Machine-enhanced CP-asymmetries in the Higgs sector

A. Bhardwaj, C. Englert, R. Hankache, A. Pilkington

LHC Higgs Working Group WG2 and WG3 Extended Higgs Sector joint meeting on CPV in Higgs sector, 11 Jan 2023

Relevant Operators in SMEFT

Dim-6 Lagrangian:

$$\mathcal{L} = \mathcal{L}_{\rm SM} + \sum_i \frac{c_i}{\Lambda^2} \widetilde{\mathcal{O}}_i \, ,$$

Higgs-Gauge EW boson interactions:

$$\mathcal{O}_{\Phi\widetilde{B}} = \Phi^\dagger \Phi B^{\mu\nu} \widetilde{B}_{\mu\nu} \, , \qquad \mathcal{O}_{\Phi\widetilde{W}} = \Phi^\dagger \Phi W^{i\,\mu\nu} \widetilde{W}^i_{\mu\nu} \, ,$$

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

Observable effects at LHC

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

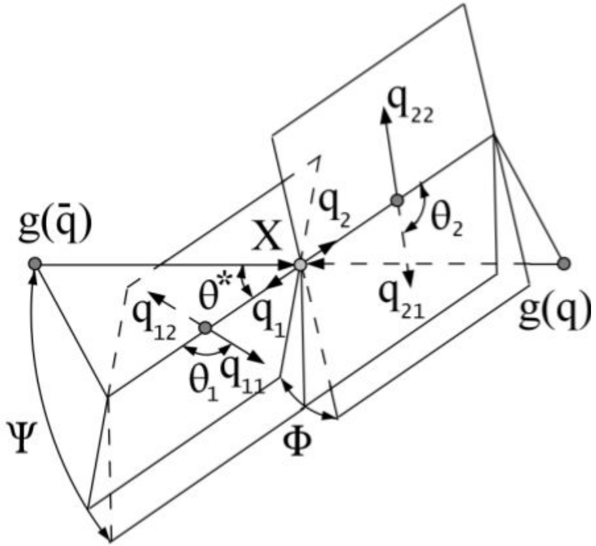
Leading order deviation from SM comes from interference terms

CP Sensitive Observables – angular observables, matrix-element method

Exploits full-kinematic information but very time/resource consuming

CP sensitive observables

$$\Phi_{4\ell} = \frac{\mathbf{q}_1 \cdot (\hat{\mathbf{n}}_1 \times \hat{\mathbf{n}}_2)}{|\mathbf{q}_1 \cdot (\hat{\mathbf{n}}_1 \times \hat{\mathbf{n}}_2)|} \times \cos^{-1}(\hat{\mathbf{n}}_1 \cdot \hat{\mathbf{n}}_2),$$



Normal vectors to the planes

$$\hat{\mathbf{n}}_1 = \frac{\mathbf{q}_{11} \times \mathbf{q}_{12}}{|\mathbf{q}_{11} \times \mathbf{q}_{12}|} \quad \text{and} \quad \hat{\mathbf{n}}_2 = \frac{\mathbf{q}_{21} \times \mathbf{q}_{22}}{|\mathbf{q}_{21} \times \mathbf{q}_{22}|}.$$

