From Machine-Enhanced to Symbolic CP Asymmetries

Akanksha Bhardwaj

IIT PALAKKAD

ML4HEP-V4 School

PART 2

Symbolic Regression in Particle Physics

Applications:

- discovering CP-sensitive observables
- jet substructure observables
- EFT parameter estimation
- anomaly detection
- interpretable ML observables

Standard Model Effective Field Theory

- Lack of experimental evidence of new physics indicate a mass gap between SM and BSM scales
- SMEFT Lagrangian

$$\mathcal{L}_{\mathrm{SMEFT}} = \mathcal{L}_{\mathrm{SM}} + \sum_{i>4} \sum_{k} \frac{C_k^{(i)}}{\Lambda^{i-4}} \mathcal{O}_k^{(i)}$$

- Any differential cross section follows:

$$d\sigma = d\sigma_{\mathrm{SM}} + \frac{C_i}{\Lambda^2} d\sigma_i + \frac{C_i C_j}{\Lambda^4} d\sigma_{ij}$$

Interference

Cross terms

- New physics effect can be seen as deformation of these distribution

Observable effects at LHC

$$|\mathcal{M}|^2 = |\mathcal{M}_{\rm SM}|^2 + \frac{c_i}{\Lambda^2} 2\Re\left[\mathcal{M}_{\rm SM}\mathcal{M}^*_{{\rm d6},i}\right] + \frac{c_i\, c_j}{\Lambda^4}\mathcal{M}_{{\rm d6},i}\mathcal{M}^*_{{\rm d6},j}\,,$$

Leading order deviation from SM comes from interference terms

CP Sensitive Observables – angular observables, matrix-element method

Exploits full-kinematic information but very time/resource consuming

Gauge-Higgs sector in SMEFT

- $\tilde{O}_i$ Operators Introduce new source of CP violation
- $$\mathscr{L} = \mathscr{L}_{\text{SM}} + \sum_i \frac{c_i}{\Lambda^2} \tilde{\mathcal{O}}_i \,,$$

Dim-6 Lagrangian

- Operators that affect the electroweak interactions of the Higgs boson

$$\mathcal{O}_{\Phi\widetilde{B}} = \frac{c_{\Phi\widetilde{B}}}{\Lambda^2} \Phi^\dagger \Phi B^{\mu\nu} \widetilde{B}_{\mu\nu} \,,$$

$$\mathcal{O}_{\Phi\widetilde{W}} = \frac{c_{\Phi\widetilde{W}}}{\Lambda^2} \Phi^\dagger \Phi W^{i\,\mu\nu} \widetilde{W}^i_{\mu\nu} \,,$$

Higgs-Gauge EW boson interactions

$$\mathcal{O}_{\Phi\widetilde{W}B} = \frac{c_{\Phi\widetilde{W}B}}{\Lambda^2} \Phi^\dagger \sigma^i \Phi \widetilde{W}^{i\,\mu\nu} B_{\mu\nu} \,,$$

