Cutting with

AEACuS,

Plotting with

RHADAMANTHUS,

& Learning with

MINOS:

ReInterpretation Applications

Joel W. Walker
Sam Houston State University

(Re)interpretation of LHC results for new physics
CERN
December 12-15, 2022

Outline of Presentation

- Introduction to Package Philosophy and Operation
- Introduction of Case Study: ATLAS 3rd Gen Vectorlike Leptons
- Structure of AEACuS Card for Analysis Replication
- Structure of RHADAManTHUS Card for Plot Generation
- Structure of MINoS Card for BDT Machine Learning
- ReInterpretation of Results as applied to a SUSY Model

Unified Work Flow

- **MadGraph (+ Others):** Matrix Element Generation
- **MadEvent (+ Others):** Hard Scattering Simulation
- **Pythia (+ Others):** Showering and Hadronization
- **DELPHES:** Detector Simulation
  (DEtector Level PHysics Emulation System)
- **AEACUS:** Statistics Computation & Cut Selection
- **RHADAMANTHUS:** Graphical Event Analysis
- **MINOS:** Machine Learning on Features

Package Notes

- AEACUS, RHADAMANTHUS, & MINOS are Perl & Python
- All Perl scripts are self contained - no libraries or installation
- RHADAMANTHUS calls MatPlotLib & Numpy
- MINOS calls XGBoost, MPL & Numpy
- Control is provided by simple reusable card files
- Code Here: `https://github.com/joelwwalker/AEACuS`
- Quick Start + Video Here: `https://pos.sissa.it/409/027/pdf`

AEACuS

(Algorithm Event Arbiter and Cut Selector)

Guiding Principles:

- **It is important to separate WHAT from HOW**
- **It is important to document UNAMBIGOUSLY**
- **It is important to streamline REPRODUCTION**

AEACuS (Goals)

- Automate model recast comparison against LHC data
- Facilitate most current search strategies for new physics
- Embody lightweight, consumer-level, standalone design
- Decouple specific usage from general functionality
- Render event cut strategies compactly & unambiguously
- Merge power & flexibility with uniformity & simplicity
- Decouple phenomenology from software maintenance

AEACuS (Function)

- Converts Delphes Root file to a lightweight extended LHCO format
- ANY information in the Delphes Root File (including substructure variables and fat-jet information) can be read and processed by AEACuS
- Event-by-event weights are read and handled consistently throughout
- Filters kinematics, geometry, isolation & overlap, charge & flavor
- Dilepton pair assembly (by like/unlike charge & flavor)
- Jet (Re)clustering (KT, C/A, Anti-KT) & Hemispheres (Lund, etc.), along with common substructure variables and extraction of jet constituents
- Missing E_T, scalar H_T, effective & invariant mass, ratios & products
- Transverse mass, 1- & 2-step asymmetric M_{T2} (with combinatorics), Tri-jet mass, α_T, Razor & α_R, Dilepton Z-balance, Lepton W-projection, $\Delta\phi$ (& biased $\Delta\phi^*$), Shape Variables (thrust & minor, spheri[o]city, F), Girth, + MORE
- Arbitrary user-described combinations of observables plus external function calls

Language Vs. Framework

AEACuS is BOTH and it is FACTORIZABLE

- The AEACuS meta language is an ideal mechanism for large experiments (CMS/ATLAS) & small phenomenology groups to unambiguously propagate an approximate rendering of internal event selection strategies
- The AEACuS software tool is an ideal agent for the rapid and uniform projection of sophisticated event cut workflows onto new physics models

LHC Case Study Example

ATLAS CONF Note

ATLAS-CONF-2022-044

July 15, 2022

Search for Third-Generation Vectorlike Leptons in pp Collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector

The ATLAS Collaboration

A search for vectorlike leptons coupling to the third generation Standard Model leptons in the multilepton (two, three, and four or more light leptons) final state with zero or more hadronic tau lepton decays is presented. The search was performed using a dataset of proton-proton collisions corresponding to an integrated luminosity of 139 fb^{-1} at a center-of-mass energy of 13 TeV recorded by the ATLAS detector at the LHC. To maximize the separation of signal from background events, a machine learning classifier was used. No excess of events was observed beyond the standard model expectation. Using a doublet vectorlike lepton model, vectorlike leptons coupling to third generation Standard Model leptons are excluded in the mass range 130 GeV to 900 GeV at the 95% confidence level, with an expected upper mass limit of 970 GeV.

This study has several features which make it a good example for demonstrating the software tools:

- Large number of observables
- Orthogonal final state channels
- S/B separation using a BDT

Signal Region Selections

Variables	BDT Training Regions						
BDT	2ℓ SSSF, 1τ	2ℓ SSOF, 1τ	2ℓ OSSF, 1τ	2ℓ OSOF, 1τ	2ℓ, $\geq 2\tau$	3ℓ, $\geq 1\tau$	4ℓ, $\geq 0\tau$
N_ℓ	2	2	2	2	2	3	≥ 4
Charge/Flavor	SSSF	SSOF	OSSF	OSOF	-	-	-
N_τ	1	1	1	1	≥ 2	≥ 1	≥ 0
N_{jet}	> 0						
$E_{\text{T}}^{\text{miss}}$ [GeV]	≥ 120	≥ 90	≥ 60	≥ 100	≥ 60	≥ 90	≥ 60

Variables	Signal Regions						
BDT	2ℓ SSSF, 1τ	2ℓ SSOF, 1τ	2ℓ OSSF, 1τ	2ℓ OSOF, 1τ	2ℓ, $\geq 2\tau$	3ℓ, $\geq 1\tau$	4ℓ, $\geq 0\tau$
BDT Score	≥ 0.15	≥ 0.1	≥ 0.1	≥ 0.1	≥ -0.11	≥ 0.08	≥ 0.08

Catalog of Observables

Variable	Description
$E_{\mathrm{T}}^{\mathrm{miss}}$	The missing transverse momentum in the event
$\mathbb{S}(E_T^{\mathrm{miss}})$	The missing transverse momentum significance in the event
L_T	The scalar sum of light leptons p_T in the event
$L_T + E_T^{\mathrm{miss}}$	The scalar sum of light leptons p_T and the missing transverse momentum in the event
$L_T + p_T(\tau)$	The scalar sum of light leptons p_T and sum of taus p_T in the event
$p_T(l_1)$	The leading light lepton p_T in the event
$p_T(l_2)$	The sub-leading light lepton p_T in the event
$p_T(j_1)$	The leading jet p_T in the event
$p_T(\tau_1)$	The leading τ p_T in the event
N_j	The number of jets in the event
N_b	The number of b-jets in the event
H_T	The scalar sum of jet p_T in the event
$L_T + H_T$	The scalar sum of light leptons p_T and sum of jets p_T in the event
M_{ll}	The invariant mass of all light leptons in the event
$M_{l\tau}$	The invariant mass of all light leptons and taus in the event
M_{lj}	The invariant mass of all light leptons and jets in the event
M_{jj}	The invariant mass of all jets in the event
$M_{j\tau}$	The invariant mass of all jets and taus in the event
M_T	The transverse mass of the leading light lepton in the event
M_{OSSF}	The invariant mass the opposite sign same flavor pair of light leptons closest to the Z mass in the event
$\Delta\phi(j_1 E_T^{\mathrm{miss}})$	$\Delta\phi$ between E_T^{miss} and the leading p_T jet in the event
$\Delta\phi(l_1 E_T^{\mathrm{miss}})$	$\Delta\phi$ between E_T^{miss} and the leading p_T light lepton in the event
$\Delta\phi(l_1 l_2)$	$\Delta\phi$ between the leading and sub-leading p_T light lepton in the event
$\Delta\phi(l_1 j_1)$	$\Delta\phi$ between the leading p_T light lepton and jet in the event
$\Delta\phi(\tau_1 E_T^{\mathrm{miss}})$	$\Delta\phi$ between E_T^{miss} and the leading p_T τ in the event
$\Delta\phi(l_1 \tau_1)$	$\Delta\phi$ between the leading p_T light lepton and τ in the event
$\Delta\phi(j_1 \tau_1)$	$\Delta\phi$ between the leading p_T jet and τ in the event
$\Delta R(j_1 E_T^{\mathrm{miss}})$	ΔR between E_T^{miss} and the leading p_T jet in the event
$\Delta R(l_1 E_T^{\mathrm{miss}})$	ΔR between E_T^{miss} and the leading p_T light lepton in the event
$\Delta R(l_1 l_2)$	ΔR between the leading and sub-leading p_T light lepton in the event
$\Delta R(l_1 j_1)$	ΔR between the leading p_T light lepton and jet in the event
$\Delta R(\tau_1 E_T^{\mathrm{miss}})$	ΔR between E_T^{miss} and the leading p_T τ in the event
$\Delta R(l_1 \tau_1)$	ΔR between the leading p_T light lepton and τ in the event
$\Delta R(j_1 \tau_1)$	ΔR between the leading p_T jet and τ in the event