$$\mathbf{q}_\alpha = \mathbf{q}_{\alpha 1} + \mathbf{q}_{\alpha 2}$$

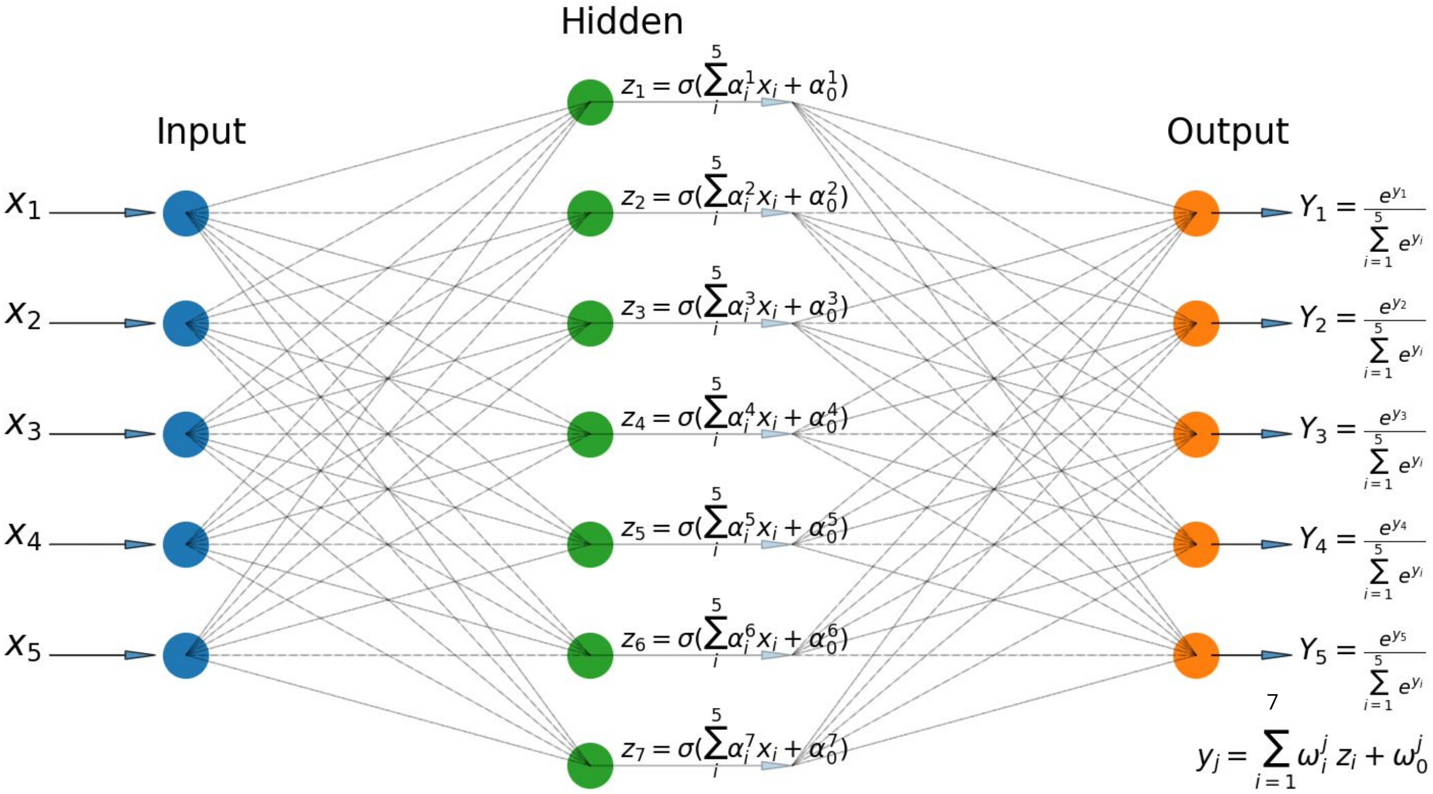
Classic Angular Variable PRD 86, 095031 (2012)

CP-odd observables from machine learning algorithms

CP-asymmetries arise from the interference between SM and CP-odd amplitudes:

- Neural networks (NN) offer an easy way to understand these asymmetries.
  - generate interference-only contribution to process (Madgraph5, SMEFTSim)
  - split sample into positive-weights and negative-weights.
  - train NN to distinguish between the two samples (binary classification)
  - easy to include Standard-Model contribution in NN (multiclass)
- Options with trained network:
  - construct observable from NN classifications, i.e
  - improve differential cross section measurements.

