Precision measurements of the Top Higgs Yukawa Coupling1 TeV - Semileptonic

T. Price

International Large Detector

txp@hep.ph.bham.ac.uk

January 25, 2013

Overview

Input Variables I

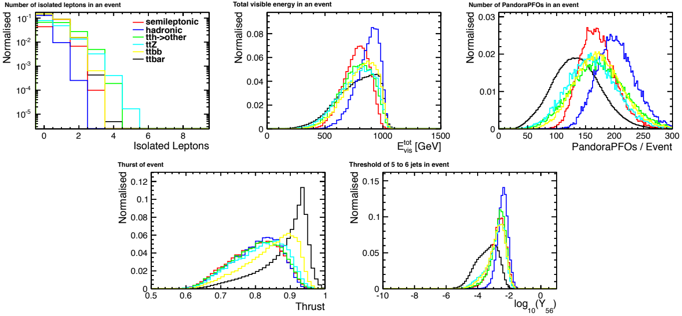


Figure: Normalised histograms of the number of identified isolated leptons (top left), total visible energy (top middle), number of PandoraPFOs (top right), thrust (bottom left), and the jet parameters (bottom right) within the events for the semileptonic (red) and hadronic (blue) modes alongside the backgrounds.

Input Variables II

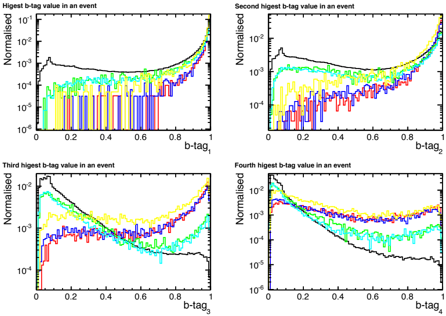


Figure: Normalised histograms for the response of the b-tagging from LCFIPlus for the highest ranked jet in an event (top left), second highest (top right), third highest (bottom left), and fourth highest (bottom right) for the semileptonic (red) and hadronic (blue) signal modes and all backgrounds.

Input Variables III

Reconstructed Masses

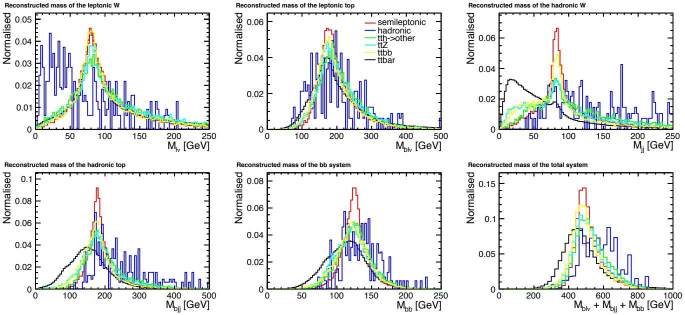


Figure: The reconstructed masses of the leptonic W boson (top left), leptonic top quark (top middle), hadronic W boson (top right), hadronic top quark (bottom left), Higgs boson (bottom middle) and total mass (bottom right) when there is exactly one identified isolated lepton in the event.

