Sensitivity to the Higgs self-coupling using full simulation

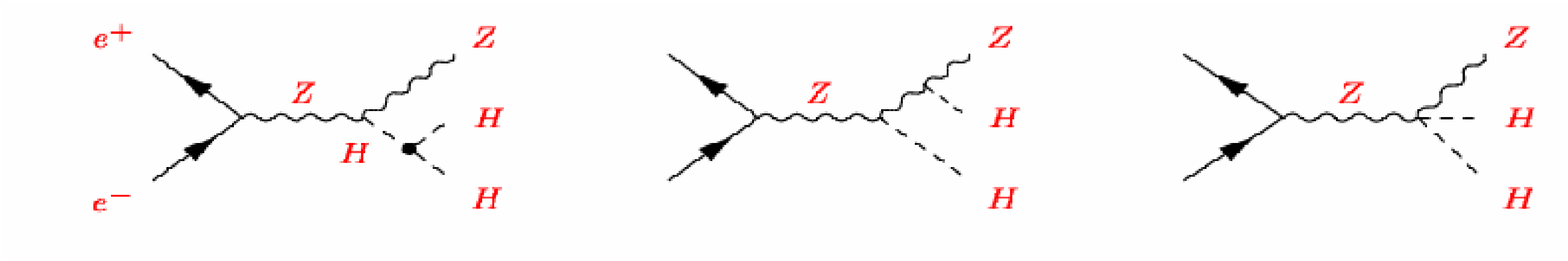
Michele Faucci Giannelli, Mike Green,
Veronique Boisvert, Fabrizio Salvatore, Tao Wu

ILD Workshop, Cambridge, 12 September 2008

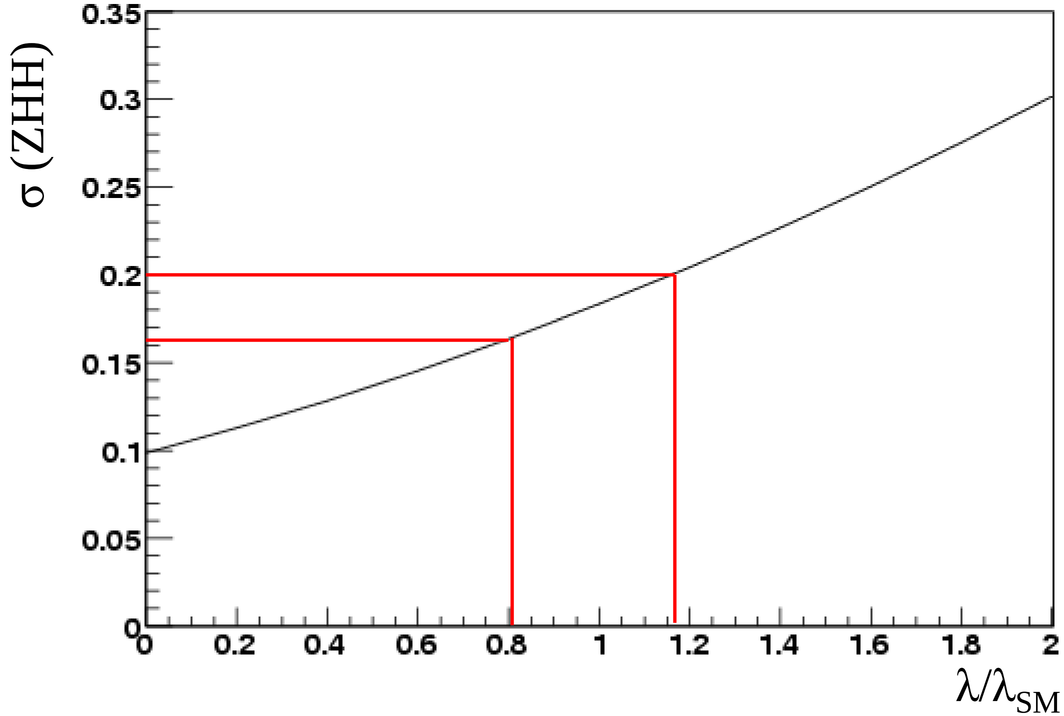
Overview

- Theory: from ZHH cross section to Higgs self-coupling
- Event simulation and reconstruction
- Jet energy resolution in 6-jet environment
- Cut-based analysis
  - B tagging importance
  - Kinematics fit
- Comparison with perfect PFA
- Conclusion

Theory



Effect of second and third diagram is that the sensitivity on the cross section and the self coupling are not linear.



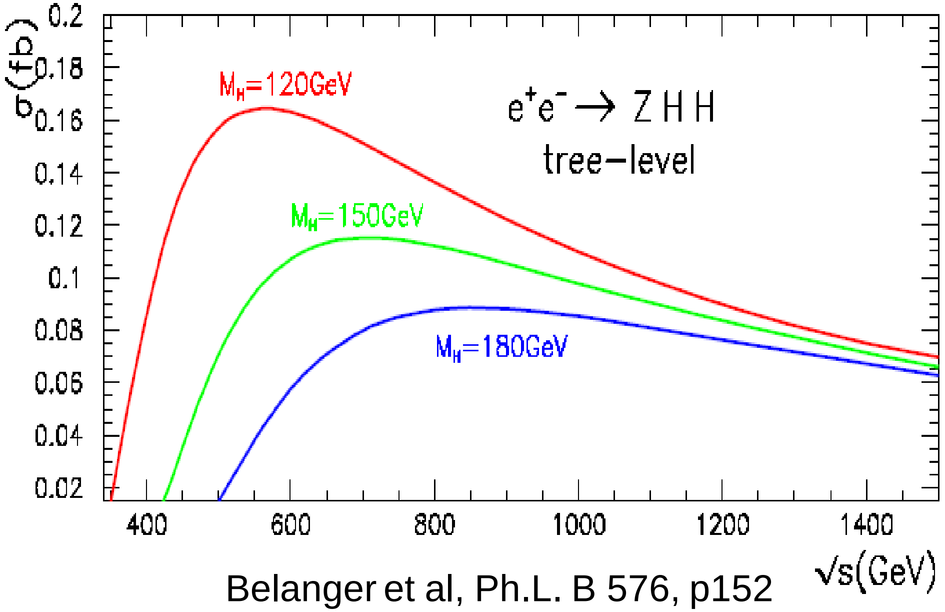
10% resolution on σ (ZHH)

↓

18% resolution on λ

Physics parameters

- This process is best studied at 500 GeV and with a Higgs mass of 120 GeV to maximize the cross section



Belanger et al, Ph.L. B 576, p152

- Given the Z and Higgs SM decays:
  - BR (Z→qq) 70%
  - BR (H→bb) 73%
- Main channel is qqbbbb (40%)
  - $\nu\nu$bbbb (16%)
  - qqbbWW (12%)
  - llbbbb (only 4.5%)

Backgrounds

- The signal cross section is 0.18 fb
  - less than 100 signals for 500 fb^{-1}
  - Only 34 qqbbbb events
- There are several background with 6-jet final state:
  - tt is the main background: σ = 710 fb
    - 160,000 hadronic events
  - WWZ
  - tbtb
  - ZZZ, ZZH
  - ZZ and ZH plus gluon emission

Fast MC results

- Two studies performed by P. Gay and T. Barklow.
- P. Gay group achieves a resolution on the cross section of ~10% using all Z decay channels
- T. Barklow reaches ~ 20% using only hadronic one.
- Both use 2000 fb^{-1} instead of 500 fb^{-1} and jet energy resolution of $30\%/\sqrt{E}$
- However these analyses do not have gluon emission
  - Barklow made a study pointing out a factor 2 worse resolution when using gluon emission
- Considering both the luminosity and the gluon emission factors, a resolution of about 80% should be achieved

