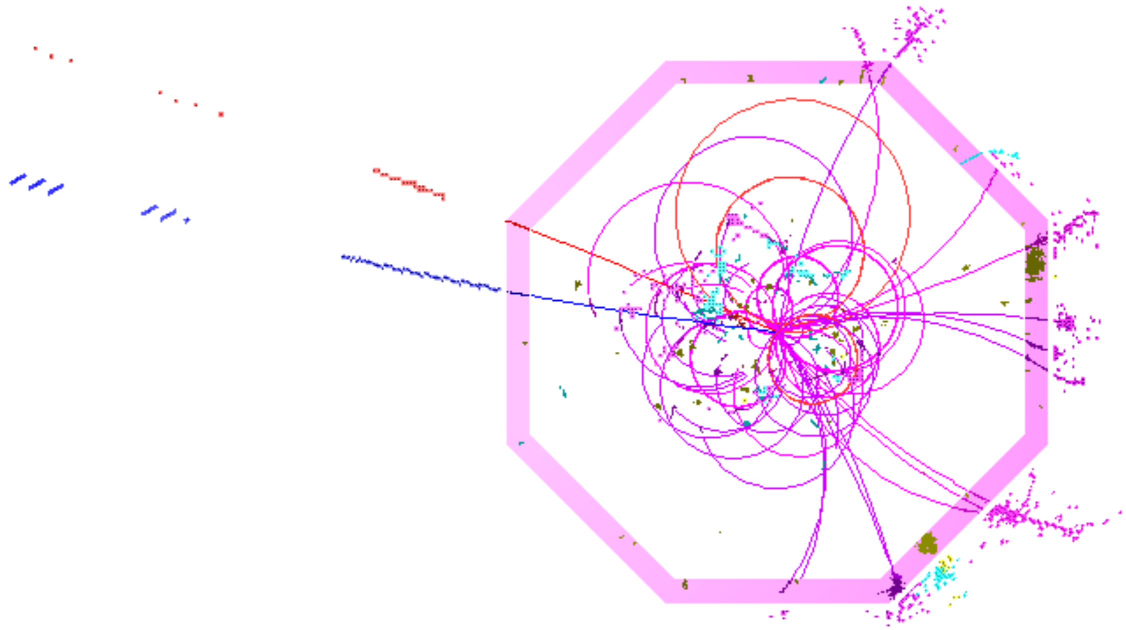


Higgs Mass and Cross-Section Measurements at CLIC

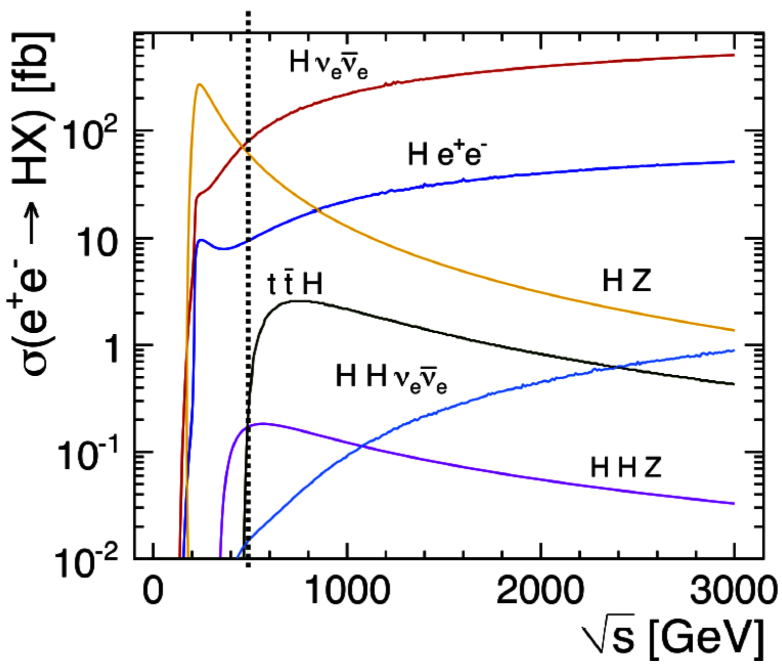
LCWS12, Arlington

J. S. Marshall, M. A. Thomson
University of Cambridge

Introduction

Aims

- Investigate a number of Higgs measurements in the first stage of possible CLIC operation
- Focus on 350 GeV and 500 GeV
  - A) Recoil mass analysis at 350 GeV
  - B) HZ → bbqq analysis at 500 GeV
  - C) Hvv → bbvv at analysis at 500 GeV
- All based on full simulation of CLIC environment/detector concept



Assumptions

$\sqrt{s}$ = 350, 500GeV
L_{int} = 500fb^{-1}
M_H = 120GeV
No polarization

Detector:
CLIC_ILD_CDR500

A) 350 GeV HZ Recoil Analysis

- The model-independent recoil analyses of the Higgsstrahlung process.
- Reconstruct the Z from its decay products, then infer the Higgs four-vector by subtracting the Z four-vector from the initial state four-vector.
- Z can be reconstructed cleanly in $Z \rightarrow \mu\mu$ ($\mu\mu X$) and $Z \rightarrow ee$ (eeX) channels. Precision depends on lepton momentum resolution and effects of beamstrahlung/ISR.

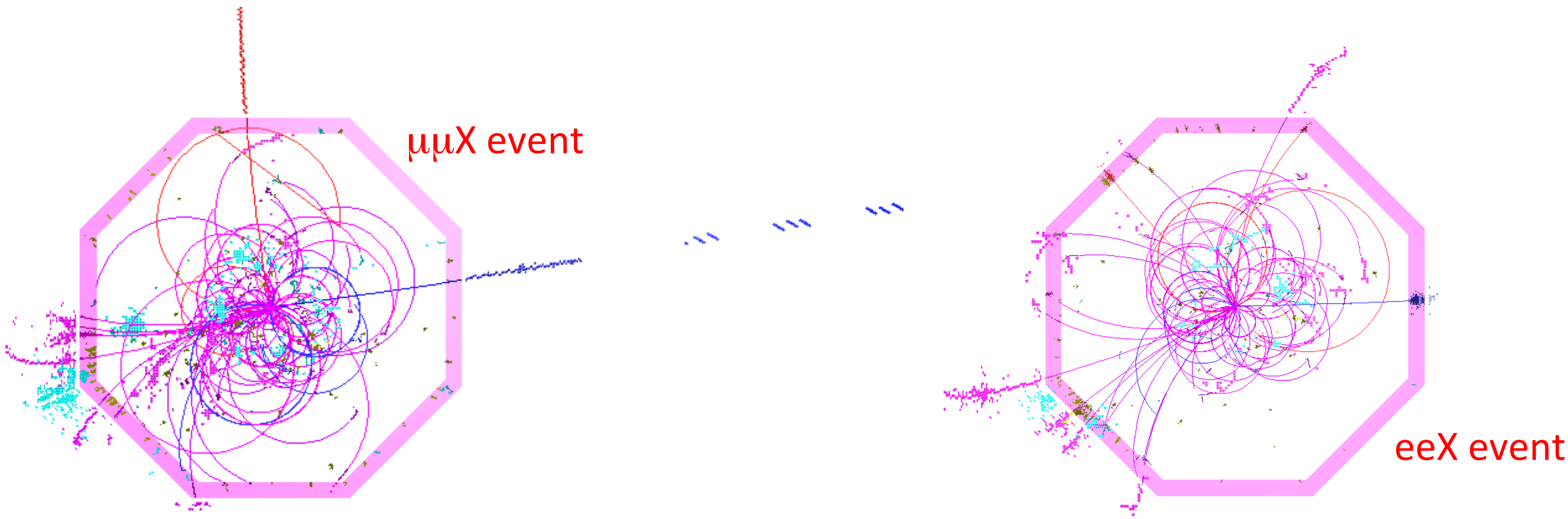
- Potential backgrounds are any processes producing a lepton pair in final state.
- Two-fermion backgrounds proved simple to remove and so neglected here.
- Four-fermion backgrounds much more difficult.

