Jet Flavour tagging

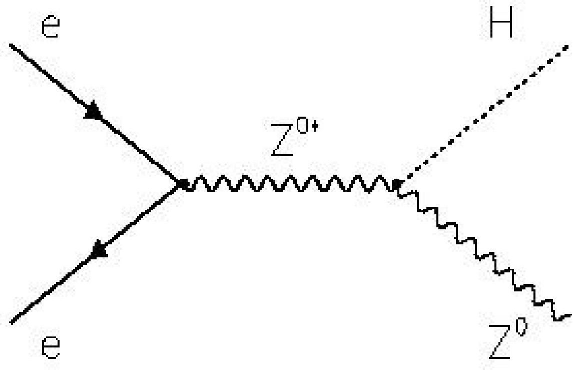
Thorsten Kuhl , Universität Mainz

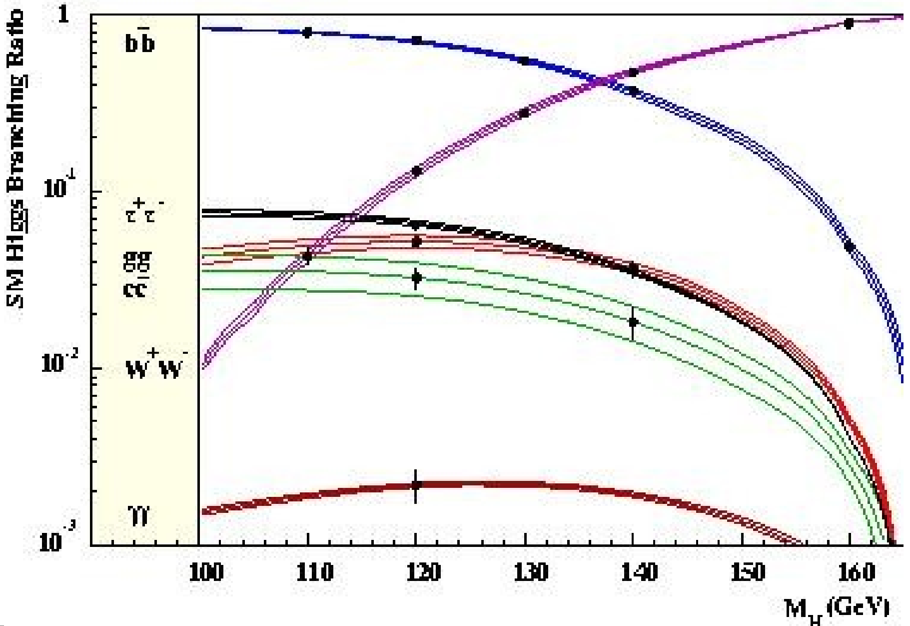
ILC Vertex Detector Workshop, Ringberg

- Analysis description
- Full simulation (Geant 3, single track analysis)
- Fast simulation (Simdet, 500fb^{-1} Higgs BR)
- Summary