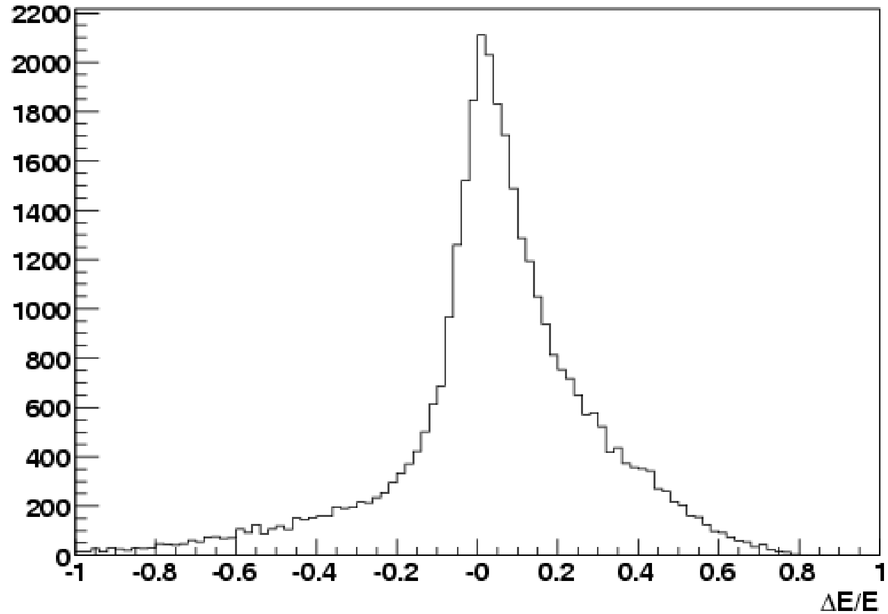
Samples generated at RHUL

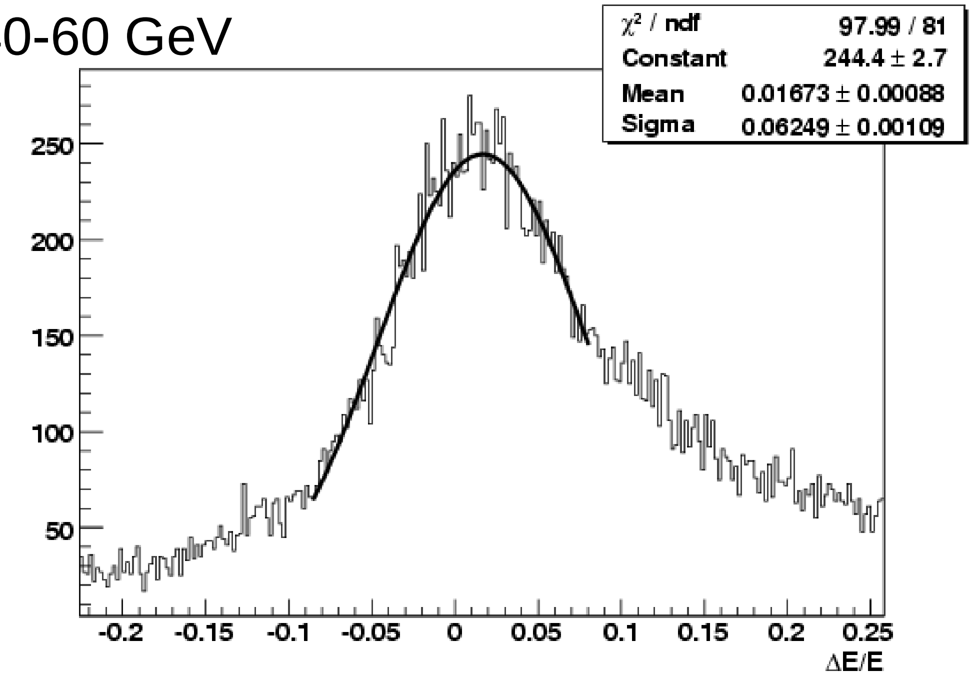
- 500 fb^{-1} simulated and reconstructed
  - Generator: Pandora Pythia and Whizard
  - Beam line: NLC500
  - Polarization: 80% (e^-), 0% (e^+)
- Mokka v06-04
  - Detector model: LDC00Sc
  - Physics list: LCPhys
- Reconstruction similar to mass production:
  - Pandora and Perfect Pandora
  - Vertex reconstruction (vertex charge included)

Jet energy resolution

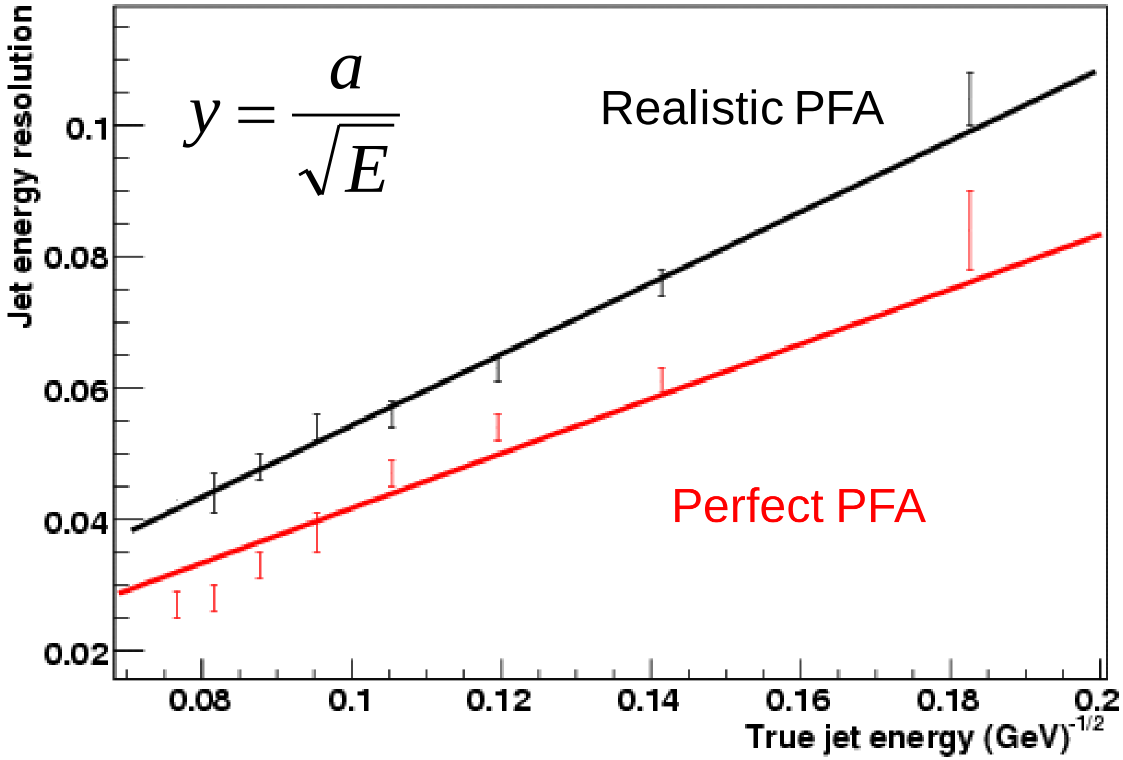
- The reconstructed jets are paired to the Monte Carlo quarks using the combination that maximize the sum of the six scalar products
- All jets are divided in bins of energy, for each bin the resolution is evaluated

