Universidad de Oviedo

MINISTERIO
DE CIENCIA, INNOVACIÓN
Y UNIVERSIDADES

Cofinanciado por
la Unión Europea

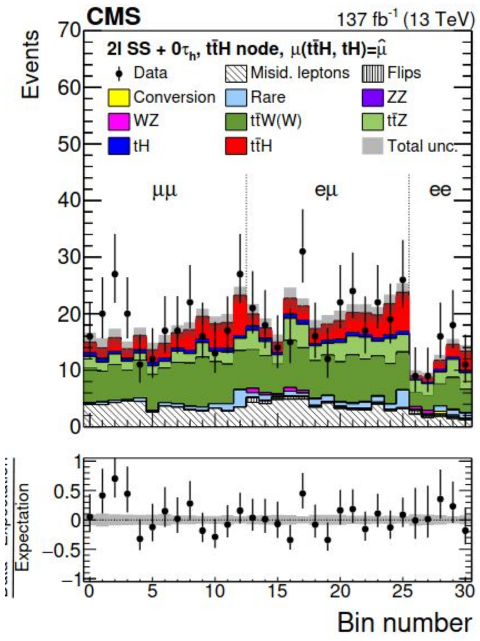
AGENCIA
ESTATAL DE
INVESTIGACIÓN

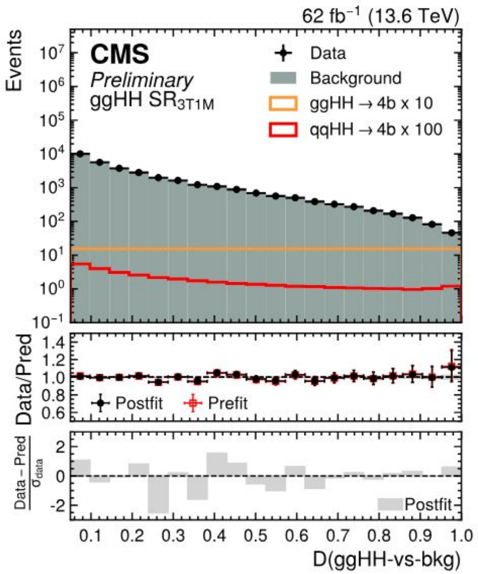
Simulation-Based Inference in CMS

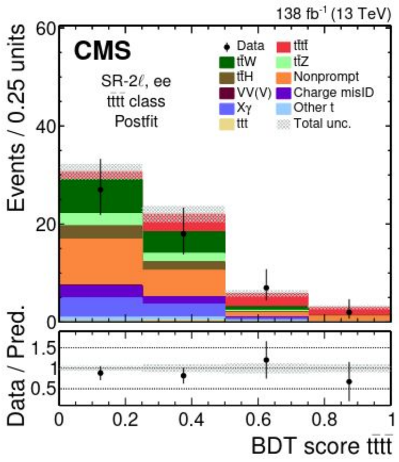
Sergio Sánchez Cruz (University of Oviedo)
26.02.2026

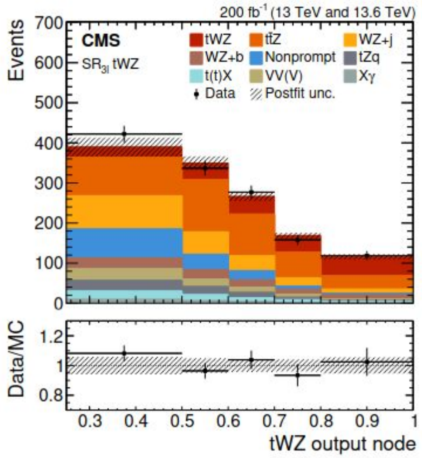
Recent activities in SBI

- Zillions of CMS analyses build observables using machine learning
- SBI is most useful for multiparameter estimation, beyond signal vs background picture
- Recent analyses using SBI to build observables sensitive to e.g. EFTs
  - Constraints on EFT in VH production (JHEP 03 (2025) 114)
  - Search for CP violation in top+Z (PLB 869 (2025) 139857)