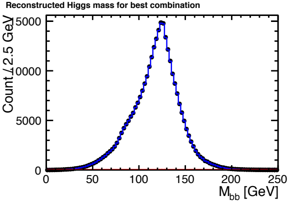
Cut Based Analysis I

Cut Flow

Cut	lept	semi	had	other	ttZ	ttbb	tt	$\frac{S}{\sqrt{S+B}}$
Total Events	151.4	628.7	652.7	1046.1	5332.4	1434.5	308800.9	1.11
N_{isolep}=1	74.6	363.5	5.0	371.8	1581.5	439.9	101295.2	1.13
$610 < E_{vis} < 1000$	49.6	338.5	4.7	312.7	1228.9	373.8	75507.1	1.21
nPFOs$>$154	15.0	235.0	4.1	195.0	589.0	194.5	12605.9	2.00
Thrust$<$0.88	12.5	205.6	3.7	168.9	492.6	140.0	6092.3	2.44
$\log_{10}(Y_{45}) > -2.25$	7.7	151.3	3.2	108.5	295.2	91.0	2067.2	2.90
$\log_{10}(Y_{56}) > -3.35$	6.9	145.1	3.2	106.2	277.6	86.0	1836.1	2.92
b-$\text{tag}_1 > 0.96$	6.7	135.1	2.8	79.8	216.9	78.3	1367.8	3.11
b-$\text{tag}_2 > 0.91$	6.1	118.2	2.3	41.2	135.1	66.9	715.2	3.59
b-$\text{tag}_3 > 0.67$	5.5	102.1	1.6	5.7	59.4	56.0	137.2	5.33
b-$\text{tag}_4 > 0.04$	5.5	100.5	1.5	5.4	58.3	54.8	128.7	5.34
$\chi^2 < 450$	5.3	100.0	1.4	5.2	56.8	53.7	126.0	5.36
$364 < M_{Total} < 808$	5.2	99.7	1.4	5.2	56.5	53.5	124.7	5.36
$98 < M_H < 234$	4.6	95.1	1.2	4.7	46.1	48.6	109.9	5.40

Table: The number of events passed each cut when the cut values are optimised to select the semileptonic signal with maximum significance. The ttH$\rightarrow$other is the background where the Higgs boson does not decay to a bb pair.

Cut Based Analysis II



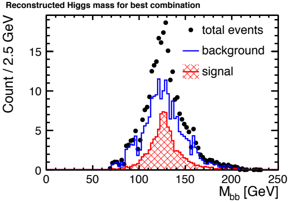


Figure: The reconstructed Higgs boson mass for the optimal combination of jets in the semileptonic decay mode for all events (left) and only the events which pass all of the cuts (right).

$$\epsilon_{sig}=15.1\%,\ \rho = 30.9\%$$

Final significance = 5.40

$$\frac{\Delta g_{ttH}}{g_{ttH}} = 9.6\% \text{ (stat)}$$

TMVA

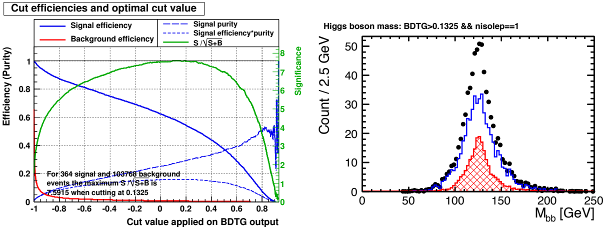


Figure: The response of the multivariate analysis for the semileptonic (left) and events passed mva cuts (right).

$$\text{Final significance} = 7.59$$

$$\frac{\Delta g_{ttH}}{g_{ttH}} = 6.9\% \text{ (stat)}$$

Conclusions

Results

Semileptonic results: 6.9% (9.6%) [TMVA (Cut based)]
Hadronic results: 5.4% (7.2%) [TMVA (Cut based)] T. Tanabe
Combined results: 4.3% for TMVA analysis

To Do

- Update results in note and send to reviewers again
- Update results in DBD
- Combine results with SiD and publish
- Consider systematics

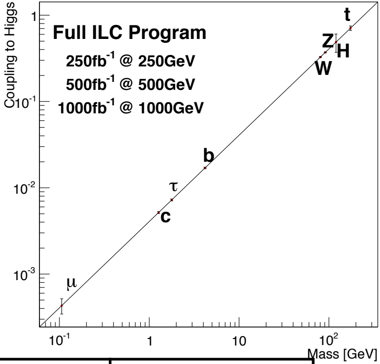
Top Yukawa Coupling at 1 TeV ILC
hadronic "8 jet" channel

Tomohiko Tanabe

January 24, 2013

Introduction

- Top quark couples strongly to the Higgs boson (top Yukawa y_t~1)
- Important probe for verification of electroweak symmetry breaking
- Many BSM models predict **large deviations in y_t** e.g. composite Higgs models