WHIZARD Process Id.	Cross-section / fb	Cross-section (gen. cuts) / fb	Events / 500fb^{-1}	Available events
hzmumu (signal)	4.855	4.855	2 427	132 867
e2e2ff	4 753	913.4	456 700	104 790
hzee (signal)	4.850	4.850	2 425	128 871
e1e1ff	4 847	1 608	804 000	134 730

Generator-level cuts for background samples e2e2ff and e1e1ff
$p_{T\,l+l-}$ > 10GeV, cut on transverse momentum, calculated from vector sum of two leptons
$\lvert\cos\theta_{l+/l-}\rvert < 0.95$, cut on angle of either of leptons

Di-lepton Selection

- Identification of charged leptons is performed by PandoraPFA; allows for a rather simple signal-selection procedure at the analysis stage:



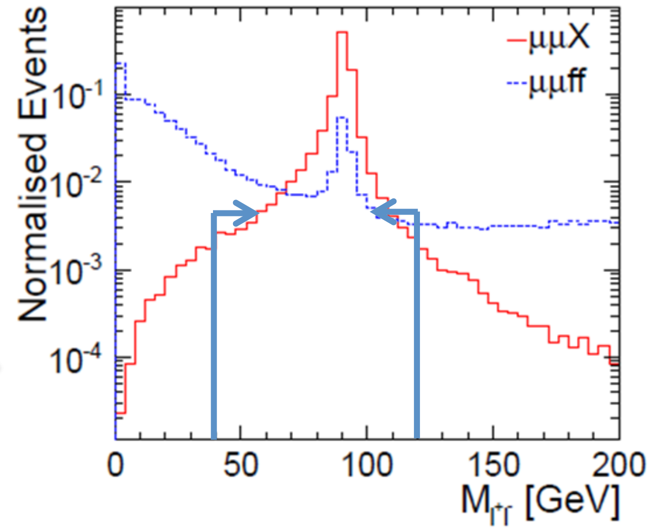
1. Loop over reconstructed particles.
2. Populate separate lists of negatively and positively charged leptons (of specified flavour).
3. If both lists are populated, event will be selected as a signal candidate.
4. If either list contains more than one entry, must investigate all possible pairings.
5. Select lepton pair producing invariant mass closest to the Z mass.

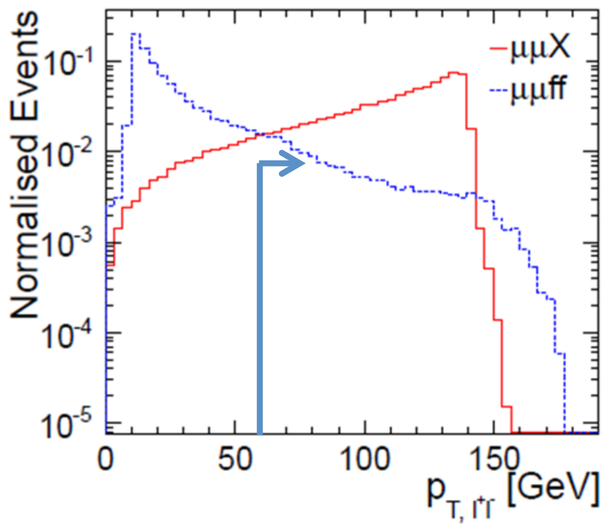
Background Rejection

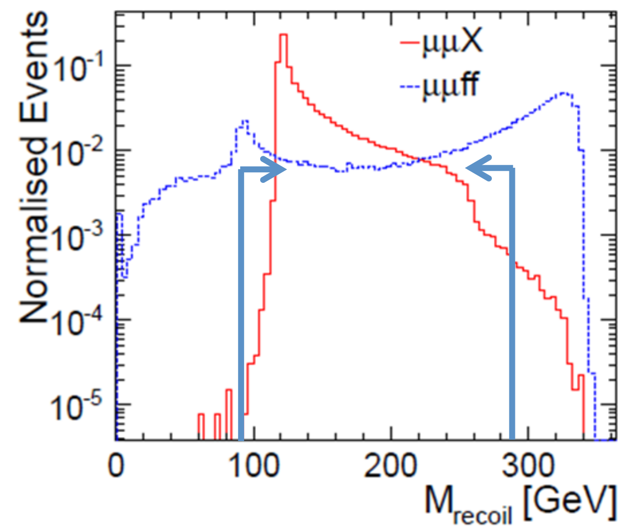
- Background rejection begins with the di-lepton selection:

WHIZARD Process Id.	*hzmumu*	*e2e2ff*	*hzee*	*e1e1ff*
Efficiency	98 %	80 %	97 %	73 %

- For both μμX and eeX channels, make the following cuts:
  - $40\,\mathrm{GeV} < M_{\ell^+\ell^-} < 120\,\mathrm{GeV}$
  - $95\,\mathrm{GeV} < M_{\mathrm{recoil}} < 290\,\mathrm{GeV}$
  - $p_{T,\ell^+\ell^-} > 60\,\mathrm{GeV}$

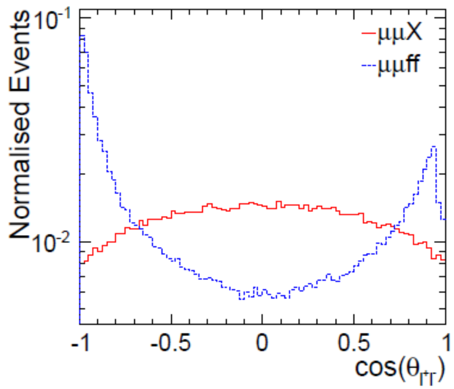




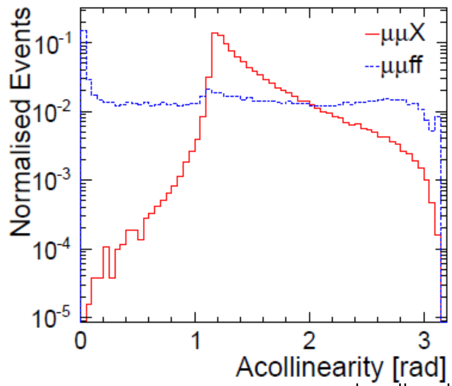


- The di-lepton mass and p_T distributions are then used as two of the six inputs to a multivariate analysis, which completes selection.

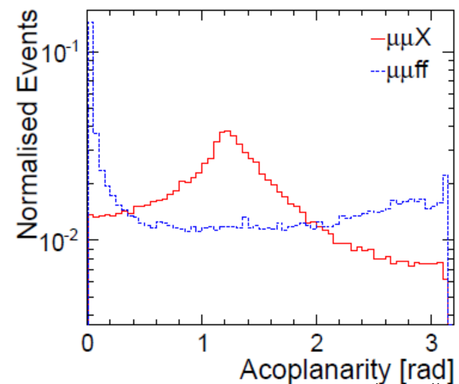
Boosted Decision Tree



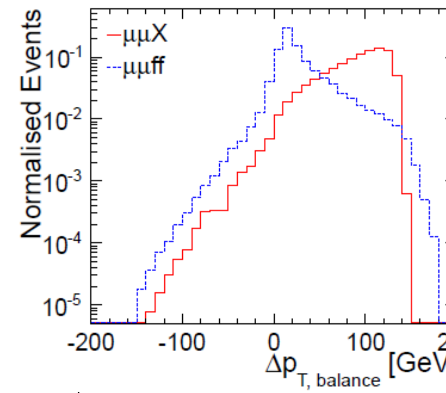
TMVA inputs:



$\mathrm{acol} = \cos^{-1}\left(\underline{p}_1.\underline{p}_2 / |\underline{p}_1||\underline{p}_2|\right)$

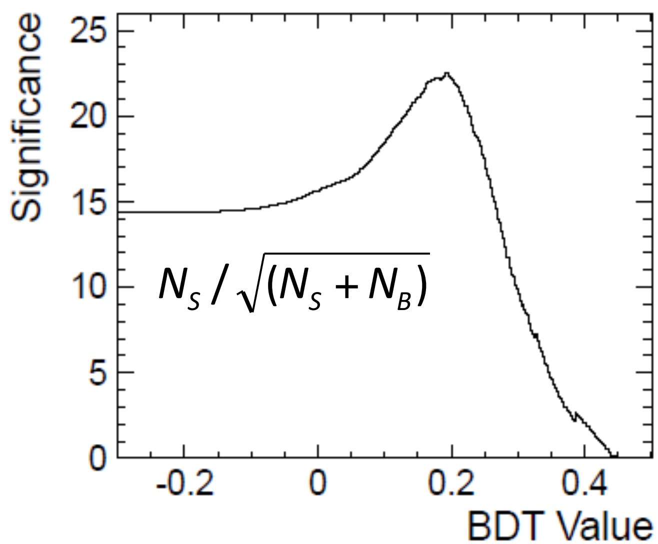


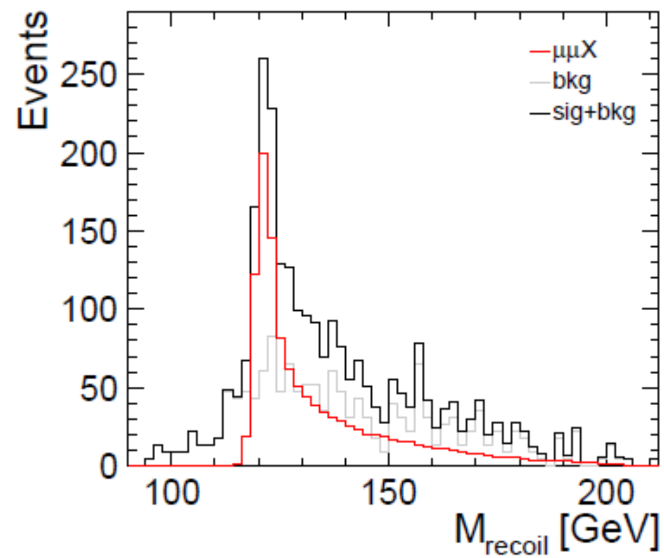
$\mathrm{acop} = \cos^{-1}\left(\underline{p}_{T1}.\underline{p}_{T2} / |\underline{p}_{T1}||\underline{p}_{T2}|\right)$



