Update on the Higgs self-coupling analysis.

Claude-Fabienne Dürig

ILD Analysis/Software Meeting, July 13th 2016

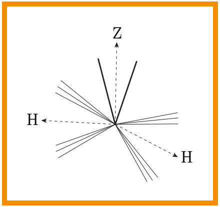
PIER
Eine Partnerschaft der Universität Hamburg und DESY

Universität Hamburg
DER FORSCHUNG | DER LEHRE | DER BILDUNG

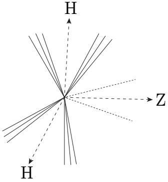
Analysis $e^+e^- \rightarrow$ ZHH (HH→bbbb) @ 500 GeV

- for m_H=125 GeV without and with overlay, with $\mathcal{L} = 2$ ab^{-1} and P(e^+e^-) = (0.3,-0.8)
- results presented at LCWS14: without overlay $\Delta\lambda = 53\%$
  with overlay $\Delta\lambda = 58\%$
- 20% relative improvement possible by improving analysis methods
- results are presented for analysis without overlay
- last updates: phone meeting Oct.28th 2015, update by M. Kurata at LCWS2015
- → Kinematic fits and ISR treatment in context of heavy flavoured jets (semi-leptonic decays)
- → event-specific ISR treatment, allows working around issues with ISR consideration
- → semi-leptonic energy correction to correct for missing energy from neutrinos

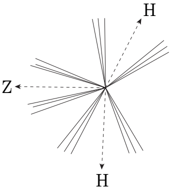
$\rightarrow l^-l^+$HH



$\rightarrow \nu\nu$HH



$\rightarrow qq$HH



ZHH @ 500 GeV - Kinematic Fitting in Analysis

slide from presentation at ALCW15 at KEK

- **useful in current analysis strategy**

  mass distributions as input for neural net training

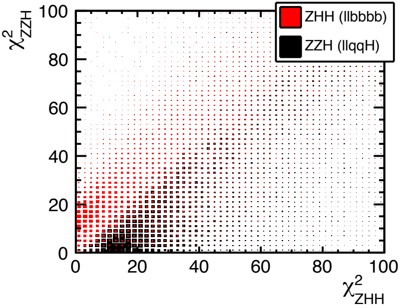
  → ZHH vs. ZZH/ZZZ
  (same final state as signal)

- **change analysis strategy**

  → test different fit hypotheses for signal and backgrounds

  → use information obtained from fit ($\chi^2, P(\chi^2)$)

  → add to analysis (neural net input or cut)



preliminary results with kinematic fit applied to llHH search mode (no overlay)

	decay mode	signal	bgrd	significance	
				excess	measurement
LCWS14	l^-l^+HH	3.0	4.3	1.16σ	0.91σ
		3.3	6.0	1.12σ	0.91σ
ALCW15		3.0	3.2	1.32σ	0.99σ
		4.4	6.8	1.43σ	1.18σ

20% improvement in llHH mode due to kinematic fitting!

Kinematic Fitting - ISR Treatment

slide from presentation ILD phone meeting Oct. 28th 2015

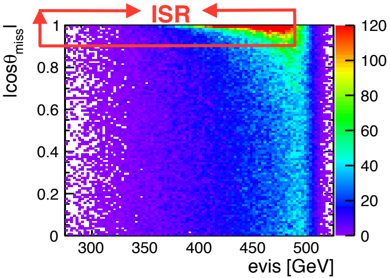
so far: $\sum \left(E_i, \vec{P}_i\right) = \left(\sqrt{s}, \vec{0}\right)$

ISR and beamstrahlung in samples

$$\sum \left(E_i, \vec{P}_i\right) = \left(\sqrt{s} - E_\gamma^{ISR}, -\vec{P}_\gamma^{ISR}\right)$$

considering ISR in fit (ISRPhotonFitobject)

- works well for light jets (diploma thesis M.Beckmann)
- in b/c jets E_{miss} due to ν → special handling



ISR not considered in fit:

- energy assigned to jets (E, $\vec{P}$ conserved)

Problem: events with ISR

- larger fitted 4-momenta of jets
- bias to large masses

ISR considered in fit:

- certain amount of energy assigned to ISR

Problem: events without ISR

- "fake" ISR, energy missing to jets
- bias to small masses

Correct treatment of ISR in fit on events by events basis using ISR characteristics

Kinematic Fitting - llHH (HH→ $b\bar{b}b\bar{b}$) - Summary

slide from presentation ILD phone meeting Oct. 28th 2015

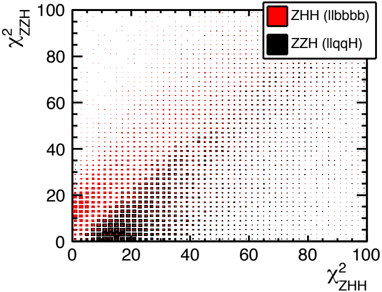
- **change analysis strategy**
  - → test different fit hypotheses for signal and backgrounds
  - → use information obtained from fit (χ^2, P(χ^2))

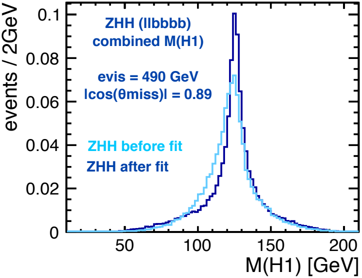
20% improvement in llHH mode due to usage of χ^2 in neural nets!

