Probing CP violation in $e^+e^- \to t\bar{t}h$ at a 550 GeV Linear Collider with Machine-learned Observable

International Workshop on Future Linear colliders 2026

Jenny List[1], Nana Ozaki[2], Aidan Robson[3], Junping Tian[4], Yuyang Zhang[1,3]

[1]DESY Hamburg, [2]Mount Holyoke College, [3]University of Glasgow, [4]Univeristy of Tokyo





November 11, 2026

Introduction: Motivation

Why study CPV in the top-Higgs sector

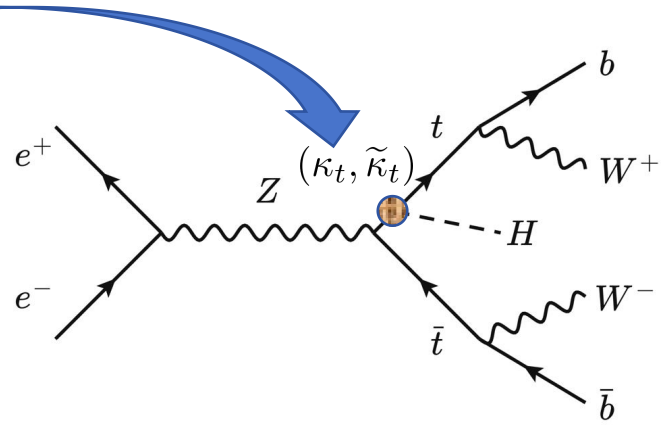
- Flaw of the SM: The SM doesn't provide enough CPV to explain the baryon asymmetry of the Universe。
  ~ 9 orders of maginitude smaller!
- Physics Chance: The top quark has the largest Yukawa coupling and a special role in many BSM scenarios.
- Collider Chance: Linear e^+e^- collider provides a clean environment with well-defined initial state to probe the CP structure of the top-Higgs interaction through a direct phycis process $e^+e^- \to t\bar{t}h$ at higher energy

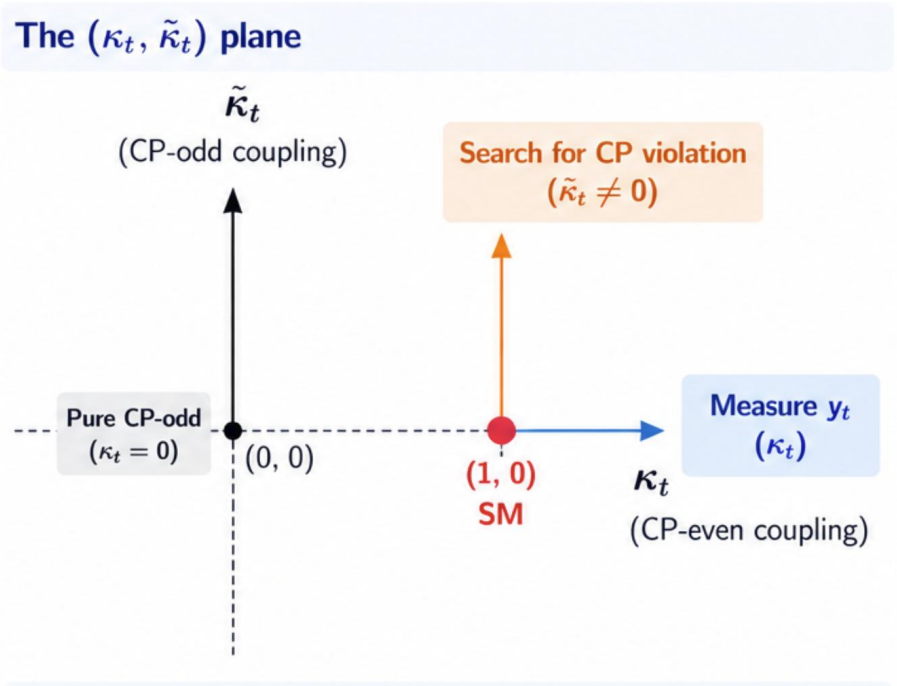
Parameterisation of the top-Higgs Yukawa coupling with CPV source

scalar coupling

$$\mathcal{L}_{t\bar{t}H} = -\frac{m_t}{v} H\,\bar{t}\left(\kappa_t + i\tilde{\kappa}_t\gamma_5\right)t$$

pseudoscalar coupling





Introduction: Research Background

The previous precision study at ILC was done by R.Yonamine et al. in 2011

《Measuring the top Yukawa coupling at the ILC at √ s = 500 GeV》

$$\mathcal{L} = 1ab^{-1}, m_H = 120GeV, (P_{e^-}, P_{e^+}) = (-0.8, +0.3)$$

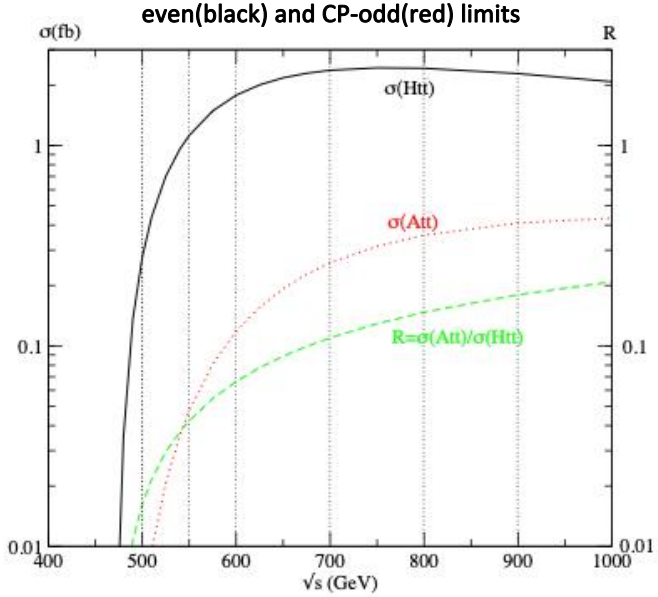
What happens after 2011?

- 125GeV Higgs boson was found in LHC
- New running energy: Study shows $\sqrt{s} = 550GeV$ is a better energy for tth CPV
- New detector design features and updated simulation tools
- New Isolated lepton tagger: Be upgraded to the TMVA based in Marlin
  - *Improve the selection efficiency by 8%, exclude 6% more hadronic background in semileptonic samples*
- New jet flavor tagging: ParT implemented ILC allow us to identify 11 class jet flavor
- New kinematic fitting algorithm: SLD neutrino correction, ErrorFLow, 4C fit...
- New S/B separation method: MVA

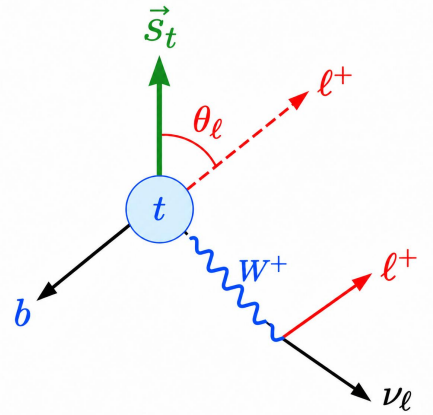
For CPV observables

- Spin-correlation based observables in this collider environment have been studied at theoretical level: 1602.00684toponia, 1712.09953 full top spin calculation, 1809.07127reference frame, K.Hagiwara, K.Ma et al.
- Mahcine Learning become widely used to build optimal observable at detector level

The total cross section of the pure CP-even(black) and CP-odd(red) limits



K. Hagiwara et al. 1712.09953



Event Generation

$$\mathcal{L}_{t\bar{t}H} = -\frac{m_t}{v} H\,\bar{t}\,(\kappa_t + i\widetilde{\kappa}_t\gamma_5)\,t$$

SM Signal: $(\kappa_t = 1, \widetilde{\kappa}_t = 0)$

- Physsim: $e^+e^- \to 6f + h$, 1M for eLpR/eRpL each
- Whzard: $e^+e^- \to t\bar{t}h$, 50K eLpR
  - *Jet-related reconstruction performance study*

CPV Signal: $(\kappa_t = 1, \widetilde{\kappa}_t = 1)$

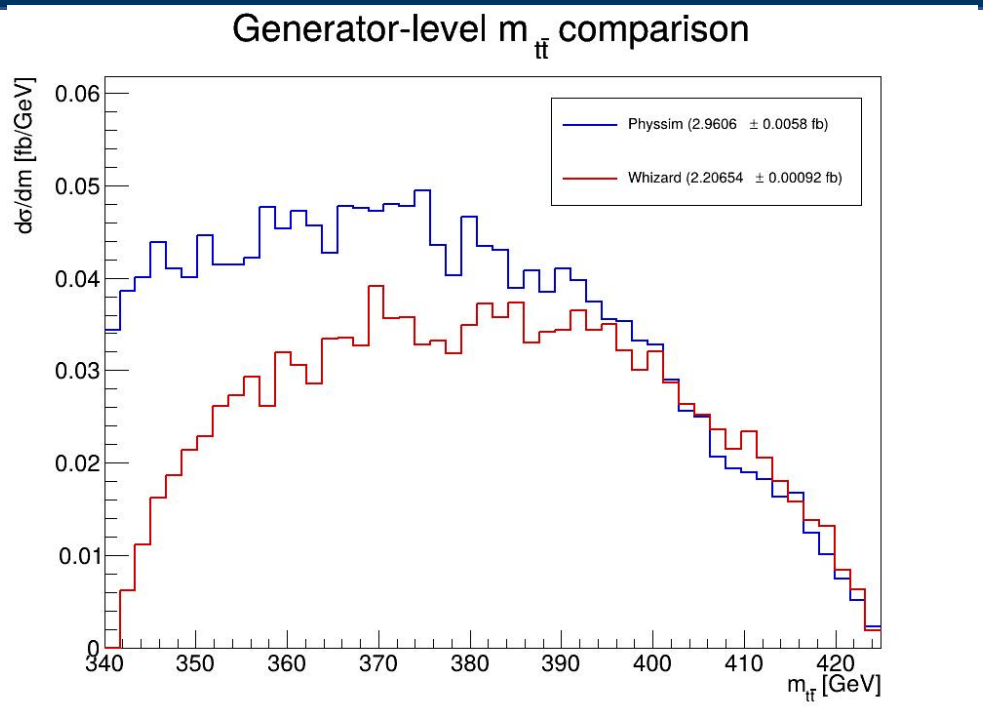
$$|M|^2 = M_S^2 + M_P^2 + 2Re(M_S^* M_P)$$

CP-odd term: change sign under CP transformation but contributes nothing to the total cross-section

- Physsim: $e^+e^- \to 6f + h$, 1M for eLpR/eRpL each

Background

- Physsim: $e^+e^- \to 6f + z$, 1M for eLpR/eRpL each
- Physsim: $e^+e^- \to 6f + g^*,\ g^* \to b\bar{b}$, 1M for eLpR/eRpL each
- Whizard: $e^+e^- \to 6f$, ~8M in total



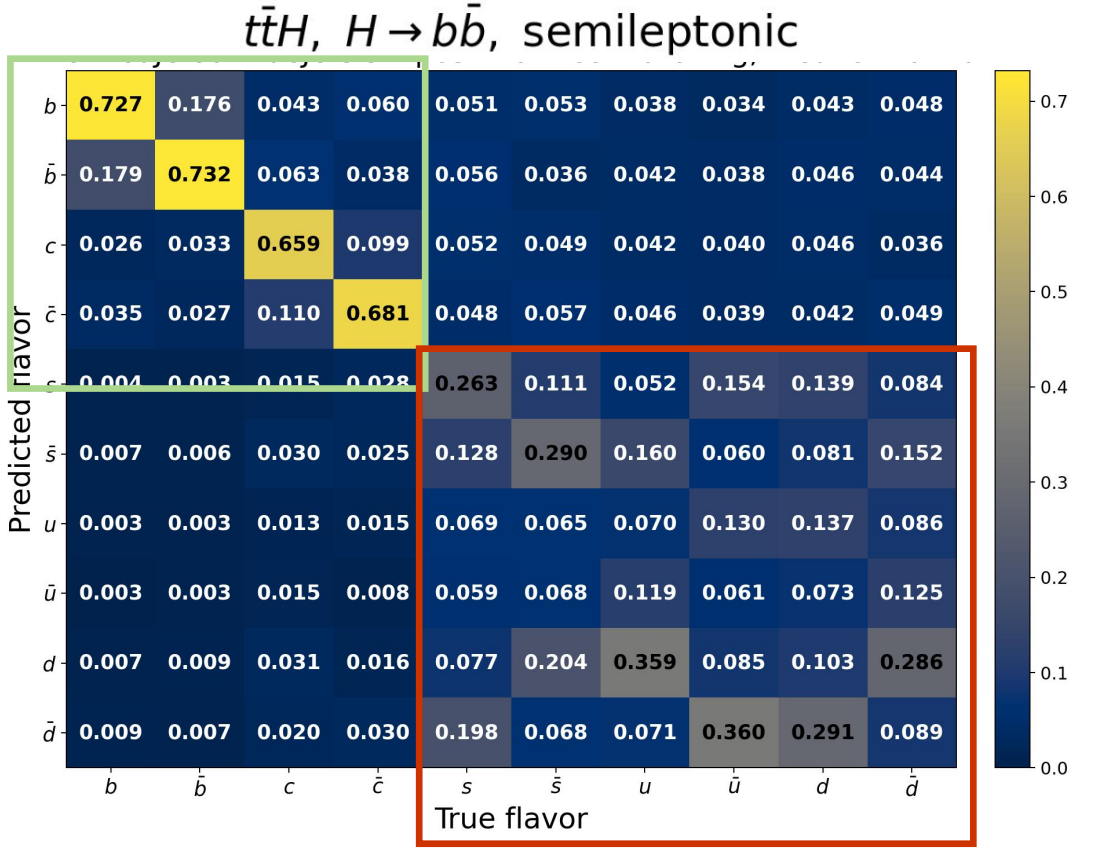
Physsim:

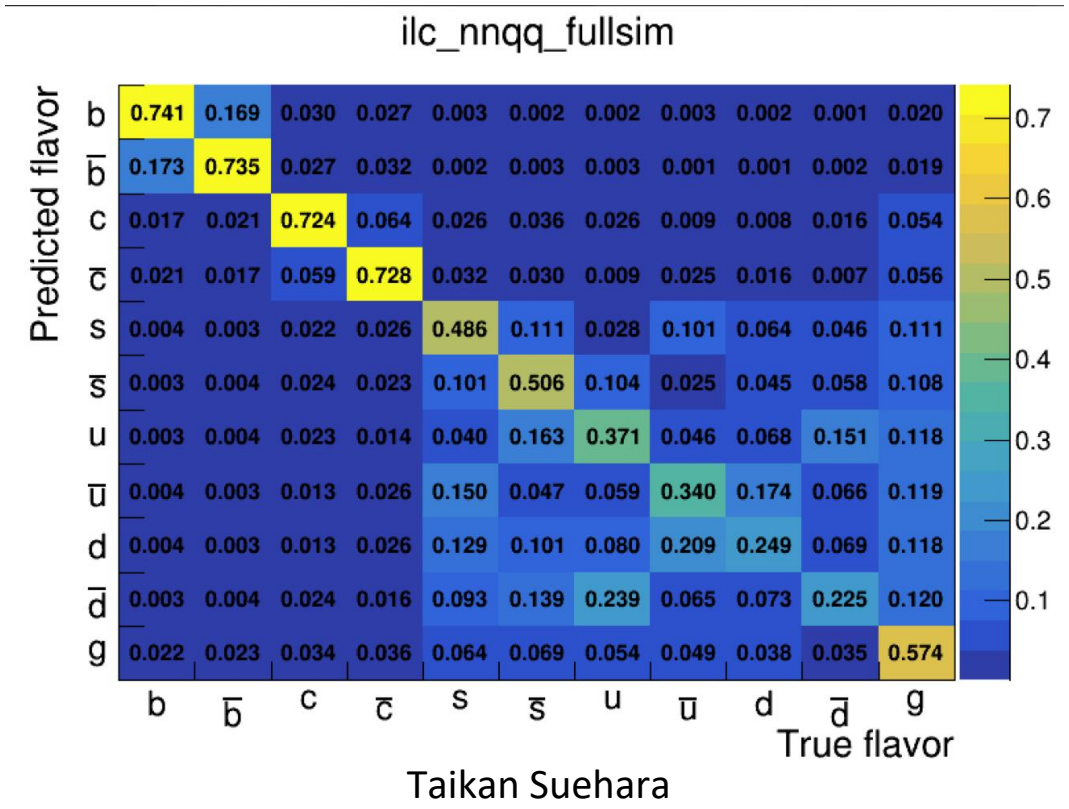
- Corrected by the $t\bar{t}$ bound effect $\sqrt{s} \sim 2m_t + m_H \sim 475GeV$
- Resonant only ($m_{bW} \sim m_{top}$)
- $e^+e^- \to Z^* + h,\ Z^* \to t\bar{t}$ is also included

Detector Simulation: All passed by SGV(Fast detector simulation) successfully.