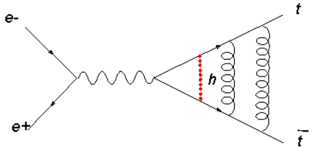


Top Yukawa coupling precision studies at LC

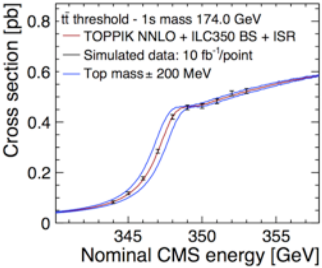
CM Energy	Simulation	Luminosity	$\Delta y_t/y_t$
500 GeV **[Phys. Rev. D 84, 014033 (2011)]**	Fast Sim	1 ab^{-1}	10%
500 GeV **[R. Yonamine, Ph.D. thesis]**	**Full Sim**	1 ab^{-1}	10%
700-800 GeV **[A. Gay, Eur.Phys.J.C49, 489 (2007)]**	Fast Sim	(various Higgs masses)	
1 TeV **[ILC TDR, 2012]**	**Full Sim**	0.5 ab^{-1} each for 2 polarizations	< 5%

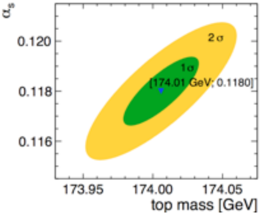
Indirect vs. direct

- **Indirect measurement** at the ttbar threshold
  - depends on theory calculation → improving
  - can be also inferred from LHC Higgs production

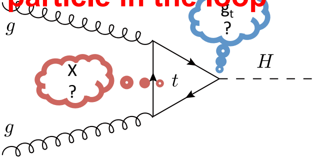


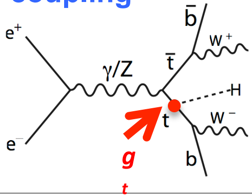
F.Simon Top Phys WS 2012





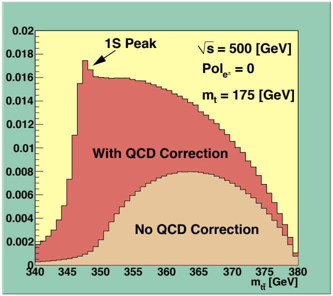
- **Direct measurement** can distinguish between **coupling anomaly** and **new particle in the loop**

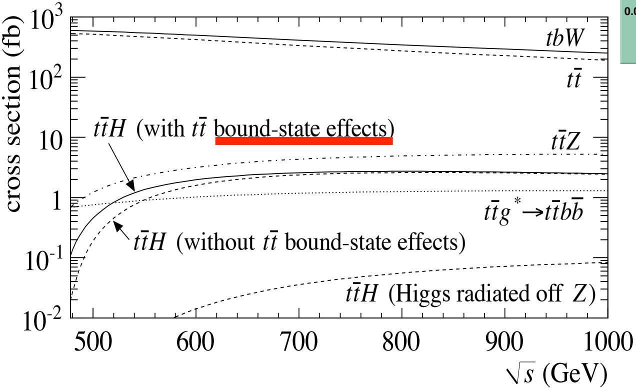


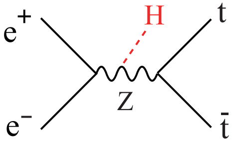


Signal and Background

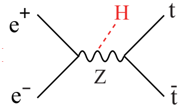
- **Signal: tth** → bWbWbb
- Irreducible backgrounds:
  - **ttZ** → bWbWbb
  - **ttg*** → bWbWbb
- Reducible background: bWbW, e.g. **tt**





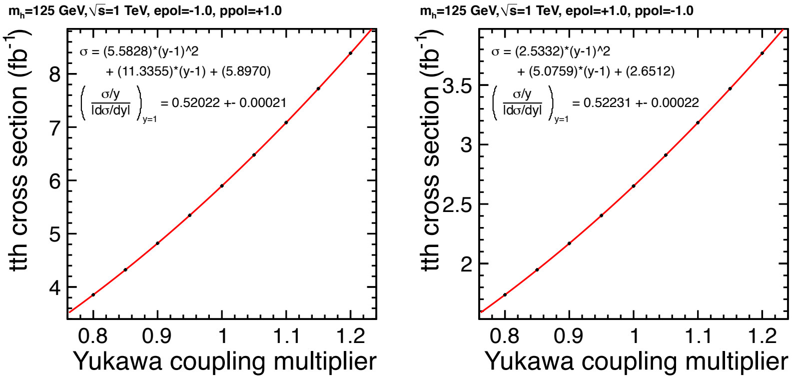


Sensitivity to top Yukawa



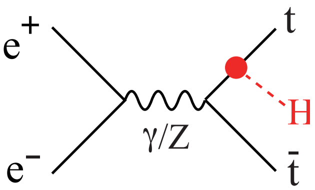
Estimate effect of non-contributing diagrams.
Dependence of cross section w.r.t. scaling of top Yukawa coupling → slope at SM value gives nominal sensitivity

$$\frac{\Delta y_t}{y_t} = (..) \frac{\Delta\sigma}{\sigma} \qquad \left(\frac{\sigma/y_t}{|d\sigma/dy_t|}\right)_{y_t=y_t(SM)} = 0.52$$



