Event Shapes in $t\bar{t}$ and QCD Events with ATLAS

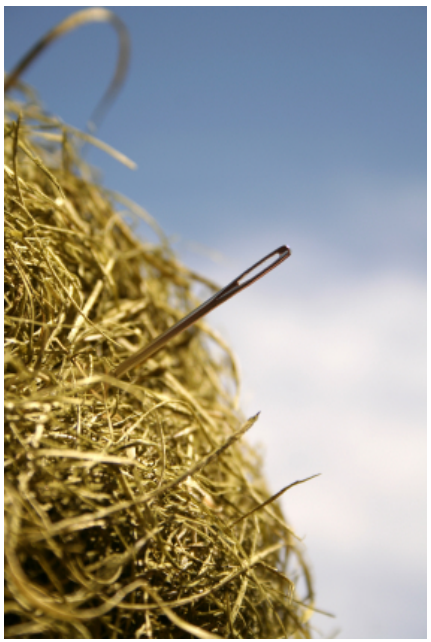
DPG Frühjahrstagung, Bonn

Martin Härtig

Max-Planck-Institut für Physik, München
Fakultät für Physik, Universität Würzburg

18. März 2010

Introduction



Problem

The $t\bar{t}$ all hadronic decay channel is very important, but suffers from large QCD multijet background

- $S/B \approx 10^{-5}$
- Preselection, B tag can improve the S/B ratio
  $\Rightarrow$ But is there room for more?
- We want to use Event Shapes to discriminate $t\bar{t}$ signal from (QCD) background

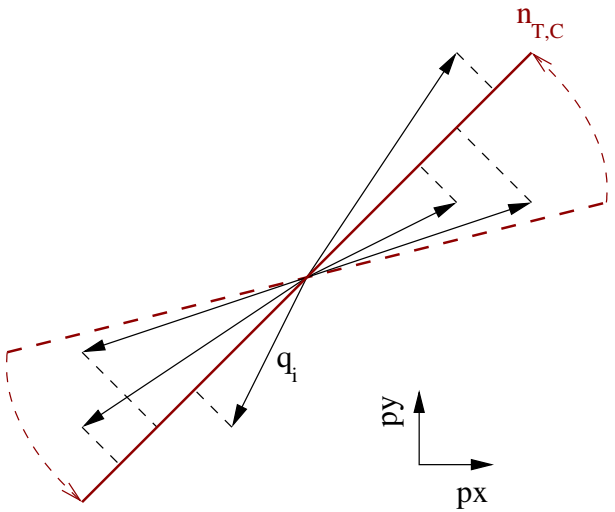
http://shrink4men.files.wordpress.com/2009/10/needle-haystack.jpg

What are Event Shapes?

Event Shapes describe the final state in a collision experiment

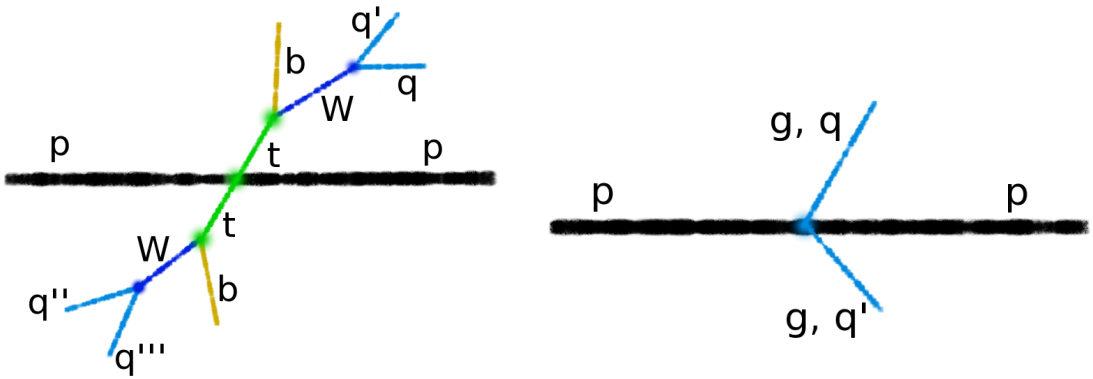
Example: Transverse Thrust

$$T_{\perp,\mathcal{C}} = \max_{\vec{n}_{T,\mathcal{C}}} \frac{\sum_{i\in\mathcal{C}} |\vec{q}_{\perp i} \cdot \vec{n}_{T,\mathcal{C}}|}{\sum_{i\in\mathcal{C}} |\vec{q}_{\perp i}|}$$



