Reconstruction of b- and c- jets at e^+e^- Higgs Factories with ParticleFlow detectors

DPG Spring Meeting, Dortmund

Yasser Radkhorrami[1,2], Jenny List [1]

[1] DESY, Hamburg
[2] Universität Hamburg, Hamburg

March 16, 2021

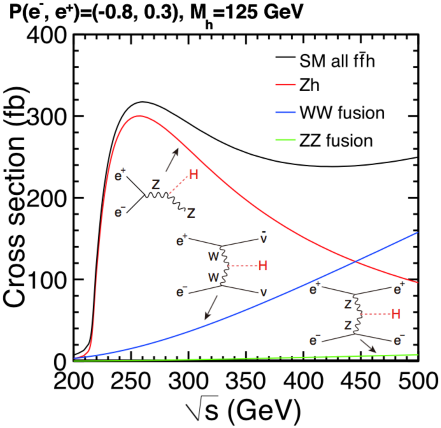
Universität Hamburg
DER FORSCHUNG | DER LEHRE | DER BILDUNG

CLUSTER OF EXCELLENCE
QUANTUM UNIVERSE

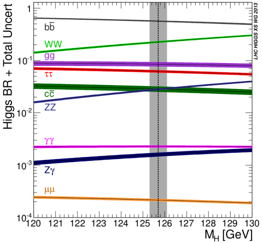
Outline

Higgs production mechanisms and decay modes at e^+e^- colliders

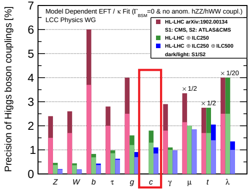
- Higgs strahlung is dominant Higgs production mechanism at 250 GeV



- Most frequent Higgs decay mode: $H \to b\bar{b}$



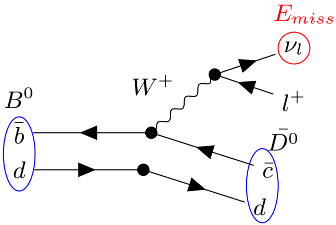
- Extremly challenging in Hadron colliders: $H \to c\bar{c}$

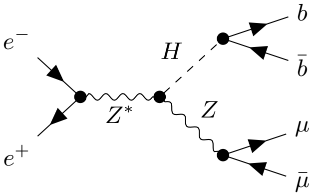


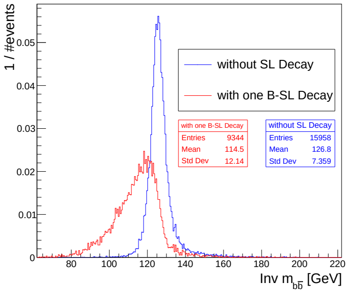
Semi-leptonic b / c decays

- Number of B-/C-hadron semileptonic decays (SLD) in $e^+e^- \to b\bar{b}$ events

		nBSLD		
		0	1	2
nCSLD	0	34%	24%	4%
	1	18%	12%	2%
	2	3%	2%	0%

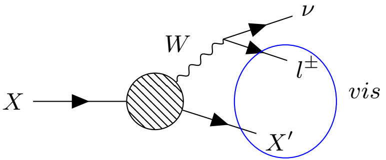






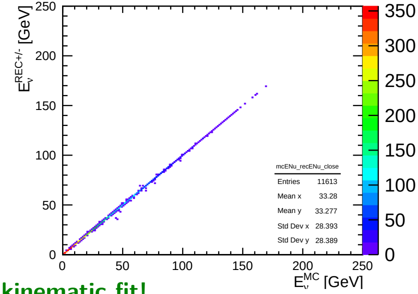
- Mis-reconstruction of $b\bar{b}$ invariant mass due to missing neutrino energy from semi-leptonic decays
  - Can the missing momentum be retrieved from event and decay kinematics?

Concept of ν-correction in a semi-leptonic decay



- Find heavy-quark jets: Identify b or c jet $\rightarrow$ flavour tag
- Find semi-leptonic decay(s): Identify lepton in jet if present $\rightarrow$ possible using detector's high granularity
- Estimate neutrino energy from decay kinematics:
  - Assign B^0 or D^0 meson mass to mother hadron.
  - Reconstruct flight direction of mother hadron from position of primary and secondary vertex.
  - Calculate neutrino momentum: up to 3-fold ambiguity.
- As proof-of-principle: CHEAT from MC truth
  - Lepton ID
  - Flavour tag
  - Mother hadron mass
  - **Associate of reconstructed particles to secondary vertex**
  - Momenta of visible decay products

- Closure test: fully cheated information ($e^+e^- \rightarrow b\bar{b}$ at $\sqrt{s} = 500$ GeV)

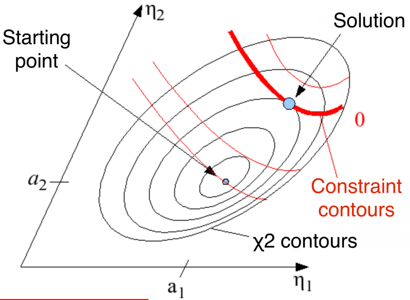


The neutrino momentum can be determined up to a two-fold ambiguity

Can we use overall event kinematics to decide between solutions? $\Rightarrow$ **kinematic fit!**

Kinematic fit

- Kinematic fit: adjustment of measured quantities under certain kinematic constraints:
  - Energy and momentum conservation
  - Invariant masses of particles



arXiv:0901.4656

Exploit well-known initial state in e^+e^- colliders

$\Rightarrow$ **need error parametrization, in particular for jets**

- Minimize χ^2:

$$\chi^2(\boldsymbol{a},\boldsymbol{\xi},\boldsymbol{f}) = (\boldsymbol{\eta}-\boldsymbol{a})^T \boldsymbol{V}^{-1}(\boldsymbol{\eta}-\boldsymbol{a}) - 2\boldsymbol{\lambda}^T \boldsymbol{f}(\boldsymbol{a},\boldsymbol{\xi})$$

$\boldsymbol{\eta}$: vector of measured kinematic variables (x)