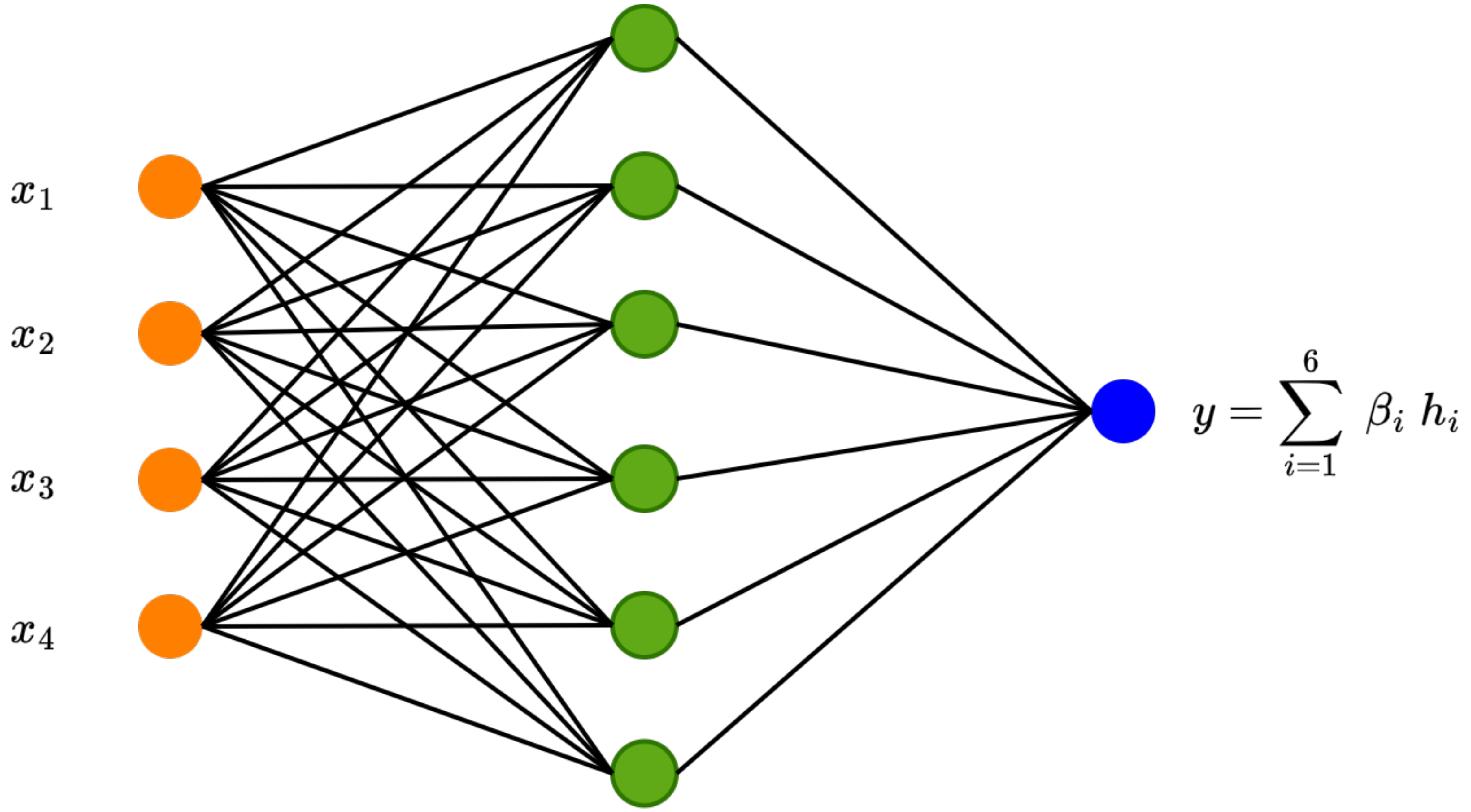
ML constructed CP-odd observable

- With the ability to learn kinematic correlations, the NN can be used to
  - Construct a near optimal CP-odd observable for each dimension-six operator
  - Design new analyses based on the correlation between the angular observables and other kinematic quantities.
- Extend to multi-class models, with the pure-SM prediction included
  - Allow the NN to learn the phase-space regions for which the SM is relatively suppressed

[Andrei V. Gritsan et al ,Phys. Rev. D 102, 056022 (2020)]

Artificial Neural Networks

Deep Learning=>Automatic feature extraction, domain-experts design data representation, architecture, training methods etc.



$$h_j = \sigma\left(\sum_{i=1}^{4} w_{ji}^{(0)} x_i + w_{j0}^{(0)}\right)$$

The core of most ML algorithms is to learn a parametrized function $f(x, \Theta)$ and tune the parameters Θ for a particular objective

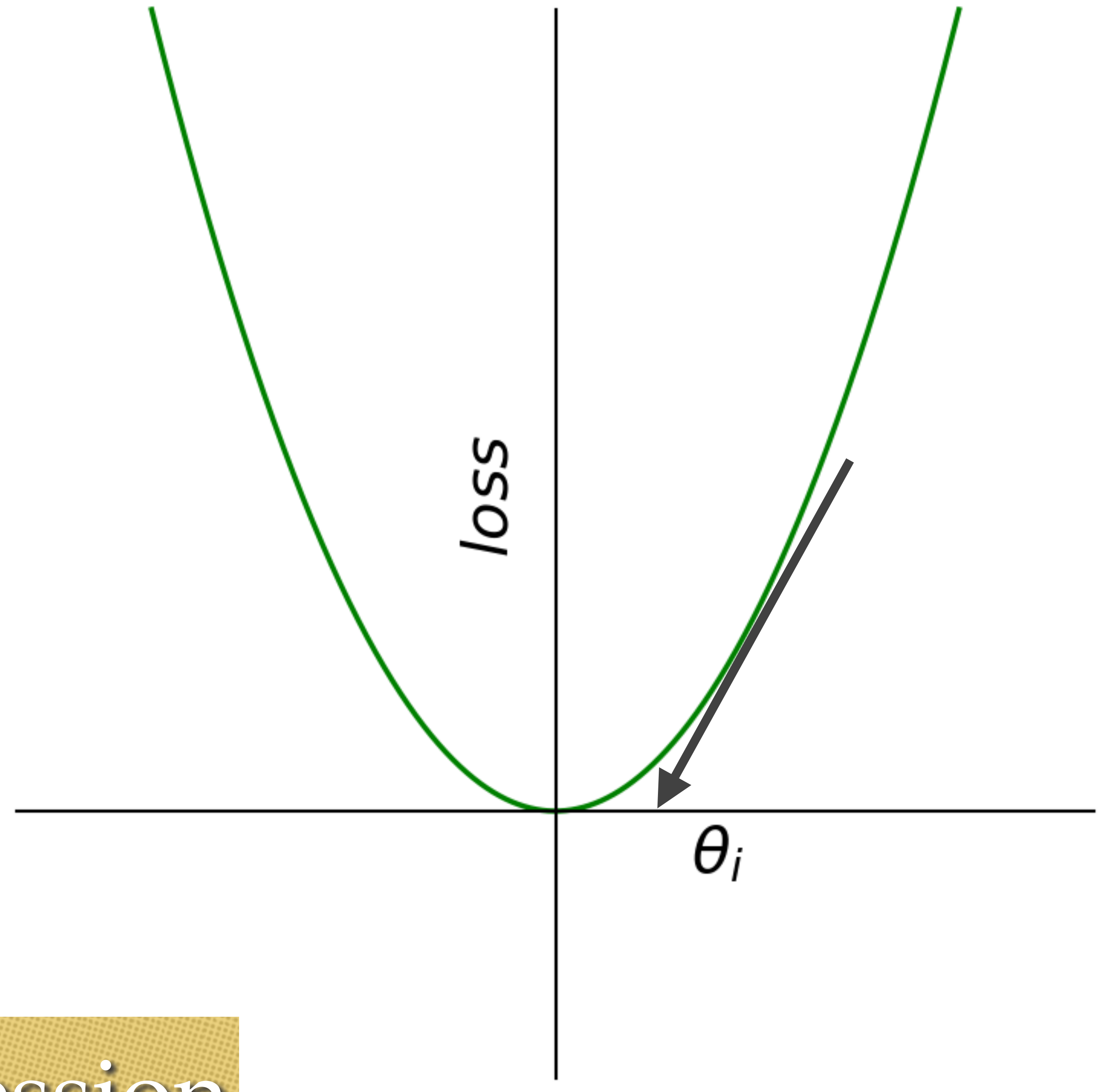
Machine Learning

Broadly, estimate a function given data samples.

