Neutrino reconstruction in semi/fully-leptonic WW scattering

J. Novak [1] **and M. Grossi** [2]

in collaboration with:

B. Kerševan, D. Rebuzzi, D. Valsecchi

[1]University of Ljubljana

[2]University of Pavia, INFN, IBM Italy

Formula Derivation for a process with one ν

$$m_{\mathrm{W}}^2 = (p_\mu + p_\nu)^2 \;\; \textit{ultra relativistic limit} \;\; \xrightarrow{m \to 0} \;\; 2p_\mu p_\nu;$$

- Let's solve for the longitudinal component of the neutrino $p_{\nu L}$;

$$\underbrace{\left(p_{lL}^2 - E_l^2\right)}_{a} p_{\nu L}^2 +$$

$$\underbrace{\left(m_{\mathrm{W}}^2 p_{lL} + 2p_{lL}\vec{p}_{lT}\vec{p}_{\nu T}\right)}_{b} p_{\nu L} +$$

$$\underbrace{\frac{m_{\mathrm{W}}^4}{4} + (\vec{p}_{lT}\vec{p}_{\nu T})^2 + m_{\mathrm{W}}^2\vec{p}_{lT}\vec{p}_{\nu T} - E_l^2\vec{p}_{\nu T}^{\,2})}_{c} = 0;$$

$$\boxed{p_{\nu L_{1,2}} = \frac{-b \pm \sqrt{\Delta}}{2a}} \quad \text{where} \quad \Delta = b^2 - 4ac$$

As a second order parametric equation, Δ determines the number of solution and their nature.

$m_W \Rightarrow$ **fixed value** $(80.385$ GeV$)$

- if $\Delta > 0 \Rightarrow$ 2 solutions $(+/-)$
- if $\Delta < 0$, from the formula:

we have two working options, we choose the first one

$$\Delta(p_L) \begin{cases} \text{put } \Delta = 0 \\ m_W = m_{WT} \Rightarrow \quad \text{correct } m_W \text{ with transverse mass} \end{cases}$$

Definitions of the samples

PHANTOM PARAMETERS for the production

- semi-leptonic: $pp \to jjjj\mu^+\nu_\mu$
- full-leptonic: $pp \to jj\mu^+\nu_\mu e^+\nu_e$
- Parton level events
- MC generator: `Phantom`
- events generated with `NNPDF30_nnlo_as_0118`
- CALCULATION TYPE: α_e^6
- SCALE CHOICE: (invariant mass of the 2 central jets and of 2 leptons)/$\sqrt{2}$

Kinematical cuts:

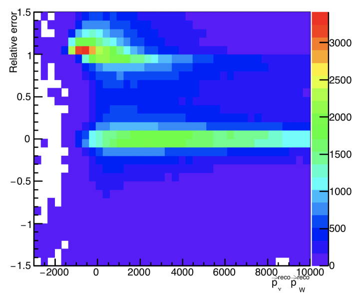
- $p_{\mathsf{T}}^{\ell} > 20$ GeV
- $|\eta^{\ell}| < 3$
- $p_{\mathsf{T}}^{min} > 30$ GeV
- $|\eta_j| < 5.4$
- $p_{\mathsf{T}}^{\text{miss}} > 20$ GeV
- $m_{jj} > 500$ GeV
- $\Delta R_{j\ell} > 0.3$

VBS Semileptonic channel

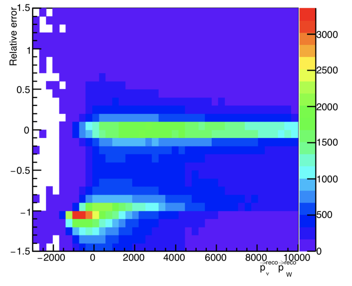
Resolving the ambiguity

- The search for a discrimination variable, which would display the phase space region with higher number of correct solutions
- Example $\vec{p}_\nu \cdot \vec{p}_W$:
  - Correct solutions tend to reside in the region of large $\vec{p}_\nu \cdot \vec{p}_W$

