A Measurement of

Top Charge Asymmetry at DØ

Work in progress

Amnon Harel

aharel@fnal.gov

University of Rochester

The April Meeting 2007 of the American Physical Society
Jacksonville, Florida, 15th April 2007

Overview

Introduction:

- The standard model top asymmetry
- Key concepts

Not a typical top analysis

Analysis Ingredients:

- Analysis strategy
- The acceptance
- The dilution
- Fit procedure
- Less than preliminary fit results

Will focus on the unusual aspects of this analysis

Summary

The Top Asymmetry

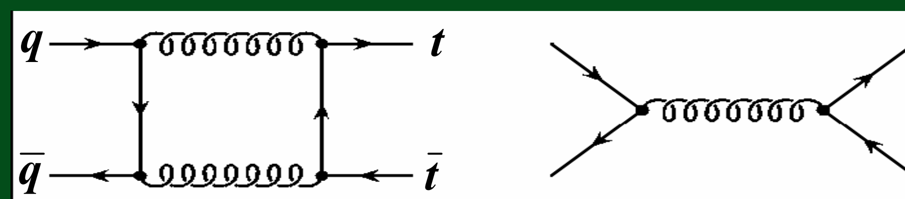
At the Tevatron the top charge production asymmetry is visible as a forward-backward asymmetry.

Much harder to measure it at the LHC (gluon fusion doesn't help)

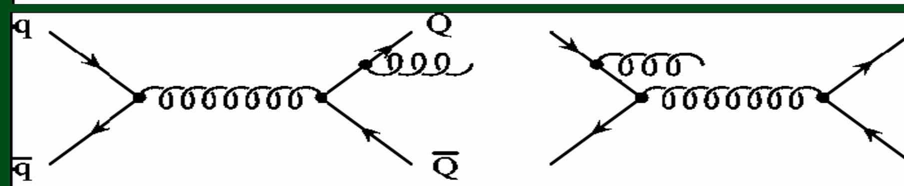
$$A_{fb} = \frac{N_+ - N_-}{N_+ + N_-}$$

The **top charge asymmetry** has several contributions.
All are **interference terms** of at least NLO:

1. An interference between:



2. An interference between:



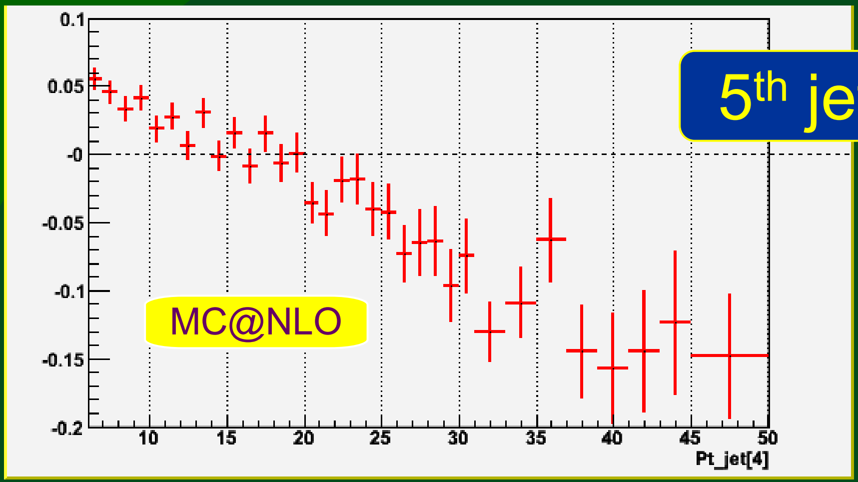
3. Higher order terms (recently evaluated, hep-ph/0703120)
4. Interferences in *qg*→*tt* diagrams (e.g. Flavor creation)
5. Interferences with mixed (electroweak neutral current + gluon) box diagrams

Some of these contributions tend to cancel out.

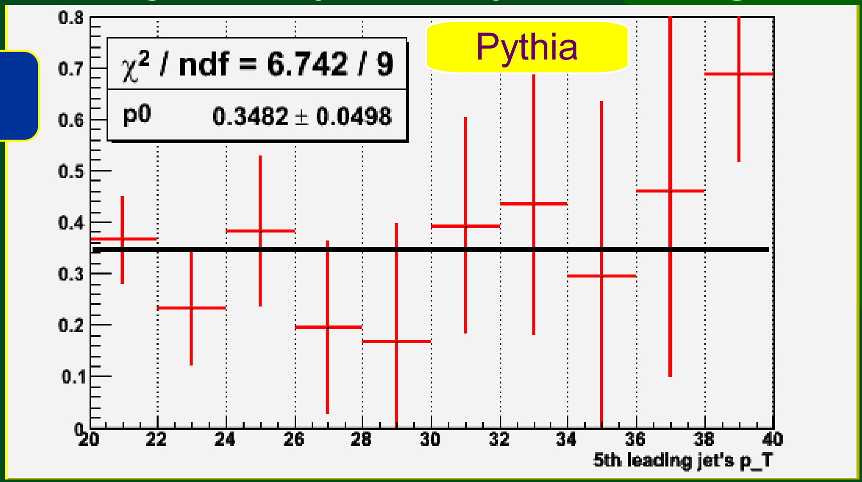
↳ The top asymmetry varies greatly throughout phase space (see next slide).
Which "*A*" are we measuring?

"Differential" Asymmetries

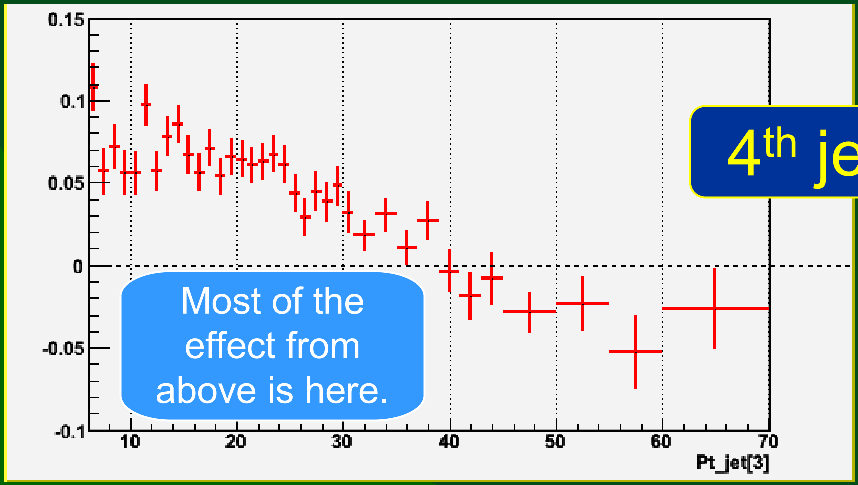
top asymmetry on the left; **W+jets** asymmetry on the right



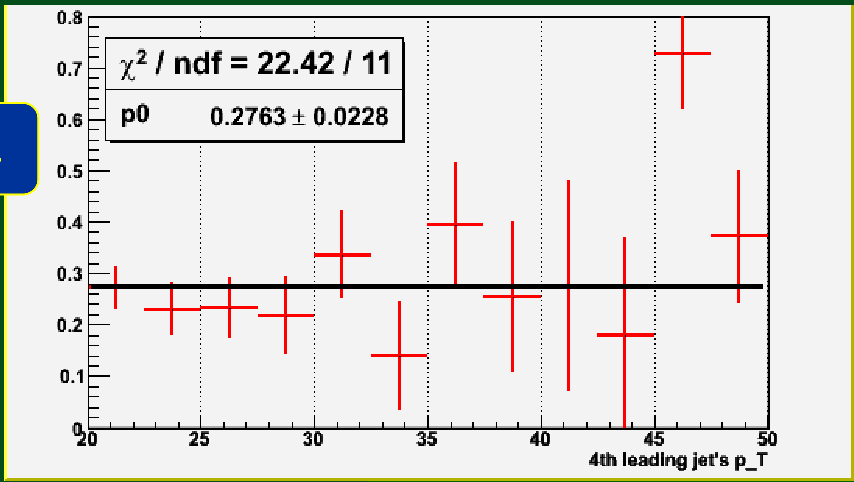
5th jet p_T



Reminiscent of the calculation by Bowen et. al. (hep-ph/0509267) stating that the *tt* and *ttj* asymmetries are approximately 6% and -7% (jet acceptance cuts: pT>20GeV, |η|<3)



4th jet p_T



Theory papers usually ignore top decays in their jets, and ignore this effect.

Key Concepts

Acceptance:

Where, in phase space, are we measuring the asymmetry?

As we saw, this changes the standard model predictions for this asymmetry.

Dilution:

How well do we reconstruct the asymmetry?

Quantitatively:

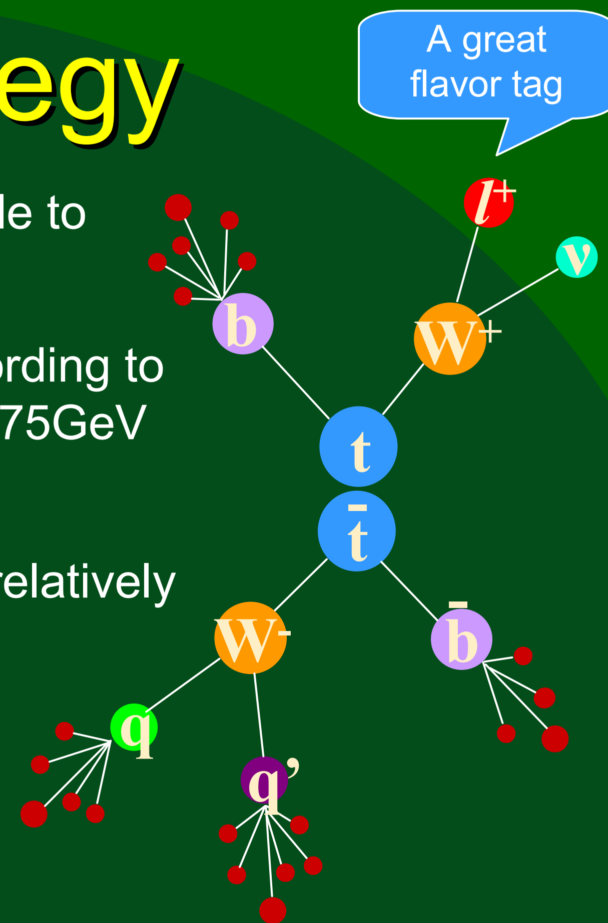
If we reconstruct the sign correctly for a fraction, *p*, of the accepted events, how much of the asymmetry is visible?

$$d = 2p - 1$$

This is a reconstruction effect, so we would like to correct for it (as is the standard practice in asymmetry measurements).