$$\{(x_1, y_1), (x_2, y_2), \ldots, (x_n, y_n)\}$$

$$\hat{y} = f(\Theta, x)$$

Optimize a Loss function $\mathscr{L}(\hat{y}, y)$



Classification:

$$\text{loss} = -\frac{1}{N_{batch}} \sum_{i=1}^{N_{batch}} y_i \ln\left(\hat{y}(x_i)\right)$$

Linear Regression

$$\hat{y} = \theta_1 x + \theta_0$$

$$\text{loss} = \frac{1}{N_{batch}} \sum_{i=1}^{N_{batch}} (\hat{y}_i - y_i)^2$$

ML constructed CP-odd observable

Binary (two-class) models

Trained to distinguish + and - interference effects

$$P_+ + P_- = 1$$

Andrei V. Gritsan,

Multi-class models

Trained to distinguish SM, +, and - interference

$$P_+ + P_- + P_{\rm SM} = 1$$

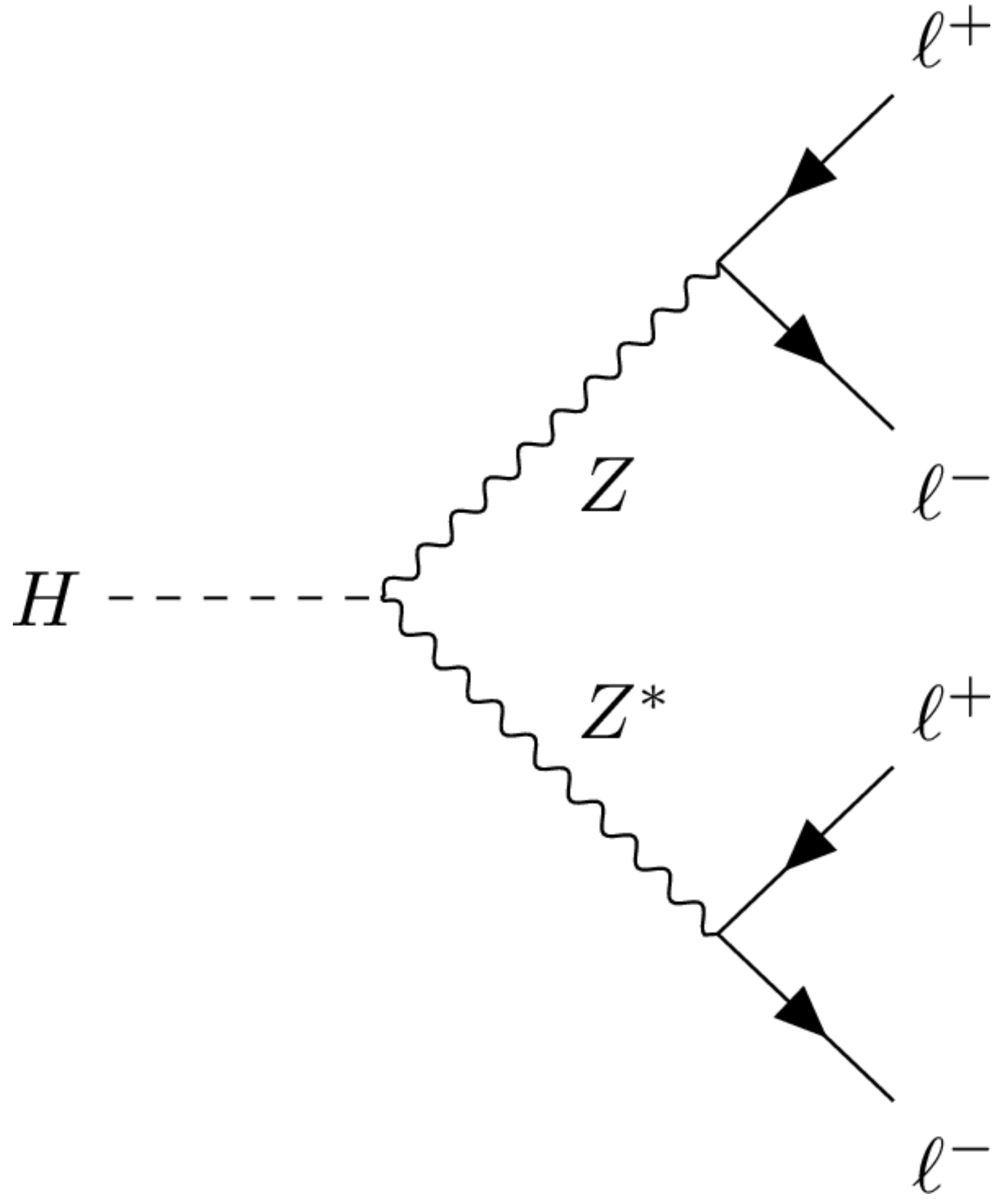
- CP observable from NN output (of either model)