Jet Flavor Tagging Performance

- In truth tth, h->bb, semileptonic channel, 6 Jets are clustered by Durham algorithm
- The overall performance of ParT is much better than LCFIplus (b/c-tag distribution...see backup)



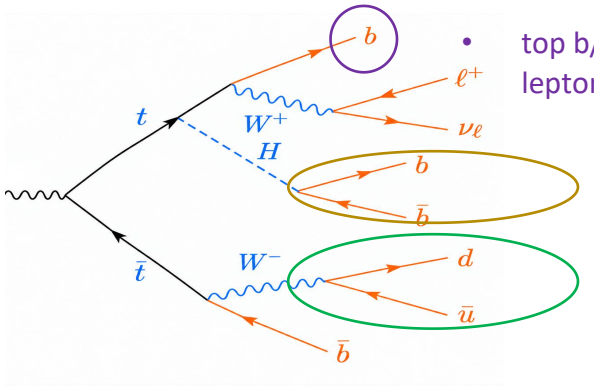


Taikan Suehara

- :) The high accuracy on distinguishing $b/\bar{b}$, $c/\bar{c}$ significantly benefits the jet assignment and signal selection
- :(The lower accuracy on the light jets cause difficulty on the construction of the CP observables, where need to distinguish up/down-type quark.

Event Flavor Score

Event flavor score q_{flavor}



- top b/$\bar{b}$ according to lepton charge
- b/$\bar{b}$ for Higgs
- W cannot decay b

$$S_W = -[ln(p_{non-b}^{w1}) + ln(p_{non-b}^{w2})]$$

$$S_H = -[ln(p_b^{H1} p_{\bar{b}}^{H2} + p_{\bar{b}}^{H1} p_b^{H2})]$$

$$S_{top} = \begin{cases} -[ln(p_b^{lep-b}) + ln(p_{\bar{b}}^{had-b})] & q_l > 0 \\ -[ln(p_{\bar{b}}^{lep-b}) + ln(p_b^{had-b})] & q_l < 0 \end{cases}$$

$$q_{flavor} = S_W + S_H + S_{top}$$

W daughter down-type score q_{down}

- option1: According to the probability of quark/anti-quark

$$q_l > 0, q_{down} = P_q(j) = \sum_{f=u,d,s,c} P_f(j),$$

$$q_l < 0, q_{down} = P_{\bar{q}}(j) = \sum_{f=u,d,s,c} P_{\bar{f}}(j)$$

- option2: According to the probability of s,d,$\bar{s}$, $\bar{d}$

$$q_{down} = \sum_{f=d,\bar{d},s,\bar{s}} P_f(j),$$

$$q_{up} = \sum_{f=u,\bar{u},c,\bar{c}} P_f(j)$$

- The jet which get the higher q_{down} will be identified as down-type quark among the two jets assigned to W
- W daughter assignment likelihood $L_{12} = q^1_{down} q^2_{up}$, $L_{21} = q^2_{down} q^1_{up}$

Kinematic Fitting

- SLD neutrino correction
- ErrorFlow-based jet uncertainties combined detector resolution, PFA confusion and neutrino

$$\sigma^2_{E_{\text{jet}}} = \sigma^2_{\text{Det}} + \sigma^2_{\text{Conf}} + \sigma^2_{\nu}, \quad V_j = \begin{bmatrix} \sigma^2_{E_{\text{jet}}} & 0 & 0 \\ 0 & \sigma^2_{\theta} & 0 \\ 0 & 0 & \sigma^2_{\phi} \end{bmatrix}$$

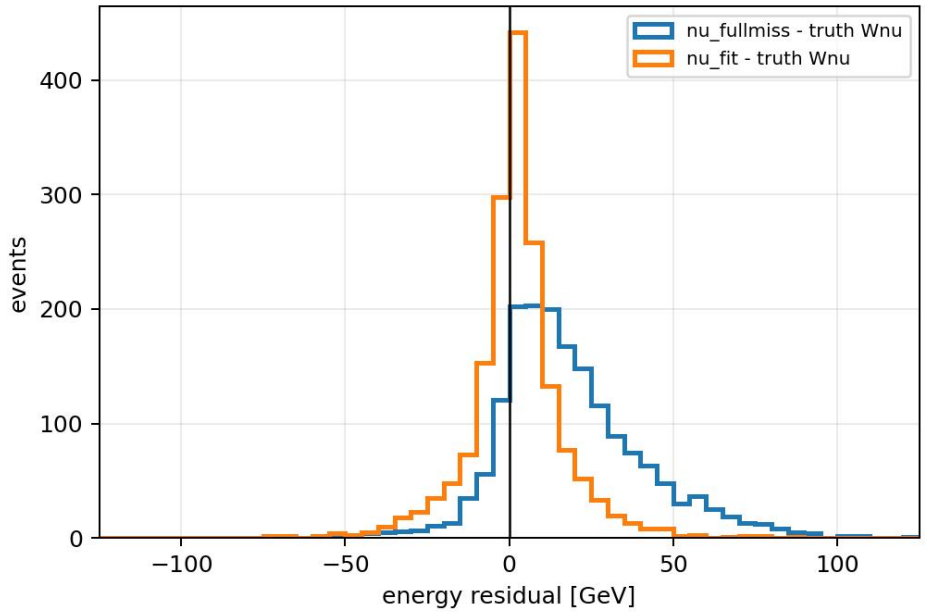
- 4C: E, px, py, pz conservation + ISR

$$\chi^2_{\text{measure}} = \sum_i \left(\mathbf{y}_i^{\text{fit}} - \mathbf{y}_i^{\text{meas}}\right)^T V_i^{-1} \left(\mathbf{y}_i^{\text{fit}} - \mathbf{y}_i^{\text{meas}}\right)$$

- Soft W/t/H mass constraints

$$\chi^2_{\text{mass}} = \left(\frac{m_W^{\text{fit}} - M_W}{\sigma_W}\right)^2 + \left(\frac{m_{t,\text{had}}^{\text{fit}} - M_t}{\sigma_{t,\text{had}}}\right)^2 + \left(\frac{m_H^{\text{fit}} - M_H}{\sigma_H}\right)^2 + \left(\frac{m_{t,\text{lep}}^{\text{fit}} - M_t}{\sigma_{t,\text{lep}}}\right)^2$$

$$\chi^2_{fit} = \chi^2_{measure} + \chi^2_{mass}$$



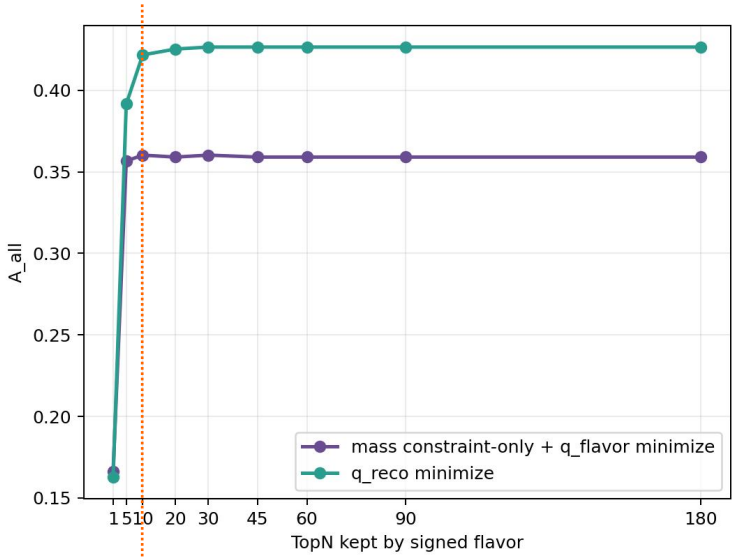
The event reconstruction score q_{reco}

$$q_{reco} = ln(1 + \chi^2_{fit}) + 0.3 q_{flavor}$$

Jet Assignment

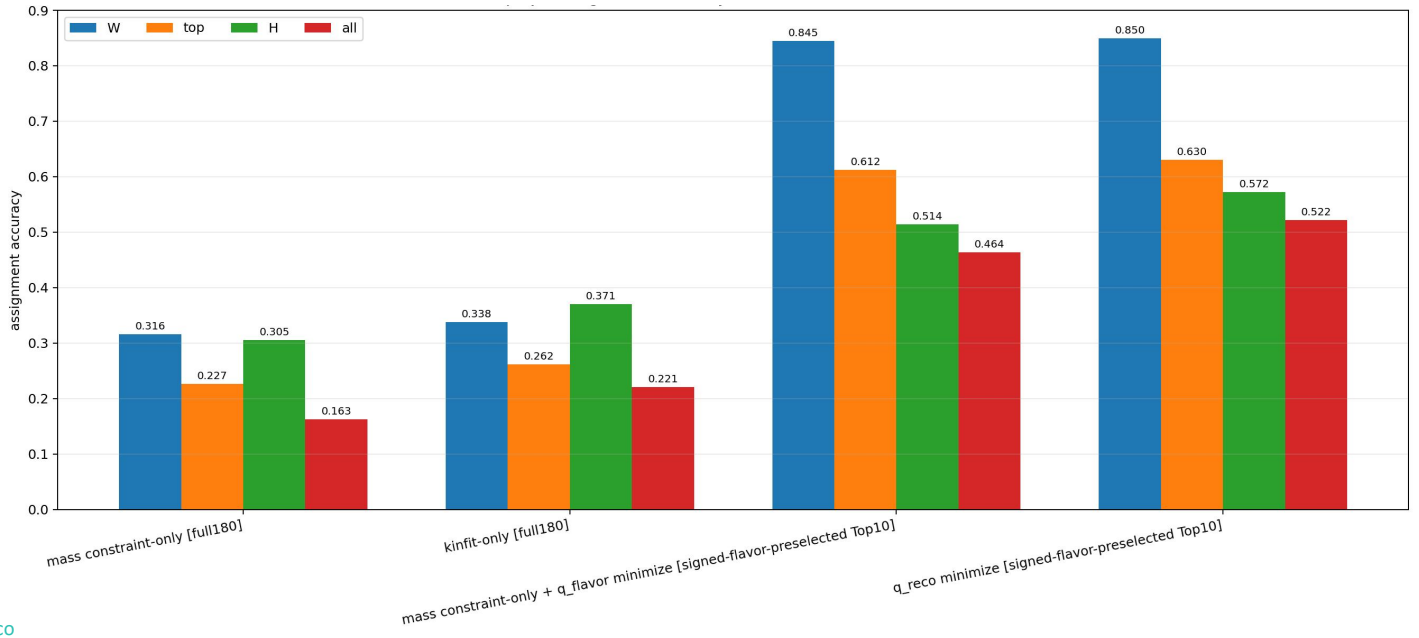
- Reduce the number of the jet permutation

  Rank the 180 permutation by the q_{flavor}

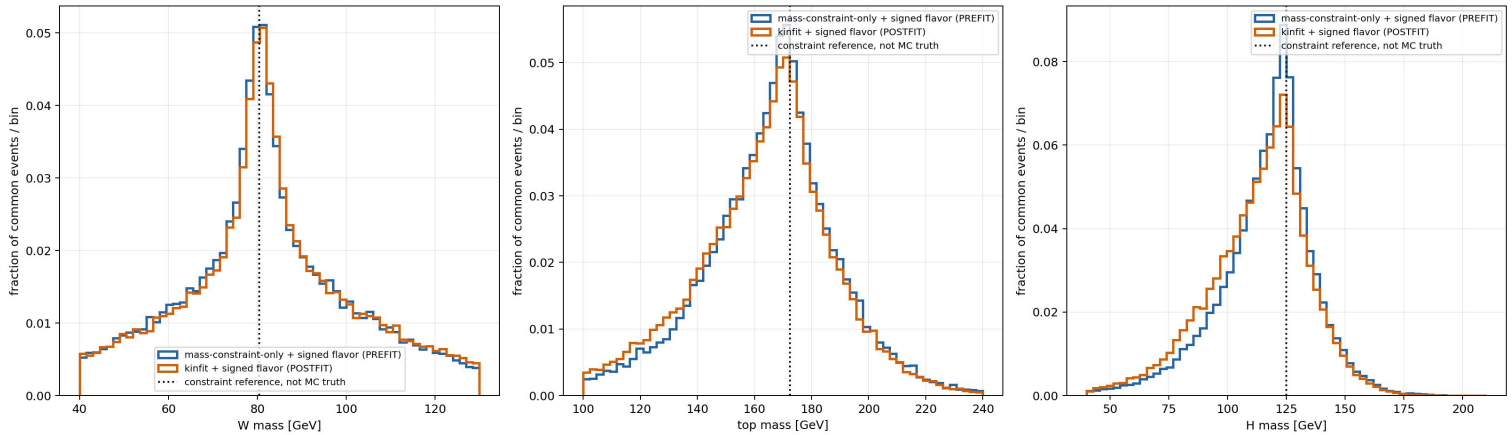


  - Initially, we choose 10 top q_{flavor} assignments to minize the q_{reco}

- Improve the accuracy of the assignment



- Mass reconstruction(Physsim)

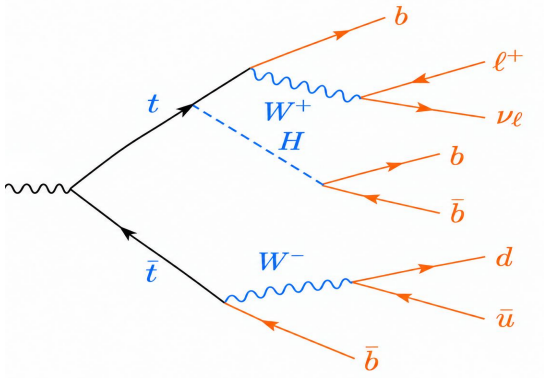


*The TopN scan was tested by 830 selected whizard events, the whole accuracy was plotted by the whole ~4000 whizard event