$+$ solutions



- solutions

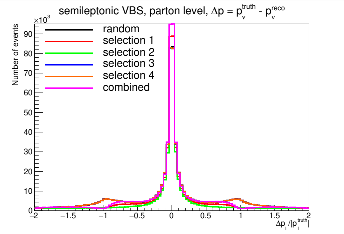


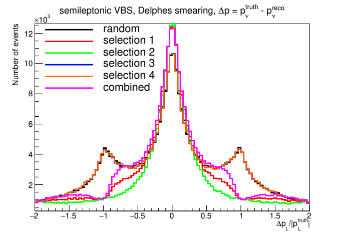
Selection criteria

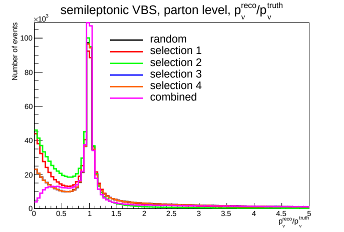
- Selection 1: $p_{\nu L} > 50$ GeV
- Selection 2: $p_{\nu L} a/b > -0.5\ \rightarrow$ choosing larger* solution
- Selection 3: $\vec{p}_\nu \cdot \vec{p}_W < 5000$ GeV^2
- Selection 4: $\vec{p}_\nu \cdot \vec{p}_W a/b < 25$ GeV
- Combined:
  - Selection of solutions with $p_{\nu L} > 50$ GeV
  - If both solutions lie above 50 GeV, Selection 3 is applied
  - If both solutions lie below 50 GeV, solution with lower value of the $|\vec{p}_\nu \cdot \vec{p}_W a/b|$ is taken.

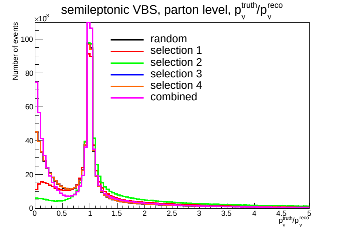
$^* p_{\nu L} a/b = -\frac{1}{2} \pm \frac{\sqrt{\Delta}}{2b}$

Selection criteria - relative errors





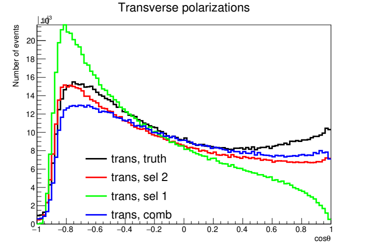


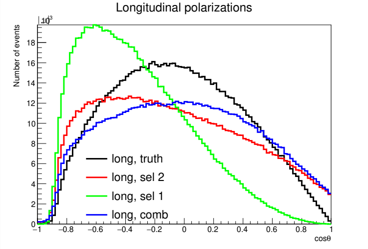


- Selection criteria typically remove wrong solutioins from one phase space region, but generate them in another

Semileptonic angular distributions

cosθ distribution of the charged lepton in the W reference frame.





Selection 4 and combined give the best results

VBS Fully leptonic channel

Reco of neutrino momentum in $pp \to jje^+\nu_e\mu^+\nu_\mu$

- 8 unknown parameters (2 x neutrino four momentum)
- 6 equations:
  - $\vec{p}_{\mathrm{T}}^{\,\nu_\mu} + \vec{p}_{\mathrm{T}}^{\,\nu_e} = \vec{p}_{\mathrm{T}}^{\,\mathrm{miss}}$ (2x)
  - $(p^\ell + p^\nu)^2 = m_W^2$ (2x)
  - $p_\nu^2 = 0$ (2x)
- Remaining 2 equations:
  1. Setting some parameters to fixed values - for example: $M_{WW}^2 = (p_e + p_{\nu_e} + p_\mu + p_{\nu_\mu})^2$ and $M_{\nu\nu}^2 = (p_{\nu_e} + p_{\nu_\mu})^2$, M_{WW} and $M_{\nu\nu}$ are fixed numbers.[1]
  2. Using of $MT2$-Assisted On-Shell (MAOS) quantites, i.e. minimization of the transverse masses of the lepton-neutrino pairs.[1]
  3. Other ideas ??

