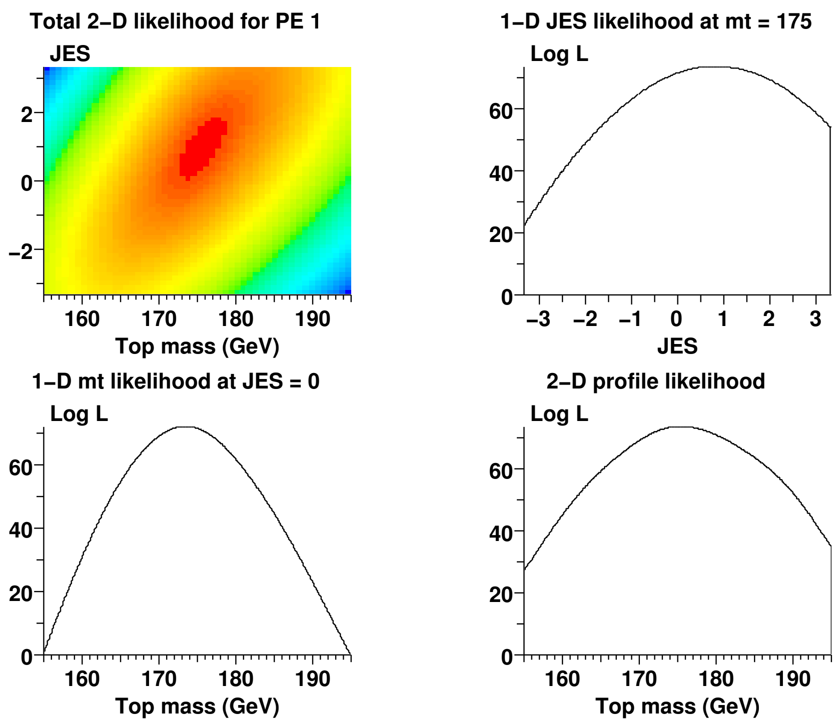
- After being integrated, we run pseudoexperiments (PEs) on the resulting 2-D likelihood curves to obtain a final mass measurement, expected error, and pull.

Sample 2D Likelihoods



- Sample 2-D likelihoods for single events
- m_t is on the x-axis, and JES is on the y-axis
- The color scale is calibrated so red is the peak of the curve, and blue is 5 units of log-likelihood below the peak (black is anything below that)
- Top 3 rows: signal; 4th row: $W + b\bar{b}$; 5th row: $W +$ light; bottom row: QCD (our three main backgrounds)

Sample PE Likelihoods



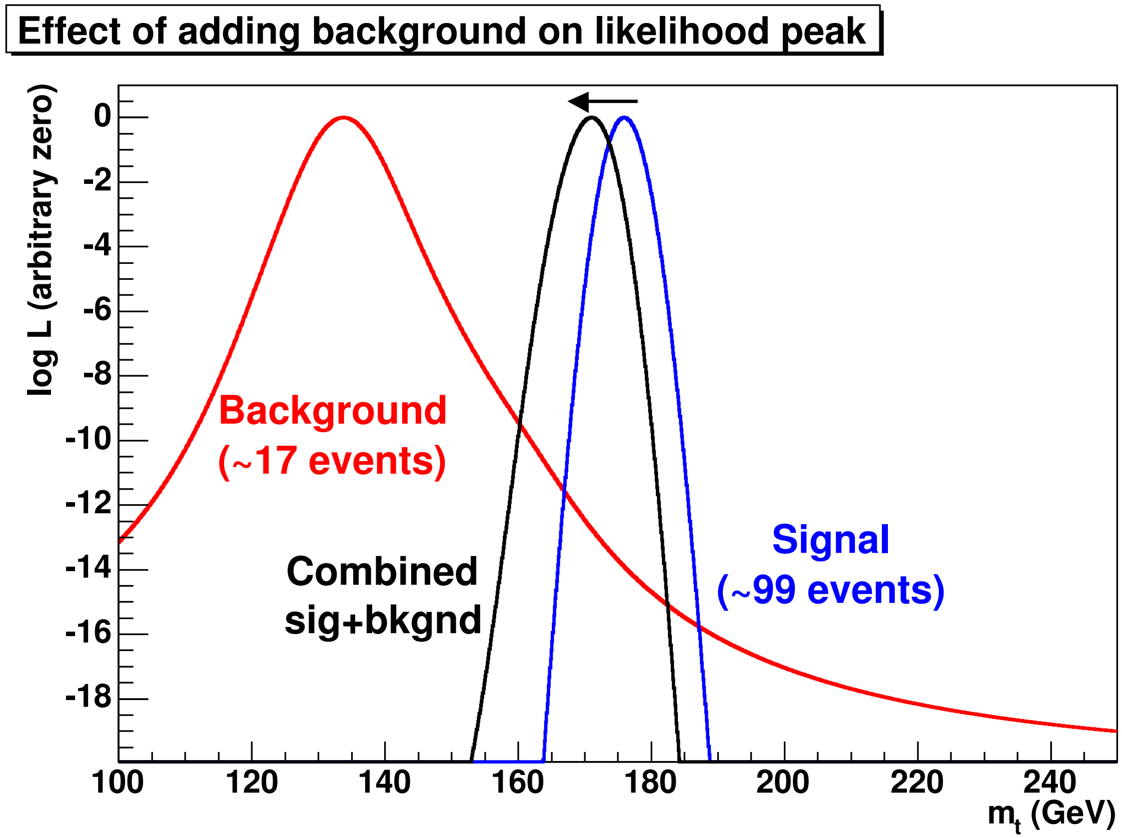
- Sample likelihoods for a single PE
- Top left: 2D likelihood
- Top right: Slice at $m_t = 175$ GeV for 1-D JES results
- Bottom left: Slice at JES $= 0$ for 1-D mass results
- Bottom right: Likelihood curve using profile likelihood

- Note how the profile likelihood curve is wider than the 1-D mass curve (essentially, because the profile likelihood is along the diagonal rather than a horizontal line). This reflects the effect of the JES systematics.

Adding Background

- Since all events are treated as $t\bar{t}$ events, adding in background events will cause a shift in the likelihood peak.