Signal mode

- Goal is to evaluate the precision of the top Yukawa coupling at $\sqrt{s}$ = 1 TeV
  - = evaluate the precision of cross section measurement
- **Higgs boson mass set to 125 GeV in light of LHC data.**
  - BR(H→bb) = 57.8%
- There are three decay modes depending on the W decay:
  - ttH → 4 jet + 2 lepton mode: BR(tt → blνblν) = 11% -- not analyzed
  - **ttH → 6 jet + lepton mode:** BR(tt → bqqblν) = 45% for l=e,μ,τ (29% for l=e,μ)
  - **ttH → 8 jet mode:** BR(tt → bqqbqq) = 44%



Data Samples

id	process	pol	xsec	ngen	weight
106427	Pttbb-all-all	eL.pR	3.429300	21000	0.047357
106428	Pttbb-all-all	eR.pL	1.517400	10600	0.041514
106429	Pttz-all-all	eL.pR	14.020600	13829	0.294018
106430	Pttz-all-all	eR.pL	4.367100	13200	0.095944
106451	Ptth-6q-hbb	eL.pR	1.552750	17620	0.025556
106452	Ptth-6q-hbb	eR.pL	0.698000	7361	0.027499
106453	Ptth-6q-hnonbb	eL.pR	1.133670	7749	0.042427
106454	Ptth-6q-hnonbb	eR.pL	0.509620	3787	0.039026
106455	Ptth-ln4q-hbb	eL.pR	1.495560	17603	0.024639
106456	Ptth-ln4q-hbb	eR.pL	0.672430	7311	0.026673
106457	Ptth-ln4q-hnonbb	eL.pR	1.091920	6684	0.047375
106458	Ptth-ln4q-hnonbb	eR.pL	0.490940	3358	0.042398
106459	Ptth-2l2nbb-hbb	eL.pR	0.360100	800	0.130536
106460	Ptth-2l2nbb-hbb	eR.pL	0.161940	400	0.117407
106461	Ptth-2l2nbb-hnonbb	eL.pR	0.262910	600	0.127073
106462	Ptth-2l2nbb-hnonbb	eR.pL	0.118230	400	0.085717
35786	P6f_yyveev	eL.pL	0.753694	10000	0.015828
35787	P6f_yyveev	eL.pR	14.262567	14263	0.289991
35788	P6f_yyveev	eR.pL	3.191048	10000	0.092540
35789	P6f_yyveev	eR.pR	0.759213	9999	0.015945
35790	P6f_yyvelv	eL.pL	1.434391	10000	0.030122
35791	P6f_yyvelv	eL.pR	22.876428	22873	0.290043
35792	P6f_yyvelv	eR.pL	6.272190	10000	0.181894
35794	P6f_yyveyx	eL.pL	4.121621	9999	0.086563
35795	P6f_yyveyx	eL.pR	67.534318	400000	0.048962
35796	P6f_yyveyx	eR.pL	18.645337	40000	0.135179
35799	P6f_yyvlev	eL.pR	22.875149	22871	0.290053
35800	P6f_yyvlev	eR.pL	6.264408	9998	0.181704
35801	P6f_yyvlev	eR.pR	1.427611	10000	0.029980
35803	P6f_yyvllv	eL.pR	41.275472	41270	0.290038
35804	P6f_yyvllv	eR.pL	12.598244	12597	0.290029
35807	P6f_yyvlyx	eL.pR	115.979040	698099	0.048179
35808	P6f_yyvlyx	eR.pL	37.306473	60000	0.180315
35811	P6f_yyxyev	eL.pR	68.502191	400000	0.049664
35812	P6f_yyxyev	eR.pL	18.659270	40000	0.135280
35813	P6f_yyxyev	eR.pR	4.163067	10000	0.087424
35815	P6f_yyxylv	eL.pR	116.426720	699144	0.048293
35816	P6f_yyxylv	eR.pL	37.321082	60000	0.180385
35819	P6f_yyuyyu	eL.pR	84.595962	500000	0.049066
35820	P6f_yyuyyu	eR.pL	27.500471	40000	0.199378
35823	P6f_yyuyyc	eL.pR	84.581774	498800	0.049175
35824	P6f_yyuyyc	eR.pL	27.508546	40000	0.199437
35827	P6f_yycyyu	eL.pR	84.426452	500000	0.048967
35828	P6f_yycyyu	eR.pL	27.483992	40000	0.199259
35831	P6f_yycyyc	eL.pR	84.975908	500000	0.049286
35832	P6f_yycyyc	eR.pL	27.584594	40000	0.199988