BTW, the "dilution factor" gives the statistical significance: $D = (2p-1)^2$

Analysis Strategy



Measurement designed to reach a sensitivity comparable to the effect expected in the Standard Model (5-10%).

Using a kinematic fitter to reconstruct $\Delta Y = Y_t - Y_{\bar{t}}$ according to the top pair hypothesis, object resolutions, M_W & m_{top}=175GeV

- 2-4 times stronger than just using the lepton's rapidity

Selecting 4 jet events in the lepton+jets channel with a relatively loose b-tag (~84% efficient)

The usual practice is to correct from the observed asymmetry to the "true" (i.e. particle-level) asymmetry. This "unfolds" the reconstruction effects.

In this measurement, statistics suffice for **only one bin**.
→Acceptance and reconstruction quality **vary greatly within the bin**.

Unfolding requires knowing the distribution of events within the bin
→ model dependence.

Instead, we'll have a simple specification of our acceptance and dilution, with which the model predictions can be "folded" to predict the results of this measurement.

The Acceptance

Goal: **A simple description of the region of phase space where signal events are accepted into the analysis.**

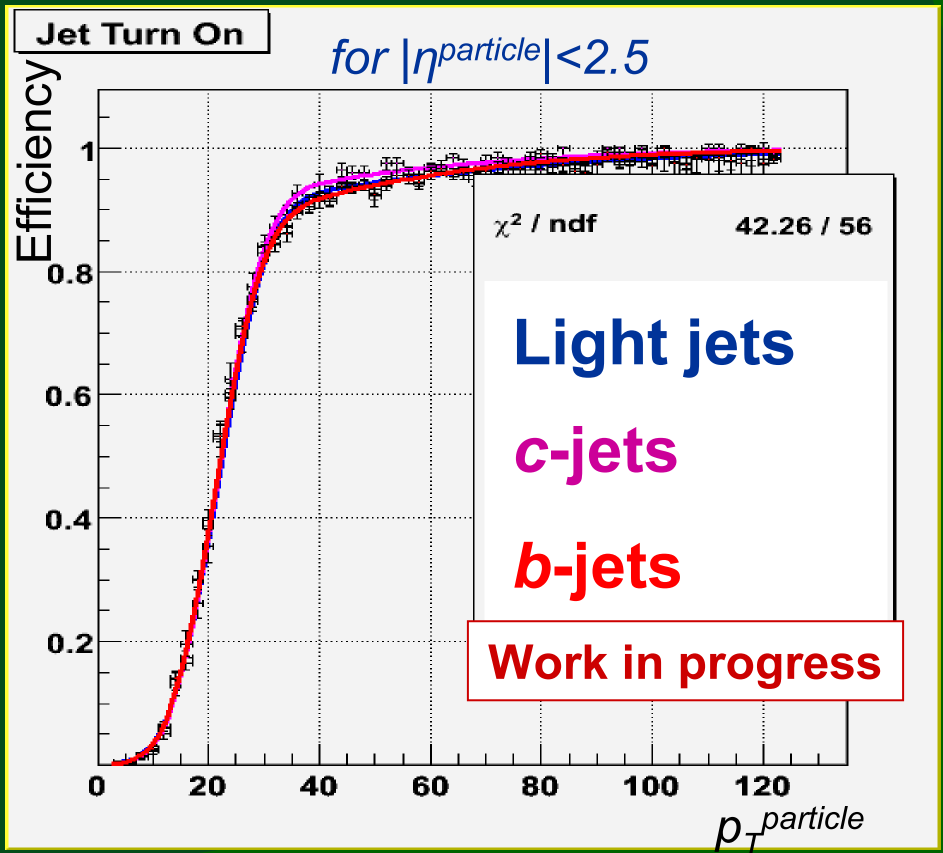
The main issue is the difference between (fully-corrected) detector jets and particle-jets. Smaller effects are neglected. More studies of *b*-tagging required.

Jet Acceptance

Was studied in fully corrected (μ correction, scale, smearing, efficiencies) Monte Carlo.

$$A = f\left(p_T\right)g\left(|\,\eta\,|\right)$$

But that's not so simple…



The Acceptance

Goal: **A simple description of the region of phase space where signal events are accepted into the analysis.**

The main issue is the difference between (fully-corrected) detector jets and particle-jets. Smaller effects are neglected. More studies of *b*-tagging required.

Jet Acceptance

Was studied in fully corrected (μ correction, scale, smearing, efficiencies) Monte Carlo.

$$A = f(p_T)g(|\eta|)$$

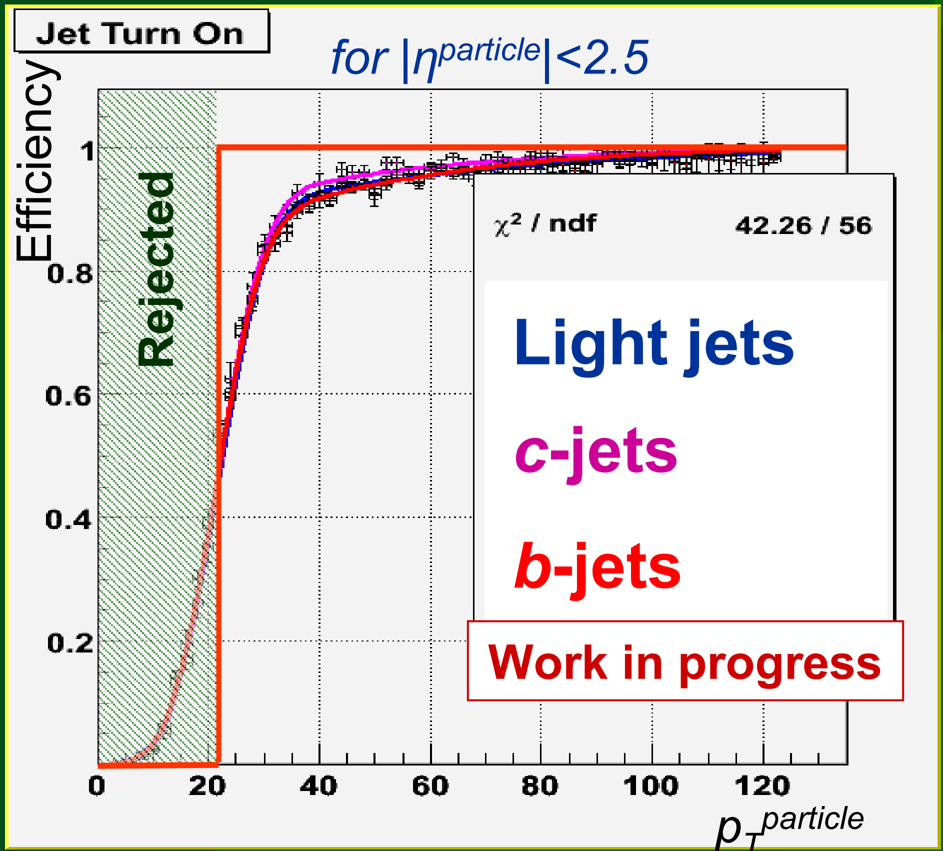
The big approximation is to replace those fitted turn on curves with simple box cuts:

p_T>21 GeV and |Y|<2.5

• built a parametrized MC on top of MC@NLO to evaluate the approximation

Conclusion:

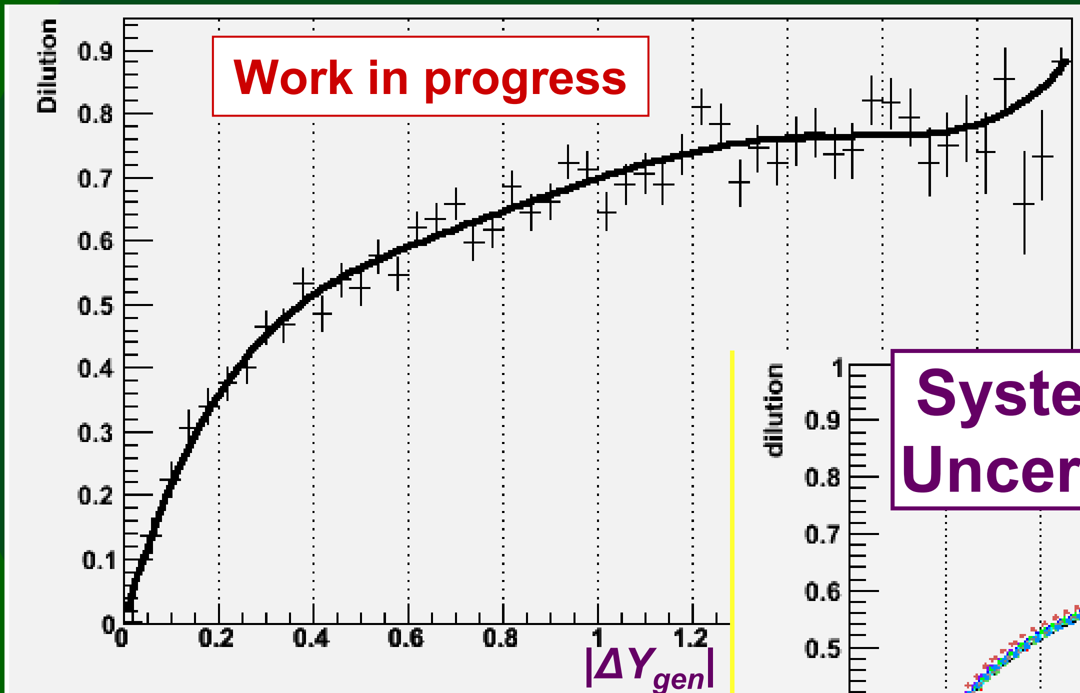
• The systematic uncertainties on the expected asymmetry from using the simple acceptance description are 1% (absolute)