- Calo Clusters are projected on Thrust Axis $\vec{n}_{T,\mathcal{C}}$
- Thrust axis is chosen so to maximize the Thrust value
- Require leading jets to be within central region $\mathcal{C}$, Calo Clusters taken from $|\eta_c| < 2.5$
- For a collimated event: $T_{\perp,\mathcal{C}} \to 1$, isotropic event: $T_{\perp,\mathcal{C}} \to 1/2$

Why Event Shapes?

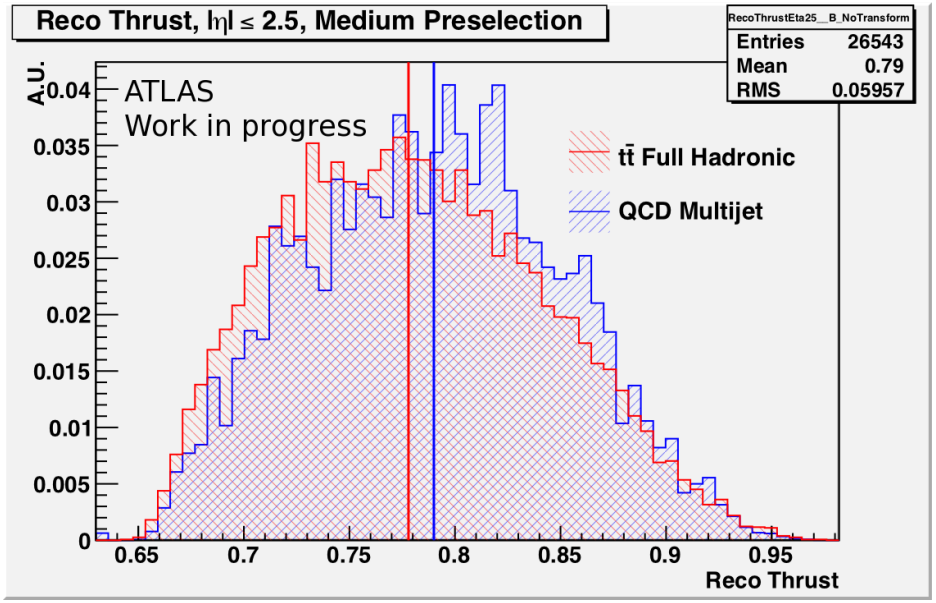


- Event Shapes are theoretically well understood and experimentally accessible
- Important for QCD and tuning Monte Carlo generators
- "Complicated" events ($t\bar{t}$) are more isotropic, "simple" events (QCD) are more collimated
- Jets with higher invariant masses have a greater opening angle