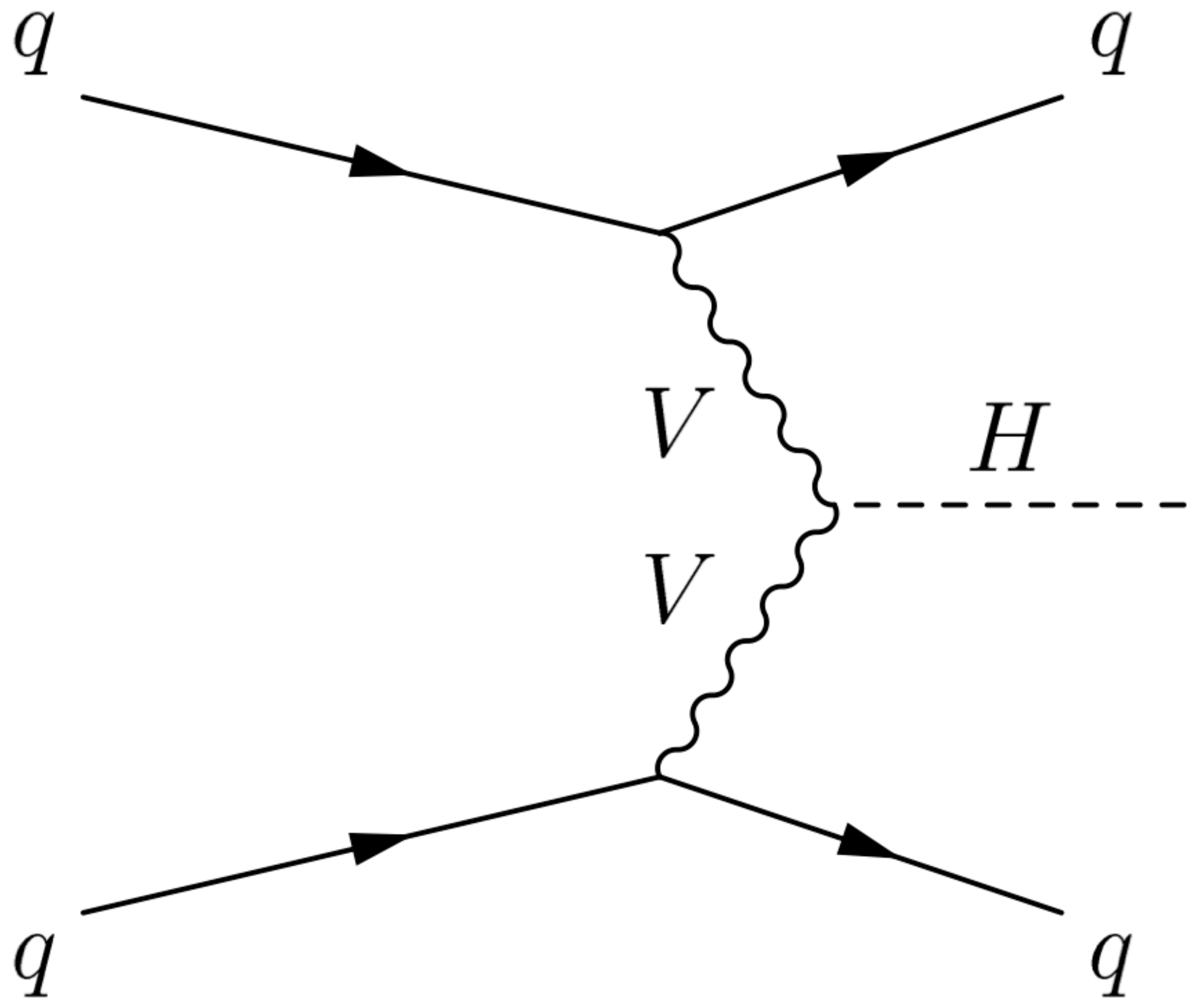
$$O_{NN} = P_+ - P_-$$

[AB, Christoph Englert, Robert Hankache, Andrew D. Pilkington,. PLB` 23]

Application of NN constructed CP-odd observables

- Two of the main search channels for CP-violation in the Higgs sector: the decay channel and in the vector-boson fusion production channel (VBF h + 2 jets).

