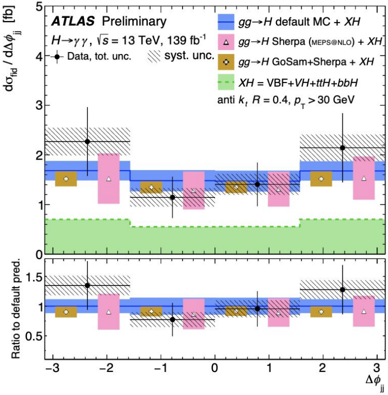
Machine-enhanced CP-asymmetries

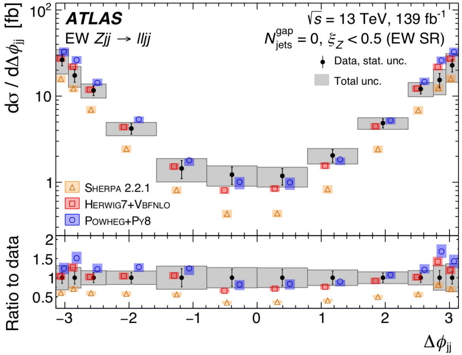
A. Bhardwaj, C. Englert, R. Hankache, A. Pilkington

LHC Higgs WG2: Discussion on CPV in Higgs boson interactions, 20th January 2022

Based on arXiv:2112.05052

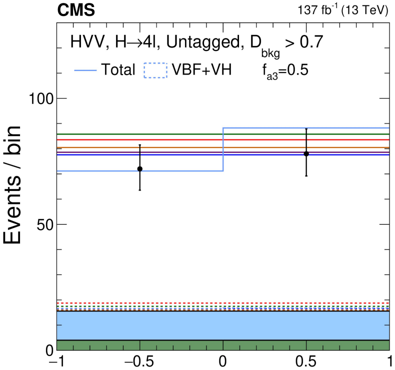
CP-sensitive observables in Higgs and EWK interactions



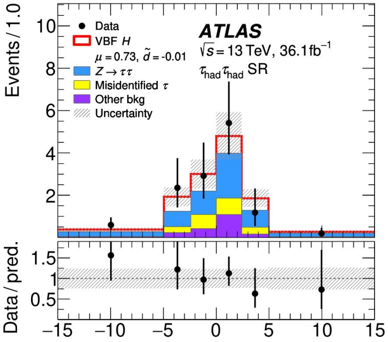


- Increasingly common to measure differential cross sections as a function of CP-odd observables, for both Higgs boson production and diboson/VBF/VBS processes.
- Advantages: model-independent, easy to combine in global fit.
- Disadvantages: sensitivity, i.e. how to optimise the phase-space?

CP-sensitive observables in Higgs: state-of-the-art



$$\mathcal{D}_{\rm int}(\boldsymbol{\Omega}) = \frac{\mathcal{P}_{\rm int}(\boldsymbol{\Omega})}{2\sqrt{\mathcal{P}_{\rm sig}(\boldsymbol{\Omega})\ \mathcal{P}_{\rm alt}(\boldsymbol{\Omega})}},$$



$$\mathcal{O}_{\rm opt} = \frac{2\,{\rm Re}(\mathcal{M}^*_{\rm SM}\mathcal{M}_{\rm CP\text{-}odd})}{|\mathcal{M}_{\rm SM}|^2}.$$

- Alternative approach is to use discriminants based on matrix-element information.
- Advantages: optimal in terms of sensitivity for a given analysis.
- Disadvantages: more complicated (i.e. time-consuming), not straightforward to combine in a global fit.

CP-odd observables from machine learning algorithms

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

$$|\mathcal{M}|^2 = |\mathcal{M}_{\rm SM}|^2 + 2\,\mathrm{Re}(\mathcal{M}_{\rm SM}^{*}\mathcal{M}_{\rm d6}) + |\mathcal{M}_{\rm d6}|^2,$$

- 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 $O_{NN} = P_+ - P_-$.
  - improve differential cross section measurements.