$\Delta p_{T,balance} = p_{Tl^+l^-} - p_{T\gamma,FSR}$

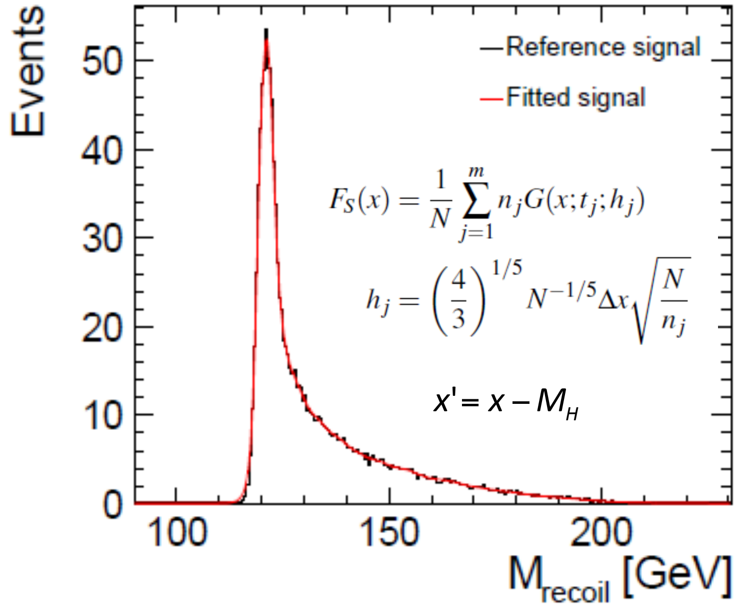
- Multivariate analysis performed by TMVA.
- BDT found to provide best signal efficiency/purity.
- Final BDT cut chosen to maximise signal significance.
- Recoil mass distribution for selected μμX events is shown:

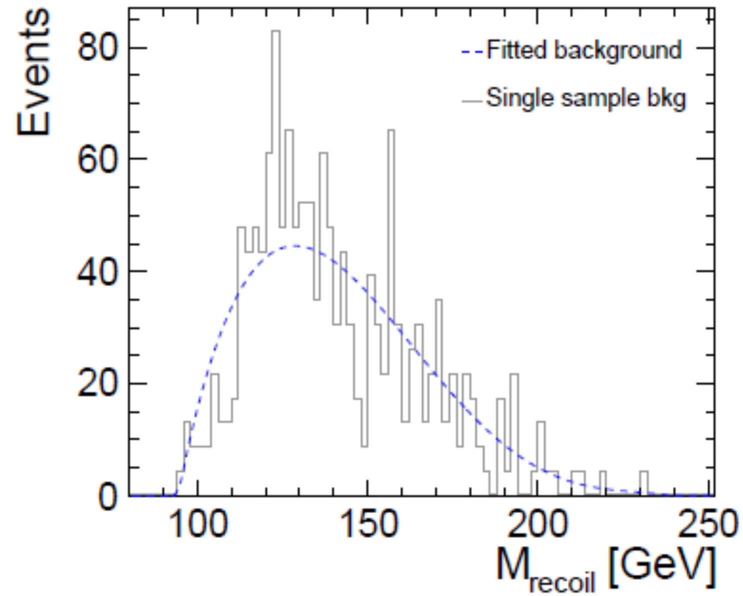




Fit Procedure

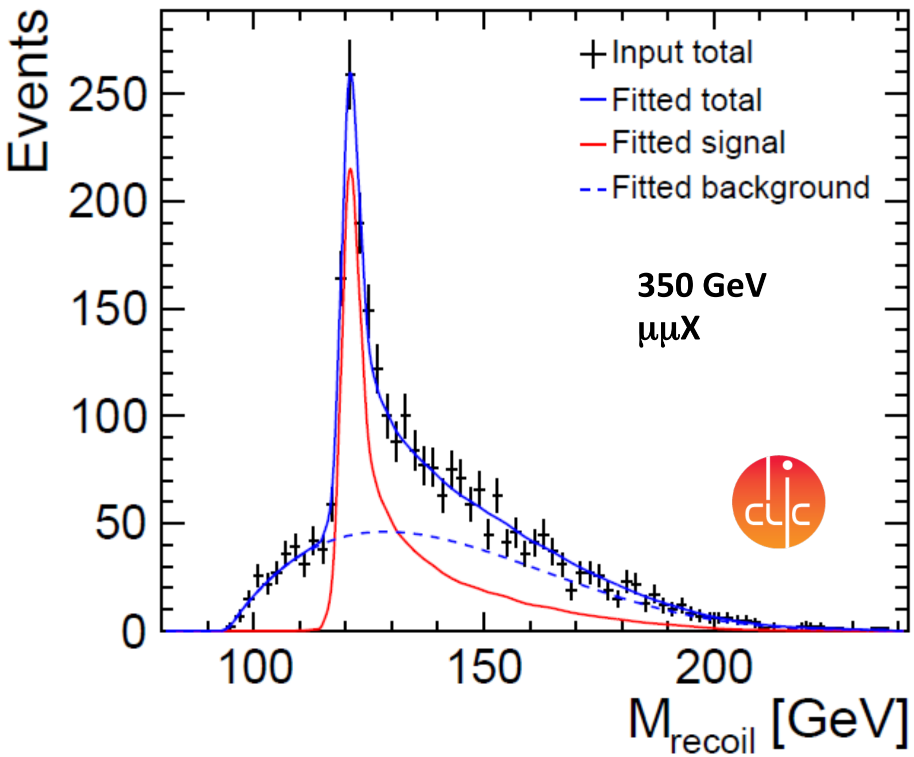
- Procedure was to develop models for the selected signal and background recoil mass distributions, allowing predicted distributions to be created for values of fit parameters: $\mathbf{M_H}$, $\mathbf{N_{Sig}}$ and $\mathbf{N_{Bkg}}$
- Compare predictions with data and calculate negative log likelihood value; MINUIT controls parameter variation so as to identify best-fit values. $-\ln\mathscr{L} = \sum_{j=1}^{n_{bins}} n_{pred,j} - n_{obs,j}\ln(n_{pred,j})$





- Use Simplified Kernel Estimation to approximate signal shape by sum of many Gaussians. Possible because have high-statistics signal samples; avoids difficulty of finding function to describe signal.
- Fit low-statistics background sample with 4th order polynomial. Tests show this is a robust strategy.

Results: μμX Channel



- To assess measurement accuracy, create 1000 representative 500fb^{-1} test “data” samples.
- Add high-statistics selected signal sample to smooth background function, then fluctuate.
- Look at distribution of best-fit values

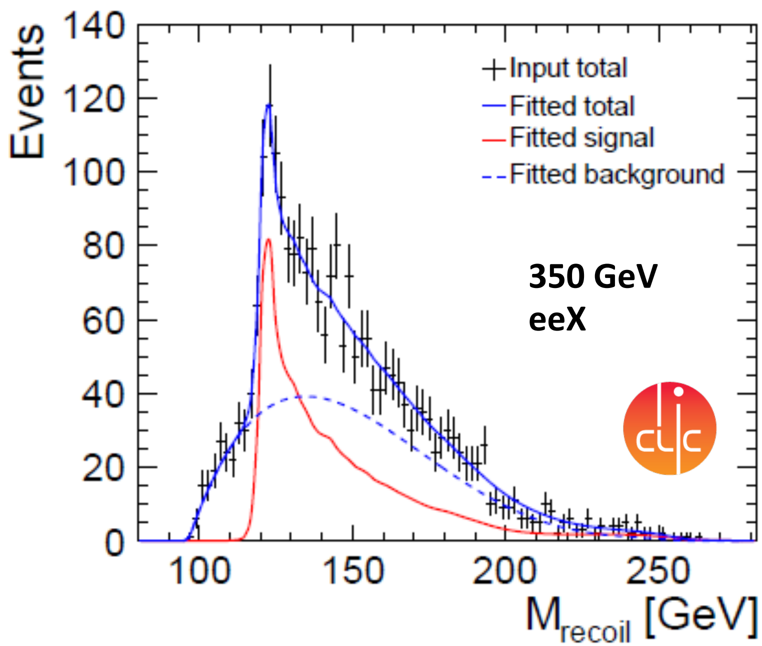
350 GeV Recoil		Mean	RMS
$\mu\mu X$	M_H	119 950.4 MeV	130.4 MeV
	ΔM_H	133.3 MeV	6.5 MeV
	ΔN_{sig}	4.91 %	0.19 %

