WW and WZ Analysis Based on Boosted Decision Trees

Hai-Jun Yang
University of Michigan
(contributed from Tiesheng Dai, Alan Wilson, Zhengguo Zhao, Bing Zhou)

ATLAS Trigger and Physics Meeting
CERN, June 4-7, 2007

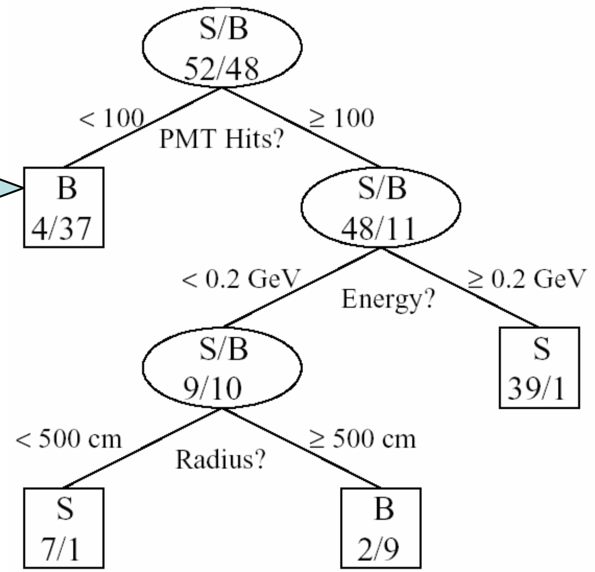
Outline

- Boosted Decision Trees (BDT)
- WW → $e\mu X$ analysis based on BDT
- WZ → $l\nu$ ll analysis based on BDT
- BDT Applications and Free Softwares
- Summary and Future Plan

Boosted Decision Trees

How to build a decision tree ?

For each node, try to find the best variable and splitting point which gives the best separation based on Gini index.
Gini_node = Weight_total*P*(1-P), P is weighted purity
Criterion = Gini_father – Gini_left_son – Gini_right_son
Variable is selected as splitter by maximizing the criterion.

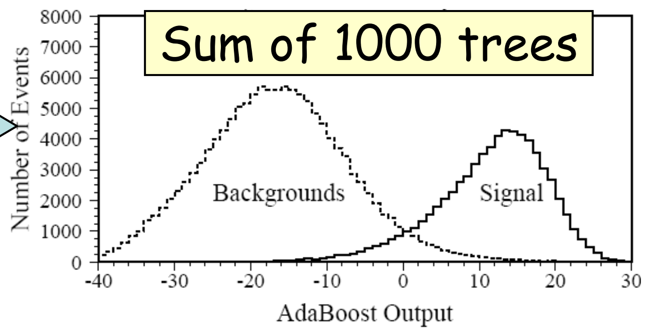


How to boost the decision trees?

Weights of misclassified events in current tree are increased, the next tree is built using the same events but with new weights. Typically, one may build few hundred to thousand trees.

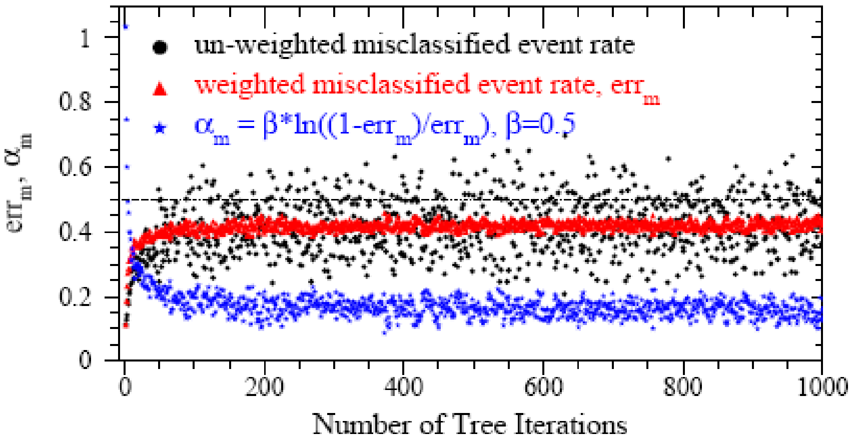
How to calculate the event score ?

For a given event, if it lands on the signal leaf in one tree, it is given a score of 1, otherwise, -1. The sum (probably weighted) of scores from all trees is the final score of the event.



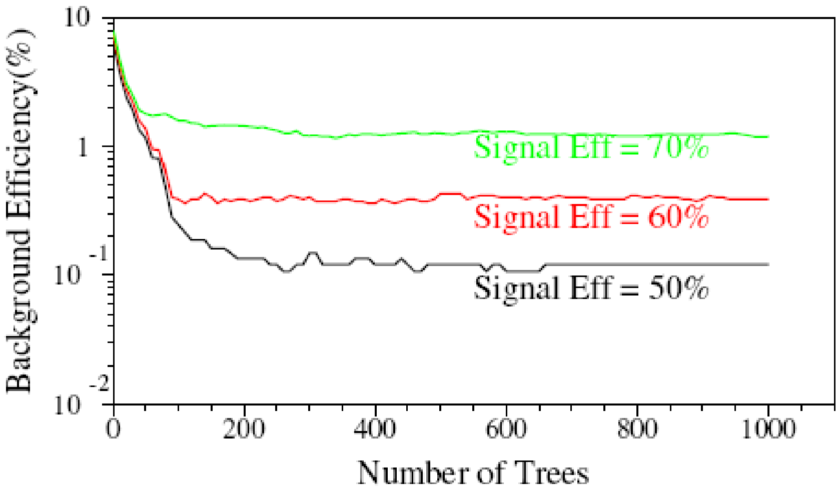
Ref: B.P. Roe, H.J. Yang, J. Zhu, Y. Liu, I. Stancu, G. McGregor, "*Boosted decision trees as an alternative to artificial neural networks for particle identification*", physics/0408124, NIM A543 (2005) 577-584.

Weak → Powerful Classifier



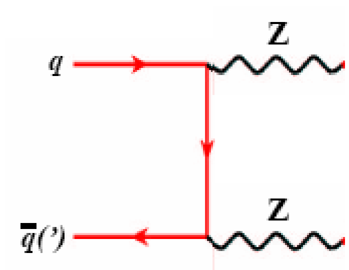
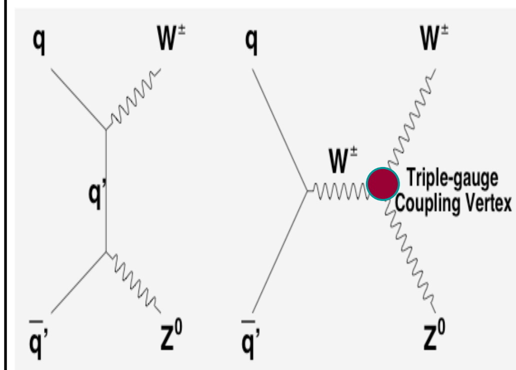
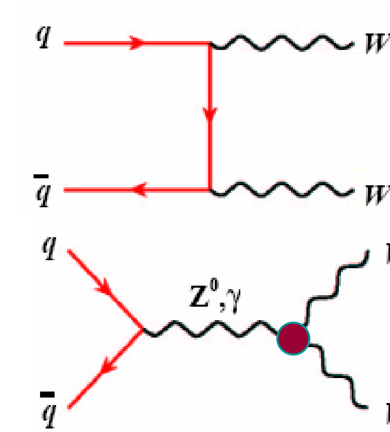
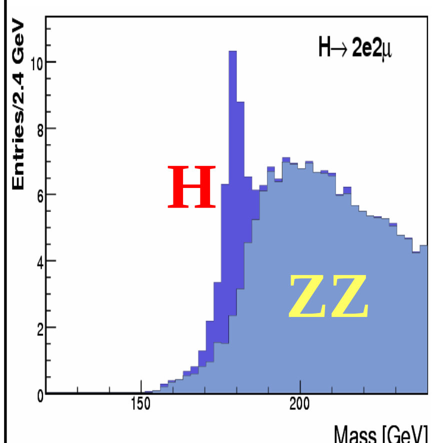
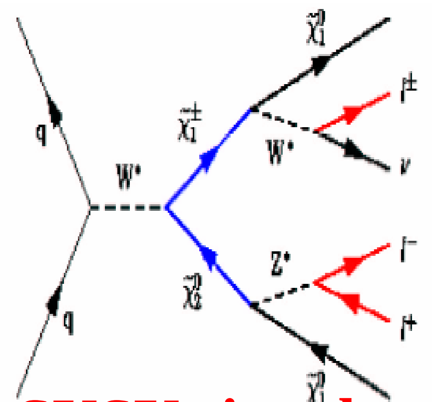
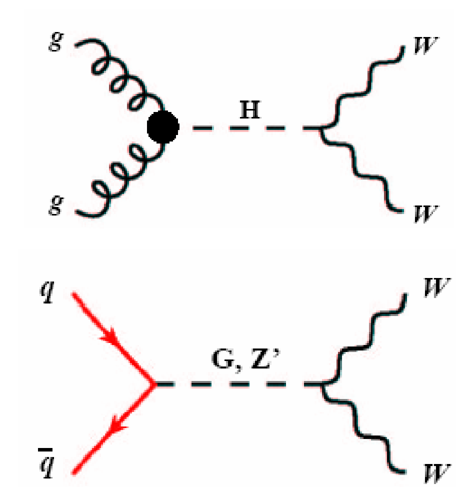
➔ Boosted decision trees focus on the misclassified events which usually have high weights after hundreds of tree iterations. An individual tree has a very weak discriminating power; the weighted misclassified event rate err_m is about 0.4-0.45.

➔The advantage of using boosted decision trees is that it combines many decision trees, "weak" classifiers, to make a powerful classifier. The performance of boosted decision trees is stable after a few hundred tree iterations.



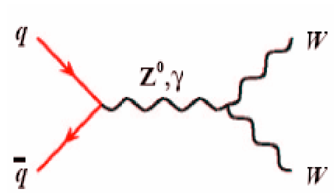
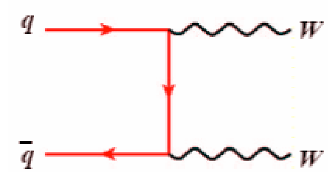
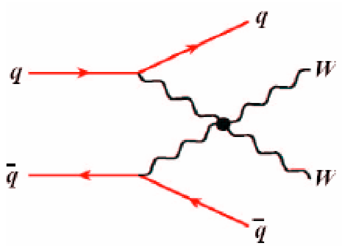
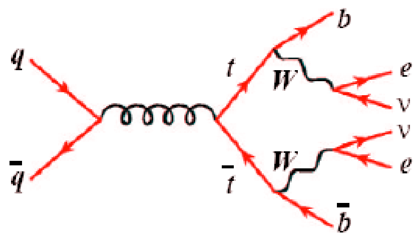
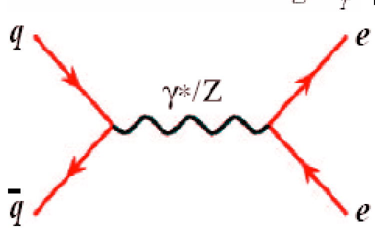
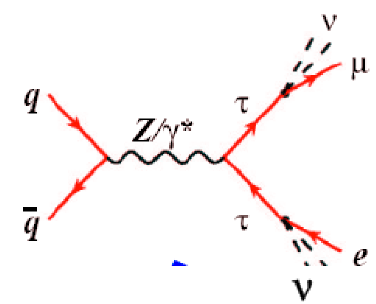
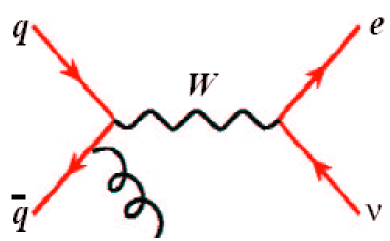
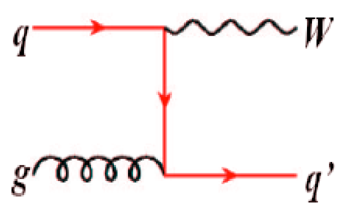
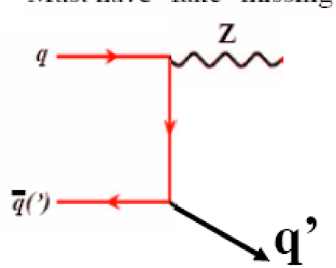
Ref1: H.J.Yang, B.P. Roe, J. Zhu, "*Studies of Boosted Decision Trees for MiniBooNE Particle Identification*", physics/0508045, Nucl. Instum. & Meth. A 555(2005) 370-385.
Ref2: H.J. Yang, B. P. Roe, J. Zhu, " *Studies of Stability and Robustness for Artificial Neural Networks and Boosted Decision Trees* ", physics/0610276, Nucl. Instrum. & Meth. A574 (2007) 342-349.