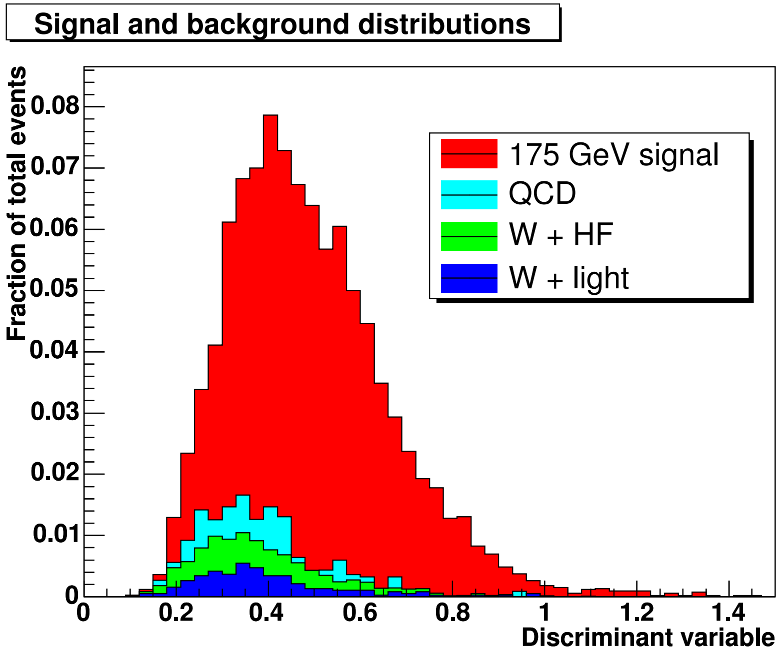
- To correct, we simply subtract out the expected background likelihood contribution, obtained from Monte Carlo.

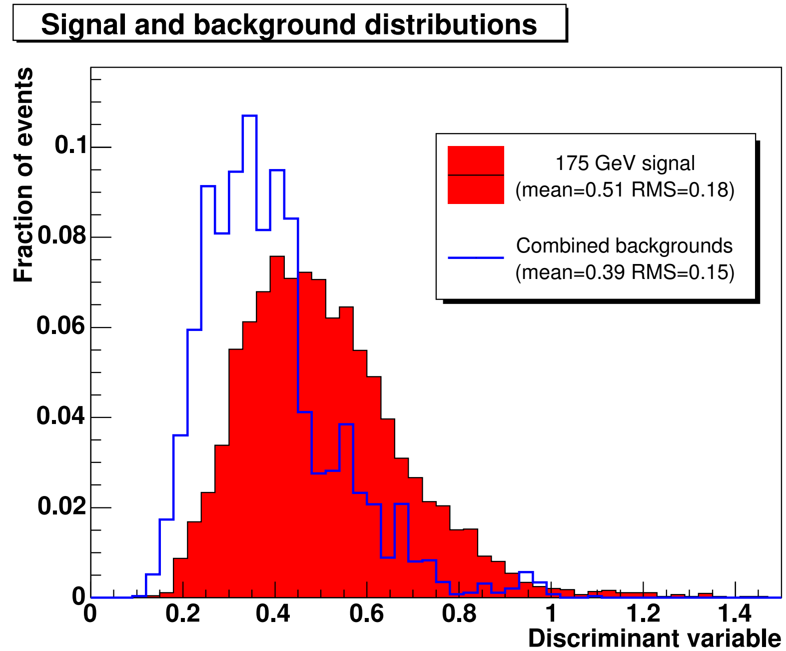


- For this to work, we need to be able to calculate the background fraction for a given event.

- To do this, we construct a discriminant variable q which has different distributions for signal and background but is independent of m_t and JES.