[1] arXiv:hep-ph/0603011, Higgs spin analysis in Collins-Soper frame using opening angles of different-flavour final state leptons

[1] arXiv:0908.0079

2. Reco with MAOS quantities - Equations

- MAOS estimations $\vec{p}_{\mathsf{T}}^{\nu_e\prime}$ and $\vec{p}_{\mathsf{T}}^{\nu_\mu\prime}$ for neutrinos transverse momentums can be obtained by minimizing the function $f(\vec{p}_1, \vec{p}_2) = \max\{M_{\mathsf{T}}^{W_1}, M_{\mathsf{T}}^{W_2}\}$, constrained by a bond $\vec{p}_1 + \vec{p}_2 = \vec{p}_{\mathsf{T}}^{\mathsf{miss}}$, where

$$M_{\mathsf{T}}^{W_1} = 2(|\vec{p}_{\mathsf{T}}^{\mu}||\vec{p}_1| - \vec{p}_{\mathsf{T}}^{\mu} \cdot \vec{p}_1), \qquad M_{\mathsf{T}}^{W_2} = 2(|\vec{p}_{\mathsf{T}}^{e}||\vec{p}_2| - \vec{p}_{\mathsf{T}}^{e} \cdot \vec{p}_2)$$

- Minimum of the function f defines quantity M_{T2}:

$$M_{\mathsf{T2}} \equiv \min_{\vec{p}_1 + \vec{p}_2 = \vec{p}_{\mathsf{T}}^{\mathsf{miss}}} f(\vec{p}_1, \vec{p}_2) = f|_{\vec{p}_{\mathsf{T}}^{\nu_e\prime}, \vec{p}_{\mathsf{T}}^{\nu_\mu\prime}} \tag{1}$$

- p_L are then determined from the m_W constraints
- Solution of the problem (1), under assumption $p_{\mathsf{T}}^{WW} \sim 0$ (approximate solution):

$$\vec{p}_{\mathsf{T}}^{\nu_e\prime} = -\vec{p}_{\mathsf{T}}^{\mu} \qquad \vec{p}_{\mathsf{T}}^{\nu_\mu\prime} = -\vec{p}_{\mathsf{T}}^{e}$$

2. Reco with MAOS quantities - Equations

Exact solution:

- $\min\left[\max\{M_T^{W_1}, M_T^{W_2}\}\right]$ can always lie only on the intersection of $M_T^{W_1}$ and $M_T^{W_2}$. $\Rightarrow$ Additional bond: $M_T^{W_1} = M_T^{W_2}$
  - It follows that:

$$\Rightarrow \quad 2(|\vec{p}_T^{\mu}||\vec{p}_1| - \vec{p}_T^{\mu} \cdot \vec{p}_1) = 2(|\vec{p}_T^{e}||\vec{p}_2| - \vec{p}_T^{e} \cdot (\vec{p}_T^{\text{miss}} - \vec{p}_1))$$

$$|\vec{p}_T^{\mu}||\vec{p}_1| - \vec{p}_T^{\ell\ell} \cdot \vec{p}_1 + \vec{p}_T^{e} \cdot \vec{p}_T^{\text{miss}} = |\vec{p}_T^{e}|\sqrt{|\vec{p}_T^{\text{miss}}|^2 - 2\vec{p}_T^{\text{miss}} \cdot \vec{p}_1 + |\vec{p}_1|^2}$$

$$(|\vec{p}_T^{\mu}||\vec{p}_1| - |\vec{p}_T^{\ell\ell}||\vec{p}_1|\cos\varphi + \vec{p}_T^{e} \cdot \vec{p}_T^{\text{miss}})^2 = |\vec{p}_T^{e}|^2|\vec{p}_T^{\text{miss}}|^2 - 2|\vec{p}_T^{e}|^2|\vec{p}_T^{\text{miss}}||\vec{p}_1|\cos(\varphi + \varphi_0) + |\vec{p}_T^{e}|^2|\vec{p}_1|^2$$