The Dilution

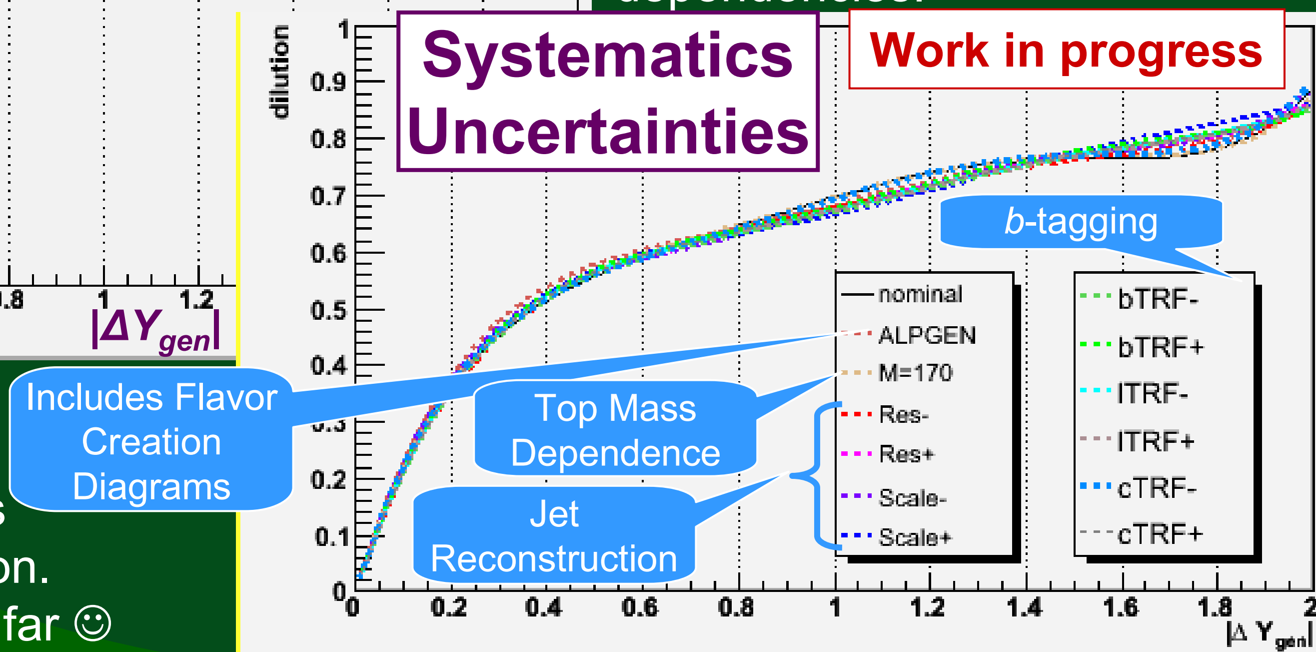
The dilution describes how well we reconstruct the asymmetry.

Some dilution from misidentifying the lepton's charge, in particular if this misidentification differs between the two sides of the detector. A small effect.



The geometric part was parametrized in $|\Delta Y_{generated}|$ and measured on Pythia Signal.

Need to handle residual dependencies.

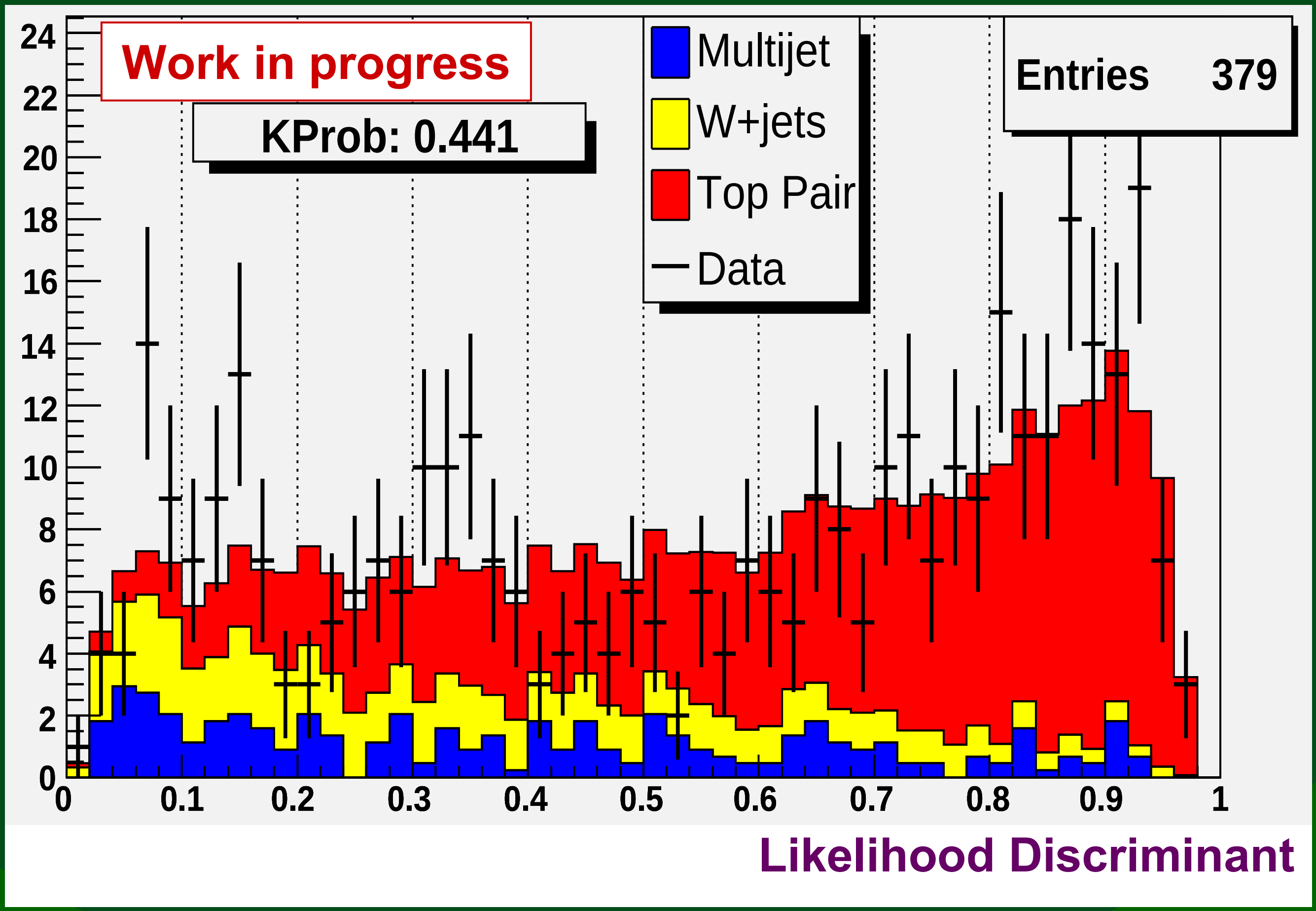


Most systematic effects enter through the dilution. They're pretty small so far ☺

The Likelihood Discriminant

A likelihood that discriminated between top and W+jets events without biasing |ΔY|

- leading *b*-jet p_T, fitter's χ^2, K^t_{min}, M_{jj} (from hadronic W according to the fitter)



Fit Procedure

- The sample composition and the asymmetry are fitted simultaneously
- a simultaneous fit over two distributions:
  - the likelihood discriminant
  - the sign of the rapidity difference
- with four templates whose sum is fitted to data:
  - **forward signal**
  - **backward signal**
  - **W+jets events**
  - **multijet events**
    - relative fraction taken from the data
    - likelihood-discriminant and asymmetry distributions taken from the data

Other fit procedures, such as using an event-by-event likelihood, have more statistical power, but they gain it by giving more weight to signal-like events and/or those where *sign(ΔY)* is well measured.

This particular fit method was chosen as it keeps the acceptance simple:

All selected events have the same contribution to the fitted asymmetry.

Less Than Preliminary Fit Results

Work in progress

Event Type	Fitted Number Of Events		
	Both Channels	e+jets	μ+jets
Signal	264 ± 22	131^{+16}_{-15}	131 ± 15
W+jets	76^{+21}_{-20}	31 ± 14	49 ± 14
Multijet	39 ± 4	$27.7^{+2.8}_{-2.7}$	$8.4^{+2.1}_{-2.0}$

The asymmetry has little correlation with the other fit parameters, up to 8%.

- the number of W+jets has little correlation as the reconstruction under the top pair hypothesis washes out the W asymmetry.

The fitted observable asymmetry: $A_{fb} = (12 \pm 8(fit) \pm x(syst))\%$

Includes errors on sample composition.

Exploring improvements that'll change the nominal value

Work in progress

MC@NLO predictions:

All A_{fb}s ± 1.00 (acceptance)	Asymmetries (in %)	
	Fitted Turn Ons	Simple Cuts
Generated Asymmetry	0.97 ± 0.34	1.94 ± 0.30
Observable Asymmetry	0.66 ± 0.27	1.27 ± 0.24

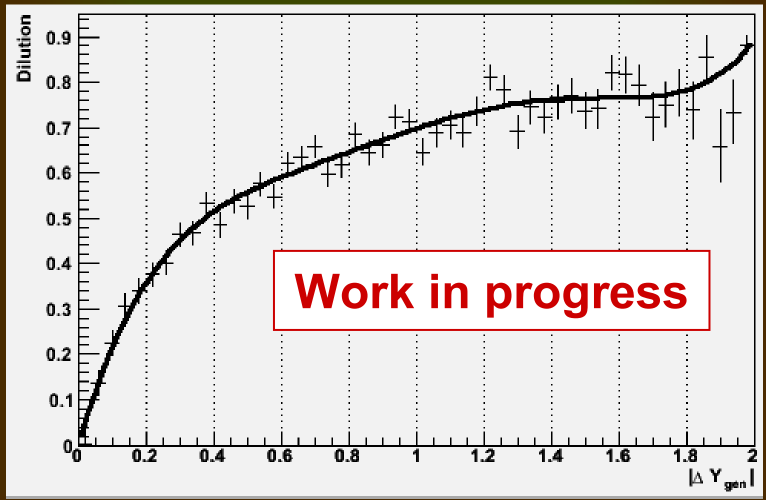
Summary

Measuring the top charge production asymmetry within the **approximate acceptance:**

- 4+ jets with p_T>21GeV and |η|<2.5
- highest jet p_T>~37GeV
- electron with p_T>15GeV and |η|<1.1 or muon with p_T>18GeV and |η|<2

The approximation cause an uncertainty of 1% (absolute) on the expected asymmetry.