Cross-Road

Physsim/Whizard → SGV → Semileptonic selection → Jet clustering+flavor tagging → KinFit+Jet Assignment

Reconstructed Physics Objects in **semileptonic channel**



Event level/kinamatic fit quantities

$m_W, m_H, m_t,$
$\chi^2_{fit}, p_{fit}, q_{flavor},$
$q_{reco...}$

Precise measurement of the coupling strength

- Fix $\tilde{\kappa}_t = 0$, measure κ_t
- Optimise signal-background separation

$$Z_{S/B} = \frac{S^{tth}_{SM}}{\sqrt{S^{tth}_{SM} + B}}$$

- MVA S/B classification → q_{SB} → Event Selection

Exploit CP-sensitive information

- Fix $\kappa_t = 1$, constraint $\tilde{\kappa}_t$
- Consider the CPV contribution dependent on the coefficient linearly

$$v_i(\tilde{\kappa}_t) \approx v^{tth}_{SM,i} + \tilde{\kappa}_t v_{cpv,i} + \mathcal{O}(\tilde{\kappa}_t^2)$$

Optimise information

$$I = \sum_i \frac{v^2_{cpv,i}}{v^{tth}_{SM,i}}$$

MVA Signal Background Separation

Model Training

Binary class, XGBoost:

Signal: $tth,\ h \rightarrow b\bar{b}\ (SM)$

Background:

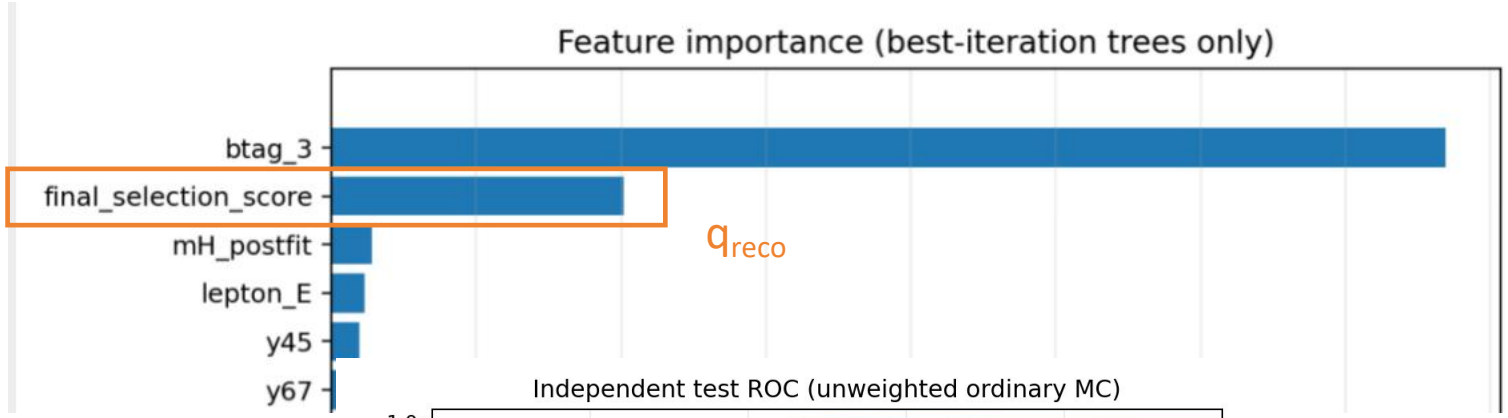
- 6f($t\bar{t}$): 6q+4q2l
- $tth, h \rightarrow non - b\bar{b}$
- $ttz,\ ttg^{*}(g^{*} \rightarrow b\bar{b})$

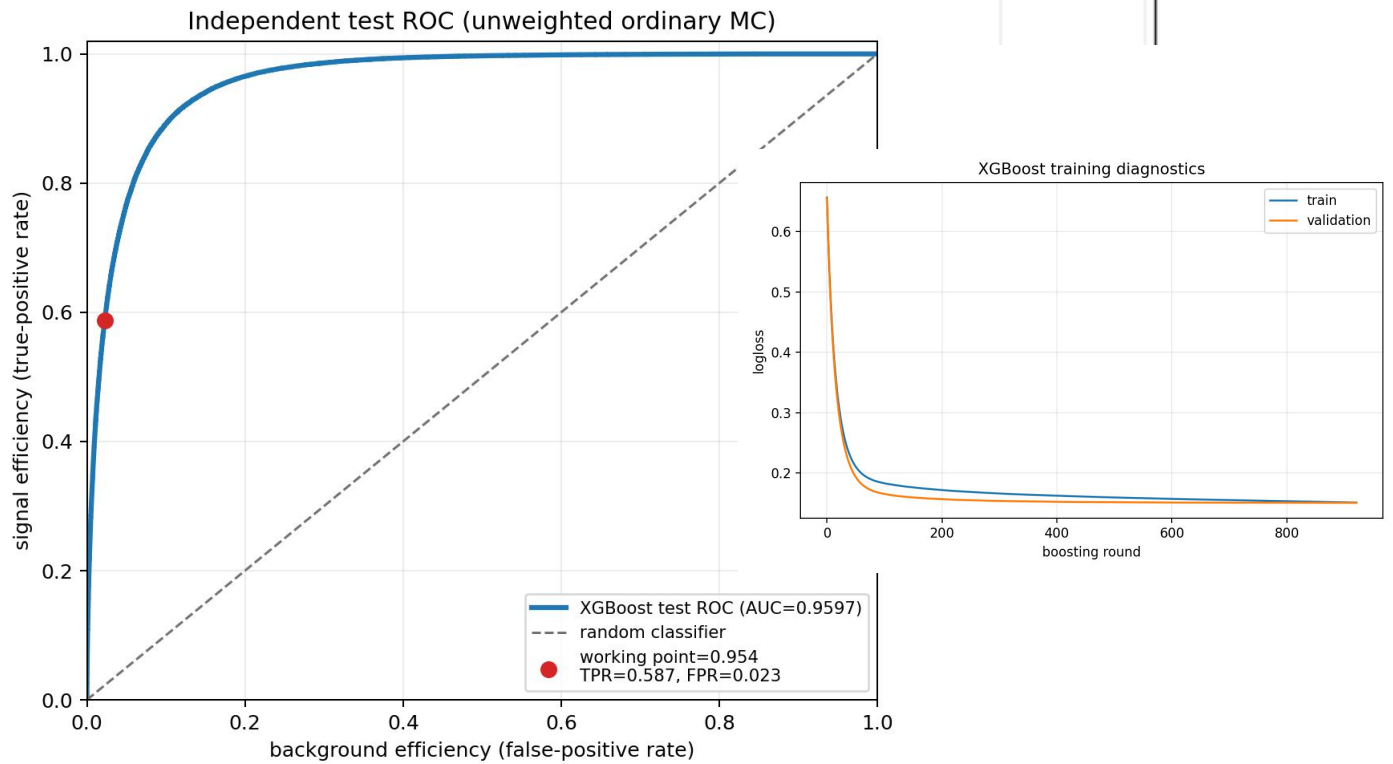
Baseline Input features

- Jet clustering Durham ymerge： y45,y56,y67
- Jet Flavor Tagging:btag1-btag4, ctag1-ctag4
- Jet assginment and kinfit score:
  χ^2_{fit} , p_{fit} , q_{flavor}, q_{reco}

Invariant Mass post kinfit: m_H, m_W, m_{top}

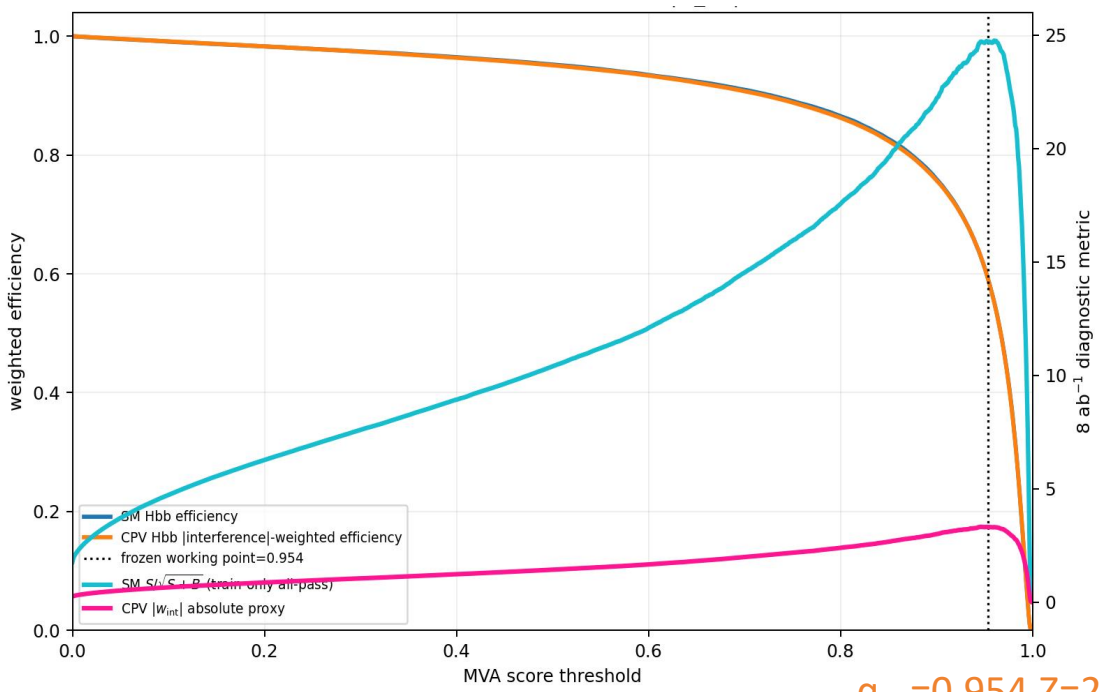
- Isolated lepton: E_{lepton}, θ_{lepton}, pt_{lepton}, lepton_charge





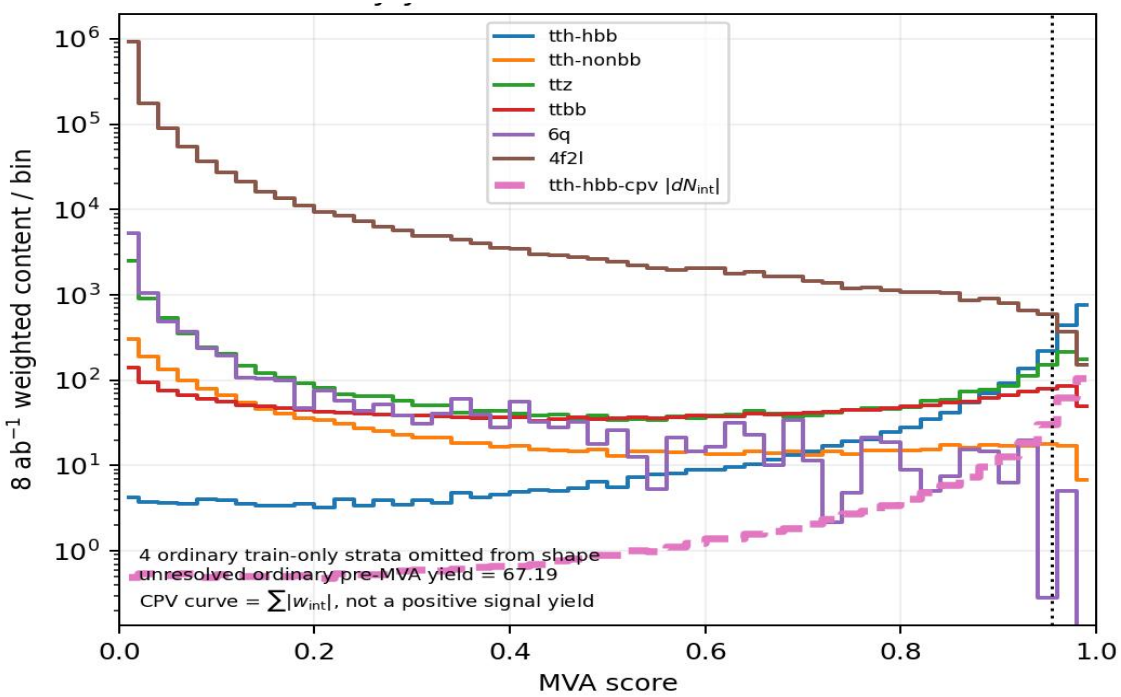
Signal Significance

Efficiency vs Significance



q_{sel}=0.954,Z=25

MVA score distribution of the signal and background



Selection Efficiency and physical yield (LCF polarization, 8 ab^{-1} in total)

sample	tth-sm/tth-hbb	tth-cpv	tth-sm/tth-nonbb	ttz	ttbb	6q	4f2l
Efficiency	58.68%	58.87%	1.67%	5.44%	6.29	0.7%	0.4%
Yield	1292	177	30	445	176	5	692

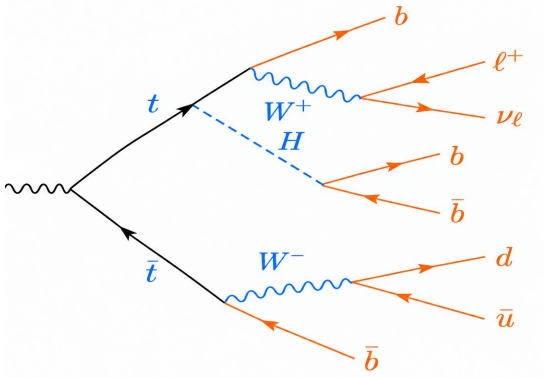
Compare with the 2011 study

- Our results in <u>semileptonic channel</u> $Z = 25 \implies |\Delta\kappa_t/\kappa_t| = 2\%$
- Generalize the 2011 study in "<u>semileptonic channel+ hadronic channel</u>" to 8ab, 550GeV in 2503.19983: $|\Delta\kappa_t/\kappa_t| = 1.9\%$
- Expect higher accuracy in hadronic channel, so the total accuracy will be smaller than 1.4%

Cross-Road

Physsim/Whizard → SGV → Semileptonic selection → Jet clustering+flavor tagging → KinFit+Jet Assignment

Reconstructed Physics Objects in **semileptonic channel**



Event level/kinamatic fit quantities

$m_W, m_H, m_t,$
$\chi^2_{fit}, p_{fit}, q_{flavor},$
$q_{reco...}$

Precise measurement of the coupling strength

- Fix $\tilde{\kappa}_t = 0$, measure κ_t
- Optimise signal-background separation

$$Z_{S/B} = \frac{S^{tth}_{SM}}{\sqrt{S^{tth}_{SM} + B}}$$

- MVA S/B classification → q_{SB} → Event Selection
- Got $|\Delta\kappa_t/\kappa_t|$=2% at in semi-leptonic channel

Exploit CP-sensitive information(eLpR, 8ab^{-1})

- Fix $\kappa_t = 1$, constraint $\tilde{\kappa}_t$
- Consider the CPV contribution dependent on the coefficient linearly

$$v_i(\tilde{\kappa}_t) \approx v^{tth}_{SM,i} + \tilde{\kappa}_t v_{cpv,i} + \mathcal{O}(\tilde{\kappa}_t^2)$$