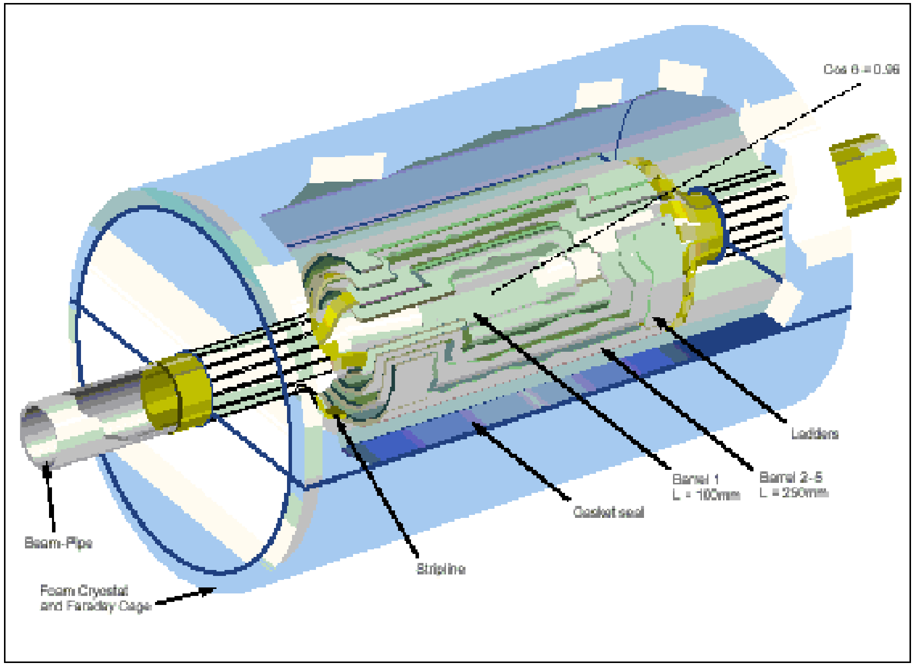
Motivation

- Important physics question
- Benchmark for detector and tool performance:
  - SI detector performance study: different CMOS configuations
  - Full trackwise vertexting (ZVTOP) in physics events
  - Very suffisticated analysis technics: realsitic lepton/photon id, likelihood selections
- Full 500 fb statistics 10 times more for reference distributions)
- Need to use „state of the art" analysis technics on all levels





Silicon Detector

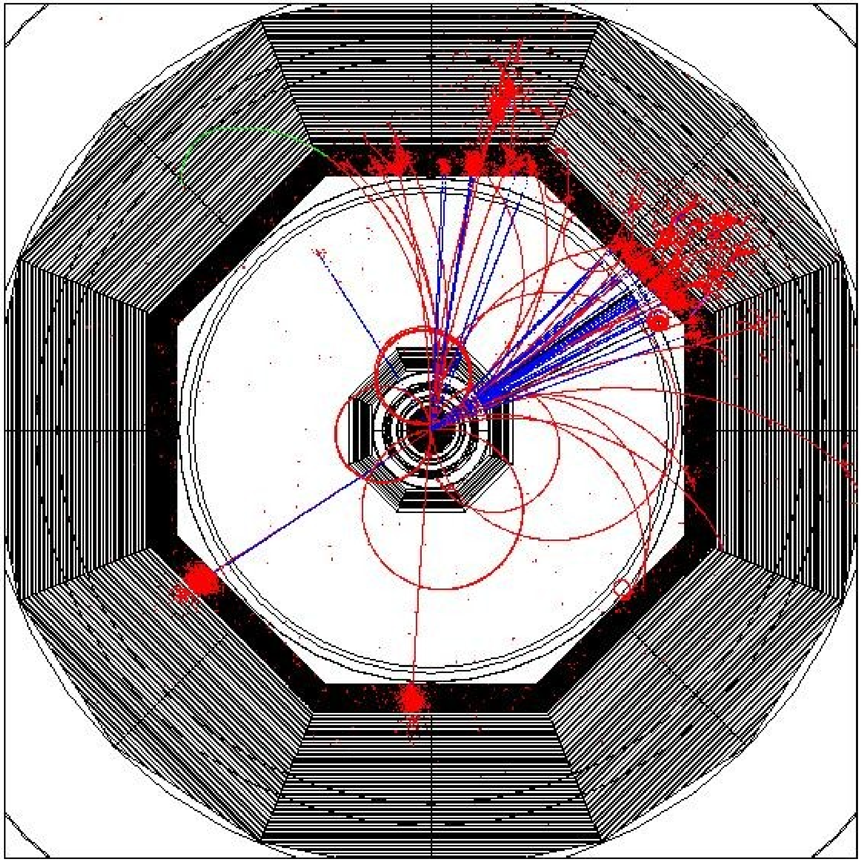


- 5 layer vertex detector
- Size:20•20μm²
- Innermost layer:
  R=1.5cm
- 900 Megapixel
- Resolution:

$$\sigma_{IPr-\phi} = \left(4.20 \oplus \frac{4.00}{(p/GeV)\sin^{3/2}(\theta)}\right)$$