Event Shapes in $t\bar{t}$

$\Rightarrow$ Event Shapes should reflect if there is a top quark

Event Shapes in Signal and Background



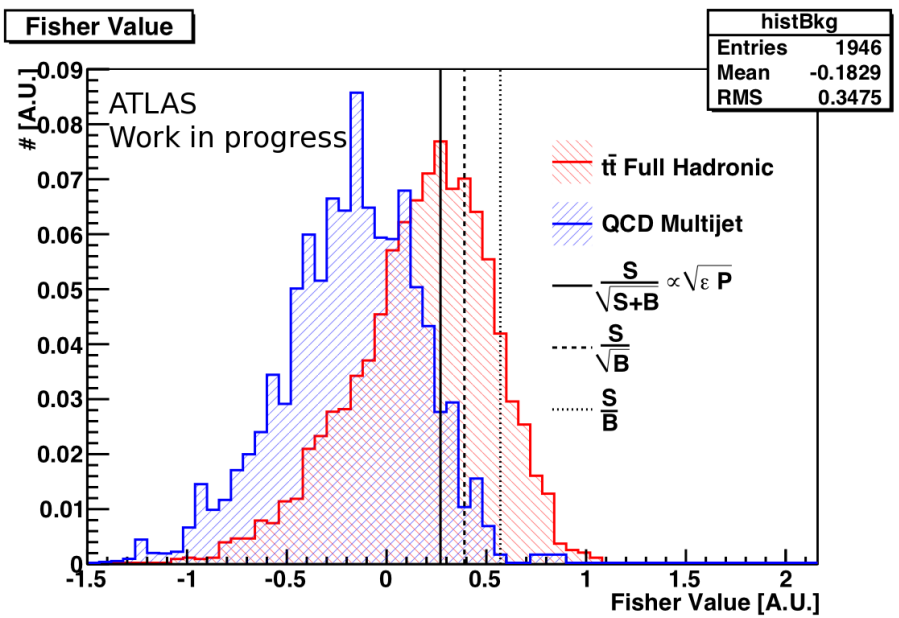
- Example:
  Calculated Thrust value in a $t\bar{t}$ All Had sample,
  QCD multijet sample
- Works well for Thrust
- Works even better for other Event Shapes

Figure: Transverse Thrust,
Vertical bars indicate sample means

- Instead of picking just one Event Shape,
  we combine the discrimination power of several

From Thrust to Multivariate Analysis

- In principle, can now use Event Shapes to discriminate Signal/Background
- One variable is not powerful enough
  $\Rightarrow$ Combine multiple Event Shapes



- Using ROOT's TMVA package, tried out different approaches, found Fisher discriminant method to work best
- Fisher discriminant uses a linear combination of input variables

Result

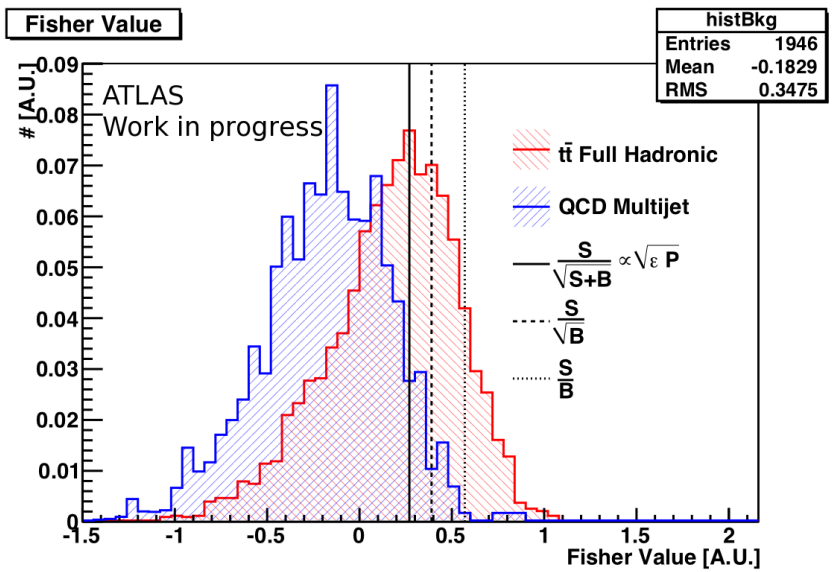
- Signal: $t\bar{t}$ Full Hadronic (MC@NLO, $\sqrt{s} = 10\,\text{TeV}$)
- Background: QCD Multijet Events (Pythia, $\sqrt{s} = 10\,\text{TeV}$)
- Preselection: Six jets, $p_\perp > 20\,\text{GeV/c}$, $|\eta| < 2.5$
  At least one B tag, missing $E_T < 20\,\text{GeV}$

Cut flow for $\mathcal{L} = 200\,\text{pb}^{-1}$, optimized for S/B:

	Before Cuts	After Preselection	After Fisher
Signal	$3.0 \cdot 10^4$	$4.3 \cdot 10^3$	580
Background	$7.6 \cdot 10^6$	$5.8 \cdot 10^4$	400
S/B	$1/255$	$1/13$	1.45