Reconstructing the asymmetry with the **dilution**:



The less than preliminary measured observable asymmetry: $A_{fb} = \left(12 \pm 8(fit) \pm x(syst)\right)\%$

Exploring improvements that'll change the nominal value

MC@NLO predictions:

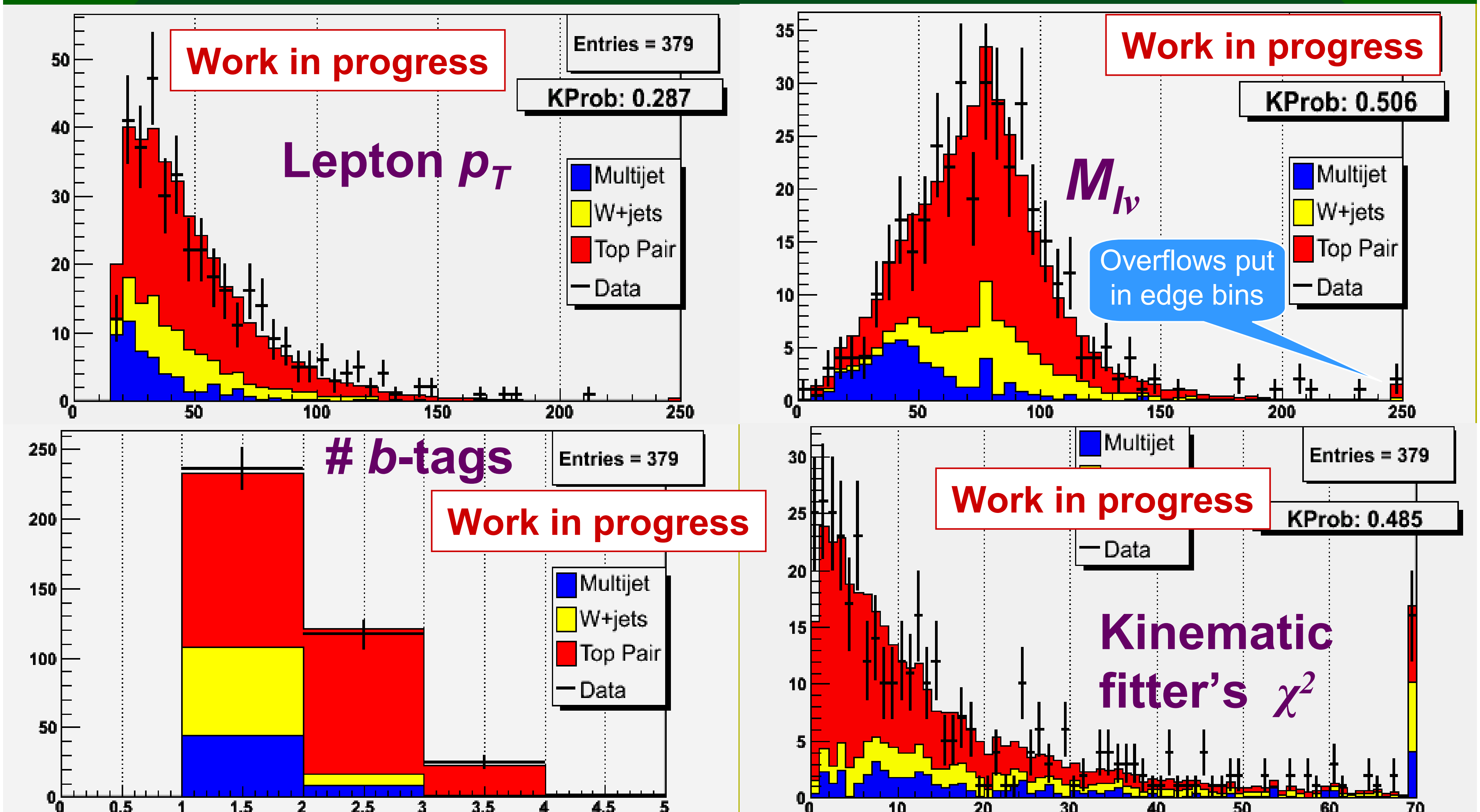
All A_{fb}s ± 1.00 (acceptance)	Asymmetries (in %)	
	Fitted Turn Ons	Simple Cuts
Generated Asymmetry	0.97 ± 0.34	1.94 ± 0.30
Observable Asymmetry	0.66 ± 0.27	1.27 ± 0.24

Back up slides

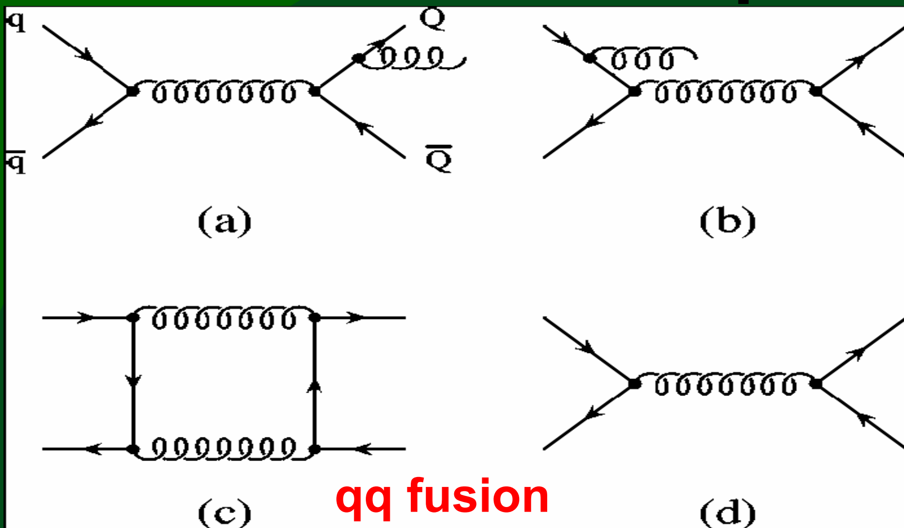
Data Modeling

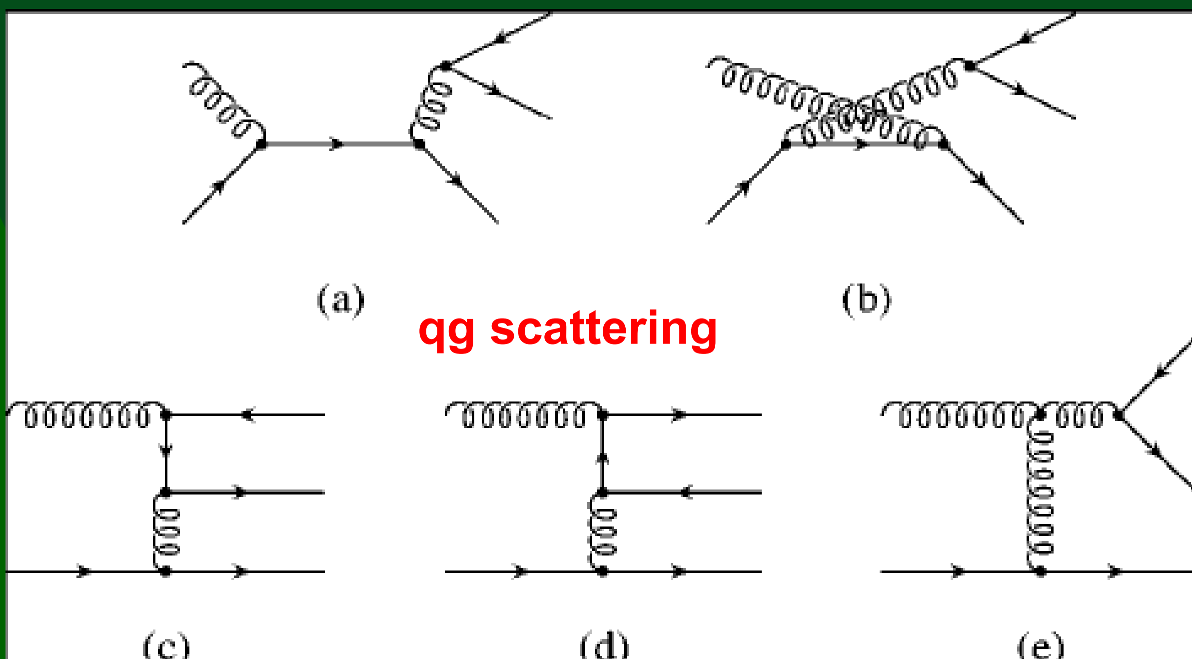
Good data modeling in both channels.

Here are a few sample sanity plots for both channels combined.