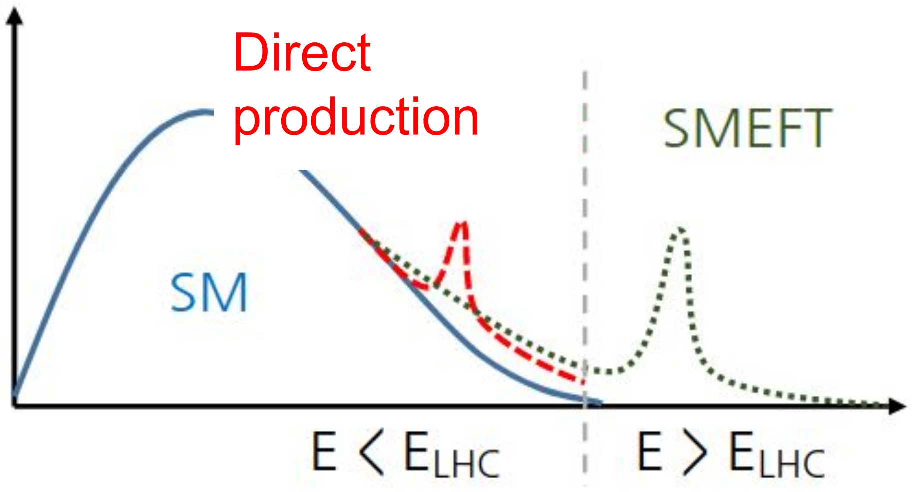






SMEFT searches

- Lack of clear evidence from direct searches
- New physics may lay outside the experimental energy scale
- SMEFT parameterizes all potential new physics contributions
  - O(1000) independent parameters

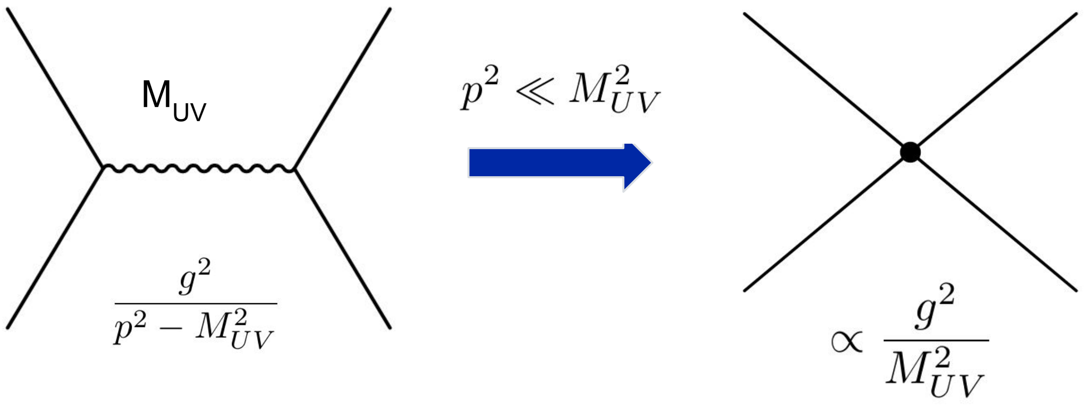


Wilson coefficient

$$\mathcal{L} = \mathcal{L}_{SM} + \sum_i \frac{\theta_i}{\Lambda^2}\mathcal{O}_i$$

UV scale

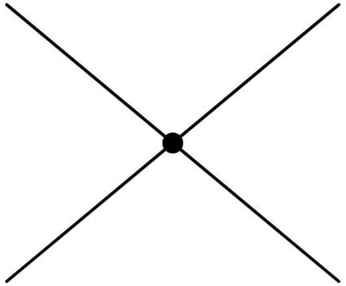
EFT operator



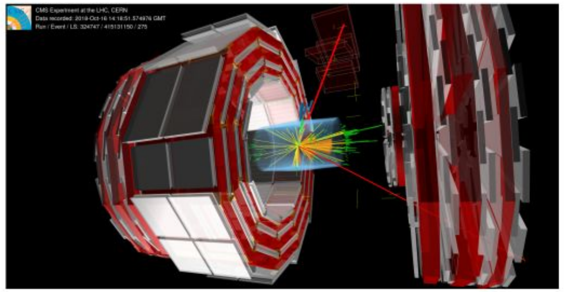
SMEFT searches in CMS

- Predictions are drawn using weighted simulated samples:

$$w(\theta) = w_{SM} + \sum \theta_i w_i(z) + \sum_{ij} \theta_i \theta_j w_{ij}(z)$$



$$p(z, x|\theta)$$



Parton-level

Can sample from the joint pdf

Detector-level

General SBI in CMS

- In a generic analysis…

$$L(\mathcal{D}|\boldsymbol{\theta}) = \frac{e^{-\mathcal{L}\sigma(\boldsymbol{\theta})}}{N!} \prod_{1\leq i\leq N} \mathcal{L}\sigma(\boldsymbol{\theta}) p(\mathbf{x}_i|\boldsymbol{\theta})$$

$$q_\theta = \mathcal{L}(\sigma(\boldsymbol{\theta}) - \sigma(\boldsymbol{\theta_0})) - \sum_{1\leq i\leq N} \log \frac{\sigma(\boldsymbol{\theta}) p(\mathbf{x}|\boldsymbol{\theta})}{\sigma(\boldsymbol{\theta_0}) p(\mathbf{x}|\boldsymbol{\theta_0})}$$

