# Background Suppression via Multivariate Analysis

Chase Peterson

August 13, 2021

Jefferson Lab
EXPLORING THE NATURE OF MATTER

# TMVA 4: Toolkit for Multivariate Data Analysis with ROOT

Using modified TMVAClassification and TMVAClassificationApplication macros:

Training

```
TMVAClassification(TString myMethodList = "BDT",
                   TString bkg_file_name = "evtgen-all-gp-to-etap-skim-4g-training-jef.root",
                   TString sig_file_name = "evtgen-ggpi0-gp-to-etap-skim-4g-training-jef.root",
                   TString output_file_name = "tmva-test.root")
```

Testing

```
TMVAClassificationApplication(TString myMethodList = "BDT",
                   TString inputName = "evtgen-all-gp-to-etap-skim-4g-testing-jef.root",
                   TString weightfile = "dataset/mva_weight/TMVAClassification_BDT.weights.xml",
                   TString outfileName = "tmva-test-app.root")
```

<u>Current Methods:</u>

Boosted Decision Trees (BDT)

Fisher Discriminants (Fisher)

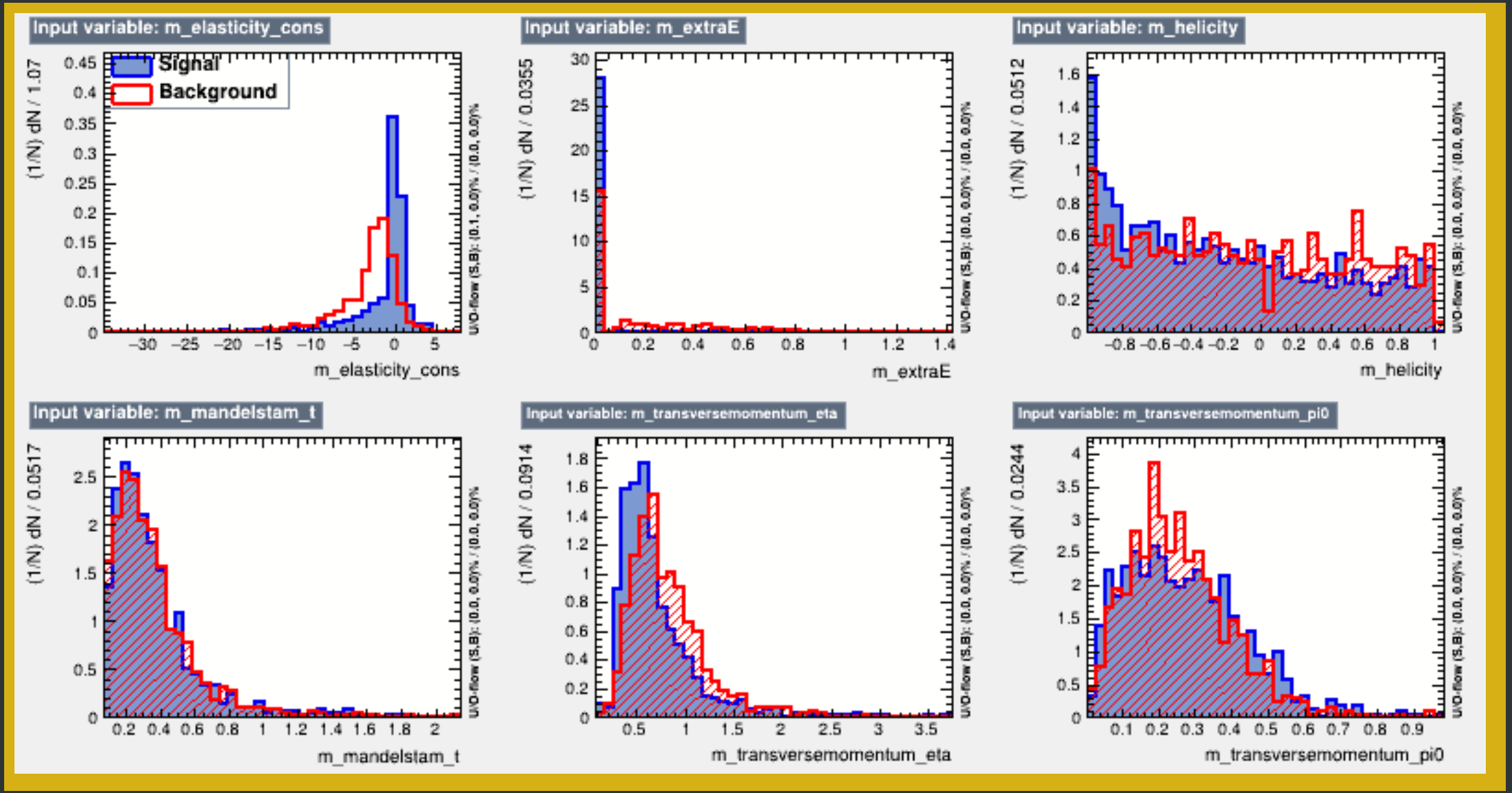
# TMVA 4: TMVAClassification Input Variables and Spectators

```
//dataloader->AddVariable( "m_chi2"    , "", "", 'F' );
//dataloader->AddVariable( "m_elasticity_meas"    , "", "", 'F' );
dataloader->AddVariable( "m_elasticity_cons"    , "", "", 'F' );
dataloader->AddVariable( "m_extraE"    , "", "", 'F' );
//dataloader->AddVariable( "m_missingmasssquare"    , "", "", 'F' );
//dataloader->AddVariable( "m_coplanarity"    , "", "", 'F' );
dataloader->AddVariable( "m_helicity"    , "", "", 'F' );
//dataloader->AddVariable( "m_insertnumber"    , "", "", 'I' );
//dataloader->AddVariable( "m_fcalnumber"    , "", "", 'I' );
dataloader->AddVariable( "m_mandelstam_t"    , "", "", 'F' );
dataloader->AddVariable( "m_transversemomentum_eta"    , "", "", 'F' );
dataloader->AddVariable( "m_transversemomentum_pi0"    , "", "", 'F' );
dataloader->AddVariable( "m_transversemomentum_2g"    , "", "", 'F' );
//dataloader->AddVariable( "m_invariantmass2g"    , "", "", 'F' );
dataloader->AddVariable( "m_shower_x0",   "", "", 'F');
dataloader->AddVariable( "m_shower_x1",   "", "", 'F');
dataloader->AddVariable( "m_shower_x2",   "", "", 'F');
dataloader->AddVariable( "m_shower_x3",   "", "", 'F');
dataloader->AddVariable( "m_shower_y0",   "", "", 'F');
dataloader->AddVariable( "m_shower_y1",   "", "", 'F');
dataloader->AddVariable( "m_shower_y2",   "", "", 'F');
dataloader->AddVariable( "m_shower_y3",   "", "", 'F');
/*
  dataloader->AddVariable( "m_shower_z0",   "", "", 'F');
  dataloader->AddVariable( "m_shower_z1",   "", "", 'F');
  dataloader->AddVariable( "m_shower_z2",   "", "", 'F');