Optimise **Fisher Information**

$$I = \sum_i \frac{v^2_{cpv,i}}{v^{tth}_{SM,i}} \qquad \tilde{\kappa}_t \sim \frac{1}{\sqrt{I}}$$

CPV observable

$$|\mathcal{M}(x)|^2 = \kappa_t^2|\mathcal{M}_S(x)|^2 + \tilde{\kappa}_t^2|\mathcal{M}_P(x)|^2 + 2\kappa_t\tilde{\kappa}_t\mathrm{Re}\left(\mathcal{M}_S^*(x)\mathcal{M}_P(x)\right)$$

CP-odd term in the $|M(x)|^2$:

$2Re(M(x)_S^* M(x)_P)$ → $\mathcal{O}_{cp}(x)$

CP $x \to \bar{x}$:
- particle →anti-particle
- $\vec{p} \to -\vec{p}$

$-2Re(M(\bar{x})_S^* M(\bar{x})_P)$ → $-\mathcal{O}_{cp}(\bar{x})$

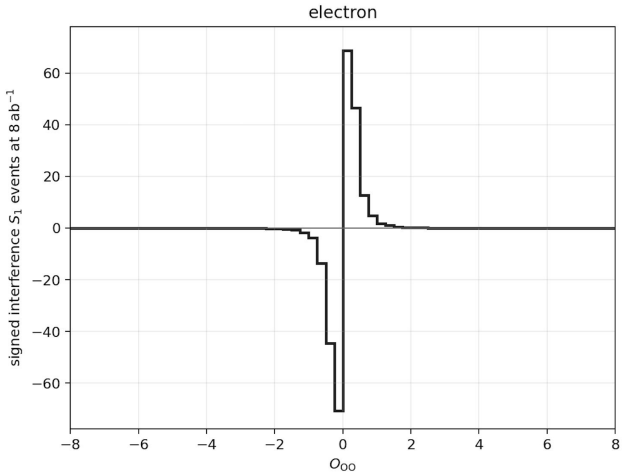
CPV observable: Theoretical Optimal Observable

$$2Re(M(x)_S^* M(x)_P) \rightarrow \mathcal{O}_{cp}(x) \rightarrow -\mathcal{O}_{cp}(\bar{x}) \leftarrow -2Re(M(\bar{x})_S^* M(\bar{x})_P)$$

CP $x \rightarrow \bar{x}$:
- particle →anti-particle
- $\vec{p} \rightarrow -\vec{p}$

Optimal Observable(OO):

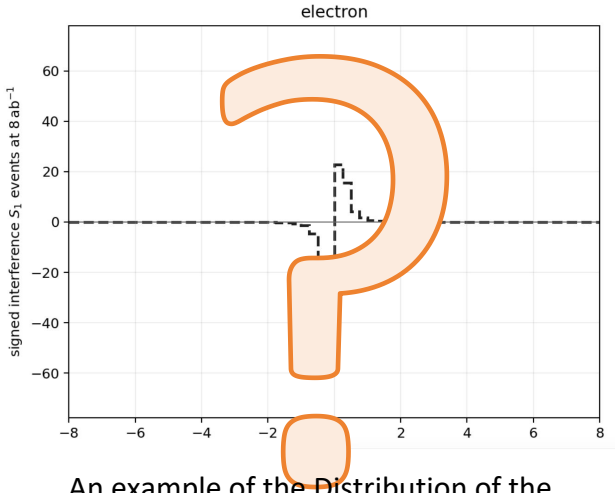
$$O_{\mathrm{OO}} = \frac{2\,\mathrm{Re}[\mathcal{M}_{\mathrm{S}}^* \mathcal{M}_{\mathrm{P}}]}{|\mathcal{M}_{\mathrm{S}}|^2}$$



Distribution of the theoretical Optimal Observable in electron channel weighted by the sign of the interference term

Reconstruction status
- Semileptonic selection: 6% lepton lost, 23% contaminated by other channels
- 50% events exist mis-matched jets
- Nearly random guess on down-type identification

$$I_{OO}^{\mu+e} = 144$$



An example of the Distribution of the theoretical Optimal Observable in electron channel weighted by the sign of the interference term (not true)

CPV observable: Top Spin correlation angular observable and ML

Angular observable

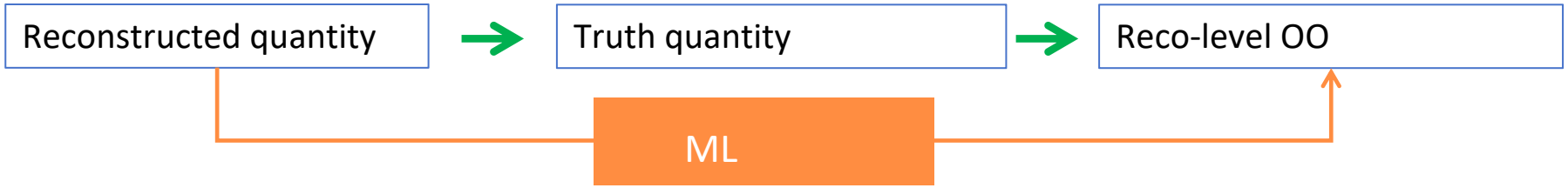
$$\mathcal{M}_{\text{int}} \sim \epsilon_{\mu\nu\rho\sigma}\, p_i^\mu p_j^\nu p_k^\rho p_l^\sigma \rightarrow \vec{p}_a \cdot (\vec{p}_b \times \vec{p}_c)$$

- Pseudo scalar coupling also alters the spin distribution of the top quark, which allow us detect the CP interference from the spin correlation of the top decay products, where down-type fermion is recognised as the maximal spin analyser.

$$\frac{1}{\Gamma}\frac{d\Gamma}{d\cos\theta_f} = \frac{1}{2}(1 + \kappa_f P_t \cos\theta_f) \qquad |\kappa_{\ell\pm}| \simeq |\kappa_{d\text{-type}}| \simeq 1$$

- Project partial information of the whole matrix element but easier and clearer, affected less from the reconstruction issue

Machine Learning:



For CPV: Train binary classifier targeted as sign($2Re(M_S^* M_P)$) ➡ P(+), P(-)

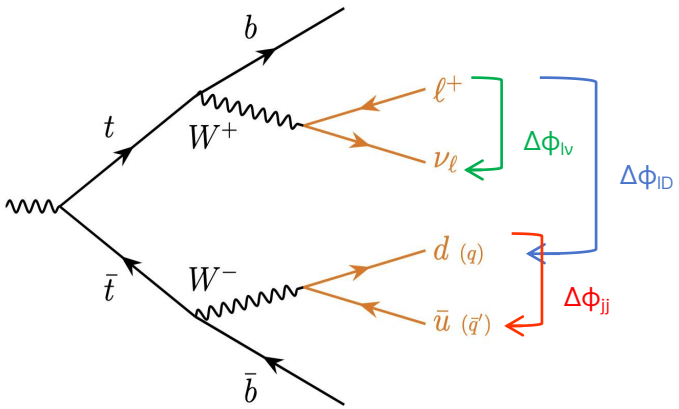
$$O_{ML} = P(+) - P(-) \in [-1,1]$$

Previous studies on the ZH coupling at 250GeV with this observable can be seen in Phys. Rev. D 113,115026

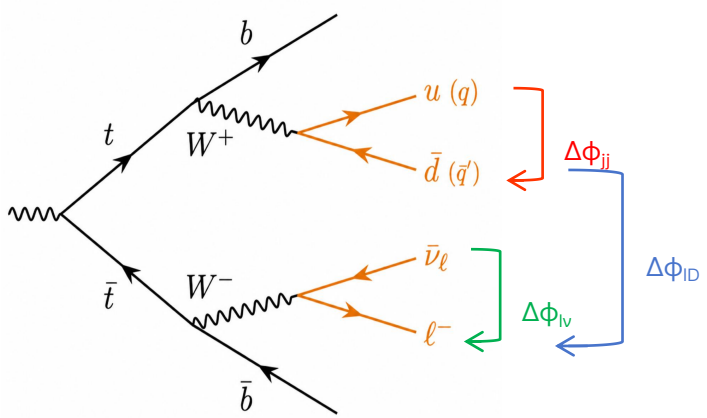
CPV Angular observable baseline at Higgs rest frame

**Higgs rest frame was studied to be the one who contains the most CPV information by Kai Ma in 1809.07127, and we also proved it with our data*

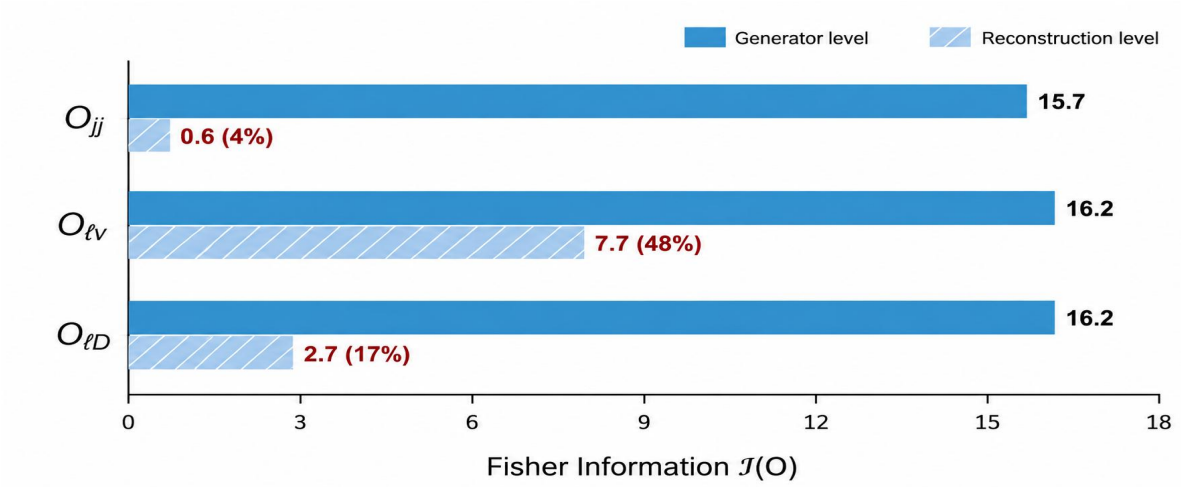
(a) lepton charge > 0



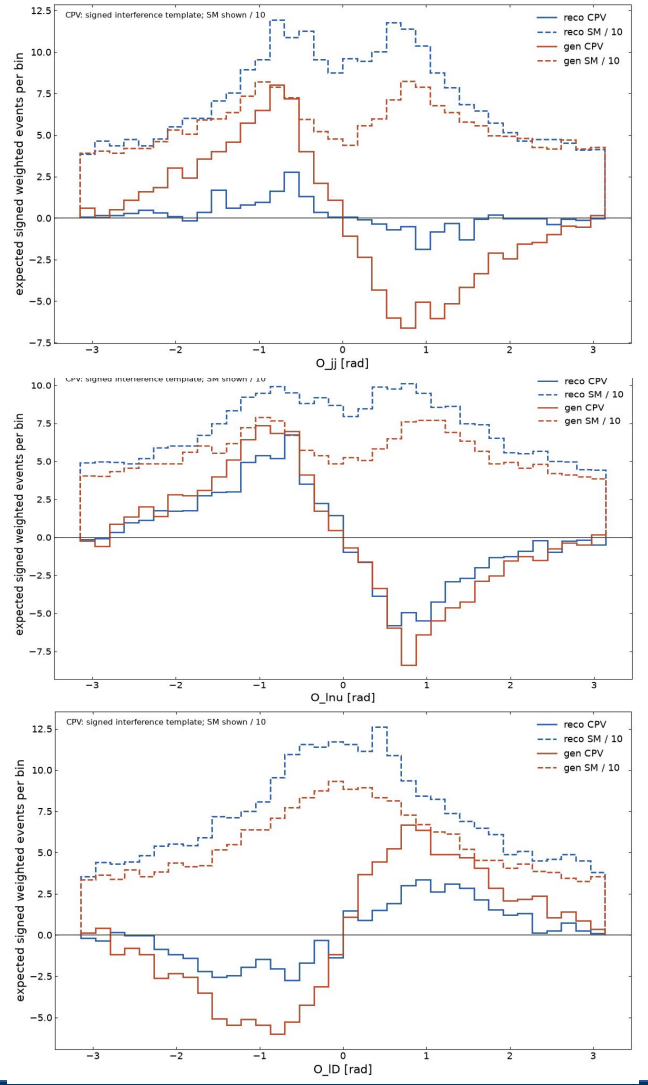
(b) lepton charge < 0



Fisher information combined muon and electron channel



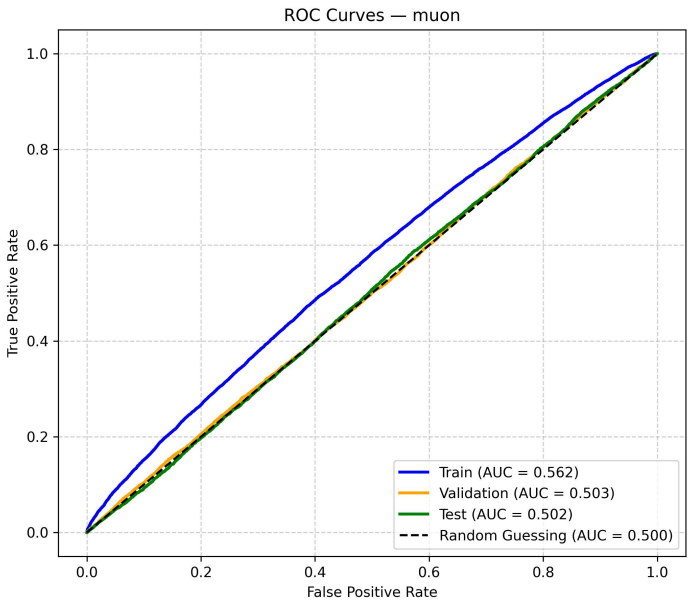
Angular distribution of the muon channel



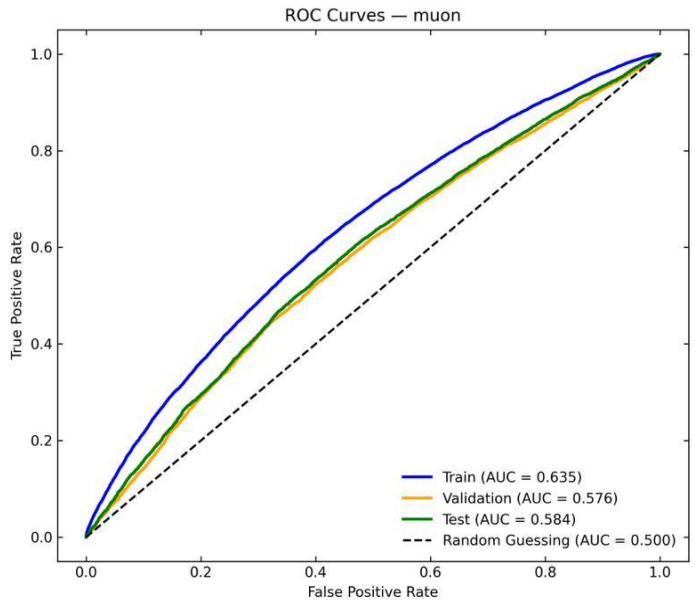
ML observable: From 0 to 1 input by particle