The Background Template

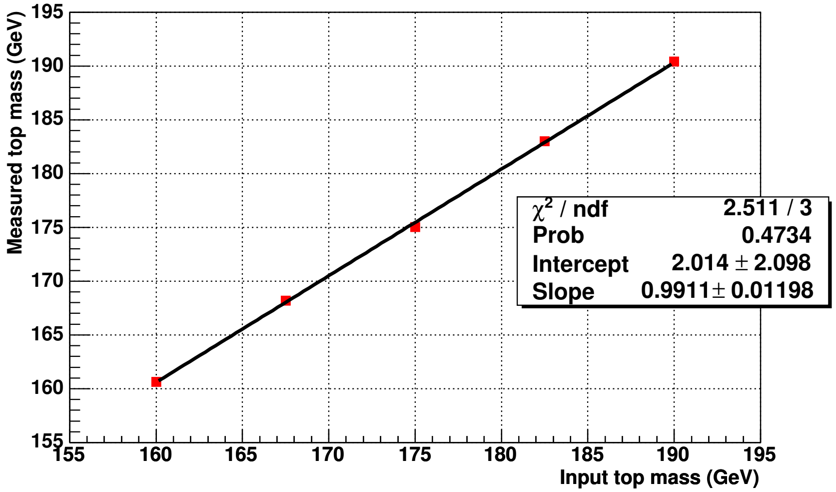


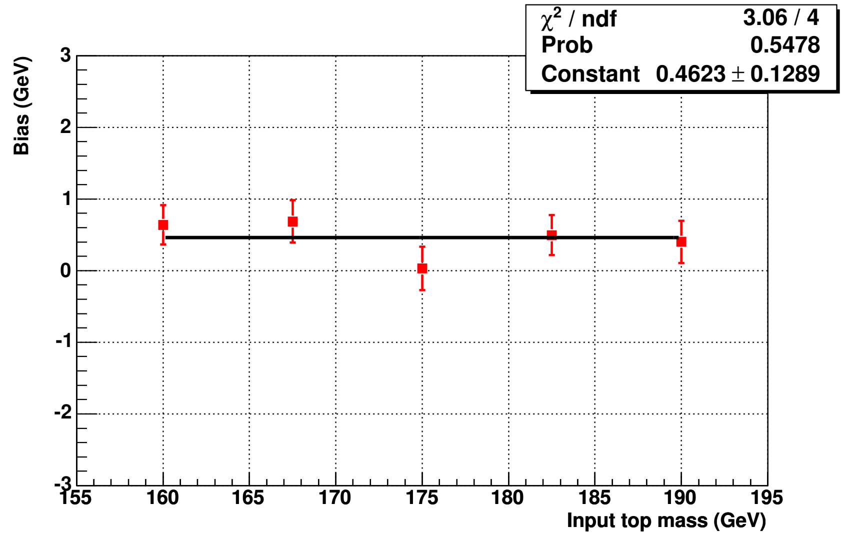


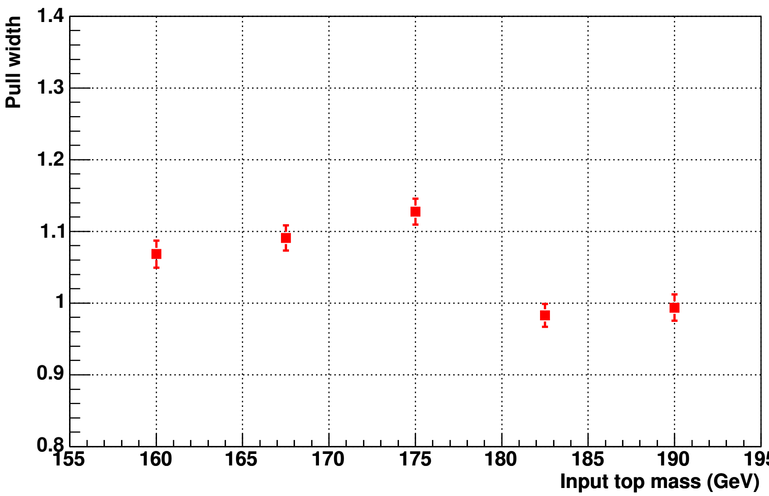
- Left: Signal and background distributions normalized to expected signal (85%) and background (15%) fractions, stacked histograms. We use this to calculate the signal probability: $f_{bg}(q) = B(q)/(B(q) + S(q))$.

- Right: Signal and background distributions normalized to 1, superimposed histograms. This illustrates the different distributions for signal and background.

Monte Carlo Results: Perfect Model Signal

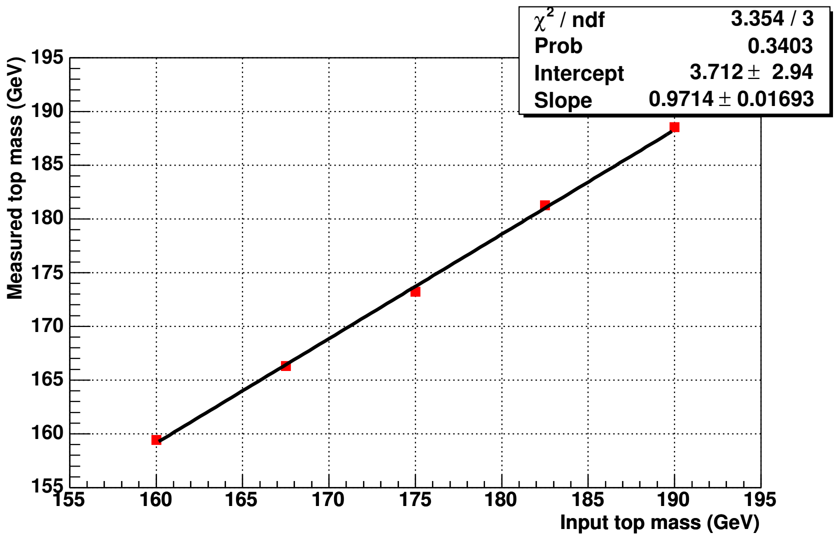


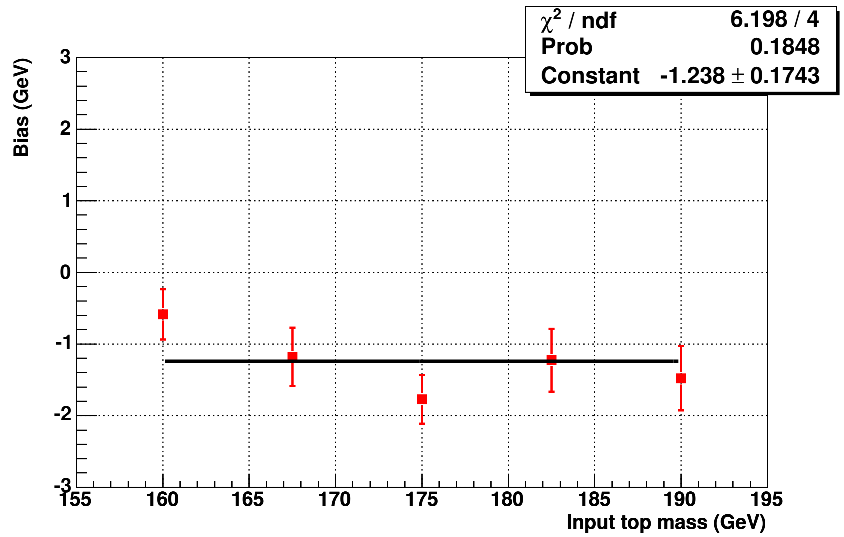


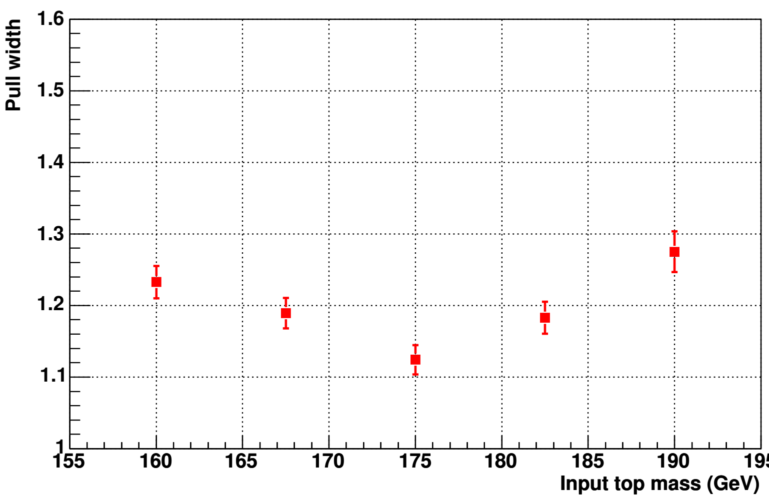


- Full simulation, 2000 PEs and 116 events/PE, signal only, 2-D profile likelihood
- Only perfectly modeled events are considered: events with poor jet-parton matching and $W \to \tau$ are rejected
- Average bias $= 0.46 \pm 0.13$ GeV
- Pull width ~ 1.05
- Slope of mass fit line $= 0.991 \pm 0.012$

