Test of CP Invariance in Higgs Boson Vector Boson Fusion Production in the $H \rightarrow \gamma\gamma$ Channel using 165 fb-1 ATLAS partial Run-3 data

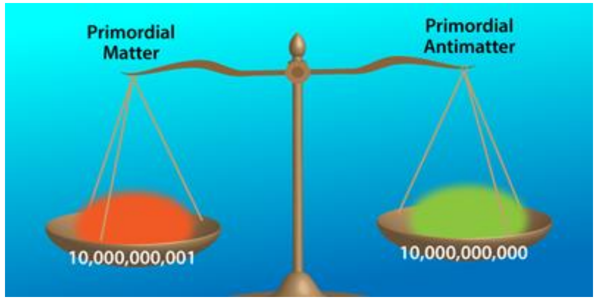
Shudong Wang, IHEP CAS

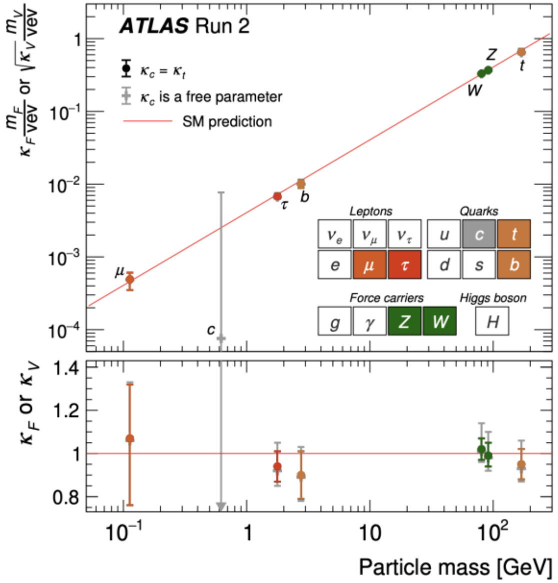
The 11th CLHCP 2025

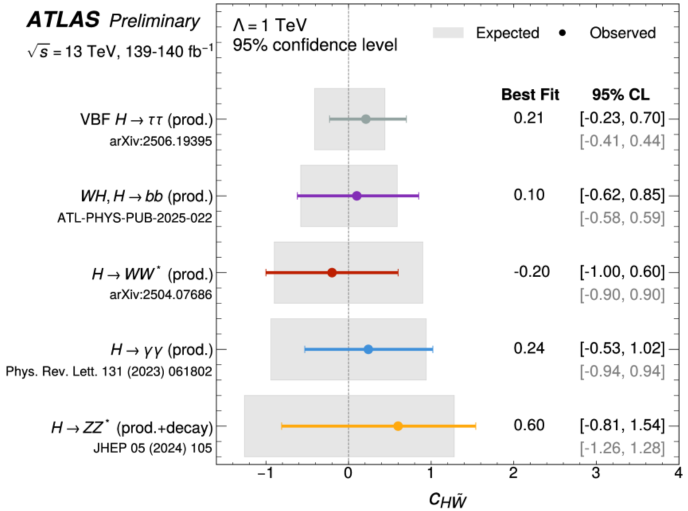
Oct 29 - Nov 02 2025, Xinxiang

Introduction

- **CP violation: one of the key condition of baryon asymmetry**
  - Current knowledge of CPV: CKM matrix, PMNS matrix, are **insufficient**
  - **Must exists some unknown mechanism!**
- **Higgs boson discovered in 2012**
  - Significant achievements on Higgs mass, width, spin, parity and couplings to gauge bosons and fermions…
  - A new field of CPV search: **Higgs coupling**
  - HVV CP property in SMEFT Warsaw basis: $c_{H\widetilde{W}}$ has been measured in several channels in Run2.
- **In this talk:**
  - Probing CP property of HVV vertex in VBF $H \rightarrow \gamma\gamma$ channel using ATLAS partial Run3 (2022-2024) data.







Theoretical Framework

- **CP properties in SMEFT:**
  - Consider **only dimension-6** operators: $\mathcal{L} = \mathcal{L}_{SM} + \sum_i \frac{c_i^{(6)}}{\Lambda^2}\mathcal{O}_i^{(6)}$.