  dataloader->AddVariable( "m_shower_z3",   "", "", 'F');
  dataloader->AddVariable( "m_shower_e0",   "", "", 'F');
  dataloader->AddVariable( "m_shower_e1",   "", "", 'F');
  dataloader->AddVariable( "m_shower_e2",   "", "", 'F');
  dataloader->AddVariable( "m_shower_e3",   "", "", 'F');
```

```
dataloader->AddSpectator( "m_chi2"    , "", "", 'F' );
dataloader->AddSpectator( "m_gammabeam_energy"    , "", "", 'F' );
dataloader->AddSpectator( "m_elasticity_meas"    , "", "", 'F' );
//dataloader->AddSpectator( "m_elasticity_cons"    , "", "", 'F' );
//dataloader->AddSpectator( "m_extraE"    , "", "", 'F' );
dataloader->AddSpectator( "m_missingmasssquare"    , "", "", 'F' );
dataloader->AddSpectator( "m_coplanarity"    , "", "", 'F' );
//dataloader->AddSpectator( "m_helicity"    , "", "", 'F' );
dataloader->AddSpectator( "m_insertnumber"    , "", "", 'I' );
dataloader->AddSpectator( "m_fcalnumber"    , "", "", 'I' );
dataloader->AddSpectator( "m_signal"    , "", "", 'I' );
dataloader->AddSpectator( "m_weight"    , "", "", 'F' );
//dataloader->AddSpectator( "m_mandelstam_t"    , "", "", 'F' );
//dataloader->AddSpectator( "m_transversemomentum_eta"    , "", "", 'F' );
//dataloader->AddSpectator( "m_transversemomentum_pi0"    , "", "", 'F' );
//dataloader->AddSpectator( "m_transversemomentum_2g"    , "", "", 'F' );
dataloader->AddSpectator( "m_invariantmass2g"    , "", "", 'F' );
dataloader->AddSpectator( "m_invariantmass4g"    , "", "", 'F' );
dataloader->AddSpectator( "m_invariantmass2gpi0"    , "", "", 'F' );
dataloader->AddSpectator( "m_rawinvariantmass4g"    , "", "", 'F' );
dataloader->AddSpectator( "m_rawinvariantmass2gpi0"    , "", "", 'F' );
/*
dataloader->AddSpectator( "m_shower_x0",   "", "", 'F');
dataloader->AddSpectator( "m_shower_x1",   "", "", 'F');
dataloader->AddSpectator( "m_shower_x2",   "", "", 'F');
dataloader->AddSpectator( "m_shower_x3",   "", "", 'F');
dataloader->AddSpectator( "m_shower_y0",   "", "", 'F');
dataloader->AddSpectator( "m_shower_y1",   "", "", 'F');
dataloader->AddSpectator( "m_shower_y2",   "", "", 'F');
dataloader->AddSpectator( "m_shower_y3",   "", "", 'F');
*/
dataloader->AddSpectator( "m_shower_z0",   "", "", 'F');
dataloader->AddSpectator( "m_shower_z1",   "", "", 'F');
dataloader->AddSpectator( "m_shower_z2",   "", "", 'F');
dataloader->AddSpectator( "m_shower_z3",   "", "", 'F');
dataloader->AddSpectator( "m_shower_e0",   "", "", 'F');
dataloader->AddSpectator( "m_shower_e1",   "", "", 'F');
dataloader->AddSpectator( "m_shower_e2",   "", "", 'F');
dataloader->AddSpectator( "m_shower_e3",   "", "", 'F');
```