Monte Carlo Results: Signal Only

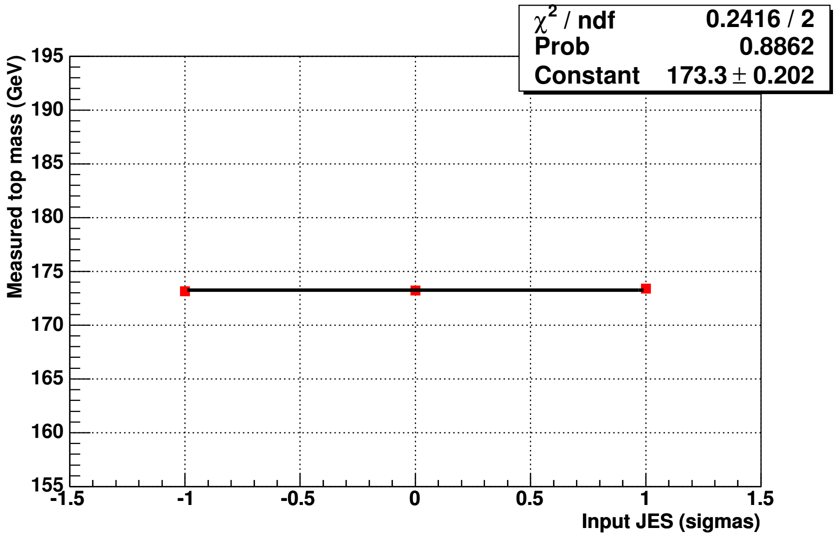


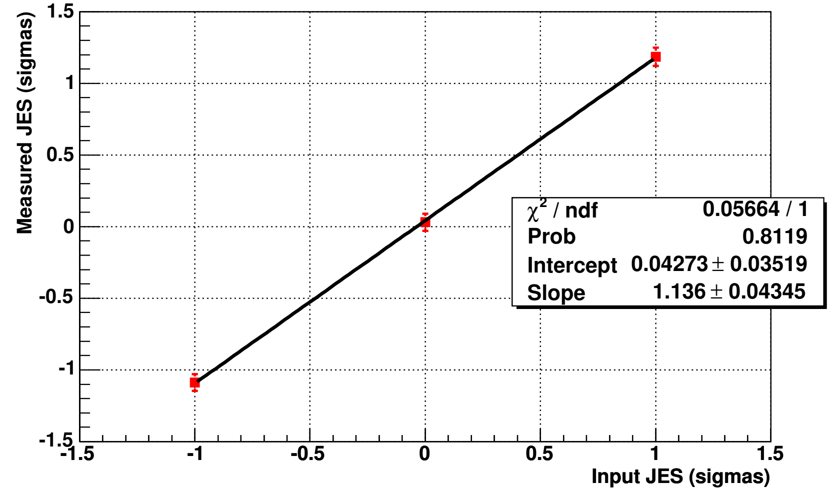




- Full simulation, 2000 PEs and 116 events/PE, signal only, 2-D profile likelihood
- Average bias = -1.24 ± 0.17 GeV
- Pull width ∼ 1.2
- Slope of mass fit line = 0.971 ± 0.017