  - φ_0 - angle between $\vec{p}_T^{\text{miss}}$ and $\vec{p}_T^{\ell\ell}$
    $\rightarrow$ Parameter of the equation: $\varphi_0 = \arccos\left(\frac{\vec{p}_T^{\ell\ell} \cdot \vec{p}_T^{\text{miss}}}{|\vec{p}_T^{\ell\ell}||\vec{p}_T^{\text{miss}}|}\right)$
  - φ - angle between $\vec{p}_T^{\ell\ell}$ and $\vec{p}_1$;

2. Reco with MAOS quantities - Equations

$$\underbrace{\left(|\vec{p}_{\mathsf{T}}^{\mu}|^2 + |\vec{p}_{\mathsf{T}}^{\ell\ell}|^2 \cos^2\varphi - 2|\vec{p}_{\mathsf{T}}^{\mu}||\vec{p}_{\mathsf{T}}^{\ell\ell}|\cos\varphi - |\vec{p}_{\mathsf{T}}^{e}|^2\right)}_{f(\varphi)} |\vec{p}_1|^2 \quad +$$

$$\underbrace{\left(2(|\vec{p}_{\mathsf{T}}^{\mu}| - |\vec{p}_{\mathsf{T}}^{\ell\ell}|\cos\varphi)\vec{p}_{\mathsf{T}}^{e} \cdot \vec{p}_{\mathsf{T}}^{\mathsf{miss}} + 2|\vec{p}_{\mathsf{T}}^{e}|^2|\vec{p}_{\mathsf{T}}^{\mathsf{miss}}|\cos(\varphi + \varphi_0)\right)}_{g(\varphi)} |\vec{p}_1| \quad +$$

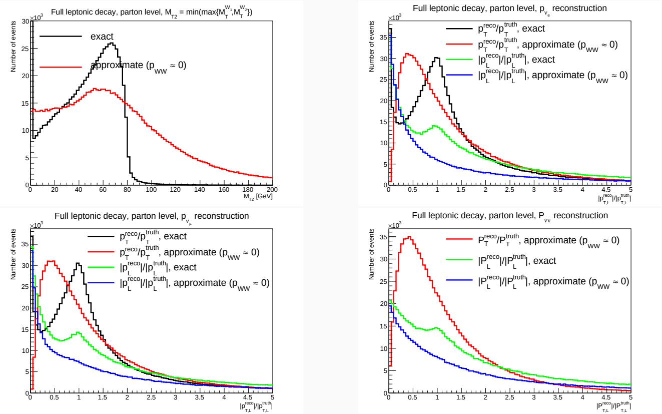
$$\underbrace{(\vec{p}_{\mathsf{T}}^{e} \cdot \vec{p}_{\mathsf{T}}^{\mathsf{miss}})^2 - |\vec{p}_{\mathsf{T}}^{\mathsf{miss}}|^2|\vec{p}_{\mathsf{T}}^{e}|^2}_{c} \quad = \quad 0$$

- Equation of the intersection curve in parametric form - x-axis of coordinate system coinciding with the $\vec{p}_{\mathsf{T}}^{\ell\ell}$ direction;

$$|\vec{p}_1| = \frac{-g(\varphi) \pm \sqrt{g(\varphi)^2 - 4cf(\varphi)}}{2f(\varphi)}, \qquad \vec{p}_2 = \vec{p}_{\mathsf{T}}^{\mathsf{miss}} - \vec{p}_1$$