- **mass distributions significantly improved**
  - → optimised strategy for ISR treatment

further improvement by using new mass reconstruction in neural nets expected

Special ISR treatment still necessary when we correct E_{miss} from semi-leptonic b/c decays?

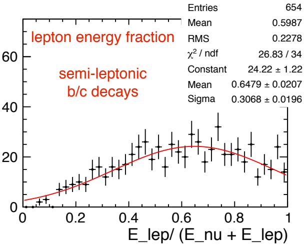




Impact of Semi-leptonic b/c Decays

slide from presentation ILD phone meeting Oct. 28th 2015

if lepton in jet, correct missing energy from neutrino: $E_{jet}^{corr} = E_{jet} + E_{\nu} = E_{jet} + x \cdot E_{lep}$



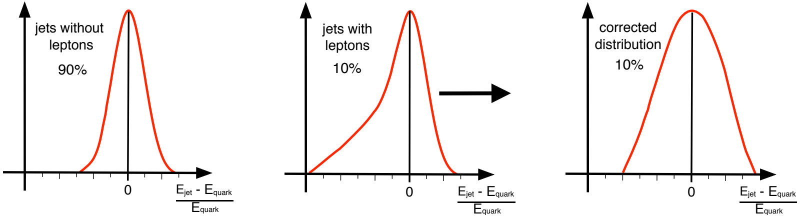
$$x = \frac{E_{lep}}{(E_{lep} + E_{\nu})} \Leftrightarrow E_{\nu} = \left(\frac{1}{x} - 1\right) E_{lep}$$

$$E_{jet}^{corr} = E_{jet} + \left(\frac{1}{\langle x \rangle} - 1\right) E_{lep}$$

$$\Delta E_{jet,corr}^2 = \Delta E_{jet}^2 + \left(\frac{\Delta \langle x \rangle}{\langle x \rangle^2}\right)^2 E_{lep}^2 + \underbrace{\cancelto{0}{\left(\frac{1}{\langle x \rangle} - 1\right) \Delta E_{lep}^2}}_{\text{track momentum resolution!}}$$

plot: $\langle x \rangle = 0.65 \rightarrow E_{jet}^{corr} = E_{jet} + 0.54 E_{lep}$

$\Delta \langle x \rangle = 0.31 \rightarrow \Delta E_{jet,corr}^2 = \Delta E_{jet}^2 + (0.73 E_{lep})^2$



Kinematic Fitting - llHH - Semi-leptonic b/c Decays

slide from presentation ILD phone meeting Oct. 28th 2015

- **add correction to kinematic fit**
  - ☆ plots correspond to fit with E, P, soft M_Z, and equal-mass constraints
  - ☆ ISR not considered in fit

- **jets without leptons**

$$E_{jet} = E_{jet}$$

$$\Delta E_{jet} = \sqrt{(0.568)^2 E_{jet} + (3.301)^2}$$

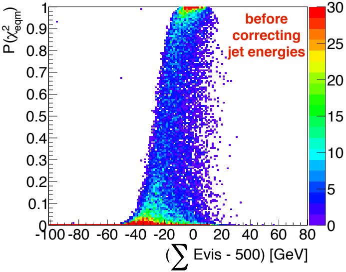
(σ_E from PhD thesis B. Hermberg)

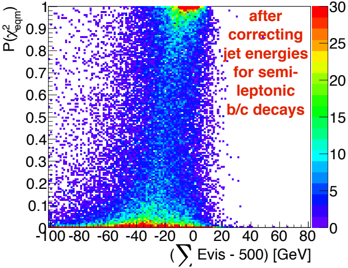
- **jets with leptons**

$$E_{jet}^{corr} = E_{jet} + 0.54 E_{lep}$$

$$\Delta E_{jet}^{corr} = \sqrt{(0.568)^2 E_{jet} + (3.301)^2 + (0.73 E_{lep})^2}$$

Improved $P(\chi^2)$ for events with E_{miss} by compensating semi-leptonic b/c decays in jets!

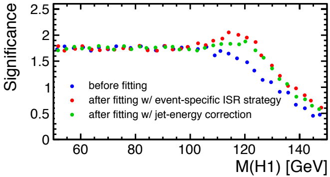




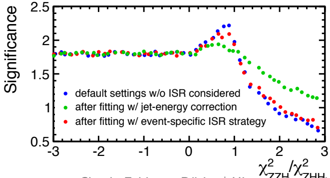
Kinematic Fitting - llHH - Semi-leptonic b/c Decays

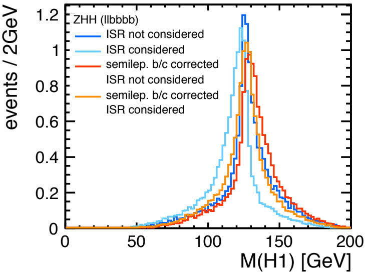
Correcting jet energies for semi-leptonic b/c decays enables automatic ISR recognition in fit as for light jets

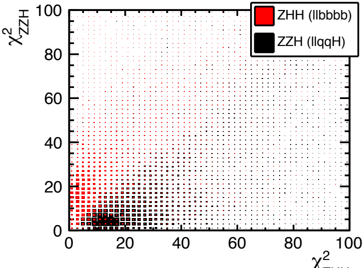
lower mass cut on signal and ZZH/Z



cut around diagonal for signal and ZZH/Z



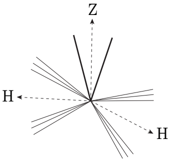




Analysis strategy $e^+e^- \to$ ZHH at $\sqrt{s} = 500$ GeV

Perform analysis for $m_H = 125$ GeV, assuming $\mathcal{L} = 2\ ab^{-1}$ of data, and $P(e^+e^-) = (0.3, -0.8)$