Monte Carlo Results: JES Systematics



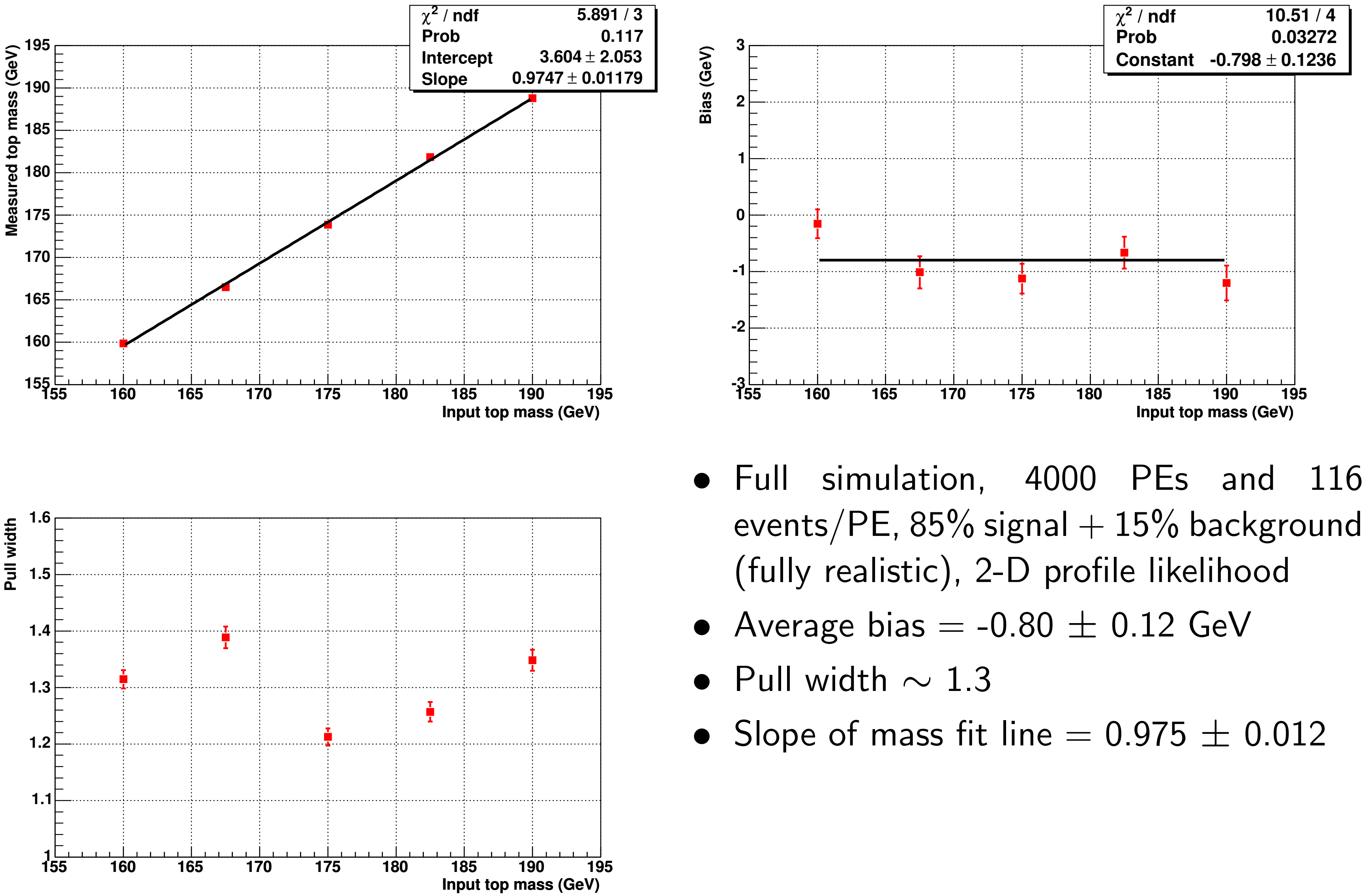


- To test that our method properly handles JES systematics, we prepare Monte Carlo samples in which all of the jets have been shifted by $\pm 1\sigma$. These plots show the effect of the JES systematic shift for a fixed input mass of 175 GeV.

- Upper left: Output mass (using 2-D profile likelihood) vs. JES systematic shift. This is completely flat, indicating that our 2-D likelihood method correctly handles JES systematics.

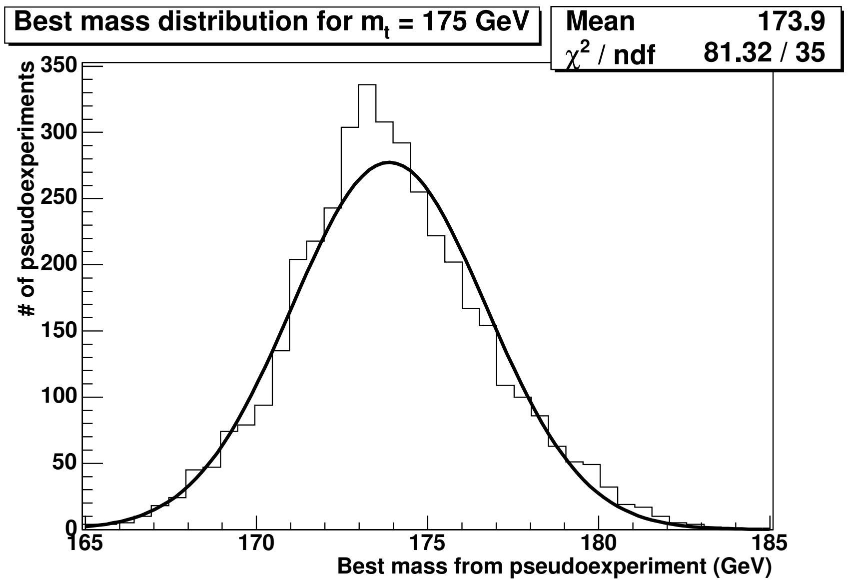
- Upper right: Output JES (using 1-D JES-only likelihood) vs. JES systematic shift.

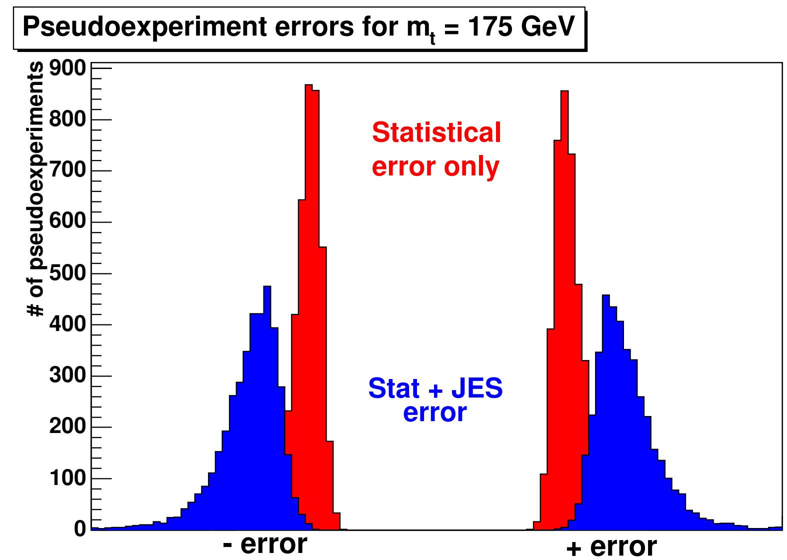
Monte Carlo Results: Signal + Background

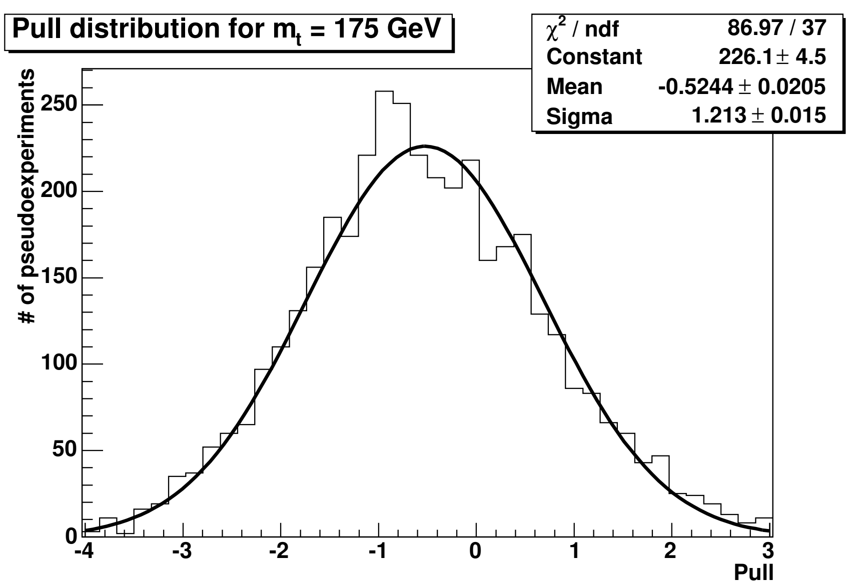


- Full simulation, 4000 PEs and 116 events/PE, 85% signal + 15% background (fully realistic), 2-D profile likelihood
- Average bias = -0.80 ± 0.12 GeV
- Pull width ∼ 1.3
- Slope of mass fit line = 0.975 ± 0.012