Signal samples

Weights are calculated assuming data samples of:

- 0.5 ab^{-1} for (−0.8, +0.2)
- 0.5 ab^{-1} for (+0.8, −0.2)

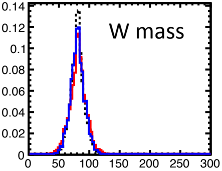
which are summed.

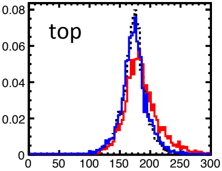
All weights << 1:
→ We have sufficient statistics.

Analysis Flow

- Start with standard reconstruction samples for DBD
- Removal of γγ→hadrons pileup background
- Isolated lepton selection
- Event selection based on
  - b-tagging, jet combination mass, etc
- Comparison of two analyses:
  - Cut-based
  - TMVA-based with Boosted Decision Trees

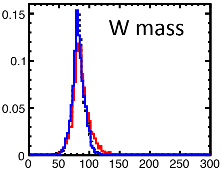
Removal of γγ→hadrons

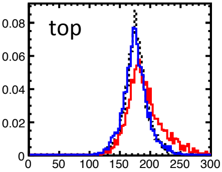




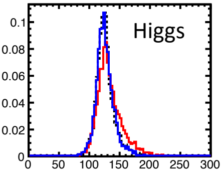
Average 4.1 events γγ→hadrons are overlaid in all simulations.

→ Degrade the mass resolution due to extra energy in the forward region.





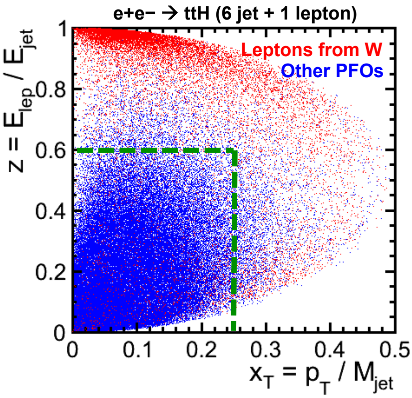
Black (dotted):
Durham (sample w/o γγ→hadrons)
Red:
Durham (sample w/ γγ→hadrons)
Blue:
Durham (sample w/ γγ→hadrons)
after removing cosθ>0.94 particles



→ **Mass resolution can be recovered.**
(PFOs are mostly central for ttH process.)

kt algorithm with R=1.2 is actually used for consistency with 6 jet + lepton analysis (studied by Tony Price)

Isolated Lepton Finding



Hard isolated leptons coming from W decay

- Useful for separating 6 jet + lepton / 8 jet / background

Selection criteria based on:

- **Lepton ID with calorimeter energies** which reduces fake leptons
- **Impact parameter significance** for reducing contamination from bottom and tau decays
- **Jet-based isolation** ("LAL Lepton Finder")
  - isolated lepton in jets tends to be "leading" or have "large p_T w.r.t jet axis"

	Efficiency	Composition			
		W→e,μ	W→τ→e,μ	Other e,μ	Fake e,μ
Electrons	84.0%	94.2%	2.9%	1.6%	2.3%
Muons	90.5%	96.3%	2.4%	1.2%	0.7%