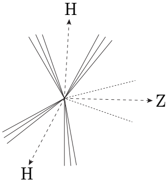
$\to l^- l^+$HH



Z $\to l\bar{l}$, HH $\to b\bar{b}b\bar{b}$

(10% × 60% × 60% ≈ 3.6%)

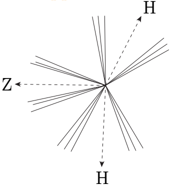
$\to \nu\nu$HH



Z $\to \nu\bar{\nu}$, HH $\to b\bar{b}b\bar{b}$

(20% × 60% × 60% ≈ 7.2%)

$\to qq$HH



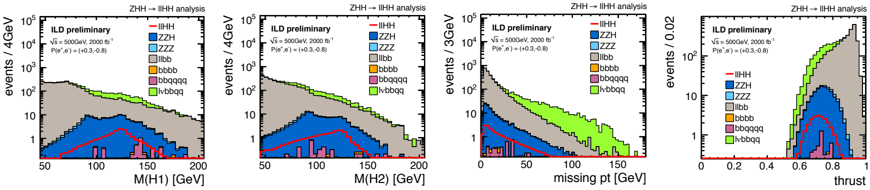
Z $\to q\bar{q}$, HH $\to b\bar{b}b\bar{b}$

(70% × 60% × 60% ≈ 25%)

Event selection:

① isolated lepton selection or rejection
② cluster particles into jets and get flavour tag information
③ pair jets to form signal bosons (smallest χ^2)
④ each dominant background is suppressed by training a separate MVA
⑤ additional precuts before MVA training

Preselection - Lepton Channel



- ➤ select events w/ isolated lepton pair consistent with $|M_{ll} - M(Z)| < 40$ GeV
- ➤ pair jets to form signal bosons $|M_{jj} - M(H)| < 80$ GeV

	$eebb$	$\mu\mu bb$	$e\nu bbqq$	$\mu\nu bbqq$	$\tau\nu bbqq$	$bbqqqq$	$bbbb$	$llbbbb$	$llqqH$	**Background**	**$llHH$ ($llbbbb$)**
expected events	$2.84 \cdot 10^5$	$4.95 \cdot 10^4$	$2.48 \cdot 10^5$	$2.46 \cdot 10^5$	$2.46 \cdot 10^5$	$6.24 \cdot 10^5$	$4.02 \cdot 10^4$	69.51	150.87	$1.73 \cdot 10^6$	40.51 (14.3)
$N_{isolep} >= 2$	$6.4 \cdot 10^4 \pm 78$	$2.1 \cdot 10^4 \pm 37$	1911 ± 22	226 ± 7	195 ± 5.9	25.5 ± 2.0	2.4 ± 0.3	21.8 ± 0.1	135 ± 0.5	$8.8 \cdot 10^4 \pm 89$	25.2 ± 0.07 (7.9)
...	....										
...	....										
thrust <0.9	603 ± 8	288 ± 4	341 ± 9	46 ± 3	12 ± 1.5	3.7 ± 0.8	0.2 ± 0.1	11.5 ± 0.1	119 ± 0.4	1424 ± 13	21.4 ± 0.06 (7.7)

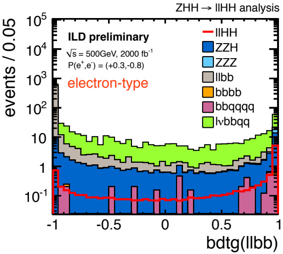
- ➤ isolated lepton tagger optimised for e/μ

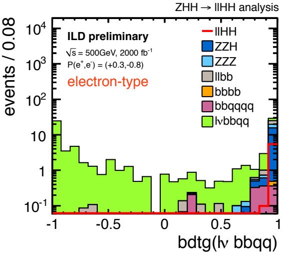
→ algorithm for hadronic τ finder (less than 10% decay leptonically)

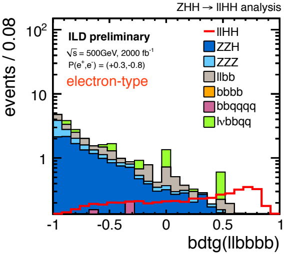
→ allows selection of $\tau\tau HH$ and supression of $\tau\nu bbqq$

→ if similar final result as in $eeHH$ and $\mu\mu HH$ then possible 5% improvement of lepton channel

Event Selection - Lepton Channel







- split into electron and muon types
- combined BDTGs for both lepton types

→ cuts optimised separately $\frac{s}{\sqrt{s+b}}$

→ trained sequentially

- MC statistics issues

→ $l\nu bbqq$: suppressed by preselection, large anyway

→ relative error: 10% for $e\nu bbqq$ (5% need $\times 8$)
20% for $\mu\nu bbqq$ (5% need $\times 30$)
35% for $\tau\nu bbqq$ (5% need $\times 180$)

- $bdtg(llbbbb)$ trained using variables from kinfit
- χ^2 most important, $M(H)$ w/o fit → no correlation

- BDTGs compared to MLPs 10% improvement
- χ^2 distributions improve classifier output by 5%
- masses w/o fit avoid correlations
- $\sim 20\%$ improvement in $llHH$
- more statistics needed ($l\nu bbqq$)