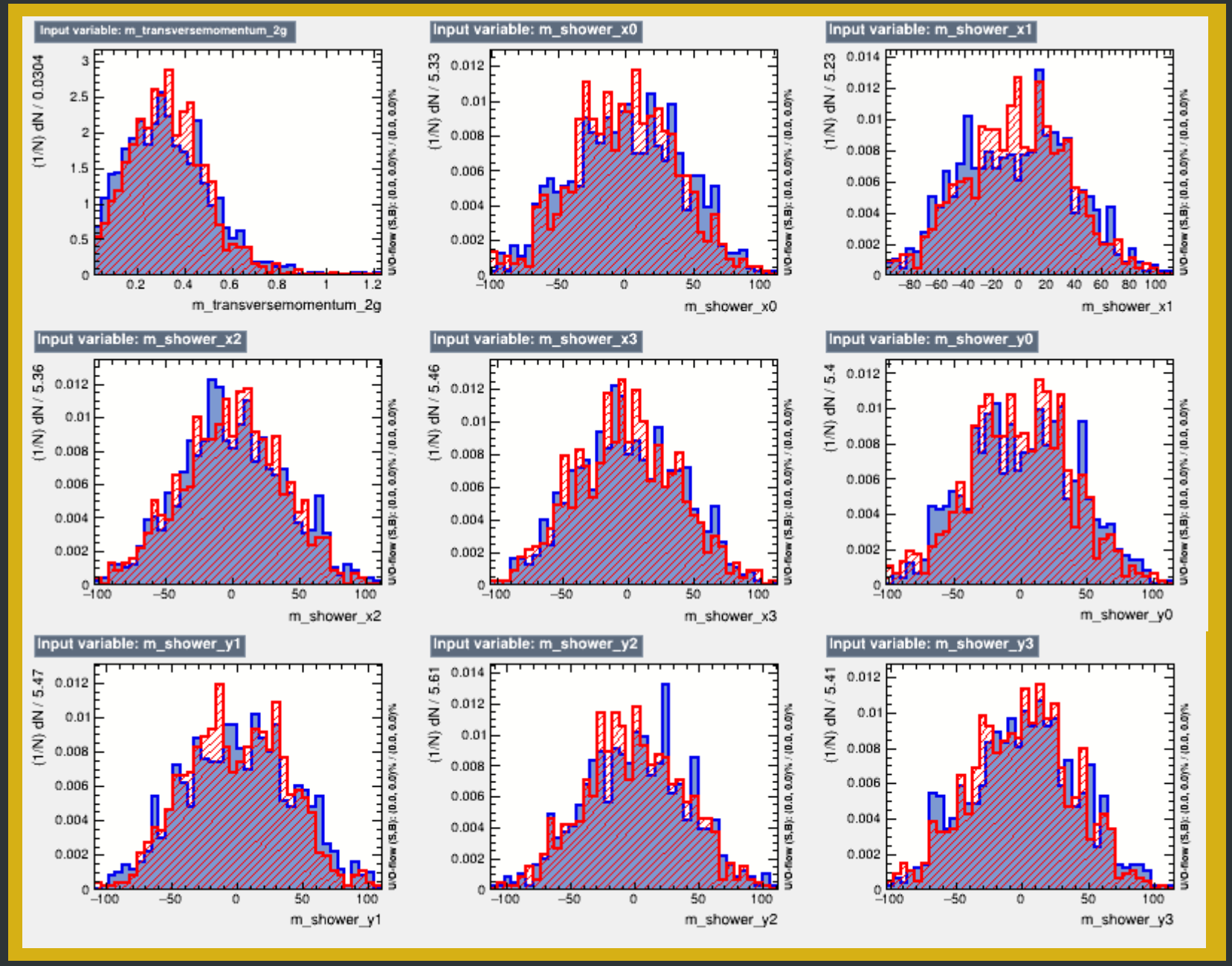
We are testing different combinations of input variables