Operator	Structure	Coupling
$O_{\Phi\tilde{W}}$	$\Phi^\dagger\Phi\tilde{W}^I_{\mu\nu}W^{\mu\nu I}$	$c_{H\tilde{W}}$
$O_{\Phi\tilde{W}B}$	$\Phi^\dagger\tau^I\Phi\tilde{W}^I_{\mu\nu}B^{\mu\nu}$	$c_{H\tilde{W}B}$
$O_{\Phi\tilde{B}}$	$\Phi^\dagger\Phi\tilde{B}_{\mu\nu}B^{\mu\nu}$	$c_{H\tilde{B}}$

Dominant in VBF process

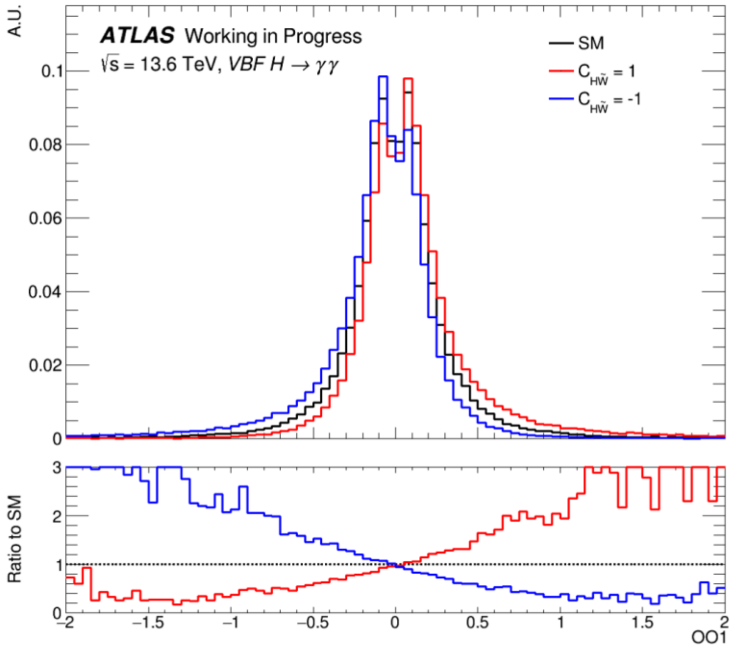
$\mathcal{O}(10)$ less sensitive but worth to probe

- **Sensitive observable: Optimal observable (OO)**
  - Defined from the **matrix element ratio**:

$$|\mathcal{M}_{MIX}(c)|^2 = |\mathcal{M}_{SM}|^2 + 2R\big(\mathcal{M}_{SM}\mathcal{M}^*_{BSM}(c)\big) + |\mathcal{M}_{BSM}|^2$$

$$OO(c) = \frac{2R(\mathcal{M}_{SM}\mathcal{M}^*_{BSM}(c))}{|\mathcal{M}_{SM}|^2}$$

  - Has been used in several CP analyses in Run 2.
  - Calculate the ME with MadGraph as ATLAS Run 2 HZZ CP analysis.



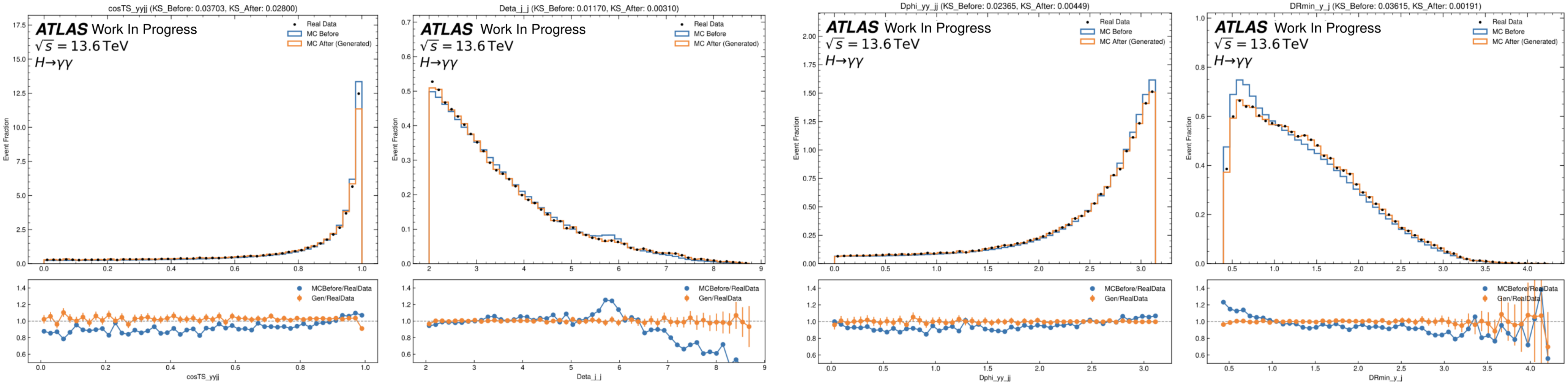
Analysis strategy

- **Data and MC**
  - using 2022-2024 Run3 data (165 fb^{-1})
  - Standard ggF Higgs, $\gamma\gamma$ continuum and Dedicated SM/BSM VBF Higgs MC.
- **Trigger, object definition and pre-selection**
  - **Common criteria used in ATLAS $H \to \gamma\gamma$ analysis**
  - Single photon and di-photon triggers
  - Common $H \to \gamma\gamma$ selections + VBF cuts