Angular O_ID vs ML[l, D]

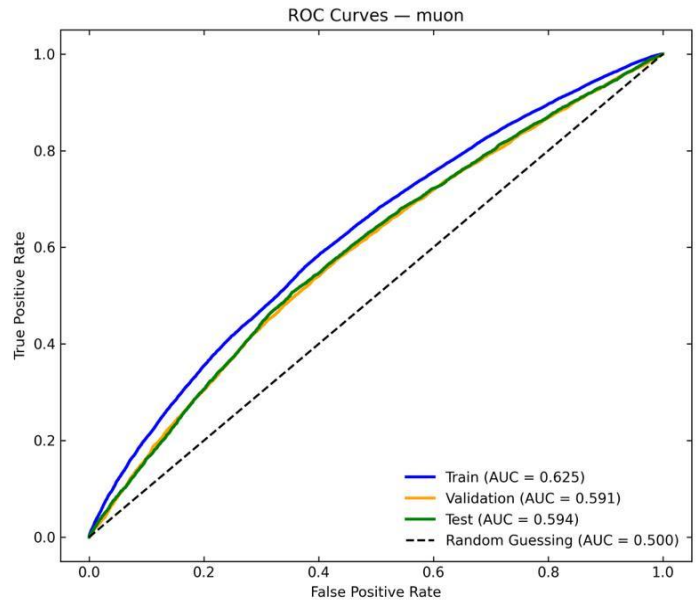
lepton + w down type daughter[E,θ,pT,ϕ)



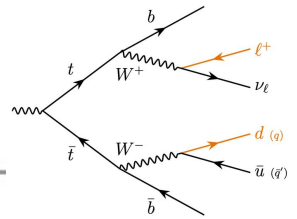
+lepton charge



->top side fermion+ anti-top-side fermion







- w assignment likelihood:

$$L_{12} = P_{down}(j_1)P_{up}(j_2),$$
$$L_{21} = P_{up}(j_2)P_{down}(j_1)$$

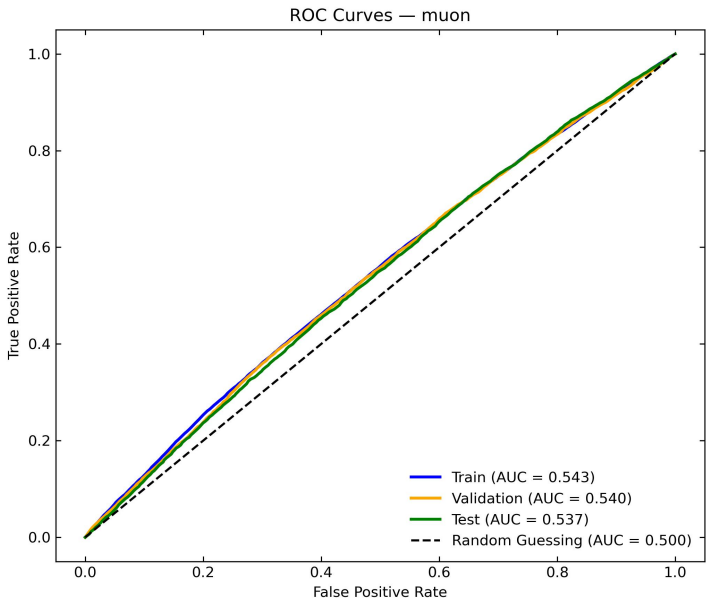
* We input both option defined in p6

* The Fisher of O_{ID} here is also evaluated by the ML test set

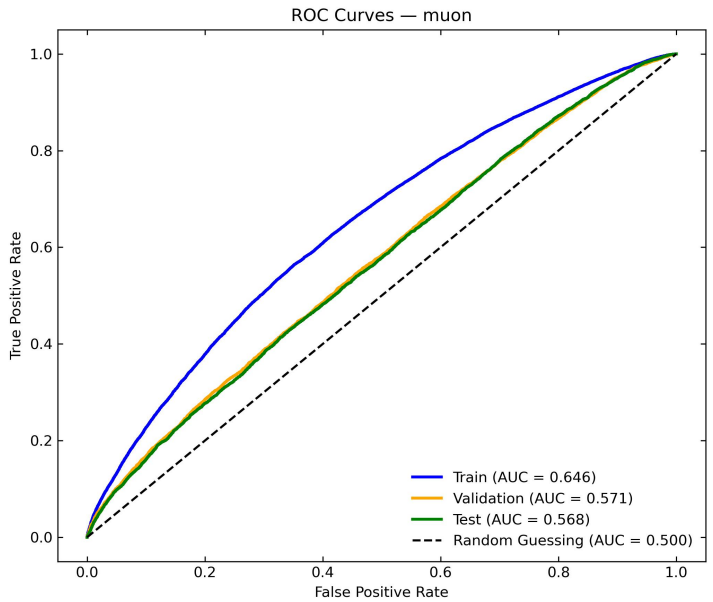
ML observable: From 0 to 1 input by azimuthal angle

Angular O_jj vs ML[O(jj)]

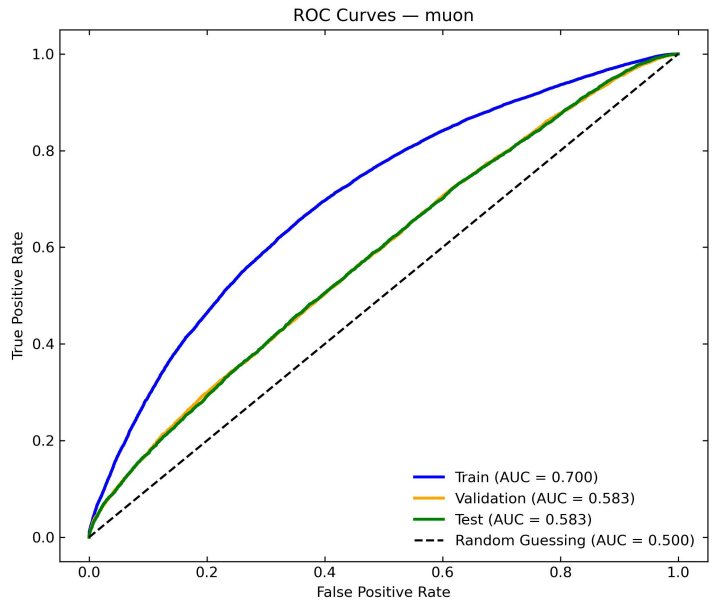
O_{jj}

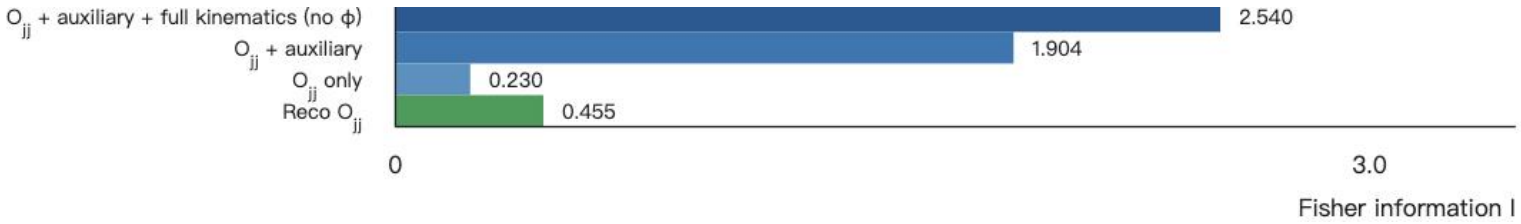


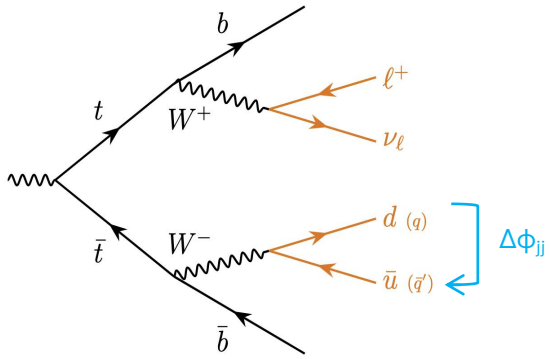
+auxilliary



+all decay kinematics without phi

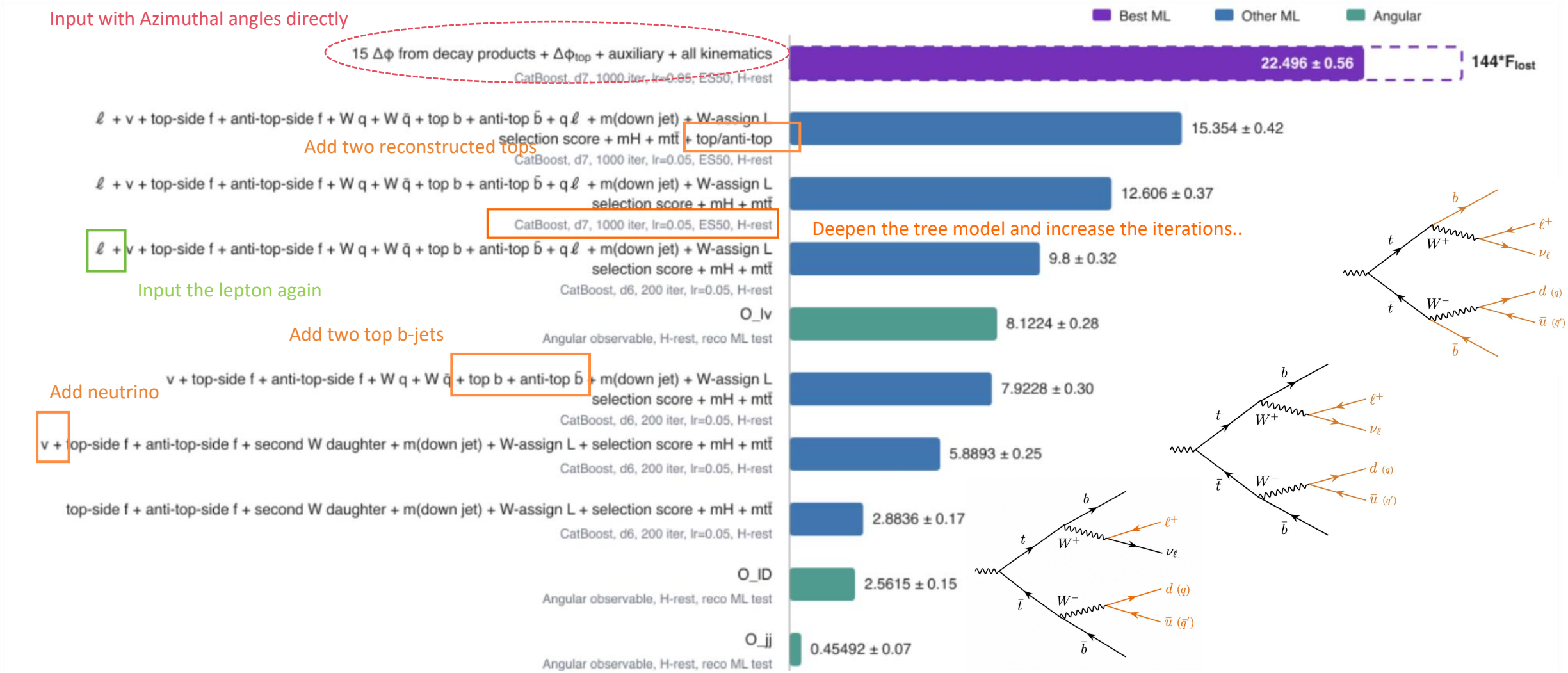






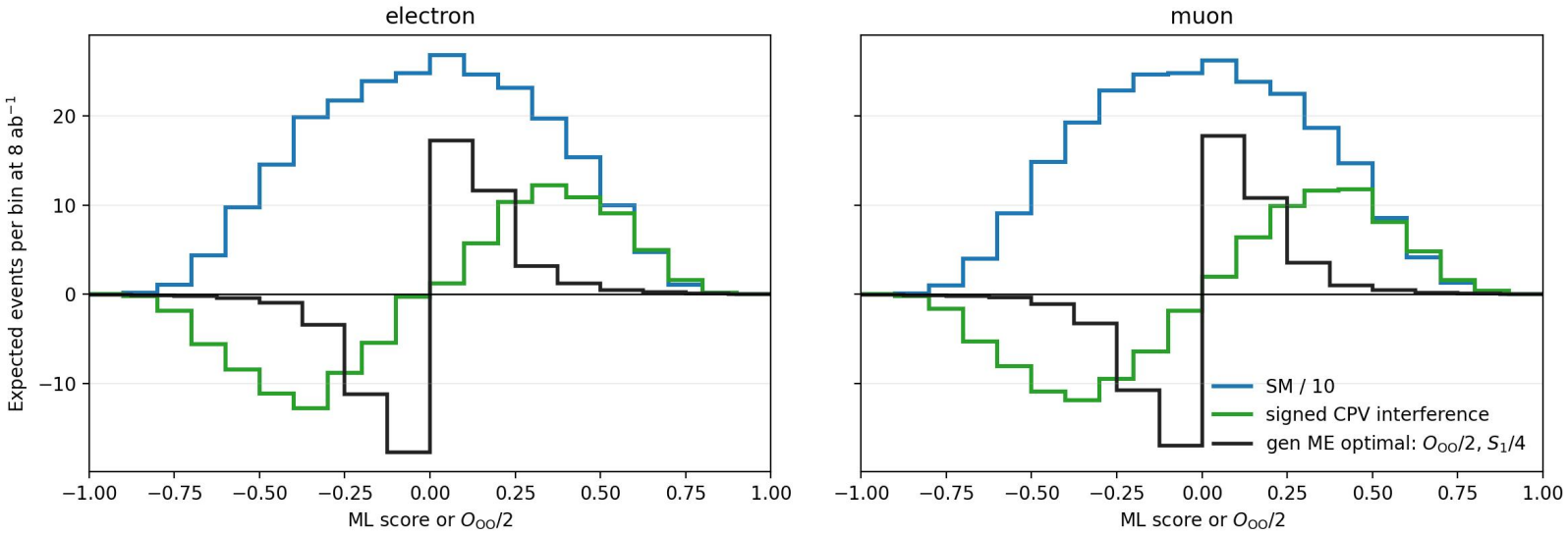
* The Fisher of O_{jj} here is also evaluated by the ML test set