- Minimum of M_{T2} on the intersection curve can be found numerically;
- The following plots are produced by evaluating M_{T2} in 2000 points;

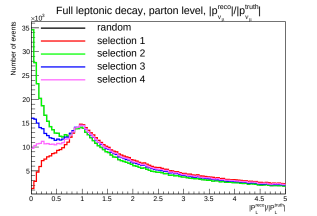
Results - random solution choice

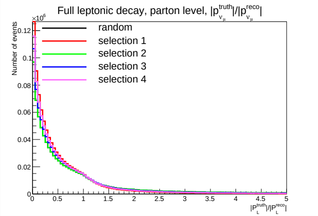


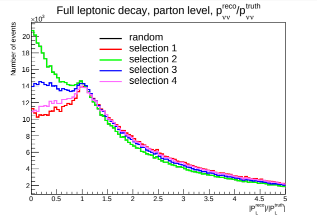
- Exact solution performs better as we have peak at 1
- M_{T2} has a sharp edge at m_W

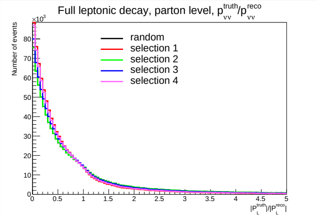
Performance of different selections

Some selections improve p^{reco}/p^{truth} ratio but degrade p^{truth}/p^{reco}



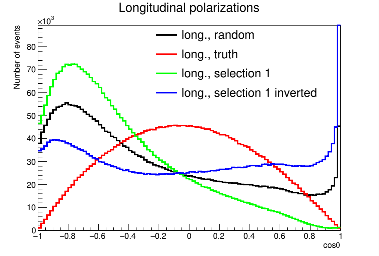


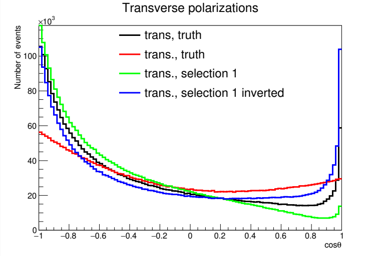




Angular distributions with selection criteria

- We are not able to reconstruct truth angular distribution accurate enough
- inverted selection 1 increase discriminating power for separation of polarization





Reconstruction with neural network in semileptonic channel

Selection of the solutions by neural network

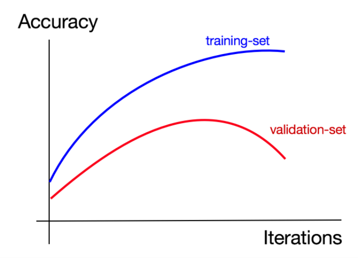
- The choice of solutions in semileptonic channel can be performed via machine learning
- We constructed a simple neural network (NN) for classification of the events in two categories:
    - Events with first solution closer to truth
    - Events with second solution closer to truth

  } Binary classification problem

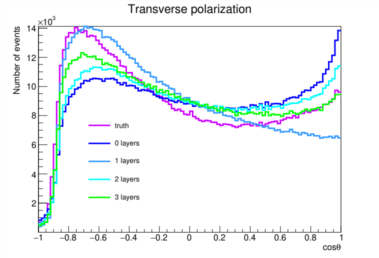
- Events with negative discriminant are discarded
- NN hyperparameters configuration:
    - 1 hidden layer, 32 neurons, 32 batch size, 0.001 learning rate, 150 epochs

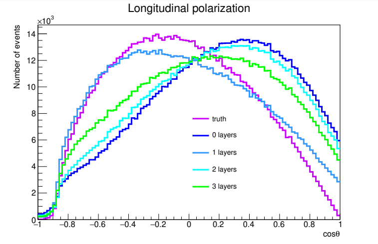
Classification with NN

- Training and validation of the NN on the unpolarized semileptonic data:
    - We splitted the data in two subsets of equal size: training and validation dataset
- Analysis:
    - Observation of the ROC curve
    - Application of the NN on the polarized samples: qualitative evaluation of the statistical significance of the separation between transverse and longitudinal, agreement with truth