Card: Object Reconstruction

```
# ATLAS-CONF-2022-044
# Third Generation Vectorlike Leptons
# sqrt(s) = 13 TeV @ 169 / fb

CHN_000 = LHC:"../../LHCO", OUT:"../../Cuts"
# Channel Zero is for global settings, e.g paths

ELE_000 = PRM:[0,2.47], PTM:30
MUO_000 = PRM:[0,2.50], PTM:30
TAU_000 = PRM:[0,2.47], PTM:20
# Zeroth order filters on pseudo-rapidity
# and transverse momentum magnitudes
# Passing objects collected on shelf LEP_000

JET_001 = SRC:+000, PRM:[0,2.50], PTM:20,
      CUT:1, OUT:[PTM_004,ETA_004]
# Define jet acceptance and require
# at least one matching object
# Output leading kinematics

JET_002 = SRC:+001, HFT:1, CUT:0
# Count jets with heavy flavor tagging
JET_003 = SRC:+001, SET:[LED,PTM,-1]
# Subset classification with the leading jet candidate,
# i.e. the last member of a momentum sort
```

Card: Object Hierarchies

```
LEP_001 = SRC:+000, EMT:-2, PRM:[1.37,1.52]
# Group non-muon leptons in the rapidity blind spot
LEP_002 = SRC:[+000,-001]
# Group object in zeroth set but not the blind spot
LEP_003 = SRC:+002, EMT:-3, CUT:2, OUT:[PTM_001-002,ETA_001-002]
# Assemble a subset with light leptons (non-taus),
# requiring at least 2 members, and outputting kinematics
LEP_004 = SRC:+002, EMT:+3, CUT:0, OUT:[PTM_003,ETA_003]
# Assemble a subset with hadronic taus
LEP_005 = SRC:+003, SET:[LED,PTM,-1]
# Collect the single leading light lepton
LEP_006 = SRC:+003, SET:[LED,PTM,-2]
# Collect the pair of leading light leptons
LEP_007 = SRC:+004, SET:[LED,PTM,-1]
# Collect the single leading hadronic tau
LEP_008 = SRC:[+005,+007]
# Collect the leading light lepton / tau pair

LEP_011 = SRC:+003, EFF:[DIL,+1,+1], CUT:0
# Count SS/SF light dilepton parings
LEP_012 = SRC:+003, EFF:[DIL,+1,-1], CUT:0
# Count SS/OF light dilepton parings
LEP_013 = SRC:+003, EFF:[DIL,-1,+1,91.2], CUT:0, OUT:MAS_001
# Count OS/SF light dilepton parings and
# output reconstructed mass closest to the Z
LEP_014 = SRC:+003, EFF:[DIL,-1,-1], CUT:0
# Count OS/OF light dilepton parings
```

Card: Observables

```
IET_000 = CUT:60
# Require at least 60 GeV of invisible
# transverse energy as reported by Delphes
RHR_001 = NUM:000, DEN:000
# Compute the MET significance, i.e. ratio
# of default (full reconstructed event)
# MET and sqrt( HT = scalar sum PT )
MHT_001 = LEP:003
# Compute HT for light leptons
MEF_001 = MET:000, MHT:001
# Compute light lepton effective mass ( HT + MET )
MHT_002 = LEP:002
# Compute HT for light leptons plus taus
MHT_003 = JET:001
# Compute HT for jets
MHT_004 = LEP:003, JET:001
# Compute HT for light leptons and jets
OIM_001 = LEP:003
# Compute object invariant mass for light leptons
OIM_002 = LEP:002
# Compute invariant mass for light leptons and taus
OIM_003 = LEP:003, JET:001
# Compute invariant mass for light leptons and jets
OIM_004 = JET:001
# Compute invariant mass for jets
OIM_005 = LEP:004, JET:001
# Compute invariant mass for taus and jets
OTM_001 = MET:000, LEP:005
# Compute transverse mass for leading light lepton
```

```
MDP_001 = MET:000, LEP:005
# Compute Delta Phi of leading light lepton to MET
MDP_002 = MET:000, LEP:007
# Compute Delta Phi of leading tau to MET
MDP_003 = MET:000, JET:003
# Compute Delta Phi of leading jet to MET
ODP_001 = LEP:006
# Compute Delta Phi between leading light lepton pair
ODP_002 = LEP:005, JET:003
# Compute Delta Phi between leading light lepton and jet
ODP_003 = LEP:008
# Compute Delta Phi between leading light lepton and tau
ODP_004 = LEP:007, JET:003
# Compute Delta Phi between leading tau and jet

VAR_001 = VAL:{NRM($1,$2),MDP_001,ETA_001}
# Compute Delta R of leading light lepton to MET
# Use custom variable to get norm of { Eta, Delta Phi }
VAR_002 = VAL:{NRM($1,$2),MDP_002,ETA_003}
# Compute Delta R of leading tau to MET
VAR_003 = VAL:{NRM($1,$2),MDP_003,ETA_004}
# Compute Delta R of leading jet to MET
ODR_001 = LEP:006
# Compute Delta R between leading light lepton pair
ODR_002 = LEP:005, JET:003
# Compute Delta R between leading light lepton and jet
ODR_003 = LEP:008
# Compute Delta R between leading light lepton and tau
ODR_004 = LEP:007, JET:003
# Compute Delta Phi between leading tau and jet
```

Card: Signal Regions

```
# Register event selection cuts to
# be applied in channel sorting
ESC_001 = KEY:LEP_003, CUT:[2,2]
# Require exactly 2 light leptons
ESC_002 = KEY:LEP_003, CUT:[3,3]
# Require exactly 3 light leptons

ESC_011 = KEY:LEP_004, CUT:[0,0]
# Require exactly 0 hadronic taus
ESC_012 = KEY:LEP_004, CUT:[1,1]
# Require exactly 1 hadronic tau

ESC_021 = KEY:LEP_011, CUT:1
# Require SS/SF light dilepton pair
ESC_022 = KEY:LEP_012, CUT:1
# Require SS/OF light dilepton pair
ESC_023 = KEY:LEP_013, CUT:1
# Require OS/SF light dilepton pair
ESC_024 = KEY:LEP_014, CUT:1
# Require OS/OF light dilepton pair

ESC_031 = KEY:IET_000, CUT:90
ESC_032 = KEY:IET_000, CUT:100
ESC_033 = KEY:IET_000, CUT:120
# Elevate invisible transverse energy cut
# to 90, 100, or 120 GeV, respectively
```

```
# Channels subscribe to event selections
CHN_001 = ESC:[+001,+012,+021,+033]
# 2 leptons (SS/SF), 1 tau, 120 GeV MET
CHN_002 = ESC:[+001,+012,+022,+031]
# 2 leptons (SS/OF), 1 tau, 90 GeV MET
CHN_003 = ESC:[+001,+012,+023]
# 2 leptons (OS/SF), 1 tau, baseline MET
CHN_004 = ESC:[+001,+012,+024,+032]
# 2 leptons (OS/OF), 1 tau, 100 GeV MET
CHN_005 = ESC:[+001,-011,-012]
# 2 leptons and 2+ (not 0 or 1) taus
CHN_006 = ESC:[+002,-011,+031]
# 3 leptons, 1+ (not 0) taus, 90 GeV MET
CHN_007 = ESC:[-001,-002]
# 4+ (not 2 or 3) light leptons
```