ML observable: Approach to the Optimal at Reco Level

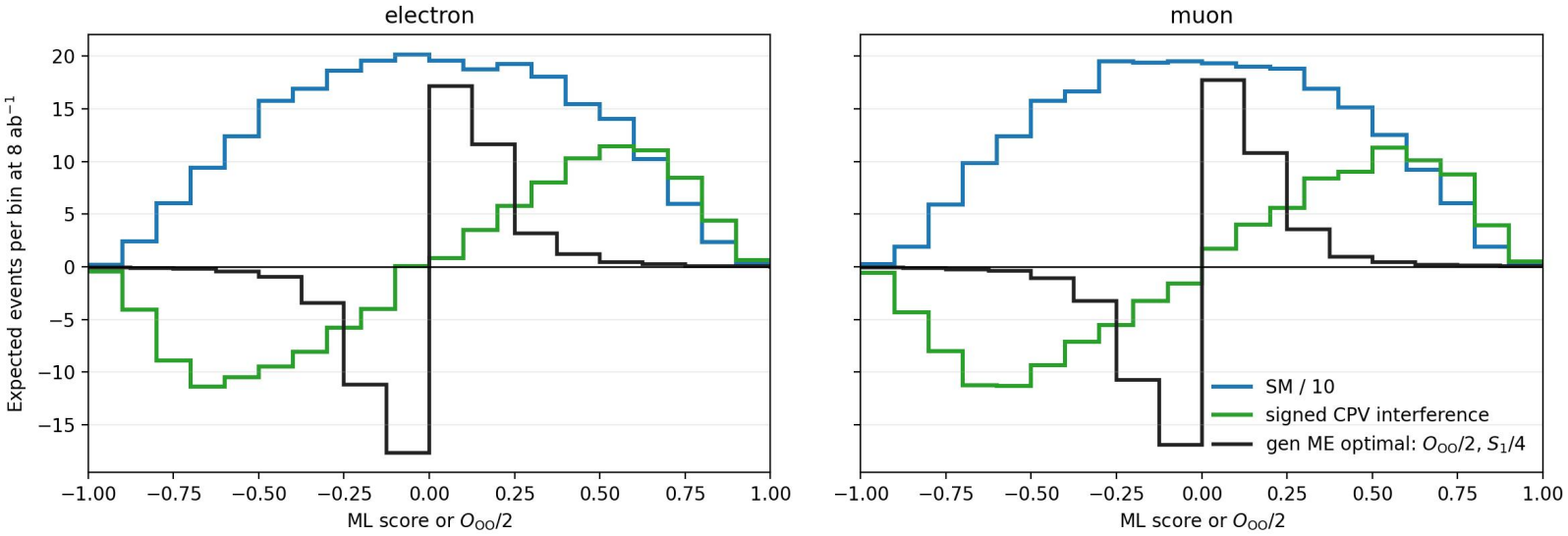


ML observable: Compare with Theoretical Optimal Observable

Optimal Particle Input Model



Optimal Angular Input Model

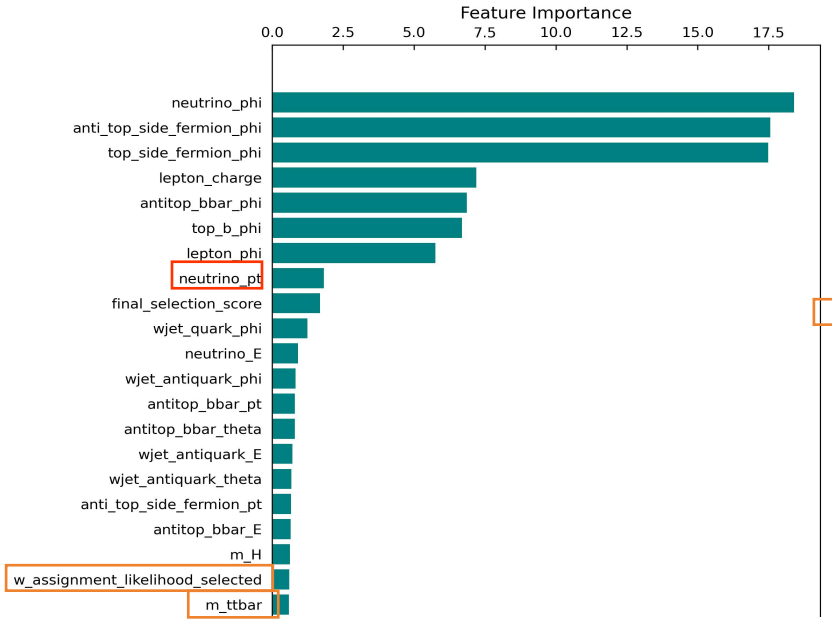


ML observable: What is the model actually learned

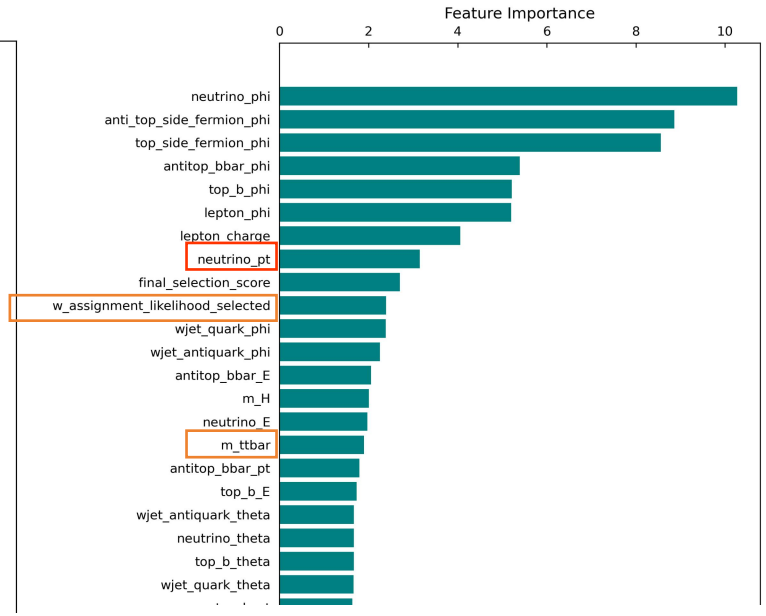
- Project the full spin correlation interference $M^*_{S,\lambda} M_{P,\lambda'}$ into a one-dimensional observable



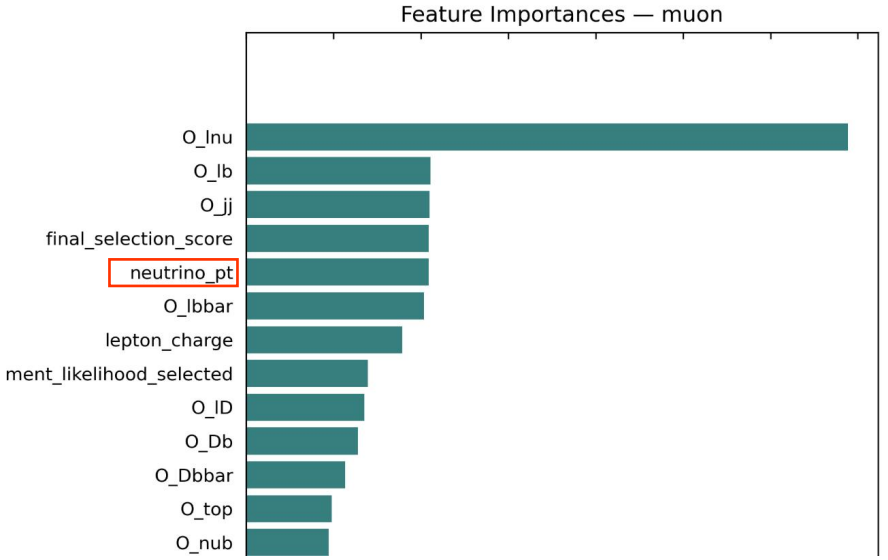
- Feature importance changing after tuning the model hyper parameters



ndepth=6, 200iter,learning rate=0.05, particle input



ndepth=7, 1000iter,learning rate=0.05, particle input



ndepth=7, 1000iter,learning rate=0.05, angular input

Other kinematics such as E, θ, pT of some particles show less relevant to the training.

1. Auxiliary variables related to the whole event topology become to gain more importance

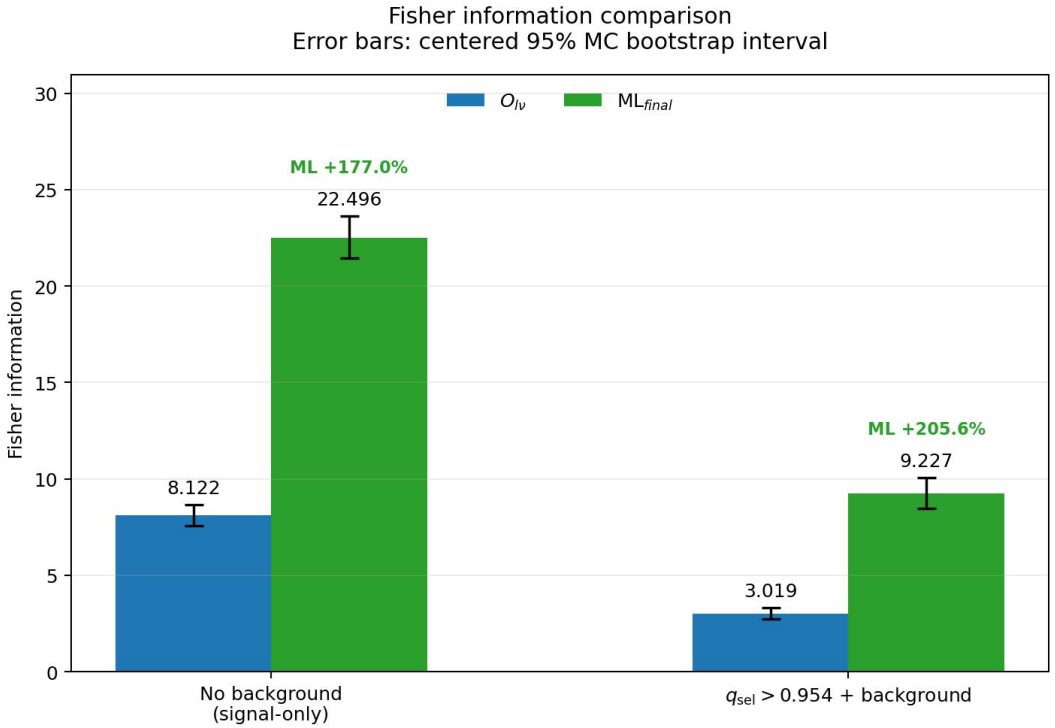
2. The distribution become more balance

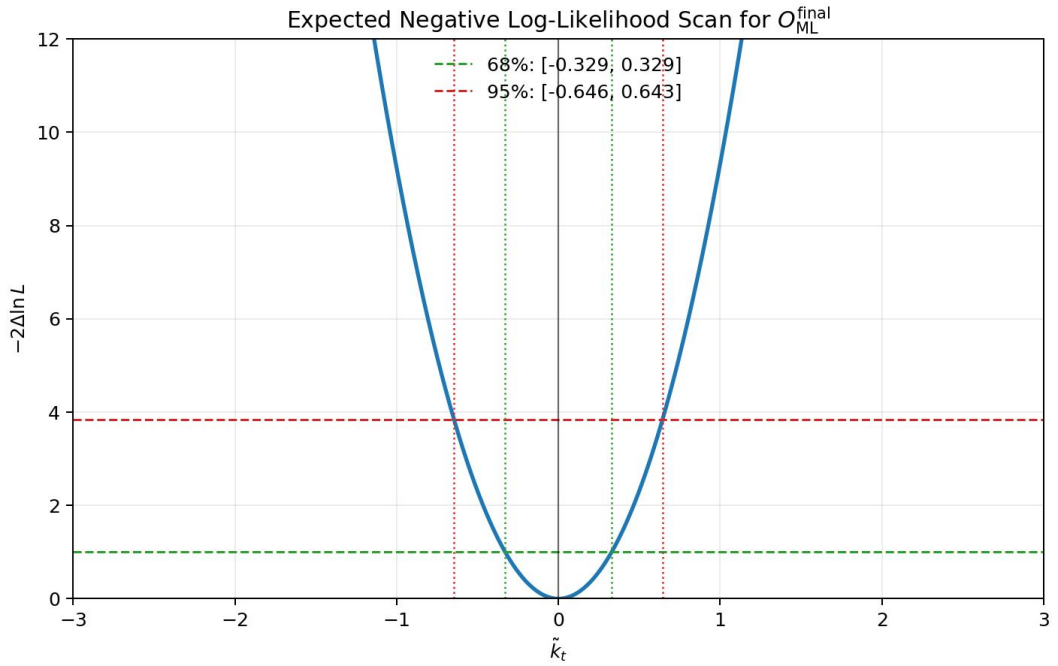
3. Neutrino pT shows special importance in all case:
   More information concentrates on the region with higher pT

ML observable: A preliminary limit setting with background:

Comparison after applying background (with the best model)

Evaluated by q_{sel}>0.954, 8ab^{-1},eLpR





Translate to other paramters

$$\mathcal{L}_{t\bar{t}H} = -\frac{m_t}{v} H\,\bar{t}\,(\kappa_t + i\tilde{\kappa}_t\gamma_5)\,t$$

- CP-mixing angle $\kappa_t = \kappa\cos\alpha = 1, \tilde{\kappa}_t = \kappa\sin\alpha$

  α≤ 0.10π at 68%, α≤0.18π at 95%,

- SMEFT $\kappa\cos\alpha = 1 - \frac{v^3}{\sqrt{2}\,m_t}\frac{C_{t\varphi}}{\Lambda^2}, \quad \kappa\sin\alpha = -\frac{v^3}{\sqrt{2}\,m_t}\frac{C^I_{t\varphi}}{\Lambda^2}.$

  $\left|\frac{C^I_{t\phi}}{\Lambda^2}\right| \leq 5.39\ \mathrm{TeV}^{-2}$ at 68% $\left|\frac{C^I_{t\phi}}{\Lambda^2}\right| \leq 10.64\ \mathrm{TeV}^{-2}$ at 95%

ML observable: What can we expect further

Learn the reconstruction ambiguity

- Fundamental side:
  - Higher accuracy of light jet tagging from ParT for tth samples
- Optimised the Binary Model：
  - Add the two b-jets assigned to Higgs, let the model choose it by itself
  - Train an additional Neural Network on jet assignment

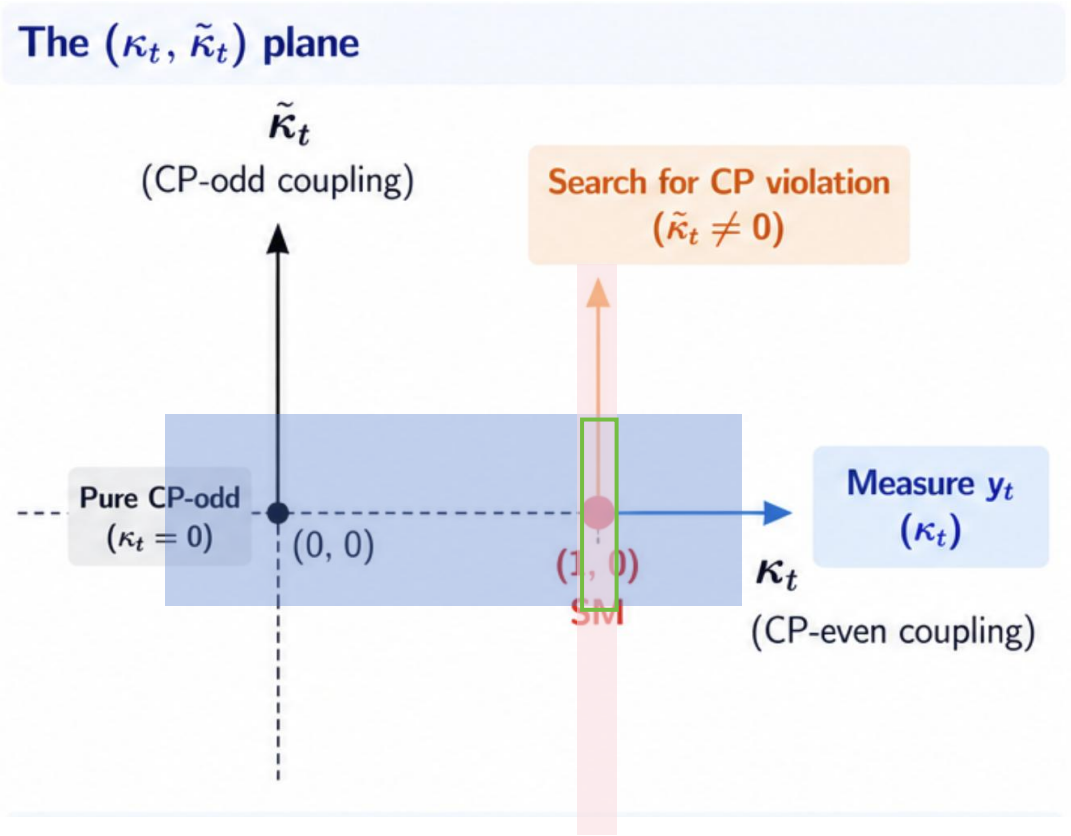
Learn SM and the background:

- Train three class model by adding addional SM class
- Add the background into the SM dataset, train them together
  - **Haven’t see significant improvement so far**

Train 4 models to combine the 4 polarisation configuraton at LCF

Summary

- After the implementation of the modern reconstruction tools at ILC with SGV fast simulation, the top-yukawa coupling strength can be constrained to accuracy of 2% in semileptonic channel, nearly 100% improves from the 2011 study.
- The ML observable has been studied to exploit the CP information from the CP-odd term directly, the best model so far constraints the tth pseudo-scalar coupling $|\tilde{\kappa}_t| \leq 0.33$ at 68% CL for the pure eLpR sample with 8ab^{-1} in the semileponic channel.
- Many potential method to improve the ML observable to the optimal level in the future.
- The likelihood constructed by kinematic fitting and ParT plays the most important role in this analysis for both side.
- The scalar coupling and the pseudo-scalar coupling will be constrained together later by fixing the $\kappa^2 + \tilde{\kappa}^2 = 1$ to compare with others.