Event Selection - Lepton Channel

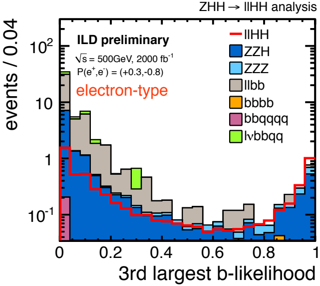
Flavour-tag information

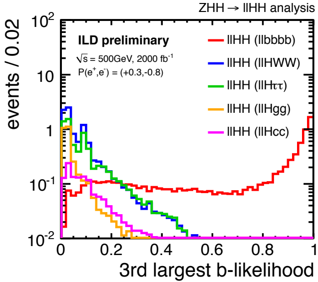
→ third largest b likelihood

→ no effect on background suppression

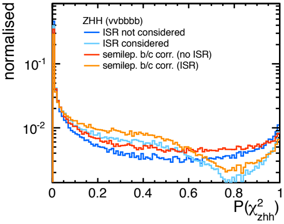
- **flavour tag of signal**
  - → at least one Higgs into other than bb
  - → $H \rightarrow WW, \tau\tau$ at small values
  - → $H \rightarrow bb$ at large values

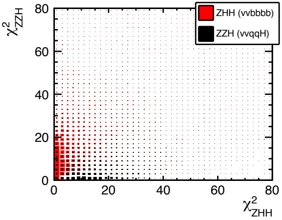
Lepton channel strategy can provide foundation for further investigations of the other Higgs decays

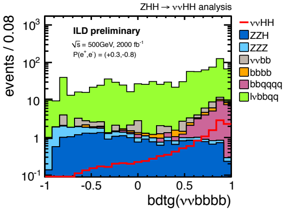




Event Selection - Neutrino Channel







- $N_{\text{iso}} = 0$, bmax3$>$0.2, E_{vis} requirement and known precuts

- BDTGs trained, cuts optimised separately

→ $bbbb$ vs ZHH $(\nu\nu bbbb)$
→ $l\nu bbqq$ vs ZHH $(\nu\nu bbbb)$ (mainly $\tau\nu bbqq$)
→ $\nu\nu bbbb$ vs ZHH $(\nu\nu bbbb)$ (kinematic fit variables)

- statistics quite okay

- Kinematic fits possible due to Zinvisible fit object!!!

→ automatic ISR recognition not possible
→ distortions in $P(\chi^2)$
→ E_{miss} requires further developments for ISR treatment

- BDTGs in neutrino channel give 5% improvement

- E and P constraint not possible
- developed Zinvisible fit object

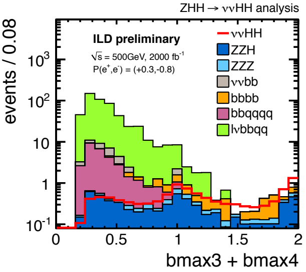
→ works like neutrino fit object
→ px, py, pz, E from $\vec{P}_{\text{miss}}$
→ mass set to M_Z

- $N_{\text{dof}} = N_{\text{constr.}} - N_{\text{unmeas.}}$
- Z from νν → 6× unmeas.
  → fit ✗
- Zinvisible FO → 3× unmeas.
  → fit ✓

Event Selection - Neutrino Channel

	$\nu\nu bb$	$e\nu bbqq$	$\mu\nu bbqq$	$\tau\nu bbqq$	$bbqqqq$	$bbbb$	$\nu\nu bbbb$	$\nu\nu qqH$	**Background**	$\boldsymbol{\nu\nu HH}$ **($\boldsymbol{\nu\nu bbbb}$)**
expected events	$2.73 \cdot 10^5$	$2.49 \cdot 10^5$	$2.46 \cdot 10^5$	$2.46 \cdot 10^5$	$6.24 \cdot 10^5$	$4.02 \cdot 10^4$	96.83	447.0	$1.68 \cdot 10^6$	80.14 (28.5)
preselection	161 ± 10	1073 ± 16	979 ± 15	$2.5 \cdot 10^4 \pm 94$	1799 ± 17	1656 ± 11	56 ± 0.3	75 ± 0.5	$3.0 \cdot 10^4 \pm 99$	25.7 ± 0.1 (21.5)
bdtg($bbbb$)>0.94	77 ± 7	621 ± 13	569 ± 11	$1.3 \cdot 10^4 \pm 69$	84 ± 4	17 ± 0.9	23 ± 0.2	49 ± 0.4	$1.5 \cdot 10^4 \pm 71$	18.8 ± 0.08 (17.1)
bdtg($l\nu bbqq$)>0.67	18 ± 3	40 ± 3	62 ± 4	716 ± 16	28 ± 2	7 ± 0.7	10 ± 0.1	22 ± 0.3	902 ± 17	13.3 ± 0.07 (11.9)
bdtg($\nu\nu bbbb$)>0.3	10 ± 2	25 ± 3	36 ± 3	414 ± 12	27 ± 2	5 ± 0.6	1 ± 0.05	6 ± 0.1	525 ± 13	10.5 ± 0.06 (9.5)
b3 + b4 > 1.08	0	0	0	2.5 ± 0.9	0	2 ± 0.4	0.5 ± 0.03	2 ± 0.1	7 ± 1	5.6 ± 0.04 (5.5)