Jets of 40-60 GeV





Jet energy resolution

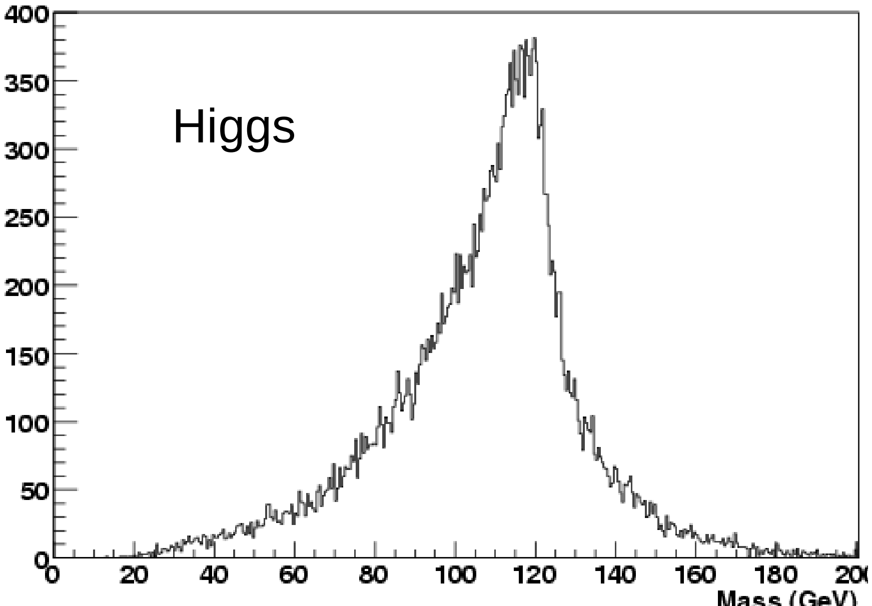


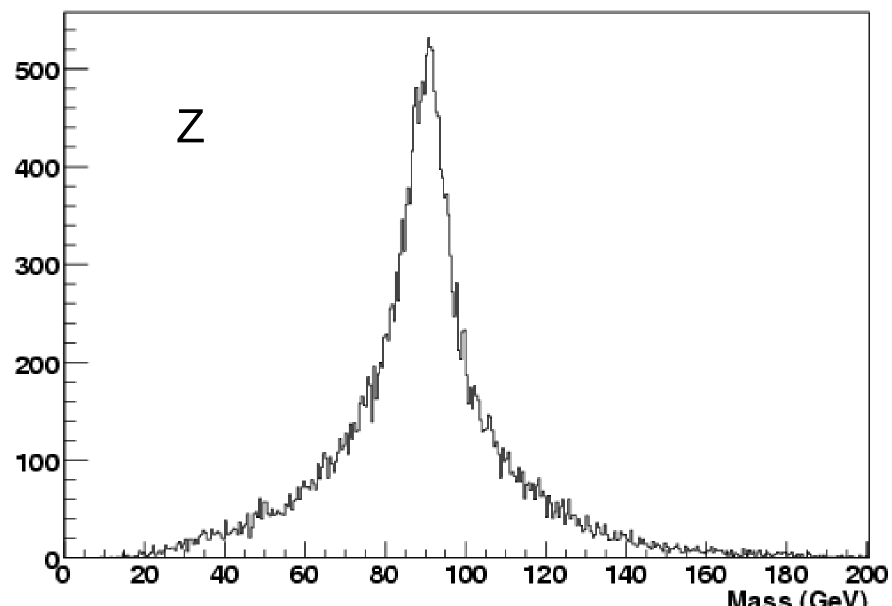
	Jet energy resolution
Realistic PFA	(54±1)%/√E
Perfect PFA	(42±1)%/√E

NB: reconstruction tested on Z→uu events and RMS agrees with M. Thomson results.

Boson mass resolution

- Using the same pairing, it is possible to reconstruct the bosons using the correct jets





Realistic PFA	Mass (GeV)	Width (GeV)
Higgs	116.6±0.3	7.2±0.4
Top	172.8±0.4	20±2
Z	90.3±0.1	6.0±0.1
W	79.9±0.1	4.8±0.1

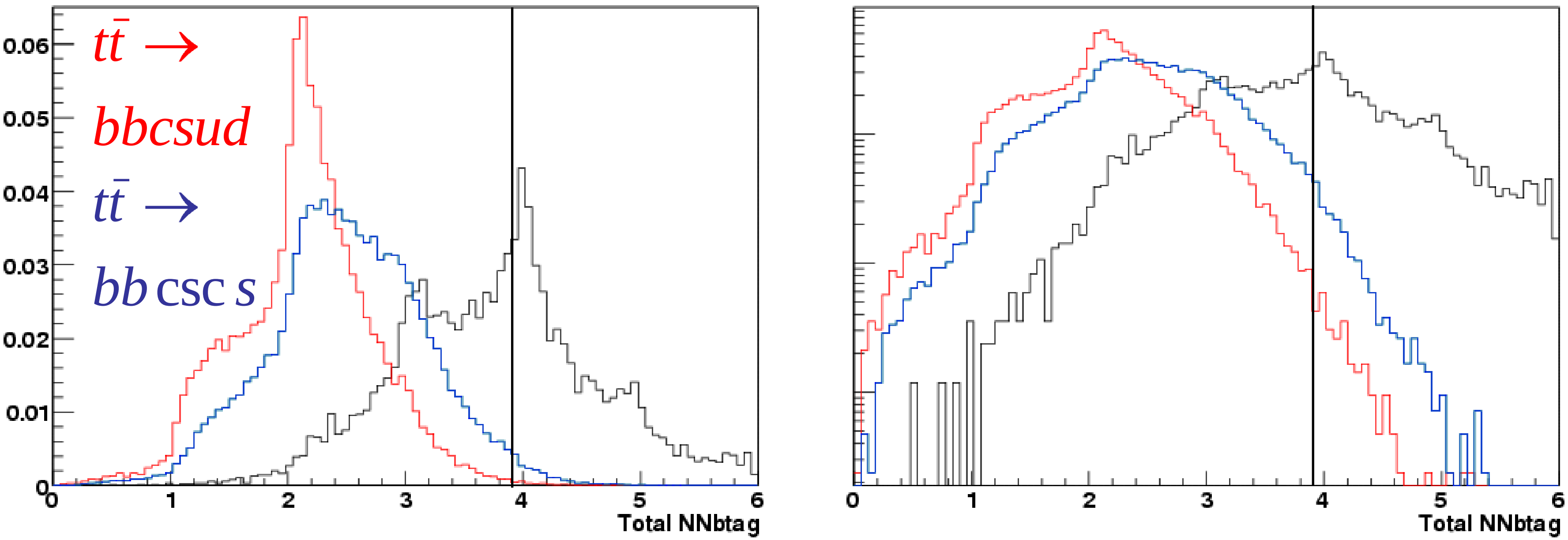
Perfect PFA	Mass (GeV)	Width (GeV)
Higgs	115.6±0.2	5.8±0.4
Top	169.1±0.3	15±2
Z	88.7±0.1	5.1±0.2
W	78.6±0.1	3.8±0.1

Analysis

- Particle flow creates the clusters and reconstructed particles
- 6 jets required in the analysis
- Jets are used by the vertex reconstruction to perform b,c tagging
- Preliminary cuts are applied
- Jets are combined to form the boson using a χ^2 variable
- The distribution of the minimized χ^2 is used to separate the signal and the backgrounds

B tagging cut

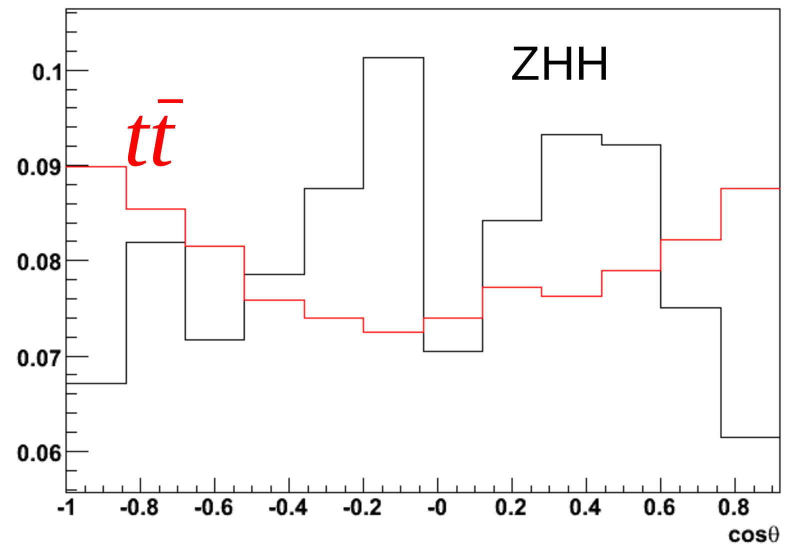
- b tagging plays a central role in reducing the background: requiring 3.9 NNBtag