Diboson analysis – Physics Motivation

Decay modes	ZZ → $\ell^+\ell^-$	ZW → $\ell^+\ell^-\nu\,\ell$	WW → $\ell^+\nu\,\ell^-\nu$
SM • **Triple-gauge-bosons couplings** • **New physics control samples**	$\ell^+\ell^-$ 		
Discovery **H→ ZZ, WW** **SUSY** **Z'→ WW** **G → WW** **ρ_T→ ZW**		 SUSY signal	

WW Analysis

- $W^+W^- \rightarrow e^+\mu^-X, \mu^+e^-X$ (CSC 11 Dataset)

WW -> $\ell^+\nu\,\ell^-\nu + X$ **Signal**			
background	 Contains additional jets.	Must have "fake" missing-E_T 	
background	 Jet must fake a lepton.	 Jet must fake a lepton.	Must have "fake" missing-E_T 

WW analysis - datasets after precuts

MC Process	ID	σ_{mc}(fb)	K	Br	N_{mc}	N_{precut}	N_{test}	Weight
ww_emx	1	0.1133E+06	1.0	0.0120	47000	18233	527.0	0.0289
ww_mex	1	0.1133E+06	1.0	0.0120	48000	18813	532.9	0.0283
ttbar	2	0.7590E+06	1.0	0.5550	688400	22849	13981.7	0.6119
ZGamma_ll	3	0.8910E+06	1.5	0.0672	149742	52	31.2	0.5998
W_enu	4	0.1580E+09	1.3	0.1072	2494958	274	2418.1	8.8254
W_munu	5	0.1580E+09	1.3	0.1072	1998396	304	3349.6	11.0183
W_taunu	6	0.1580E+09	1.3	0.1072	2493808	53	468.0	8.8294
WJET010020_lepnu	7	0.4350E+08	1.3	0.3216	400000	34	1545.9	45.4662
WJET020040_lepnu	8	0.2680E+08	1.3	0.3216	303000	72	2662.5	36.9787
WJET040080_lepnu	9	0.1180E+08	1.3	0.3216	300000	123	2022.7	16.4445
WJET080120_lepnu	10	0.2160E+07	1.3	0.3216	299000	113	341.3	3.0202
WJET120_lepnu	11	0.9080E+06	1.3	0.3216	296000	85	109.0	1.2825
ZJET010020_2e	12	0.1360E+08	1.3	0.0336	597281	123	122.3	0.9946
ZJET020040_2e	13	0.8670E+07	1.3	0.0336	398697	221	209.9	0.9499
ZJET040080_2e	14	0.4120E+07	1.3	0.0336	397524	468	211.9	0.4527
ZJET080120_2e	15	0.8270E+06	1.3	0.0336	397009	408	37.1	0.0910
ZJET120_2e	16	0.3830E+06	1.3	0.0336	198652	157	13.2	0.0842
ZJET010020_2mu	17	0.1360E+08	1.3	0.0336	597413	491	488.2	0.9944
ZJET020040_2mu	18	0.8670E+07	1.3	0.0336	396793	489	466.7	0.9544
ZJET040080_2mu	19	0.4120E+07	1.3	0.0336	776793	1365	316.2	0.2317
ZJET080120_2mu	20	0.8270E+06	1.3	0.0336	396856	813	74.0	0.0910
ZJET120_2mu	21	0.3830E+06	1.3	0.0336	194832	638	54.8	0.0859
ZJET010020_2tau	22	0.1360E+08	1.3	0.0336	598783	1883	1868.1	0.9921
ZJET020040_2tau	23	0.8670E+07	1.3	0.0336	399076	1688	1601.8	0.9490
ZJET040080_2tau	24	0.4120E+07	1.3	0.0336	398972	2487	1121.8	0.4511
ZJET080120_2tau	25	0.8270E+06	1.3	0.0336	396671	3582	326.2	0.0911
ZJET120_2tau	26	0.3830E+06	1.3	0.0336	199046	2984	250.8	0.0840
ZoG030081_2lep	27	0.4220E+07	1.3	0.1010	599000	160	148.0	0.9250
ZoG081100_2lep	28	0.4610E+08	1.3	0.1010	499000	649	7872.4	12.1301
ZoG100_2lep	29	0.1750E+07	1.3	0.1010	493000	1145	533.7	0.4661
WGamma_lnu	30	0.1420E+07	1.0	0.2144	1996438	462	70.5	0.1525
WGamma_taunu	31	0.1420E+07	1.0	0.1072	687999	30	6.6	0.2213
WpZ_lnull.v11004206	32	0.3673E+05	1.0	0.0144	27000	4815	94.3	0.0196
WmZ_lnull.v11004206	33	0.2099E+05	1.0	0.0144	17700	3537	60.4	0.0171
ZZ_llll.v11004206	34	0.1886E+05	1.0	0.0045	36400	5341	12.5	0.0023

Breakdown of MC samples for WW analysis after precuts

MC Process	ID	σ_{mc}(fb)	K	Br	N_{mc}	N_{precut}	N_{test}	Weight
ww_emx	1	0.1133E+06	1.0	0.0120	47000	18233	527.0	0.0289
ww_mex	1	0.1133E+06	1.0	0.0120	48000	18813	532.9	0.0283
ttbar	2	0.7590E+06	1.0	0.5550	688400	22849	13981.7	0.6119
ZGamma_ll	3	0.8910E+06	1.5	0.0672	149742	52	31.2	0.5998
W_enu	4	0.1580E+09	1.3	0.1072	2494958	274	2418.1	8.8254
W_munu	5	0.1580E+09	1.3	0.1072	1998396	304	3349.6	11.0183
W_taunu	6	0.1580E+09	1.3	0.1072	2493808	53	468.0	8.8294

$$N_{test} = \sigma \cdot K \cdot Br \cdot \frac{N_{precut}}{N_{MC}} \underbrace{\int \mathscr{L} dt}_{1\ \mathrm{fb}^{-1}}$$

events in 1 fb^{-1}

11 beam-days @ 10^{33}

$$Weight = \frac{N_{test}}{N_{preselect}}$$

Event Selection for WW-> eμX

Event Pre-selection

- At least one electron + one muon with Pt>10 GeV
- Missing Et > 15 GeV
- Signal efficiency is 39%

Final Selection

- **Simple cuts** based on Rome sample studies
- **Boosted Decision Trees** with 15 input variables

Select of WW -> $e\mu$ / μe + Missing E_T

Simple cuts used in Rome studies

- Two isolated di-lepton P_T > 20 GeV;
  at least one P_T > 25 GeV
- Missing E_T > 30 GeV
- $M_{e\mu}$ > 30 GeV; Veto M_Z (ee, $\mu\mu$)
- E_T(had)=$|\mathrm{Sum}\,(\vec{\ell}_T)+\text{missing}\ \vec{E}_T|$ < 60 GeV,
  Sum Et(jet) < 120 GeV;
- Number of jets < 2
- $P_T(\ell^+\ell^-)$ > 20 GeV
- Vertex between two leptons: ΔZ < 1mm, ΔA < 0.1 mm

For 1 fb^{-1}, 189 signal and 168 background

BDT Training Procedure

- 1st step: use all 48 variables for BDT training, rank variables based on their gini index contributions or how often they were used as tree splitters.

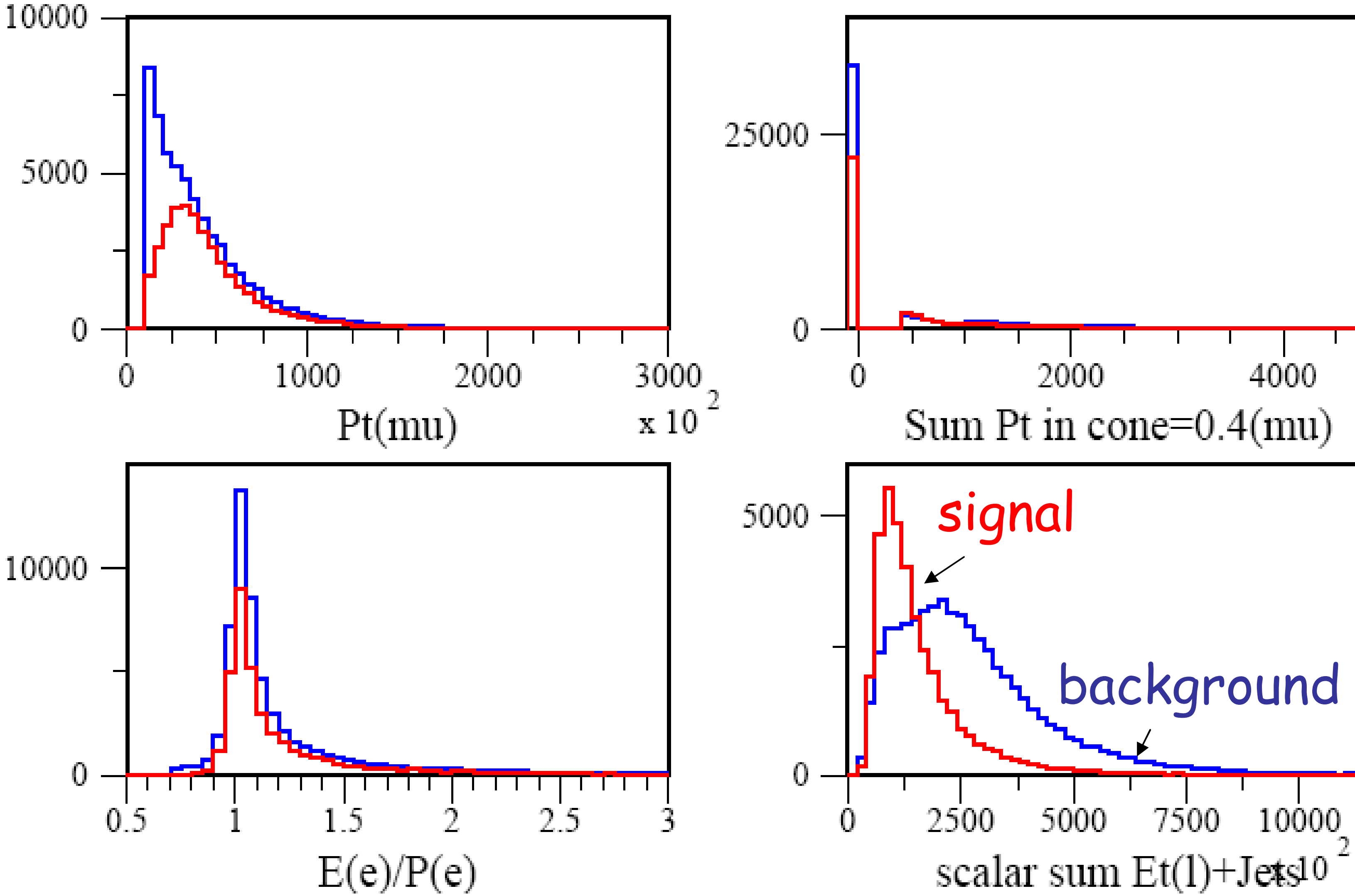
- 2nd step: select 15 powerful variables

- 3rd step: re-train BDT based on 15 selected good variables

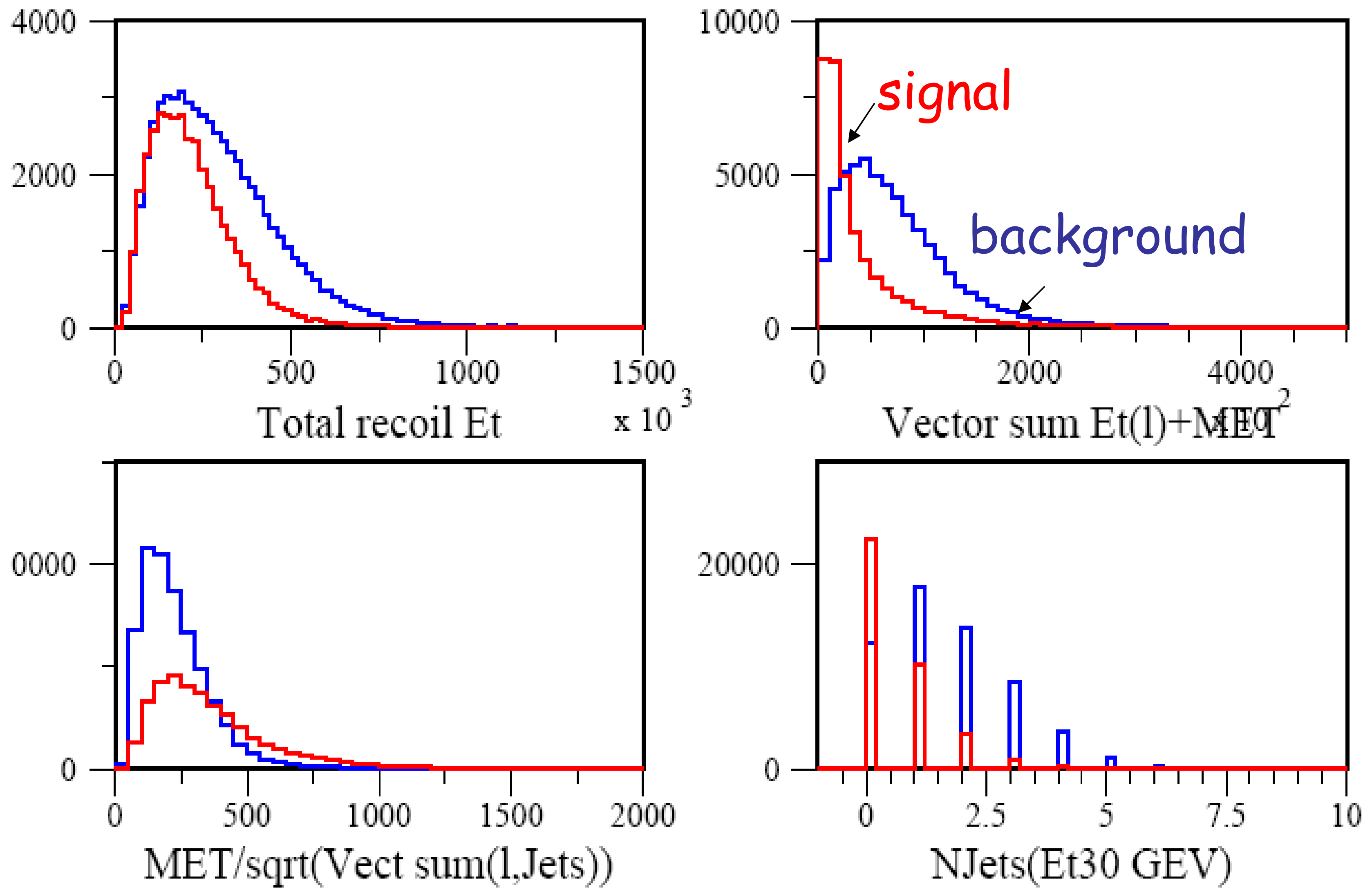
Variables after pre-selection used in BDT