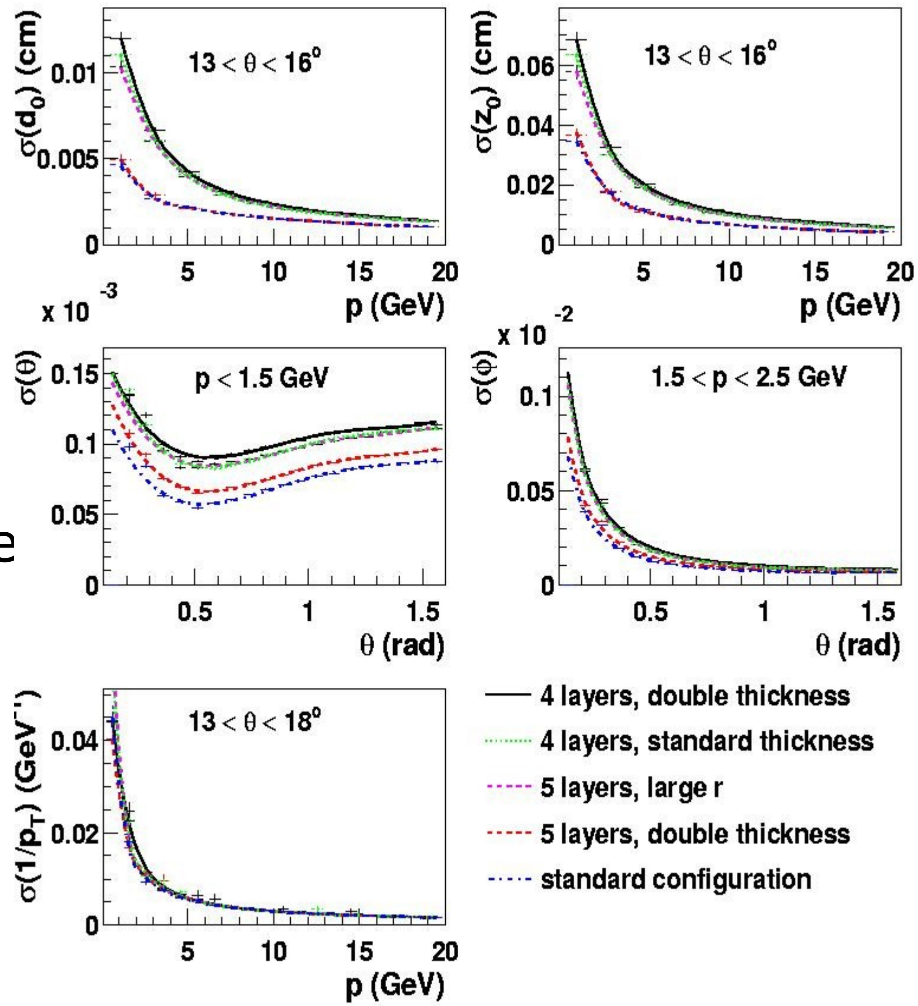
Experimental set up

- “Tesla” like detector
- Full Simulation:
  - Brahms
  - Based on Geant 3
- Fast Simulation
  - Simdet 3
  - Smearing based on parametrisation extracted from Brahms
  - No machine related backgrund (no beamstahlung!)



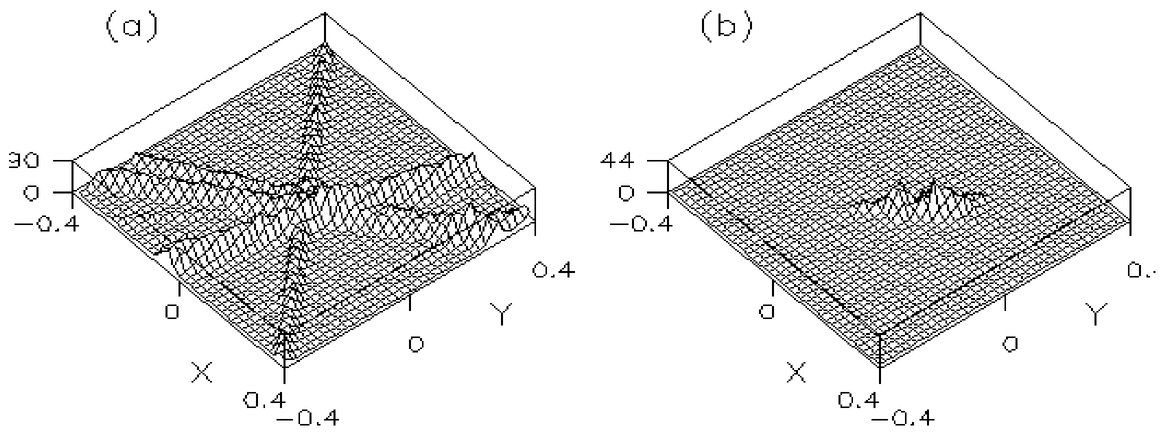
Vertex detector resolution study

- Comparison of CCD (Brahms)
  - Standard 5 layer
  - 5 layer, double wafer thickness
  - Innermost layer removed
  - Innermost layer removed, double thickness
  - 5layer, large r
- Material does not play a big role for tracking (but for calorimeter!)
- Position of innermost layer: very important !
- 4 layer large radius as good as 5 layer large radius