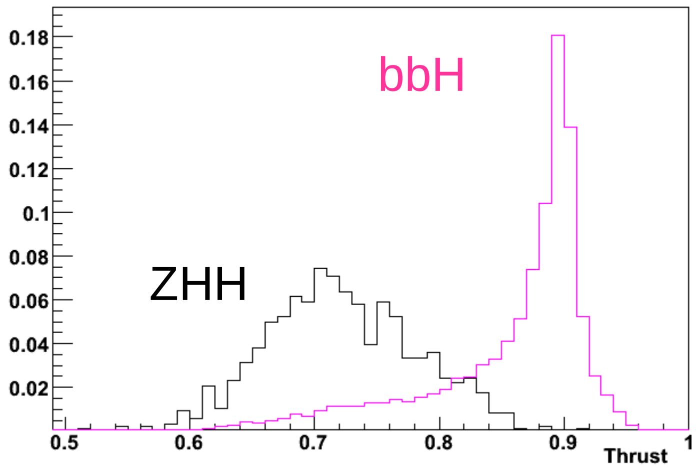


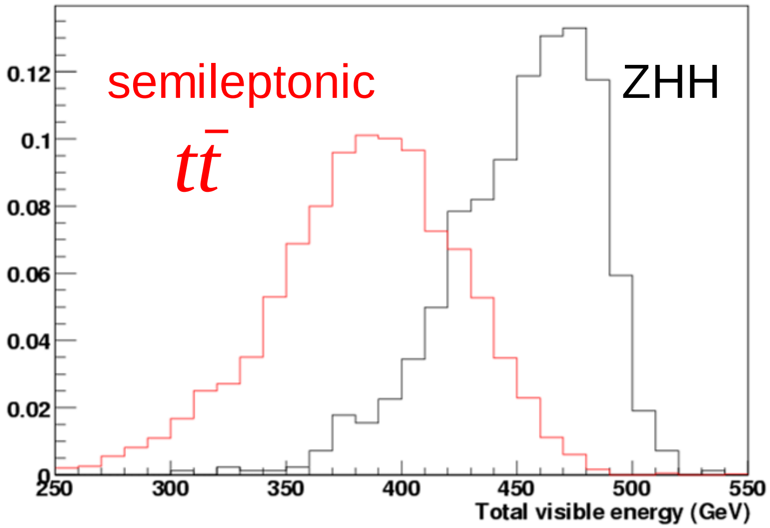
- After the cut $S/\sqrt{(S+B)} = 0.24$

Used cuts

- Topological cuts:
  - $\cos(\theta_{thrust})$
  - Thrust
  - Fox-Wolfram moments
- Missing energy:
  - Total reconstructed energy

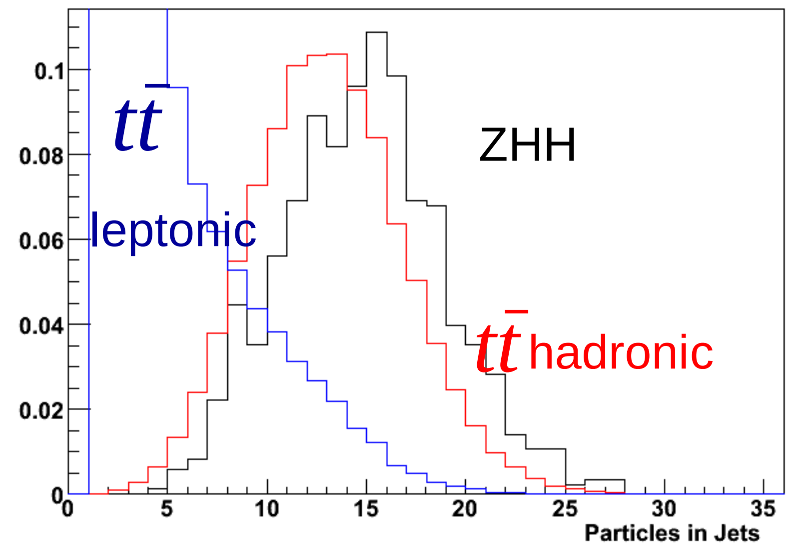






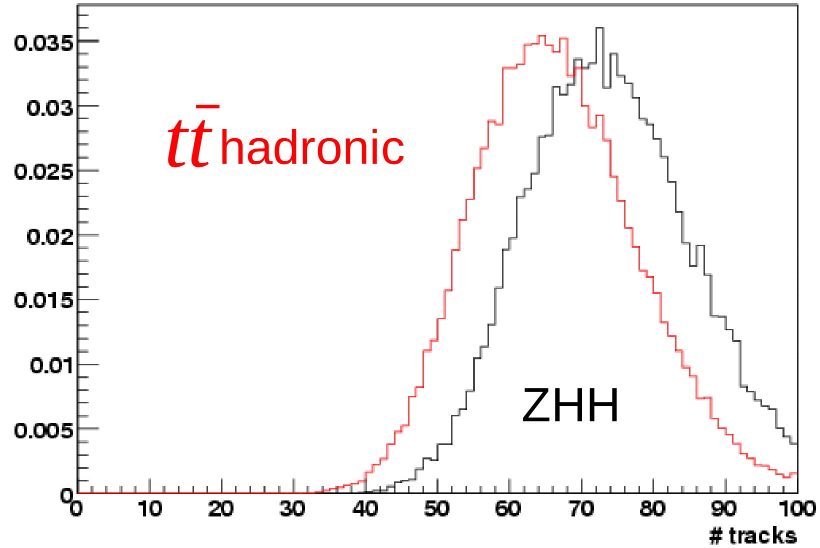
Used cuts

- 2 and 4 jets events can be rejected using:
  - Jets EnergyEM/Energy
  - Jet number of particles
  - Y_6
  - Number of charged tracks



- Multi variables optimization performed to maximize S/√(S+B)

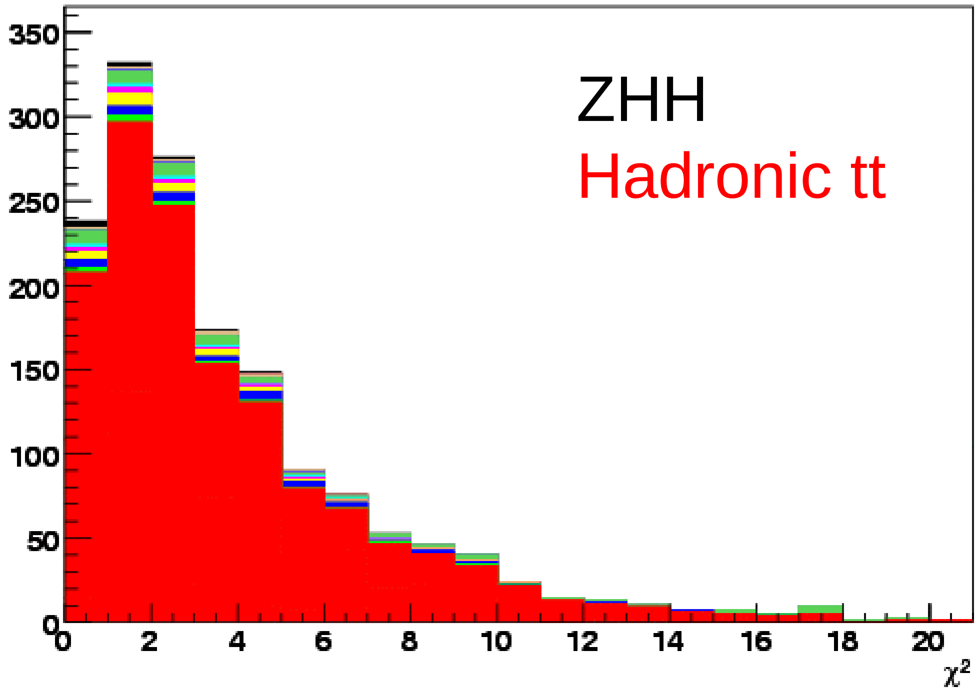
- After cuts S/√(S+B) = 0.36



Jet pairing

- The jets are combined in all 45 possible permutations
- For each permutation a χ^2 is evaluated
- The combination that give the minimum χ^2 is chosen

$$\chi^2 = \frac{(M_{12} - M_Z)^2}{\sigma_Z^2} + \frac{(M_{34} - M_H)^2}{\sigma_H^2} + \frac{(M_{56} - M_H)^2}{\sigma_H^2}$$