$\boldsymbol{a}$: vector of fitted quantities

$\boldsymbol{\xi}$: vector of unmeasured kinematic variables

$\boldsymbol{V}$: **covariance matrix**

$\boldsymbol{\lambda}$: Lagrange multipliers

$\boldsymbol{f}(\boldsymbol{a},\boldsymbol{\xi})$: vector of constraints

- Measures of performance:

$F(\chi^2; \text{ndf})$: cumulative χ^2 distribution for a certain ndf

$P(\chi^2)$: fit probability

Jet specific energy resolution

Parametrize sources of uncertainties (assumed uncorrelated) in jet energy measurements (`ErrorFlow`):

$$\sigma_{E_{jet}} = \sigma_{Det} \oplus \sigma_{Conf} \oplus \sigma_{\nu} \oplus \sigma_{Clus} \oplus \sigma_{Had}$$

- Detector resolution using track and cluster parameters, σ_{Det}
- Particle confusion in Particle Flow Algorithm, σ_{Conf}
  Estimated based on jet energy and neutral hadron / photon energy fractions
- Semi-leptonic decays: error propagation from neutrino correction, σ_{ν}
  currently none since cheating
- Misassignment of particles in the jet clustering, σ_{Clus}
- QCD effects in parton shower and hadronization, σ_{Had} has not been included yet

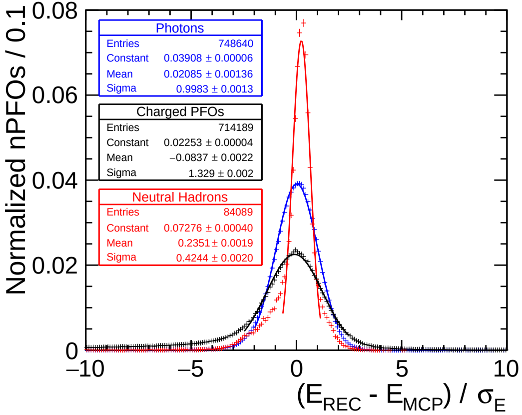
DESY-THESIS-2017-045

ErrorFlow: Jet Error Parametrisation from Particle Flow Objects (PFO)

Energy

Error estimation in PFO level:

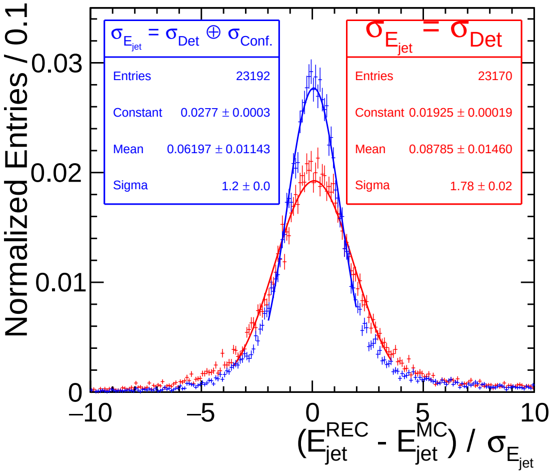
- Photons: energy error is perfectly modeled.
- **Charged PFOs**: uncertainties propagated from track fit covariance matrix
  - uncertainties 30% too small
  - possible future improvement from track refitting with specific mass hypothesis after particle ID
- Neutral Hadrons: energy and energy error are significantly overestimated.



Uncertaities in jet-level: Energy

Propagation of errors from PFOs to jets:

- Transform the covariance matrix of each PFO (E,x,y,z for clusters, track parameters for charged) to (E,p_x,p_y,p_z)
- Add up covariance matrices of all PFOs
- Add confusion term for jet energy
  - calculate using jet energy composition
- Transform to (E,θ,ϕ,m)



Confusion term improves the estimate of the jet energy uncertainty, but not quite enough ⇒ need adjustment

Application of kinematic fit to $e^+e^- \to ZH \to \mu\bar{\mu}b\bar{b}$ events

Parameters of jets and leptons are variated within their uncertainties to satisfy 5 constraints:
Conservation of momentum (hard constraints):

- p_x: e^+e^- crossing angle: 14 mrad
  $\Sigma p_x = \sqrt{s} \times \sin 0.007 \approx 1.75$ GeV
- p_y: $\Sigma p_y = 0$
- p_z: $\Sigma p_z = 0$

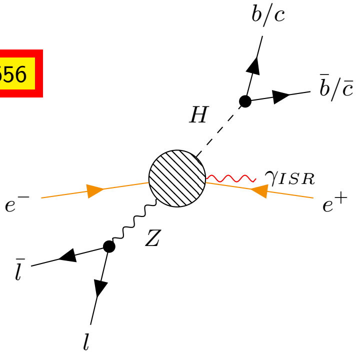
Conservation of total energy (hard constraint):

- $E_{lab} = 2\sqrt{(\frac{\sqrt{s}}{2})^2 + (\Sigma p_x)^2}$

Constrain di-muon mass to agree with m_Z within its natural width (soft constraint):

- $m_Z = 91.2$ GeV , $\sigma_{m_Z} = \frac{2.4952}{2}$

arXiv:0901.4656

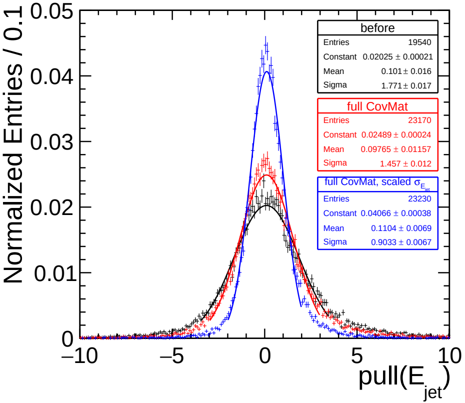


Kinematic fit performance

$e^+e^- \to ZH \to \mu\bar{\mu}b\bar{b}$ at $\sqrt{s} = 250$ GeV (without semi-leptonic decays)