Event-level selection requirements	
Number of selected photons	$N_\gamma \geq 2$
Number of selected jets	$N_j \geq 2$
Invariant di-photon mass	$105\,\text{GeV} < m_{\gamma\gamma} \leq 160\,\text{GeV}$
Transverse momentum of the leading photon relative to the invariant di-photon mass	$p_{\text{T},\gamma_1}/m_{\gamma\gamma} > 0.35$
Transverse momentum of the subleading photon relative to the invariant di-photon mass	$p_{\text{T},\gamma_2}/m_{\gamma\gamma} > 0.25$
Pseudorapidity gap between the two leading jets	$\Delta\eta_{jj} > 2.0$
Zeppenfeld observable	$\lvert\eta^{\text{Zepp}}\rvert < 5.0$

Analysis strategy

- **Discriminate VBF events from ggF and non-resonant background.**
  - Using a **3-class DNN**
  - Input features: **16 features** pruned from 25 investigated features.
- **Input feature modeling in continuum $\gamma\gamma$ background**
  - The sherpa MC + data-driven γ-j cannot perfectly describe the sideband data.
  - Use **flow-matching model** to generate the feature distributions from sideband data to avoid performing training on mis-modelled inputs.

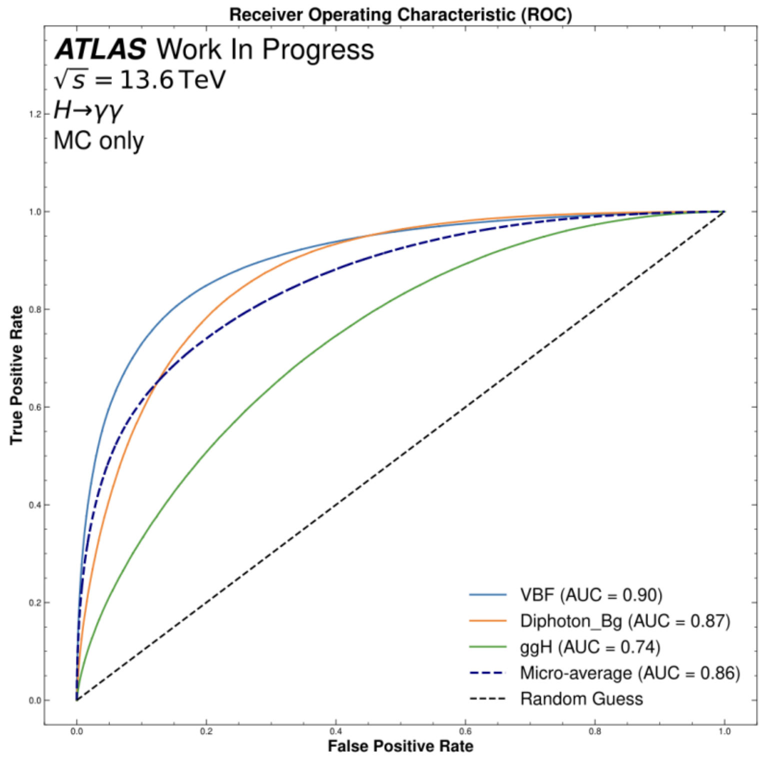


Analysis strategy

- **Model training: DNN trained with PyTorch**
  - 2-fold training to avoid bias from overfitting.

Layer #	Layer Type	Input Dim	Output Dim	Activation	Normalization	Dropout
1	Linear (Input)	16	128	ReLU	LayerNorm	No
2	Linear	128	64	ReLU	LayerNorm	Yes (0.1)
3	Linear	64	32	ReLU	LayerNorm	Yes (0.1)
4	Linear (Output)	32	3	—	—	No