Result

- Optimization of the Fisher cut value with respect to different estimators yield:



$$\frac{S}{\sqrt{S+B}} \approx 23 \qquad (\text{Fi} = 0.27)$$

$$\frac{S}{\sqrt{B}} \approx 30 \qquad (\text{Fi} = 0.39)$$

$$\frac{S}{B} \approx 1.45 \qquad (\text{Fi} = 0.57)$$

Problem

Low $p_\perp$ multijet samples were only available with small $\mathcal{L}$, too few events pass the preselection

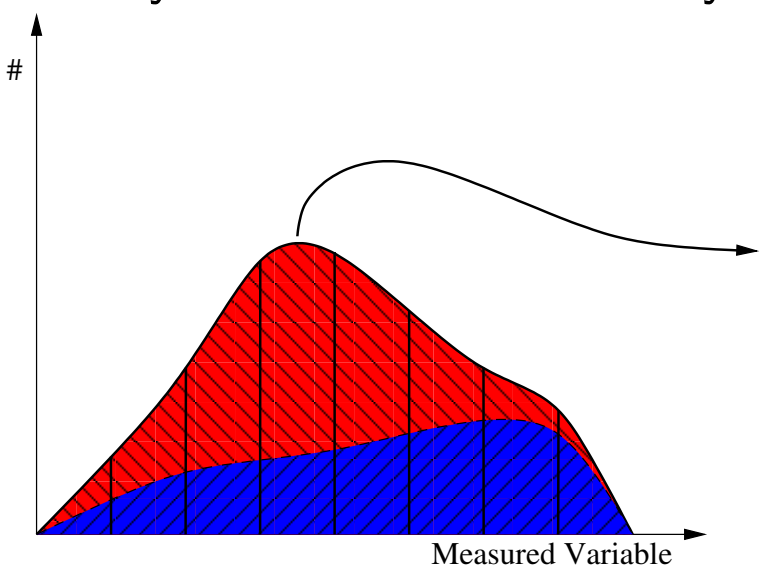
Measuring Event Shapes (Work in Progress)

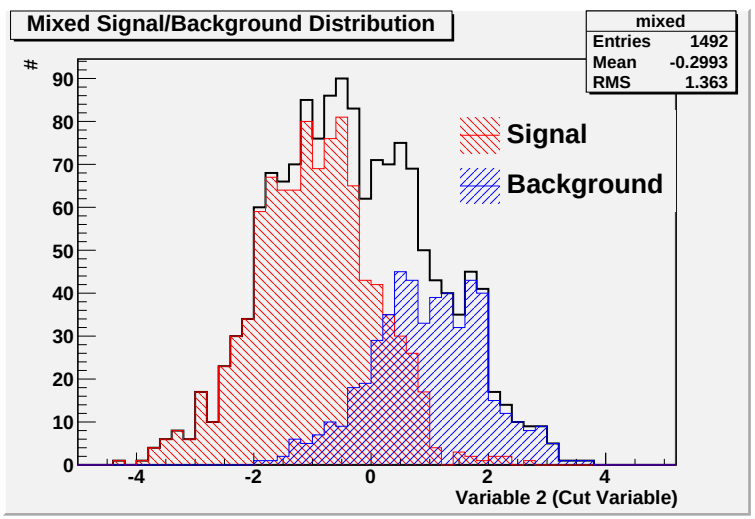
Next, we want to measure an Event Shape distribution

Problem

Template methods use the shape of the distribution,
but we expect that the Monte Carlo generators are not very reliable