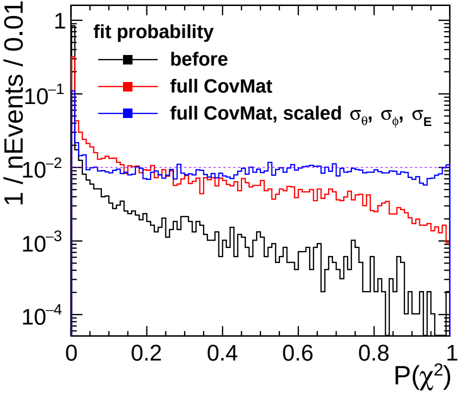
- pull distribution

$$pull(x) = \frac{x_{fitted} - x_{measured}}{\sqrt{\sigma^2_{fitted} - \sigma^2_{measured}}}$$



- fit probability

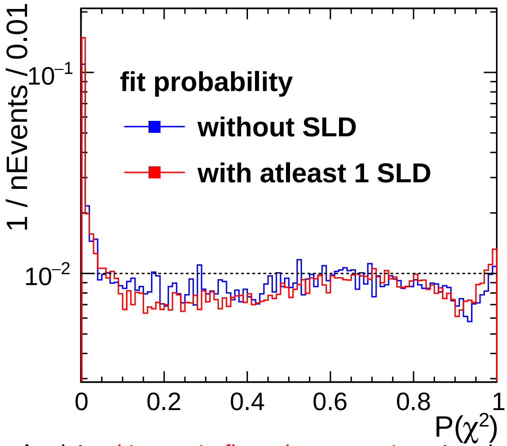
$$P(\chi^2) = 1 - F(\chi^2; \text{ndf})$$

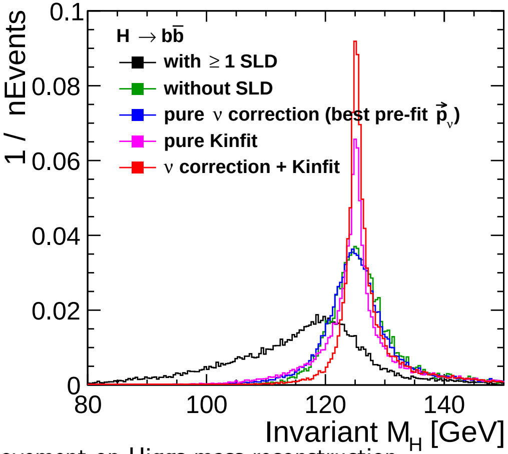


Improved kinematic fit performance with full CovMat of jets + scaled jet energy uncertainty

Higgs mass in presence of SLDs

ν-correction and kinematic fit on $H \to b\bar{b}$





- Applying kinematic fit and ν-correction gives huge improvement on Higgs mass reconstruction

Conclusions

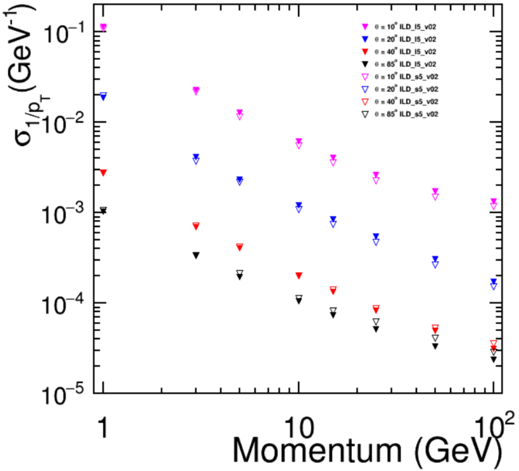
- Higgs mass reconstruction essential eg in ZZH vs ZHH separation (Higgs self-coupling measurement)
- Heavy flavour jets are essential for Higgs physics
- Correction of semi-leptonic decays of heavy flavour jets is important for Higgs mass reconstruction
  - Neutrino momentum can be reconstructed up to a sign ambiguity
  - Ambiguity can be resolved by kinematic fit
  - Next: remove the partial cheating from the neutrino correction
- Kinematic fit exploits well-known initial state in e^+e^- colliders and requires excellent understanding of jet measurement
- ILD as a Particle Flow detector provides full detail for estimating jet measurement uncertainties

BACKUP

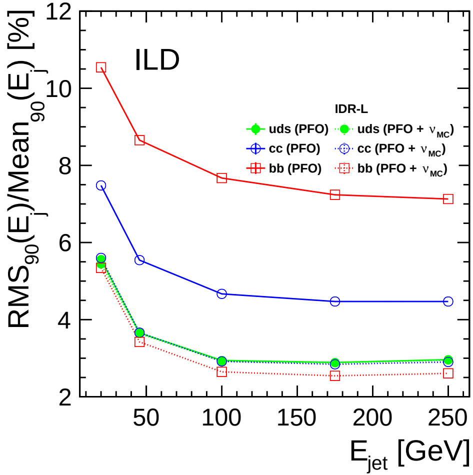
International Large Detector (ILD)

arXiv:2003.01116

- Momentum Resolution



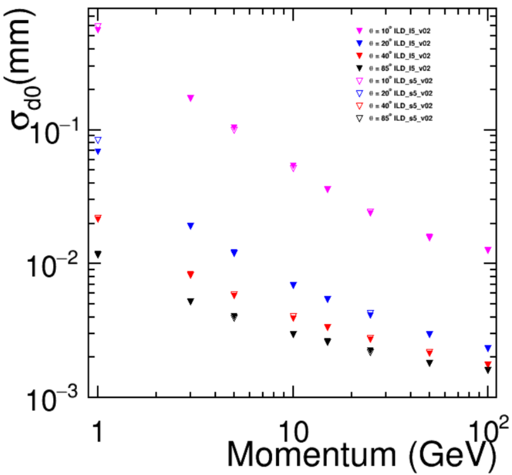
- Jet Energy Resolution $(E_{PFO} + E_{\nu}^{MC})$