- Sum P_t in cone=0.4(μ)
- E(e)/P(e)
- scalar sum E_t(l)+Jets
- Total recoil E_t
- Vector sum E_t(l)+MET
- MET/sqrt(Vect sum(l,Jets))
- NJets(E_t >30 GeV)
- $\Delta\phi(e, \mu)$
- $P_t(e + \mu)$
- Inv.mass(e,μ)
- Trans.mass(WW)
- $\Delta\phi(e\mu, MET)$
- $\Delta Z(e, \mu)$
- $\Delta A(e, \mu)$

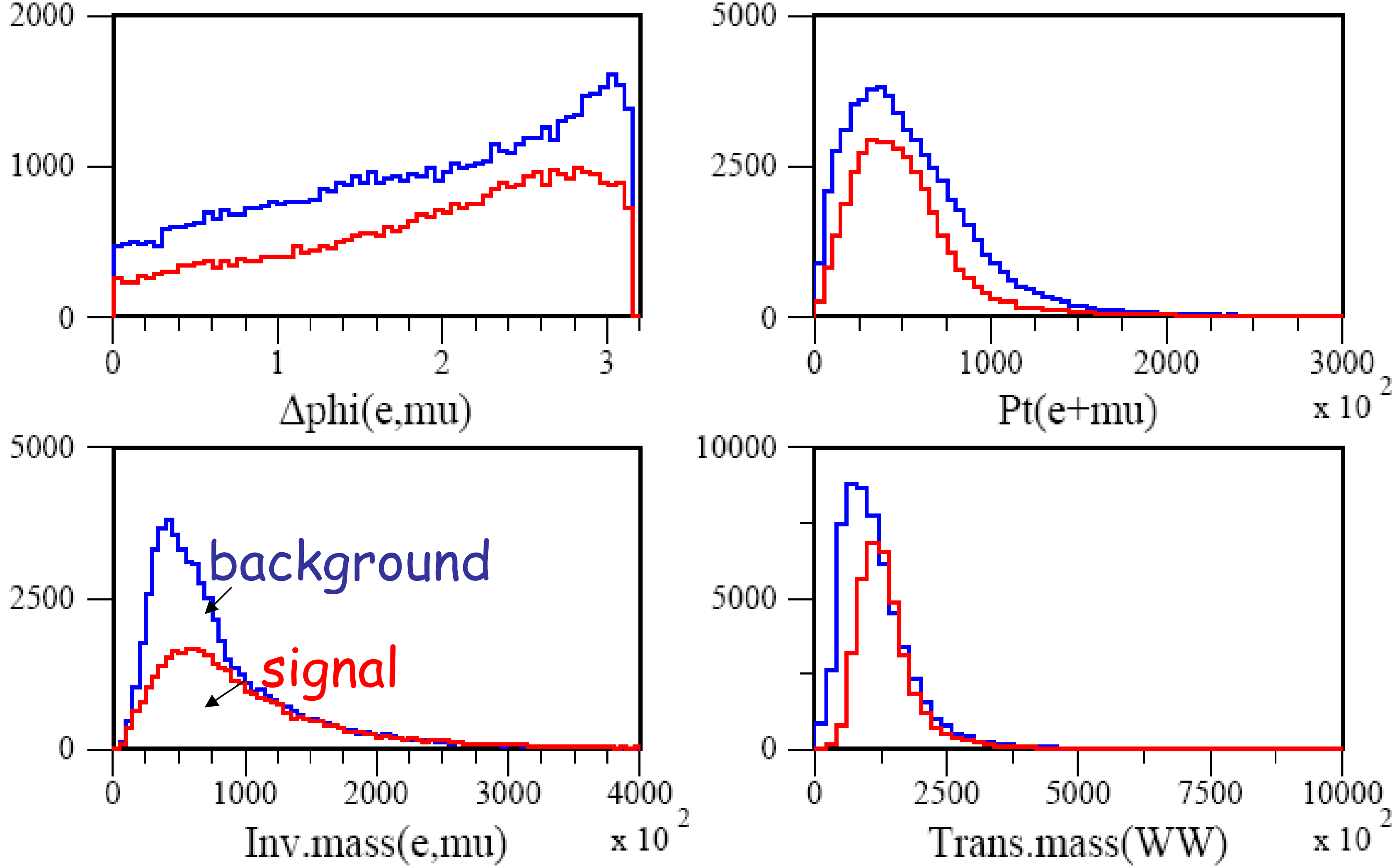
Variable distributions after pre-selection



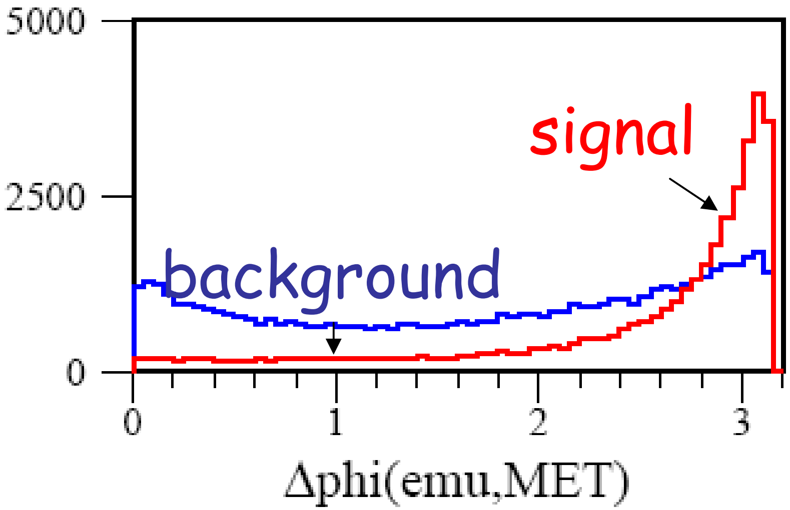
Variable distributions after pre-selection

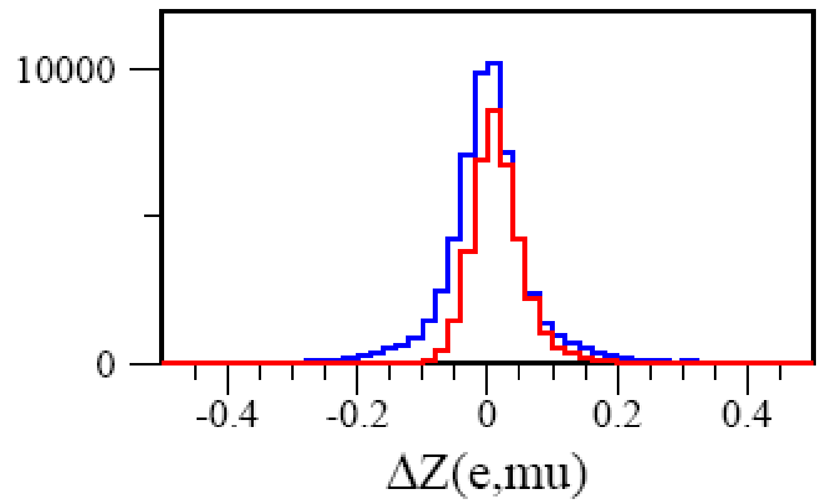


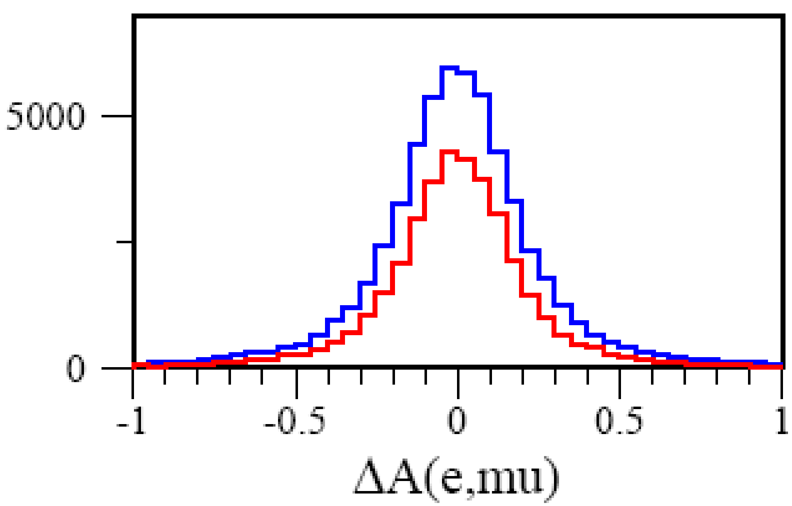
Variable distributions after pre-selection



Variable distributions after pre-selection







BDT Training Tips

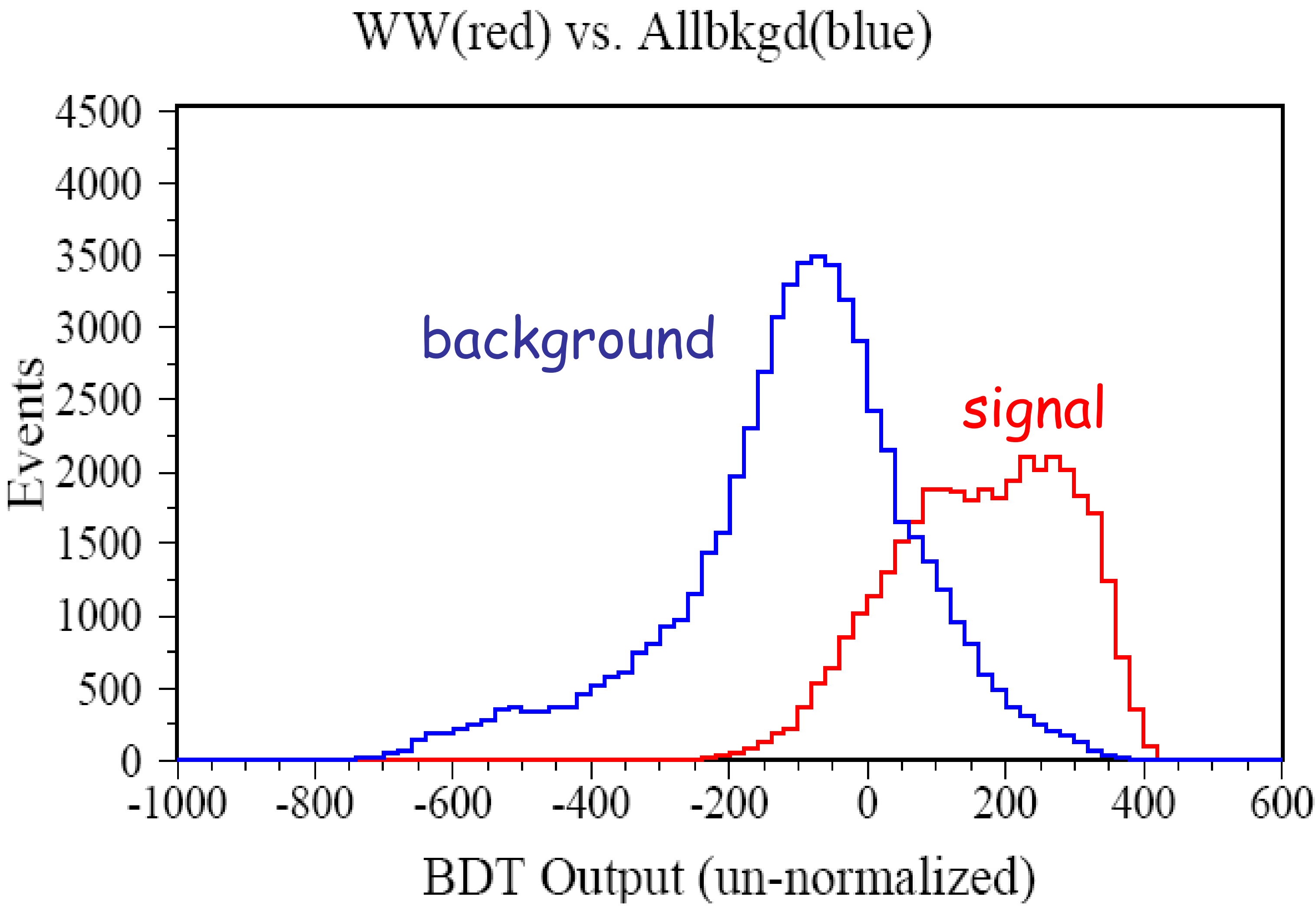
- Epsilon-Boost (epsilon=0.01)

$$w_i \leftarrow w_i \times exp(2\epsilon I(y_i \neq T_m(x_i))), \text{ i=1, 2,...,n}$$