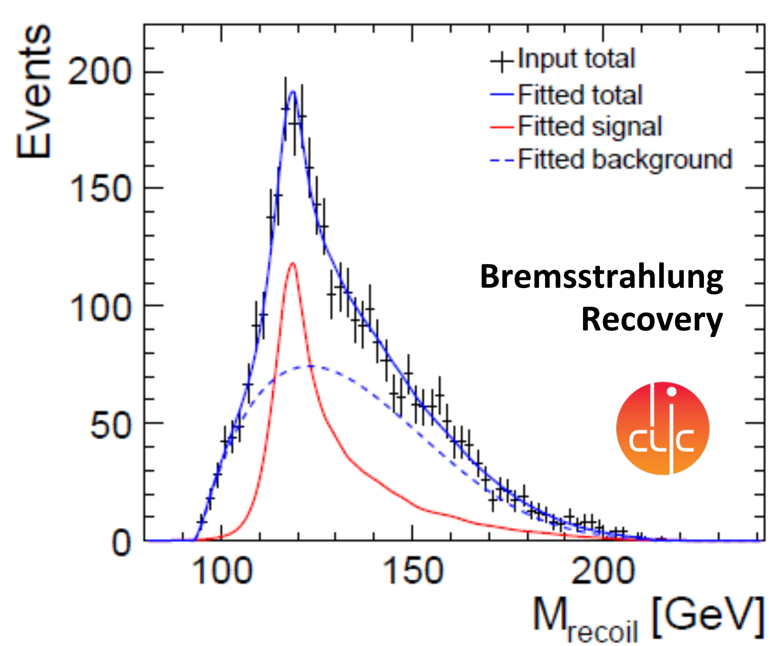
2.5 % meas. of g_{HZZ}

Results: eeX Channel

- eeX channel introduces a new complication: Bremsstrahlung of final state electrons.
- Attempt to find Bremsstrahlung photons and adjust relevant electron four-momenta.
- This procedure increases number of events in recoil mass peak, but increases peak width.

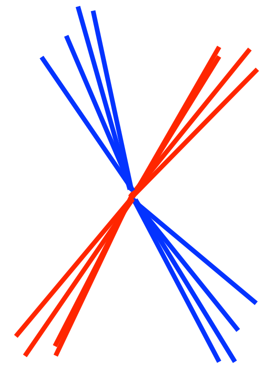
350 GeV Recoil		Mean	RMS
	M_H	119 915.3 MeV	302.5 MeV
eeX	ΔM_H	299.8 MeV	30.4 MeV
	ΔN_{sig}	8.08 %	0.64 %
eeX	M_H	120 012.0 MeV	397.2 MeV
Bremsstrahlung	ΔM_H	394.3 MeV	36.1 MeV
recovery	ΔN_{sig}	7.93 %	0.59 %





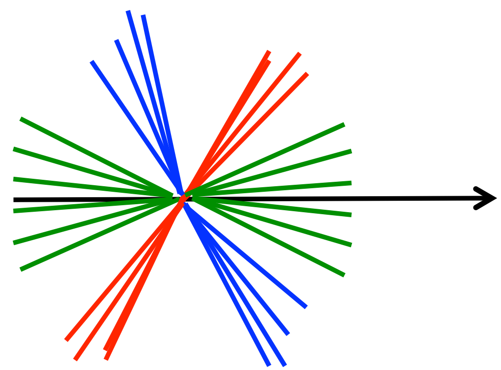
B) 500 GeV HZqq Analysis

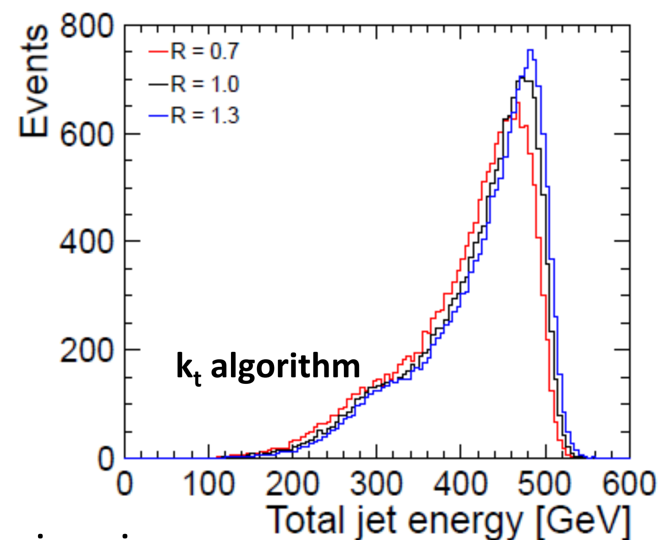
- A 500 GeV LC: measure Higgs cross-sections in a model-dependent manner.
- First HZ → bbqq
  - Force events into four jets, then use kinematic fit to assign jet-pairs to Z and H. Backgrounds have final states that plausibly contain four jets.



Beam backgrounds…

- Jet reconstruction via k_t algorithm (consider three R values);





Analysis

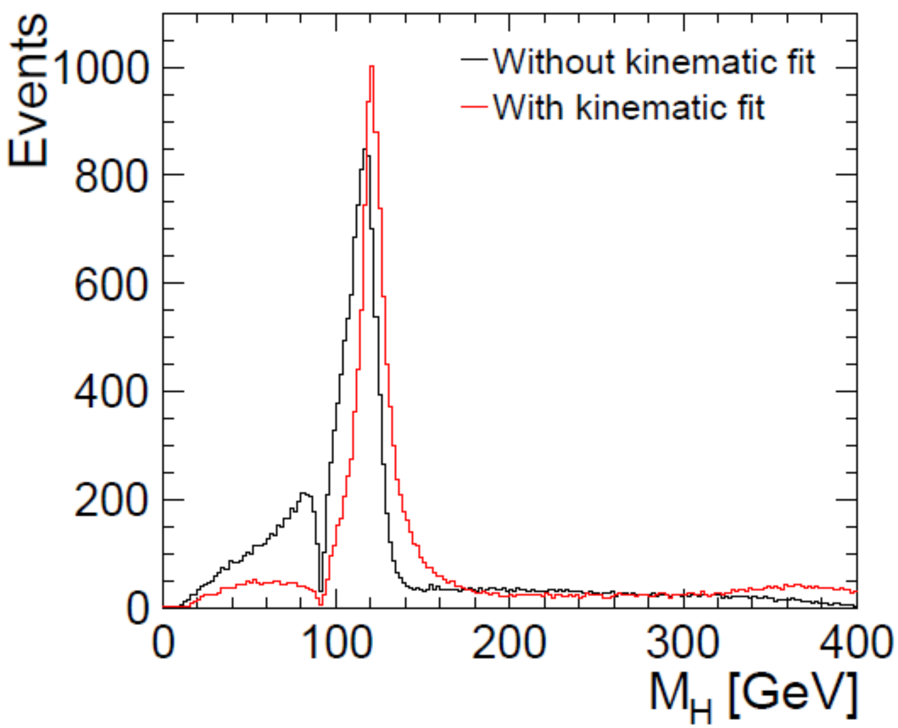
- b-tagging via LCFI. Reject background with multivariate analysis using jet-shape and b-tagging info.

Kinematic Fit

- Kinematic fit using MarlinKinFit package. Inputs to the fit were the four jets in an event. The six possible unique assignments of the four jets to the H and Z were considered.

- Constraints:
  - $\sum_i E_i = 500\ \text{GeV}$
  - $\sum_i (p_{x,i}, p_{y,i}, p_{z,i}) = (5,0,0)\ \text{GeV}$ (note beam-crossing 20 mrad)
  - Mass of one pair of jets equal to Z mass.