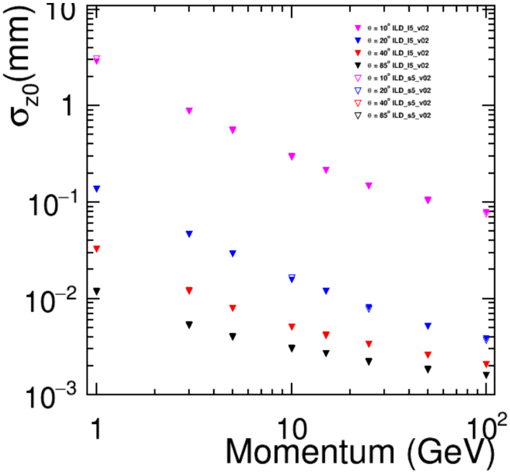
International Large Detector (ILD)

arXiv:2003.01116

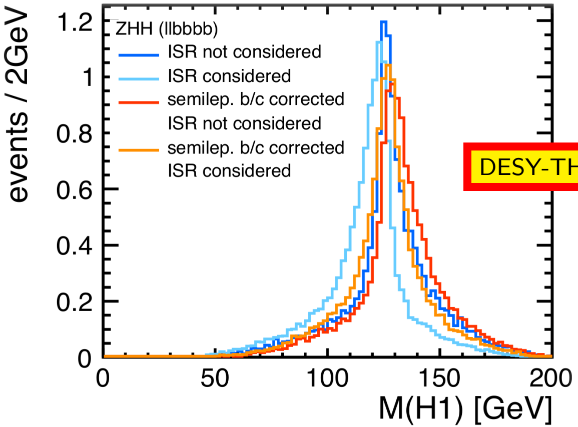
- Impact Parameter Resolution, d_0



- Impact Parameter Resolution, z_0



Simple neutrino correction for Higgs mass reconstruction



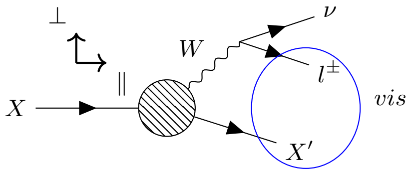
DESY-THESIS-2016-027

- Bias and assymetry in m_H is removed by correcting jet energy and adding ISR

correcting neutrino energy

4-vector based approach

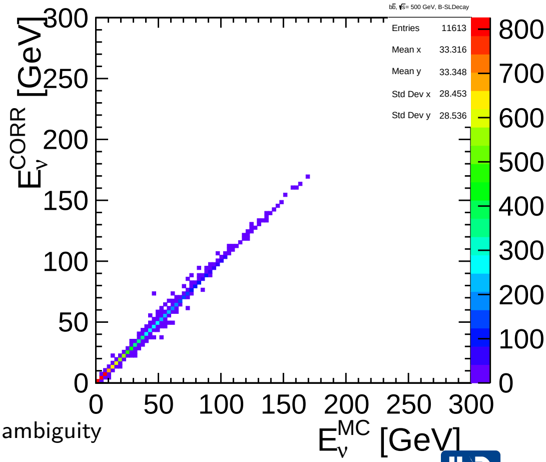
- $(E, \vec{p})$-based approach



$$\vec{p}_{\nu,\perp} = -\vec{p}_{vis,\perp}$$
$$\vec{p}\nu, \| = \frac{1}{2D}(-A \pm \sqrt{A^2 - BD})\hat{n}$$
$$A = p_{vis,\|}(2p^2_{vis,\perp} + m^2_{vis} - m^2_X)$$
$$B = 4p^2_{vis,\perp}E^2_{vis} - (2p^2_{vis',\perp} + m^2_{vis} - m^2_X)^2$$
$$D = E^2_{vis} - p^2_{vis,\|}$$
$$\hat{n} = \frac{\vec{p_{vis,\|}}}{|\vec{p_{vis,\|}}|}$$

The neutrino momentum can be determined up to a two-fold ambiguity

- closure test: apply correction with fully cheated information and compare with true neutrino energy

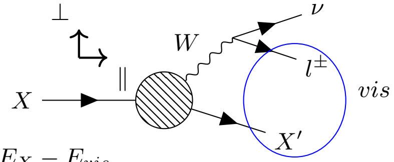


Correcting neutrino energy

Rapidity based approach

Rapidity under Lorents-transformations $\sim$ velocity under Galileo-transformations: $\omega = \omega_X + \omega'$; $\omega = \frac{1}{2} ln \frac{E+p'_{\parallel}}{E-p'_{\parallel}}$

ω: rapdity in lab frame , ω': rapdity in rest frame of X , ω_X: rapdity of X in lab frame



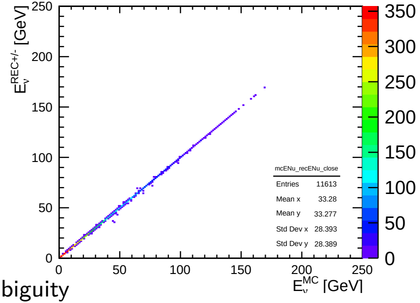
$$E_\nu = E_X - E_{vis}$$

$$E_X = \frac{E_{vis}E'_{vis} - p_{vis_\parallel} p'_{vis_\parallel}}{m^2_{vis} + p^2_{vis_\perp}} m_X$$

$$E'_{vis} = \frac{m^2_X + m^2_{vis}}{2m_X}$$

$$p'_{vis_\parallel} = \pm\sqrt{\left(\frac{m^2_X - m^2_{vis}}{2m_X}\right)^2 - p^2_{vis_\perp}}$$

- Closure test: fully cheated information ($e^+e^- \to b\bar{b}$ at $\sqrt{s} = 500$ GeV)



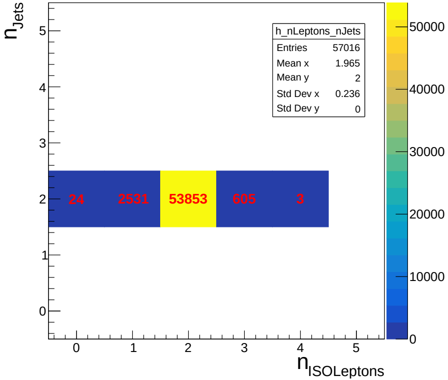
The neutrino momentum can be determined up to a two-fold ambiguity

Can we use overall event kinematics to decide between solutions? $\Rightarrow$ kinematic fit!