# TMVA 4: Comparison between Signal and Background

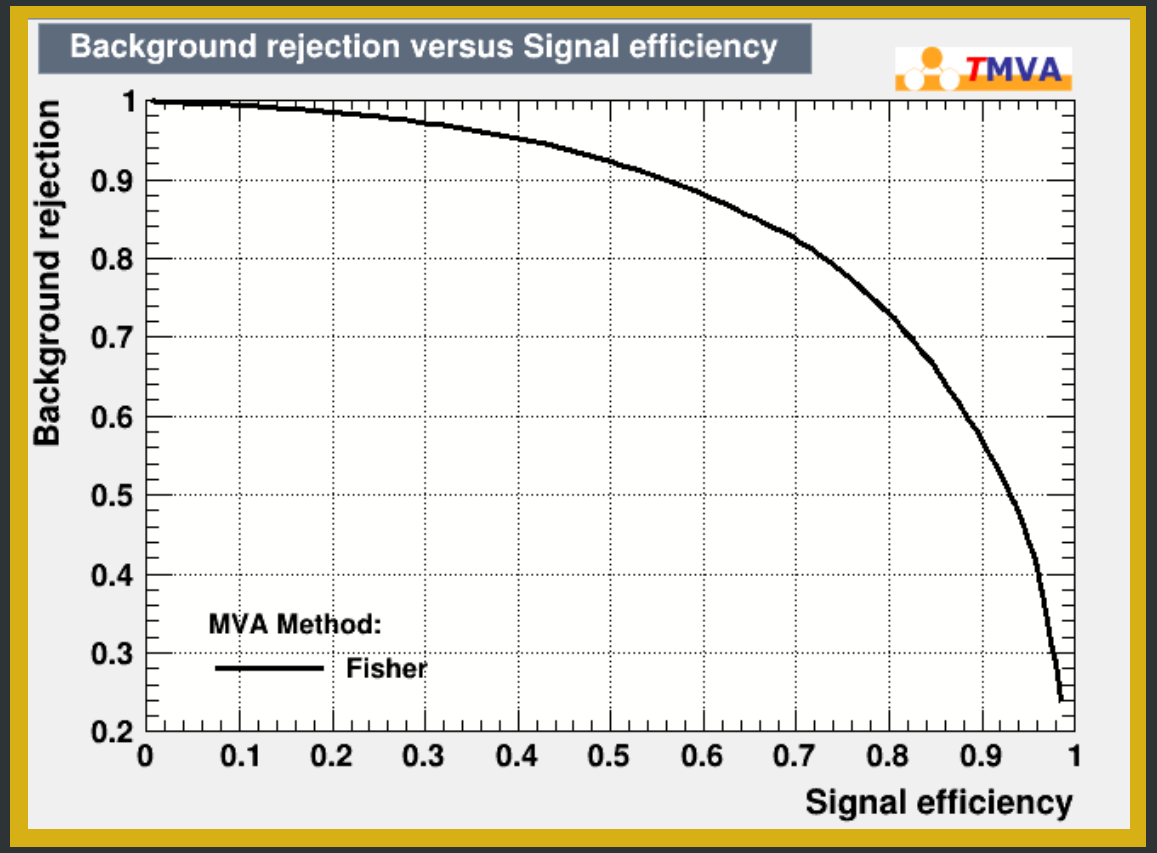



Classification by