Monte Carlo Results: Signal + Background, 175 GeV







- Full simulation, 4000 PEs and 116 events/PE, 85% signal + 15% background (fully realistic), 2-D profile likelihood
- Top left: best mass distribution for PEs at $m_t = 175$
- Top right: expected error distribution for PEs at $m_t = 175$
- Bottom left: pull distribution for PEs at $m_t = 175$

Results Summary

Configuration	Average bias	Pull widths	Mass slope
Perfect model	0.46 $\pm$ 0.13	$\sim$ 1.05	0.991 $\pm$ 0.012
Signal, no τ	-0.24 $\pm$ 0.17	$\sim$ 1.15	0.982 $\pm$ 0.015
Signal only	-1.24 $\pm$ 0.17	$\sim$ 1.2	0.971 $\pm$ 0.017
Sig + bkgnd	-0.80 $\pm$ 0.12	$\sim$ 1.3	0.975 $\pm$ 0.012

- As we can see, filtering out τ events from our signal-only sample brings the bias much closer to 0 and the pulls closer to 1.
- Hopefully when we finish our framework for dealing with τ, we will see a similar improvement.

Conclusions

- The method has demonstrated a lot of success and prospects look good for a quality result.
- Still working on some improvements which will hopefully improve our error and reduce our pull width:
  - Better treatment of events with τ
  - Better handling of our error introduced by background events
  - Allowance for the jet energy systematics to vary on a jet-by-jet basis (currently we use a fixed systematic for all jets)
- Lots of systematics to be treated.
- Hope to have a good result soon!

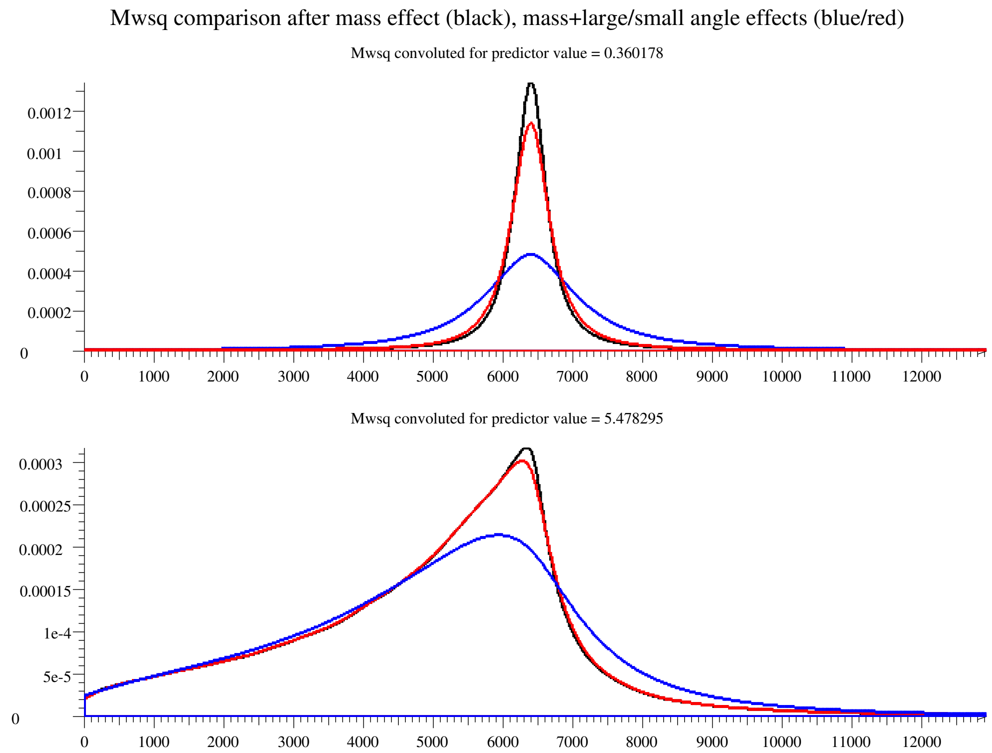
BACKUP SLIDES

Building Effective Propagators

- How do we build these effective propagators? We take advantage of the way Herwig handles the t and W decay.
  - First, Herwig decays the t and W into massless decay products.
  - Then, Herwig "fudges" the decay products by adding masses (conserving the overall 4-momentum, of course) so that it can begin a parton shower.
- By taking the results of the first step, and then rotating these into our final detector angles, we will get partons ("effective partons") which adhere to our integration assumptions.
- Then we can build M_W^2 and M_t^2 distributions from these partons to use as our effective propagators.
- We also include terms for the angular resolution and jet mass effects.