Event selection

Select $e^+e^- \rightarrow ZH \rightarrow \mu\bar{\mu}b\bar{b}$ events at $\sqrt{s} = 250$ GeV with (exactly) 2-leptons + 2-jets final state:

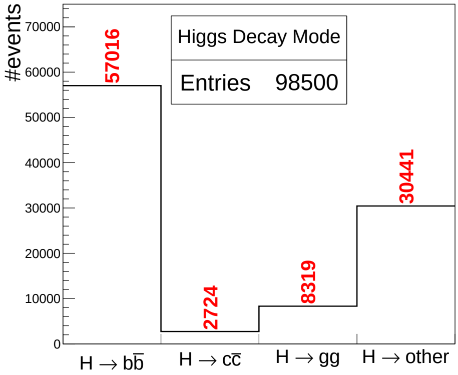
- IsolatedLeptonTagging
  Training for the IDR 500 GeV samples is used,
  1. Lepton ID: $\mu^{\pm}$
     Deposited energy in subdetectors
  2. Vertex: primary or secondary
     Significance of impact parameters (d_0 , z_0)
  3. Isolated: not belong to jets
- FastJetProcessor
  - Exclusive k_t (Durham) algorithm (no overlay)
  - Find smallest of (d_{ij},d_{iB})
    $d_{ij} = 2\min(E_i^2, E_j^2)(1 - \cos\theta_{ij})$
    i,j: particles, B: Beam
  - $d_{ij} < d_{iB}$: combine i&j as pseudojet(p): $p_i + p_j$
  - $d_{iB} < d_{ij}$: remove particle i from list
  - Repeat iteration until d_{ij} or $d_{iB} > d_{cut}$(threshold)



IsolatedLeptonTagging has not been trained for new software at 250 GeV yet!

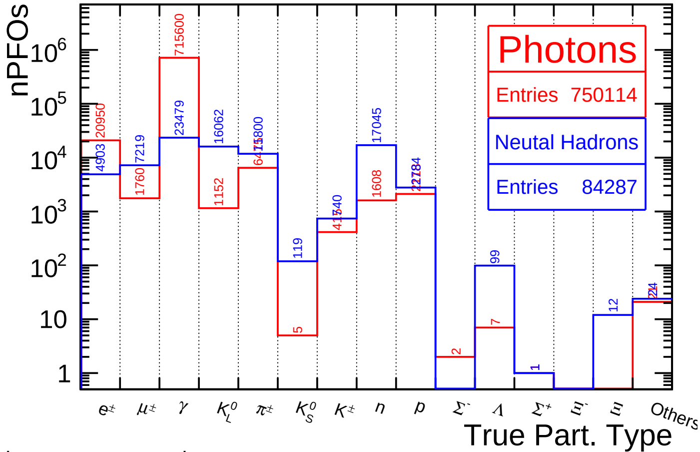
event selection

separate Higgs decay modes: $H \to b\bar{b}$, cheat from MCTruth



$\frac{2}{3}$ of $b\bar{b}$ jets contain at-least one semi-leptonic decay $\Rightarrow$ Frequent $H \to b\bar{b}$ needs neutrino correction.

Neutral PFO identification by Pandora



Majority of identified photons are true photons.

No explicit decision for mass of identified neutral hadrons due to their multiplicity.

Pandora treatment with Neutral Hadrons

What Pandora does:

- Cluster energy is assigned to PFO(massless) energy
  $E_{PFO} = |\vec{p}_{PFO}| = E_{cluster}$
- Neutral Hadrons are identified as neutron
- neutron mass is set for PFO $\Rightarrow$ **incosistent 4-momentum!**
- CovMat of Neutral PFO is calculated (using inconsistent 4-momentum):
  CovMat$(\vec{p}, E) = J^T$ CovMat$(\vec{x}_{clu}, E_{clu})\ J$

$$J = \begin{pmatrix} \frac{\partial p_x}{\partial x_c} & \frac{\partial p_y}{\partial x_c} & \frac{\partial p_z}{\partial x_c} & \frac{\partial E}{\partial x_c} \\ \frac{\partial p_x}{\partial y_c} & \frac{\partial p_y}{\partial y_c} & \frac{\partial p_z}{\partial y_c} & \frac{\partial E}{\partial y_c} \\ \frac{\partial p_x}{\partial z_c} & \frac{\partial p_y}{\partial z_c} & \frac{\partial p_z}{\partial z_c} & \frac{\partial E}{\partial z_c} \\ \frac{\partial p_x}{\partial E_c} & \frac{\partial p_y}{\partial E_c} & \frac{\partial p_z}{\partial E_c} & \frac{\partial E}{\partial E_c} \end{pmatrix}$$

CovMat$(\vec{p}, E)$ of Neutral PFOs depend on the mass assumption.
Suggestion: Take consistent 4-momentum of massive neutral hadrons for CovMat calculations.

CovMat of Neutral PFOs

- Current CovMat calculation (MarlinReco/Analysis/AddClusterProperties)

  $$E_{PFO} = |\vec{p}_{PFO}| = E_{clu}\ ,\ p_x = E_{clu}\frac{x}{r}\ ,\ p_y = E_{clu}\frac{y}{r}\ ,\ p_z = E_{clu}\frac{z}{r}$$

- Alternative CovMat calculation (taking consistent 4-momentum of neutral hadrons)

  $$E_{PFO} = \sqrt{|\vec{p}_{PFO}|^2 + m_{PFO}^2} = \sqrt{E_{clu}^2 + m_n^2}$$

$$J = \begin{pmatrix} E_{clu}\frac{r^2-x^2}{r^3} & -E_{clu}\frac{xy}{r^3} & -E_{clu}\frac{xz}{r^3} & 0 \\ -E_{clu}\frac{xy}{r^3} & E_{clu}\frac{r^2-y^2}{r^3} & -E_{clu}\frac{yz}{r^3} & 0 \\ -E_{clu}\frac{xz}{r^3} & -E_{clu}\frac{yz}{r^3} & E_{clu}\frac{r^2-z^2}{r^3} & 0 \\ \frac{x}{r} & \frac{y}{r} & \frac{z}{r} & 1 \end{pmatrix} \rightarrow J = \begin{pmatrix} E_{clu}\frac{r^2-x^2}{r^3} & -E_{clu}\frac{xy}{r^3} & -E_{clu}\frac{xz}{r^3} & 0 \\ -E_{clu}\frac{xy}{r^3} & E_{clu}\frac{r^2-y^2}{r^3} & -E_{clu}\frac{yz}{r^3} & 0 \\ -E_{clu}\frac{xz}{r^3} & -E_{clu}\frac{yz}{r^3} & E_{clu}\frac{r^2-z^2}{r^3} & 0 \\ \frac{E}{E_{clu}}\cdot\frac{x}{r} & \frac{E}{E_{clu}}\cdot\frac{y}{r} & \frac{E}{E_{clu}}\cdot\frac{z}{r} & 1 \end{pmatrix}$$