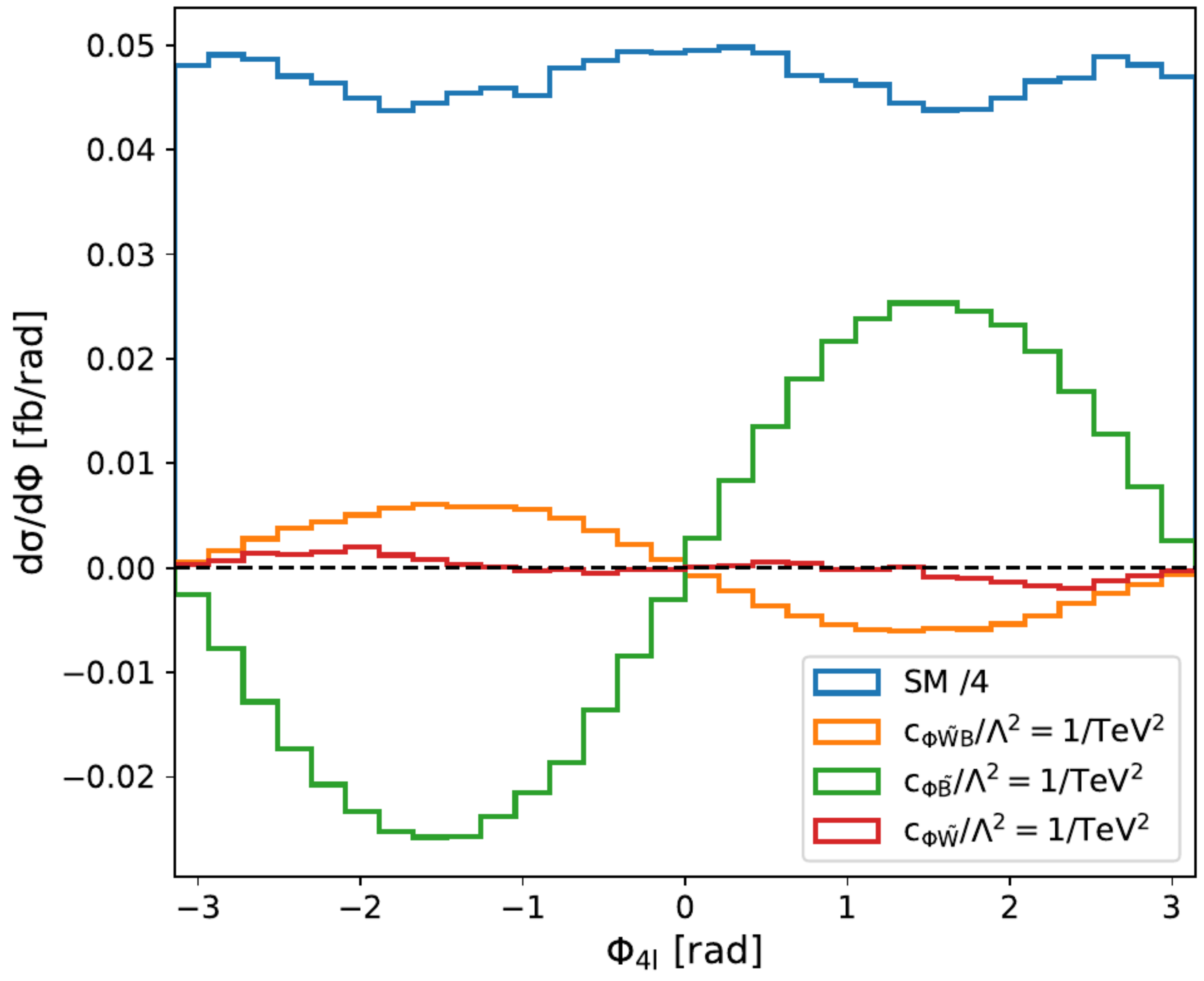


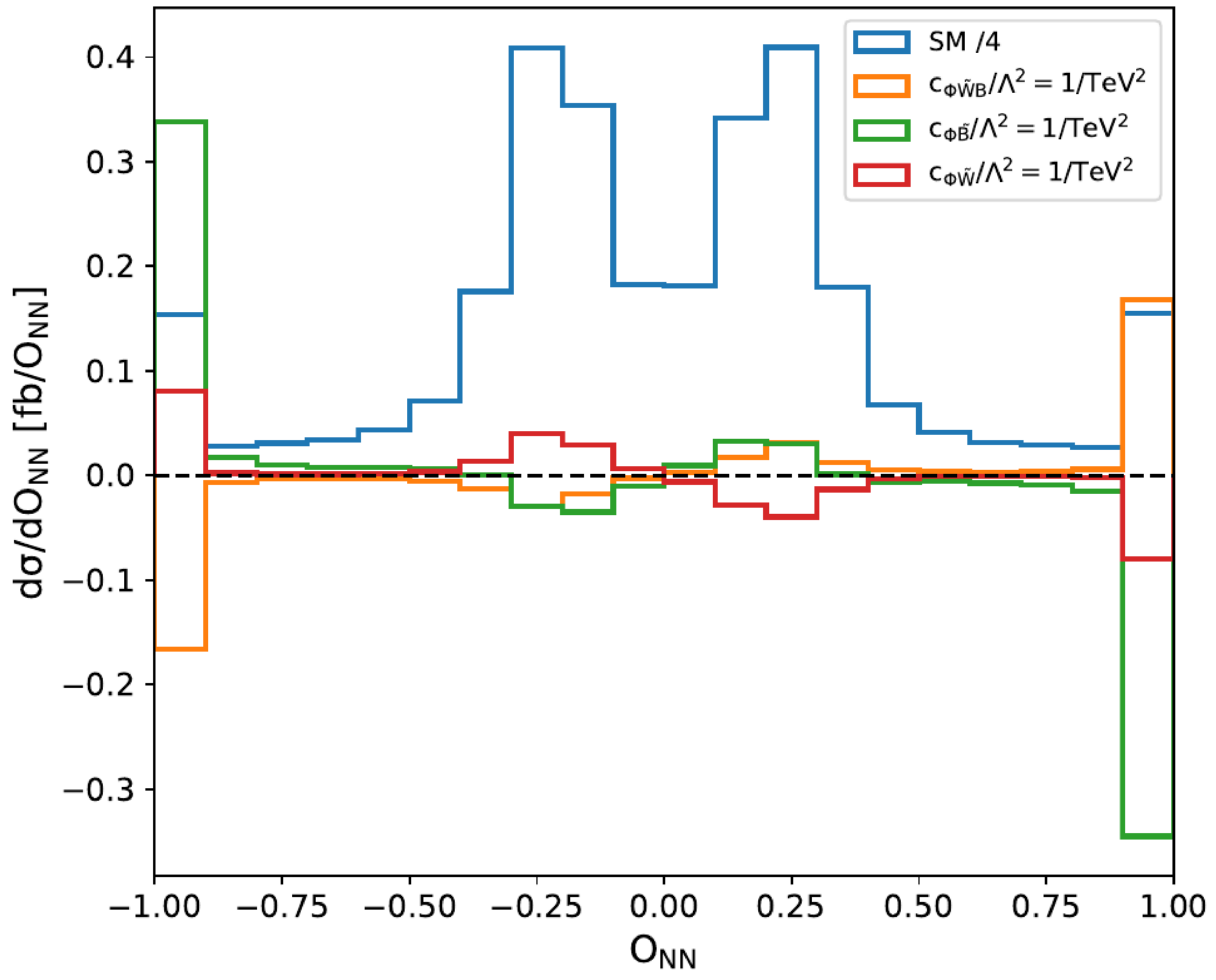


Application to $h \to ZZ^* \to 4\ell$

Differential cross section as fn of

Binary class NN output

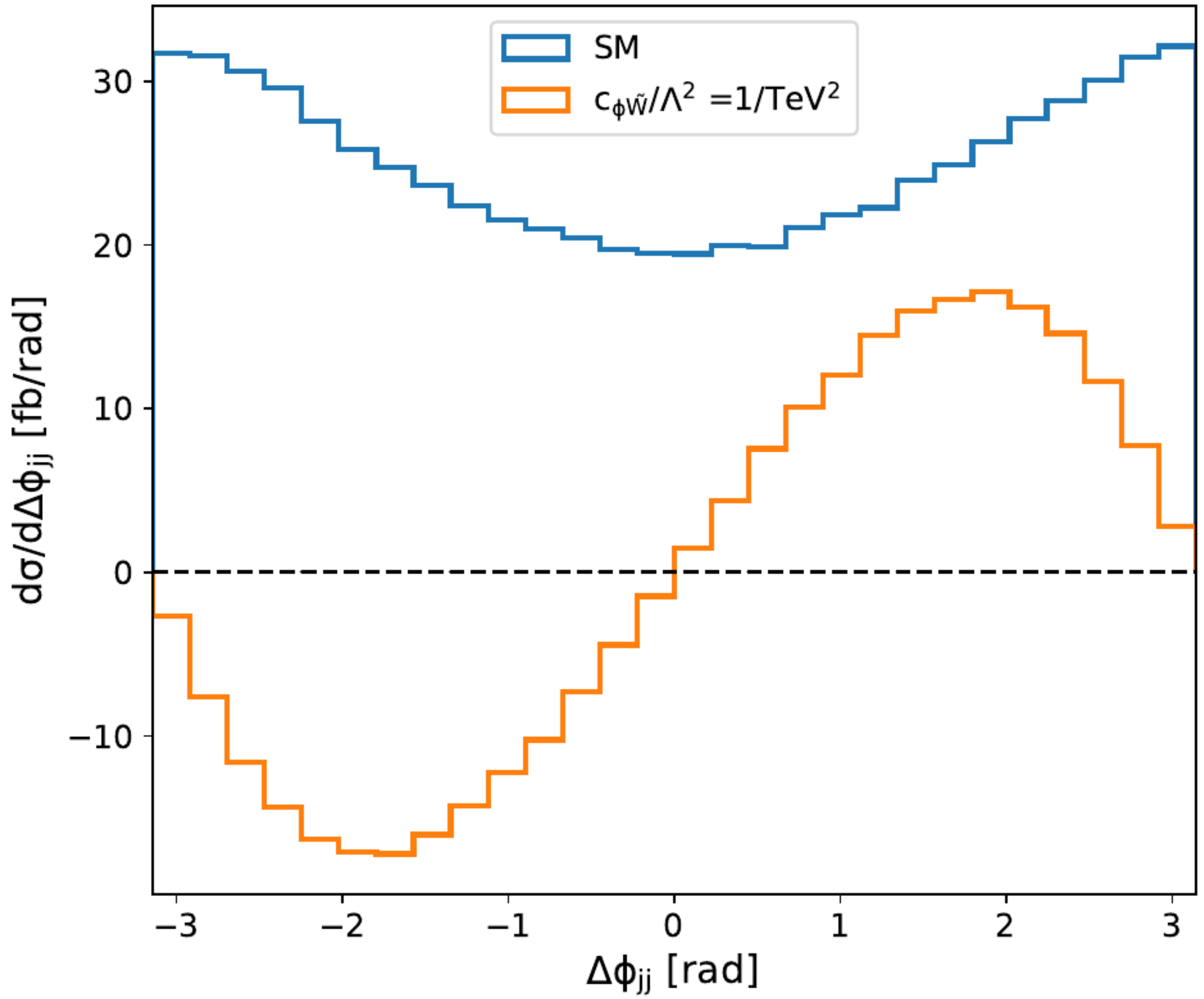


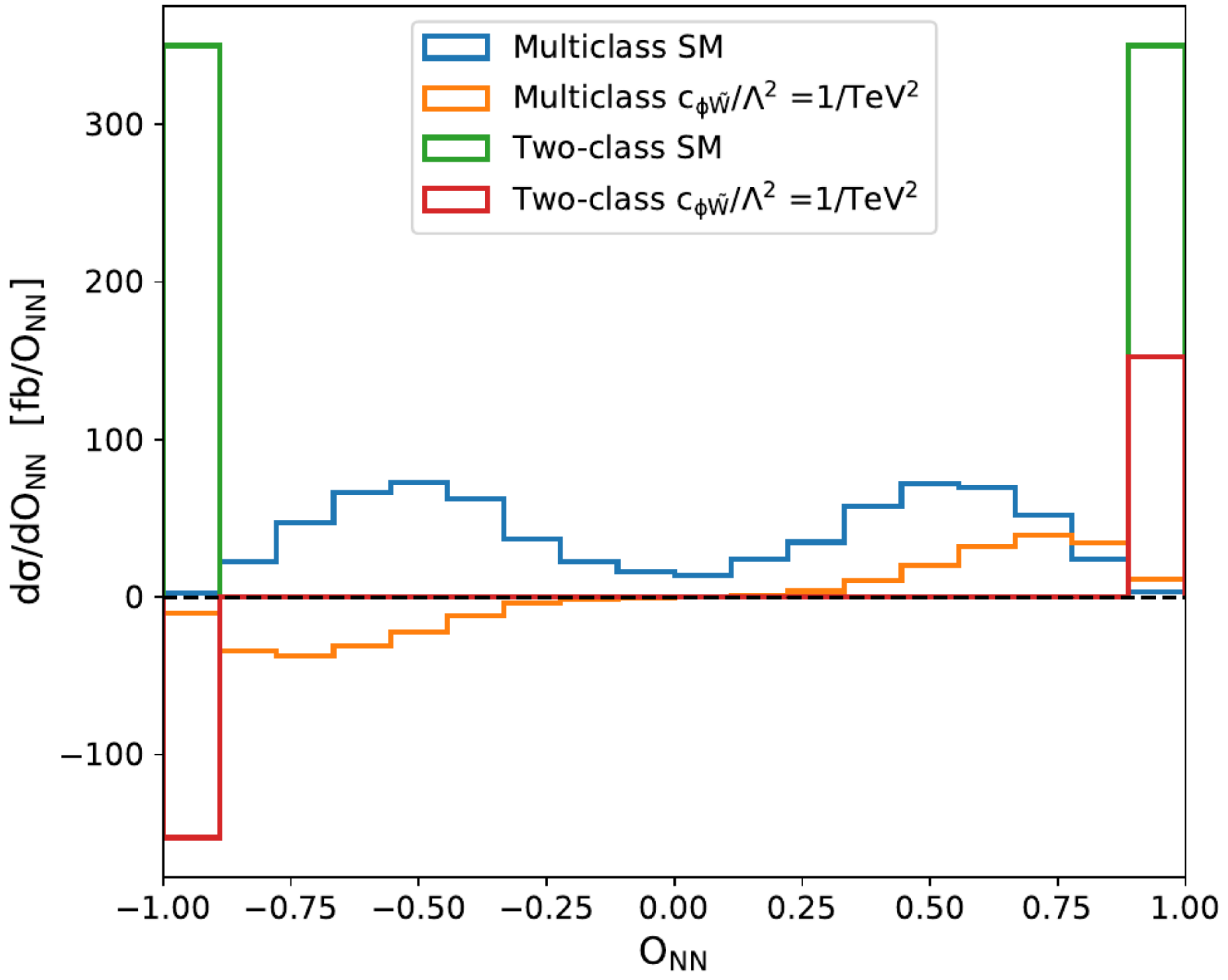


Analysis carried out in the *Higgs Mass* region of the ATLAS inclusive *4l* measurement (JHEP 07, 005 (2021) for H→2e2μ events.

[AB, Christoph Englert, Robert Hankache, Andrew D. Pilkington,. PLB` 23]

Application to $h + 2$ jets





- Analysis carried out in the *VBF_1* region of the ATLAS H→ττ analysis (ATLAS-CONF-2021-044)
- Classic CP-odd variable: $\Delta\phi_{jj} = \phi(j_1) - \phi(j_2)$

[AB, Christoph Englert, Robert Hankache, Andrew D. Pilkington,. PLB`23]

Limits on CP-odd operators for $h \to ZZ^* \to 2e2\mu$

CP-odd observable	$c_{\Phi\widetilde{W}B}/\Lambda^2$ [TeV^{-2}]	$c_{\Phi\widetilde{B}}/\Lambda^2$ [TeV^{-2}]	$c_{\Phi\widetilde{W}}/\Lambda^2$ [TeV^{-2}]