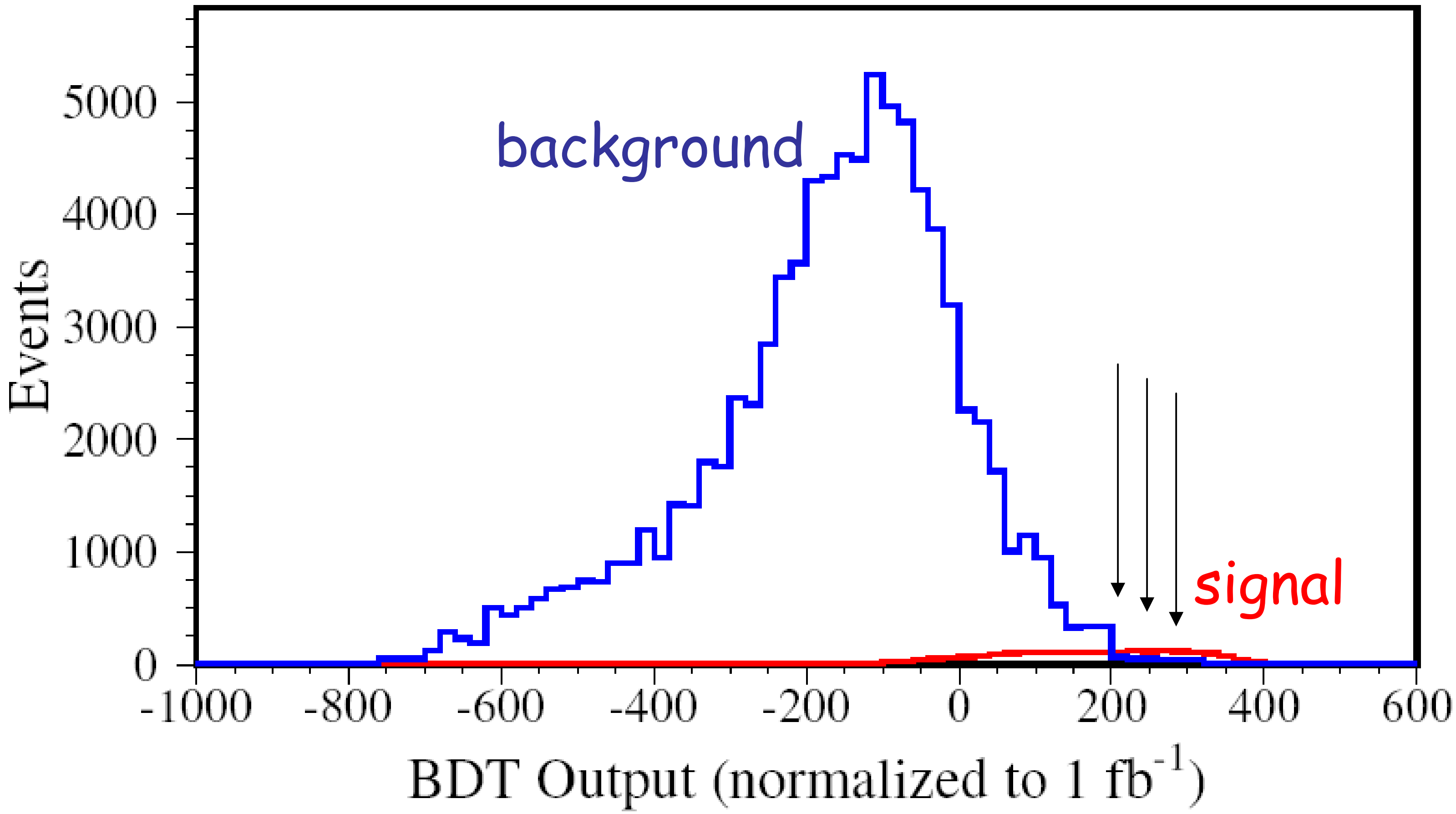
- exp(2*0.01) = 1.0207, I = 1 if training events are misclassified, otherwise I = 0
- 1000 tree iterations, 20 leaves/tree

- The MC samples are split into two halves, one for training, the other for test; then reverse the training and testing samples. The average of testing results are regarded as the final results.

Boosted Decision Trees output



Boosted Decision Trees output



Signal (WW) and Backgrounds for 1 fb^{-1}

Simple Cuts	$Eff_{WW}(\%)$	N_{WW}	N_{BKGD}	N_{WW}/N_{BKGD}	N_σ
	17.9	188.9 ± 2.3	168.0 ± 32.1	1.1	14.6
BDT Cut	$Eff_{WW}(\%)$	N_{WW}	N_{BKGD}	N_{WW}/N_{BKGD}	N_σ
≥ 220	38.5	405.8 ± 3.4	103.2 ± 22.7	3.9	39.9
≥ 230	35.6	375.7 ± 3.3	88.7 ± 22.4	4.2	39.9
≥ 240	32.8	345.6 ± 3.1	79.5 ± 22.3	4.4	38.8
≥ 250	30.1	317.3 ± 3.0	62.2 ± 19.3	5.1	40.2
≥ 260	27.3	287.9 ± 2.9	56.8 ± 19.3	5.1	38.2
≥ 270	24.5	258.4 ± 2.7	42.6 ± 15.8	6.1	39.6
≥ 280	21.6	227.7 ± 2.6	36.6 ± 15.7	6.2	37.6
≥ 290	19.0	200.5 ± 2.4	33.9 ± 15.7	5.9	34.4
≥ 300	16.2	170.2 ± 2.2	20.3 ± 11.1	8.4	37.8
≥ 310	13.6	143.8 ± 2.0	7.2 ± 1.5	19.9	53.4
≥ 320	11.2	117.9 ± 1.8	4.5 ± 1.1	26.2	55.6

MC breakdown with all cuts for 1 fb-1

ww_emx	ID	Simple Cuts	BDT≥220	230	240	250	260	270	280	290	300	310	320
ww_mex	1	188.9	405.8	375.7	345.6	317.3	287.9	258.4	227.7	200.5	170.2	143.8	117.9
ttbar	2	27.5	21.4	15.9	12.9	9.8	8.6	7.3	5.5	4.3	3.7	3.1	1.8
ZGamma_ll	3	1.2	0.6	0.6	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
W_enu	4	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
W_munu	5	77.1	44.1	44.1	44.1	33.1	33.1	22.0	22.0	22.0	11.0	0.0	0.0
W_taunu	6	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
WJET010020_lepnu	7	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
WJET020040_lepnu	8	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
WJET040080_lepnu	9	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
WJET080120_lepnu	10	0.0	3.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
WJET120_lepnu	11	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
ZJET010020_2e	12	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
ZJET020040_2e	13	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
ZJET040080_2e	14	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
ZJET080120_2e	15	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
ZJET120_2e	16	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
ZJET010020_2mu	17	4.0	1.0	1.0	1.0	1.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
ZJET020040_2mu	18	7.6	1.0	1.0	1.0	1.0	1.0	1.0	0.0	0.0	0.0	0.0	0.0
ZJET040080_2mu	19	7.9	2.3	1.6	1.2	1.2	0.7	0.5	0.2	0.2	0.0	0.0	0.0
ZJET080120_2mu	20	2.2	0.5	0.5	0.4	0.3	0.3	0.2	0.1	0.1	0.0	0.0	0.0
ZJET120_2mu	21	0.5	0.1	0.1	0.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
ZJET010020_2tau	22	1.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
ZJET020040_2tau	23	1.9	1.9	0.9	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
ZJET040080_2tau	24	3.6	1.4	0.9	0.5	0.5	0.5	0.5	0.0	0.0	0.0	0.0	0.0
ZJET080120_2tau	25	2.3	1.1	1.1	0.8	0.7	0.7	0.6	0.5	0.3	0.2	0.1	0.1
ZJET120_2tau	26	0.8	0.7	0.3	0.3	0.3	0.2	0.1	0.1	0.1	0.1	0.1	0.1
ZoG030081_2lep	27	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
ZoG081100_2lep	28	12.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
ZoG100_2lep	29	0.9	4.2	3.3	1.9	1.4	0.9	0.9	0.5	0.5	0.5	0.5	0.0
WGamma_lnu	30	1.7	1.7	1.4	1.4	1.1	0.6	0.6	0.5	0.5	0.3	0.2	0.2
WGamma_taunu	31	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
WpZ_lnull.v11004206	32	9.2	10.1	8.8	7.7	6.7	5.8	5.0	4.0	3.3	2.6	1.9	1.4
WmZ_lnull.v11004206	33	5.8	8.1	7.1	6.2	5.3	4.5	3.8	3.3	2.6	1.9	1.5	1.0
ZZ_llll.v11004206	34	0.6	0.3	0.2	0.2	0.1	0.1	0.1	0.1	0.0	0.0	0.0	0.0

Summary for WW Analysis

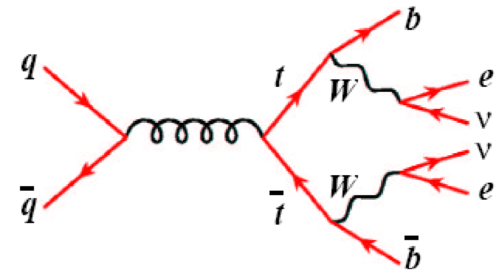
- ➔Background event sample compared to Rome sample increased by a factor of ~10; compared to post Rome sample increased by a factor of ~2.
- ➔Simple Cuts: **S/B ~ 1.1**
- ➔Boosted Decision Trees with 15 variables: **S/B = 5.9**
- ➔The major backgrounds are W-> $\mu\nu$(~50%), ttbar, WZ
- ➔W-> $\mu\nu$ (event weight = 11.02) needs more statistics(x5) if possible.

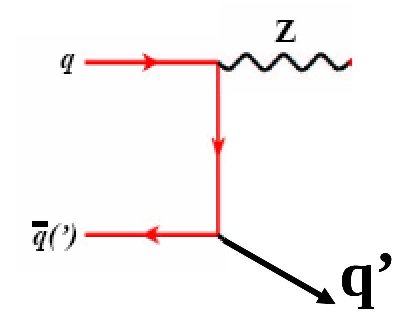
WZ→lνll analyis

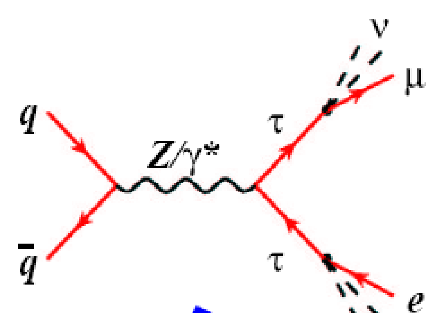
- Physics Goals
  - Test of SM couplings
  - Search for anomalous triple gauge boson couplings (TGCs) that could indicate new physics
  - WZ final state would be a background to SUSY and technicolor signals.
- WZ event selection by two methods
  - Simple cuts
  - Boosted Decision Trees

WZ selection – Major backgrounds

- Major backgrounds
  - pp → t tbar
    - Pair of leptons fall in Z mass window
    - Jet produces lepton signal
  - pp → Z+jets
    - Fake missing E_T
    - Jet produces third lepton signal
  - pp → Z/γ → ee, mm
    - Fake missing E_T and third lepton
  - pp → ZZ→ 4 leptons
    - Lose a lepton







Pre-selection for WZ analysis

- Pre-selection
  - Identify leptons and require
    - $p_T > 5$GeV, one with $p_T > 20$GeV
  - Require missing $E_T > 15$GeV
  - Find e^+e^- or m^+m^- pair with inv. mass closest to Z peak
    - must be within 91.18 ± 20 GeV.
  - Third lepton with $p_T > 15$GeV and $10 < M_T < 400$GeV
- Eff(W^+Z) = 25.8%, Eff(W^-Z) = 29.3%

- Compute additional variables(invariant masses, sums of jets, track isolations …), 67 variables in total