Backup

Semileptonic Selection(electron,muon)

Pre-selected by the total muon and electron number (BDT Isolated lepton Tagger)

	tth total	tth, truth semi-lep(e,mu)	ttz	ttg	6f
IsoMu+IsoE==1	37%	94%	34%	36.4%	31%
Isolep==1(old Finder)	-	88%	-	-	-

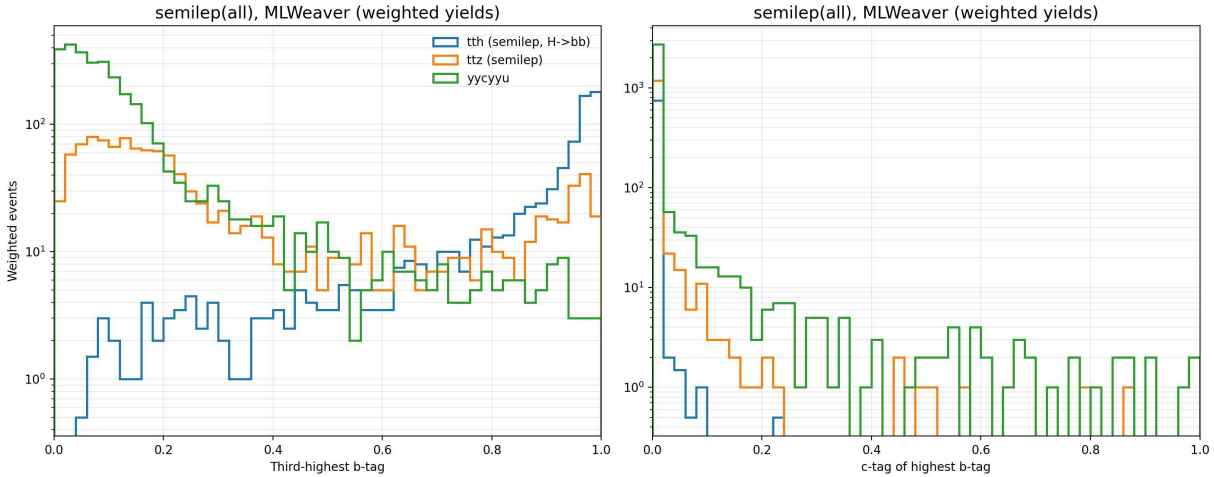
In the selected tth semileptonic events

	tth, truth semi-lep(e,mu)	tth, truth semi-lep(tau)	tth, truth hadronic	tth, truth dileptonic
BDT Tagger	77%	10%	1%	12%
Cut Finder	71%	11%	7%	11%

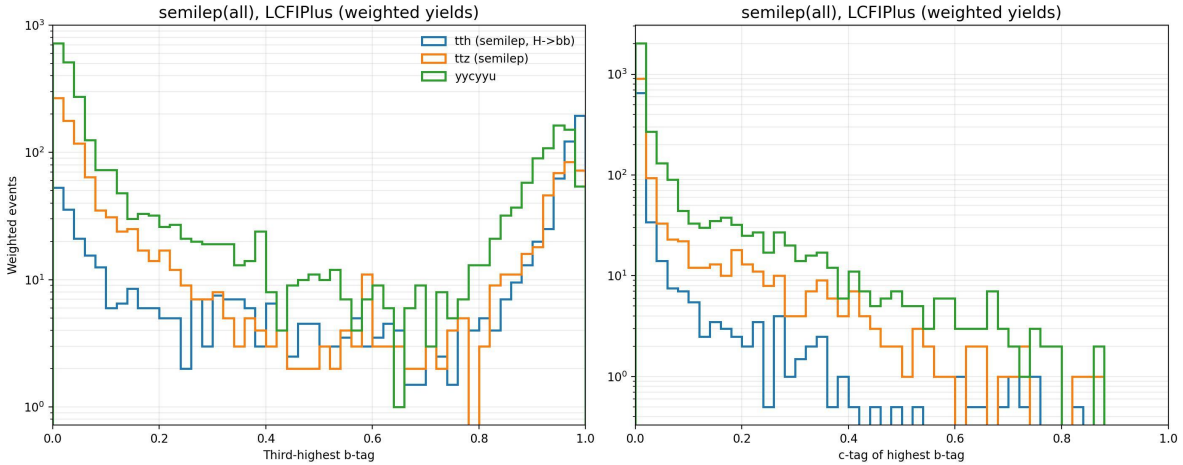
- The BDT tagger improved the efficiency by 6%
- BDT tagger has a higher ability on excluding the hardonic background, cannot perform better on excluding tau.