using error propagation, PFO angular uncertainties are calculated directly from cluster position error:

$$\sigma_\theta^2 = (\frac{\partial\theta}{\partial x})^2\sigma_x^2 + (\frac{\partial\theta}{\partial y})^2\sigma_y^2 + (\frac{\partial\theta}{\partial z})^2\sigma_z^2 + \frac{\partial\theta}{\partial x}\frac{\partial\theta}{\partial y}\sigma_{xy} + \frac{\partial\theta}{\partial x}\frac{\partial\theta}{\partial z}\sigma_{xz} + \frac{\partial\theta}{\partial y}\frac{\partial\theta}{\partial z}\sigma_{yz}$$

$$\sigma_\phi^2 = (\frac{\partial\phi}{\partial x})^2\sigma_x^2 + (\frac{\partial\phi}{\partial y})^2\sigma_y^2 + \frac{\partial\phi}{\partial x}\frac{\partial\phi}{\partial y}\sigma_{xy}$$

MUST: angular and energy uncertainties remain unchanged!

CovMat of Jets

- `AddClusterProperties/FourMomentumCovMat`: CovMat(cluster/track) $\rightarrow$ CovMat($\vec{p}, E$)
  - Current CovMat calculation (inconsistent 4-momentum of neutral hadrons):

    $E_{PFO} = |\vec{p}_{PFO}| = E_{clu}$, $p_x = E_{clu}\frac{x}{r}$, $p_y = E_{clu}\frac{y}{r}$, $p_z = E_{clu}\frac{z}{r}$, $m_{PFO} = m_n$
  - Alternative CovMat calculation (taking consistent 4-momentum of neutral hadrons)

    $E_{PFO} = \sqrt{|\vec{p}_{PFO}|^2 + m_{PFO}^2} = \sqrt{E_{clu}^2 + m_n^2}$

$$J_{(wrong)} \rightarrow J_{(right)}$$

$$\begin{pmatrix} E_{clu}\frac{r^2-x^2}{r^3} & -E_{clu}\frac{xy}{r^3} & -E_{clu}\frac{xz}{r^3} & 0 \\ -E_{clu}\frac{xy}{r^3} & E_{clu}\frac{r^2-y^2}{r^3} & -E_{clu}\frac{yz}{r^3} & 0 \\ -E_{clu}\frac{xz}{r^3} & -E_{clu}\frac{yz}{r^3} & E_{clu}\frac{r^2-z^2}{r^3} & 0 \\ \frac{x}{r} & \frac{y}{r} & \frac{z}{r} & 1 \end{pmatrix}_{wrong} \rightarrow \begin{pmatrix} E_{clu}\frac{r^2-x^2}{r^3} & -E_{clu}\frac{xy}{r^3} & -E_{clu}\frac{xz}{r^3} & 0 \\ -E_{clu}\frac{xy}{r^3} & E_{clu}\frac{r^2-y^2}{r^3} & -E_{clu}\frac{yz}{r^3} & 0 \\ -E_{clu}\frac{xz}{r^3} & -E_{clu}\frac{yz}{r^3} & E_{clu}\frac{r^2-z^2}{r^3} & 0 \\ \frac{E}{E_{clu}}\cdot\frac{x}{r} & \frac{E}{E_{clu}}\cdot\frac{y}{r} & \frac{E}{E_{clu}}\cdot\frac{z}{r} & 1 \end{pmatrix}_{right}$$

- `ErrorFlow`:

$$\text{CovMat}(\vec{p}_{jet}, E_{jet}) = \sum_{PFO} \text{CovMat}(\vec{p}, E) \quad : \quad \sigma_{E_{jet}}^2 = \sigma_{conf}^2 + \sum_{PFO} \sigma_{E_{PFO}}^2$$

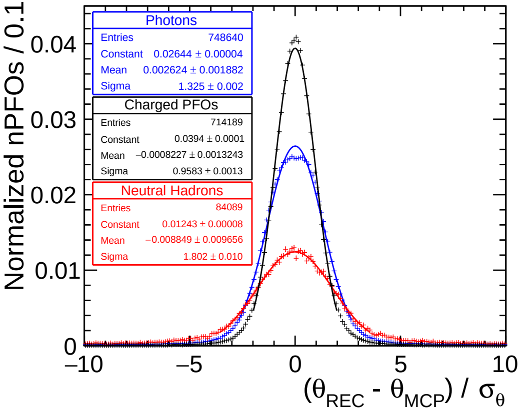
- `MarlinKinfitProcessors`:

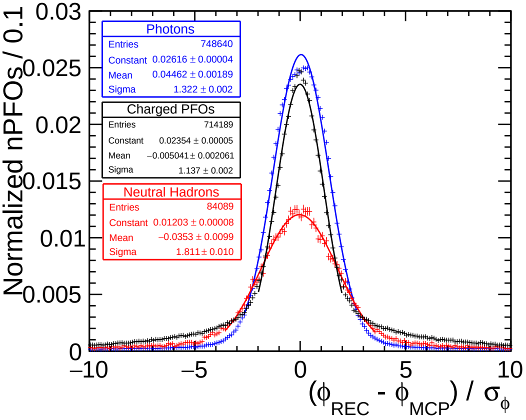
$$\text{CovMat}(\vec{p}_{jet}, E_{jet}) \rightarrow (\sigma_{\theta_{jet}} , \sigma_{\phi_{jet}} , \sigma_{E_{jet}})$$

ErrorFlow: Jet Error Parametrisation from Particle Flow Objects (PFO)