Some Top Pair Diagrams



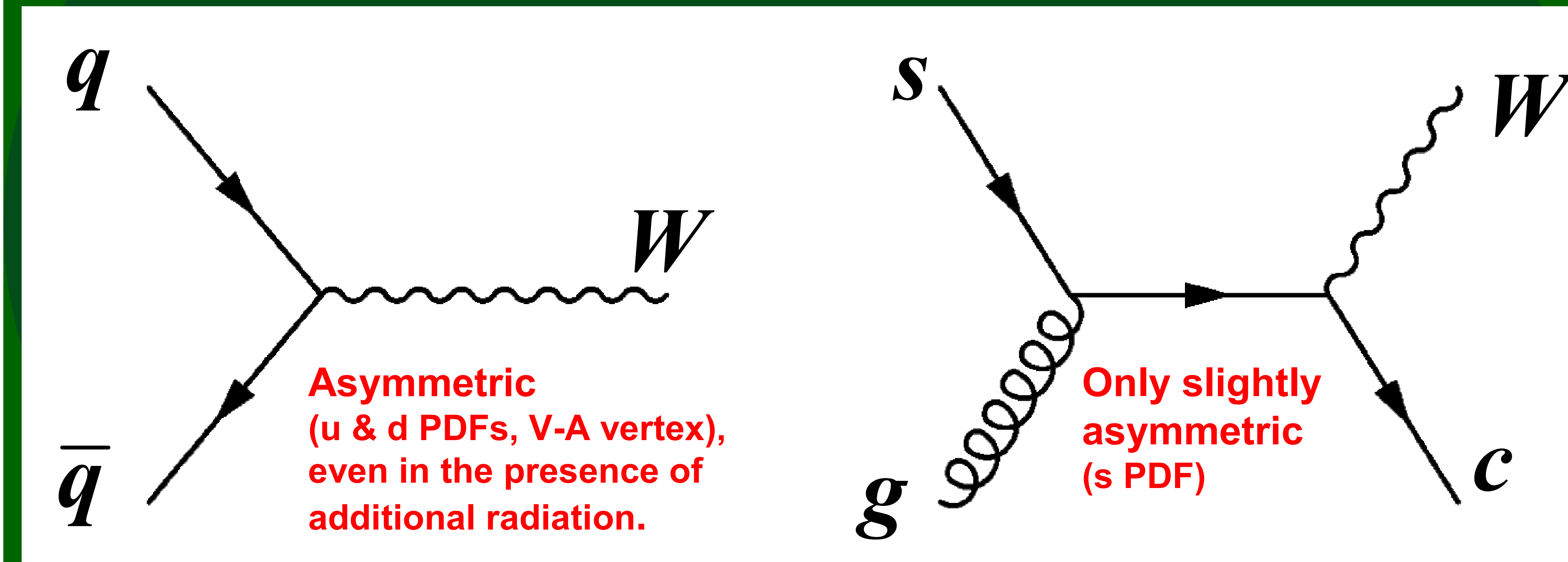


The top asymmetry arises from NLO contributions to top pair production.

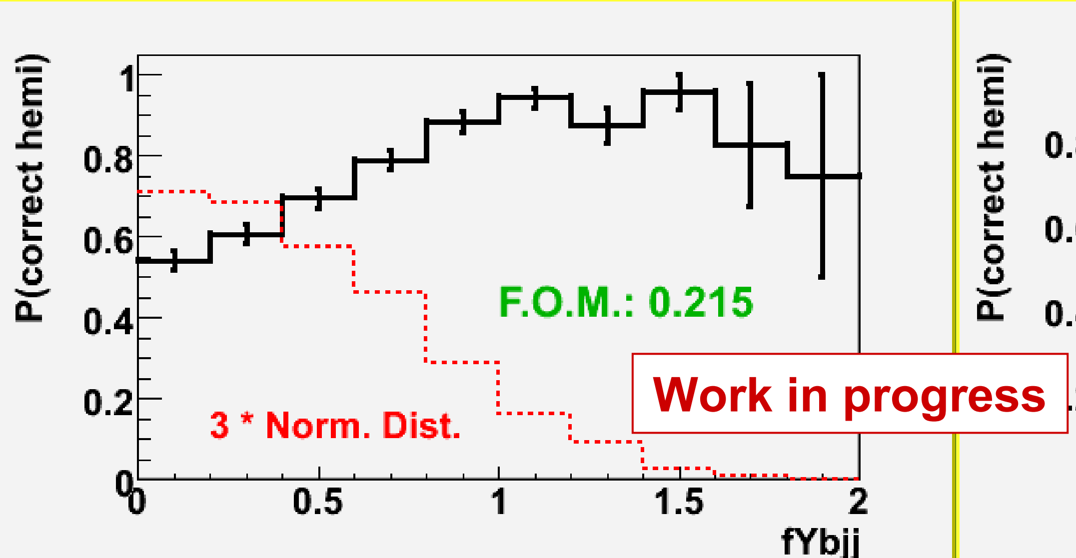
Plots are from Kuhn & Rodrigo hep-ph/9807420

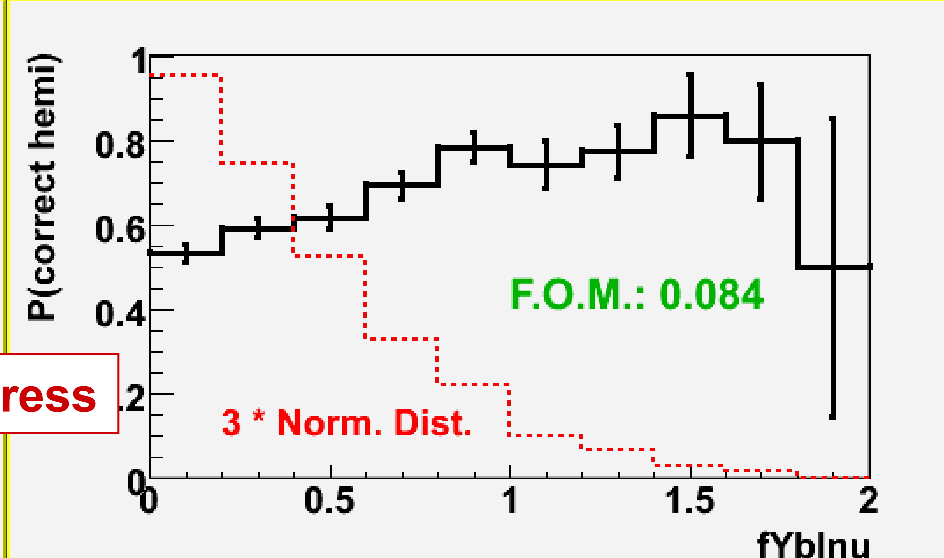
The W+jets Asymmetry

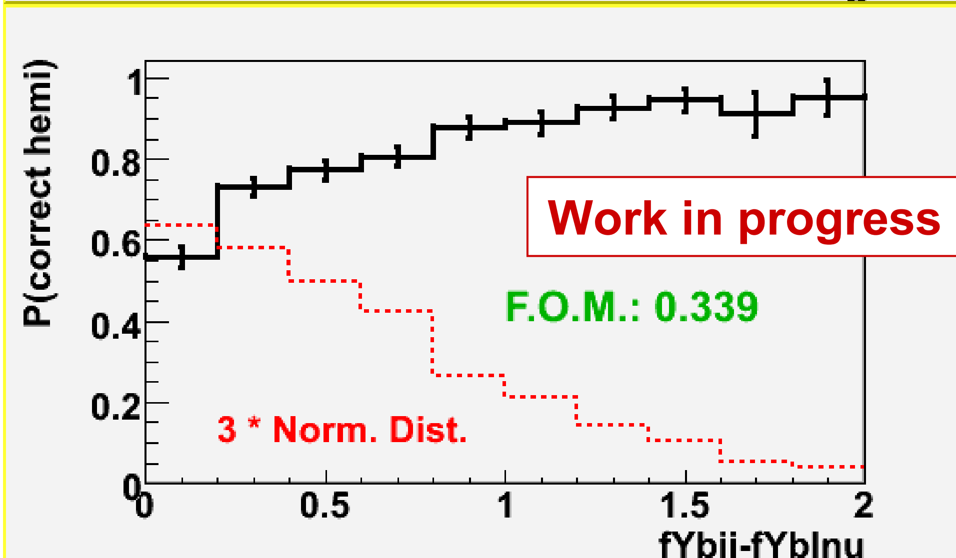
The main background is W+jets and is also asymmetric!



What to measure?







$$F.O.M. = \int f(x)\left(2p(x)-1\right)^2 dx$$

Where p(x) is the tag's purity, in this case, the probability to choose the correct hemisphere

Reconstructed W+jets Asymmetry

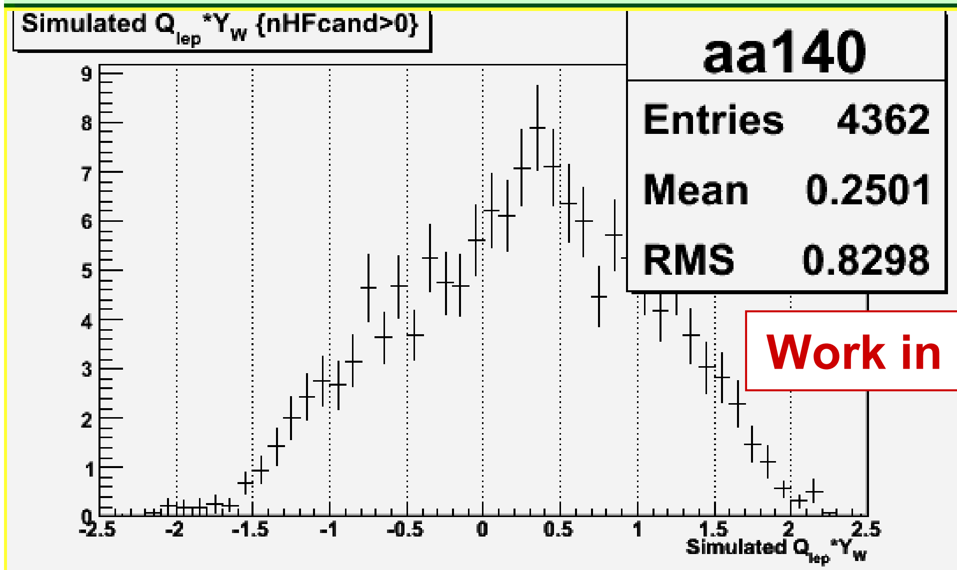
Breakdown may be useful for systematics (e.g. b-tagging RFs)