Easy

Intractable

- For EFTs…

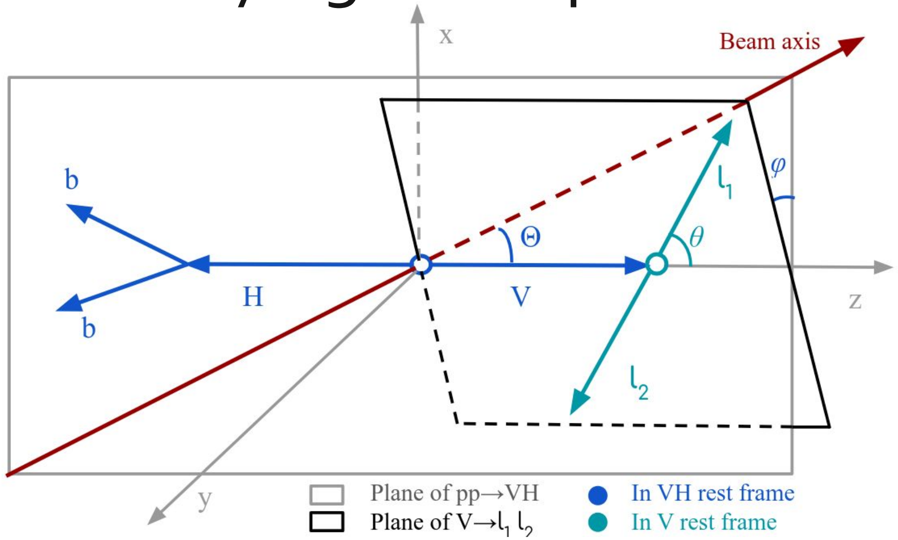
$$R(\mathbf{x}|\boldsymbol{\theta},\boldsymbol{\theta}_0) = 1 + \sum_{1\leq i\leq M} (\theta_i - \theta_0) R_i(\mathbf{x}) + \sum_{1\leq i\leq j\leq M} \frac{1}{2}(\theta_i - \theta_0)(\theta_j - \theta_0) R_{i,j}(\mathbf{x})$$

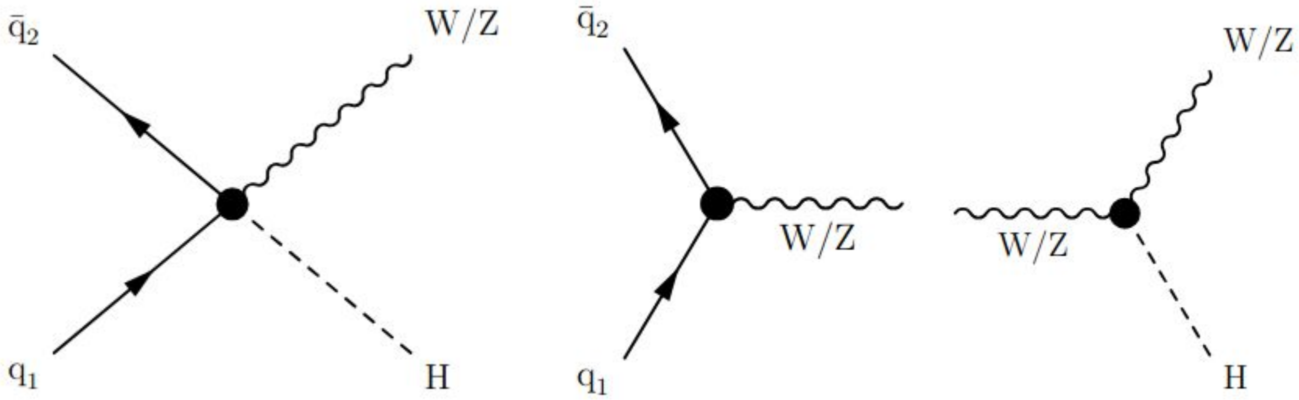
- Both analyses learn (part of) R(x|theta) to build optimal observables

VH(bb) production - analysis in a nutshell

- Exploring decay channels:
  - 0 leptons: Z → nn
  - 1 lepton: W → ln
  - 2 leptons: Z → ll
- Studying 6 independent POIs