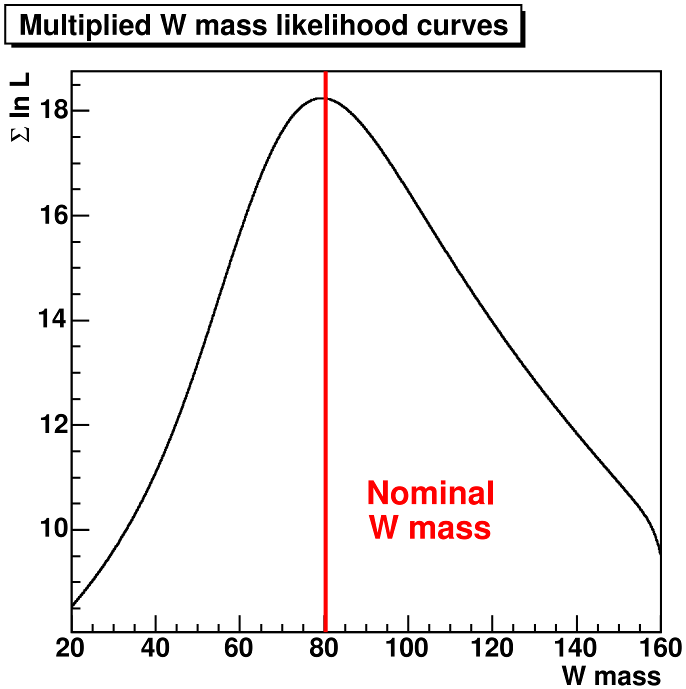
Event-by-Event Propagator Adjustment

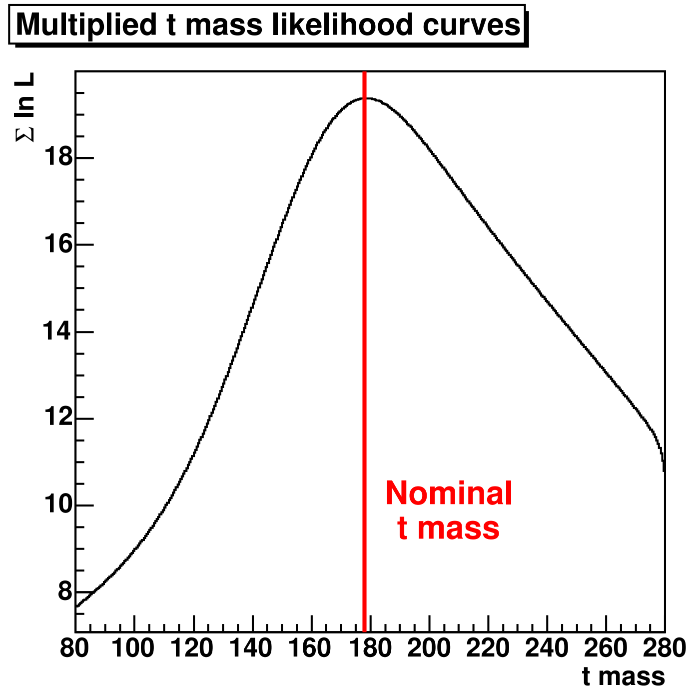
- We adjust the width of the hadronic propagator based on the kinematics of the event.
- The jet mass effects and angular resolution effects affect the width of the propagator differently for different events. We compute this uncertainty using the partial derivatives of the W mass with respect to these variables for the event.



- The top shows the M_W^2 propagator for an event for which the effect of the jet mass is small, and the bottom shows a large jet mass effect.
- Black = jet mass effect only, red = jet mass effect + small angular resolution effect, blue = jet mass effect + large angular resolution effect

Transfer Function Validation





- To verify the transfer functions, we reconstruct the hadronic W and t mass using the transfer functions (but no other parts of our full integration).
- We integrate over the parton momenta using the transfer functions and a prior distribution of $\vec{P}_W$ and $\vec{P}_t$ from Monte Carlo.
- The agreement shows that the transfer functions are working as desired.

Permutation Weighting

- We employ the following procedure to weight the permutations in our integration. First, we use the following parameterization of the tag efficiencies for b-jets, c-jets, and light jets:
  - $P(\mathrm{tag}|b, E_T, \eta) = (0.108 + 0.0175E_T - 3.47 \cdot 10^{-4}E_T^2 + 3.32 \cdot 10^{-6}E_T^3 - 1.58 \cdot 10^{-8}E_T^4 + 2.93 \cdot 10^{-11}E_T^5)(1.05 - 0.517\eta + 1.457\eta^2 - 1.20\eta^3 + 0.0466\eta^4 + 0.0895\eta^5)$
  - $P(\mathrm{tag}|c, E_T, \eta) = 0.22 \cdot P(\mathrm{tag}|b, E_T, \eta)$
  - $P(\mathrm{tag}|l, E_T, \eta) = (0.00355 - 2.63 \cdot 10^{-4}E_T + 1.18 \cdot 10^{-5}E_T^2 - 1.41 \cdot 10^{-7}E_T^3 + 7.53 \cdot 10^{-10}E_T^4 - 1.55 \cdot 10^{-12}E_T^5)(0.821 + 0.452\eta + 0.437\eta^2 - 0.555\eta^3)$

- Next, we have the *a priori* probabilities for a jet in a $t\bar{t}$ event, assuming $W \to ud$ and $W \to cs$ each have a probability of 0.5: $P(b) = \frac{1}{2}, P(c) = \frac{1}{8}, P(l) = \frac{3}{8}$

- Hence, using Bayes' Theorem, $P(b|\mathrm{tag}) = \frac{P(\mathrm{tag}|b) \cdot P(b)}{P(\mathrm{tag})}$, where $P(\mathrm{tag}) = P(\mathrm{tag}|b)P(b) + P(\mathrm{tag}|c)P(c) + P(\mathrm{tag}|l)P(l)$.

- We use these probabilities to weight accordingly.

Expected Backgrounds

These backgrounds are from published results for 4 tight jets for 318 pb^{-1}. These are rescaled to 940 pb^{-1}. We also rescale by a fraction of 0.648 to account for the effect of our 0 loose jet cut.