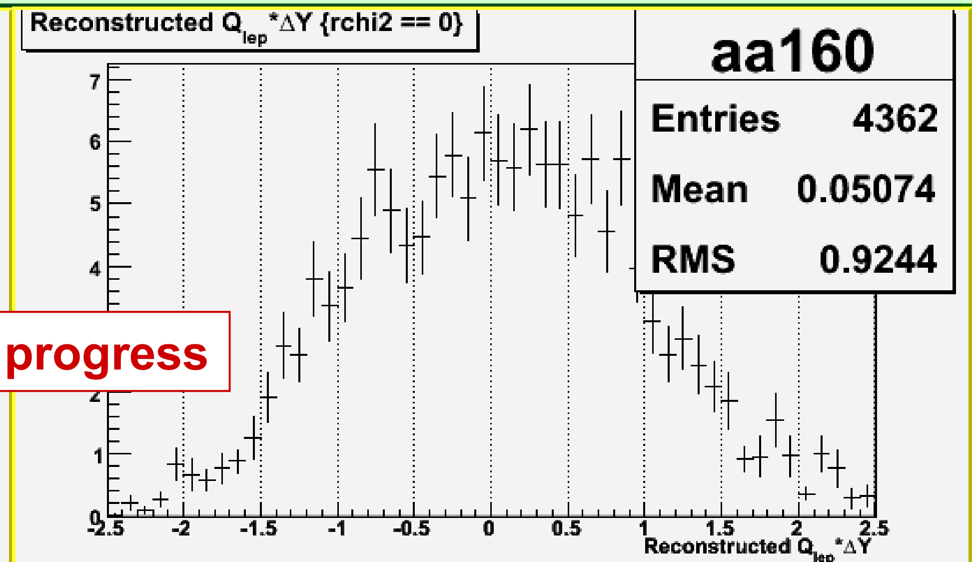
Sample	Fraction	Asymmetries (in %)		
		Simulated	Reconstructed	MC@NLO-particle-level
W plus jets	1.0	25.7 ± 2.3	3.9 ± 2.4	
Wjjjj	0.487 ± 0.012	28 ± 4	5 ± 4	21.98 ± 0.34
Wcjjj	0.140 ± 0.008	18 ± 6	-1 ± 6	0.2 ± 1.3
WccJJ	0.154 ± 0.007	27 ± 5	9 ± 5	
WbJJJ	0.092 ± 0.004	25.3 ± 3.2	-0.2 ± 3.4	13 ± 4
Wbbxx	0.127 ± 0.004	21.6 ± 2.9	-0.4 ± 3.0	

Work in progress

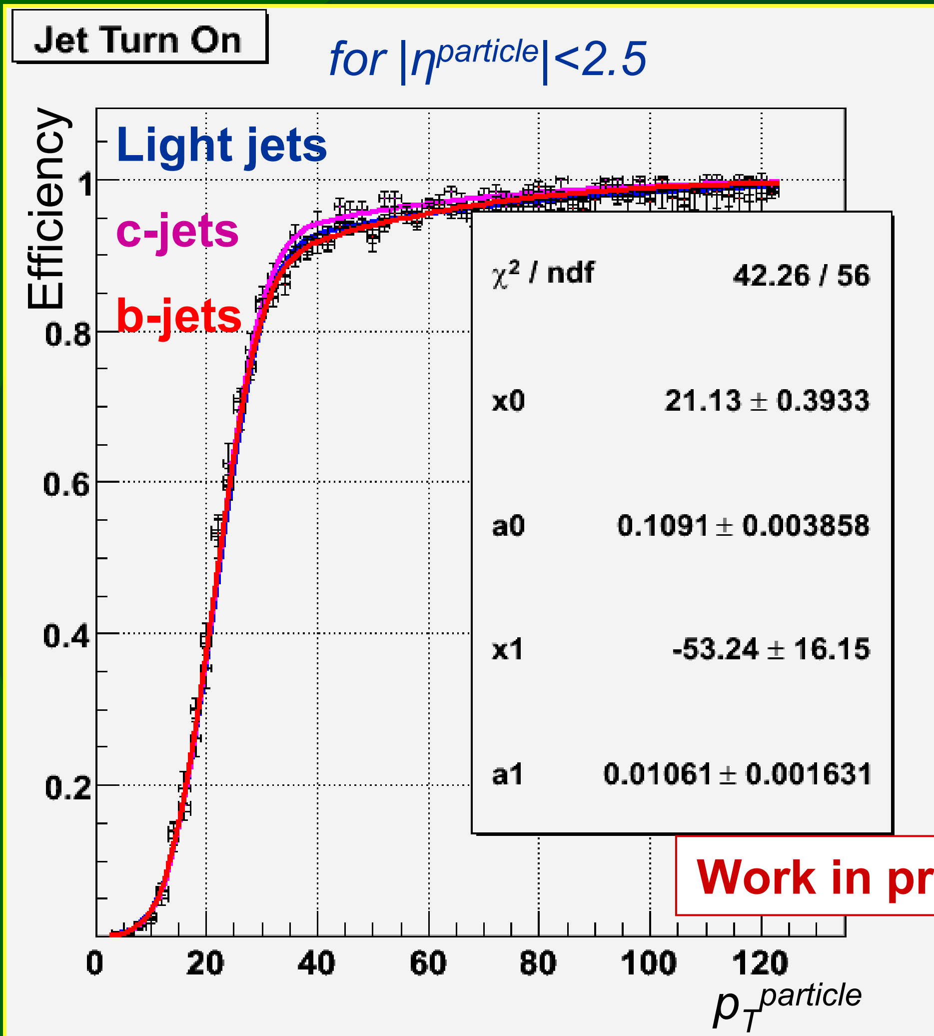
Reconstruction as a top pair washes it out ☺

Table 2: Simulated W asymmetries and reconstructed top pair asymmetries in W plus jets events, broken down by heavy flavor composition. The labels for additional jets are 'j' for a light (u,d,s, and g) jet, 'J' for a light or charm jet, and 'x' for any jet. The last column offers a comparison with the MC@NLO prediction, which was made with significantly different cuts as explained in the text.





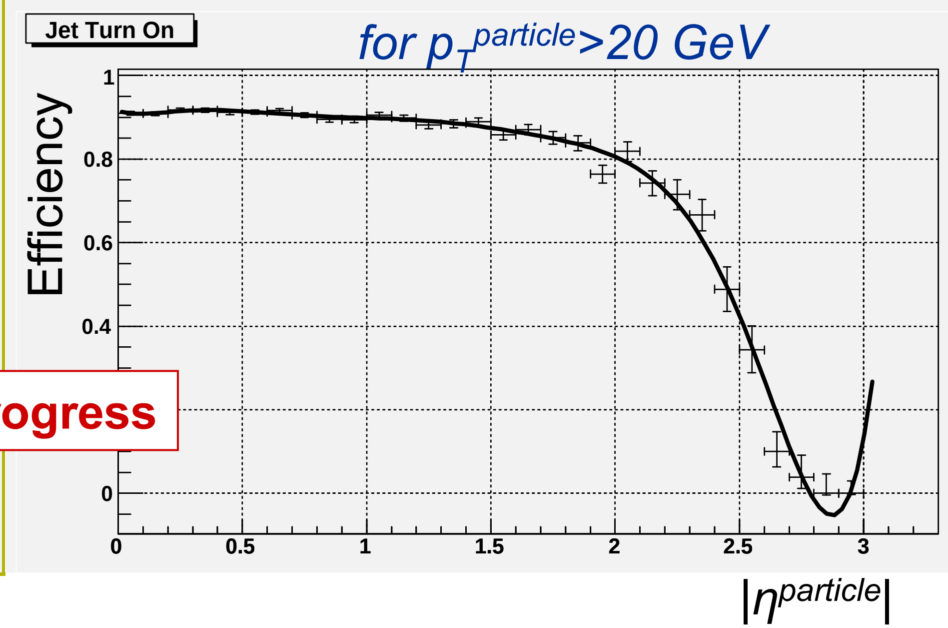
Jet Turn-On Curves



$$f(p_T) = \frac{1}{4}\left[1 + erf\left(a_0(p_T - x_0)\right)\right] \cdot \left[1 + erf\left(a_1(p_T - x_1)\right)\right]$$

Particle jets are reconstructed from stable particles excluding neutrinos (and W decay products in W+jets events) using the PXCONE algorithm.
(with R=0.5, OVLIM=0.5, p_T>3GeV).

This yields flavor-independent turn ons, as desired.



Looking at $g(|\eta|)$ plots with various p_T thresholds: factorization works much better than...

Full Closure Test Results

Sample	Asymmetries (in %) Fitted Turn Ons	Asymmetries (in %) Simple Cuts	Difference (absolute, in %)
W boson + > 1 jets	23.62 ± 0.14	23.40 ± 0.13	0.22 ± 0.06
W boson + > 2 jets	23.20 ± 0.24	23.13 ± 0.23	0.07 ± 0.19
W boson + > 3 jets	24.4 ± 0.7	24.2 ± 0.7	0.1 ± 0.7
W boson + > 4 jets	30 ± 4	29 ± 4	1 ± 4
$t\bar{t}$ for > 4 jets	1.31 ± 0.30	2.16 ± 0.26	-0.85 ± 0.19
$t\bar{t}$ for > 5 jets	-7.1 ± 0.7	-7.0 ± 0.6	0.0 ± 0.6
$t\bar{t}$ for > 6 jets	-16.5 ± 1.8	-15.5 ± 1.6	-1.0 ± 1.7
diluted W boson + > 1 jets	17.83 ± 0.12	17.73 ± 0.11	0.10 ± 0.05
diluted W boson + > 2 jets	17.59 ± 0.20	17.63 ± 0.20	-0.04 ± 0.16
diluted W boson + > 3 jets	18.6 ± 0.6	18.3 ± 0.6	0.3 ± 0.6
diluted W boson + > 4 jets	21.7 ± 3.1	19.7 ± 3.4	2 ± 4
diluted $t\bar{t}$ for > 4 jets	0.83 ± 0.24	1.44 ± 0.21	-0.61 ± 0.16
diluted $t\bar{t}$ for > 5 jets	-4.9 ± 0.5	-4.7 ± 0.5	-0.2 ± 0.4
diluted $t\bar{t}$ for > 6 jets	-10.7 ± 1.5	-10.3 ± 1.3	-0.5 ± 1.4

Full error propagation from six independent samples

Work in progress

Table 2: Effect of using simple box cuts to evaluate a model's asymmetry. The last column lists the absolute difference for each sample, given in percent. The jets considered are those that pass either the fitted turn ons (in the second column) or the simple box cuts (in the third column). In the last two section "diluted" refers to samples weighted by the dilution factor, thus predicting the observable asymmetry after reconstruction effects.