$\Phi_{4\ell}$	[-6.2,6.2]	[-1.4,1.4]	[-30,30]
$\Phi_{4\ell}$, m_{12}	[-1.9,1.9]	[-0.85,0.85]	[-3.7,3.7]
O_{NN} (binary)	[-1.5,1.5]	[-0.75,0.75]	[-3.0,3.0]
O_{NN} (multi-class)	[-1.4,1.4]	[-0.71,0.71]	[-2.7,2.7]

[AB, Christoph Englert, Robert Hankache, Andrew D. Pilkington,. PLB` 23]

Sensitivity to specific operators established using the Profile Likelihood method, after normalising the MC samples to the number of events observed in the ATLAS analyses.

Main observations:

- NN-based observables offer the best sensitivity.
- Multiclass models offers 5-10% improvements over binary classification
- Double-differential analysis of Φ_{4l} and m_{Z1} captures most of the sensitivity gained by NN

Symbolic Regression

Input Space $x \in \mathbb{R}^d$

Training dataset $\{(x_i, y_i)\}_{i=1}^N$

NN selects and combines mathematical primitives $f_\theta(x)$ is built from $\{\sin, \tanh, \times, /, +\}$

Linear mixing

$$z^{(\ell)} = W^{(\ell)} h^{(\ell-1)} + b^{(\ell)}, \qquad h^{(0)} = x$$

Primitive function library

$$g(z) \in \{\sin z,\ \tanh z,\ z_i z_k,\ z_i / z_k\}$$

Symbolic feature vector

$$\phi^{(\ell)} = [g_1(z^{(\ell)}),\ g_2(z^{(\ell)}),\ \dots]$$

Information Flow (Forward Pass)

Sparse symbolic combination

$$h^{(\ell)} = A^{(\ell)} \phi^{(\ell)}$$

Deep symbolic network (compact form)

$$h^{(\ell)} = A^{(\ell)} g\left(W^{(\ell)} h^{(\ell-1)} + b^{(\ell)}\right)$$

Output score

$$f_\theta(x) = w_{\text{out}}^\top h^{(L)} + b_{\text{out}}$$

Training & Interpretability

Classification probability

$$\hat{p}(x) = \sigma\left(f_\theta(x)\right)$$

Loss with sparsity

$$\mathcal{L} = -\frac{1}{N} \sum_{i=1}^{N} y_i \log \hat{p}(x_i) + \lambda \sum_{\ell} \|A^{(\ell)}\|_1$$