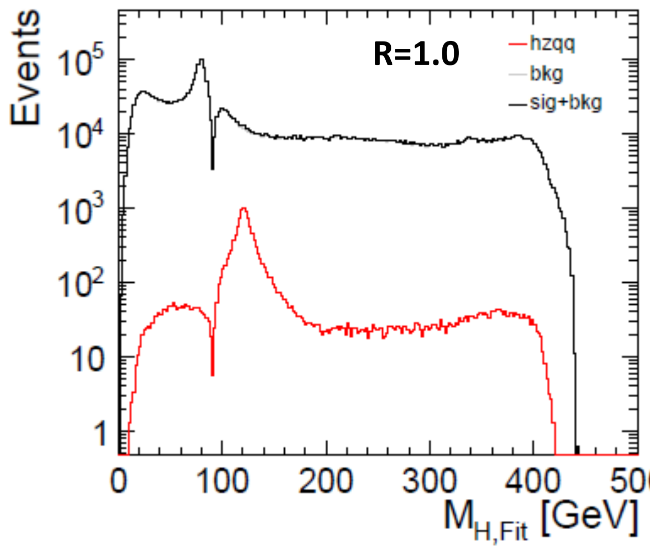
- Fit probability used to finalise assignment of jets to Z and H. The Z mass constraint was then removed to calculate final Z and H four-vectors.

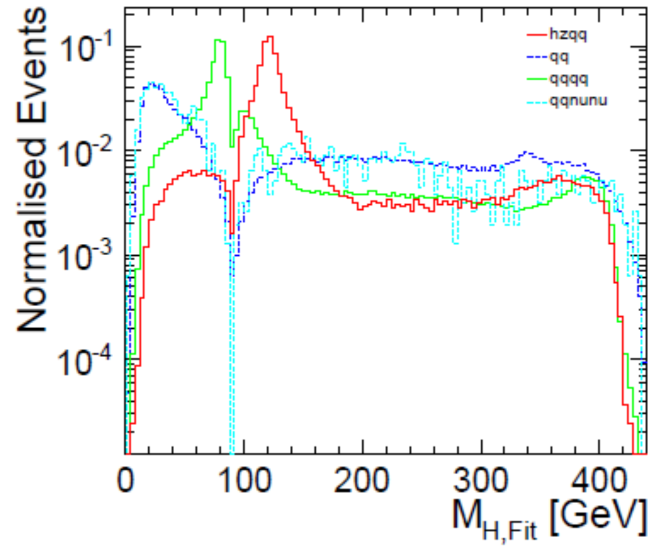


WHIZARD Process Id.	*hzqq*	*qqqq*	*qq*	*qq_nunu*
Events passing kinematic fit	89 %	78 %	21 %	1 %

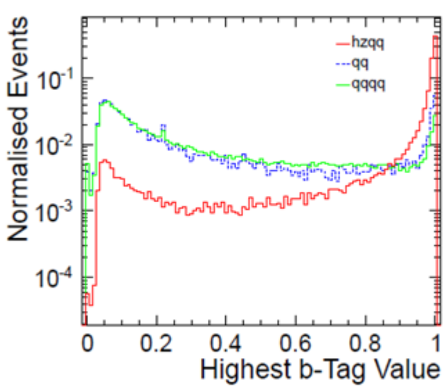
Background Rejection

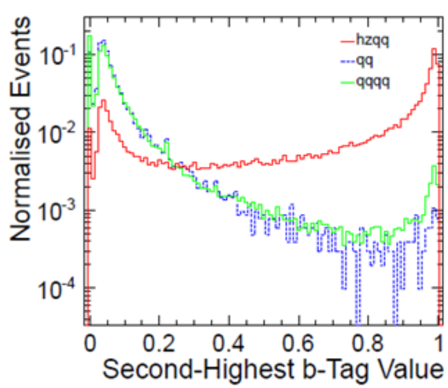
- Distribution to right shows fitted Higgs mass distributions for signal and background events.
- Also show shapes of fitted Higgs mass distributions for each of the individual event samples.
- Background rejection via TMVA, using jet-shape variables and info from jet clustering and b-tagging.

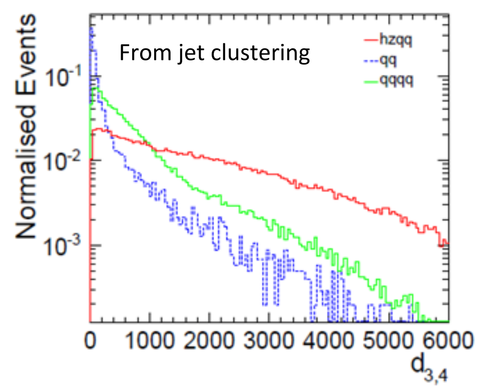


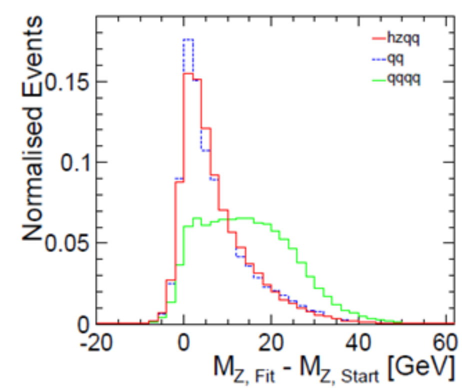


TMVA inputs:

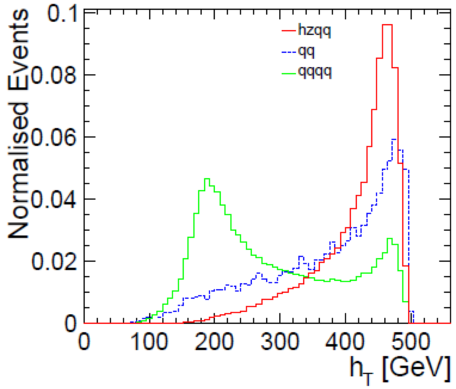




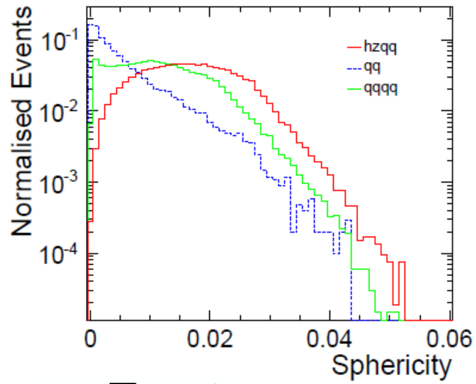




Boosted Decision Tree



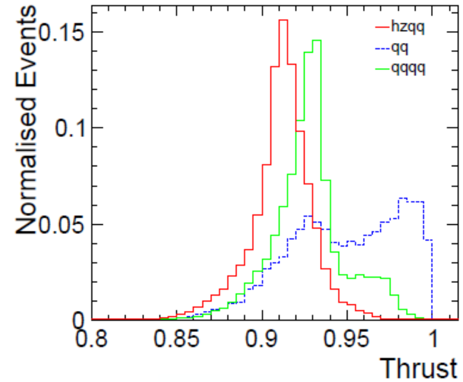
$$ht = \sum_i p_{T,i}$$



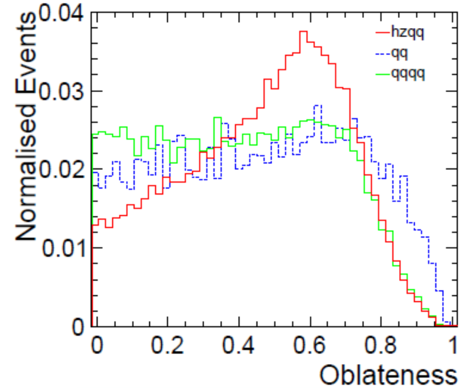
$$S^{\alpha\beta} = \frac{\sum_i p_i^\alpha p_i^\beta}{\sum_i |\mathbf{p}_i|^2}$$

Eigenvalues $\lambda_1 \geq \lambda_2 \geq \lambda_3$

Sphericity $S = \frac{3}{2}(\lambda_2 + \lambda_3)$



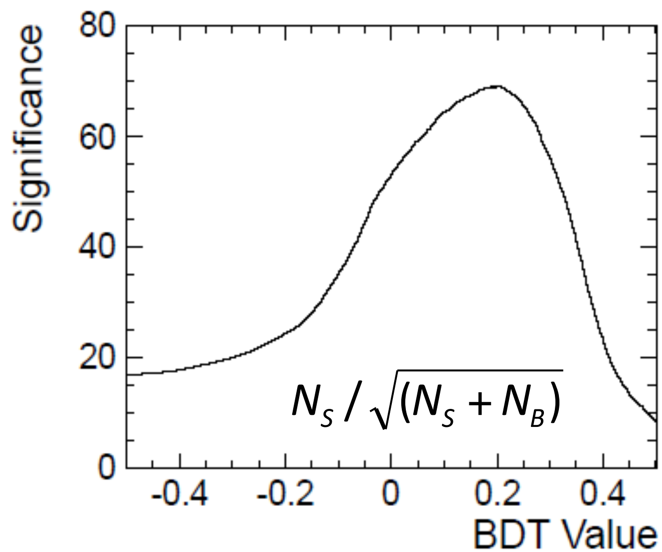
$$T = \max_{|\mathbf{n}|=1} \frac{\sum_i |\mathbf{n} \cdot \mathbf{p}_i|}{\sum_i |\mathbf{p}_i|}$$