The Best Variables

For each I list their names in my plots and how well each variable does for the main criteria on a scale of 1-5: (separation from W+jets, modeling, no $|\Delta Y|$ bias).

(5,4,2) evLBPt – the leading b-jet's p_T

(5,4,3) chi2 – hitfit's χ^2 for the best jet assignment

(5,5,5) Mbb – invariant mass of the two b-jets selected in the best assignment

- This can find g→bb, but as hitfit's assignment might be random in background, it probably works because in top decays the b-jets tend to be roughly back to back.

(3,5,4) lowestMqq – the event's lowest invariant mass of two non b-tagged jets.

(3,5,3) Mjj – of the jets hitfit assigned to the W in the best assignment

(3,3,3) Ktmin – dR*min(pT) of the two closest jets

- separates better than the minimal P_{Trel}. Why?
- can avoid overall-JES dependence by normalizing in jets' H or H_T. But this reduces separation. Using the 5-object H keeps some separation, but prefers central jets as H_T/H is centrality!

(2,3,5) evH – energy sum of the jets and lepton

- Bad data-MC agreement with 15 GeV jets, reasonable with 20GeV jets.
- H_T=centrality*H offers a much better separation and an acceptable bias. Will use if 100% needed.

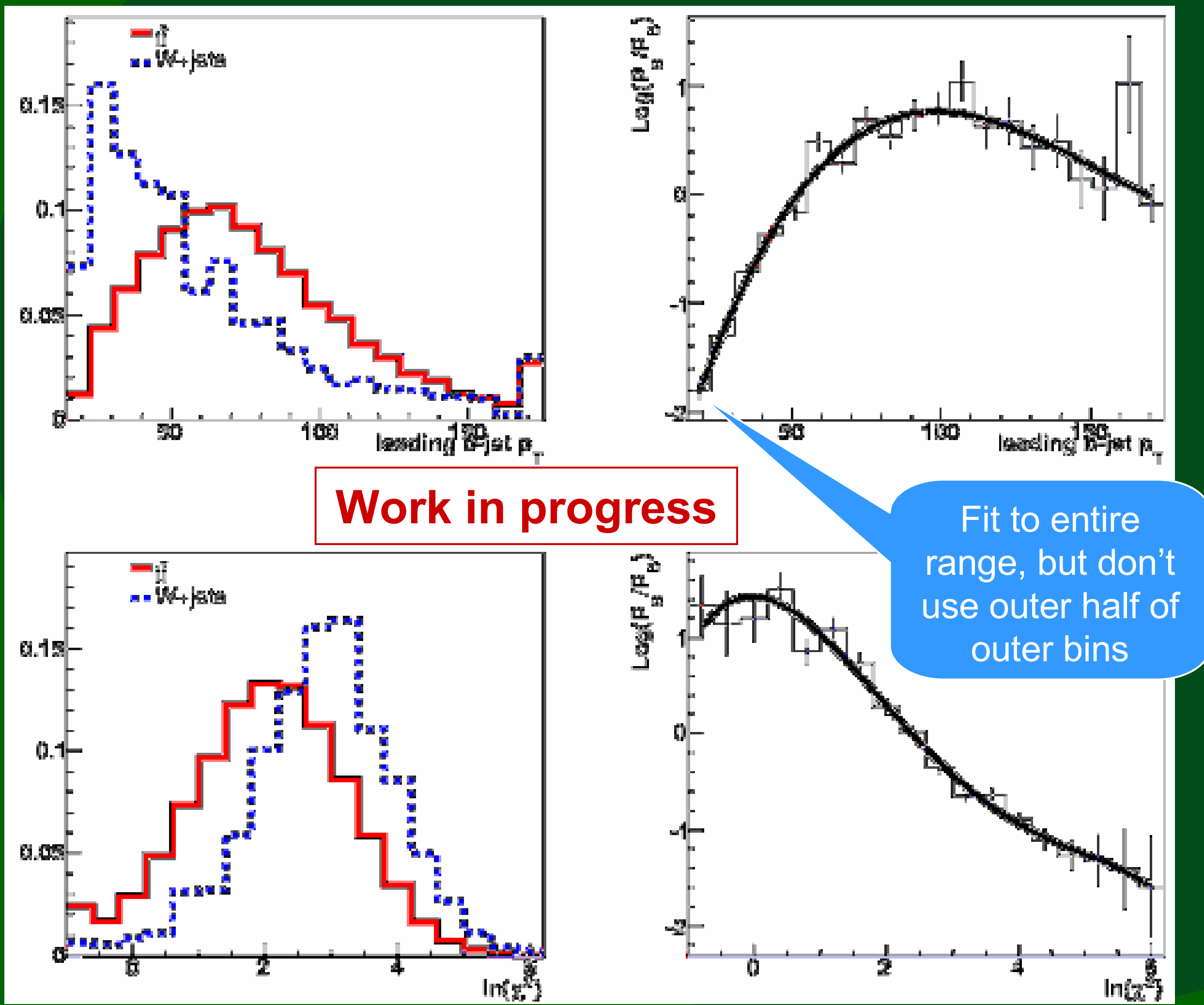
(1,4,3) aBPt – the pT asymmetry between those two jets.

- Barely any separation, looking at it just because it has few correlations

BTW: Good variables that choose central tops: $\Delta Y_{j1,j2}$, $\cos\theta^{*1}$, dRmax, dRmax5, Ktmin/H

Fits to Likelihood Variables

Here are the two best variables:



Can also calculate a rough F.O.M. for the separation each variable gives using the fit and the distributions.

Variable	on s+b	on b.
evLBPt	0.111	0.137
log(chi2)	0.114	0.105
log(lowestMqq)	0.051	0.051
Mbjj	0.070	0.095
log(evH)	0.029	0.039
Ktmin	0.071	0.068
aBPt	0.007	0.008
Mbb	0.040	0.047
Mjj	0.040	0.044
log(highestMqq)	0.031	0.043

$$p = \frac{P(x \mid signal)}{P(x \mid signal) + P(x \mid bkg.)}$$

$$d = 2p - 1 \quad D = d^2 = (2p-1)^2$$

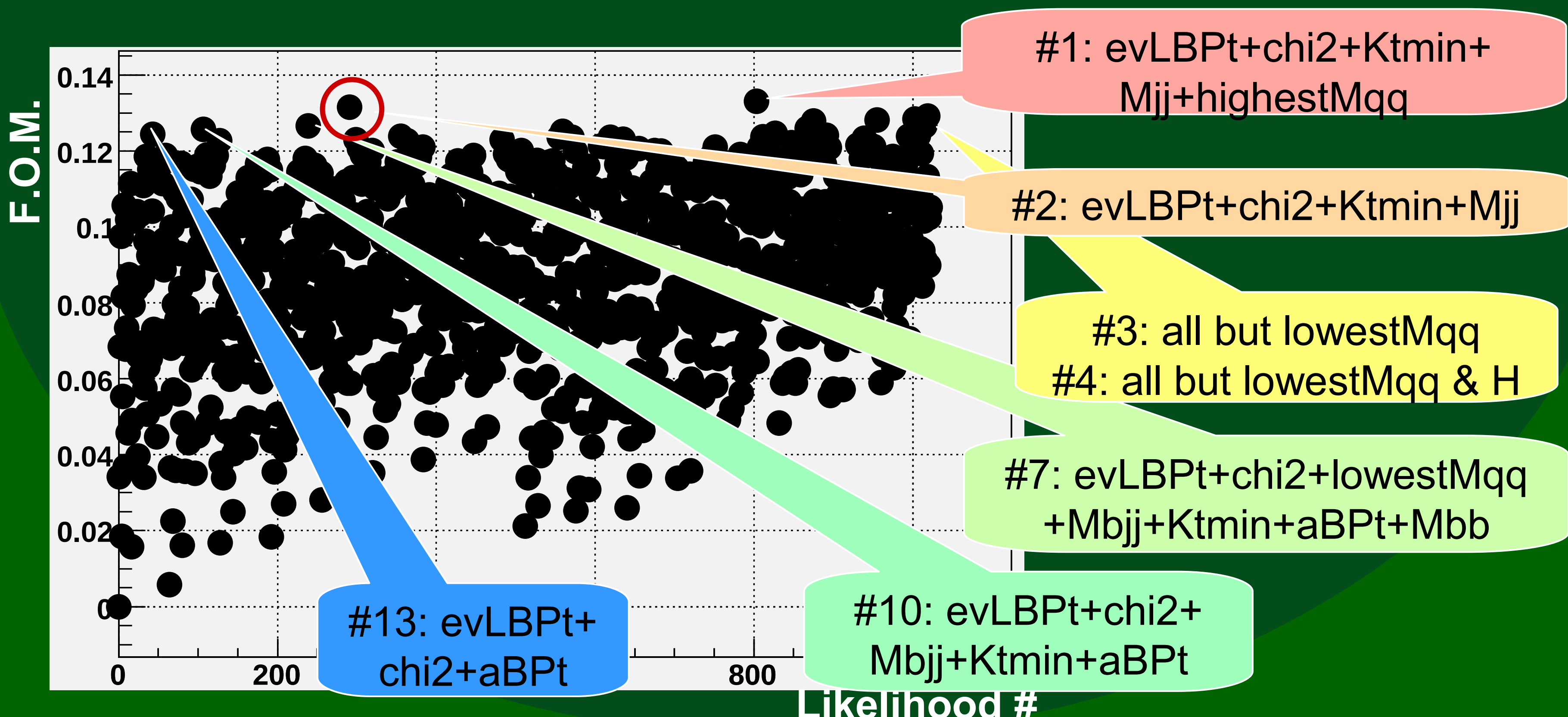
$$F.O.M. = \int D(x) f(x) dx$$

Choosing By F.O.M.

Can also choose a likelihood by trying them all.

I just reuse the same Figure Of Merit (at least it's fast: 10 seconds).

To account for correlations, the probability comes from the distributions of L, not L itself; the separation is evaluated using the fractions I fitted to data in the previous iteration (with $M_{l\nu}$ & evLBPt).