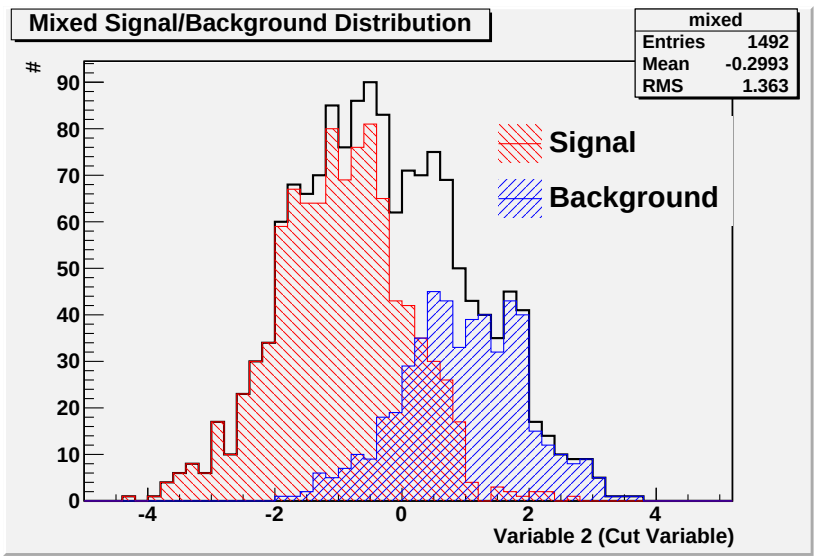
We try to circumvent this by looking at individual bins:





Measuring Event Shapes (Work in Progress)

- Assuming the distribution across one bin is predicted right and efficiencies obtained from MC simulations are accurate, then



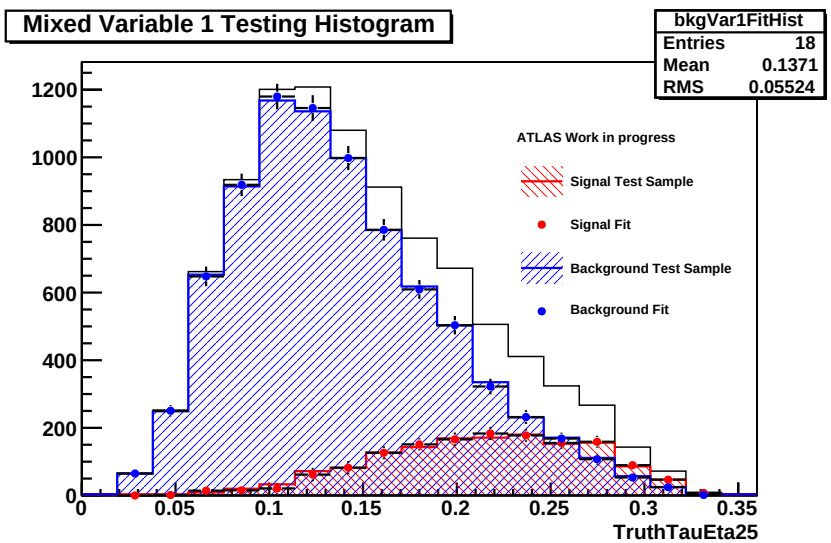
$$N_{\text{obs}}^{(\text{cut } 1)} = \varepsilon_{\text{sig}}^{(1)} \cdot N_{\text{sig}} + \varepsilon_{\text{bkg}}^{(1)} \cdot N_{\text{bkg}}$$
$$N_{\text{obs}}^{(\text{cut } 2)} = \varepsilon_{\text{sig}}^{(2)} \cdot N_{\text{sig}} + \varepsilon_{\text{bkg}}^{(2)} \cdot N_{\text{bkg}}$$

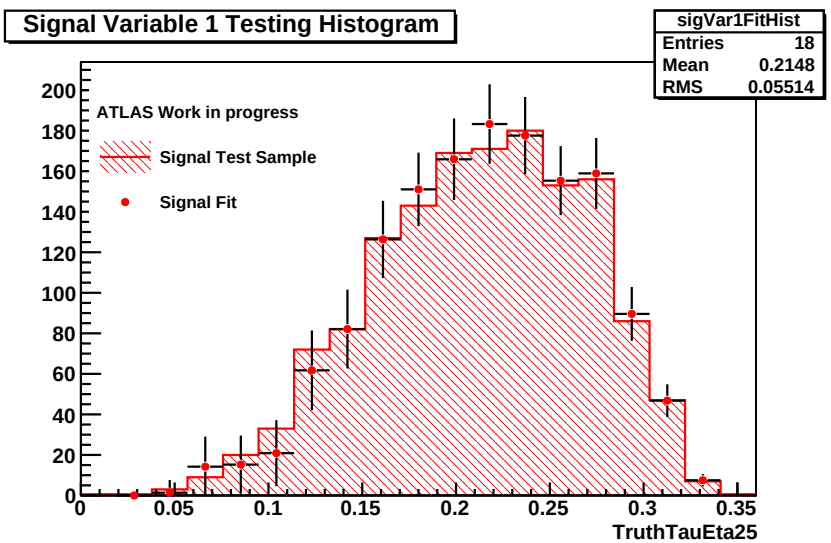
- This can be solved for two cuts
- For more then two cuts, use a least squares fit
- Repeat the fit in every bin