WZ analysis – datasets after precuts

MC Process	ID	σ_{mc}(fb)	K	Br	N_{mc}	N_{precut}	N_{test}	Weight
WpZ_5941	0	0.3673E+05	1.0	0.0144	26550	6848	136.4	0.0199
WmZ_5971	1	0.2099E+05	1.0	0.0144	17450	5118	88.7	0.0173
Z_2mu_0001	2	0.4610E+08	1.3	0.0336	1153109	0	0.0	1.7463
ZGamma_ll_4190	3	0.8910E+06	1.5	0.0672	999742	111	10.0	0.0898
ZJET010020_2e_v10_4270	4	0.1360E+08	1.3	0.0336	597281	0	0.0	0.9946
ZJET020040_2e_v10_4271	5	0.8670E+07	1.3	0.0336	398697	0	0.0	0.9499
ZJET040080_2e_v10_4272	6	0.4120E+07	1.3	0.0336	397524	0	0.0	0.4527
ZJET080120_2e_v10_4273	7	0.8270E+06	1.3	0.0336	397009	0	0.0	0.0910
ZJET120_2e_v10_4274	8	0.3830E+06	1.3	0.0336	198652	0	0.0	0.0842
ZJET010020_2tau_v10_4275	9	0.1360E+08	1.3	0.0336	598783	0	0.0	0.9921
ZJET020040_2tau_v10_4276	10	0.8670E+07	1.3	0.0336	399076	0	0.0	0.9490
ZJET040080_2tau_v10_4277	11	0.4120E+07	1.3	0.0336	398972	0	0.0	0.4511
ZJET080120_2tau_v10_4278	12	0.8270E+06	1.3	0.0336	396671	0	0.0	0.0911
ZJET120_2tau_v10_4279	13	0.3830E+06	1.3	0.0336	199046	0	0.0	0.0840
ZJET010020_mumu_4290	14	0.1360E+08	1.3	0.0336	2996413	492	97.5	0.1983
ZJET020040_mumu_4291	15	0.8670E+07	1.3	0.0336	1995792	789	149.7	0.1898
ZJET040080_mumu_4292	16	0.4120E+07	1.3	0.0336	1189793	1516	229.3	0.1513
ZJET080120_mumu_4293	17	0.8270E+06	1.3	0.0336	397856	1105	100.3	0.0908
ZJET120_mumu_4294	18	0.3830E+06	1.3	0.0336	199832	1133	94.9	0.0837
ZoG030081_2lep_4295	19	0.4220E+07	1.3	0.1010	1000000	16	8.9	0.5541
ZoG081100_2lep_4296	20	0.4610E+08	1.3	0.1010	3284999	406	748.1	1.8426
ZoG100_2lep_4297	21	0.1750E+07	1.3	0.1010	971000	271	64.1	0.2366
JimmyZmumuM150_5115	22	0.1750E+07	0.8	0.0336	43000	33	36.1	1.0940
Zmumu_5145	23	0.1497E+07	1.0	1.0000	48000	0	0.0	31.1875
PythiaZmumuJet_5181	24	0.8270E+06	0.8	0.0336	35000	20	12.7	0.6351
PythiaZee_pt100_5185	25	0.8270E+06	0.8	0.0336	46000	11	5.3	0.4833
PythiaZmumu_pt100_5186	26	0.8270E+06	0.8	0.0336	33000	42	28.3	0.6736
PythiaZtautau_pt100_5187	27	0.8270E+06	0.8	0.0003	32000	41	0.3	0.0069
T1_McAtNlo_Jimmy_5200	28	0.7590E+06	1.0	0.5550	604750	1071	746.0	0.6966
PythiaZPhoton25_5900	29	0.4510E+05	1.0	0.0672	46800	43	2.8	0.0648
WpWm_enuenu_5921	30	0.1133E+06	1.0	0.0120	41950	9	0.3	0.0324
WpWm_enumunu_5922	31	0.1133E+06	1.0	0.0120	45900	22	0.7	0.0296
WpWm_enutaunu_5923	32	0.1133E+06	1.0	0.0120	71000	7	0.1	0.0191
WpWm_munuenu_5925	33	0.1133E+06	1.0	0.0120	47000	18	0.5	0.0289
WpWm_munumunu_5924	34	0.1133E+06	1.0	0.0120	48950	30	0.8	0.0278
WpWm_munutaunu_5926	35	0.1133E+06	1.0	0.0120	44000	8	0.2	0.0309
WpWm_taunuenu_5928	36	0.1133E+06	1.0	0.0120	47700	2	0.1	0.0285
WpWm_taunumunu_5929	37	0.1133E+06	1.0	0.0120	45800	8	0.2	0.0297
WpWm_taunutaunu_5927	38	0.1133E+06	1.0	0.0120	34850	0	0.0	0.0390
ZZ_llll_5931	39	0.1886E+05	1.0	0.0045	35700	8597	20.4	0.0024
Pythiagamgam_5960	40	0.7100E+05	1.0	1.0000	45300	0	0.0	1.5673

Final selection – simple cuts

Based on pre-selected events, make further cuts

- Lepton selection
  - Isolation: leptons have tracks totaling < 8 GeV within DR<0.4
  - Z leptons have p_T > 6 GeV
  - W lepton p_T > 25 GeV
  - E ~ p for electrons: 0.7 < E/p < 1.3
  - Hollow cone around leptons has little energy: [E_T(DR<0.4) – E_T(DR<0.2)] / E_T < 0.1
- Leptons separated by DR > 0.2
- Exactly 3 leptons
- Missing E_T > 25 GeV
- Few jets
  - No more than one jet with E_T > 30 GeV in |h| < 3
  - Scalar sum of jet E_T < 200 GeV
- Leptonic energy balance
  - | Vector sum of leptons and missing E_T | < 100 GeV
- Z mass window: ±9 GeV for electrons and ±12 GeV for muons
- W mass window: 40 GeV < M_T(W) < 120 GeV

Simple cuts – results (Alan)

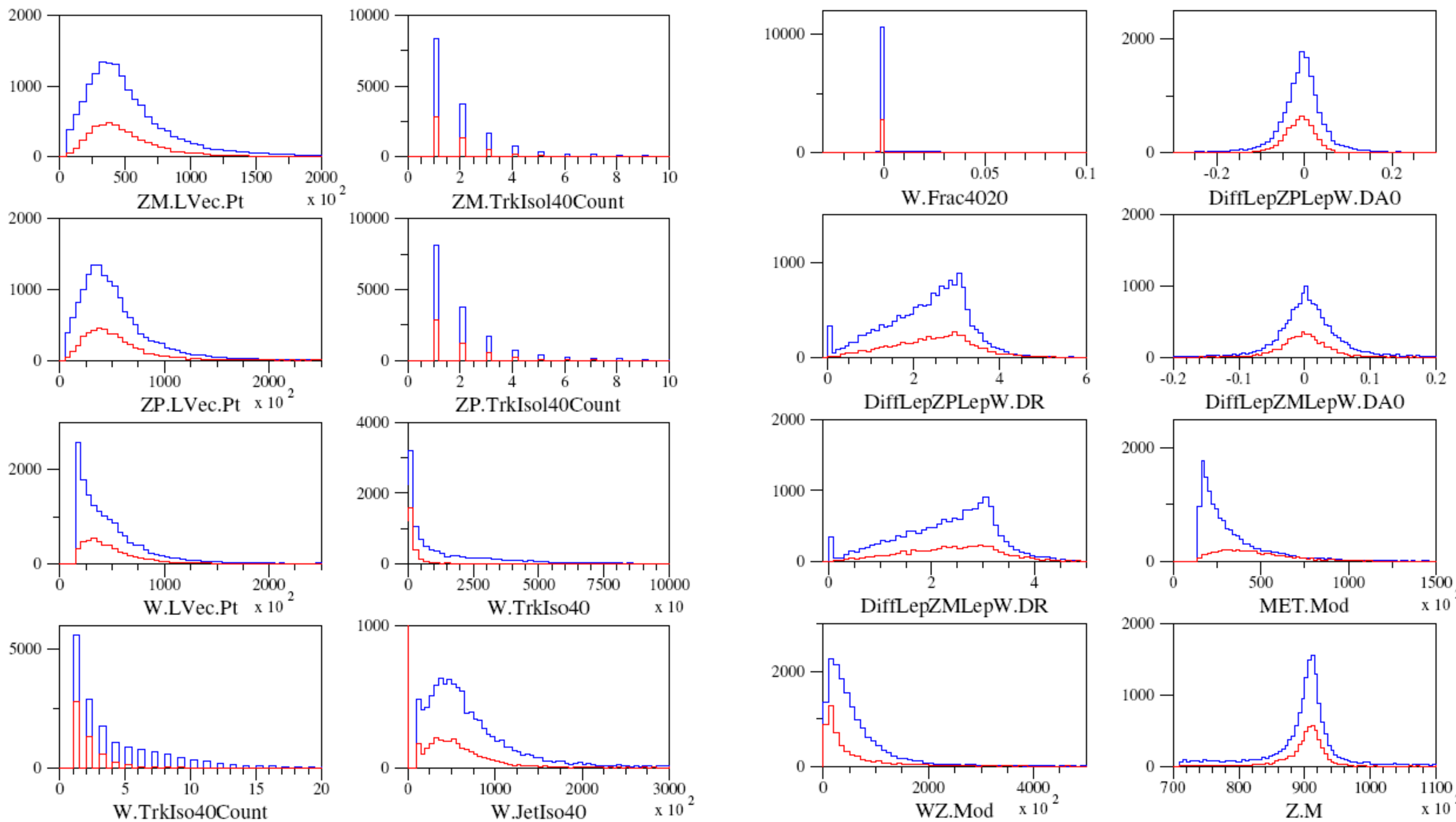
	Dataset	Pre-selection			Simple Cuts						
		nevents	%	in 1fb-1	nevents	%	in fb-1	eee	eem	mme	mmm
5971	WmZ	5118	29.3295	88.6	1587	9.09456	27.5	2.04	4.62	7.39	13.41
5941	WpZ	6848	25.7928	136.4	1964	7.39736	39.1	2.99	6.02	11.22	18.91
	TOTAL	11966	27.1955	225	3551	8.07045	$\boxed{\textbf{66.6}}$ signal	5.03	10.64	18.61	32.32
4190	ZGamma_ll	111	0.0111	10	3	0.0003	0.3	0.09	0	0.18	0
4270	ZJET010020_2e_v10	100	0.01674	99.5	0	0	0	0	0	0	0
4271	ZJET020040_2e_v10	128	0.0321	121.7	0	0	0	0	0	0	0
4272	ZJET040080_2e_v10	330	0.08301	159.4	1	0.00025	0.5	0.48	0	0	0
4273	ZJET080120_2e_v10	702	0.17682	63.8	0	0	0	0	0	0	0
4274	ZJET120_2e_v10	631	0.31764	53	0	0	0	0	0	0	0
4290	ZJET010020_mumu	492	0.01642	97.5	7	0.00023	1.4	0	0	0.59	0.79
4291	ZJET020040_mumu	789	0.03953	149.8	16	0.0008	3	0	0	0.76	2.28
4292	ZJET040080_mumu	1516	0.12742	244.6	11	0.00092	1.8	0	0	0.65	1.13
4293	ZJET080120_mumu	1105	0.27774	100.3	7	0.00176	0.6	0	0	0	0.64
4294	ZJET120_mumu	1133	0.56698	94.7	0	0	0	0	0	0	0
4295	ZoG030081_2lep	16	0.0016	8.9	0	0	0	0	0	0	0
4296	ZoG081100_2lep	406	0.01236	748.1	6	0.00018	11.1	0	5.53	0	5.53
4297	ZoG100_2lep	271	0.02791	64.2	3	0.00031	0.7	0	0	0	0.71
5200	T1_McAtNlo_Jimmy	1071	0.1771	816.4	5	0.00083	3.8	0	0	0	3.81
5931	ZZ_llll	8597	24.0812	20.5	1394	3.90476	3.3	0.01	0.79	0.11	2.41
	TOTAL	17420	0.09865	2856	1453	0.00823	$\boxed{\textbf{26.5}}$ background	0.1	6.8	2.29	17.3

WZ - Boosted Decision Trees

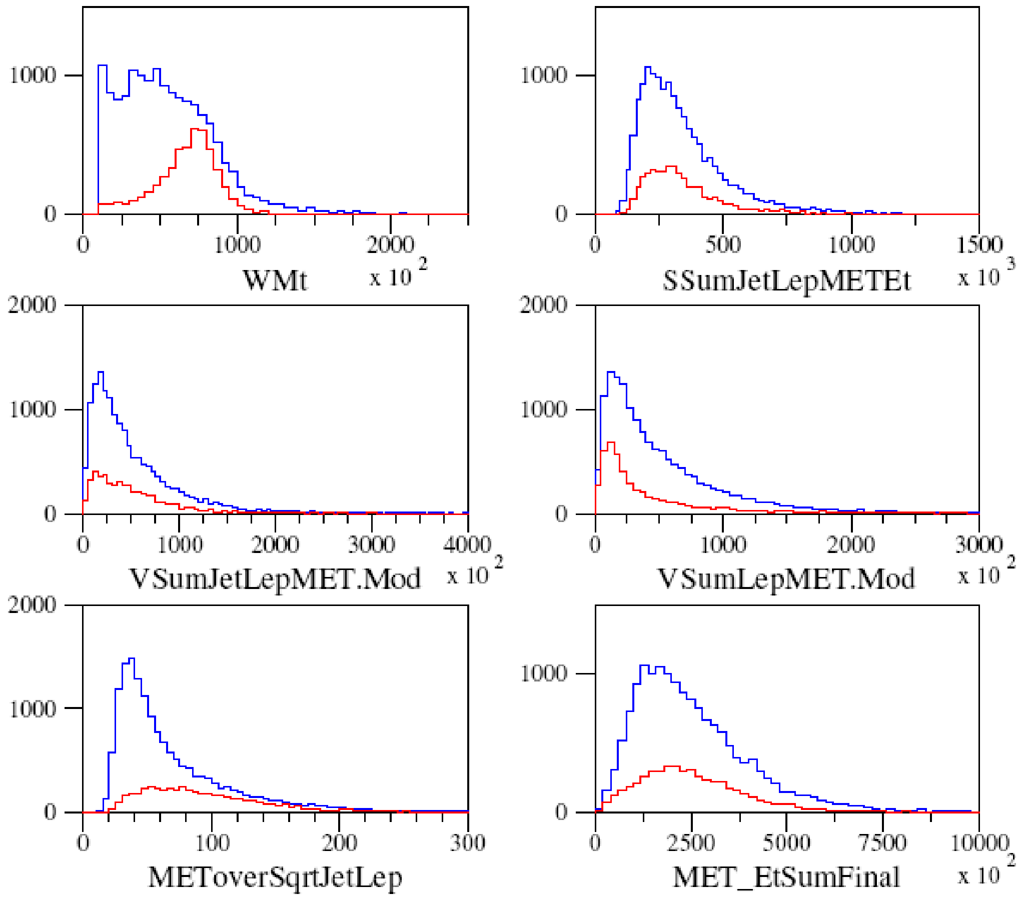
- Select 22 powerful variables out of 67 total available variables for BDT training.
- Rank 22 variables based on the gini index contributions and the number of times used as tree splitters.
- W lepton track isolation, M(Z) and $M_T(W)$ are ranked highest.