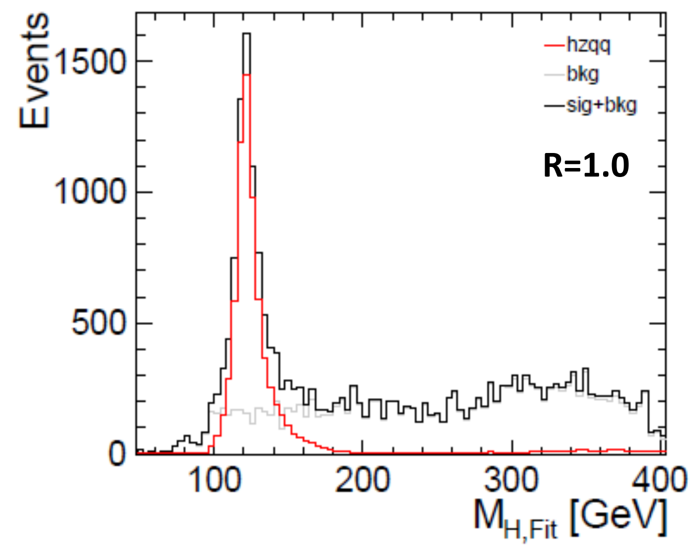


$$T_{major} = \max_{|\mathbf{n}|=1,\, \mathbf{n}\cdot\mathbf{v}_1=0} \frac{\sum_i |\mathbf{n} \cdot \mathbf{p}_i|}{\sum_i |\mathbf{p}_i|}$$

Oblateness $O = T_{major} - T_{minor}$

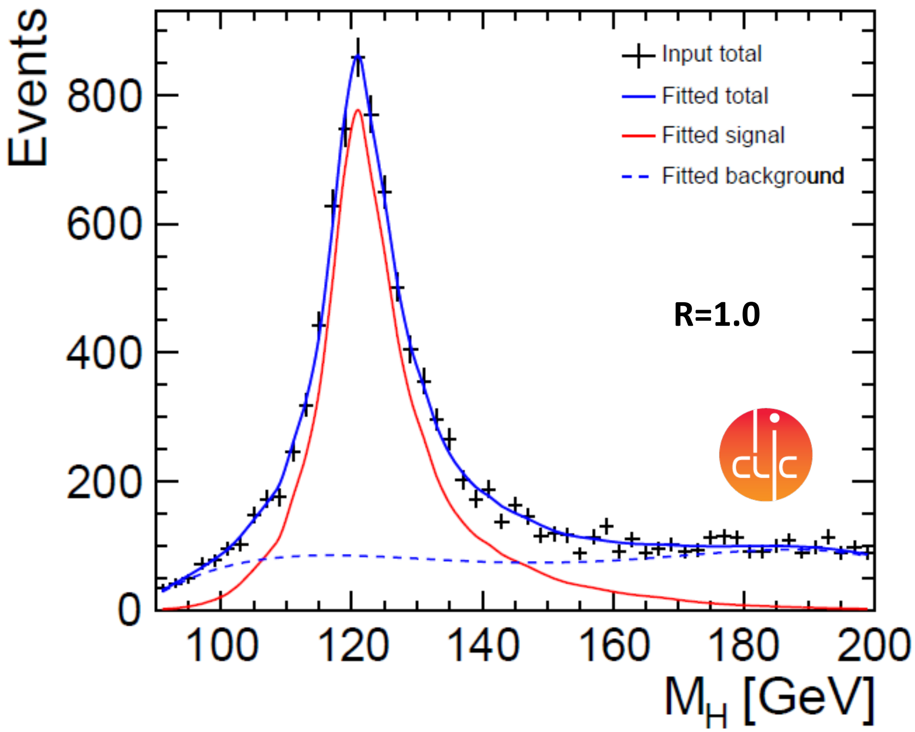
- BDT found to provide the best signal efficiency/purity.
- Final BDT cut chosen to maximise signal significance.
- Fitted Higgs mass distribution is as shown.
- Background approximated by 4th order polynomial.





Results

- Examine best-fit values for 1000 representative 500fb^{-1} test "data" samples
  - Consider different cone sizes

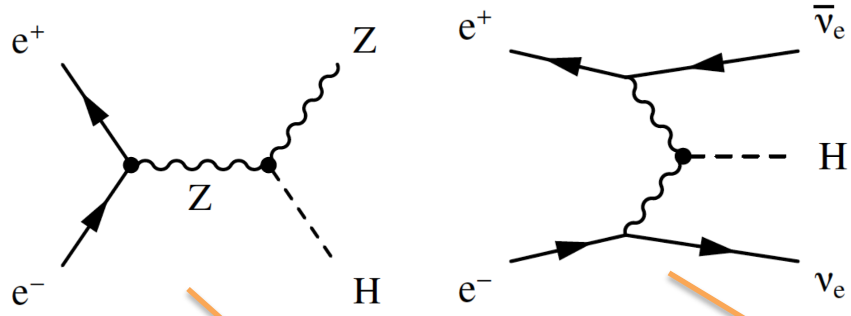


500 GeV $HZqq$		Mean
R=0.7	M_H	119 996.4 MeV
	ΔM_H	97.8 MeV
	ΔN_{sig}	1.572 %
R=1.0	M_H	119 995.1 MeV
	ΔM_H	103.7 MeV
	ΔN_{sig}	1.568 %
R=1.3	M_H	119 997.4 MeV
	ΔM_H	121.8 MeV
	ΔN_{sig}	1.574 %

- 1.6 % measurement of σ x BR

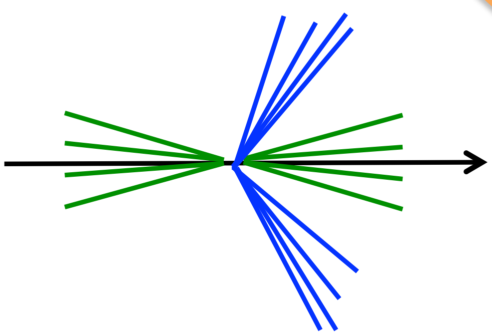
C) 500 GeV Hνν Analysis

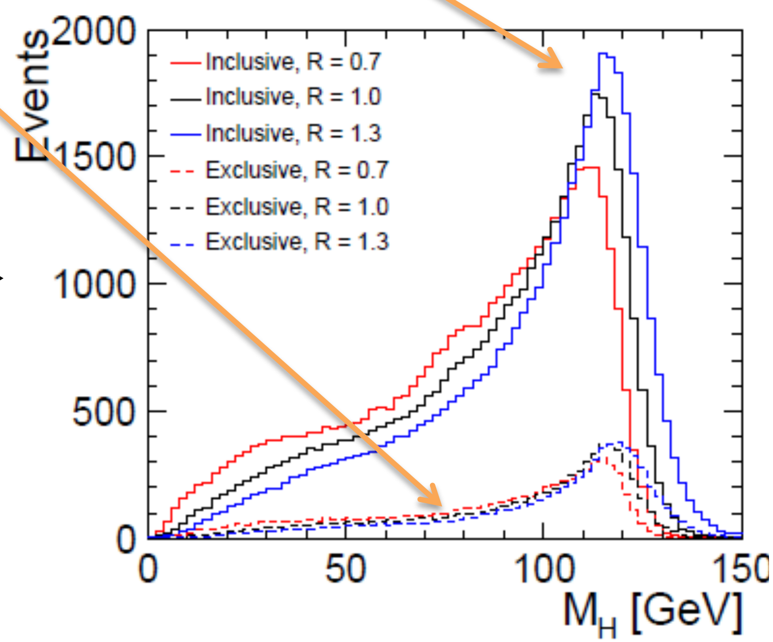
- Hνν at √s =500 GeV: contributions from Higgsstrahlung and WW-fusion.