Artificial Neural Networks



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)

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

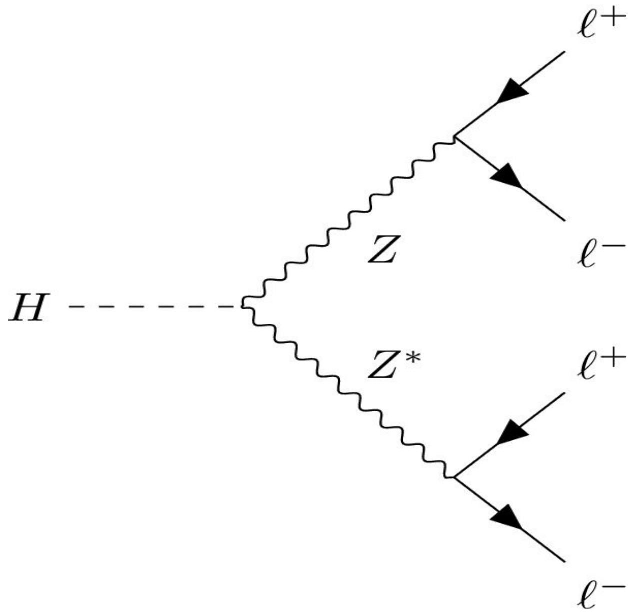
$$P_+ + P_- + P_{\mathrm{SM}} = 1$$

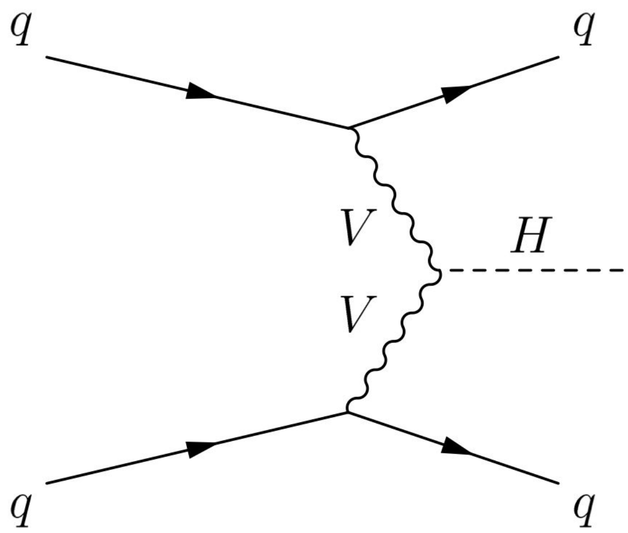
CP observable from NN output (of either model)

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

Application of NN constructed CP-odd observables

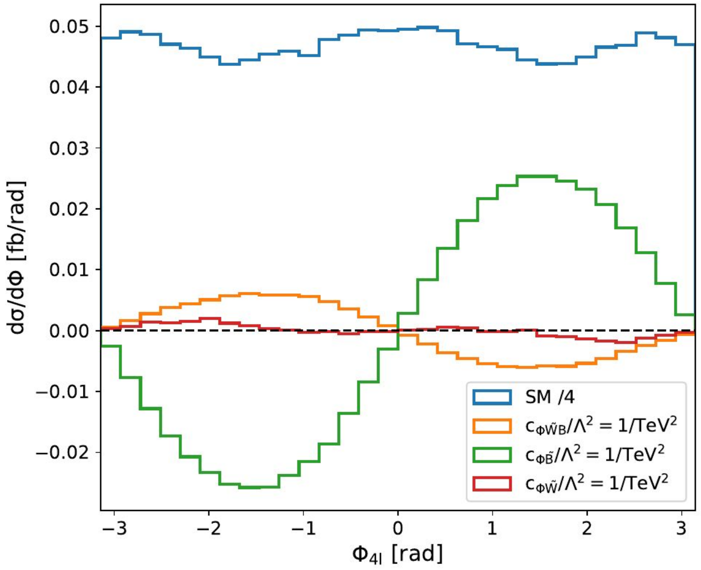
Two of the main search channels for CP-violation in the Higgs sector: the decay channel and in the vector-boson fusion production channel (VBF h + 2 jets).