Name of Variable	Rank (tree splitters)	Rank (gini index)
ZM.LVec.Pt	12	14
ZM.TrkIsol40Count	14	6
ZP.LVec.Pt	16	17
ZP.TrkIsol40Count	11	4
W.LVec.Pt	10	8
W.TrkIso40	2	2
W.TrkIso40Count	5	1
W.JetIso40	15	15
W.Frac4020	17	16
DiffLepZPLepW.DA0	6	11
DiffLepZPLepW.DR	8	13
DiffLepZMLepW.DA0	4	7
DiffLepZMLepW.DR	9	12
MET.Mod	13	9
WZ.Mod	18	18
Z.M	1	3
WMt	3	5
SSumJetLepMETEt	21	20
VSumJetLepMET.Mod	20	21
VSumLepMET.Mod	22	22
METoverSqrtJetLep	19	19
MET_EtSumFinal	7	10

Input Variables(1-16)



Input Variables(17-22)



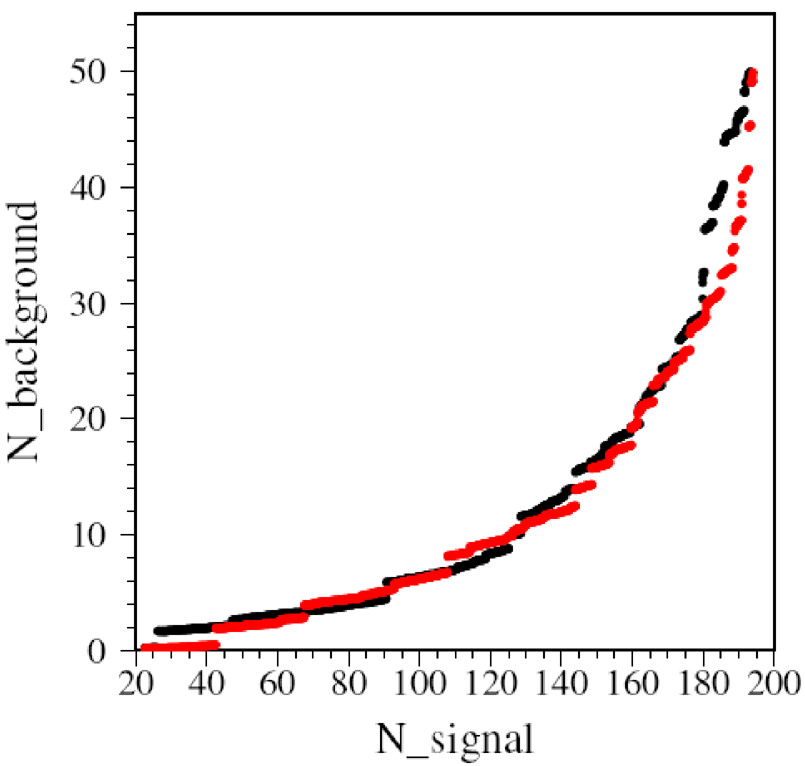
BDT Training Tips

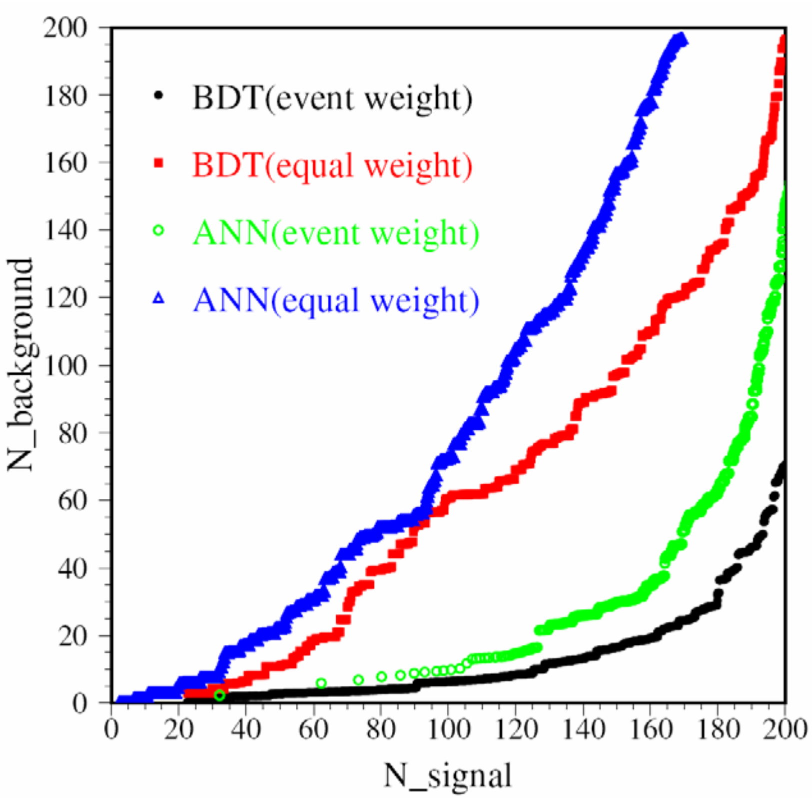
- In the original BDT training program, all training events are set to have same weights in the beginning (the first tree). It works fine if all MC processes are produced based on their production rates.
- Our MCs are produced separately, the event weights vary from various backgrounds. e.g. assuming 1 fb^{-1} wt (ZZ_llll) = 0.0024, wt (ttbar) = 0.7, wt(DY) = 1.8
- We made two BDT trainings. One based on equal event weights for all training MC; the other based on their correct event weights for the 1st tree training.
- BDT performance with correct event weights for training works better than that with equal weights.

BDT Tuning

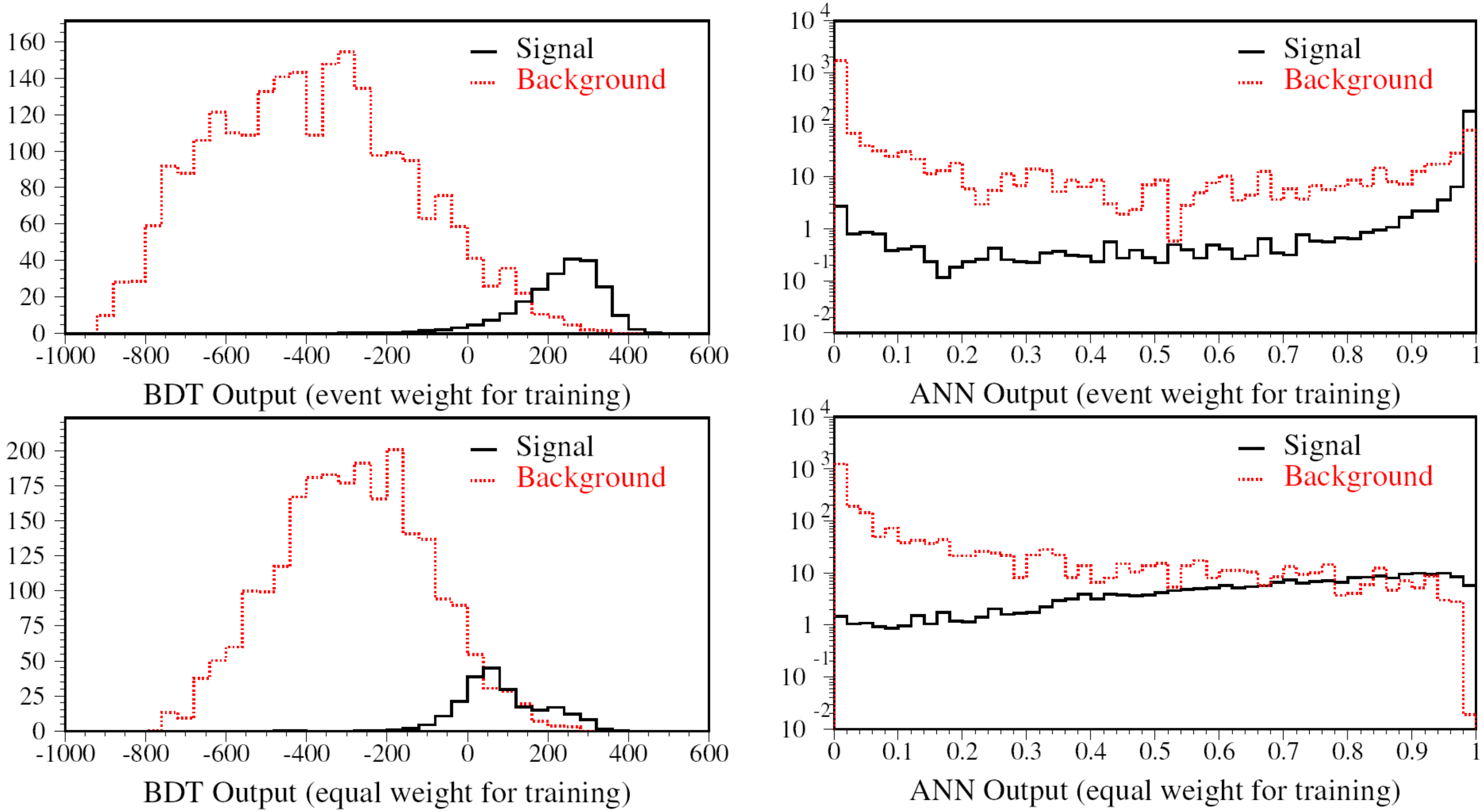
- ➔ BDT with 22 and 67 variables have comparable performance
- ➔ ANN and BDT training with correct event weights works significantly better than that with equal event weights

BDT with 22 Var(black) vs 67 Var(red)





ANN/BDT Tuning



ANN/BDT Comparison

➔Event weight training technique works better than equal weight training for both ANN(x5-7) and BDT(x6-10)

➔BDT is better than ANN by reducing more background(x1.5-2)

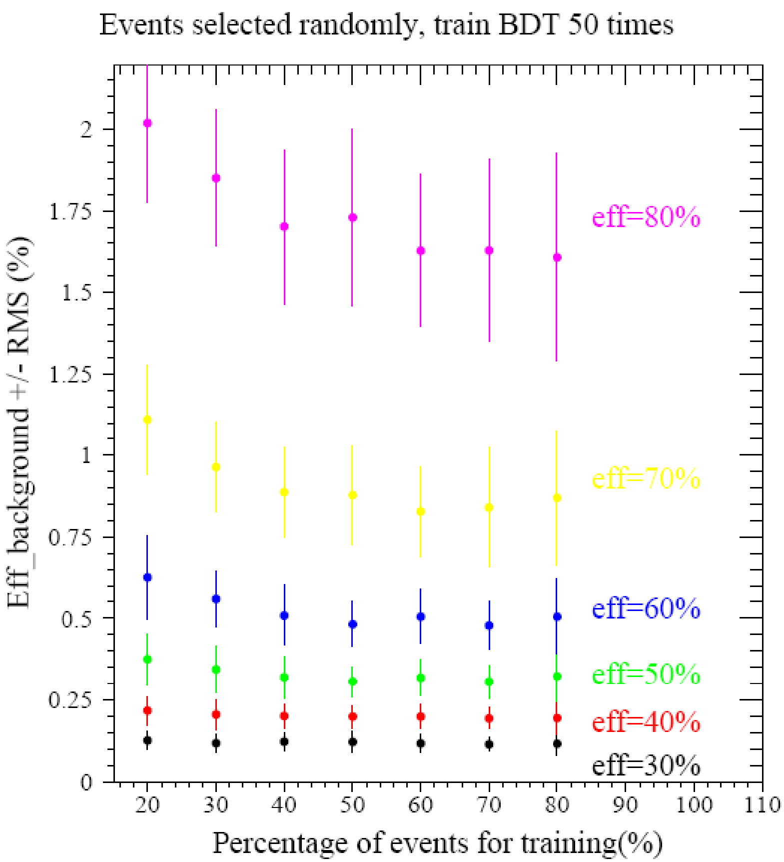
➔ A note to describe the event weight training technique in detail will be available shortly.