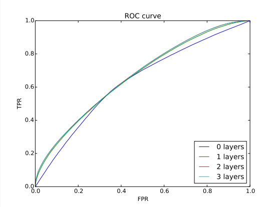
Dependence on the number of layers





$$\mathrm{TPR} = \frac{\mathrm{TP}}{\mathrm{TP}+\mathrm{FN}}$$

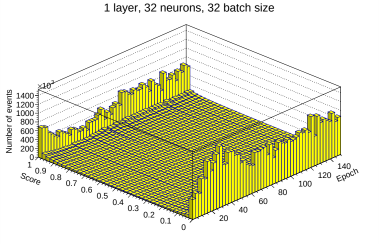
$$\mathrm{FPR} = \frac{\mathrm{FP}}{\mathrm{TN}+\mathrm{FP}}$$

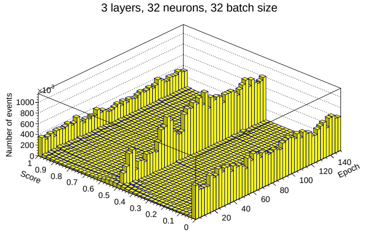


- Smaller networks perform better than larger
- No significant difference in ROC curve (AUC $\sim 66\%$)
  - Larger networks classify better in the regions of less importance

Score vs. epoch

- Score = probability for + solution to be correct

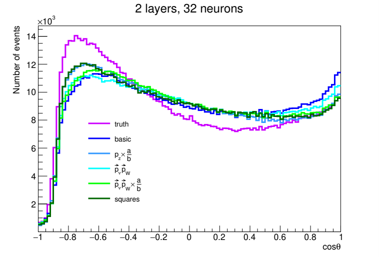


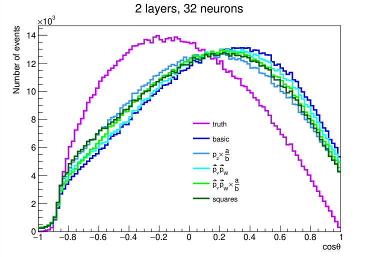


- With increasing number of layers, events begin to cluster at 0.5 score
- Fluctuations due to low statistics

Rich dataset

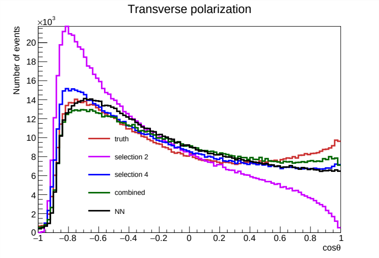
- We trained the network with 4 dataset extensions:
    - selection 2 + selection 1: $p_\nu^\pm \frac{a}{b}$, $p_\nu^\pm$
    - selection 3 + selection 1: $\vec{p}_\nu^\pm \cdot \vec{p}_W^\pm$, $p_\nu^\pm$
    - selection 4 + selection 1: $\vec{p}_\nu^\pm \cdot \vec{p}_W^\pm \frac{a}{b}$, $p_\nu^\pm$
    - squares of all the variables from basic dataset

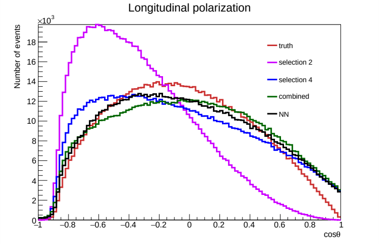




- No significant improvement from dataset extensions

Comparison of NN to selection criteria





- Reconstruction with NN performs better than any selection criteria

CONCLUSIONS: WW reconstruction

- Longitudinal momentum of neutrino can be reconstructed in the semileptonic channel up to a two fold ambiguity
- Selection criteria improve performance, but reconstruction is not very accurate
- We showed that the reconstruction of neutrino can be tackled with the NN approach and that it gives better results than the selection criteria
- MAOS algorithm offers technique for neutrino reconstruction in fully leptonic channel, but only in rough approximation $\rightarrow$ More accurate technique is needed