- Large background after selection using BDTG
  - → flavour tagging crucial
  - → $bmax3 + bmax4 > 1.08$
  - → $\tau\nu bbqq$ suppressed by 99%
- Isolated-lepton tagger optimised for e and μ
  - → $\tau\nu bbqq$ hardy suppressed by isolated-lepton veto
  - → mainly $\tau\nu bbcs$ → τ and c affect flavour tagging
  - → less than 10% from leptonic τ decays
  - → algorithm to identify hadronic τ decays needed

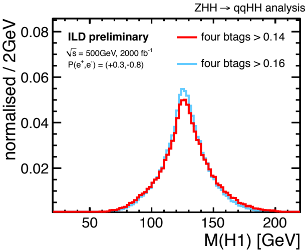


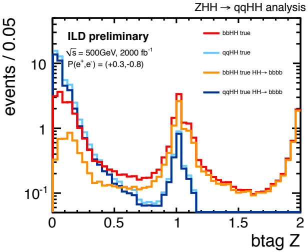
Event Selection - Hadron Channel

- very challenging
  - → ZHH → 45 permutations
  - → difficulties to get correct permutation
  - → prone to errors from combinatorics
  - → b tag requirement supports χ^2 pairing and kinematic fit
  - → at least four jets with $btag > 0.16$
- very strict $\rightarrow \sim 15\%$ of $ZHH \rightarrow qqbbbb$ events rejected
- more loose requirement would:
  - → more background could pass requirement
  - → more $H \nrightarrow bb$ pass, degraded mass resolution
  - → relative gain in signal events lost in event selection
  - → since larger backgrounds need tighter cuts
  - → tighter cuts 20% degradation

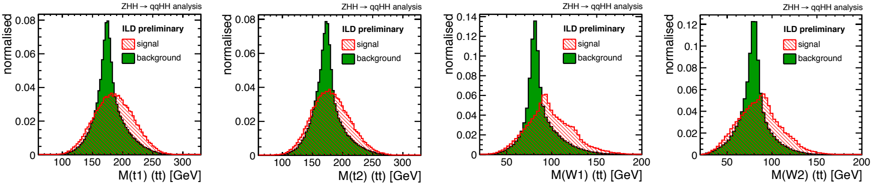
- same investigations as for llHH
- Kinematic fit allows ISR treatment
- slight improvements with semi-leptonic energy correction
- dominated by mis-clustering

- in selection: split into two categories according to $b$tag of Z





Event Selection - Hadron Channel



- Hadron channel event selection similar to previous channels
  - → $bdtg(bbbb)$
  - → $bdtg(bbqqqq)$
  - → $bdtg(qqbbbb)$
- limited statistics for $tt \rightarrow bbqqqq$ at the training stage of corresponding BDTG
  - → relative MC statistical error of 18% on $bbuddu$
- in $bdtg(bbqqqq)$ and $bdtg(qqbbbb)$ most important inputs are mass distributions
  - → $tt \rightarrow l\nu/bbqqqq$ dominant in all three signal channels
  - → kinematic fits not investigated for tt
  - → χ^2 could improve BDTG output by $\sim 5\%$ as in lepton channel
  - → could enhance the significance of hadron channel by $\sim 4\%$

Event Selection - Summary

- MVA Classifier: ☆ use of BDTGs gives 10% improvement in $llHH$ and 5% in $\nu\nu HH$
  - ☆ less sensitive to overtraining compared to MLP
  - ☆ BDTGs are trained sequentially → MC statistics very limited ($tt \to bbqqqq/l\nu bbqq$)
- Kinematic fit and semi-leptonic energy correction:
  - ☆ boson masses crucial for background suppression
  - ☆ energy correction improved jet four-momenta and jet pairing
  - ☆ χ^2 from kinematic fit improves BDTG response by 5%
  - ☆ masses with fitting show enhanced separation power
  - ☆ correlations degrade χ^2 power in BDTG → masses w/o fitting used
  - ☆ future task: kinematic fit for tt backgrounds → could improve $qqHH$ selection by 4%
- Isolated lepton tagging: ☆ efficiency to select isolated lepton pairs 90%
  - ☆ veto efficiency for events with one isolated lepton 90%
  - ☆ not optimised for τ identification
  - ☆ $\tau \to$ leptonically $< 10\%$ → algorithm to find hadronic τ can improve selection
    - → by 5% in $\nu\nu HH$ and $qqHH$
    - → $\tau\tau HH$ do not pass iso. lep. requirement in $llHH$
    - → by 5% if included and results as in $ee/\mu\mu HH$
- Flavour tagging: ☆ optional in $llHH$ → foundation for investigation of other Higgs decays channels
  - ☆ in $\nu\nu HH$ crucial to reject background, especially $\tau\nu bbcs$
  - ☆ in $qqHH$ used in jet pairing

Analysis Results - No Overlay

Results without overlay

decay channel	signal	background	significance	
			excess	measurement
ZHH → $l^- l^+$HH	3.9 ± 0.03 (2.6)	7 ± 0.6	1.29σ	1.07σ
	5.1 ± 0.03 (2.8)	9 ± 0.5	1.48σ	1.26σ
ZHH → $\nu\bar{\nu}$HH	5.6 ± 0.04 (5.5)	7 ± 1.0	1.78σ	1.50σ
ZHH → $q\bar{q}$HH	8.5 ± 0.1 (8.0)	22 ± 1.3	1.75σ	1.57σ
	12.6 ± 0.1 (10.9)	55 ± 2.0	1.65σ	1.55σ