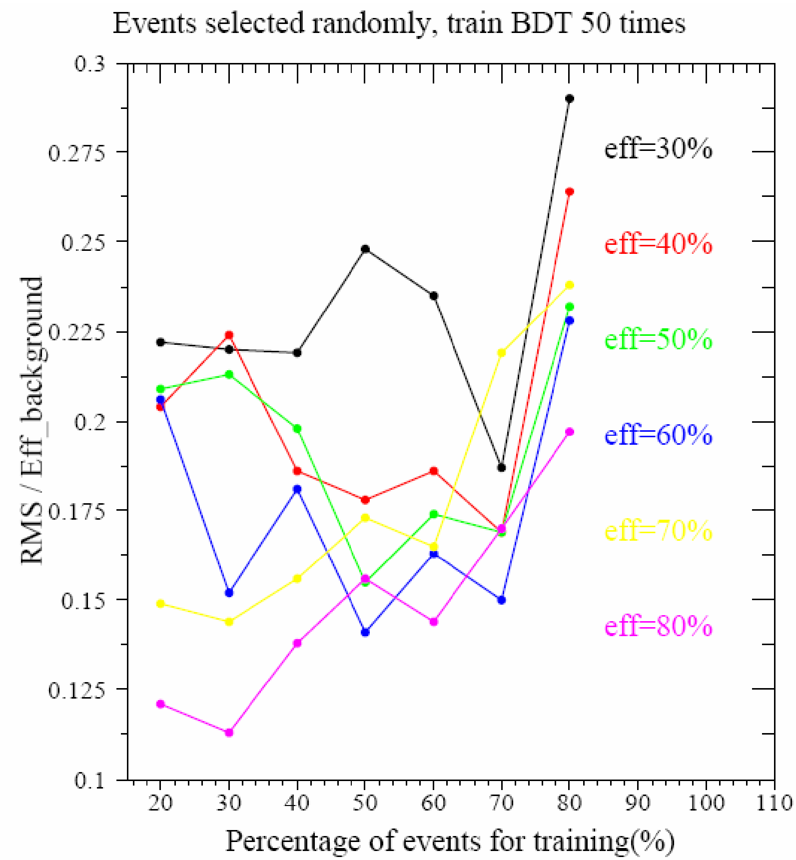
N_{signal}	60	80	100	120	140	160
N_{bg1} for ANN-equal-weight	30.5	51.9	72.4	104.7	133.3	177.6
N_{bg2} for ANN-event-weight	5.8	7.7	9.8	14.7	25.9	34.9
$Ratio = N_{bg1}/N_{bg2}$ for ANN	5.3	6.7	7.4	7.1	5.1	5.1
N_{bg3} for BDT-equal-weight	18.5	39.4	60.7	69.1	88.9	110.1
N_{bg4} for BDT-event-weight	3.1	4.0	6.3	8.4	13.2	19.3
$Ratio = N_{bg3}/N_{bg4}$ for BDT	6.0	9.9	9.6	8.2	6.7	5.7
$Ratio = N_{bg2}/N_{bg4}$ for ANN/BDT	1.90	1.93	1.56	1.75	1.96	1.81

Eff_bkgd/RMS vs Training Events

➔BDT is not well trained with few training events, it will result in large Eff_bkgd for a set of given Eff_sig

➔RMS errors tend to be larger with fewer events for testing. For a limited MC, it's better to balance training & testing events.





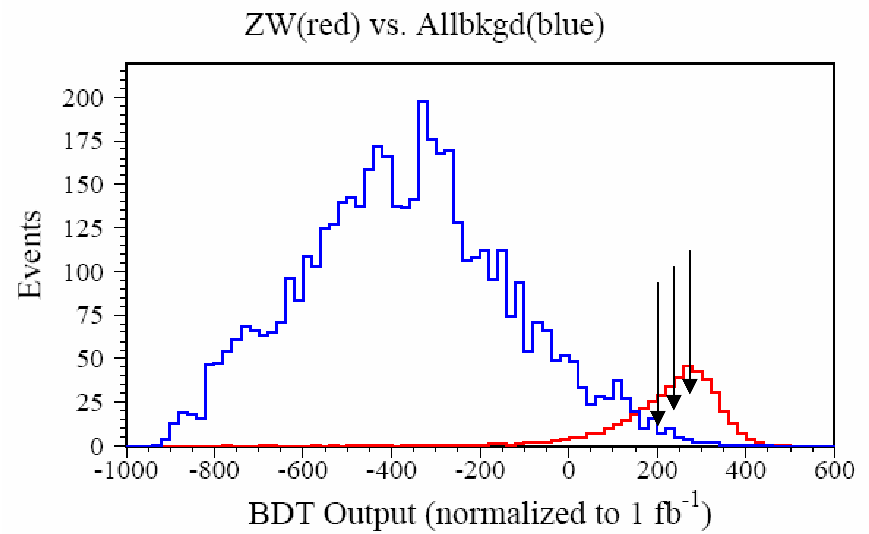
WZ - Boosted Decision Trees

For 1 fb^{-1} , BDT Results

- N_{signal} = 150 to 60
- Significance ($N_{signal}/\sqrt{N_{bkg}}$) ~ 40

- BDT, S/BG ~ 10 to 24
- Simple cuts

  S/BG ~ 2-2.5



Simple Cuts	$Eff_{ZW}(\%)$	N_{ZW}	N_{BKGD}	N_{ZW}/N_{BKGD}	N_σ
Alan	28.4	66.6	27.0	2.5	12.8
Bing/Haijun	31.9	74.8	36.6	2.0	12.4
Zhengguo	22.2	52.0	46.3	1.1	7.6
BDT Cut	$Eff_{ZW}(\%)$	N_{ZW}	N_{BKGD}	N_{ZW}/N_{BKGD}	N_σ
≥ 200	65.1	152.6 ± 1.7	16.1 ± 2.5	9.5	38.1
≥ 210	62.2	145.7 ± 1.7	14.6 ± 2.4	10.0	38.1
≥ 220	59.0	138.3 ± 1.6	12.6 ± 2.3	11.0	39.0
≥ 230	55.5	130.1 ± 1.6	10.7 ± 2.2	12.1	39.7
≥ 240	51.8	121.4 ± 1.5	7.7 ± 1.1	15.7	43.7
≥ 250	47.7	111.9 ± 1.5	6.6 ± 1.1	17.0	43.6
≥ 260	43.5	102.0 ± 1.4	5.5 ± 1.0	18.5	43.5
≥ 270	38.7	90.6 ± 1.3	4.1 ± 0.8	22.1	44.8
≥ 280	33.7	79.0 ± 1.2	3.5 ± 0.7	22.6	42.2
≥ 290	29.4	68.8 ± 1.1	3.0 ± 0.7	23.1	39.9
≥ 300	24.8	58.0 ± 1.0	2.4 ± 0.7	24.1	37.4

ZW → eee, eeμ, μμe, μμμ

Simple Cuts	eee N_{ZW}/N_{BG}	$ee\mu$ N_{ZW}/N_{BG}	$\mu\mu e$ N_{ZW}/N_{BG}	$\mu\mu\mu$ N_{ZW}/N_{BG}	All 4 Channels N_{ZW}/N_{BG}
Alan	5.03/ 0.58	10.6/ 6.8	18.6/ 2.29	32.3/ 17.3	66.6/ 27.0
Zhengguo	11.7/ 4.07	13.4/ 5.16	13.6/ 23.5	13.3/ 10.7	52.0/ 46.3
BDT Cut	eee N_{ZW}/N_{BG}	$ee\mu$ N_{ZW}/N_{BG}	$\mu\mu e$ N_{ZW}/N_{BG}	$\mu\mu\mu$ N_{ZW}/N_{BG}	All 4 Channels N_{ZW}/N_{BG}
≥ 200	31.7/ 4.0	34.9/ 2.7	39.5/ 3.9	46.6/ 5.5	152.6/ 16.1
≥ 210	30.5/ 3.9	33.2/ 2.5	37.9/ 3.5	44.1/ 4.7	145.7/ 14.6
≥ 220	29.4/ 3.1	31.6/ 2.3	36.0/ 3.0	41.2/ 4.2	138.3/ 12.6
≥ 230	28.1/ 3.0	29.7/ 2.1	34.3/ 1.9	37.9/ 3.7	130.1/ 10.7
≥ 240	26.7/ 0.6	27.8/ 1.9	32.8/ 1.8	34.2/ 3.3	121.4/ 7.7
≥ 250	25.2/ 0.5	25.4/ 1.7	30.6/ 1.5	30.6/ 2.9	111.9/ 6.6
≥ 260	23.7/ 0.2	23.3/ 1.5	28.3/ 1.2	26.7/ 2.6	102.0/ 5.5
≥ 270	21.8/ 0.2	20.6/ 1.3	25.7/ 1.1	22.5/ 1.5	90.6/ 4.1
≥ 280	19.9/ 0.2	17.5/ 1.1	23.0/ 1.0	18.6/ 1.2	79.0/ 3.5
≥ 290	18.2/ 0.2	14.8/ 0.8	20.6/ 1.0	15.2/ 1.0	68.8/ 3.0
≥ 300	16.4/ 0.1	11.6/ 0.7	18.1/ 1.0	11.9/ 0.7	58.0/ 2.4

MC breakdown with all cuts for 1 fb-1

MC Process	ID	Simple Cuts	BDT$\geq$ 200	210	220	230	240	250	260	270	280	290	300
WpZ_5941	0	33.5	91.3	87.2	82.5	77.8	72.6	66.9	60.7	54.0	47.1	41.0	34.8
WmZ_5971	1	41.3	61.3	58.5	55.8	52.3	48.7	45.0	41.3	36.5	31.9	27.8	23.2
Z_2mu_0001	2	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
ZGamma_ll_4190	3	0.4	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.0
ZJET010020_2e_v10_4270	4	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
ZJET020040_2e_v10_4271	5	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
ZJET040080_2e_v10_4272	6	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
ZJET080120_2e_v10_4273	7	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
ZJET120_2e_v10_4274	8	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
ZJET010020_2tau_v10_4275	9	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
ZJET020040_2tau_v10_4276	10	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
ZJET040080_2tau_v10_4277	11	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
ZJET080120_2tau_v10_4278	12	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
ZJET120_2tau_v10_4279	13	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
ZJET010020_mumu_4290	14	2.6	0.6	0.6	0.2	0.2	0.2	0.2	0.2	0.2	0.2	0.2	0.2
ZJET020040_mumu_4291	15	3.8	0.4	0.2	0.2	0.2	0.2	0.0	0.0	0.0	0.0	0.0	0.0
ZJET040080_mumu_4292	16	3.3	0.8	0.5	0.5	0.2	0.2	0.2	0.2	0.2	0.2	0.2	0.2
ZJET080120_mumu_4293	17	0.9	0.6	0.5	0.4	0.2	0.2	0.0	0.0	0.0	0.0	0.0	0.0
ZJET120_mumu_4294	18	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.0	0.0	0.0	0.0	0.0
ZoG030081_2lep_4295	19	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
ZoG081100_2lep_4296	20	16.6	1.8	1.8	1.8	1.8	0.0	0.0	0.0	0.0	0.0	0.0	0.0
ZoG100_2lep_4297	21	0.7	0.2	0.2	0.2	0.2	0.2	0.2	0.0	0.0	0.0	0.0	0.0
JimmyZmumuM150_5115	22	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Zmumu_5145	23	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
PythiaZmumuJet_5181	24	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
PythiaZee_pt100_5185	25	0.0	0.5	0.5	0.5	0.5	0.0	0.0	0.0	0.0	0.0	0.0	0.0
PythiaZmumu_pt100_5186	26	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
PythiaZtautau_pt100_5187	27	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
T1_McAtNlo_Jimmy_5200	28	4.2	2.8	2.8	2.1	1.4	1.4	1.4	1.4	0.7	0.7	0.7	0.7
PythiaZPhoton25_5900	29	0.2	0.5	0.3	0.3	0.3	0.3	0.2	0.1	0.0	0.0	0.0	0.0
WpWm_enuenu_5921	30	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
WpWm_enumunu_5922	31	0.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
WpWm_enutaunu_5923	32	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
WpWm_munuenu_5925	33	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
WpWm_munumunu_5924	34	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
WpWm_munutaunu_5926	35	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
WpWm_taunuenu_5928	36	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
WpWm_taunumunu_5929	37	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
WpWm_taunutaunu_5927	38	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
ZZ_llll_5931	39	3.7	7.7	6.9	6.2	5.6	4.9	4.3	3.6	3.0	2.4	1.8	1.4
Pythiagamgam_5960	40	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0

Summary for WZ Analysis

- ➔ Simple Cuts

  S/BG = 2 ~ 2.5

- ➔ Boosted Decision Trees with 22 variables

  S/BG = 10 ~ 24

- ➔ The major backgrounds are (BDT>=200):