N_{sig}-A_{fb} Correlations

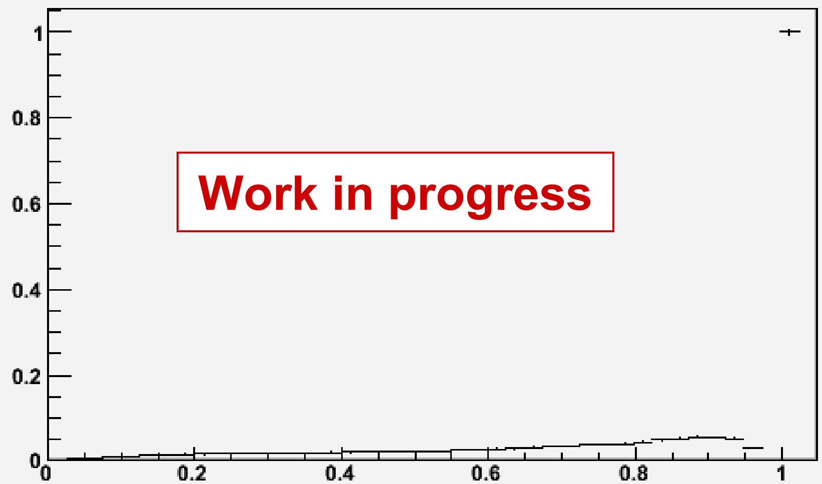
N_{sig} and A_{fb} are correlated in two ways:

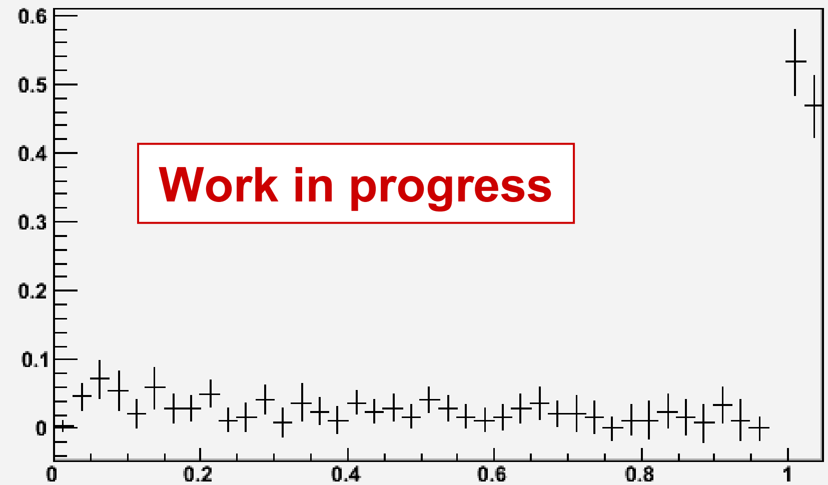
- A larger N_{sig} implies the observed asymmetry must be assigned to more events
- Since $N_{sig}=N_{+}+N_{-}$, a statistical fluctuation in, e.g., N_{+} would increase both.
    - If $N_{+}+N_{-}$ this cancels out with the similar effect from N_{-}. But they differ.

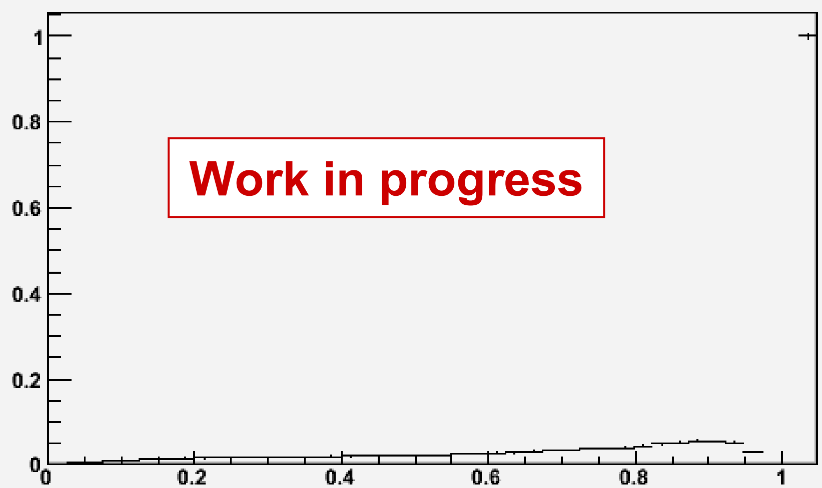
Untangling it by doing a simultaneous fit.

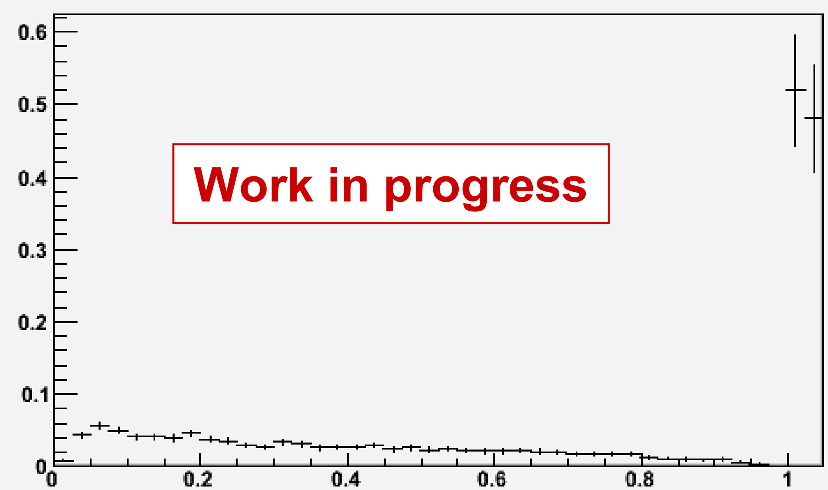
Need ensemble tests to prove whether cutting on the likelihood discriminate is acceptable.

Templates with asymmetry

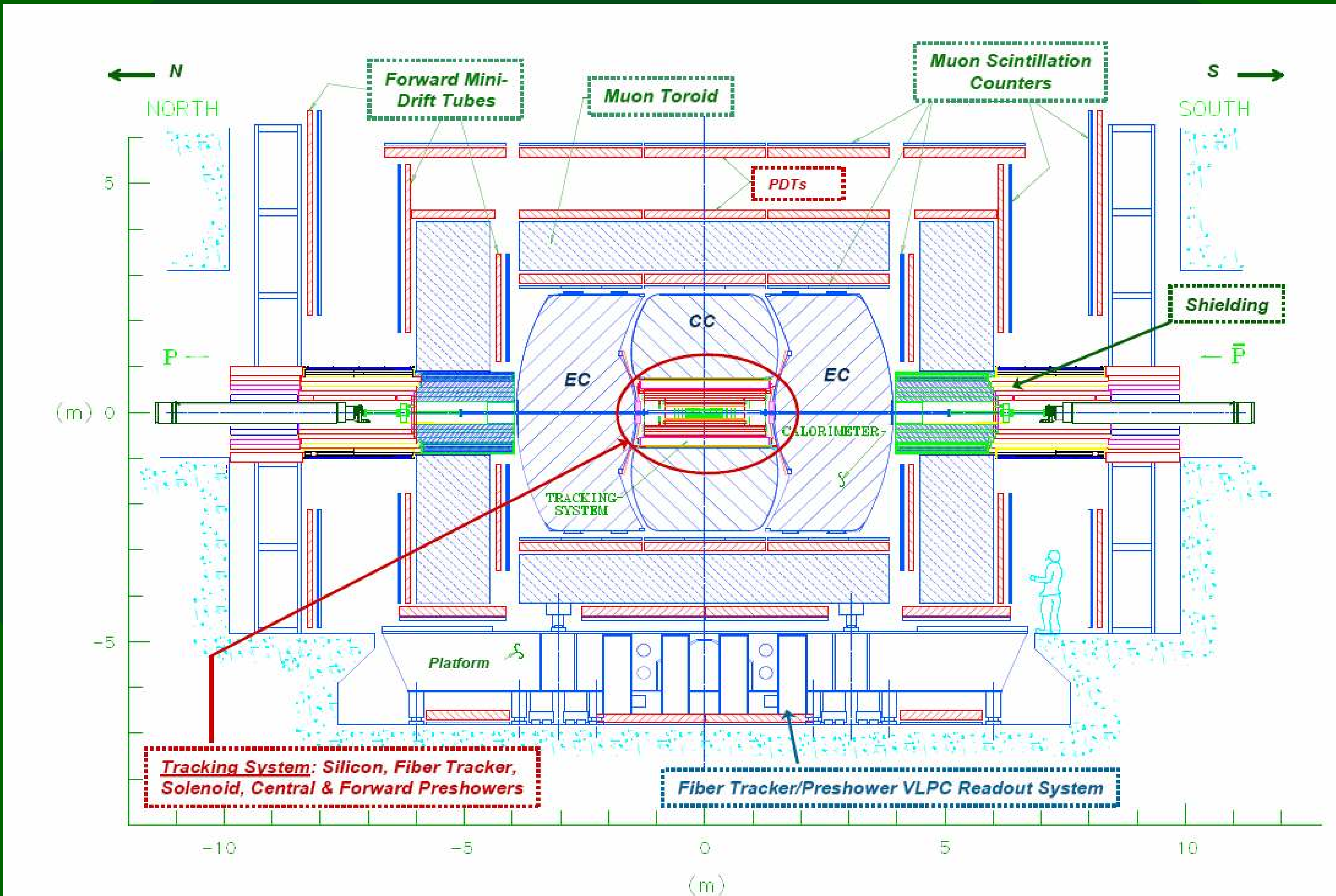








The DØ Detector



Matched MC

Good at generating hard, large-angle processes (calculates interference)

Weak on the “texture” of the QCD radiation

counting of

Good at generating the details within a jet

Multijet events don’t describe data well (and are hard to generate)

Matrix element generator

Alpgen / Sherpa

LO calculations for 2→N hard processes

Hard scatter partons

Parton shower generator

Pythia / Herwig

Only 2->2 hard processes

Resummed soft, collinear radiation.

Particles

Detector simulation