BG Simulation

- We simulate MC data with MadGraph/MadEvent, Pythia8, Delphes
- Our purpose is demonstration of the tools … Several Caveats:
- We generate just 3 leading SM backgrounds: TTbar, WZ, ZZ
- BGs are +1 inclusive jet, LO only, without k-factor correction
- Fake rates are substantially under-modeled here

Signal ReInterpretation

- We select a SUSY signal model matching targeted leptonic final states
- The model is picked for tool demonstration (not necessarily physics)
- Initial state is stau / tau-sneutrino pair production (500 GeV)
- Decays cascade through chargino-1 / neutralino-2 (400 GeV)
- … as well as sleptons / lepton sneutrinos (200 GeV)
- … and terminate with neutralino-1 (100 GeV)
- Cross sections are rescaled by (10^{-2} — 10^{-5}) to hit edge of visibility

AEACuS Output

```
NNN KEY_NNN EPW_NNN ENW_NNN EZW_NNN       XPB_NNN       ABS_NNN       ERR_NNN       IPB_NNN       EFF_NNN       PRD_NNN
000 CHN_000  124459       0       0 +6.40526E+02 +6.40526E+02 +1.60155E-01 +1.94307E+02 1.0000000000 1.0000000000
001 TAU_000  106652       0       0 +5.48883E+02 +5.48883E+02 +1.48256E-01 +1.94307E+02 0.8569247704 0.8569247704
002 OBJ_001     355       0       0 +1.82700E+00 +1.82700E+00 +8.55348E-03 +1.94307E+02 0.0033285827 0.0028523449
003 JET_002     316       0       0 +1.62629E+00 +1.62629E+00 +8.06998E-03 +1.94307E+02 0.8901408451 0.0025389887
004 JET_006      54       0       0 +2.77910E-01 +2.77910E-01 +3.33600E-03 +1.94307E+02 0.1708860759 0.0004338778
005 MET_000      48       0       0 +2.47031E-01 +2.47031E-01 +3.14521E-03 +1.94307E+02 0.8888888889 0.0003856692
006 ESC_001      24       0       0 +1.23504E-01 +1.23504E-01 +2.22389E-03 +1.94326E+02 0.4999534471 0.0001928165

OBJ_001 LEP_001 JET_001
LEP_001 0.00000 0.00000
JET_001 0.00000 0.00000

EID_000 TAU_000 LEP_001 JET_001 JET_002 JET_006 ETA_001 PTM_001 MAS_001 ETA_002 PTM_002 ETA_003 PTM_003 MAS_003 ETA_004 PTM_004 CAL_000 MET_000 MHT_000
MEF_000 MT2_001 TTM_001 ODP_001 MDP_001 CTS_001 MT2_002 ODP_002 MDP_002 ODP_003 MDP_003 ODP_004 MDP_004 ODP_005 ODP_006     VAR_001     VAR_002     VAR_003
VAR_004     VAR_005     VAR_006     VAR_011     VAR_012     VAR_013     VAR_014     VAR_021     WGT_000
0000898       0       2       1       0       0  -0.029    28.3    46.5  -1.027    17.5   0.631    83.7     8.3   UNDEF   UNDEF    62.0    81.9   154.1
236.0   108.1   -75.4   2.261   0.112   0.461     5.1   2.971   2.373   UNDEF   3.083   0.710   UNDEF   UNDEF   UNDEF +7.608E-01 +5.784E-01 +2.899E-02
+9.299E-01 +7.727E-01 +5.587E-01 +3.459E-01 +2.140E-01 +1.023E+00 +0.000E+00 +1.594E+00 +5.146E-03
0003430       0       2       1       0       0  -1.221   100.8   206.5  -0.308    90.4  -0.089    53.4     5.8   UNDEF   UNDEF    96.2   120.6   269.0
389.7   133.2   242.0   2.676   1.948   0.427    72.5   1.083   1.659   UNDEF   3.031   1.593   UNDEF   UNDEF   UNDEF +7.226E-01 +8.117E-01 +8.399E-01
+2.156E-01 +2.986E-01 +8.877E-02 +8.354E-01 +7.496E-01 +4.428E-01 +0.000E+00 +4.587E-01 +5.146E-03
0004304       0       2       1       0       0  -0.521   108.2   190.9   0.475    73.8   1.873    41.5     6.1   UNDEF   UNDEF    18.7    31.1   223.5
254.7   131.7   109.3   2.409   2.400   0.461    41.9   2.393   1.475   UNDEF   0.007   1.481   UNDEF   UNDEF   UNDEF +7.599E-01 +9.835E-01 +4.785E-01
+8.849E-01 +4.422E-01 +9.539E-01 +3.475E+00 +2.368E+00 +1.334E+00 +0.000E+00 +7.566E-01 +5.146E-03
0027442       0       2       1       0       0  -2.246    55.2    29.9  -1.594    28.1  -1.994    41.0     7.7   UNDEF   UNDEF    82.7    88.9   148.7
237.6   144.5   -83.1   0.370   2.237   0.315    76.2   1.940   2.607   UNDEF   2.106   1.570   UNDEF   UNDEF   UNDEF +5.730E-01 +2.468E-01 +9.779E-01
+3.799E-01 +9.208E-01 +9.636E-01 +6.214E-01 +3.157E-01 +4.615E-01 +0.000E+00 +5.839E-01 +5.146E-03
0027631       0       2       1       0       0   0.596    97.8   105.7  -0.860    35.8  -0.858    46.3     8.2   UNDEF   UNDEF    47.3    61.8   312.7
374.5   153.3  -100.9   0.844   2.019   0.622    74.1   1.206   2.863   UNDEF   0.813   2.050   UNDEF   UNDEF   UNDEF +8.969E-01 +8.965E-01 +5.342E-01
+2.000E-03 +6.963E-01 +6.952E-01 +1.583E+00 +5.794E-01 +7.500E-01 +0.000E+00 +7.193E-01 +5.146E-03
0030072       0       2       1       0       0   1.802    31.4    55.9   1.888    24.9   1.267    40.0     7.3   UNDEF   UNDEF    59.3    64.4   119.9
184.3   101.6   276.5   3.008   2.502   0.043     0.4   0.550   0.506   UNDEF   3.052   2.725   UNDEF   UNDEF   UNDEF +8.579E-02 +4.892E-01 +9.470E-01
+5.518E-01 +9.552E-01 +8.530E-01 +4.880E-01 +3.873E-01 +6.215E-01 +0.000E+00 +4.038E+00 +5.146E-03
```

- Running on a MC sample => tables reporting requested statistics & cut fractions
- It is often convenient to make limited cuts at the lowest level, and just compute
- Names such as "JET_001" have no invariant meaning - they are defined in a card_file