Operator	Definition	Operator	Definition
$\mathcal{O}_{\mathrm{Hq}}^{(1)}$	$i\mathrm{H}^\dagger \overleftrightarrow{D}_\mu \mathrm{H}\bar{\mathrm{q}}_\mathrm{L}\gamma^\mu \mathrm{q}_\mathrm{L}$	$\mathcal{O}_{\mathrm{HWB}}$	$\mathrm{H}^\dagger\sigma^a\mathrm{H}\mathrm{W}^a_{\mu\nu}\mathrm{B}^{\mu\nu}$
$\mathcal{O}_{\mathrm{Hq}}^{(3)}$	$i\mathrm{H}^\dagger\sigma^a \overleftrightarrow{D}_\mu \mathrm{H}\bar{\mathrm{q}}_\mathrm{L}\sigma^a\gamma^\mu \mathrm{q}_\mathrm{L}$	$\mathcal{O}_{\mathrm{H\widetilde{W}B}}$	$\mathrm{H}^\dagger\sigma^a\mathrm{H}\mathrm{W}^a_{\mu\nu}\widetilde{\mathrm{B}}^{\mu\nu}$
$\mathcal{O}_{\mathrm{Hu}}$	$i\mathrm{H}^\dagger \overleftrightarrow{D}_\mu \mathrm{H}\bar{\mathrm{u}}_\mathrm{R}\gamma^\mu \mathrm{u}_\mathrm{R}$	$\mathcal{O}_{\mathrm{HW}}$	$(\mathrm{H}^\dagger\mathrm{H})\,\mathrm{W}^a_{\mu\nu}\mathrm{W}^{a\mu\nu}$
$\mathcal{O}_{\mathrm{Hd}}$	$i\mathrm{H}^\dagger \overleftrightarrow{D}_\mu \mathrm{H}\bar{\mathrm{d}}_\mathrm{R}\gamma^\mu \mathrm{d}_\mathrm{R}$	$\mathcal{O}_{\mathrm{H\widetilde{W}}}$	$(\mathrm{H}^\dagger\mathrm{H})\,\mathrm{W}^a_{\mu\nu}\widetilde{\mathrm{W}}^{a\mu\nu}$
$\mathcal{O}_{\mathrm{HD}}$	$(\mathrm{H}^\dagger D^\mu\mathrm{H})^*(\mathrm{H}^\dagger D_\mu\mathrm{H})$	$\mathcal{O}_{\mathrm{HB}}$	$(\mathrm{H}^\dagger\mathrm{H})\,\mathrm{B}_{\mu\nu}\mathrm{B}^{\mu\nu}$
$\mathcal{O}_{\mathrm{H}\Box}$	$(\mathrm{H}^\dagger\mathrm{H})\Box(\mathrm{H}^\dagger\mathrm{H})$	$\mathcal{O}_{\mathrm{H\widetilde{B}}}$	$(\mathrm{H}^\dagger\mathrm{H})\,\mathrm{B}_{\mu\nu}\widetilde{\mathrm{B}}^{\mu\nu}$





VH production - BIT strategy

- Boosted Information Tree* trained to minimize the cross-entropy

$$L\left[g(\mathbf{x}|\boldsymbol{\theta})\right] = -\left(\int \mathrm{d}x\mathrm{d}z \frac{\sigma(\boldsymbol{\theta})}{\sigma(\boldsymbol{\theta}_0)} \frac{p(\mathbf{x},\mathbf{z}|\boldsymbol{\theta})}{p(\mathbf{x},\mathbf{z}|\boldsymbol{\theta}_0)} \log(1 - g(\mathbf{x}|\boldsymbol{\theta})) + \log(g(\mathbf{x}|\boldsymbol{\theta}))\right)$$

$$R(\mathbf{x}|\boldsymbol{\theta},\boldsymbol{\theta}_0) = 1 + \sum_{1\le i\le M} (\theta_i - \theta_0) R_i(\mathbf{x}) + \sum_{1\le i\le j\le M} \frac{1}{2}(\theta_i - \theta_0)(\theta_j - \theta_0) R_{i,j}(\mathbf{x})$$