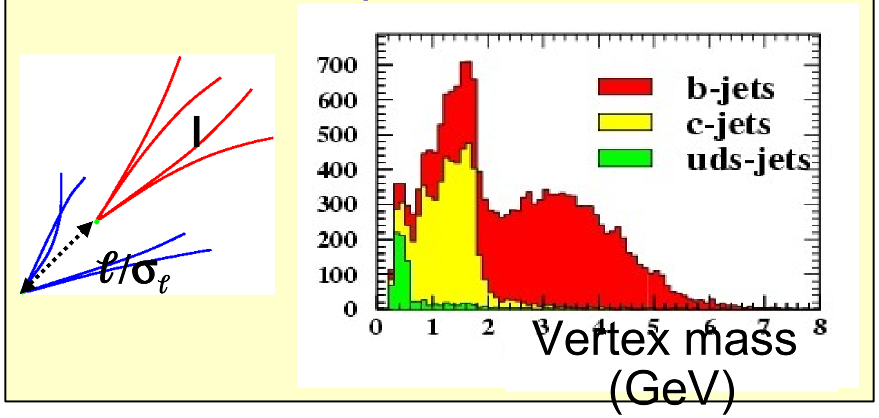


Vertexing: ZVTOP

- Vertex reconstruction
  - Tracks interpreted as probability tool
  - Vertex: crossing of this tubes
- Neural Net for flavor separation
  - B-net
  - C-net
- Training:
  - 91 GeV: qq
  - 350 GeV: HZ--> qqll
- Most important variable:
  - Vertex mass

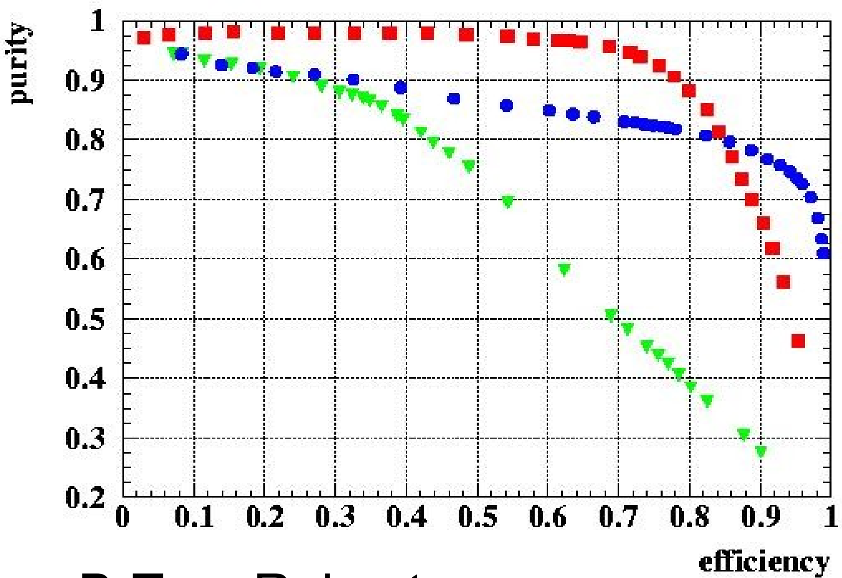


Neural Net input: Vertex mass

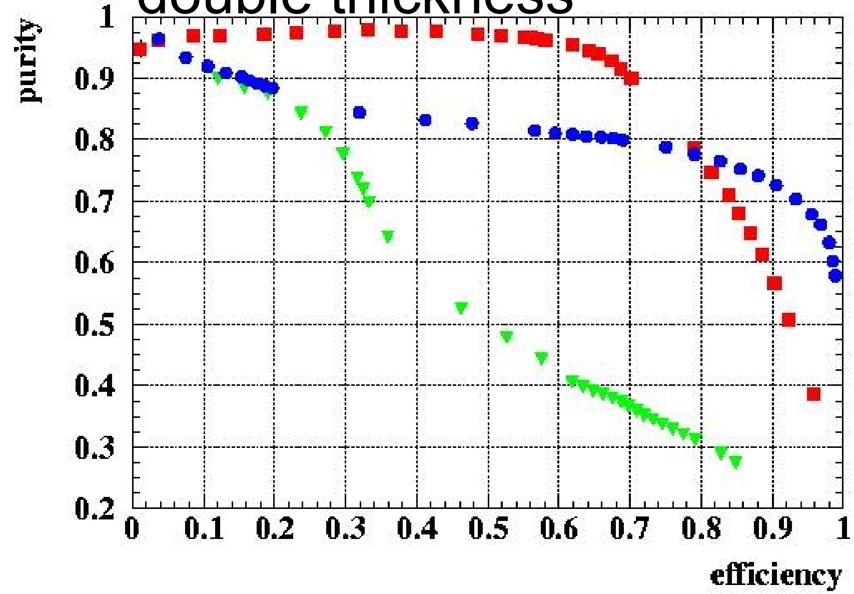


Detector dependence (Brahms)