Suppression Power

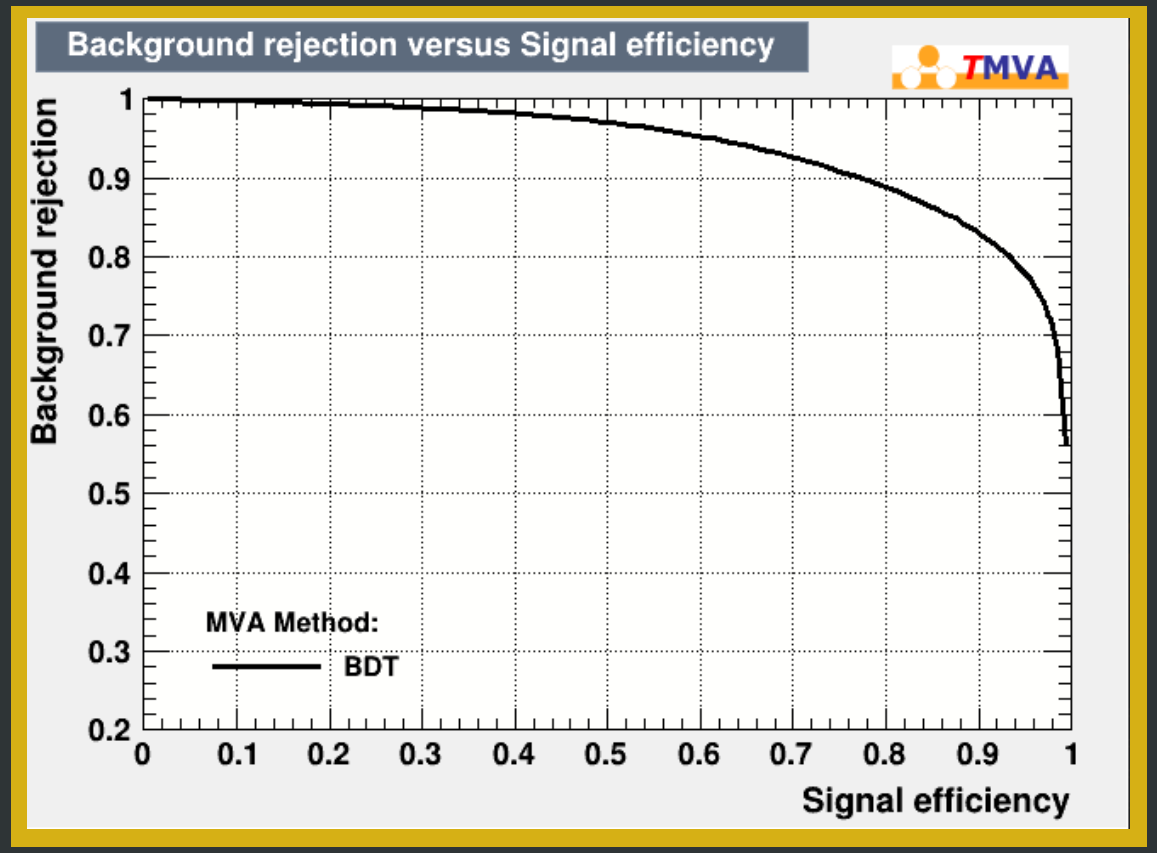
# TMVA 4: Comparison between Signal and Background