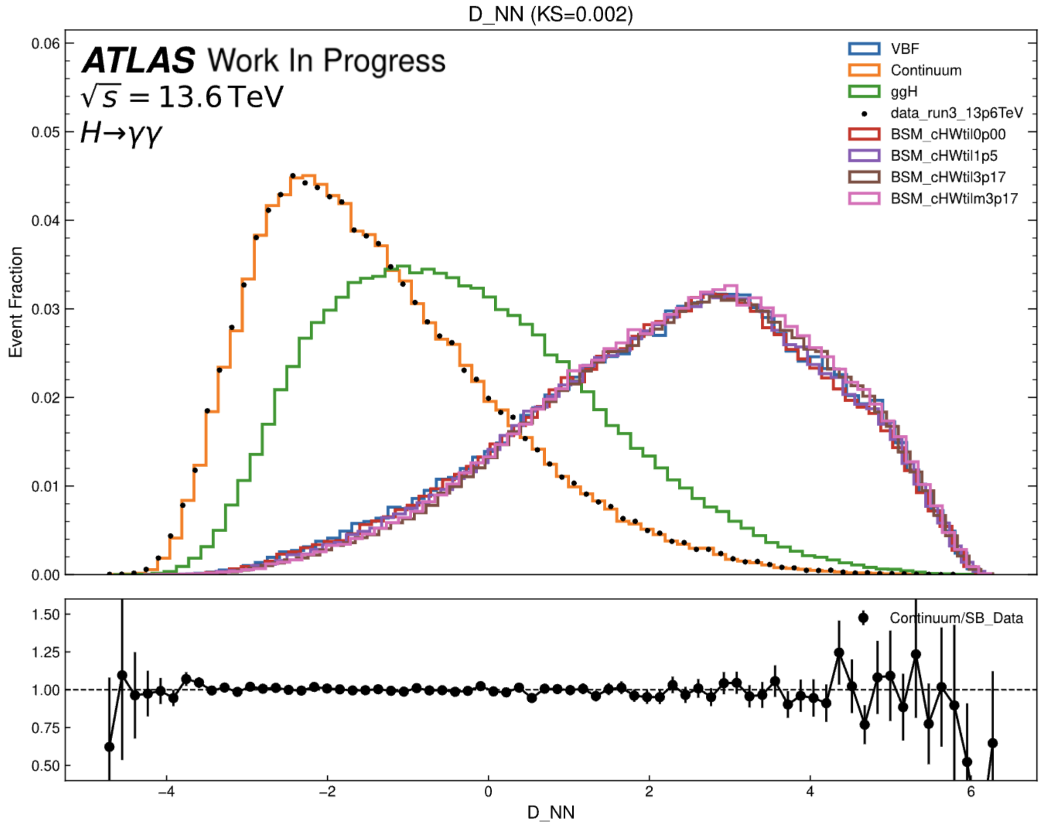
Table 7.4: Model Architecture with $in\_len = 16$, $num\_classes = 3$, and $dropout\_rate = 0.1$



Analysis strategy

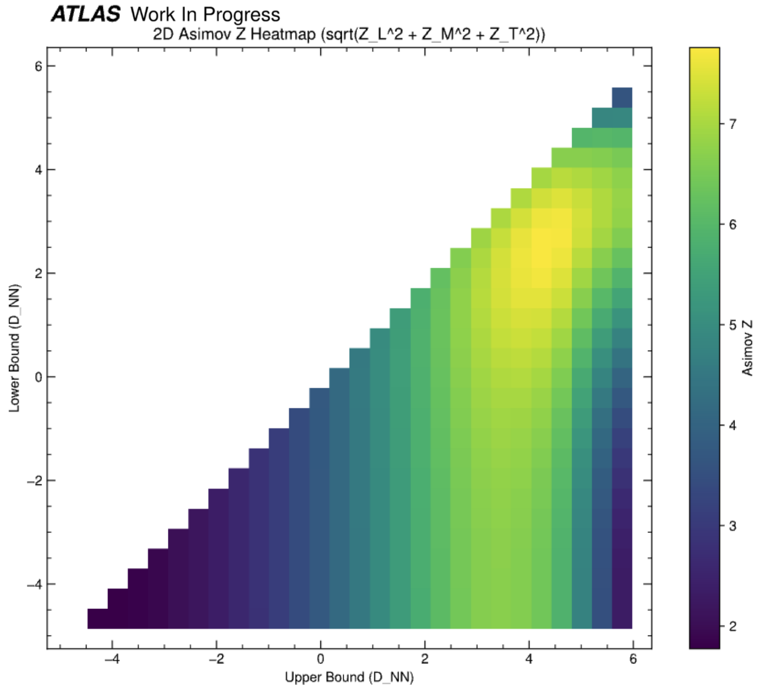
- **Event categorization: for high VBF significance**
  - The NN output provides the probability score for each class.
  - Define a discriminator:

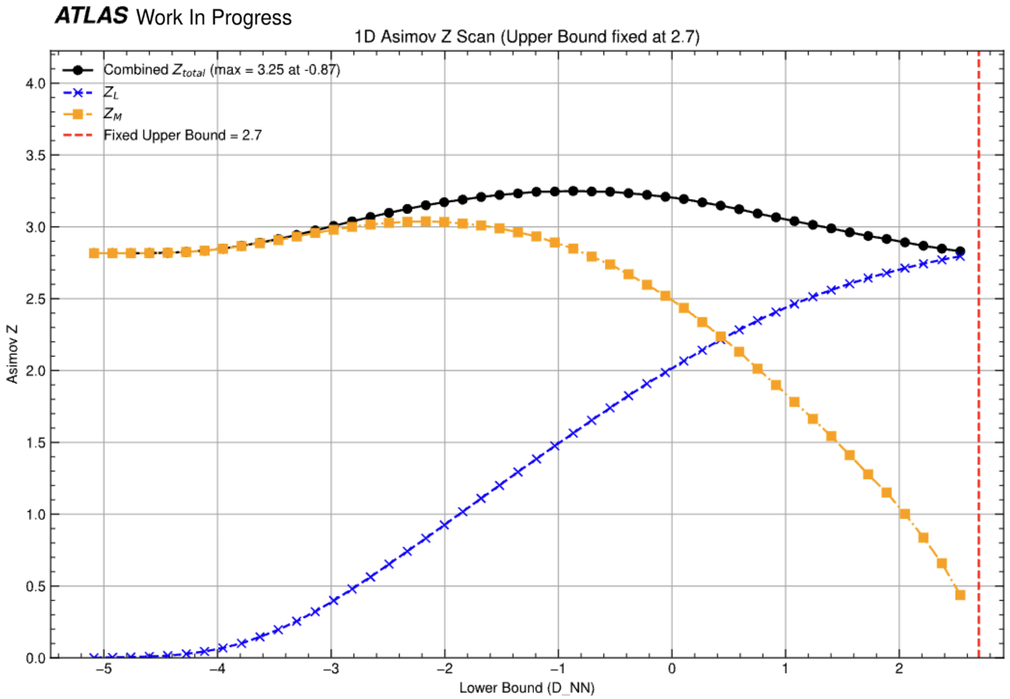
$$D_{\mathrm{NN}} = \ln\frac{P_{VBF}}{f_{ggH}\cdot P_{ggH} + f_{continum}\cdot P_{continuum}},\ f_{ggH} = \frac{N_{ggH}}{N_{ggH} + N_{continuum}},\ f_{continuum} = 1 - f_{ggH}$$



Analysis strategy

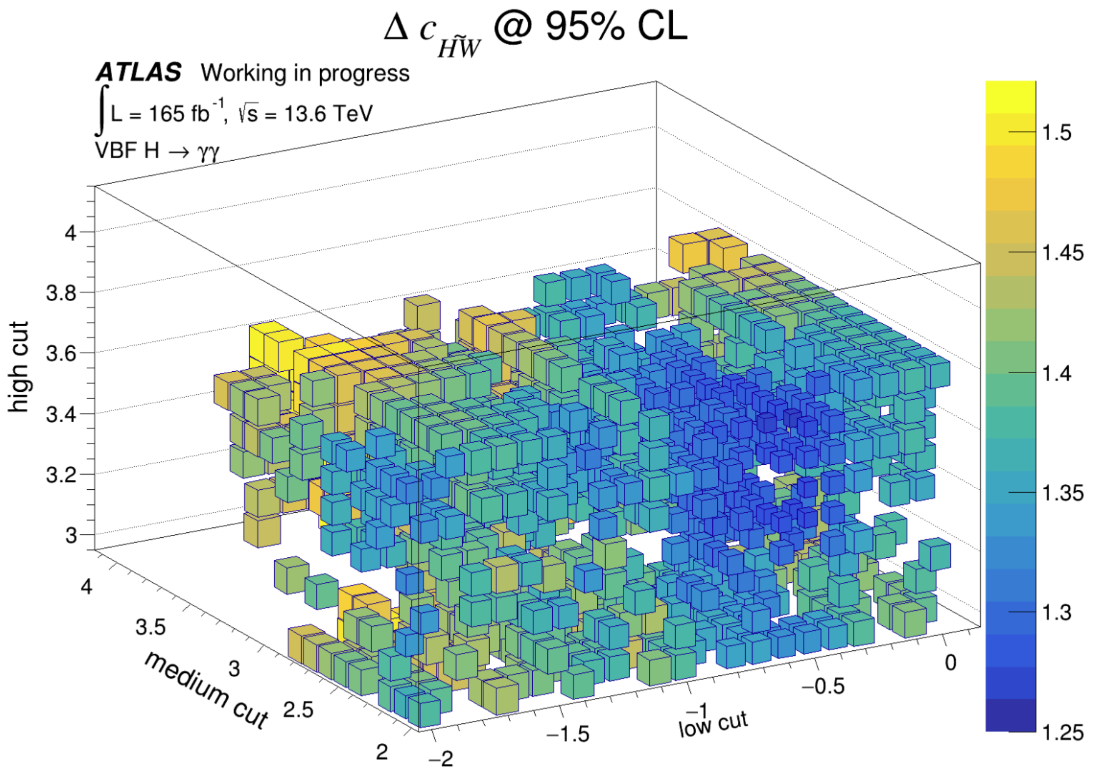
- **Event categorization – Maximize VBF significance**
  - First, split events into 3 categories with D_{NN}: **Tight, Medium, Loose** by a **2D significance scan**.
  - $Z = \sqrt{2[(s+b)\cdot\ln\left(1+\frac{s}{b}\right) - s]}, Z^{combined} = \sqrt{(Z^L)^2 + (Z^M)^2 + (Z^T)^2}$
  - Loose is further split into 2 categories: **Loose** and **VeryLoose** by a **1D scan**, while optimizing the ggF significance.
  - Gives a rough estimation of categorization boundaries which will be further optimized.