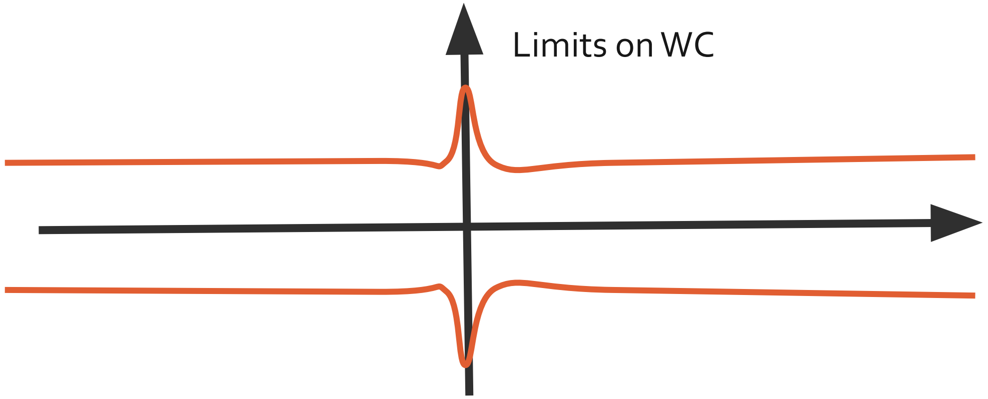
- Likelihood ratio obtained from the likelihood ratio “trick”:

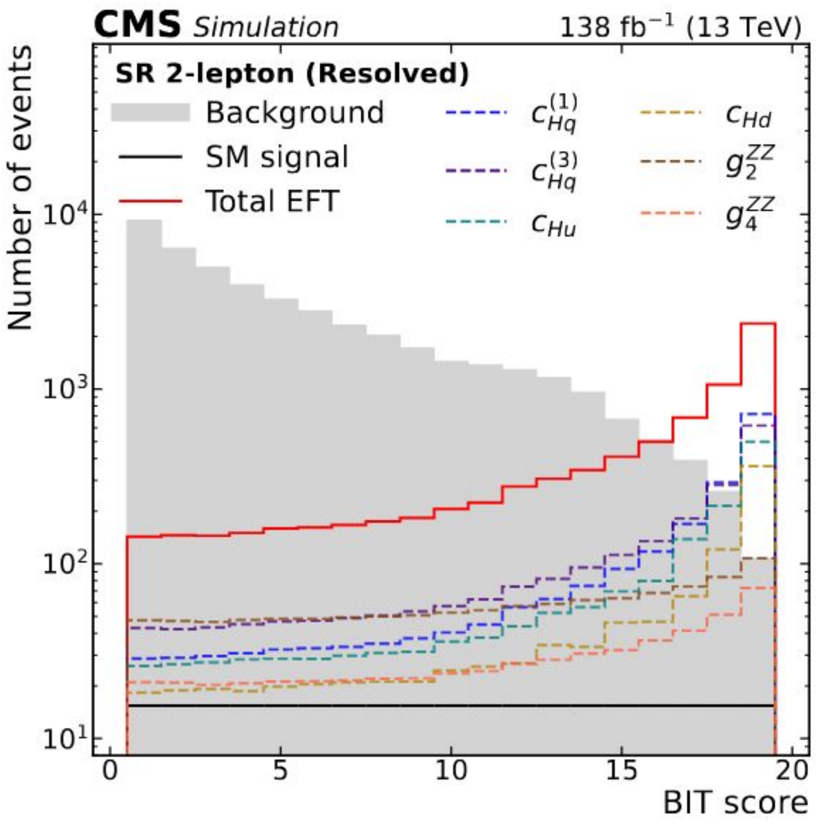
$$g^*(\mathbf{x}|\boldsymbol{\theta}) = \frac{1}{1 + R(\mathbf{x}|\boldsymbol{\theta},\boldsymbol{\theta}_0)}$$