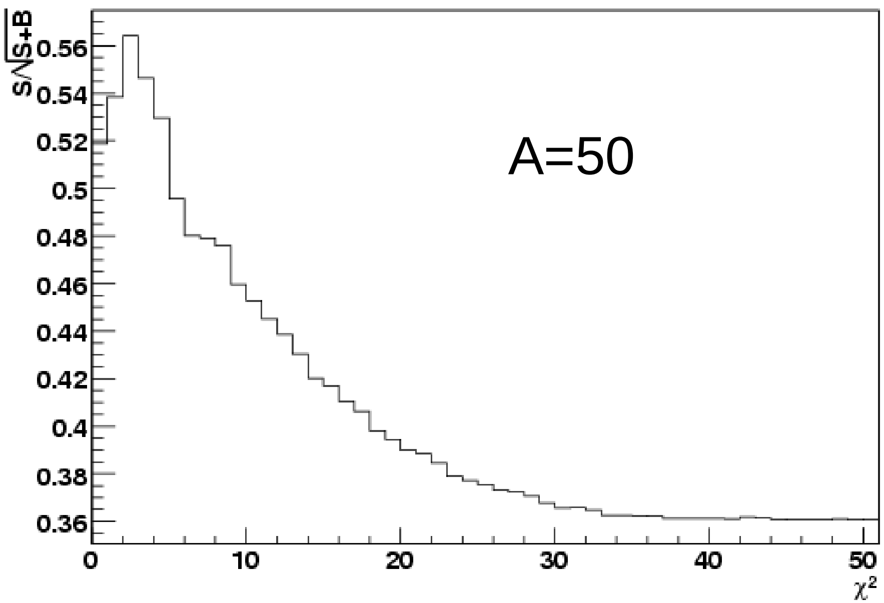
No separation

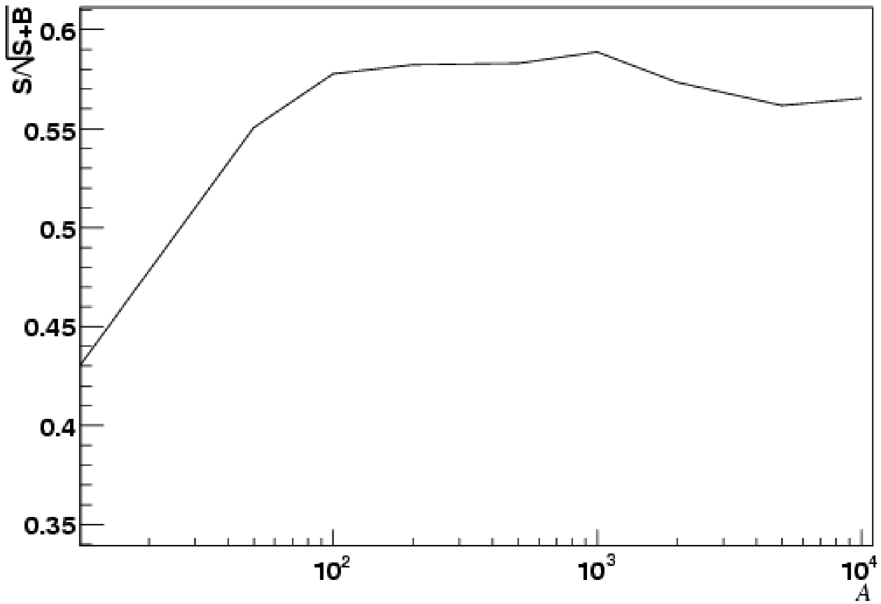
Combining b tagging

- A second χ^2 has been used combining the masses (PFA dependent) and b tagging

$$\chi^2 = \frac{(M_{12} - M_Z)^2}{\sigma_Z^2} + \frac{(M_{34} - M_H)^2}{\sigma_H^2} + \frac{(M_{56} - M_H)^2}{\sigma_H^2} + \sum_{i=3,4,5,6} A(NNbtag(i) - 1)^2$$

- The constant A in b tagging term is optimized maximizing S/√(S+B)





Kinematic fitting

- Kinematics fitting performed varying the jet energy E^{fit}

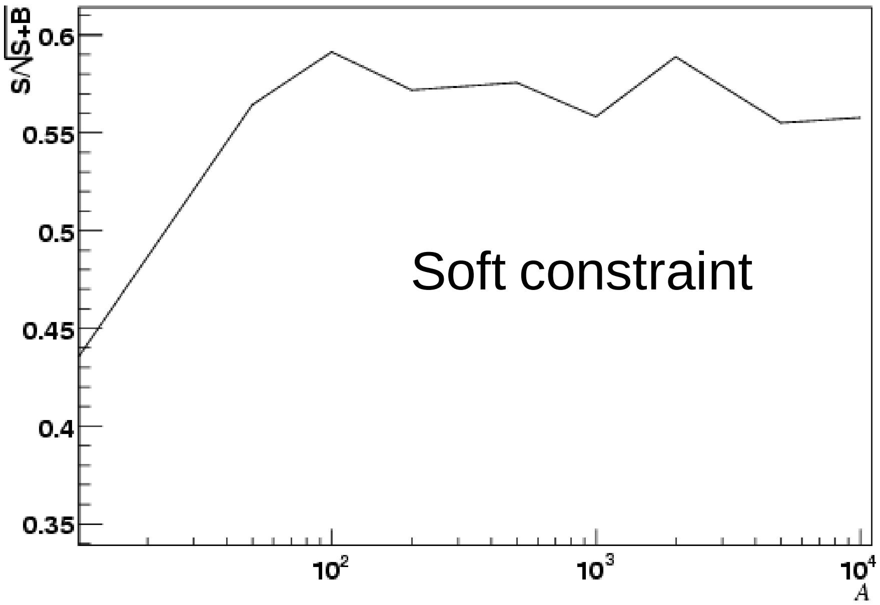
$$\chi^2 = \frac{\left(M_{12}^{fit} - M_Z\right)^2}{\sigma_Z^2} + \frac{\left(M_{34}^{fit} - M_H\right)^2}{\sigma_H^2} + \frac{\left(M_{56}^{fit} - M_H\right)^2}{\sigma_H^2} +$$

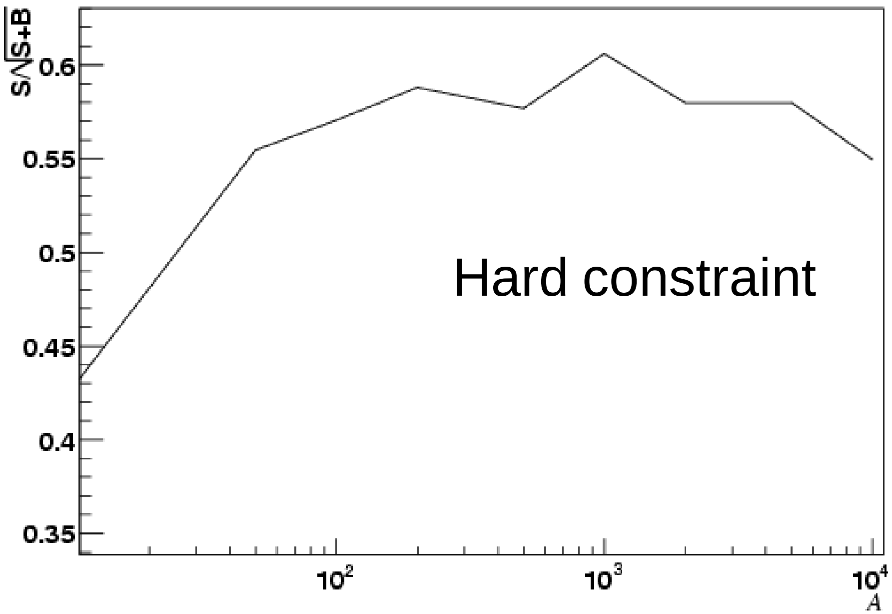
$$\sum_{i=3,4,5,6} A\left(NNbtag(i) - 1\right)^2 + \sum_{i=1}^{6} \frac{\left(E^{fit} - E^{reco}\right)^2}{\sigma_{ene}^2}$$