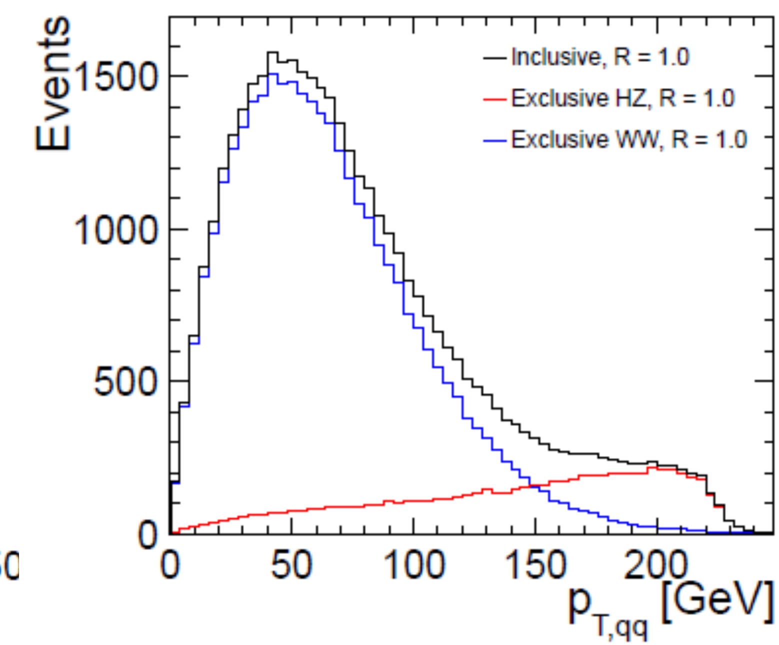


- Use an inclusive signal sample, comprising both Higgsstrahlung and WW-fusion. Analysis: use k_t algorithm to force event into two jets, then search for H→bb final states.

Exclusive = HZ

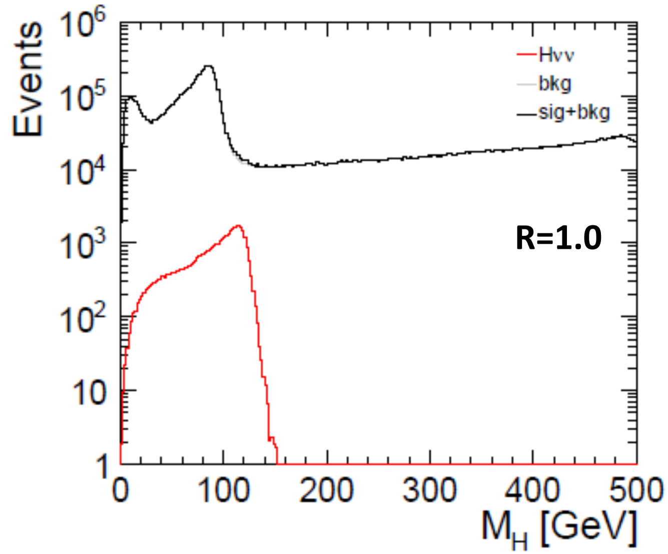


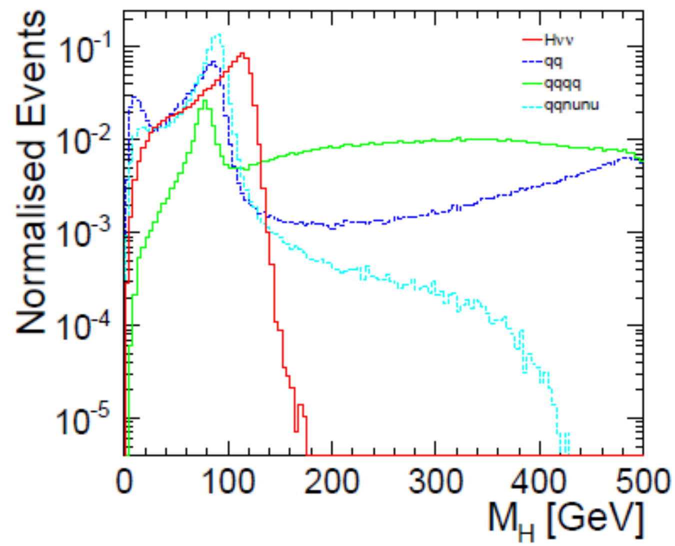




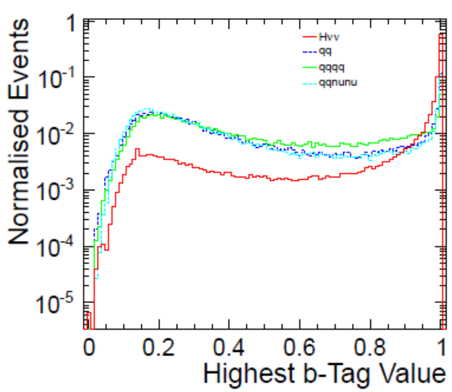
Background Rejection

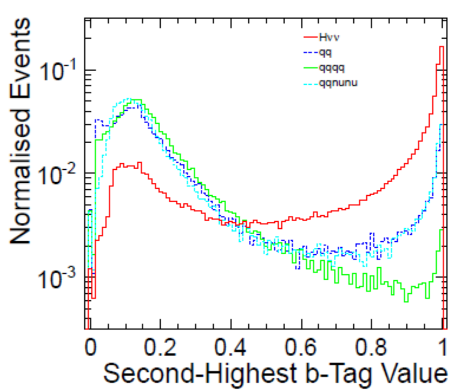
- Distribution to right shows reconstructed Higgs mass distributions for signal and background events.
- Also show shapes of Higgs mass distributions for each of the individual event samples.
- Background rejection via TMVA, using jet-shapes and info from jet clustering and b-tagging.

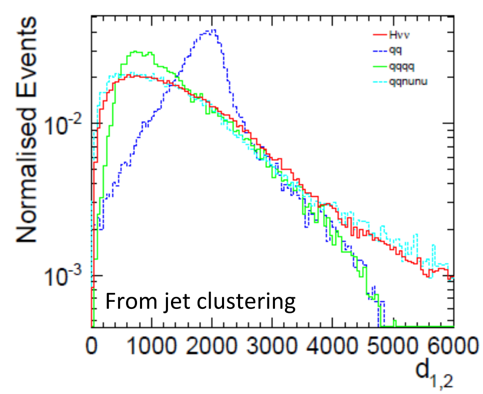


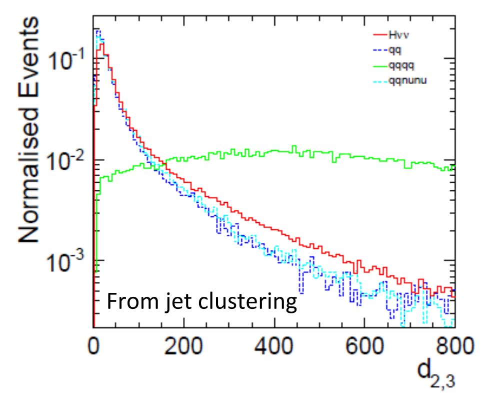


TMVA inputs:

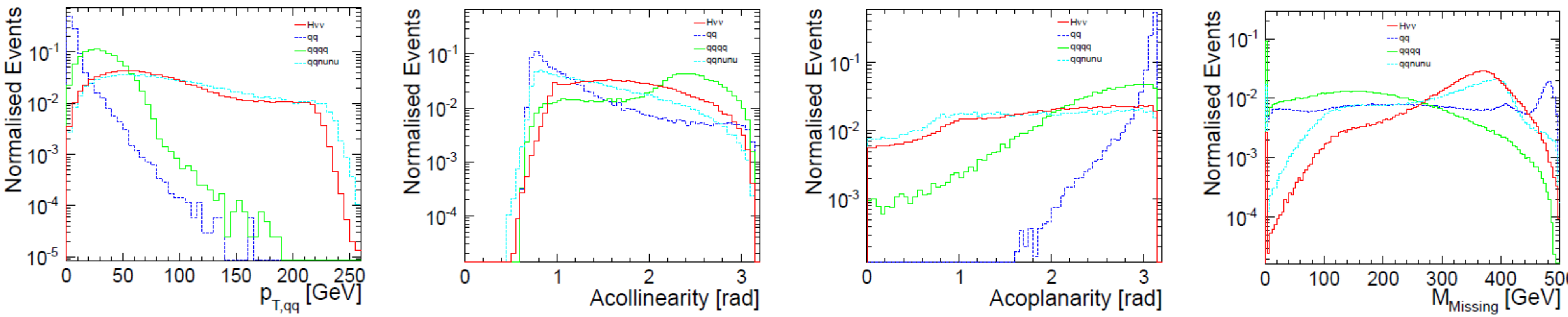




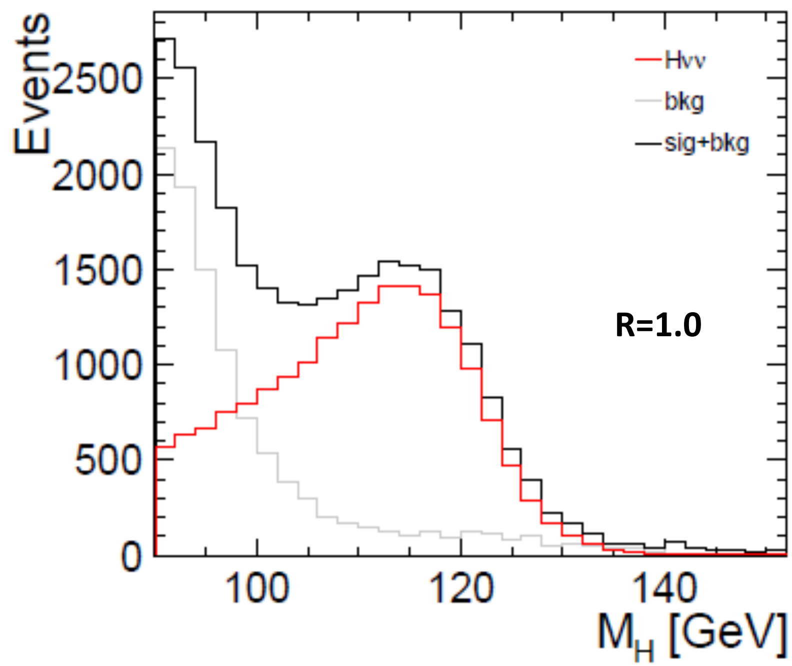




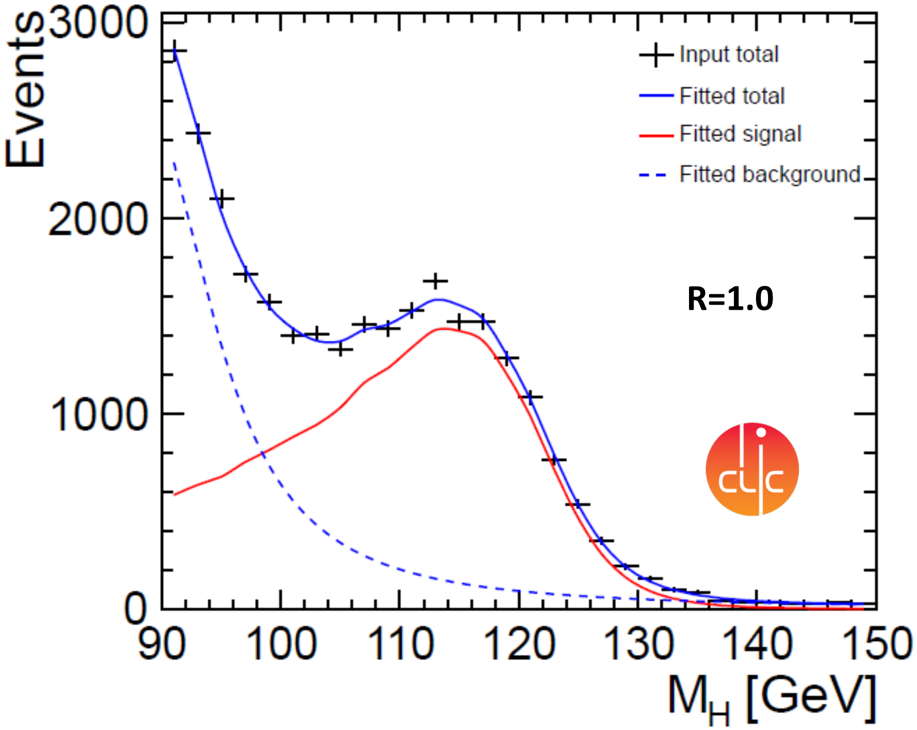
Boosted Decision Tree



- BDT found to provide the best signal efficiency/purity.
- Final BDT cut chosen to maximise signal significance.
- Higgs mass distribution for selected events is as shown.



Results



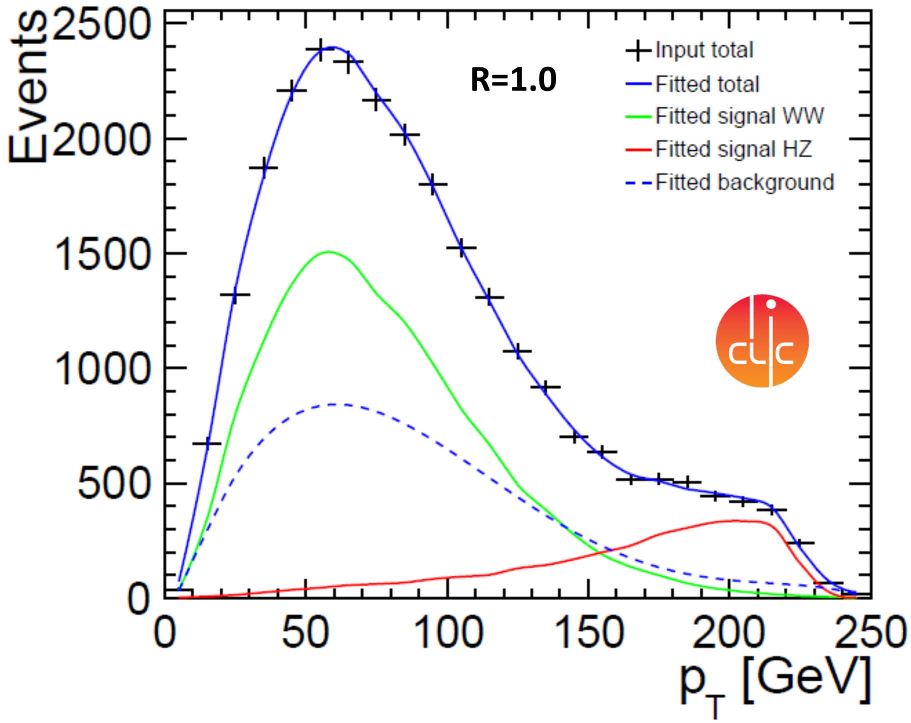
- Examine best-fit values and precisions for 1000 representative 500fb^{-1} test “data” samples:

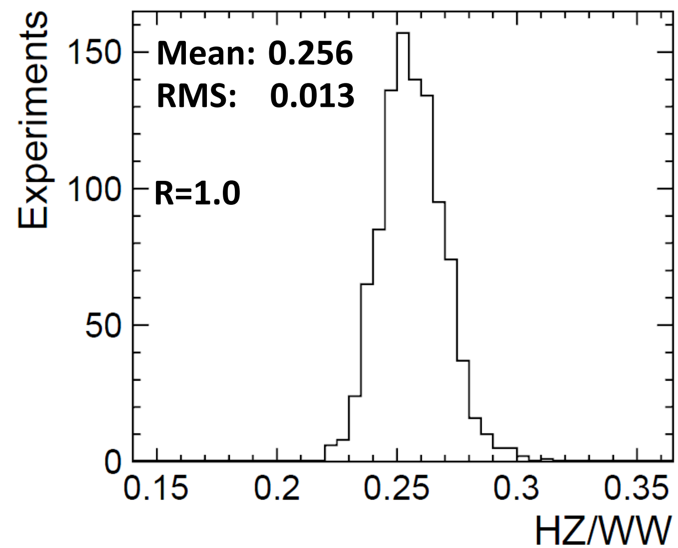
500 GeV	$H\nu\nu$	Mean	RMS
	M_H	120 034.3 MeV	94.7 MeV
R=0.7	ΔM_H	95.9 MeV	1.9 MeV
	ΔN_{sig}	1.061 %	0.007 %
	M_H	120 015.0 MeV	98.4 MeV
R=1.0	ΔM_H	97.2 MeV	1.6 MeV
	ΔN_{sig}	1.010 %	0.007 %
	M_H	120 006.5 MeV	109.7 MeV
R=1.3	ΔM_H	110.8 MeV	2.4 MeV
	ΔN_{sig}	1.057 %	0.008 %

- 1.0 % measurement of σ x BR

Ratio of Couplings g_{HZZ}/g_{HWW}

- Examine p_T distribution for selected events in order to extract a measurement of the relative Higgsstrahlung and WW-fusion normalisations.
- Same selection + M_H > 95 GeV





- Relative HZ/WW normalisations can be determined with precision of 5.1%

- 2.5 % measurement of g_{HWW}/g_{HZZ}

Summary

- ★ $\mu\mu X$ at 350 GeV: — 2.5 % meas. of g_{HZZ}
- ★ bbqq at 500 GeV: — 1.6 % meas. of σ x BR
- ★ $bb\nu\nu$ at 500 GeV: — 1.0 % meas. of σ x BR

2.5 % meas. of g_{HWW}/g_{HZZ}

Assumptions L_{int} = 500 fb^{-1}, M =120GeV, No polarization