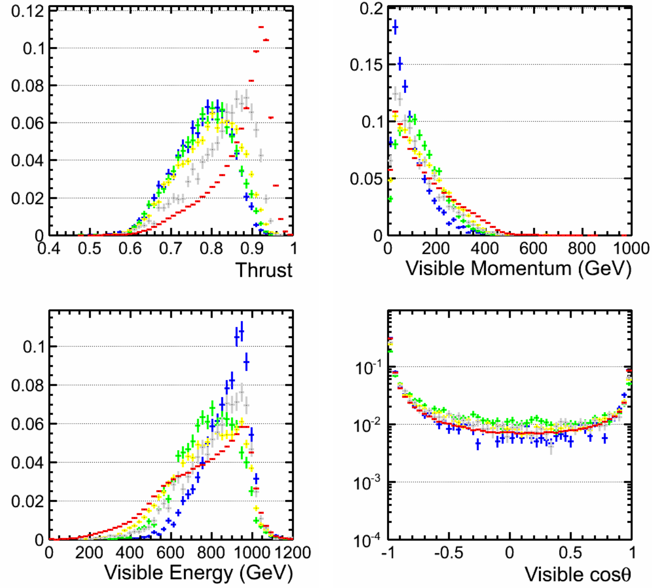
Performance of isolated lepton finder for tth 6 jets + 1 lepton sample

Variables (1)

ttH 8 jet
ttH other
ttZ
ttbb
tt

$$T = \max_{|\hat{n}|=1} \frac{\sum_i |\hat{n} \cdot \vec{p}_i|}{\sum_i |\vec{p}_i|}$$

Thrust definition
- dijet-like → 1
- Isotropic → 0

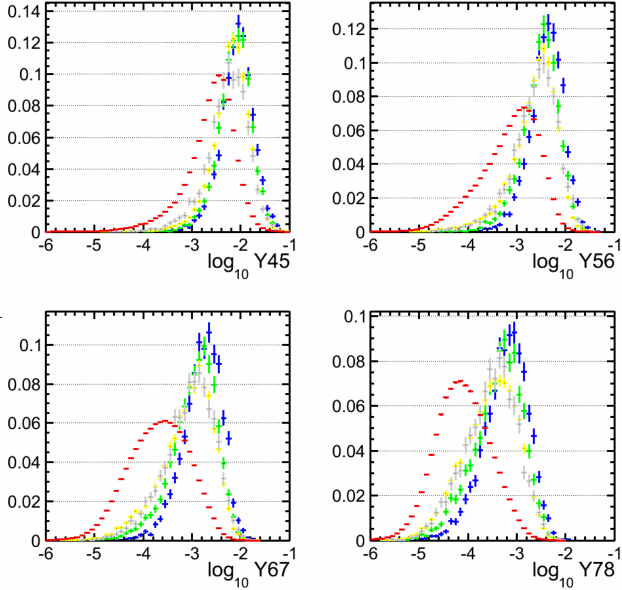


Variables (2)

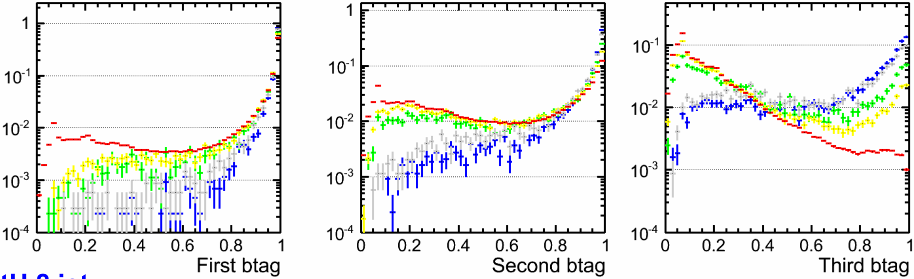
ttH 8 jet
ttH other
ttZ
ttbb
tt

Jet Finder "Y" variables

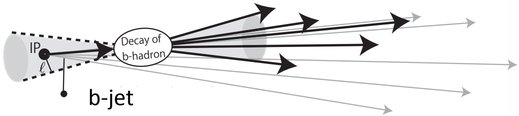
$$Y_{ij} = \frac{\max(E_i^2, E_j^2)(1 - \cos\theta_{ij})}{E_{\mathrm{CM}}^2}$$