Future plans:

- Testing of different neural network topologies in the semileptonic channel
- Train on polarized data, evaluate on unpolarized
- Tackle fully leptonic channel with the NN approach

References

[1] arXiv:hep-ph/0603011

[2] arXiv:0801.3359 [hep-ph]

[3] arXiv:1710.09339

[4] arXiv:1205.2484

[5] arXiv:hep-ph/9406381

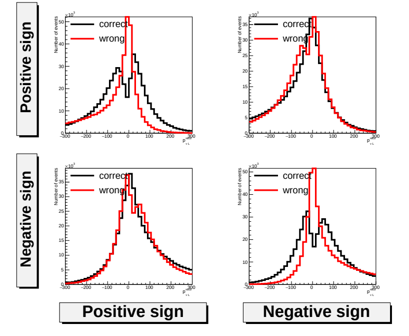
[6] arXiv:hep-ph/9406381

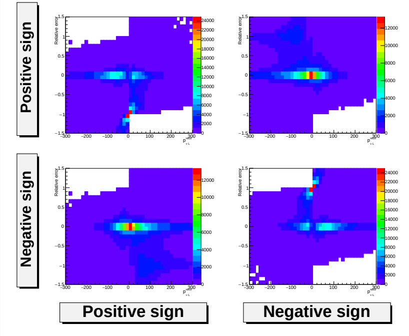
[7] arXiv:0908.0079

[8] B. Hoonhout, K. Oussoren, S. Bentvelse: *Higgs spin analysis in Collins-Soper frame using opening angles of different-flavour final state leptons* (link)

backup

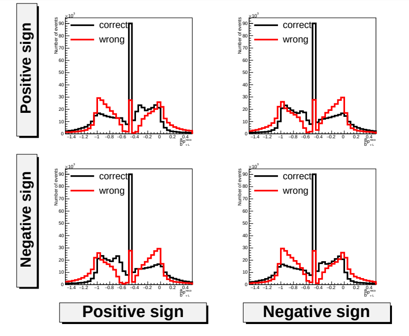
Criteria Selection 1 at Parton Level

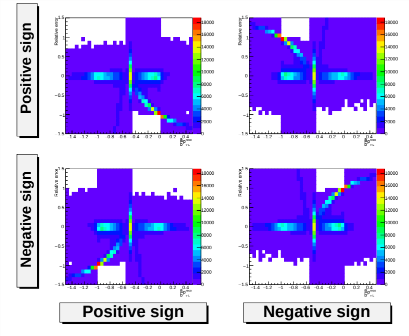




Selection 1 - algorithm discards the solutions which have absolute value smaller than 50 GeV, if both solutions lie under or above 50 GeV, random solution is chosen.

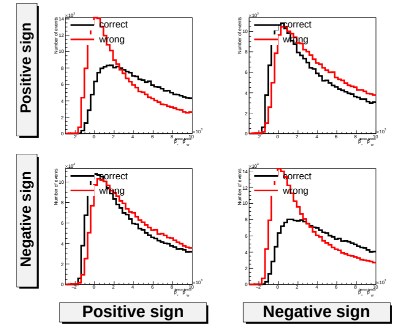
Criteria Selection 2 at Parton Level

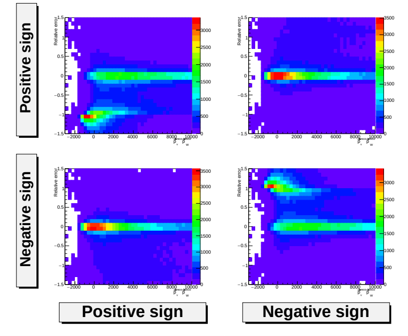




selection 2 - algorithm discards all the solutions for which $-p_L * a/b < 0.5$. If both solutions pass or don't pass this criterium, random solution is chosen.