Angles

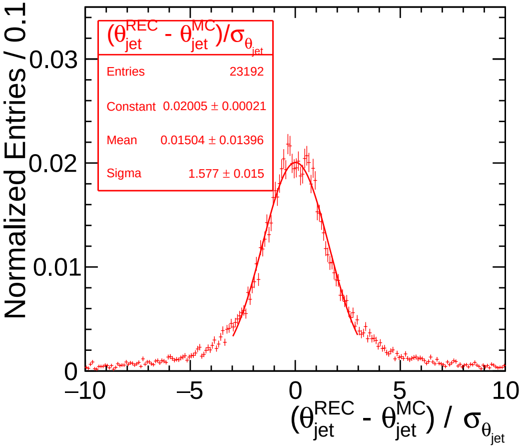
The angular uncertainties obtained directly from track parameters / cluster position errors

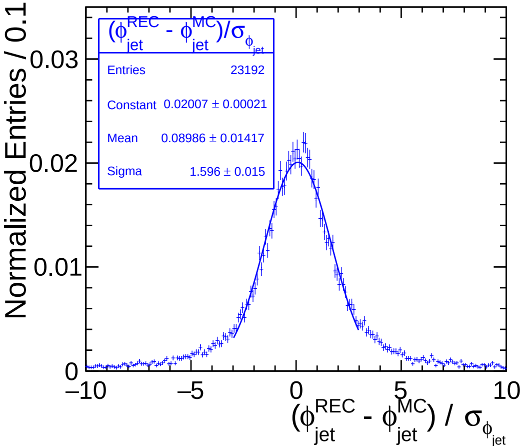




⇒ Scale σ_θ and σ_ϕ by factor ~1.3 (for photons) and ~1.8 (for neutral hadrons)

Uncertaities in jet-level: θ & ϕ





Jet angular uncertainties need scaling factor $\sim$1.6

Neutrino correction hypothesis

- Assign semi-leptonic decays to jets
- Add neutrino momentum to 4-momentum of assigned jet:
  Test three hypothesis for neutrino energy in each semi-leptonic decay: E_ν^+ , E_ν^- , 0
  3^{nSLD} combination of E_ν's for adding to jet 4-momentum:
  Number of semileptonic decays in a jet: nSLD = nSLDB + nSLDC
  Example:
  If an event contains two jets: jet-1 contains 2 semi-leptonic decays and jet-2 contains 1 semi-leptonic decay,
  27($=3^2 \times 3^1$) combinations of E_ν's are available for neutrino correction in the event:

- jet-1:

comb.	1	2	3	4	5	6	7	8	9
$\vec{p}_{\nu,1}$	-	+	0	-	+	0	-	+	0
$\vec{p}_{\nu,2}$	-	-	-	+	+	+	0	0	0

- jet-2:

comb.	1	2	3
$\vec{p}_{\nu,3}$	-	+	0

$\vec{p}_{\nu,1} + \vec{p}_{\nu,2}$ is added to 4-momentum of jet-1 and $\vec{p}_{\nu,3}$ is added to 4-momentum of jet-2.
$\vec{p}_{\nu,1} + \vec{p}_{\nu,2} + \vec{p}_{\nu,3} = 0$ allows fitter to neglect neutrino correction
Combination with highest fit probability is chosen as best neutrino correction.

Simple neutrino correction for Higgs mass reconstruction

- Neutrino energy correction:

  Estimating neutrino energy as a fraction of corresponding lepton energy:

$$E_{jet}^{corr} = E_{jet} + E_{\nu} = E_{jet} + (\frac{1}{x} - 1)E_{lep}$$

- Uncertainty on jet energy parametrised as:

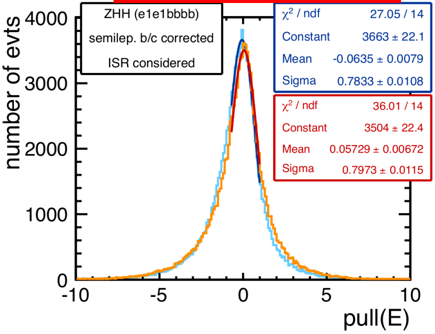
$$\sigma_{E_{jet}}^{corr} = \frac{100\%}{\sqrt{E_{jet}}} \oplus \sigma_{\nu}$$

$$\sigma_{\nu}^{2} = (\frac{\sigma_{\langle x \rangle}}{\langle x \rangle^{2}})^{2} E_{lep}^{2} + (\frac{1}{x} - 1)\Delta E_{lep}^{2}$$

- Fixed uncertainties on angles:

$$\Delta\theta_{jet} = \Delta\phi_{jet} = 100\,\text{mrad}$$

DESY-THESIS-2016-027



~~Blue~~: before neutrino energy correction

~~Orange~~: After neutrino energy correction

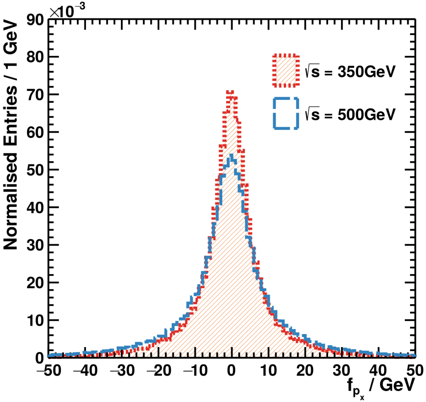
Simple correction to jet energy improves jet energy pull distribution as a measure of fit performance.

fit constraints

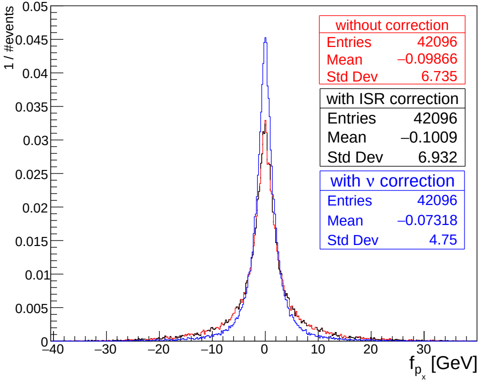
momentum conservation: p_x

ISR is initialized to satisfy momentum conservation on x direction

- by error flow on jet energy (Ali)