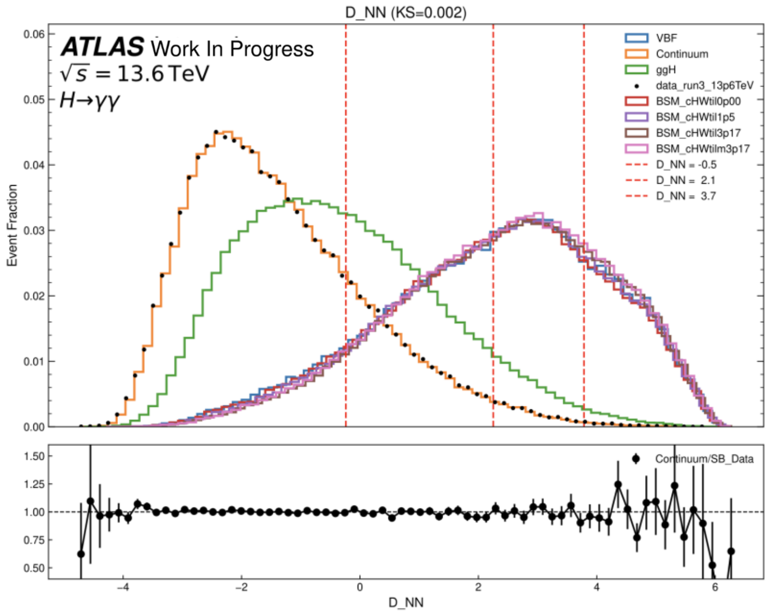
Analysis strategy

- **Event categorization – Minimize the $c_{H\widetilde{W}}$ 95% interval**
  - A more direct way for the optimization, as $c_{H\widetilde{W}}$ is the POI of analysis.
  - Create the workspace with dummy signal/bkg modeling.
  - Simultaneous 3D scan for 3 cuts, based on Z_{VBF} results.



Requiring sideband data > 9 in any bins.
Optimized cut:

- Low cut: -0.5
- Medium cut: 2.1
- High cut: 3.7



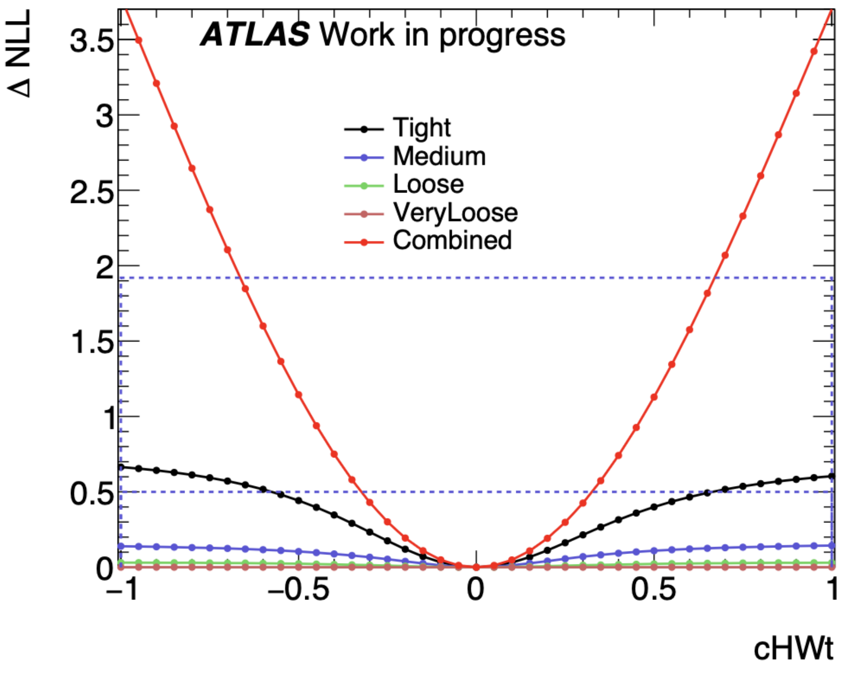
Statistics model

- **Parameterization of CPV effects: based on morphing.**
  - $\sigma(c) = \sigma_{SM} + \sigma_{int} + \sigma_{quad} = \sigma_{SM} + A \cdot c + B \cdot C^2$.
  - Derive A and B from VBF base samples with different $c_{H\widetilde{W}}$.
- **Nominal systematics schemes**
- **Fit $m_{\gamma\gamma}$ in each OO bins to derive the OO shape**
  - Likelihood constructed from the fit in all bins.
  - Floating μ_{VBF} to have a shape only measurement.
  - Template fit for different $c_{H\widetilde{W}}$ hypotheses.
  - POI: $c_{H\widetilde{W}}, \mu_{VBF}, \mu_{\mathrm{ggF}}$