RHADAMANTHUS

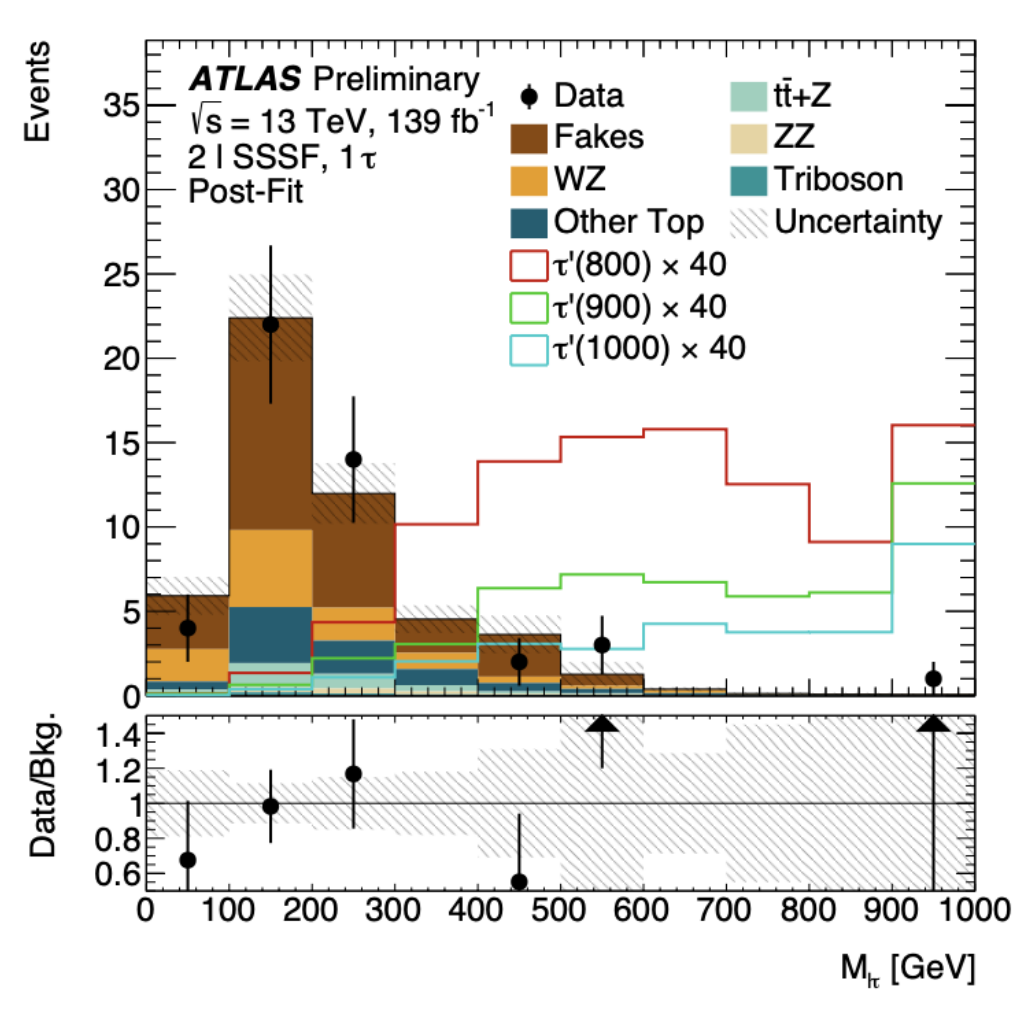
(Recursively Heuristic Analysis, Display, And MANipulation: The Histogram Utility Suite)

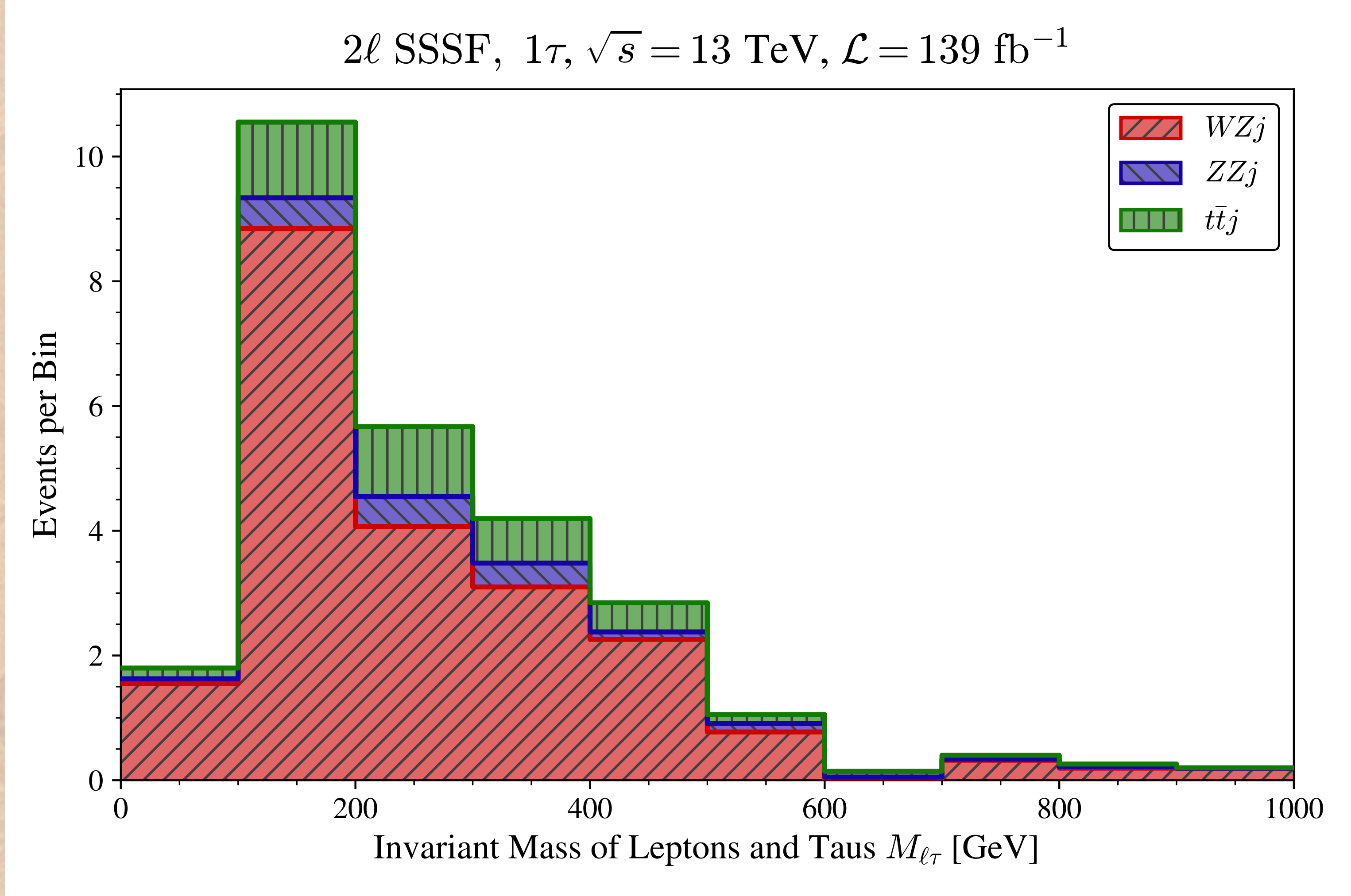
- RHADAManTHUS plots observables computed by AEACuS
- It correctly combines distinct / over-sampled MC by cross section
- It generates 1- and 2-dimensional plots with per-event weighting
- Event selections and functional transformations are made easy
- We will validate the analysis with a few example plots
- Expect variations from k-factors, statistics, and shortcuts, etc.