# TMVA 4: Methods



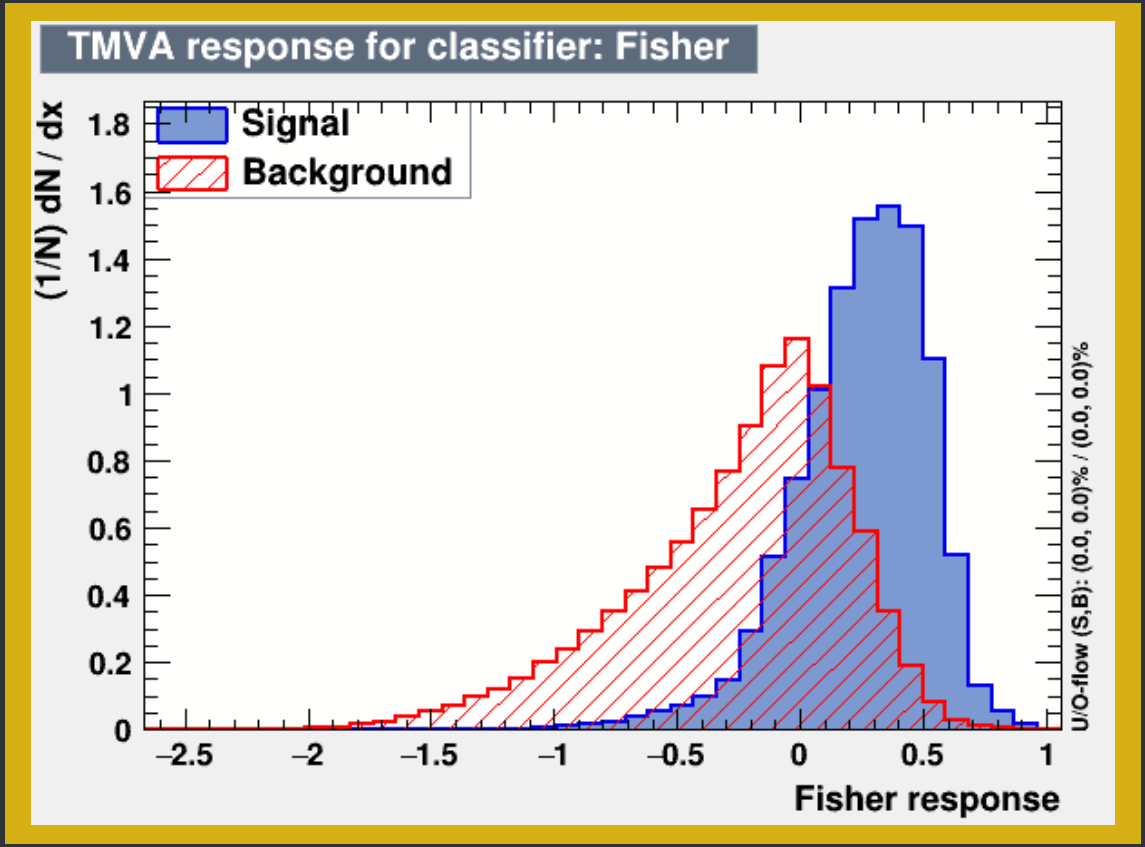

Fisher



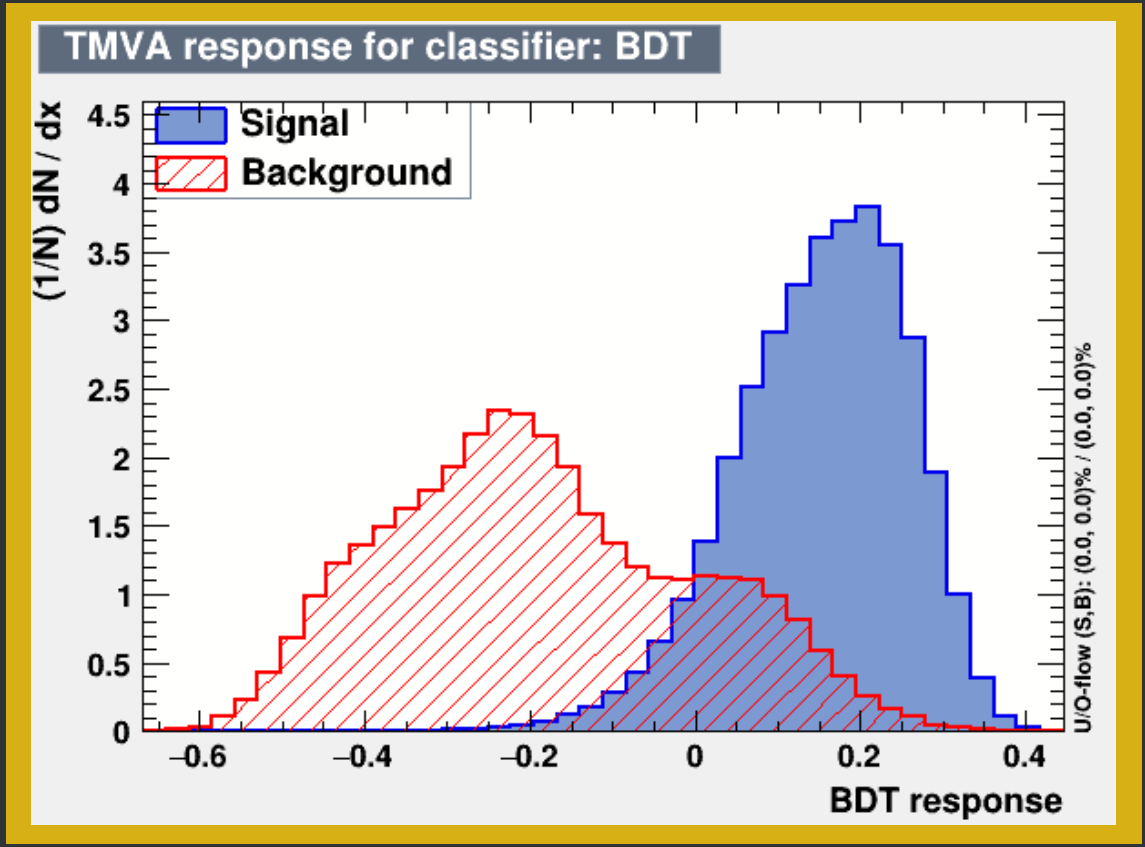

BDT

BDT Background Rejection vs Signal Efficiency more ideal than Fisher