- N BITs are trained to solve each term in the quadratic expansion

VH production - building optimal observables

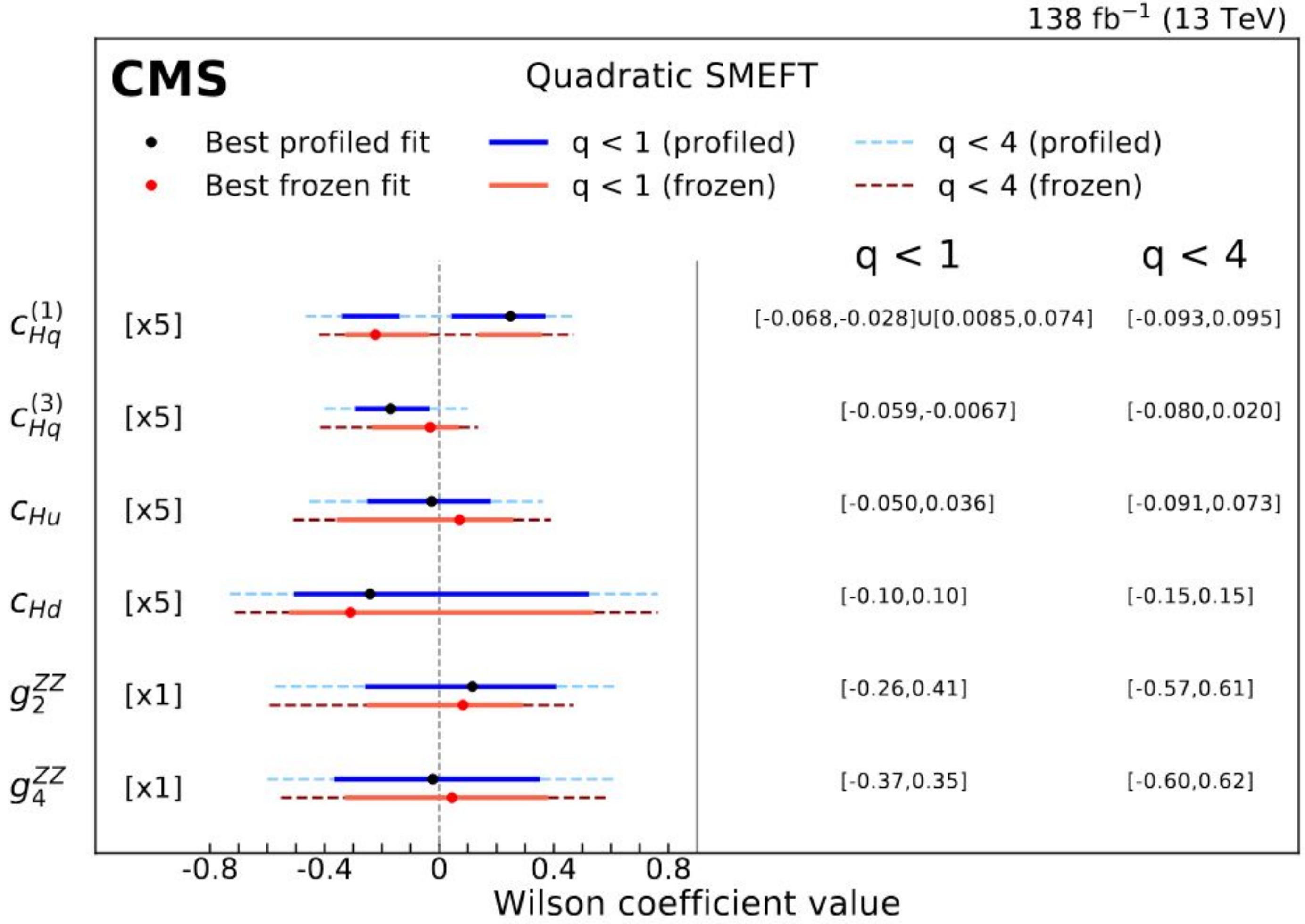
- BIT score is parametric on the WCs
- Optimal discriminator for each WC chosen
  - How to choose the best one?