Plot Card Example

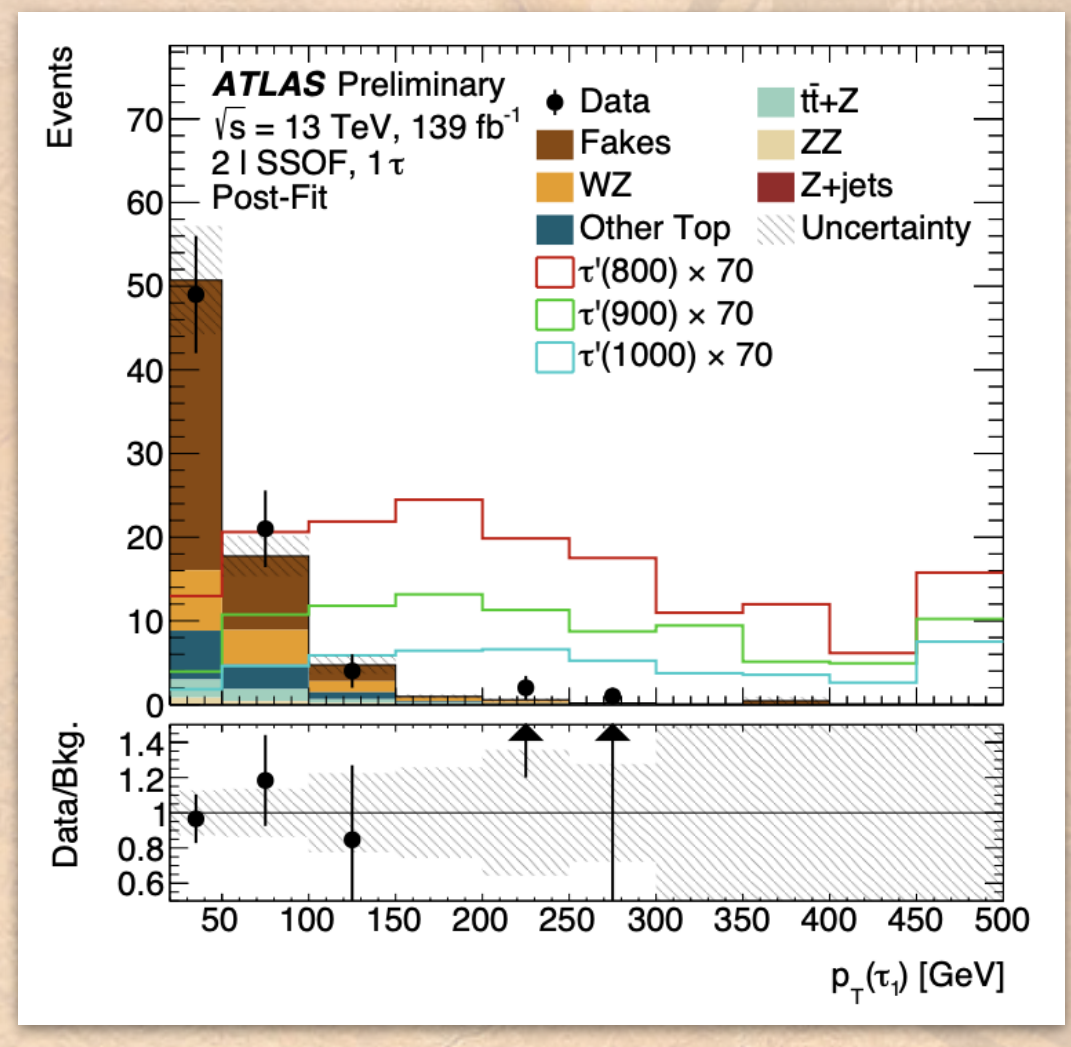
```
# Construct Data Sets From Files
DAT_101 = DIR:"./Cuts/CHN_001", FIL:"SMBG_W_Z_J_*"
DAT_102 = DIR:"./Cuts/CHN_001", FIL:"SMBG_Z_Z_J_*"
DAT_103 = DIR:"./Cuts/CHN_001", FIL:"SMBG_TTBar_Z_J_*"

# Construct Channels from Data Sets
CHN_001 = DAT:[101,102,103], KEY:OIM_002

# Construct Histograms from Channels
HST_001 = CHN:001, IFB:139, NAM:"CHN_001", FMT:"PDF",
      STK:+1, LFT:100, RGT:1000, SPN:100, MIN:0,
      LGD:[ "$WZj$", "$ZZj$" , "$t\bar{t}j$" ],
      TTL:"$2\ell~{\rm SSSF},~1 \tau$,
            $\sqrt{s} = 13~$TeV,
            $\mathcal{L} = 139~{\rm fb}^{-1}$",
      LBL:["Invariant Mass of Leptons and Taus
            $M_{\ell\tau}$ [GeV]", "Events per Bin"]
```

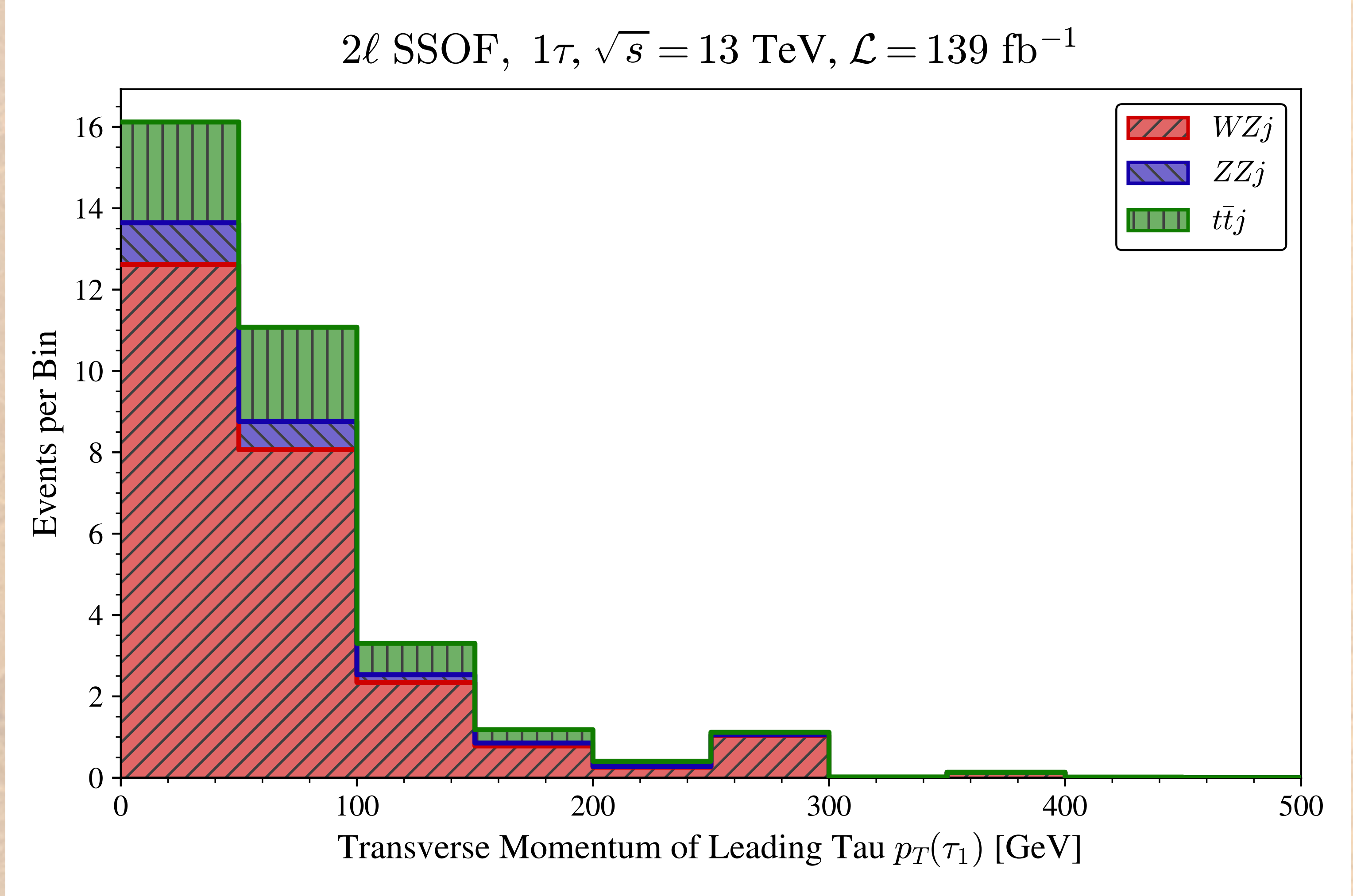
Sample Plot Comparison





Another Plot Comparison





MINOS

(Machine Intelligent Optimization of Significance)

- MInOS automates BDT Machine Learning in a Collider Context
- It reads event features computed by AEACuS
- It correctly combines distinct / over-sampled MC by cross section
- It trains for optimal Signal/BG discrimination (XGBoost backend)
- It generates density, significance, feature importance, & ROC plots (MatPlotLib backend) from validation data (1/3 by default)
- It lets Pheno Projects skip overhead & get answers QUICKLY
- We will reinterpret the ATLAS analysis with an example training

Why BDTs for Physics?