- 5 layer: silicon detector



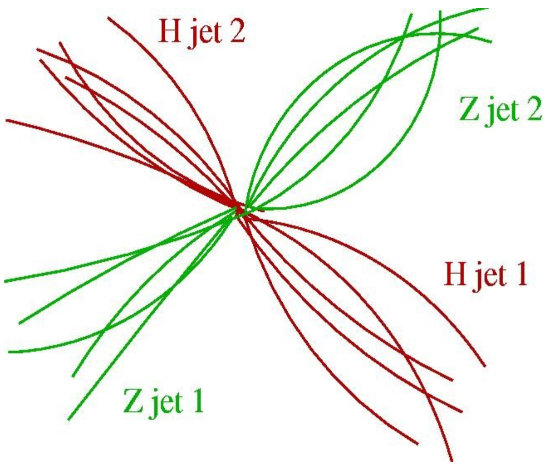
- 4 layer wo innermost layer+ double thickness



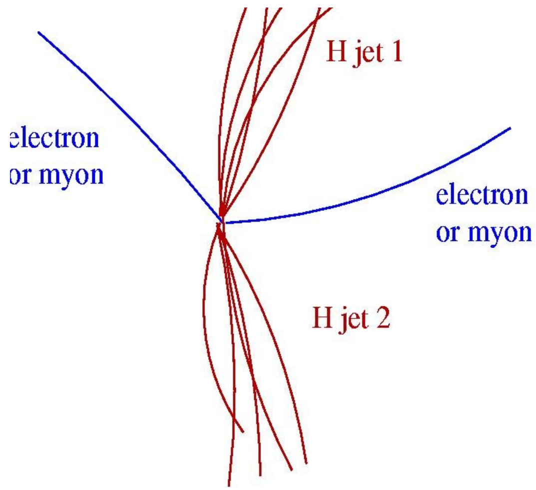
- B-Tag: Robust
- C-Tag suffers because loosing short lived D^0
  - Mostly effect of radius for innermost layer
  - Material is not that important for vertexing
- Jet-charged study: performance differences in the same order of magnitude

Higgs branching ratio analyis

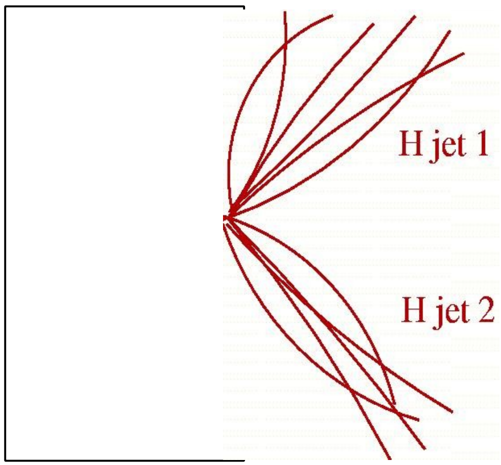
- Huge number of events ---> use fast simulation
- 3 signal topologies:



ZH$\rightarrow$qqqq



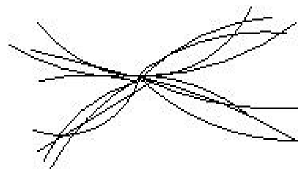
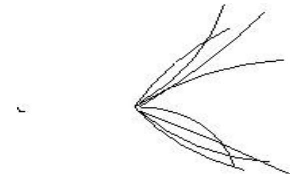
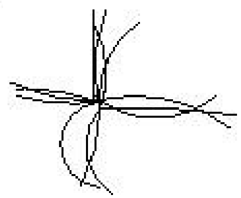
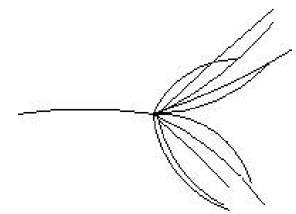
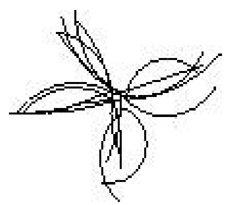
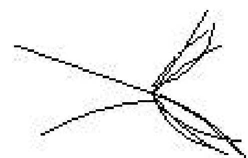
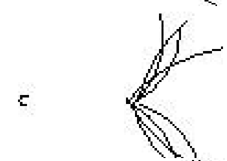
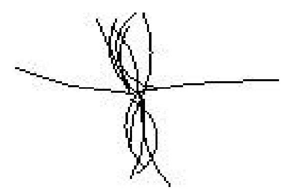
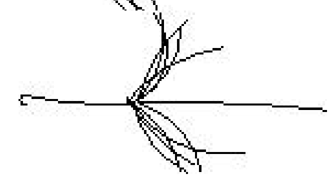
ZH$\rightarrow$qqll