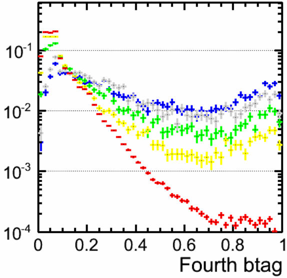


Variables (3)



ttH 8 jet
ttH other
ttZ
ttbb
tt

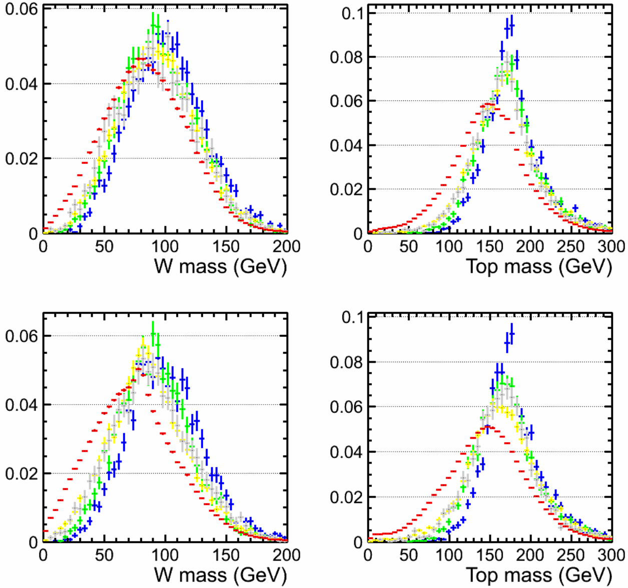




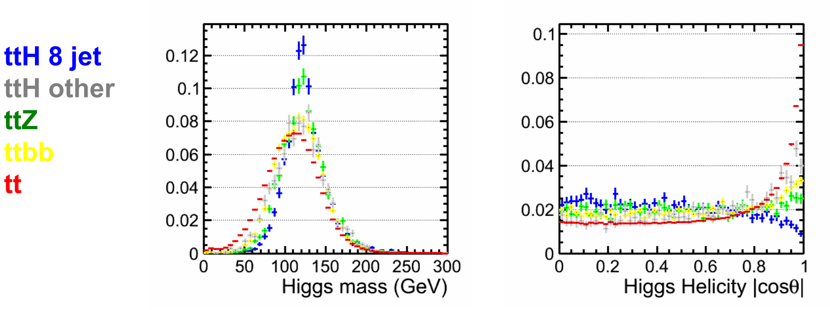
b-jet tagging: displaced tracks & secondary vertices
Multivariate analysis of variables (LCFIPlus)

Variables (4)

ttH 8 jet
ttH other
ttZ
ttbb
tt



Variables (5)



$$\chi^2 = \frac{(M_{t_1} - M_t)^2}{\sigma_{t_1}^2} + \frac{(M_{t_2} - M_t)^2}{\sigma_{t_2}^2} + \frac{(M_{b\bar{b}} - M_H)^2}{\sigma_H^2}$$

Jet combination is based on chi-squared minimization method.
Jets with 4 lowest b-tags selected for W candidates.

Tomohiko Tanabe (tomohiko@icepp.s.u-tokyo.ac.jp)