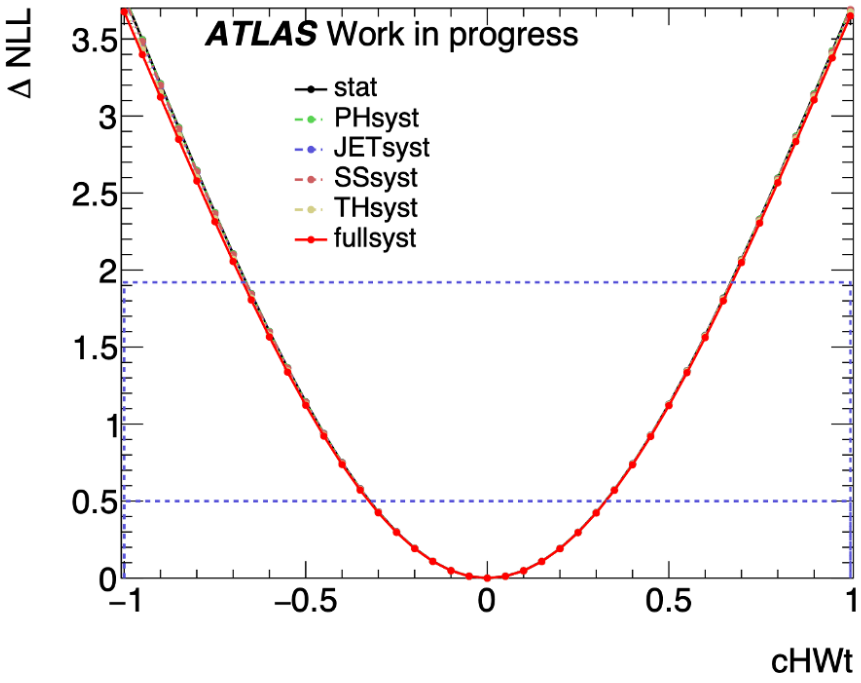
	systematic source	Implementation
Exp.	lumi	N_{tot} log-normal
	jet	N_t asymmetric
	photon Isolation	N_t asymmetric
	photon Identification	N_t asymmetric
	photon Scale	N_t asymmetric
	photon Resolution	N_t asymmetric
	PU	N_t asymmetric
Theo.	PDF	N_t log-normal
	α_s	N_t log-normal
	QCD	N_t log-normal
	shower	N_t Gaussian
Mass.	ATLAS-CMS m_H	μ_{CB} Gaussian
	Photon Scale	μ_{CB} Gaussian
	Photon Resolution	σ_{CB} log-normal
Bkg.	spurious signal	N_{spur} Gaussian

Expected results

- **Expected constrain on $c_{H\widetilde{W}}$**
  - Estimate ggF from **data**.
    - **Floating** μ_{ggF} in the fit.
    - For less dependence on SM.
  - Still **statistical dominant**.



stat-only



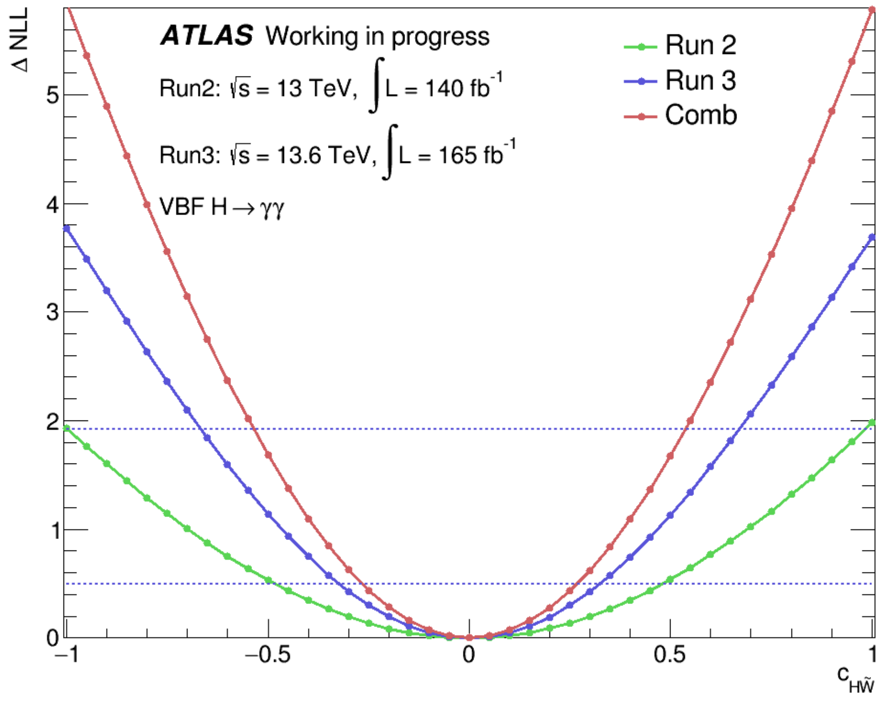
syst

Combination of Run2 and Run3

- **Workspace-level combination**
  - Combined likelihood

$$\mathcal{L}(\boldsymbol{\mu}, \boldsymbol{\theta}; \text{data}) = \prod_{c=1}^{N_{\text{cats}}} \mathcal{L}_c(\boldsymbol{\mu}, \boldsymbol{\theta}; \text{data}) \prod_{k=1}^{N_{\text{cons}}} \mathcal{G}(\theta_k; \tilde{\theta}_k)$$

  - Stat-only results for now.