ZH$\rightarrow\upsilon\upsilon$qq

- Strategy:
  - Divide in 3 samples
  - Cut based selection
  - Likelihood selection
  - Simultaneous fit of branching ratios

Background samples

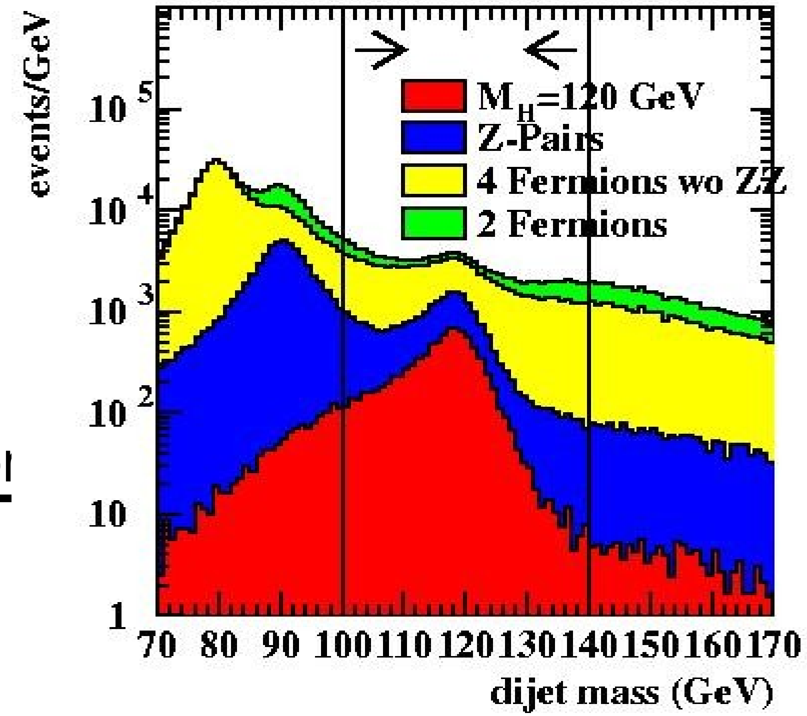
	500 fb^{-1}	HZ-->qqqq	HZ-->qqll	HZ-->qqυυ
2Fermion	14 M		NO	
W-Paare	6 M		Minor	
Z-Paare	0.2 M			
Single W/Z	8 M	Minor		

Reference distributions: 10 times this statistics !!!

Example: neutrino channel

- Most complexed analysis
- Preselection
  - Minimum number of tracks
  - No hard photon, remove soft ones
  - Raw and rescaled mass ($m_{recoil}=m_Z$)
  - Likelihood selection

	HZ	SM BG	WW-Fusio
Class	13955	8.6M	13150
Preselection	7158	84k	4321
Likelihood	3854	1777	358



Clean sample with moderate statistics!

ε_{signal} : 31%

$s/\sqrt{(s+b)}$: 50

Selection results

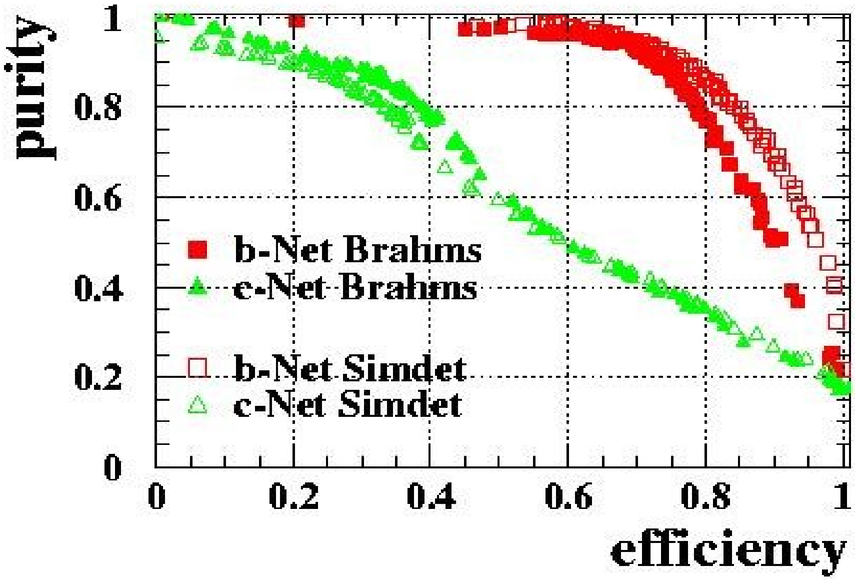
- Leptonic channel: Easy but no statistics
  - Tag Z-boson and look on Higgs properties
- Hadronic channel: Dirty but high statistics
  - Probelms because of confusion in jet finding, background is badly reconstrcuted qq events
- Neutrino channel: Complex medium statistic but clean sample
  - Sensitifity close to hadronic channel: Higgs background from WW-fusion