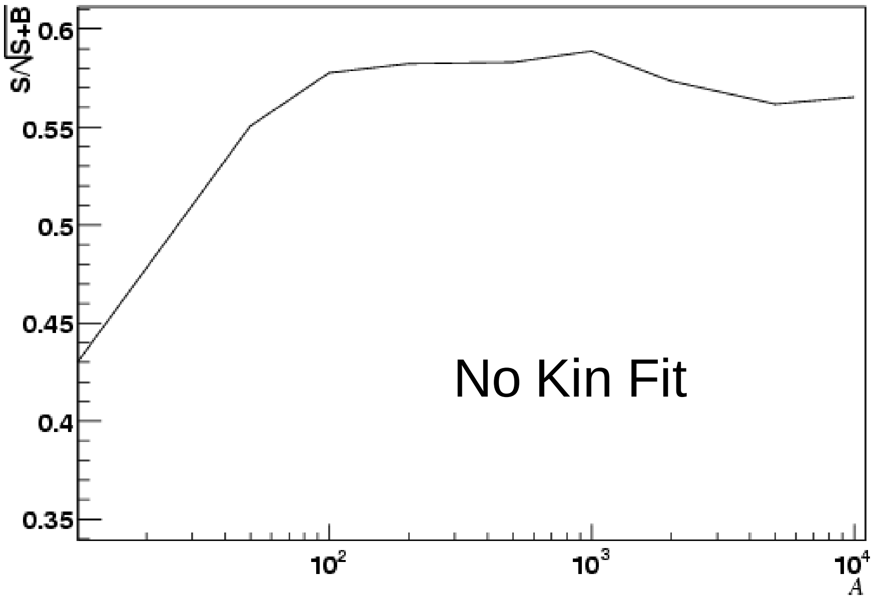
σ

- Two possibilities to constrain the Higgs:
  - Hard constraint → σ_H = 30 MeV
  - Soft constraint → σ_H = 7.2 GeV
- *A* optimized as before

S/√(S+B)







Similar result for kin fit and without it

Best S/√(S+B) is 0.59

Perfect PFA

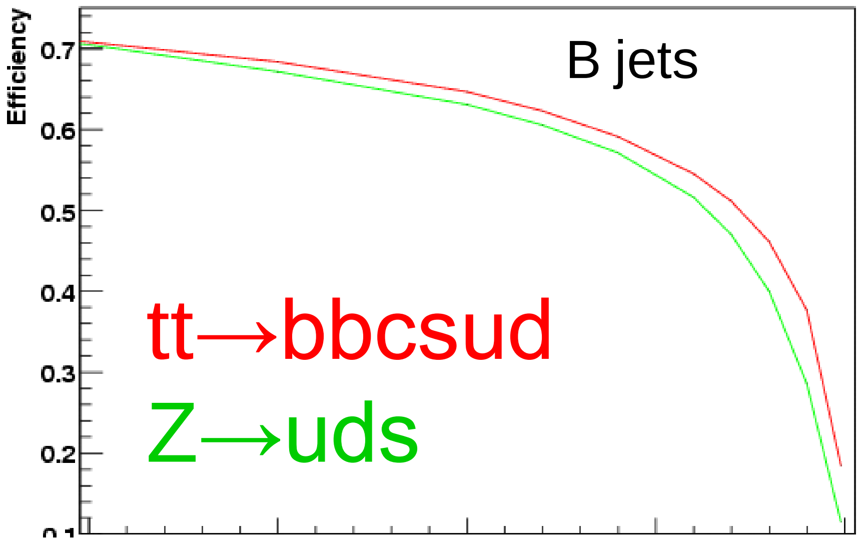
- Realistic PFA
- Best S/√(S+B): 0.59
- 3.6 signal events
- 33 background events:
  - 23 hadronic tt
  - 3 tbtb
  - 2 wwz
  - 1 ttz
  - 1 tth
  - 1 semileptonic tt

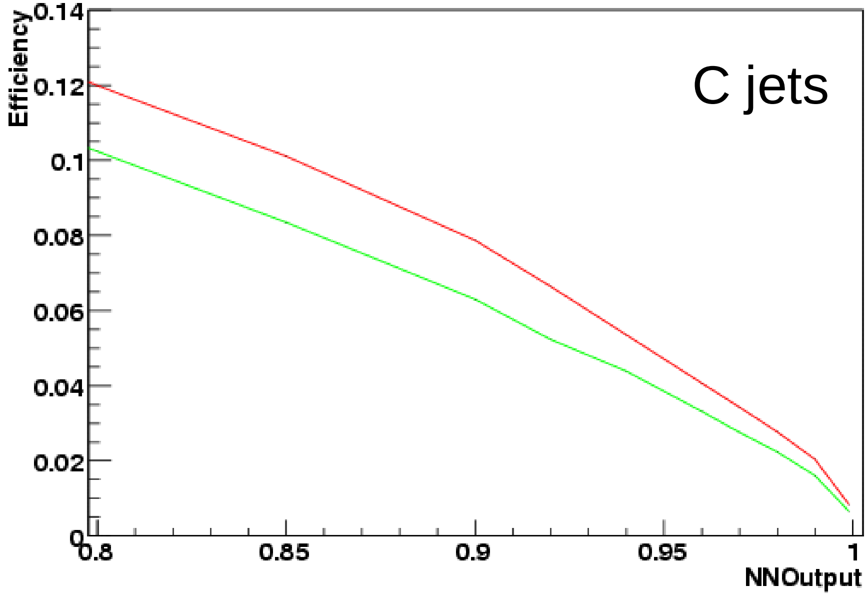
- Perfect PFA
- Best S/√(S+B): 0.59
- 3.5 signal events
- 31.2 background events:
  - 18 hadronic tt
  - 5 tbtb
  - 2 ttz
  - 2 tth
  - 1 zzh
  - 1 wwz
  - 1 semileptonic tt

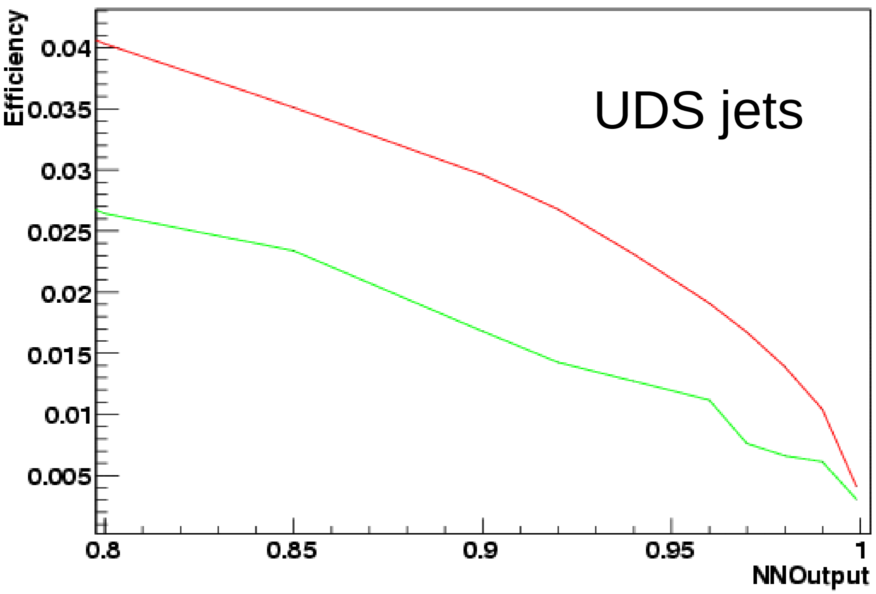
Summary of analysis

- The process ZHH→qqbbbb has been studied using full simulation of 500 fb^{-1} using LDC00Sc
- The resolution ($\frac{\sqrt{S+B}}{S}$) to such process is 170%
  - About a factor 2 worse than fast MC studies
- Since Perfect PFA has similar performances of realistic PFA, the main difference should be in b-tagging performances
- NN analysis almost complete
  - Preliminary results do not show significant improvement

B tagging







Better b tagging efficiency for b jets
Fake rate for c jets is 25% higher
Fake rate is almost double for light jets

Difference is only due to environment, can it be corrected?