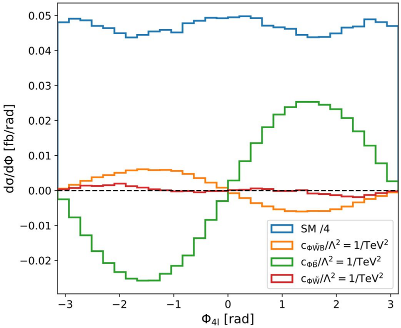
Some details of the theory framework and analysis

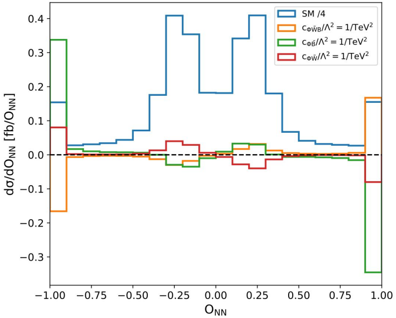
- Use Madgraph5_aMC@NLO to generate the events, with the SMEFTSim 3.0 package to incorporate dimension-six operators.
- Lagrangian then given by:

$$\mathcal{L} = \mathcal{L}_{\mathrm{SM}} + \sum_i \frac{c_i}{\Lambda^2} \widetilde{\mathcal{O}}_i \qquad \text{with} \qquad \begin{aligned} \mathcal{O}_{\Phi\widetilde{B}} &= \Phi^\dagger \Phi B^{\mu\nu} \widetilde{B}_{\mu\nu}\,, \\ \mathcal{O}_{\Phi\widetilde{W}} &= \Phi^\dagger \Phi W^{i\,\mu\nu} \widetilde{W}^i_{\mu\nu} \\ \mathcal{O}_{\Phi\widetilde{W}B} &= \Phi^\dagger \sigma^i \widetilde{W}^{i\,\mu\nu} B_{\mu\nu} \end{aligned}$$

- Neural net sensitivity compared to simple angular observables for two processes:
  - inclusive H→4l production
  - VBF Higgs production

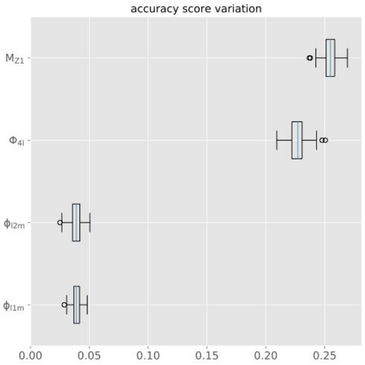
Application to H→4l

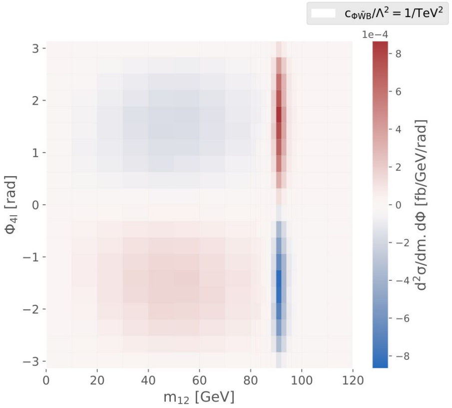




- Analysis carried out in the *Higgs Mass* region of the ATLAS inclusive *4l* measurement (JHEP 07, 005 (2021) for H→2e2μ events.
- Simple CP-odd variable (PRD 86, 095031 [2012]): $$\Phi_{4\ell} = \frac{\mathbf{q}_1 \cdot (\hat{\mathbf{n}}_1 \times \hat{\mathbf{n}}_2)}{|\mathbf{q}_1 \cdot (\hat{\mathbf{n}}_1 \times \hat{\mathbf{n}}_2)} | \times \cos^{-1}(\hat{\mathbf{n}}_1 \cdot \hat{\mathbf{n}}_2),$$
- NN trained using the interference induced by the $\mathcal{O}_{\Phi\widetilde{W}B}$ operator in the Warsaw basis.

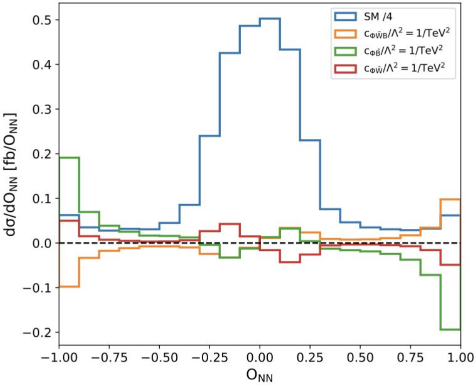
Feature importance





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

Multiclass models



- Multiclass = including the SM prediction as a third class in the training ($P_+ + P_- + P_{sm} = 1$).
- Optimises the separation of the interference contributions for a process, by accounting for any kinematic differences between the SM and the interference.