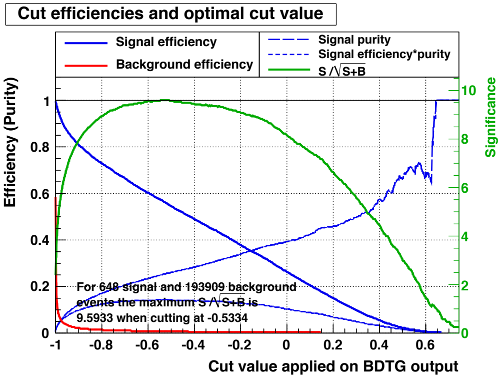
Cut-based Analysis

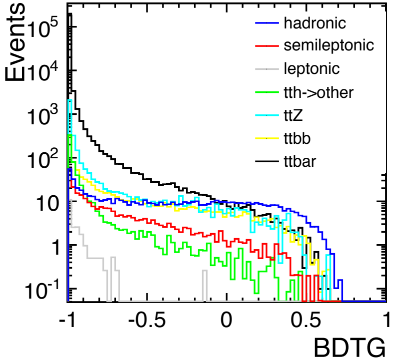
	$t\bar{t}h$ (4J)	$t\bar{t}h$ (6J)	$t\bar{t}h$ (8J)	$t\bar{t}h$ ($h \not\to b\bar{b}$)	$t\bar{t}Z$	$t\bar{t}b\bar{b}$	$t\bar{t}$	Sig.
No cuts	151.39	628.73	652.77	1046.10	5332.52	1434.53	306238.26	1.16
$N_{\mathrm{iso}} = 0$	20.87	261.17	647.92	556.71	3226.14	932.49	188911.38	1.47
$E_{\mathrm{vis}} > 650$ GeV	9.83	220.97	636.16	497.45	2743.54	849.34	157389.56	1.58
Thrust< 0.87	8.09	187.75	577.60	440.06	2219.68	540.88	46916.14	2.56
$Y_{78} > 0.0001$	3.65	143.55	549.52	415.51	1926.58	474.59	27472.09	3.12
$btag_4 > 0.38$	1.89	80.98	275.02	17.55	230.04	209.60	680.62	7.11
$\lvert\cos\theta_{\mathrm{hel}}\rvert < 0.9$	1.63	73.80	263.71	16.48	215.91	189.19	584.92	7.19
$m_t > 120$ GeV	1.50	68.09	255.38	15.58	207.81	178.53	530.93	7.20

Cut-based analysis result (8 jet mode only):
Statistical significance = 7.2 sigma

Multivariate Analysis

- Use TMVA Boosted Decision Trees with Gradient boost
- 18 variables (shown earlier) were used



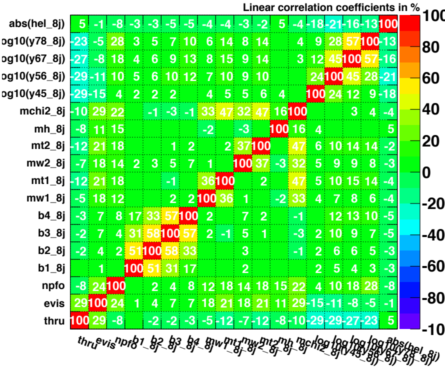


TMVA analysis result (8 jet mode only):
Statistical significance = 9.6 sigma

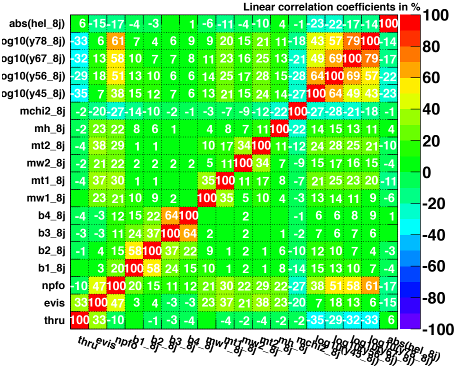
Error in normalization of the events was found and fixed.

Correlation Matrices

Correlation Matrix (signal)



Correlation Matrix (background)



Difference between cut-based and TMVA-based analysis will be studied.
<u>The input variables are correlated.</u>

Systematic Uncertainties

So far our results give the statistical precision only.

For **O(1)%** measurements, need to address **systematic uncertainties** such as:

- Background normalization
- Jet energy scale
- Luminosity spectrum
- B-tagging efficiency
- Lepton isolation criteria
- Lepton ID performance
- …

→ *These are the next steps…*

Conclusions

- Hadronic ("8 jet") analysis:
  - Cut-based analysis: 7.2 sigma
  - TMVA-based analysis: 9.6 sigma
- Semileptonic ("6 jet") analysis (**Tony Price**):
  - Cut-based analysis: 5.4 sigma
  - TMVA-based analysis: 7.6 sigma
- **Combined: 12.2 sigma, 4.3% precision in $\Delta y_t/y_t$ (TMVA)**
  - for 0.5 ab^{-1} (-0.8, +0.2) and 0.5 ab^{-1} (+0.8, -0.2)

- To-do:
  - Include final results into DBD
  - Combine ILD + SiD works into publication
  - Consideration of systematic uncertainties