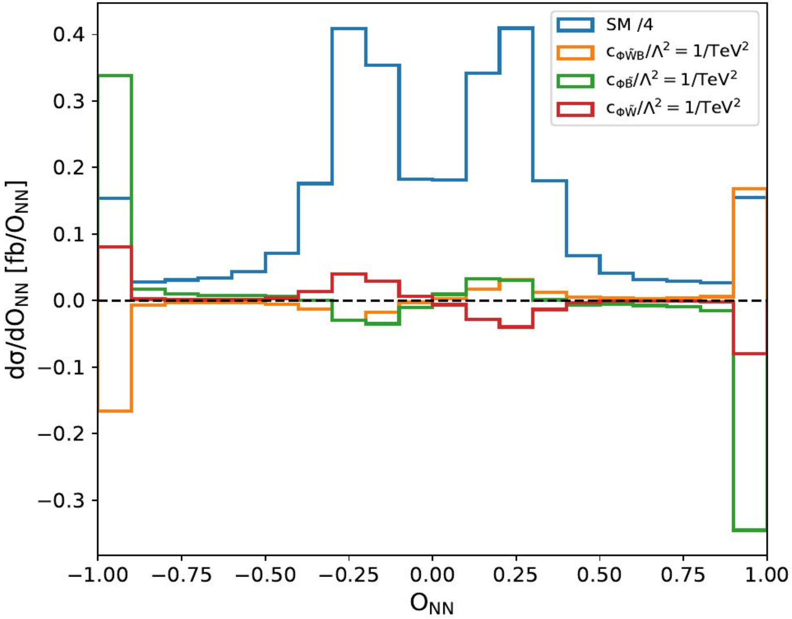


Application to $h \to ZZ^* \to 4\ell$

Differential cross section as fn of



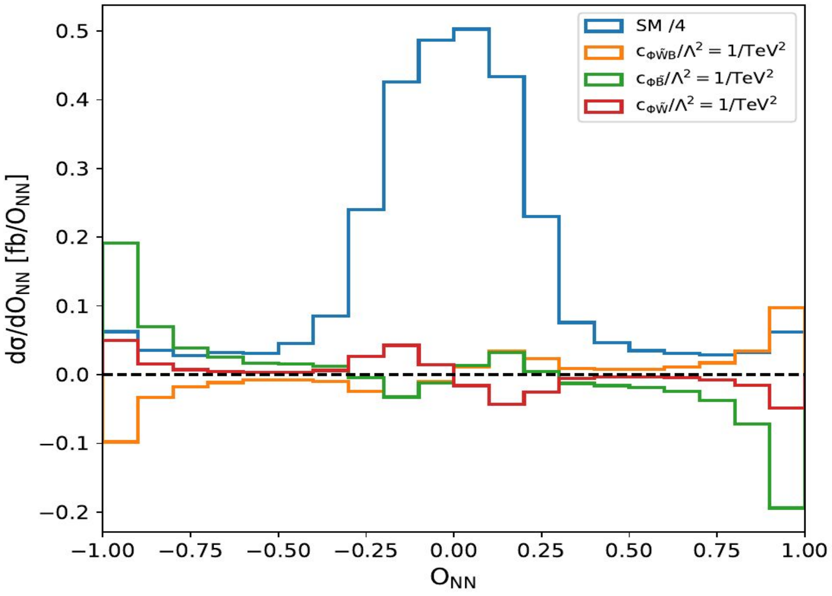
Binary class NN output