cross section: $\frac{\Delta\sigma_{\text{ZHH}}}{\sigma_{\text{ZHH}}} = 30.3\%$

Higgs self-coupling: $\frac{\Delta\lambda}{\lambda} = 49.1\%$

→ by including analysis improvements cross section is improved by 10%

→ this results in 8% improvement of λ_{SM}

Results with overlay

decay channel	signal	background	significance	
			excess	measurement
ZHH → $l^- l^+$HH	3.5 ± 0.02 (2.4)	10 ± 0.8	1.10σ	0.93σ
	3.9 ± 0.03 (2.5)	7 ± 0.4	1.29σ	1.07σ
ZHH → $\nu\bar{\nu}$HH	5.2 ± 0.04 (5.1)	8 ± 1.0	1.57σ	1.34σ
ZHH → $q\bar{q}$HH	8.5 ± 0.06 (8.0)	23 ± 1.4	1.61σ	1.46σ
	6.8 ± 0.06 (6.1)	33 ± 1.8	1.13σ	1.03σ

cross section: $\frac{\Delta\sigma_{\text{ZHH}}}{\sigma_{\text{ZHH}}} = 35.8\%$

Higgs self-coupling: $\frac{\Delta\lambda}{\lambda} = 58\%$

→ degradation of 15%

→ results have to be taken with care, since modelling and understanding the overlay in ongoing

→ more sophisticated tools needed

BACKUP SLIDES

Jet energy resolution on ZHH events

PhD Thesis B.Hermberg (Defence)

Jet parametrization and jet resolution:

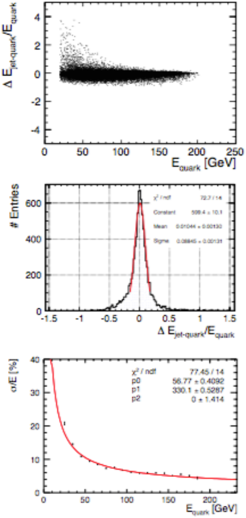
- Hadronic jets are parametrized with E, θ, ϕ
- Angular dependence is neglected

Determination of jet energy resolution:

- Comparing energy of primary quarks with jet energy
- Splitting energy range in 25 slices
- Fitting each slice with gaussian in core area determines σ
- Non-gaussian tails are caused by semi-leptonic decays

Consequences:

- For lower energies, jets are less collimated (jet mis-clustering error is dominant)
- For higher energies, jets are stronger collimated (smaller error in jet mis-clustering)



Application of Kinematic Fits

- tool to improve jet energy and invariant mass resolution
- precise initial states at ILC are ideal for application of kinematic fits
- number of 4-vectors, representing final state particles, is fitted under constraints

$$\begin{aligned}&E,\ \vec{P}\ \text{conservation}\\ &\text{invariant mass: } M_{ij} = X\ \text{GeV},\\ &\qquad\qquad\qquad M_{ij} = M_{lk},\\ &\qquad\qquad\qquad M_{ij} = (X \pm Y)\ \text{GeV}\end{aligned}$$

- variation of measured quantities → with respect to allowed uncertainties
  → under condition that constraints are fulfilled

- minimise deviation between measured and fitted parameters by

$$\chi^2 = (\eta - a)^T C^{-1} (\eta - a)$$

a = vector of measured quantity
η = vector of varied quantity
C = covariance matrix

- C^{-1} contains uncertainties of measured parameters

Higgs Self-Coupling Analysis Strategy

Strategic difficulties:

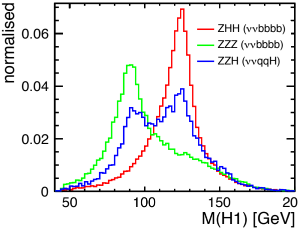
- very small number of signal events
- irreducible diagrams: degrade self-coupling sensitivity
- flavour tagging and isolated lepton selection: high efficiency and purity
- Higgs mass reconstruction: misclustering, wrong jet pairing
- neural net training: separate neural nets, need large statistics

there are several key points for potential improvement in analysis
expect 20%* improvement of Higgs self-coupling precision

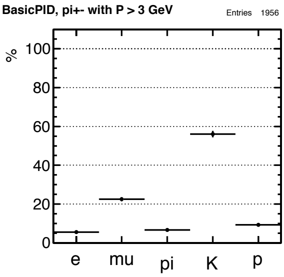
$\star$ with respect to Snowmass

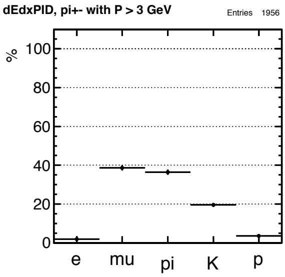
One possibility is use of kinematic fits:

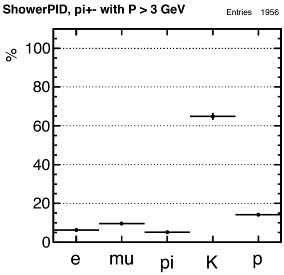
- find right jet pairing
  → improved jet/boson assignment
- improved mass resolution
  → clearer signal/background separation
- testing of fit hypothesis
  → fit information for signal/background separation