Limits on CP-odd operators for H→2e2μ

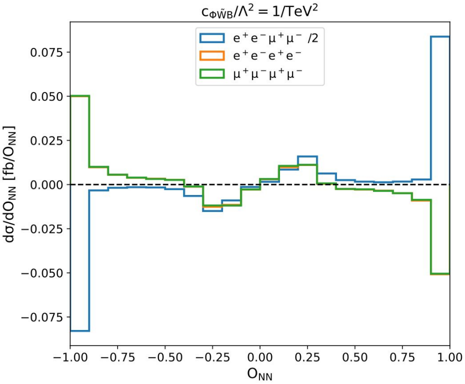
CP-odd observable	$c_{\Phi\widetilde{W}B}/\Lambda^2$ [TeV^{-2}]	$c_{\Phi\widetilde{B}}/\Lambda^2$ [TeV^{-2}]	$c_{\Phi\widetilde{W}}/\Lambda^2$ [TeV^{-2}]
$\Phi_{4\ell}$	[-6.2,6.2]	[-1.4,1.4]	[-30,30]
$\Phi_{4\ell}$, m_{12}	[-1.9,1.9]	[-0.85,0.85]	[-3.7,3.7]
O_{NN} (binary)	[-1.5,1.5]	[-0.75,0.75]	[-3.0,3.0]
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 improve sensitivity w.r.t binary classification, i.e. networks learn the difference between the SM and the interference contributions.
- Double-differential analysis of Φ_{4l} and m_{Z1} captures most of the sensitivity gained by NN

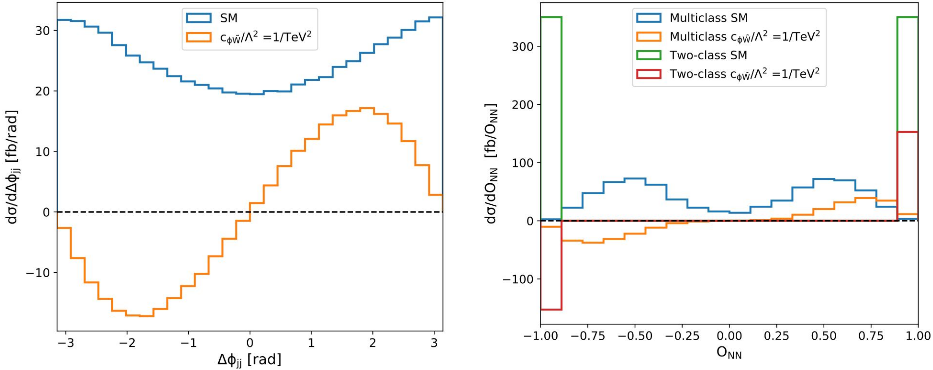
Decay channel considerations:



- Difference in sign and magnitude of interference depending on channel (2e2μ, 4e, 4μ)
  - Mispairing of leptons in 4e and 4μ channels when both pairs are off-shell.
  - Additional diagrams in the 4e/4μ channels.

→ Channels need to be measured independently.

Application to VBF Higgs production



- Analysis carried out in the *VBF_1* region of the ATLAS H→ττ analysis (ATLAS-CONF-2021-044)
- Classic CP-odd variable: $\Delta\phi_{jj} = \phi(j_1) - \phi(j_2)$
- NN trained using the interference induced by the $\mathcal{O}_{\Phi\widetilde{W}}$ operator in the Warsaw basis.

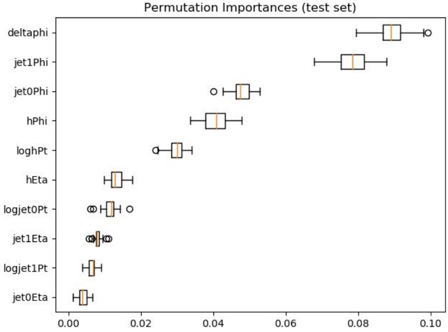
VBF Higgs results

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}]
$\Delta\phi_{jj}$	[-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]

Main conclusions:

- multiclass training more important
- matching NN sensitivity using differential cross sections will be trickier

Both of these features arise because VBF is a multiscale process.



Summary and outlook

Neural networks offer a simple approach to studying CP asymmetries:

- can distinguish between the positive and negative interference contributions predicted for a particular operator.
- CP-odd observable can be directly constructed from the network output.
- Origin of CP-asymmetries can be easily explored and used to improve differential cross section measurements
- Multiclass networks can be used to optimise an analysis, by exploiting differences in kinematics between the interference and Standard-Model predictions.