Measuring Event Shapes (Work in Progress)

As an example:

- Signal: $t\bar{t}$ Full Hadronic, background: QCD Multijet
- Observation variable: $\tau = 1 - T_{\perp,\mathcal{C}}$
- Cut variable: Number of B tags in event (≥ 0, ≥ 1, ≥ 2)
- Efficiencies from $100\,\mathrm{pb}^{-1}$ signal, $70\,\mathrm{pb}^{-1}$ background
- Tested with $50\,\mathrm{pb}^{-1}$ subsample





Conclusion

- Event Shapes are a valuable tool for discriminating top quark events from background
- Also valuable for measuring QCD properties and tuning Monte Carlo generators
- We found an optimal S/B value of 1.45 in the all hadronic decay channel using the Fisher method
- Currently working on a method to measure signal and background distributions of Event Shapes.
  Sub-sample tests look promising, but still have to compare strongly differing samples

Samples, BTagging

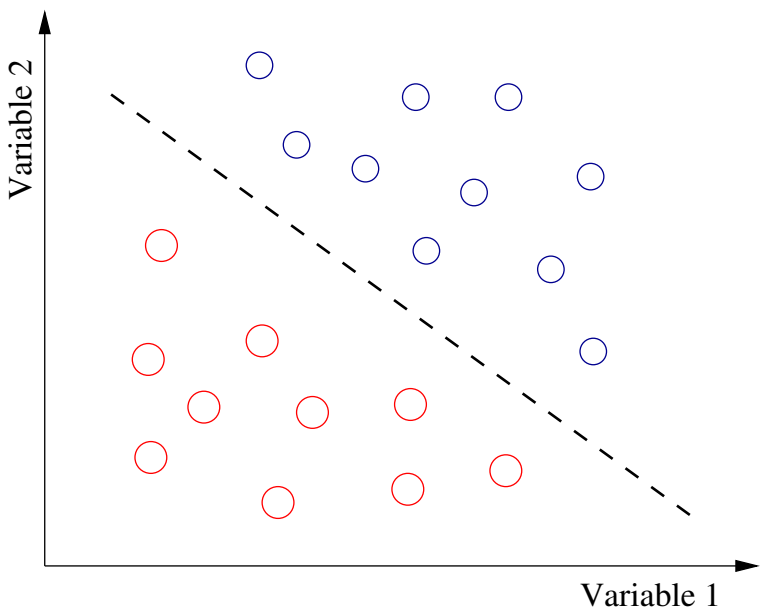
Samples used

- 105204 tt MC@NLO Jimmy e363 s462 r563 skimmed
- 105013 J4 Pythia JetJet e344 s475 r586 skimmed
- 105014 J5 Pythia JetJet e344 s475 r586 skimmed
- 105015 J6 Pythia JetJet e344 s475 r586 skimmed

- Cluster: CaloCalTopoCluster
- Jets: AntiKt4LCTopoJets
- Missing Et: MET_MuonBoy + MET_LocHadTopoObj
- B Tag: Cone4H1TowerJets