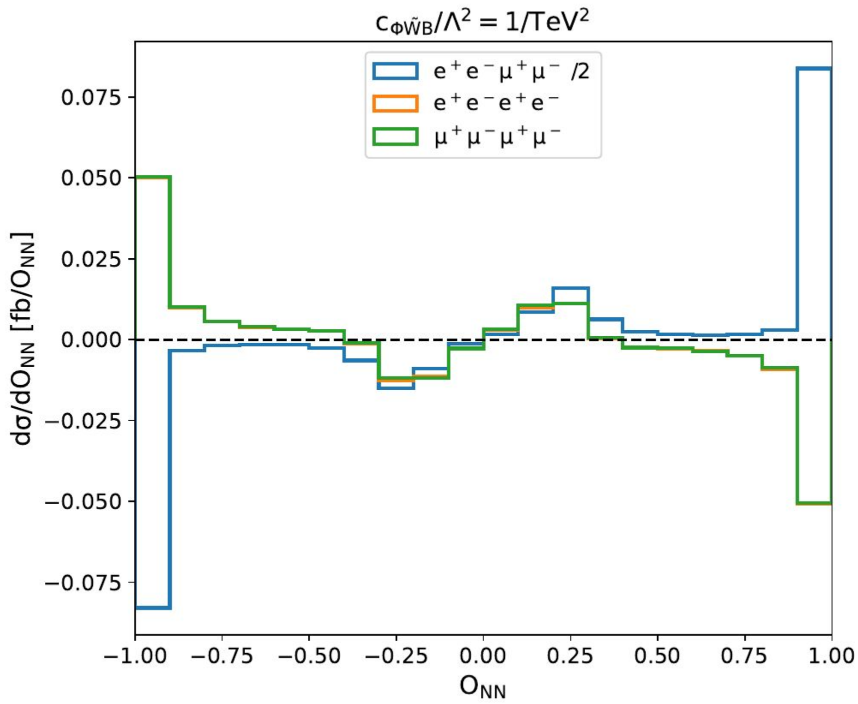
Analysis carried out in the *Higgs Mass* region of the ATLAS inclusive *4l* measurement (JHEP 07, 005 (2021) for H→2e2μ events.