# TMVA 4: Methods



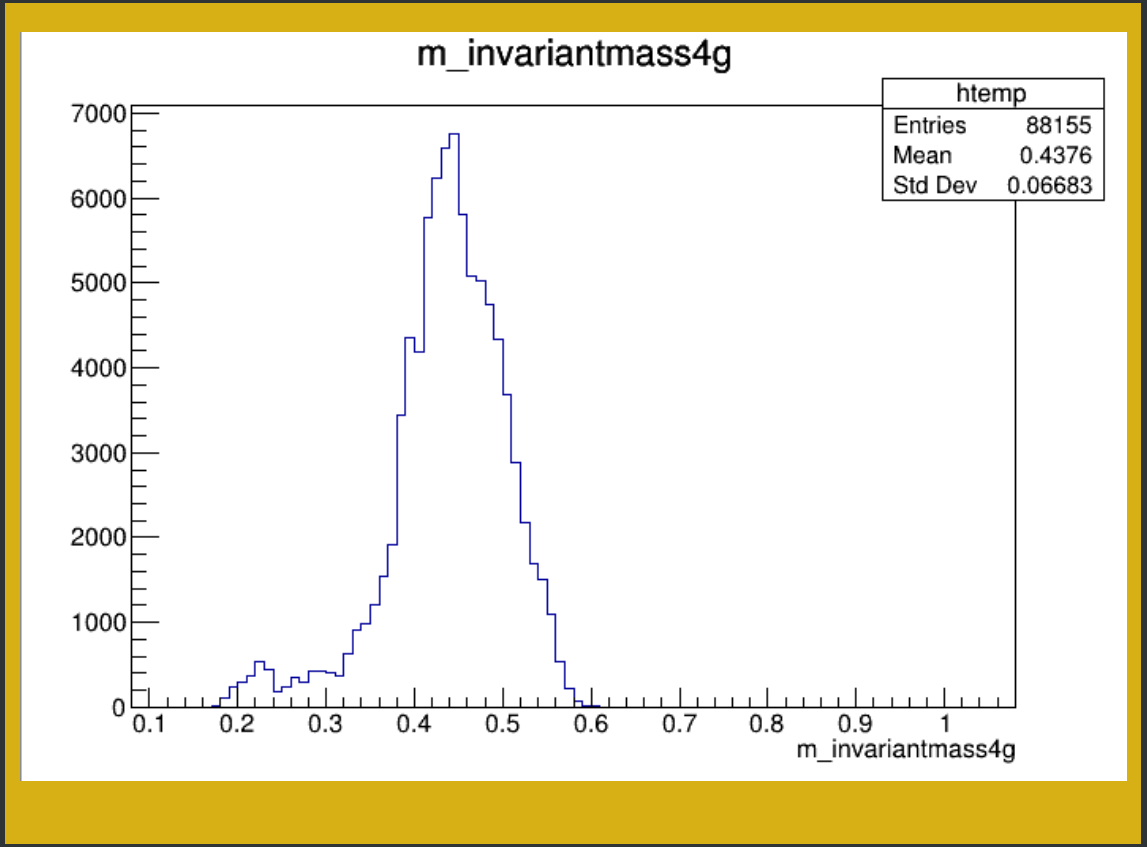

Fisher

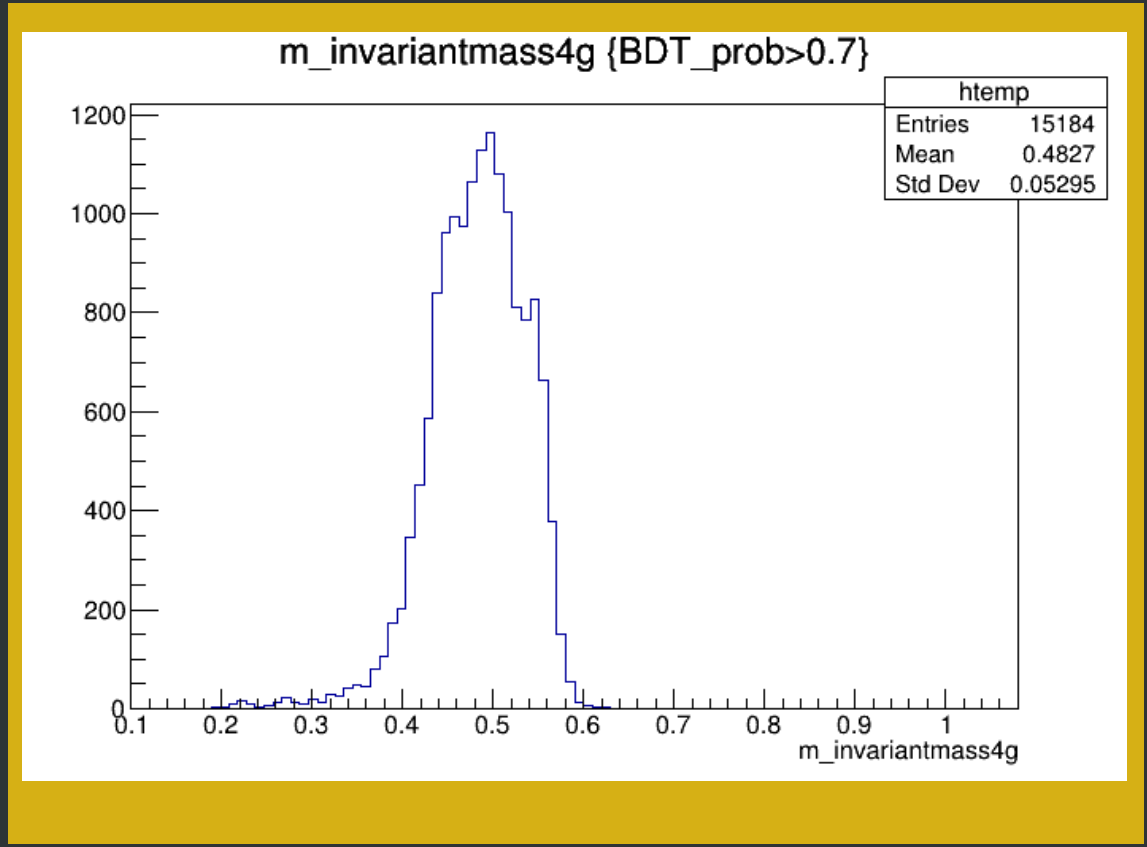



BDT

BDT separates signal from Background more efficiently

# TMVA 4: Training example




Raw

BDT w/ Default Parameters

Training removes most of the background at 0.2-0.3

# Conclusion

- Boosted Decision Tree method shows better results
- We can further use BDT parameter cuts
- With 15+ variables we are testing different combinations of input variables