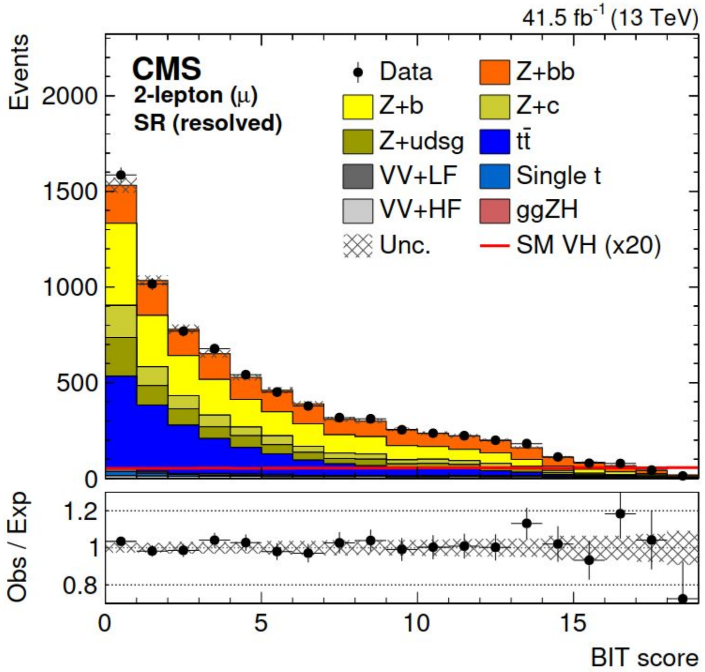


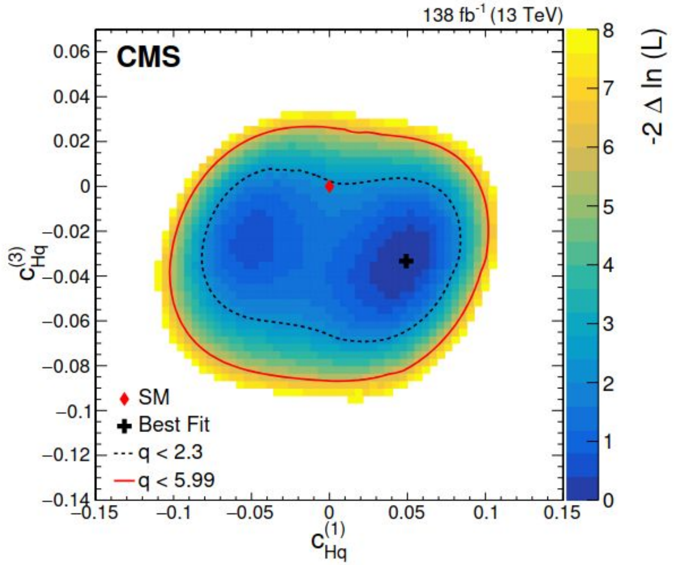


- Discriminator chosen using Bayesian optimization to minimize the volume spanned by the exclusion contours
  - Sizeable improvements in overall limits with the optimized choice

VH production - results







CP violation in EFT

SM contribution

Pure BSM contribution

SM-BSM interference

- Mostly CP-invariant in the top/Higgs sector
  - Particularly interesting from the theory standpoint
  - Odd under CP transformations

- To do: ejemplo de una distribucion CP-odd y decir por que son buenas

Phys. Rev. D 110, 096023 (2024)

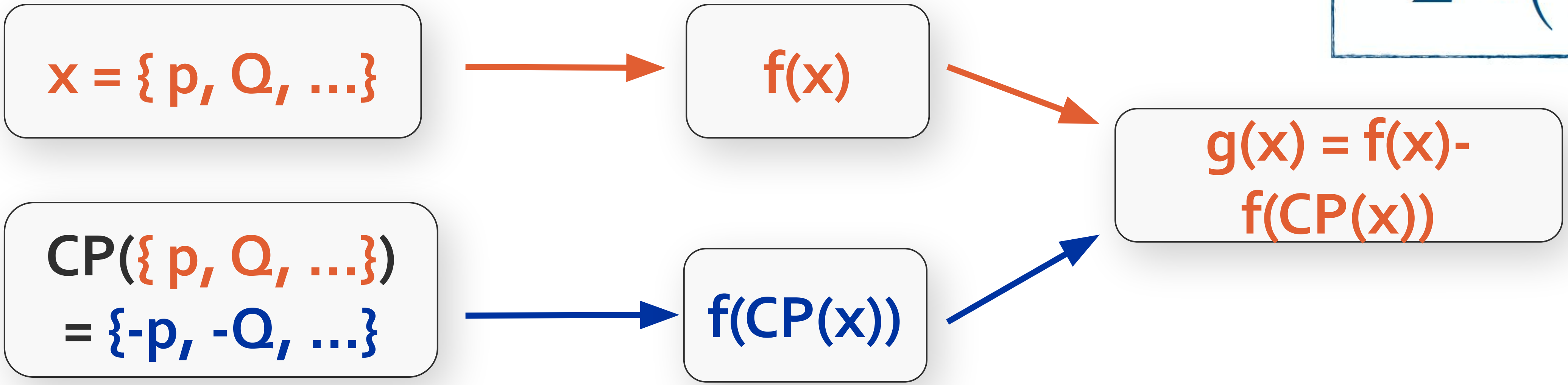
Where symmetries meet SBI…

- Using an equivariant model to learn the score

$$R(\mathbf{x}|\boldsymbol{\theta},\boldsymbol{\theta}_0) = 1 + \sum_{1\leq i\leq M} (\theta_i - \theta_0) R_i(\mathbf{x}) + \sum_{1\leq i\leq j\leq M} \frac{1}{2}(\theta_i - \theta_0)(\theta_j - \theta_0) R_{i,j}(\mathbf{x})$$