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

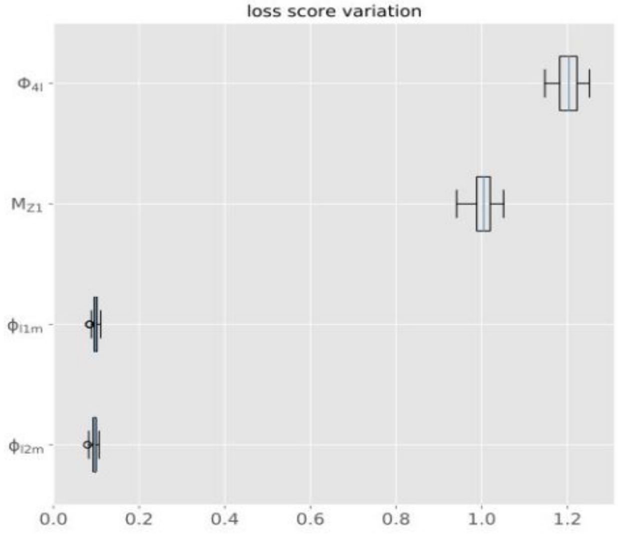
Multi-class NN output

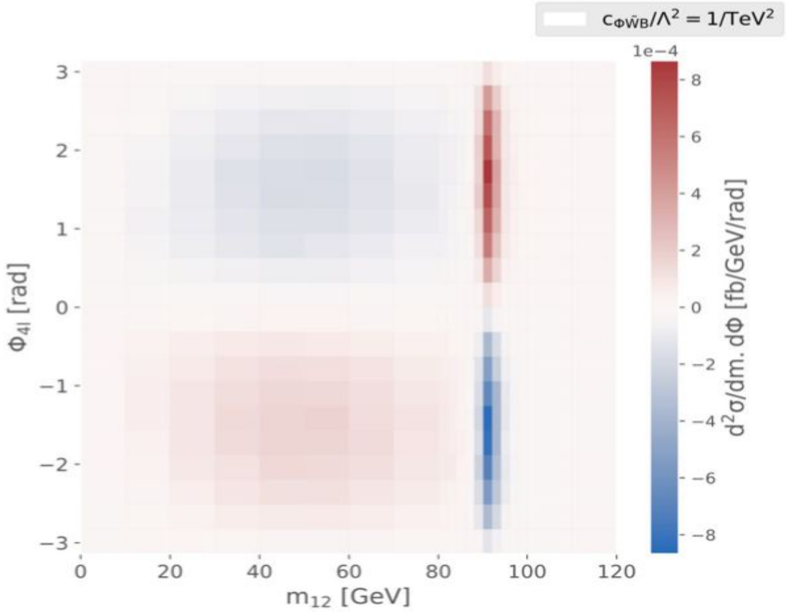


Decay channel consideration



Feature Importance





- Origin of extra sensitivity investigated using feature importance techniques, whereby the change in loss score is evaluated after decorrelating input variables in the trained network.
- Clear interplay between Φ_{4l} and m_{Z1} (highest mass of e^+e^- or $\mu^+\mu^-$ pair).

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

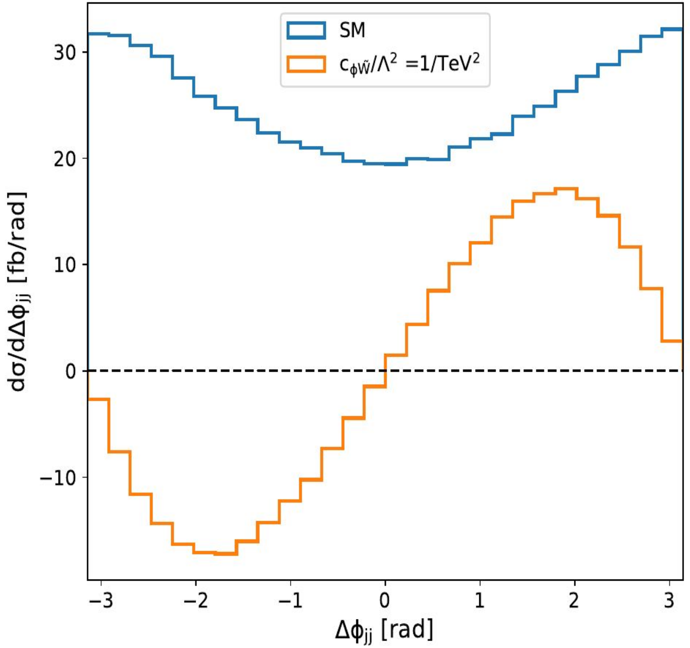
CP-odd observable	$c_{\Phi\widetilde{W}B}/\Lambda^2$ [TeV^{-2}]	$c_{\Phi\widetilde{B}}/\Lambda^2$ [TeV^{-2}]	$c_{\Phi\widetilde{W}}/\Lambda^2$ [TeV^{-2}]
$\Phi_{4\ell}$	[-6.2,6.2]	[-1.4,1.4]	[-30,30]
$\Phi_{4\ell}$, m_{12}	[-1.9,1.9]	[-0.85,0.85]	[-3.7,3.7]
O_{NN} (binary)	[-1.5,1.5]	[-0.75,0.75]	[-3.0,3.0]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