- by error flow on CovMatrix (new)



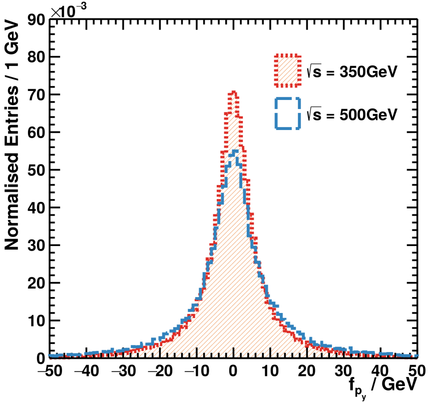
angular resolution for individual jets: improved constraint on momentum conservation

fit constraints

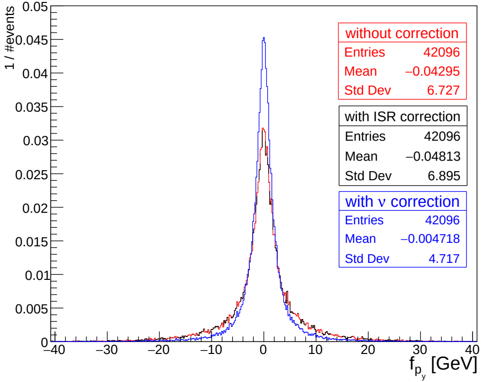
momentum conservation: p_y

ISR is initialized to satisfy momentum conservation on z direction

- by error flow on jet energy (Ali)



- by error flow on CovMatrix (new)



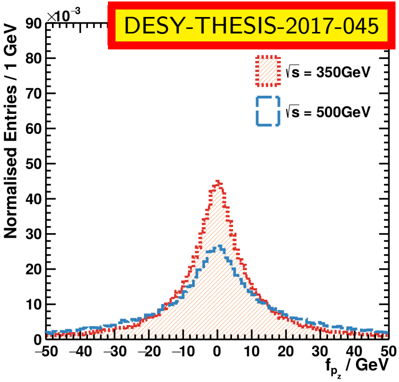
angular resolution for individual jets: improved constraint on momentum conservation

Fit constraints

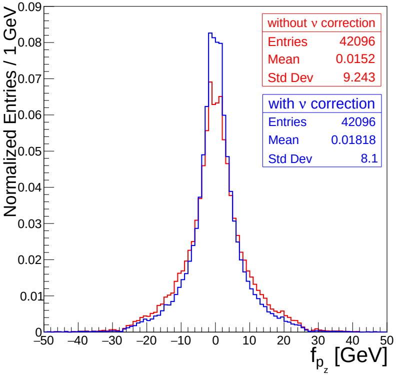
Momentum conservation: p_z

Adding 4-momentum of neutrino improves jet fit object initialization

- DBD 350/500 GeV samples



- MC-2020 250 GeV prod. samples



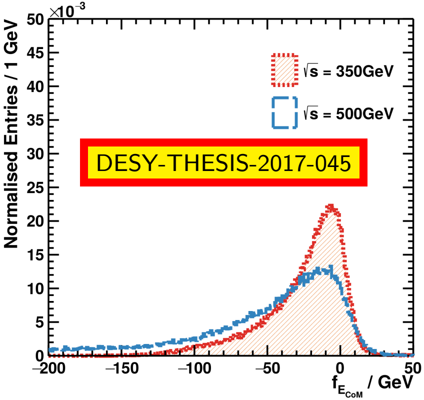
Proper neutrino correction for jets: improved constraint on momentum

fit constraints

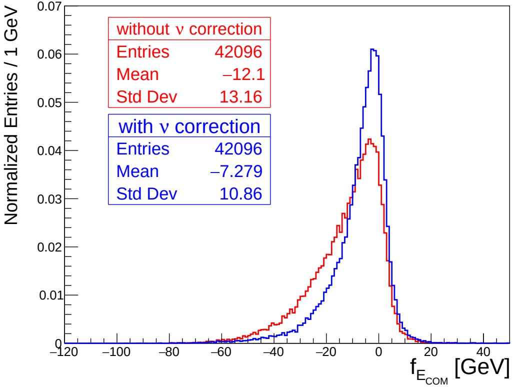
energy conservation: E

Neutrino correction (best pre-fit $\vec{p}_\nu$ for succesful fits) improves start values ⇒ better fit object initialization

- DBD 350/500 GeV samples



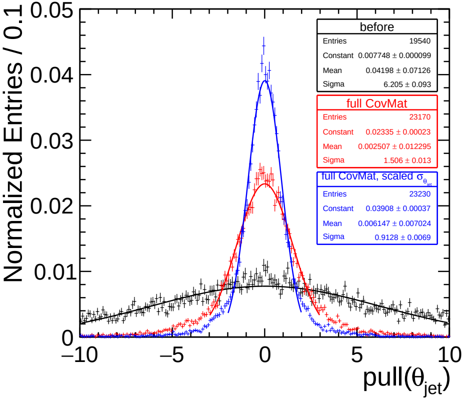
- MC-2020 250 GeV prod. samples

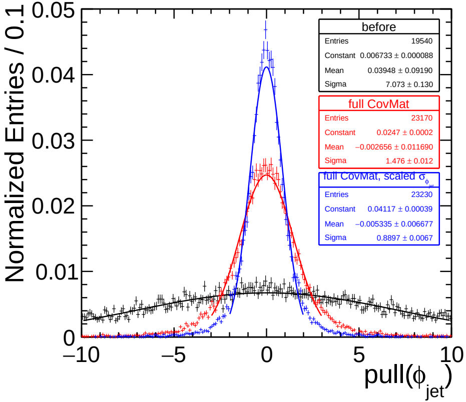


By neutrino correction, inital value of constraint function closer to target ⇒ fit should work better!

Kinematic fit performance (pull)

$e^+e^- \to ZH \to \mu\bar{\mu}b\bar{b}$ at $\sqrt{s} = 250$ GeV (without semi-leptonic decays)





Improved kinematic fit performance with full CovMat of jets + scaled jet angular uncertainties

fit probability