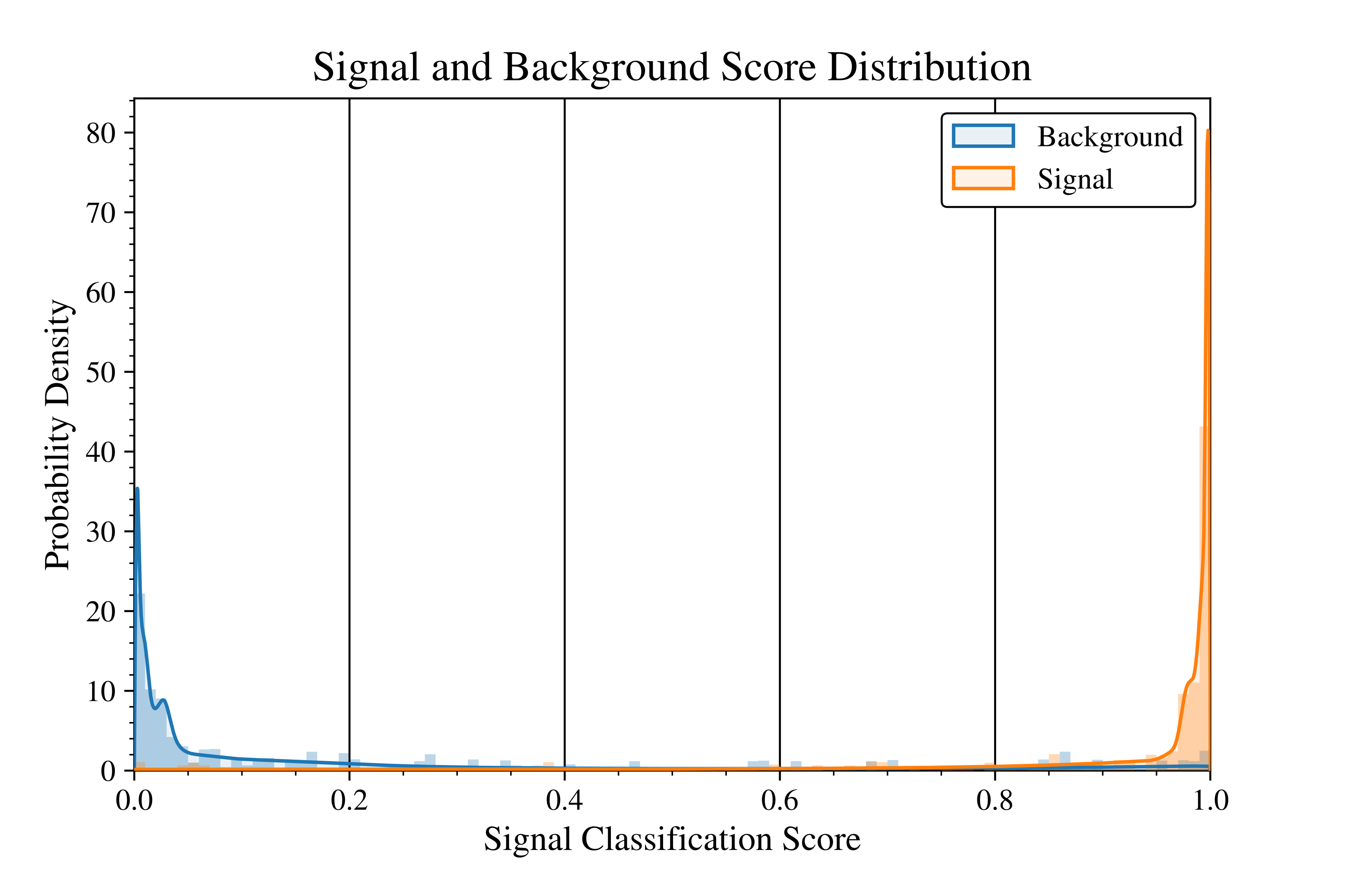
- Binary classification problems (Signal vs. Background) are common
- We want to maximize discrimination power
- We want to eliminate bias and work efficiently
- We want to incorporate domain knowledge & expertise
- We want to understand what the machine learning learned

BDTs balance POWER with TRANSPARENCY

MInOS Card Example

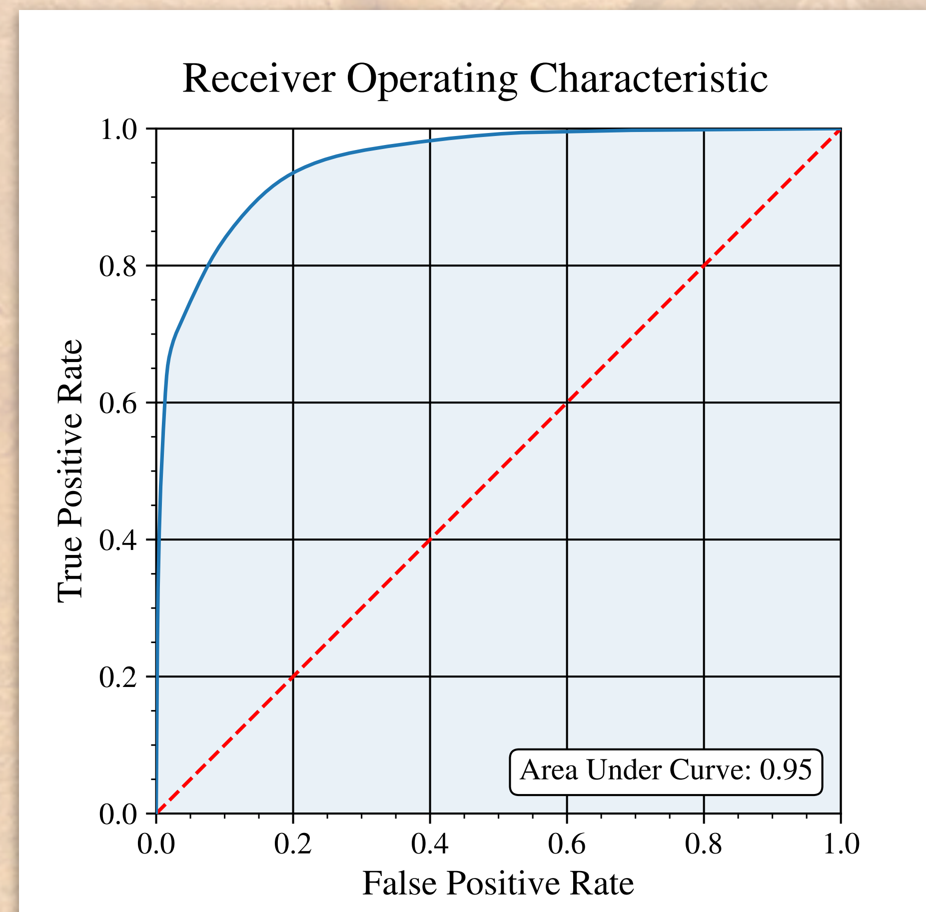
```
# Construct Data Sets From Files
DAT_201 = DIR:"./Cuts/CHN_002", FIL:[
      "SMBG_W_Z_J_*", "SMBG_Z_Z_J_*", "SMBG_TTBar_Z_J_*" ]
DAT_211 = DIR:"./Cuts/CHN_002", FIL:"SUSY_SNUT_SNUT_rif_500_*"
DAT_212 = DIR:"./Cuts/CHN_002", FIL:"SUSY_SNUT_STAU_rif_500_*"
DAT_213 = DIR:"./Cuts/CHN_002", FIL:"SUSY_SNUT_STAU_LL_rif_500_*"
DAT_214 = DIR:"./Cuts/CHN_002", FIL:"SUSY_STAU_STAU_LL_rif_500_*"

# Training 000 is for defaults
TRN_000 = IFB:139, INC:[
      # Specify training features for inclusion
      IET_000, RHR_001, MHT_001, MEF_001, MHT_002, PTM_001,
      PTM_002, PTM_003, PTM_004, JET_001, JET_002, MHT_003,
      MHT_004, OIM_001, OIM_002, OIM_003, OIM_004, OIM_005,
      OTM_001, MDP_001, MDP_002, ODP_001, ODP_002, ODP_003,
      ODP_004, VAR_001, VAR_002, VAR_003, ODR_001, ODR_003,
      ODR_004 ], TEX: [
      # Specify LaTeX Labels
      IET_000, "$E_T^{\rm miss}$",
      RHR_001, "$\mathcal{S}(E_T^{\rm miss})$",
      MHT_001, "$L_T$",
      # ... Skipping for brevity
      ODR_004, "$\Delta R(j_1 \tau_1)$" ]

# Construct Channels from Data Sets
CHN_201 = DAT:201, LBL:0
CHN_211 = DAT:211, LBL:1, WGT:{$1/100000,WGT_000}
CHN_212 = DAT:212, LBL:1, WGT:{$1/100,WGT_000}
CHN_213 = DAT:213, LBL:1, WGT:{$1/10000,WGT_000}
CHN_214 = DAT:214, LBL:1, WGT:{$1/10000,WGT_000}

# Construct Trainings from Channels
TRN_201 = CHN:[201,211,212,213,214]
```