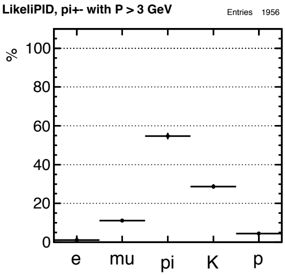


Fake Rate


BasicPID, pi+- with P > 3 GeV
Entries 1956
%
0
20
40
60
80
100
e
mu
pi
K
p

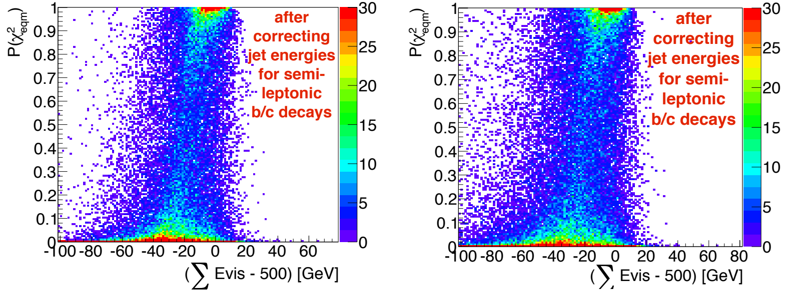

dEdxPID, pi+- with P > 3 GeV
Entries 1956
%
0
20
40
60
80
100
e
mu
pi
K
p


ShowerPID, pi+- with P > 3 GeV
Entries 1956
%
0
20
40
60
80
100
e
mu
pi
K
p

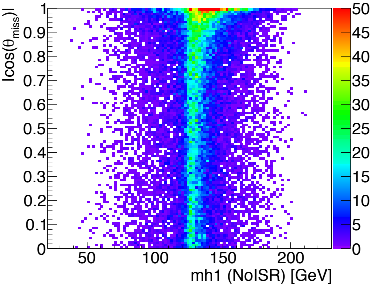

LikeliPID, pi+- with P > 3 GeV
Entries 1956
%
0
20
40
60
80
100
e
mu
pi
K
p

cheated leptons and mum

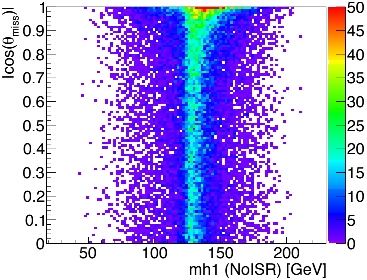
cheated leptons and p>3 GeV



cheated leptons and mum

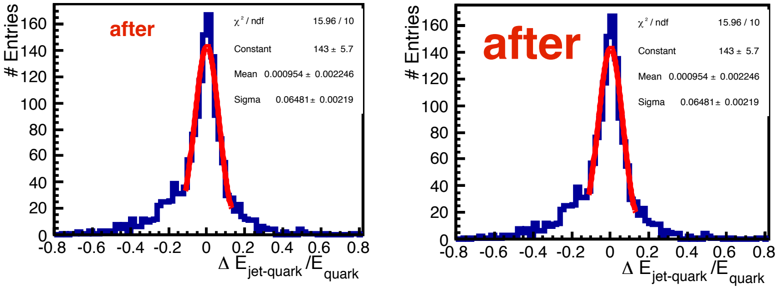


cheated leptons and p>3 GeV

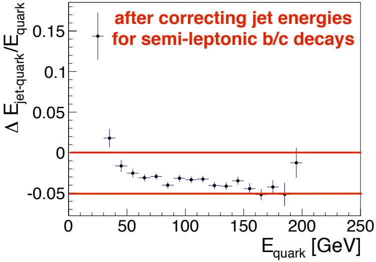


cheated leptons and mum

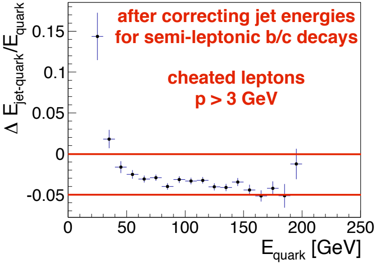
cheated leptons and p>3 GeV

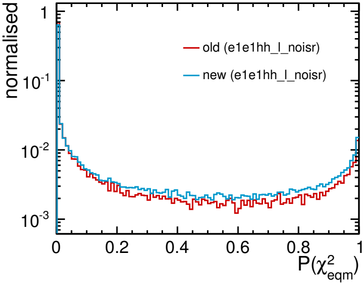


cheated leptons and mum

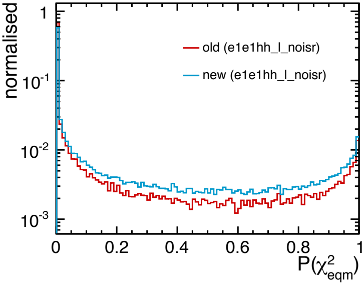


cheated leptons and p>3 GeV

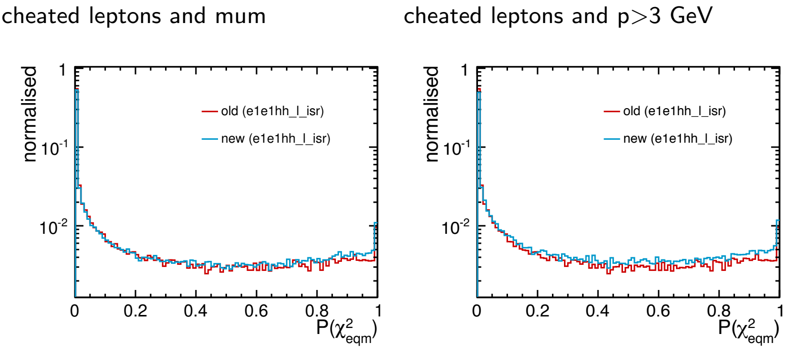




cheated leptons and mum

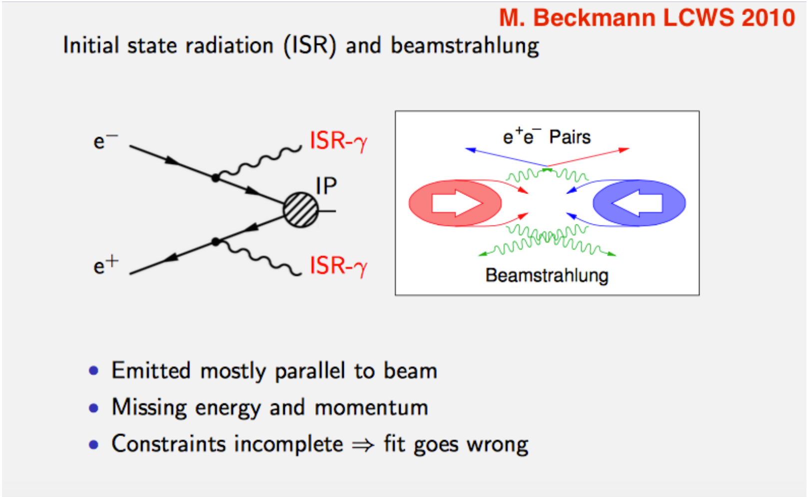


cheated leptons and p>3 GeV


cheated leptons and mum

cheated leptons and p>3 GeV

ISR



M. Beckmann LCWS 2010

Initial state radiation (ISR) and beamstrahlung

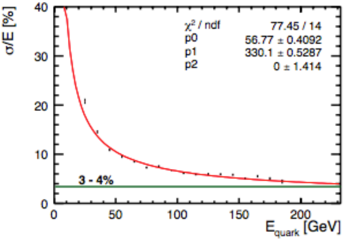
- Emitted mostly parallel to beam
- Missing energy and momentum
- Constraints incomplete ⇒ fit goes wrong