  ZZ -> 4 l (47.8%)
  ZJet -> 2μ X (15.5%)
  ttbar (17.4%)
  Drell-Yan -> 2 l (12.4%)

Applications of BDT in HEP

- Boosted Decision Trees (BDT) has been applied for some major HEP experiments in the past few years.
  - MiniBooNE data analysis (BDT reject 20-80% more background than ANN)
    - physics/0408124 (NIM A543, p577), physics/0508045 (NIM A555, p370),
    - physics/0610276(NIM A574, p342), physics/0611267
    - " A search for electron neutrino appearance at dm^2 ~ 1 eV^2 Scale", hep-ex/0704150 (submitted to PRL)
  - ATLAS Di-Boson analysis, ww, wz, wγ, zγ
  - ATLAS SUSY analysis – hep-ph/0605106 (JHEP060740)
  - LHC B-tagging, physics/0702041, for 60% b-tagging eff, BDT has 35% more light jet rejection than that of ANN.
  - BaBar data analysis
    - "Measurement of CP-violating asymmetries in the B0->K+K-K0 dalitz plot", hep-ex/0607112
    - physics/0507143, physics/0507157
  - D0 data analysis
    - hep-ph/0606257, Fermilab-thesis-2006-15,
    - "Evidence of single top quarks and first direct measurement of |Vtb|", hep-ex/0612052 (to appear in PRL), BDT better than ANN, matrix-element likelihood
  - More are underway …

BDT Free Softwares

- http://gallatin.physics.lsa.umich.edu/~hyang/boosting.tar.gz

- TMVA toolkit, CERN Root V5.14/00
  http://tmva.sourceforge.net/
  http://root.cern.ch/root/html/src/TMVA__MethodBDT.cxx.html

Summary and Future Plan

- ➔ WW and WZ analysis results with Simple cuts and BDT are presented
- ➔ BDT works better than ANN, it is a very powerful and promising data analysis tool

- ➔ Redo WW/WZ analysis with CSC12 MC
- ➔ BDT will be applied for WW -> 2μX, H -> ZZ,WW and tautau etc.

BACKUP SLIDES
for
Boosted Decision Trees

Decision Trees & Boosting Algorithms

- ➔ Decision Trees have been available about two decades, they are known to be powerful but unstable, i.e., a small change in the training sample can give a large change in the tree and the results.

  Ref: L. Breiman, J.H. Friedman, R.A. Olshen, C.J.Stone, "Classification and Regression Trees", Wadsworth, 1984.

- ➔ The boosting algorithm (AdaBoost) is a procedure that combines many "weak" classifiers to achieve a final powerful classifier.

  Ref: Y. Freund, R.E. Schapire, "Experiments with a new boosting algorithm", Proceedings of COLT, ACM Press, New York, 1996, pp. 209-217.

- ➔ Boosting algorithms can be applied to any classification method. Here, it is applied to decision trees, so called "Boosted Decision Trees", for the MiniBooNE particle identification.

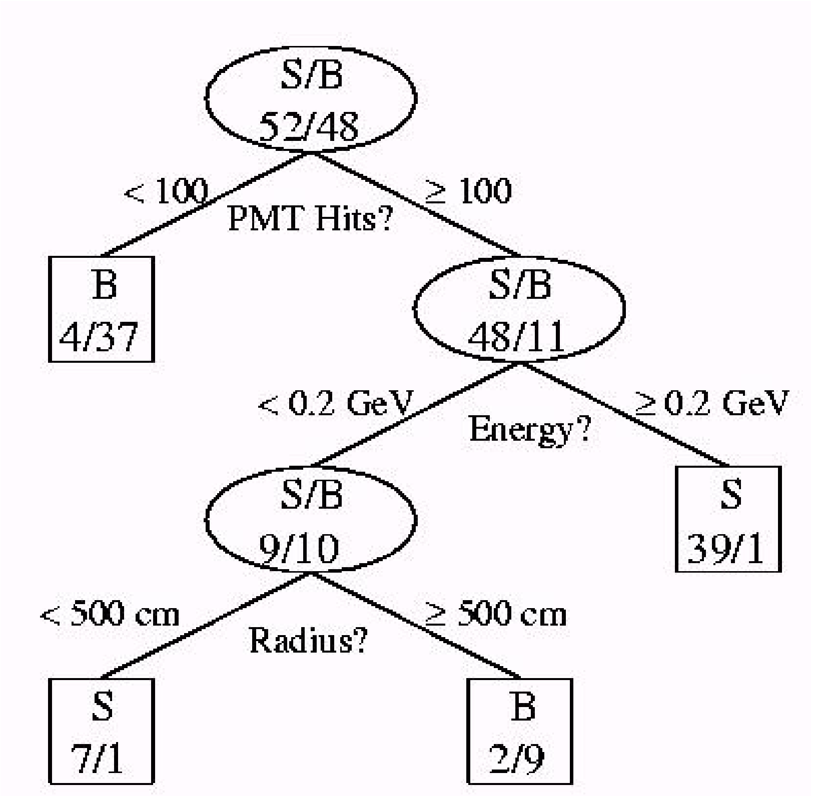
  * Hai-Jun Yang, Byron P. Roe, Ji Zhu, " Studies of boosted decision trees for MiniBooNE particle identification", physics/0508045, NIM A 555:370,2005

  * Byron P. Roe, Hai-Jun Yang, Ji Zhu, Yong Liu, Ion Stancu, Gordon McGregor," Boosted decision trees as an alternative to artificial neural networks for particle identification", NIM A 543:577,2005

  * Hai-Jun Yang, Byron P. Roe, Ji Zhu, "Studies of Stability and Robustness of Artificial Neural Networks and Boosted Decision Trees", NIM A574:342,2007

How to Build A Decision Tree ?

1. Put all training events in root node, then try to select the splitting variable and splitting value which gives the best signal/background separation.
2. Training events are split into two parts, left and right, depending on the value of the splitting variable.
3. For each sub node, try to find the best variable and splitting point which gives the best separation.
4. If there are more than 1 sub node, pick one node with the best signal/background separation for next tree splitter.
5. Keep splitting until a given number of terminal nodes (leaves) are obtained, or until each leaf is pure signal/background, or has too few events to continue.



* If signal events are dominant in one leaf, then this leaf is signal leaf (+1); otherwise, backgroud leaf (score= -1).

Criterion for “Best” Tree Split

- Purity, *P*, is the fraction of the weight of a node (leaf) due to signal events.
- Gini Index: Note that Gini index is 0 for all signal or all background.

$$Gini = (\sum_{i=1}^{n} W_i)P(1-P)$$

- The criterion is to minimize

  Gini_left_node+ Gini_right_node.

Criterion for Next Node to Split

- Pick the node to maximize the change in Gini index. Criterion =

  $$\text{Gini}_{\text{parent\_node}} - \text{Gini}_{\text{right\_child\_node}} - \text{Gini}_{\text{left\_child\_node}}$$

- We can use Gini index contribution of tree split variables to sort the importance of input variables.
- We can also sort the importance of input variables based on how often they are used as tree splitters.

Signal and Background Leaves

- Assume an equal weight of signal and background training events.
- If event weight of signal is larger than ½ of the total weight of a leaf, it is a signal leaf; otherwise it is a background leaf.
- Signal events on a background leaf or background events on a signal leaf are misclassified events.

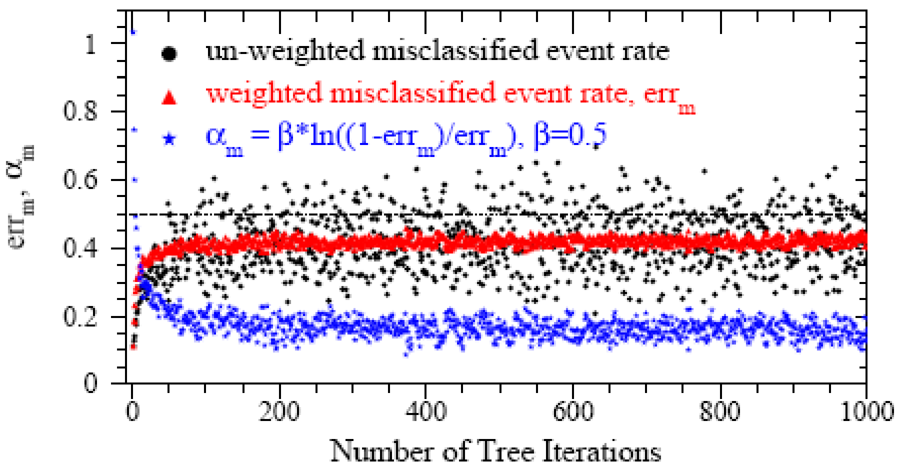
How to Boost Decision Trees ?

- For each tree iteration, same set of training events are used but the weights of misclassified events in previous iteration are increased (boosted). Events with higher weights have larger impact on Gini index values and Criterion values. The use of boosted weights for misclassified events makes them possible to be correctly classified in succeeding trees.

- Typically, one generates several hundred to thousand trees until the performance is optimal.

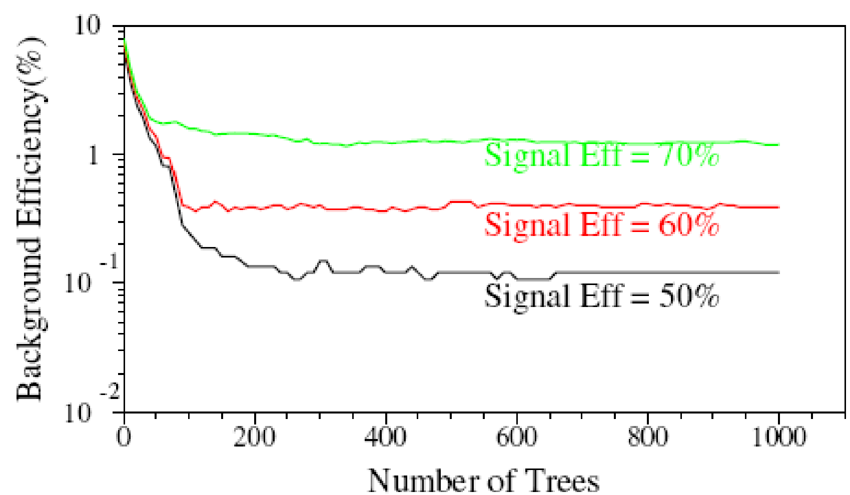
- The score of a testing event is assigned as follows: If it lands on a signal leaf, it is given a score of 1; otherwise -1. The sum of scores (weighted) from all trees is the final score of the event.

Weak → Powerful Classifier



➔ The advantage of using boosted decision trees is that it combines many decision trees, "weak" classifiers, to make a powerful classifier. The performance of BDT is stable after few hundred tree iterations.

➔ Boosted decision trees focus on the misclassified events which usually have high weights after hundreds of tree iterations. An individual tree has a very weak discriminating power; the weighted misclassified event rate err_m is about 0.4-0.45.



Two Boosting Algorithms

- AdaBoost Algorithm:

1. Initialize the observation weights $w_i = 1/\text{n}$, i = 1, 2,..., n
2. For m = 1 to M:
   2.a Fit a classifier $T_m(x)$ to the training data using weights w_i
   2.b Compute

$$err_m = \frac{\sum_{i=1}^{n} w_i I(y_i \neq T_m(x_i))}{\sum_{i=1}^{n} w_i}$$

I = 1, if a training event is misclassified; Otherwise, I = 0

   2.c Compute $\alpha_m = \beta \times log((1 - err_m)/err_m)$
   2.d Set $w_i \leftarrow w_i \times exp(\alpha_m I(y_i \neq T_m(x_i)))$, i=1, 2,...,n
   2.e Re-normalize $w_i = w_i / \sum_{i=1}^{n} w_i$
3. Output T(x) = $\sum_{m=1}^{M} \alpha_m T_m(x)$

- $\epsilon-$boosting Algorithm:

1. Initialize the observation weights $w_i = 1/\text{n}$, i = 1, 2,..., n
2. For m = 1 to M:
   2.a Fit a classifier $T_m(x)$ to the training data using weights w_i
   2.b Set $w_i \leftarrow w_i \times exp(2\epsilon I(y_i \neq T_m(x_i)))$, i=1, 2,...,n
   2.c Re-normalize $w_i = w_i / \sum_{i=1}^{n} w_i$
3. Output T(x) = $\sum_{m=1}^{M} \epsilon T_m(x)$

Example

- **AdaBoost: the weight of misclassified events is increased by**
  - error rate=0.1 and $\beta = 0.5$, $\alpha_m = 1.1$, $\exp(1.1) = 3$
  - error rate=0.4 and $\beta = 0.5$, $\alpha_m = 0.203$, $\exp(0.203) = 1.225$
  - Weight of a misclassified event is multiplied by a large factor which depends on the error rate.
- **e-boost: the weight of misclassified events is increased by**
  - If **ε = 0.01,** $\exp(2*0.01) = 1.02$
  - If **ε = 0.04,** $\exp(2*0.04) = 1.083$
  - It changes event weight a little at a time.

➔ AdaBoost converges faster than ε-boost. However, the performance of AdaBoost and ε-boost are comparable with sufficient tree iterations.