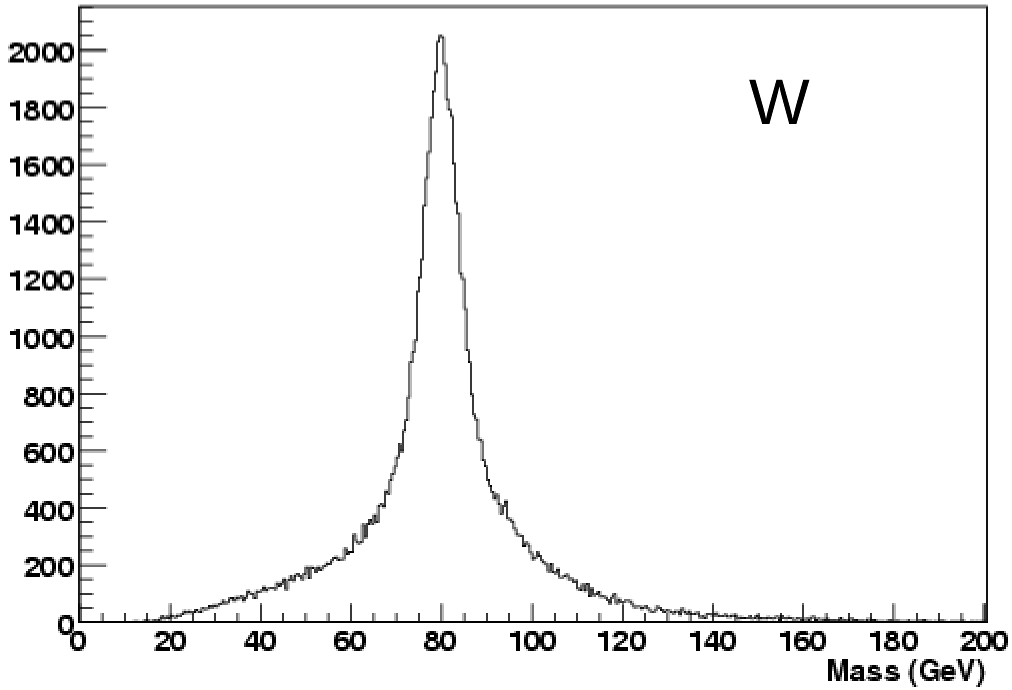
tt always higher than Z, can it be energy dependent?

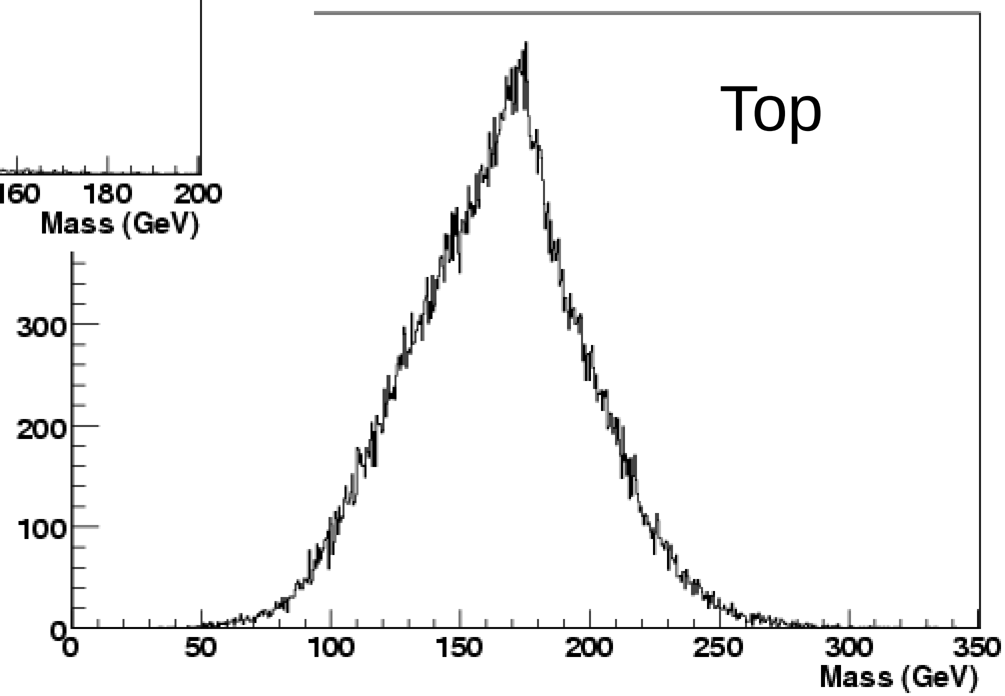
Outlook

- Double check NN analysis
- Release a LC note with final results
- At RHUL the analysis will continue with a master student
  - Focus on DST files for detector optimization
- Improvement on b tagging
  - train network for 6 jet environment

Backup Slides

Mass W and Top





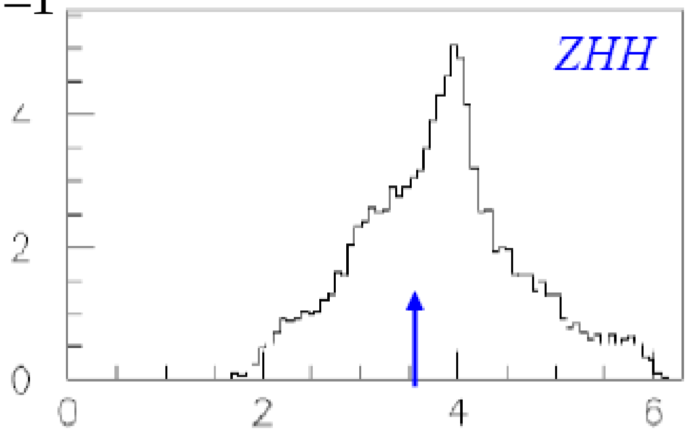
Cross sections

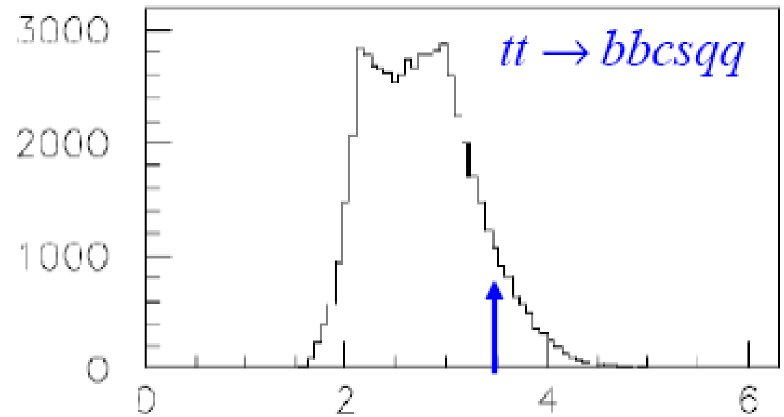
Event type	σ (fb)	Events/500fb^{-1}	Generated events (PP)	Simulated events (Mokka)	% of available events/500fb^{-1}
Zhh (tot)	0.16	80			
Zhh→qqbbbb	0.0593	34	1000	1000	3375
ttbar (lept)	73	36500	100000	30000	82
ttbat (mixed)	310	155000	100000	45000	29
ttbar (cqcq)	82	41000	200000	41000	100
ttbar (uquq)	82	41000	200000	15000	37
ttbar (cquq)	164	82000	300000	41000	50
bbh	10.6	5300	30000	16000	302
ZZh	0.174	87	1000	1000	1150
ZZZ	1.05	525	0	0	0
WWZ	35.3	17650	0	0	0
tth	0.15	75	0	0	0
ttZ	0.7	350	0	0	0
tbW	16.8	8400	0	0	0