Channel	Total BG	HZ	Efficiency	s/√(s+b)
Z→ll	807	1873	39%	36
Z→qq	7973	8827	20%	68
Z→nn	1777	4212	31%	50

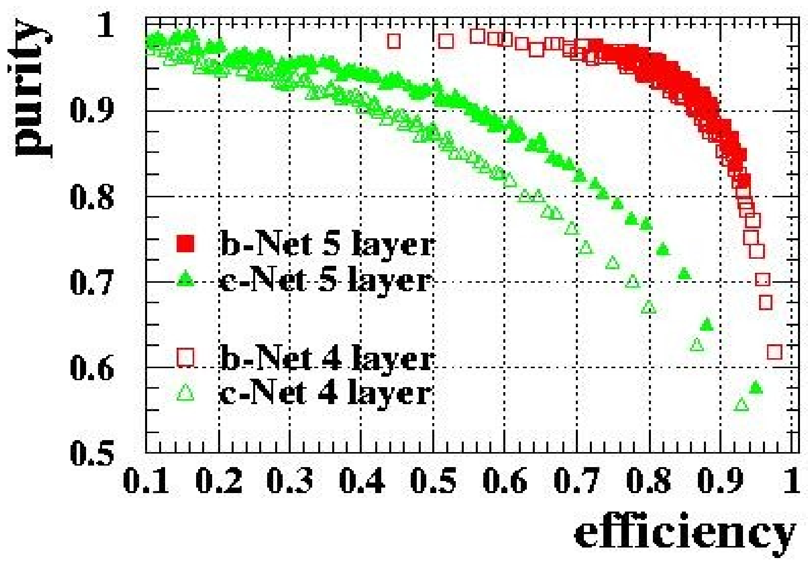
- Higgs backgrounds less than 15% of signal (but still too large!)
- Large improvement vs TDR in leptonic and neutrino channel

Neural Net performance

- Comparison of fast Simdet and full Brahms simulation
  - Simdet tunes with this Brahms version (91.2 GeV)
  - C-Tag: Very good agreement
  - B-Tag: reasonable, missing

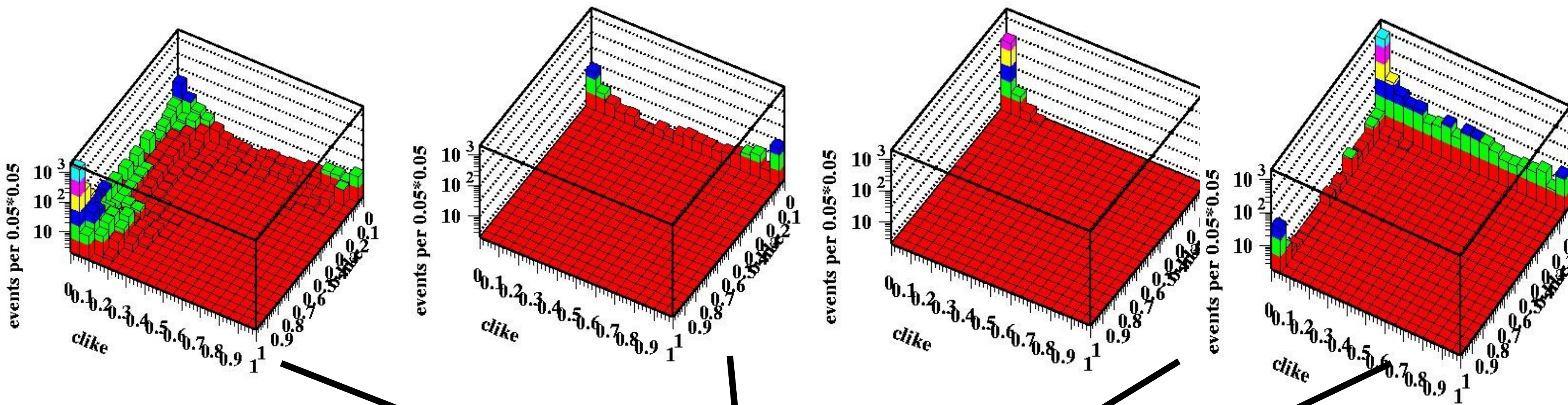


- Vertex detector performance
  - Two jet tag for Higgs events at 350 GeV
  - with/without innermost layer
  - B-Tag: very robust
  - C-tag: senistiv benchmark