More Flavor Tagger performance： ParT vs LCFiPlus

ParT

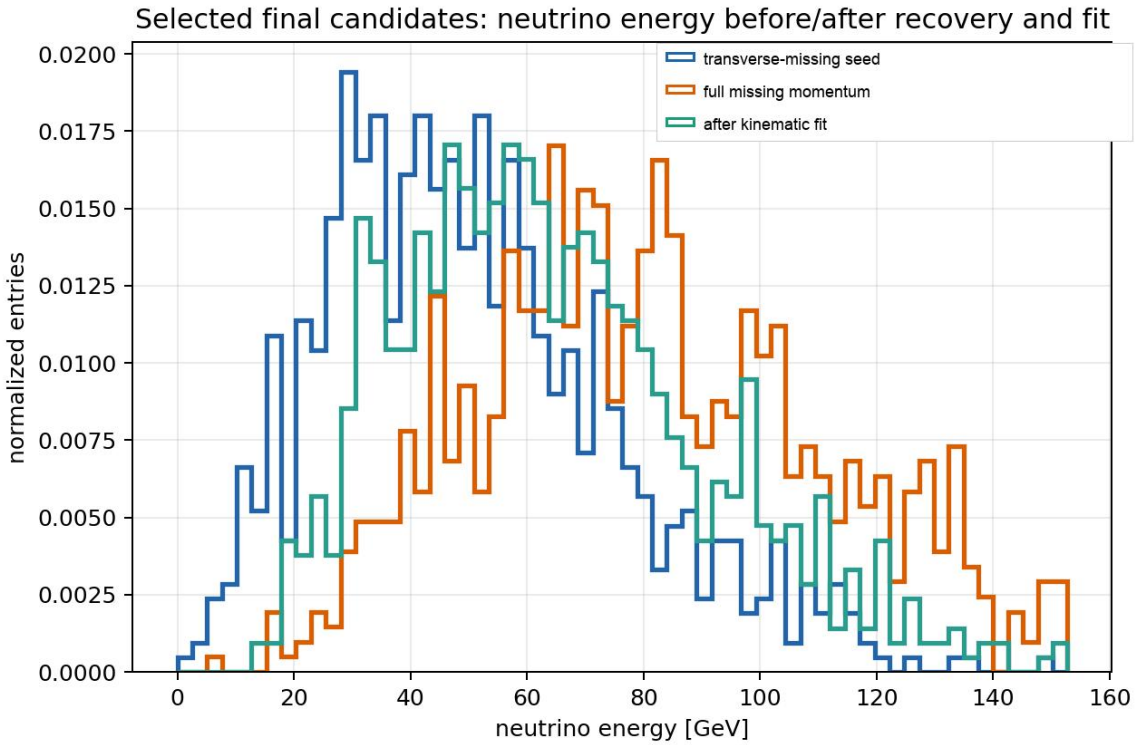


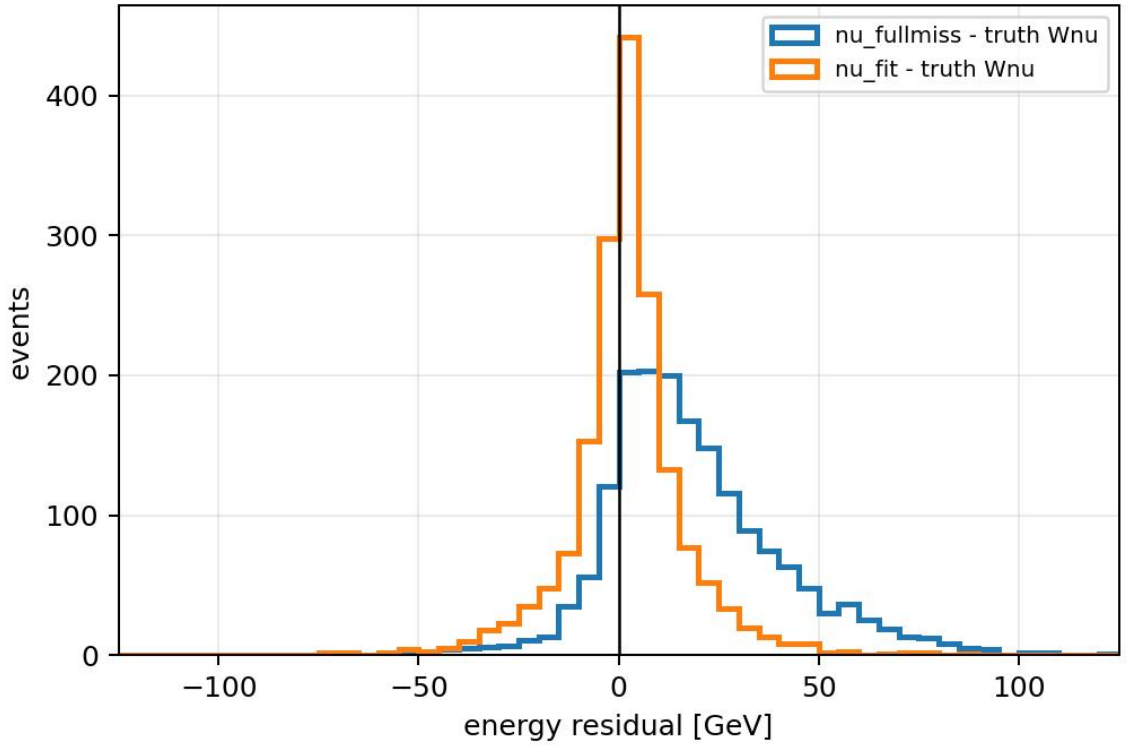
LCFiPlus



More Flavor Tagger performance and Kinfit detials

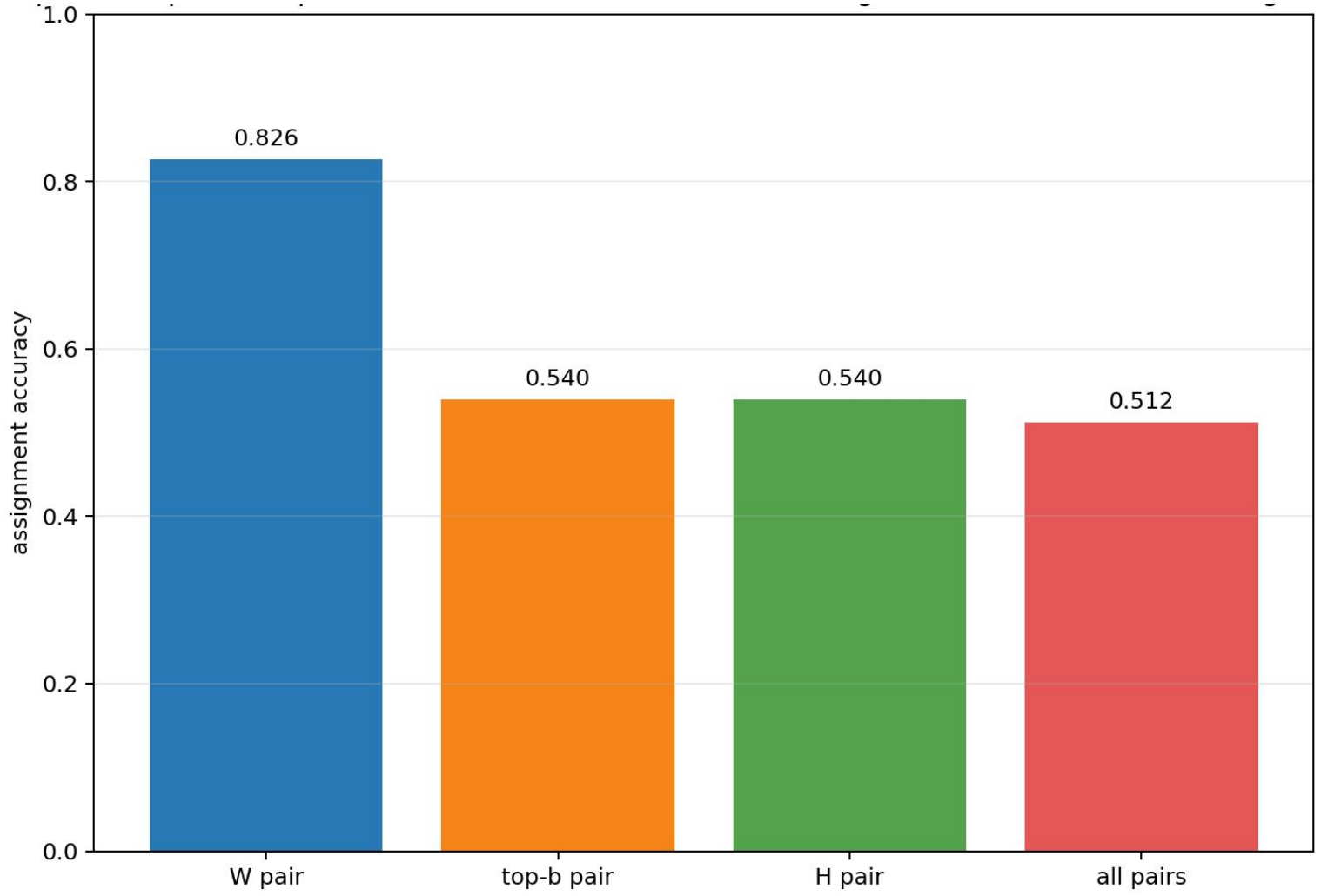
Kinematic Fitting: Neutrino Correction



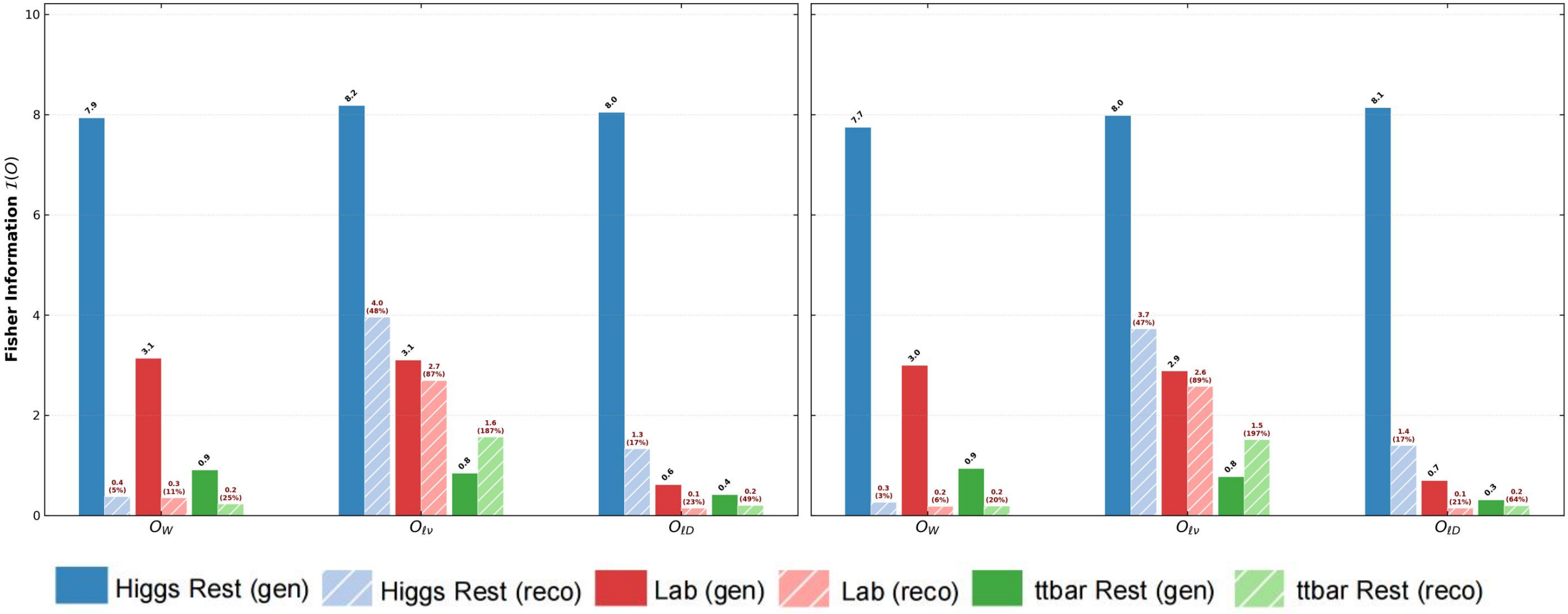


Benefit the observable O_{ln}

Physsim Jet Assignment Accuracy



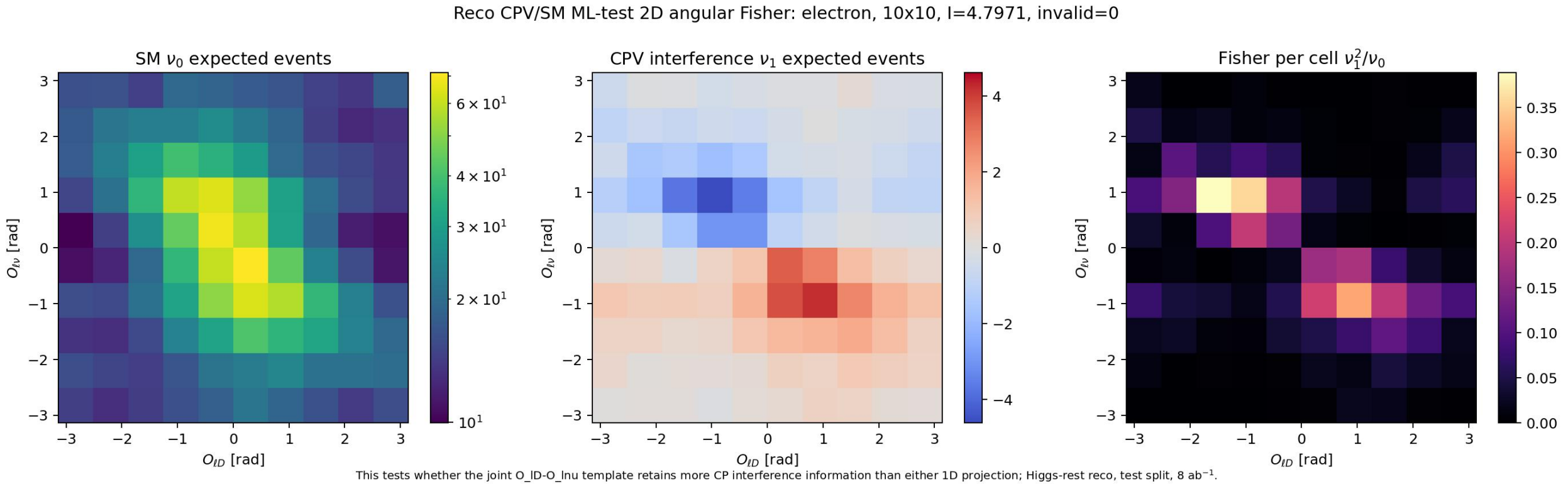
Angular observable baseline in different reference frame



$$\mathcal{M}_P(\lambda_f = s) \propto \begin{cases} \gamma_\psi \beta_\psi & \text{in } R_\psi\,, \\ \gamma_h \beta_h & \text{in } R_h\,, \end{cases} \qquad \frac{\gamma_h \beta_h}{\gamma_\psi \beta_\psi} = \frac{m_\psi}{m_h} \gtrsim 2.77\,,$$

$$m_\psi = m_{t\bar{t}} \approx 370 GeV$$
$$m_H \approx 125 GeV$$

Information fusion of the Angular observables



I_2D_electron=4.79, muon=4.41 I_2D=9.2

I_lD+I_lnu=10.4 > I_2D

CPV observable: Derivation of the interference-sign binary classifier

1 Signed-interference training sample

$$S_1(x) = 2\,\mathrm{Re}[\mathcal{M}_S^*(x)\mathcal{M}_P(x)]\,, \qquad y(x) = \mathrm{sgn}\,S_1(x), \quad d\mu_{\mathrm{train}} \propto |S_1(x)|\,dx$$

Truth label y = ± 1; effective training measure proportional to $|S_1|$.

2 Class-weight densities after reconstruction

$$q_\pm(z) = \int dx\, R(z|x)\,|S_1(x)|\,\Theta(\pm S_1(x))$$

R(z|x) includes smearing, combinatorics and selection.

3 Weighted cross-entropy minimization

$$\mathcal{L}_{\mathrm{BCE}} = -\int dz\,\left[q_+(z)\ln P_+(z) + q_-(z)\ln P_-(z)\right]$$

$$\frac{\partial \mathcal{L}_{\mathrm{BCE}}}{\partial P_+} = 0 \implies P_+(z) = \frac{q_+(z)}{q_+(z)+q_-(z)}, \quad P_-(z) = \frac{q_-(z)}{q_+(z)+q_-(z)}$$

4 Optimal classifier output under these training weights

$$O_{\mathrm{ML}}(z) = P_+(z) - P_-(z) = \tanh\left[\tfrac{1}{2}\ln\frac{q_+(z)}{q_-(z)}\right]$$

$$\boxed{O_{\mathrm{ML}}(z) = \frac{\int dx\, R(z|x)\,S_1(x)}{\int dx\, R(z|x)\,|S_1(x)|} \in [-1,1]}$$

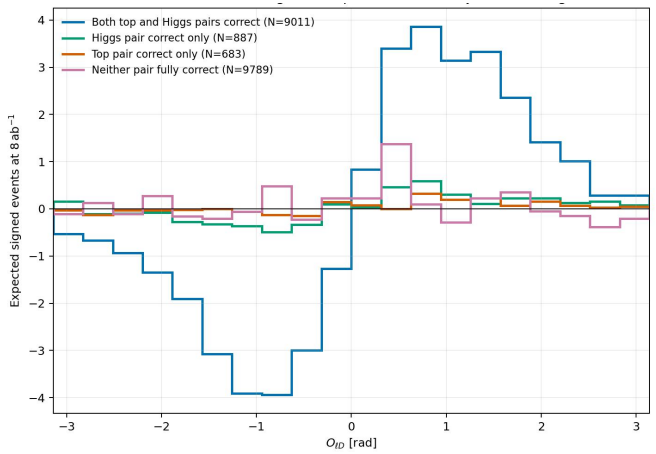
Similiar to the local likelihood score for SALLY:

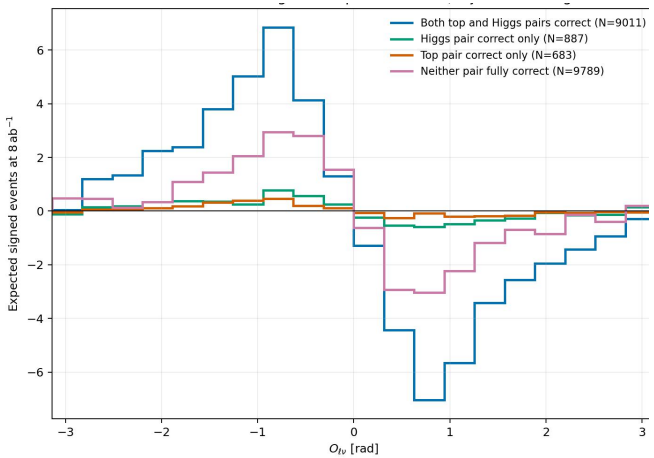
$$s_{\mathrm{reco}}(z) = \frac{\int dx\, R(z|x)\,S_1(x)}{\int dx\, R(z|x)\,S_0(x)}$$

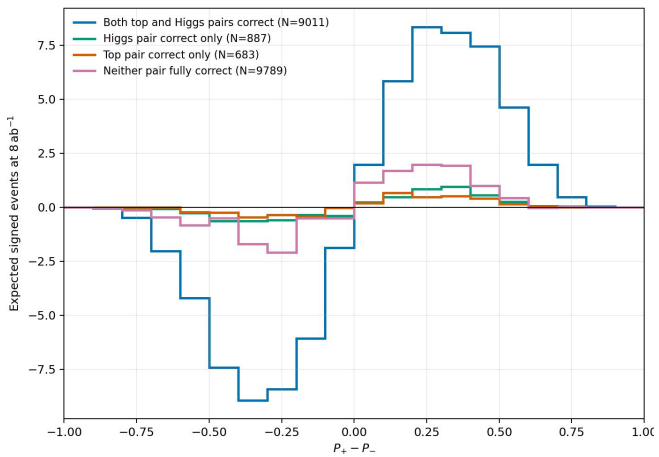
x: truth phase space | z: reconstructed inputs | S_1 : signed CP interference | S_0 : SM contribution | Θ: step function | $P_+ + P_- = 1$

Assumption: unaltered relative class weights (no independent class balancing); ideal calibrated binary classifier.

Comparison of truth-matched jet set





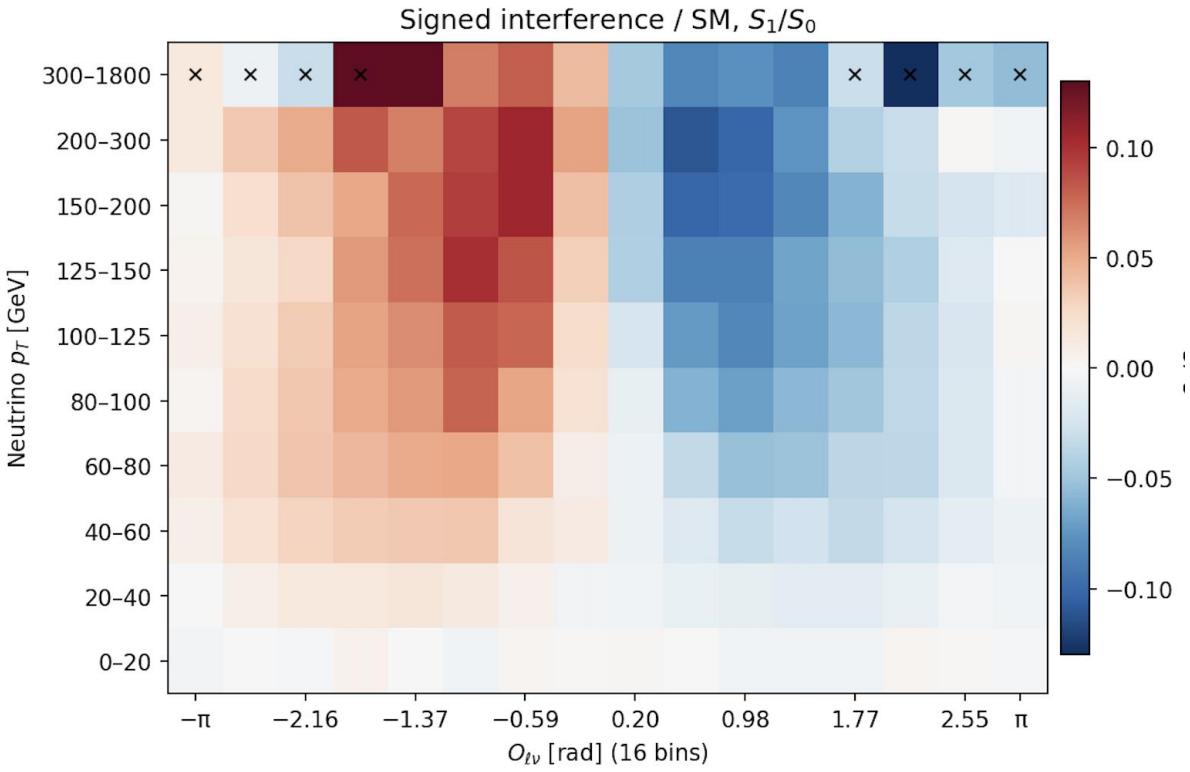


fisher	O_lD	O_lnu	ML_best
Blue	1.93	3.79	7.97
Pink	0.04	0.70	0.27

- The ML fisher value of Blue Line is higher than the value we get by the full event set/2 （half corretly assigned) , which actually shows the mis-matched jet events cancel with the correct events

ML observable: What is the model actually learned with pT

Reco



Gen