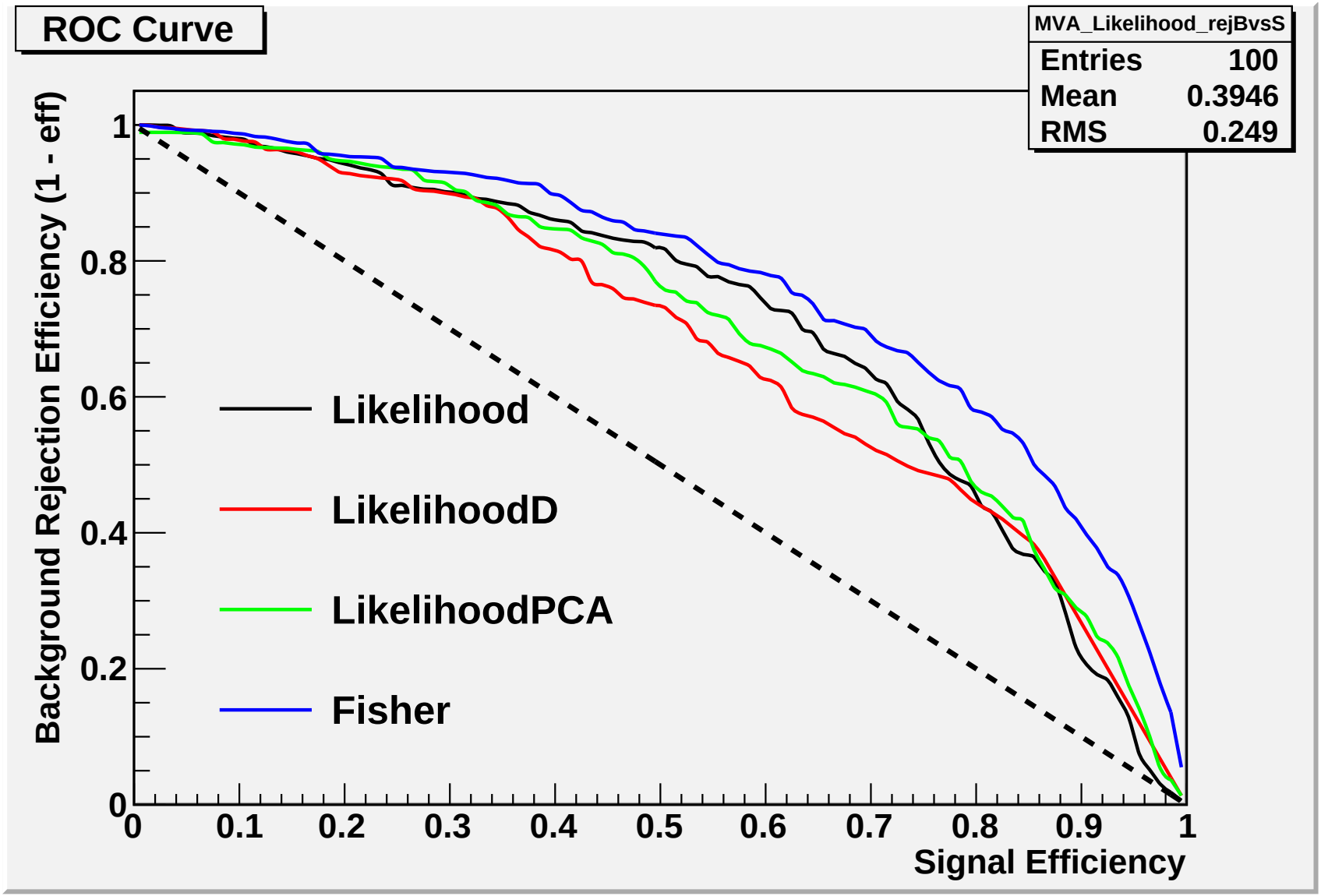
Fisher Discriminant

- Mathematically, transforms the variable space
- A linear combination of all input variables
- Determines a hyper plane which separates the signal and background as far as possible



- Works best on uncorrelated gaussian distributions with different mean values

Results


ROC Curve
MVA_Likelihood_rejBvsS
Entries 100
Mean 0.3946
RMS 0.249
Background Rejection Efficiency (1 - eff)
Signal Efficiency
Likelihood
LikelihoodD
LikelihoodPCA
Fisher
0
0.2
0.4
0.6
0.8
1
0.1
0.3
0.5
0.7
0.9