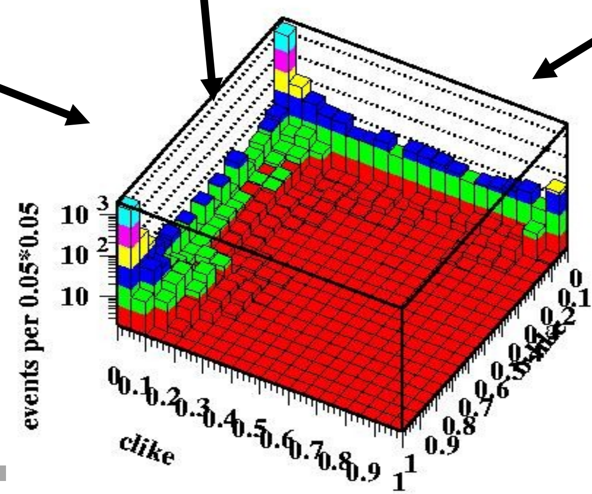


Braching ratio extraction

ZH→νν(ll,qq)bb+ZH→νν(ll,qq)cc +HZ→νν(ll,qq)gg   + background=



Binned maximum likelihood Fit of b-like/c-like distribution to 'Data'



Fit details:

- simultaneous fit of all distributions for all 3 felection channels (ll,qq,νν)
- 5ab^{-1} for reference distributions

Fit Results

- Combined result (all 3 channels, m_H=120 GeV, ECM=350 GeV, 500fb^{-1})
  - $\Delta(\sigma BR(H\rightarrow bb)/\sigma_{SM})$: 68.20± 0.75%
  - $\Delta(\sigma BR(H\rightarrow cc)/\sigma_{SM})$: 3.01 ± 0.36%
  - $\Delta(\sigma BR(H\rightarrow \upsilon\upsilon)/\sigma_{SM})$: 6.70 ± 0.55%

	Z→all	Z→qq	Z→ll	Z→νν
Δ(σ BR)/σ BR(bb)	1,1%	1,5%	3,0%	2,1%
Δ(σ BR)/σ BR(cc)	12,1%	17,5%	33,0%	20,5%
Δ(σ BR)/σ BR(gg)	8,3%	14,4%	18,5%	12,3%

- Error checked with 10 independent samples: no bias
- Different background composition for different channels is helpfull
- Missing: Corrrect treatment of other Higgs decay channels (WW, ττ, ZZ), currently treaten as fixed background
- No detector variation done (needs 300M events per set up)

Comparison with other analyses

	$\Delta(\sigma BR)/(\sigma BR)(bb)$	$\Delta(\sigma BR)/(\sigma BR)(cc)$	$(\sigma BR)/(\sigma BR)(gg)$
This analysis	1,1%	12,1%	8,3%
TDR(Battaglia) *	0,9%	8,0%	5,1%
Snowmass(Brau,Potter)	1,6%	19,0%	10,4%

- A lot of these anaysis were done scaling 1 or 2 channels
- TDR analysis:
  - Error 50% larger
  - But selection much more efficient in qqnn and qqll
  - Difference: optimistic jet-based flavor tag parametrisation based on monojets: no jet-jet confusion, no gluon splitting ...
- Snowmass (Brau/Potter)
  - Flavor tag a bit better (1.2cm inner radius)
  - Similar analysis using only qqqq and qqll
  - Cuts for flavor separation instead 2-dim fit
- J-C Brient (direct measurement from recoil events): error in binominal error treatment, extraction should be redo.

Conclusion

- Results for flavor tagging
- Detector:
  - Most important: innermost layer
- Flavor tag:
  - B-Tag robust (LEP Btagger as good as „Tesla“)
  - C-Tag is the benchmark!
- Use realsistic and comparable physics and tools
- Several analysis for higgs Branching ratios/jet flavour tagging done
  - Differences are well understood
  - Mostly differences in analysis technics
- No Beam related background studied (the biggest difference between different detector concepts ?)
- 5 year develoment (Hawkins, Xella, Wing, de Groot, Raspereza, Desch, Kuhl...)

Conclusion