Summary

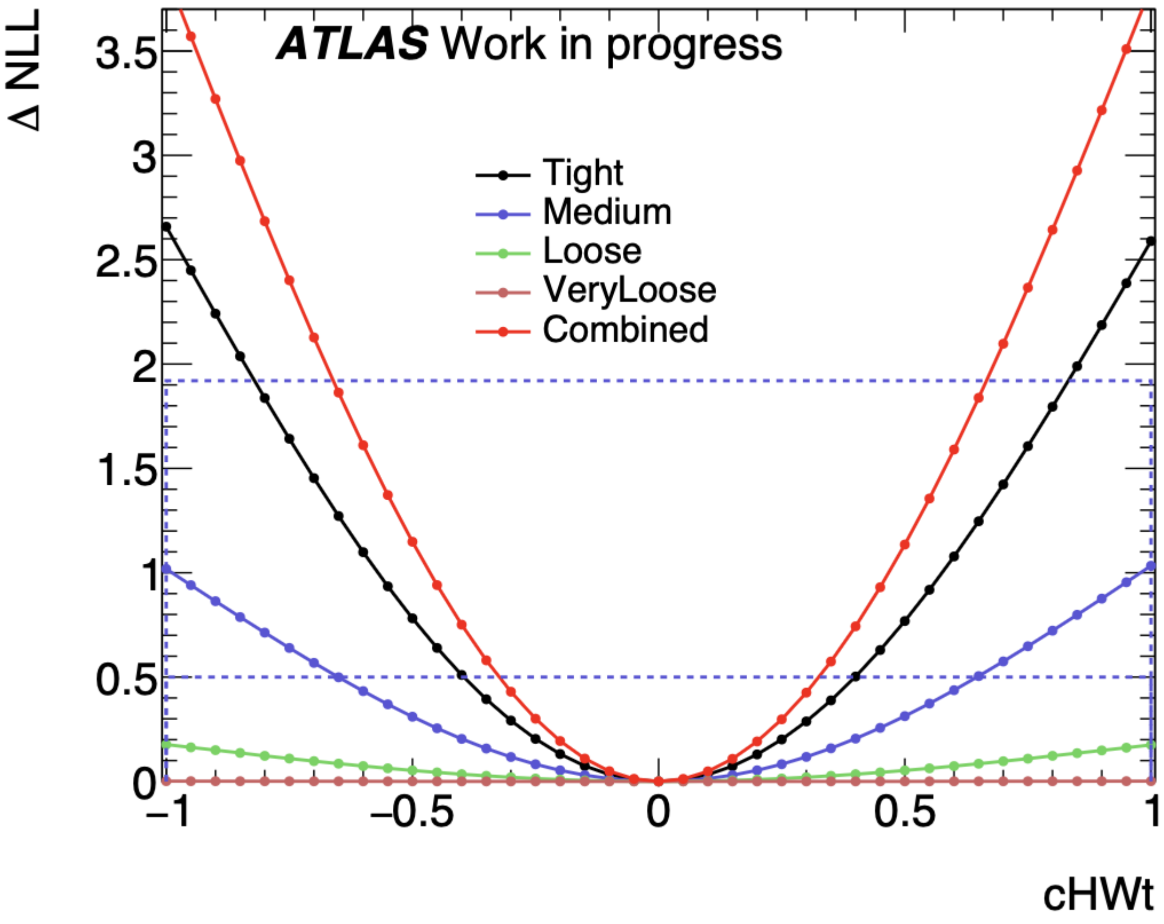
- **Searching for new CP violation sources beyond SM**
  - Motivated by baryon asymmetry puzzle.
  - HVV vertex provides a sensitive window to search new physics.
  - Great achievements from ATLAS Run2 HVV CP property studies.

- **An analysis probing the HVV CP property with ATLAS partial Run3 data is converging**
  - Preliminary result for $c_{H\tilde{W}}$ is derived (statistical dominant).
  - Stat-only result for the combination of Run2 and Run3 CP study is derived.

- **Future Plan:**
  - **Complete results.**

Backup

Backup


ATLAS Work in progress
Δ NLL
3.5
3
2.5
2
1.5
1
0.5
0
−1
−0.5
0
0.5
1
cHWt
Tight
Medium
Loose
VeryLoose
Combined