CP-Analyses with Symbolic Regression

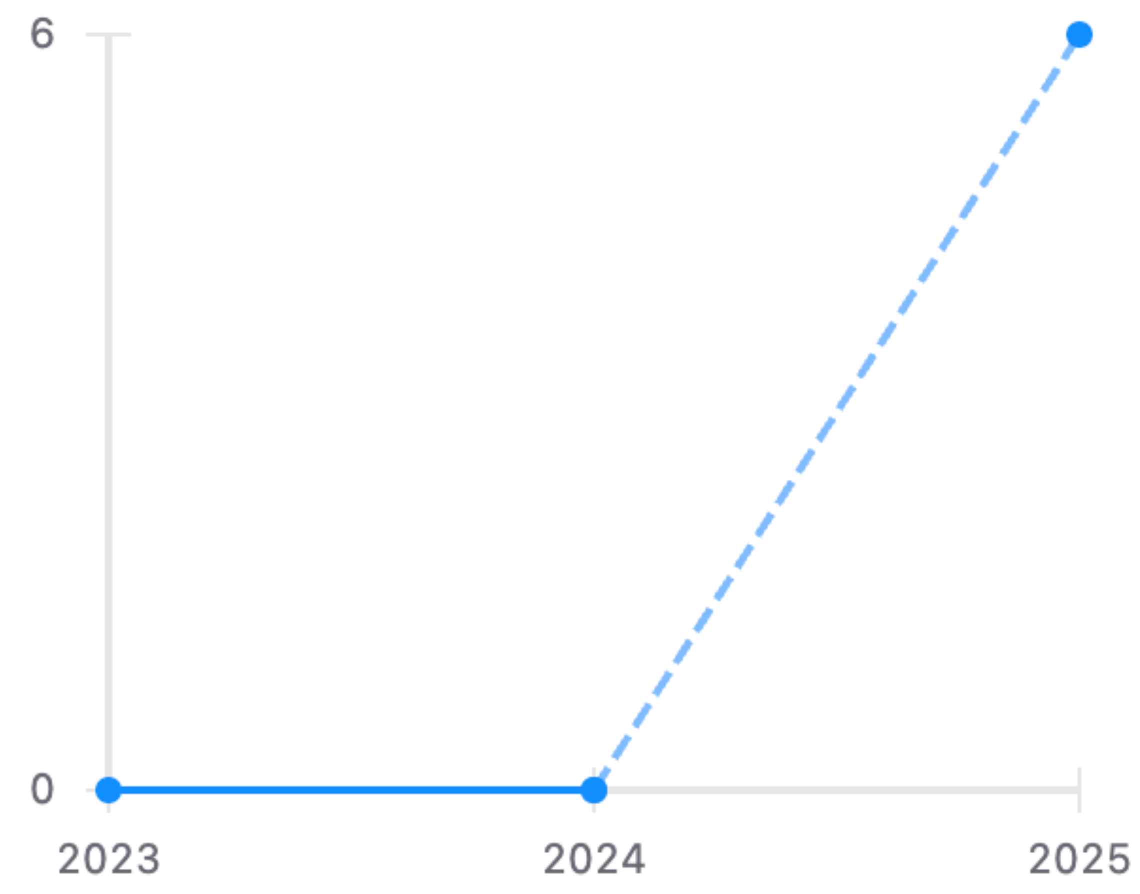
$\mathcal{CP}$-Analyses with Symbolic Regression

Henning Bahl (U. Heidelberg, ITP), Elina Fuchs (Leibniz U., Hannover and Braunschweig, Phys. Tech. Bund. and DESY), Marco Menen (Leibniz U., Hannover and Braunschweig, Phys. Tech. Bund.), Tilman Plehn (U. Heidelberg, ITP and U. Heidelberg (main))

Jul 8, 2025

37 pages
e-Print: 2507.05858 [hep-ph]
View in: ADS Abstract Service

Citations per year



Abstract: (arXiv)
Searching for $\mathcal{CP}$ violation in Higgs interactions at the LHC is as challenging as it is important. Although modern machine learning outperforms traditional methods, its results are difficult to control and interpret, which is especially important if an unambiguous probe of a fundamental symmetry is required. We propose solving this problem by learning analytic formulas with symbolic regression. Using the complementary PySR and SymbolNet approaches, we learn $\mathcal{CP}$-sensitive observables at the detector level for WBF Higgs production and top-associated Higgs production. We find that they offer advantages in interpretability and performance.

Note: 37 pages, 15 figures, 5 tables

CP-Analyses with Symbolic Regression

$$d^{\texttt{SymbolNet}} = 0.715\Delta\phi_{jj}\Big[-0.348(0.561x_{j_2} - 0.25\Delta\eta_{jj} + 0.0315x_{jj} + 0.746x_h) - 0.27\Big]$$
$$\cdot\Big[0.0493(0.603\Delta\eta_{jj} - 0.0811x_{jj} - x_h)^2$$
$$- 0.654|0.463x_{j_2} + 0.477\Delta\eta_{jj} + 0.373x_h|^{0.5}$$
$$- 0.134\sin(-0.555\Delta\eta_{jj} + 0.345x_{jj} + 0.443x_h)$$
$$- 1.82\cos(0.642\Delta\phi_{jj})\Big], \quad (32)$$

Now the ML output looks more interpretable to a physicists !!

Opens Many possible direction to study!!