Background	318 pb^{-1}	940 pb^{-1}	0 loose jets
non-W QCD	3.07 ± 1.06	9.08 ± 3.13	5.88 ± 2.18
W + light (mistag)	2.27 ± 0.45	6.71 ± 1.33	4.34 ± 1.00
diboson (WW, WZ, ZZ)	0.39 ± 0.08	1.15 ± 0.24	0.75 ± 0.18
Sum of above 2	2.66 ± 0.53	7.86 ± 1.57	5.09 ± 1.18
W $b\bar{b}$	1.70 ± 0.79	5.03 ± 2.33	3.25 ± 1.24
W $c\bar{c}$, W c	1.31 ± 0.63	3.87 ± 1.86	2.51 ± 1.21
Single top	0.41 ± 0.09	1.21 ± 0.27	0.78 ± 0.20
Sum of above 3	3.43 ± 1.41	10.11 ± 4.16	6.54 ± 2.80
Total Background	9.16 ± 1.83	27.05 ± 5.41	17.51 ± 4.05
Expected top (m_t=175)	42 ± 5	124 ± 15	82 ± 10
Events observed	63	179	116

Total expected background fraction $\sim$ 85%

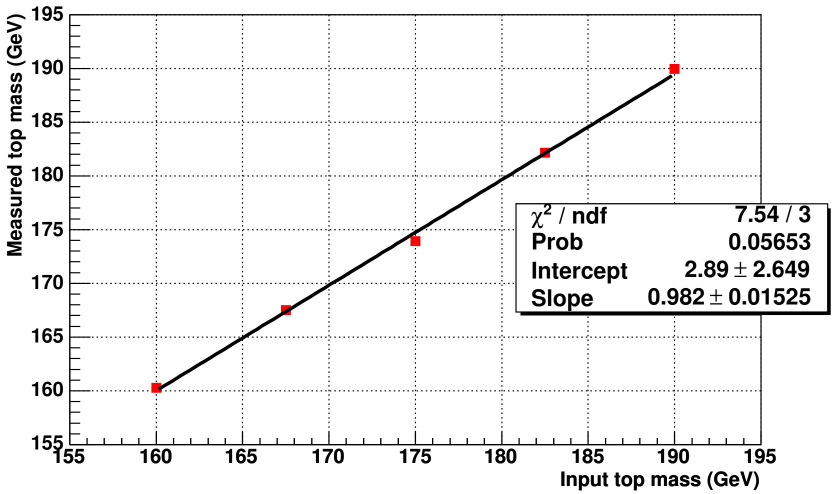
Background Discrimination Variable

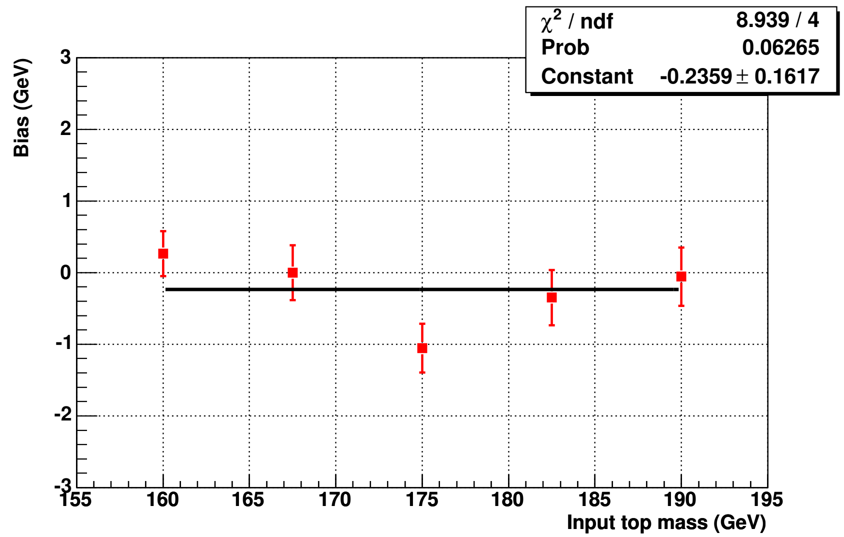
- Our discriminant variable is called a "hybrid" variable, as it is constructed out of a linear combination of three topological quantities:
  - Aplanarity $= \frac{3}{2}Q_1$, where Q_1 is the smallest eigenvalue of the momentum tensor
  - D_R, the minimum jet-jet ΔR weighted by the momentum ratio of the smaller jet momentum to the lepton $= \Delta R_{ij}^{\min} \cdot \min(p_T^{(i,j)})/p_T^{\ell}$
  - H_{TZ}, the scalar sum of all jet P_T except the leading jet over the scalar sum of all P_z for jets, lepton, and neutrino (using smaller solution) $= \sum_{i=2}^{4} |p_T^{(i)}| / (\sum_{i=1}^{4} |p_z^{(i)}| + |p_z^{\ell}| + |p_z^{\nu}(\min)|)$

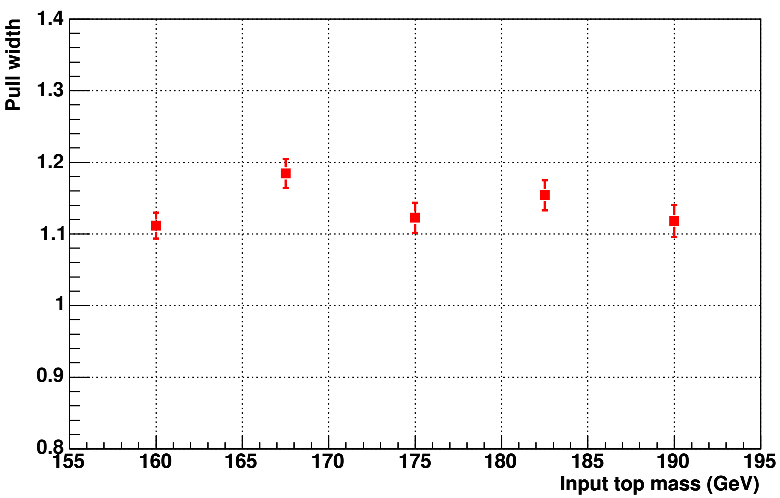
- While our inputs into our "hybrid" variable may have some m_t or JES dependence individually, by creating the proper linear combination we can ensure that the total dependence cancels out.

- By varying the weights of the three inputs into our "hybrid" variable, we can achieve these goals very well – we get excellent m_t and JES stability and decent S/B discrimination.

Results: Signal Only, No τs







- Full simulation, 2000 PEs and 116 events/PE, signal only, 2-D profile likelihood
- Events with $W \rightarrow \tau$ are rejected
- Average bias = -0.24 $\pm$ 0.17 GeV
- Pull width = 1.15 $\pm$ 0.01
- Slope of mass fit line = 0.982 $\pm$ 0.015