Differences with Fast Simulation

$$\sum_{i=1}^{i=6} NNbtag(i)$$

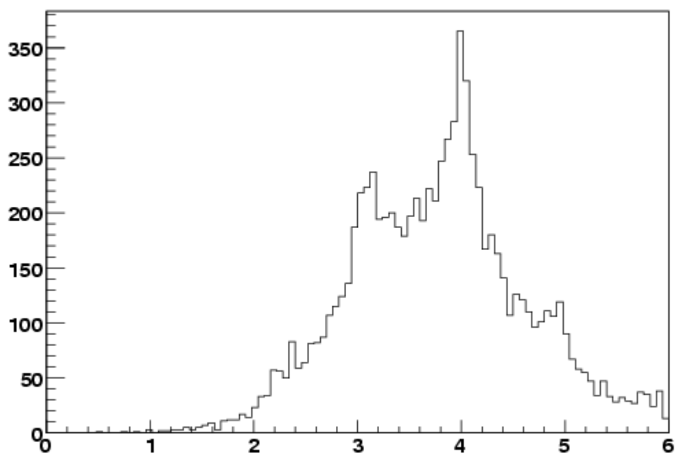
Fast simulation

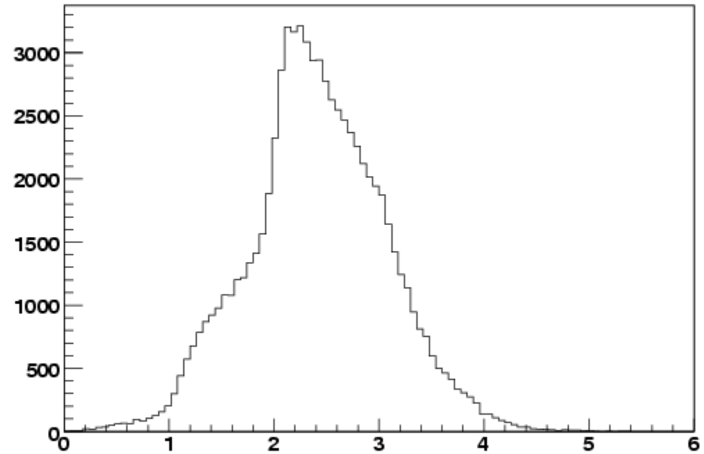




by T. Barklow (LCWS07)

Full simulation





Neural Network analysis

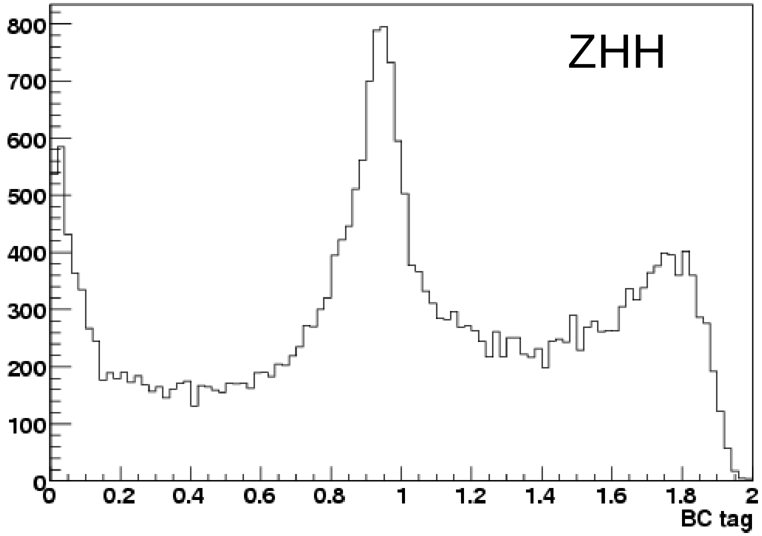
- The neural network package developed for the vertex reconstruction is used.
- A separate sample has been generated to train the network (~125 fb^{-1})
- The training is performed using the back propagation algorithm and 300 epochs.
- Several combination of inputs and nodes in the hidden layer have been tested
  - From 5 to 35 inputs variables
  - From 1 to 45 nodes in hidden layer
- The best result is achieved when 35 variables are used.
- The best network structure is 35:40:1

Most relevant input variables

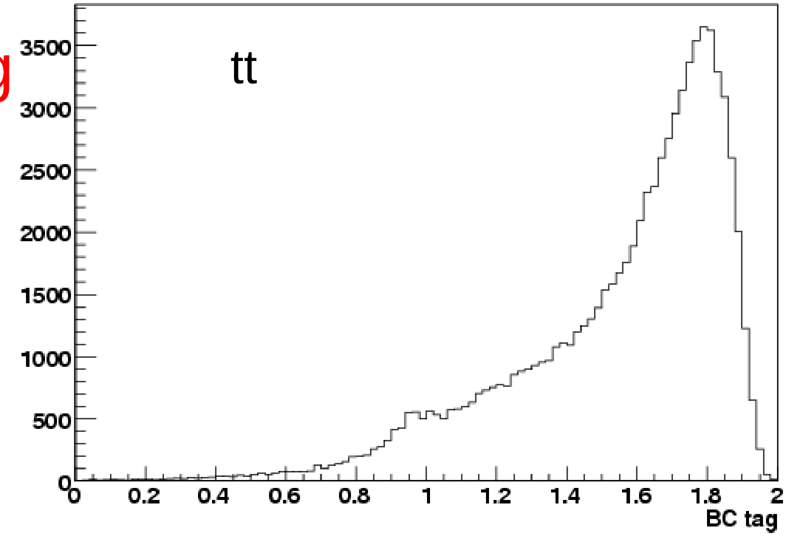
$$\chi_{ZHH}^{2} = \frac{\left(M_{12}^{fit} - M_Z\right)^2}{\Gamma_Z^2} + \frac{\left(M_{34}^{fit} - M_H\right)^2}{\Gamma_H^2} + \frac{\left(M_{56}^{fit} - M_H\right)^2}{\Gamma_H^2}$$

$$\chi_{tt}^{2} = \frac{\left(M_{12}^{fit} - M_W\right)^2}{\Gamma_W^2} + \frac{\left(M_{45}^{fit} - M_W\right)^2}{\Gamma_W^2} + \frac{\left(M_{123}^{fit} - M_T\right)^2}{\Gamma_T^2} + \frac{\left(M_{456}^{fit} - M_T\right)^2}{\Gamma_T^2}$$

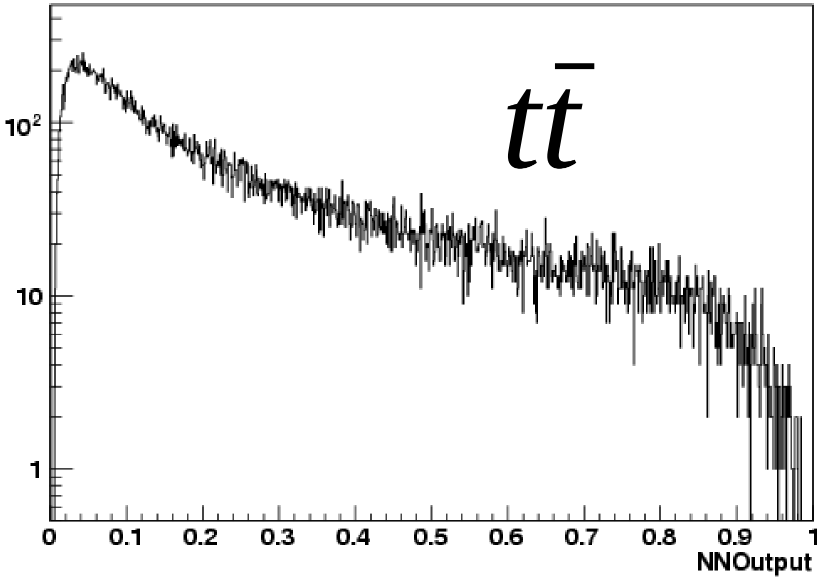
Since the high combinatorial in pairing to form the bosons is the reason the background simulate ZHH events, the most effective χ^2 variables are those in which the number of combination are reduced imposing the b jets to for the Higgs or to be the jets not coming from the W.



BC tag



NN output



Separation

- The separation improves within the statistical error
- 0.60 0.15 is the best $S/\sqrt{(S+B)}$ achieved
  - Cut based analysis reached 0.59 0.06