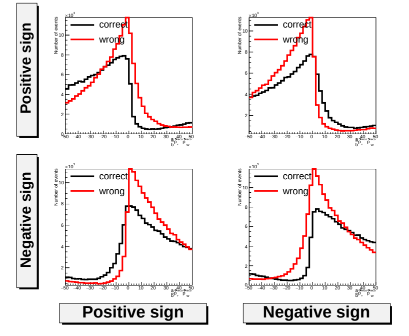
Criteria Selection 3 at Parton Level

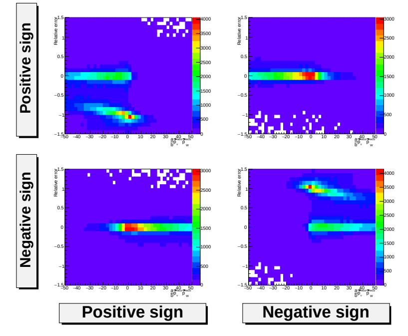




Selection 3 - if the scalar product of the reconstructed neutrino three-momentum (solution is taken for the longitudinal component) with the reconstructed W three-momentum is smaller than $2500 GeV^2$, the solution is discarded. If both solutions lie under or above $2500 GeV^2$, random solution is chosen.

Criteria Selection 4 at Parton Level





selection 4- if the value of the scalar product of the reconstructed neutrino three-momentum (solution is taken for the longitudinal component) with the reconstructed Ws three-momentum (each threaded separately), multiplied by a/b, is smaller than 30 GeV or larger than 25 GeV, the solution is discarded. a and b are the parameters of the quadratic equation. If both solutions pass or don't pass the criterion, random solution is chosen.

Selection criteria - efficiencies

- **Efficiencies**: count number of "correct solutions"
  - correct solution: the one which lies closest to the truth
  - wrong solution: the one which lies further from to the truth
- Solutions with negative discriminant are not taken into account

Fixed m_W (on-shell)

- Random: 49.9 %
- Selection 1: 56.3 %
- Selection 2: 61.3 %
- Selection 3: 46.0 %
- Selection 4: 52.6 %
- Combined selection 44.6 %

Truth m_W (off-shell)

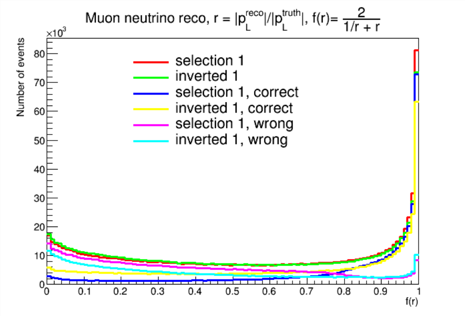
- Random: 50.0 %
- Selection 1: 53.7 %
- Selection 2: 58.4 %
- Selection 3: 48.1 %
- Selection 4: 51.4 %
- Combined selection 46.4 %

It turns out that efficiency is not always a good measure for selection performance

Selection performance

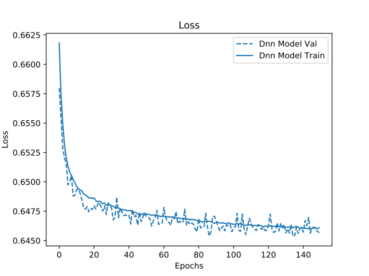
- inverted selection 1 has better efficiency
- inverted selection 1 has worst relative error
- inverted selection 1 has selects better in case of events with bad reconstruction

- Random: 50.0 %
- Selection 1: 42.2 %
- Selection 2: 50.0 %
- Selection 3: 48.2 %
- Selection 4: 45.1 %



Learning curve

1 layer



3 layers