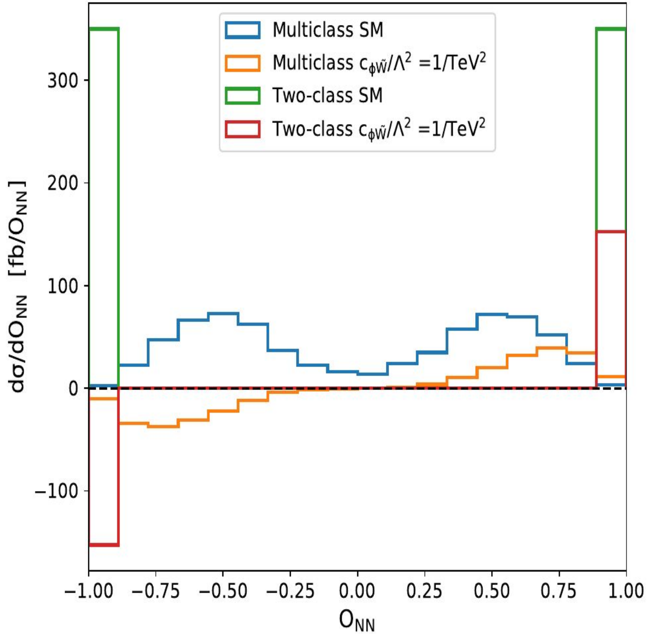
Sensitivity to specific operators established using the Profile Likelihood method, after normalising the MC samples to the number of events observed in the ATLAS analyses.

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

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: $\Delta\phi_{jj} = \phi(j_1) - \phi(j_2)$

Limits on CP-odd operators for

CP-odd observable	$C_{\Phi\widetilde{W}B}/\Lambda^2$	$C_{\Phi\widetilde{B}}/\Lambda^2$	$C_{\Phi\widetilde{W}}/\Lambda^2$
$\Delta\phi_{jj}$	[-21,21]	[-149,149]	[-0.60,0.60]
$O_{NN}(binary)$	[-11,11]	[-43,43]	[-0.66,0.66]
$O_{NN}(multi-class)$	[-10,10]	[-36,36]	[-0.42,0.42]

Conclusion

- A method to directly construct CP-odd observables using the output of neural networks
- NN optimise the separation of +ve and -ve interference using the full kinematic information
- Demonstrated the NN CP-odd observables for h-> 4l decay channel and the VBF production mechanism
- large improvements in sensitivity to CP-violating effects in the Higgs sector, demonstrated this using dimension-six effective field theory predictions for the interference contributions.
- Improving the sensitivity to CP-violating effects in $h \to 4l$ and VBF Higgs production is particularly important for the self-consistency of the dimension-six approach

Thank you

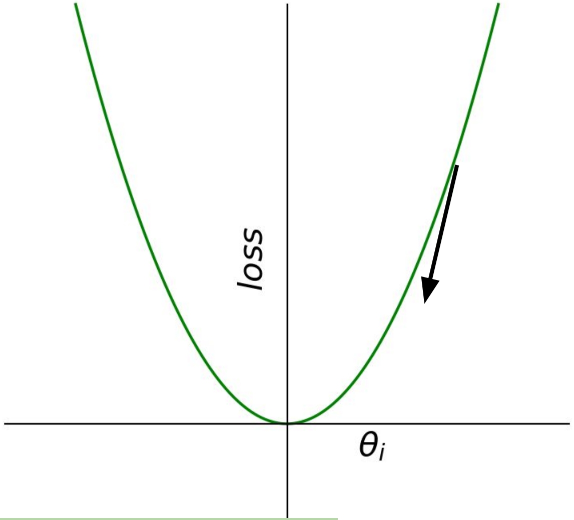
Machine Learning

Broadly, estimate a function given data samples.

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

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

$$\text{Optimise a loss function:} L(\hat{y}, y)$$



Linear Regression

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

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

Classification:

$$\boldsymbol{loss = -\frac{1}{N_{batch}} \sum_{i=1}^{i=N_{batch}} y_i \ \ln(\hat{y}(\vec{x}_i))}$$

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

CP-odd observable	$C_{\Phi\widetilde{W}B}/\Lambda^2$	$C_{\Phi\widetilde{B}}/\Lambda^2$	$C_{\Phi\widetilde{W}}/\Lambda^2$
$\Phi_{4\ell}$	[-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]

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.