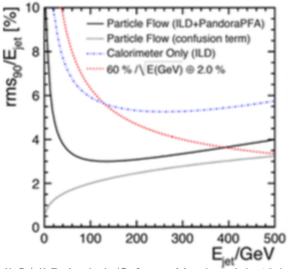
Jet energy resolution on ZHH events

PhD Thesis B. Hermberg (Defence)

Comparison with Particle Flow Benchmarks:

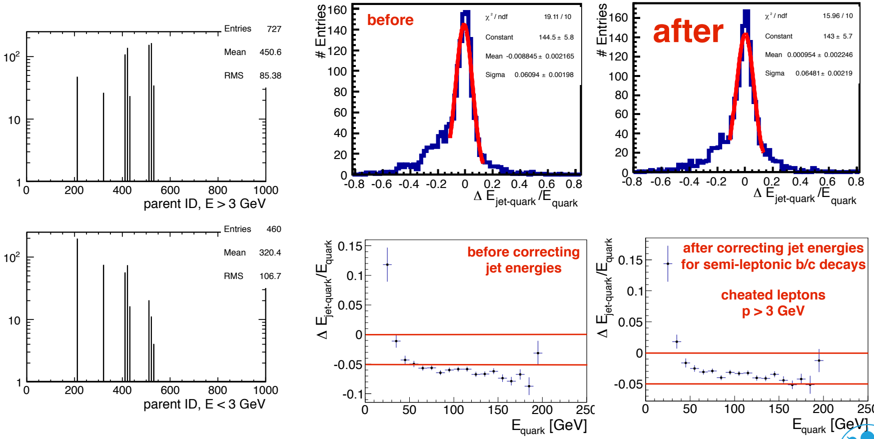
- Particle Flow considers only $e^+e^- \rightarrow u, d, s$
- Total energy is taken and divided by 2 (no need for jet finding)
- Use of RMS90 (exclude long tails in energy distribution) → RMS90 is ~20% smaller for a gaussian compared to σ
- Jet energy resolution includes additional effects (semi-leptonic decays, jet-misclustering)
  → At ~200 GeV it gets comparable to Particle Flow



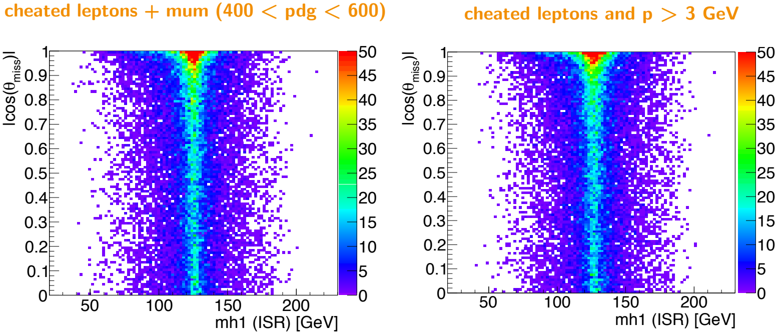


Kinematic Fitting - llHH - Semi-leptonic b/c Decays

remark: samples with standard DBD reconstruction → ParticleID tools not included
following studies with cheated e/μ, with $p > 3$ GeV



Kinematic Fitting - Semi-leptonic b/c Decays



identification of semi-leptonic b/c decays not an issue if leptons can be found - even w/o impact parameter/ secondary vertex.

(in backup more plots for cheated leptons + cheated mum ($400 < pdg < 600$))

Kinematic Fitting - Semi-leptonic b/c Decays

samples with standard DBD reconstruction → ParticleID tools not included
shown studies with cheated e/μ, with $p > 3$ GeV

benefit from new ParticleID tools in new reconstruction to get best result

electrons

muons