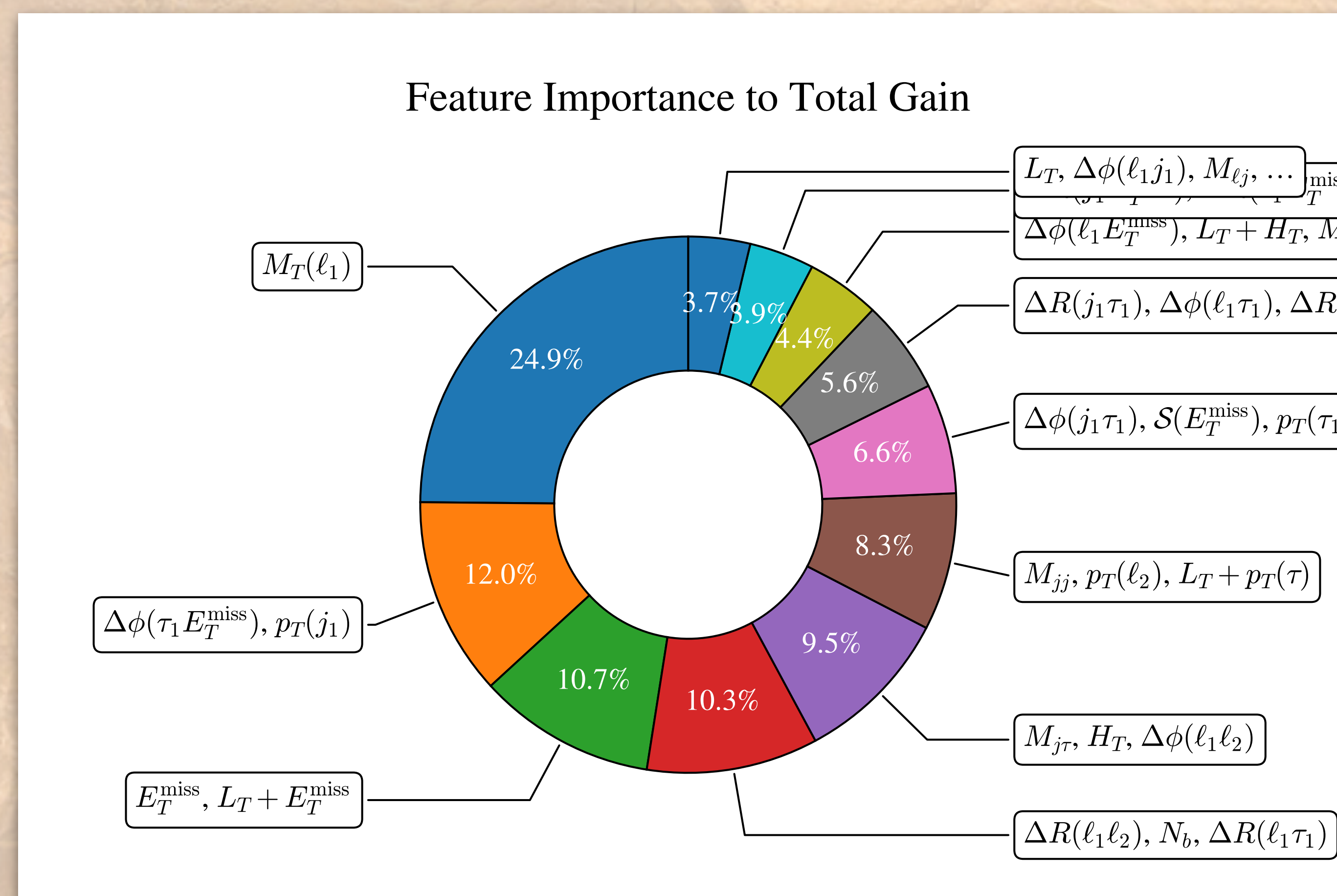
MInOS Output



- Signal & Background Probability Density Visualizes Separation

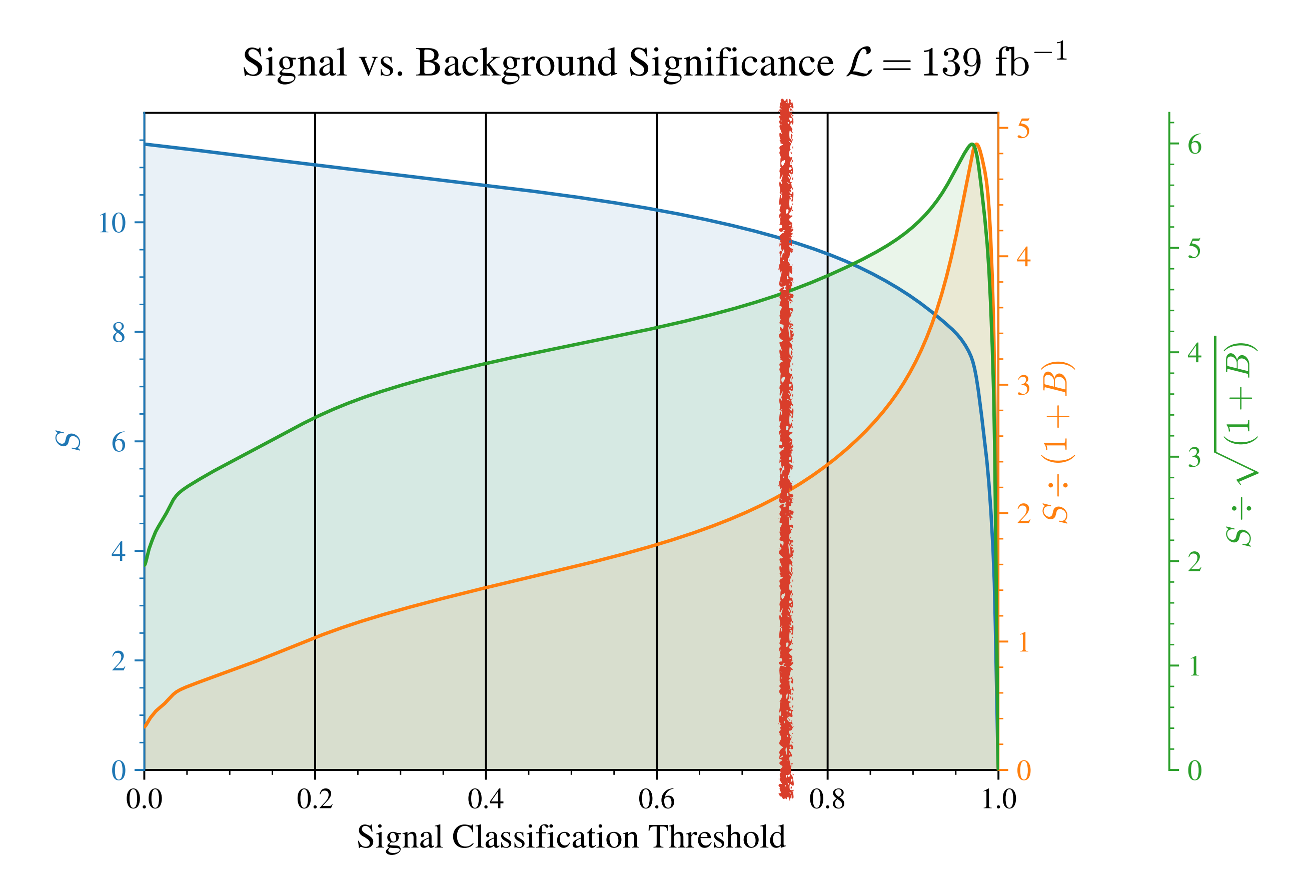
MInOS Output





- The ROC curve is a standard metric of S/B separability
- A feature importance chart clarifies what is going on inside the BDT

MInOS Output



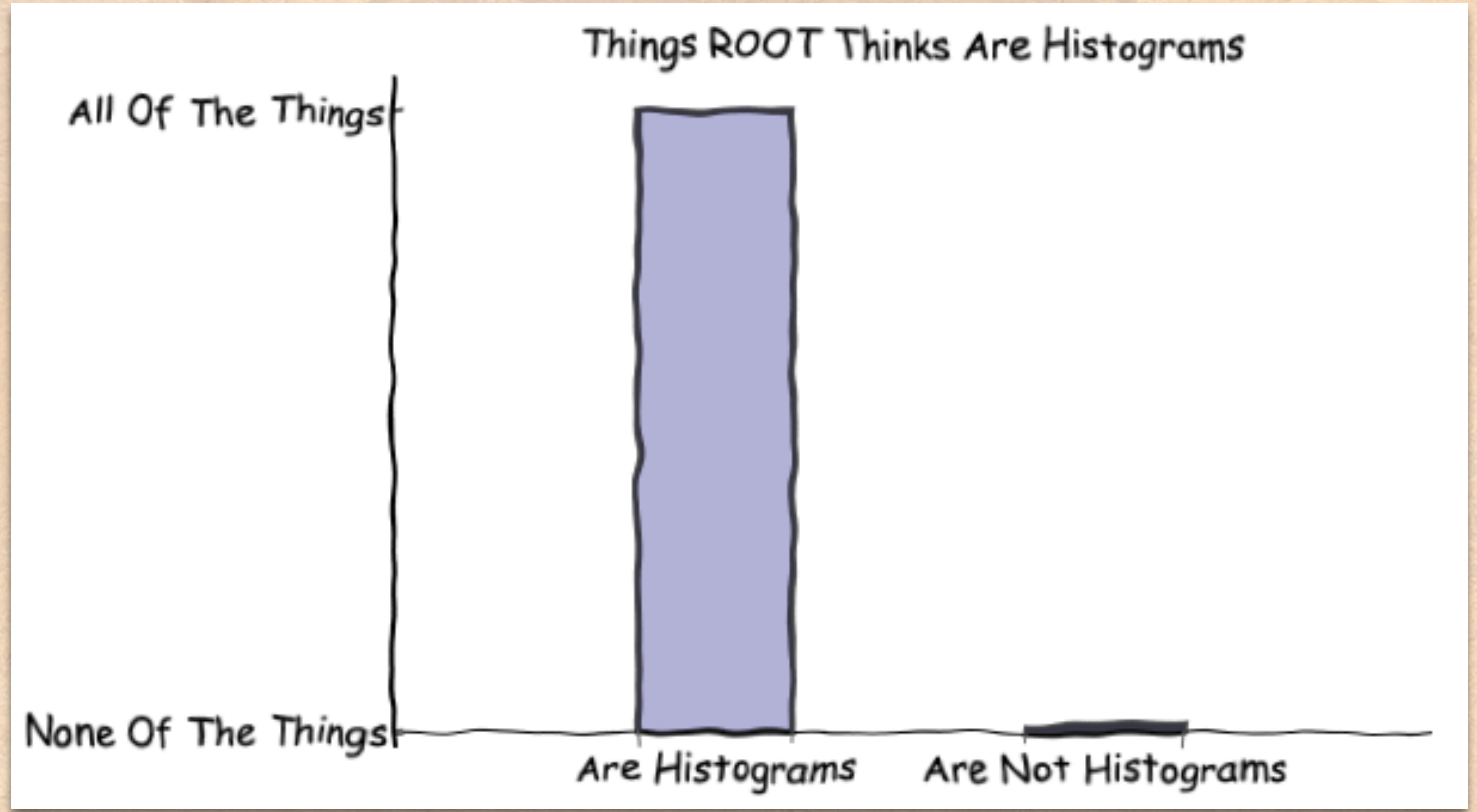
Signal Regions	$2\,\ell$ SSOF, $1\,\tau$
Observed Events	3
Total Background	3.70 ± 0.40
Other Top	0.85 ± 0.24
$t\bar{t}$+Z	0.28 ± 0.07
ZZ	0.13 ± 0.02
WZ	1.28 ± 0.19
Triboson	< 0.01
Fakes	1.18 ± 0.17

- Survival fraction of S, B as a function of the classification threshold are used to show achievable significance (at specified luminosity)
- At the working point (~ red line) of the BDT (B = 3.7) we have Signal ~ 9-10, so this model benchmark would be ruled out

AEACuS, RHADAMaNTHUS & MINoS

- The joint package is now ready to use, available at GitHub
- https://github.com/joelwwalker/AEACuS
- Please contact author directly: `jwalker@shsu.edu`
- If you are interested in teaming up, borrowing features, building an analysis library, or doing validations, please Let Me Know!

Thank You!



Things ROOT Thinks Are Histograms
All Of The Things
None Of The Things
Are Histograms
Are Not Histograms

"Then spake Zeus: … 'The cases are now indeed judged ill and it is because … many … who have wicked souls are clad in fair bodies and ancestry and wealth, and … the judges are confounded …, having their own soul muffled in the veil of eyes and ears and the whole body. … They must be stripped bare of all those things …, beholding with very soul the very soul of each immediately. … [I] have appointed sons of my own to be judges; two from Asia, **Minos** and **Rhadamanthus**, and one from Europe, **Aeacus**. These … shall give judgement in the meadow at the dividing of the road, whence are the two ways leading, one to the Isles of the Blest …, and the other to Tartaros.' "

– Plato, *Gorgias* (trans. Lamb)