Loss function

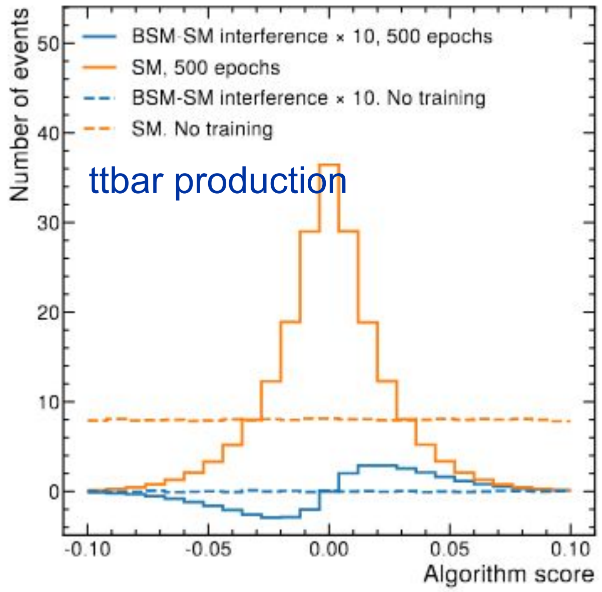
$$L = \sum w_{SM}\left(\frac{w_{linear}}{w_{SM}} - g(X)\right)^2$$

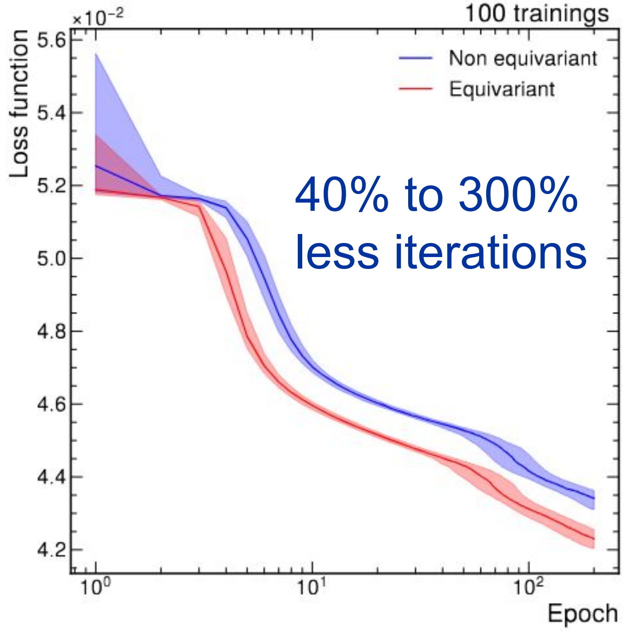


Phys. Rev. D 110, 096023 (2024)

A few pheno examples

- The resulting model is equivariant wrt CP transformations
  - g(CP(x)) = - g(x)
- Score distribution symmetric for SM
- Score distribution odd for SM-BSM interference
  - Analyses based on asymmetries
  - Systematic uncertainties heavily suppressed
  - Better numerical convergence

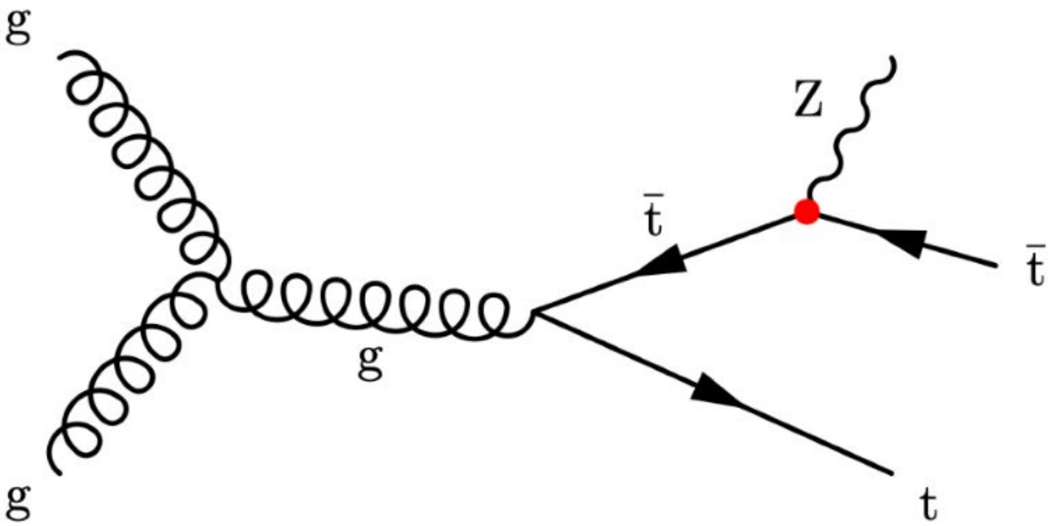


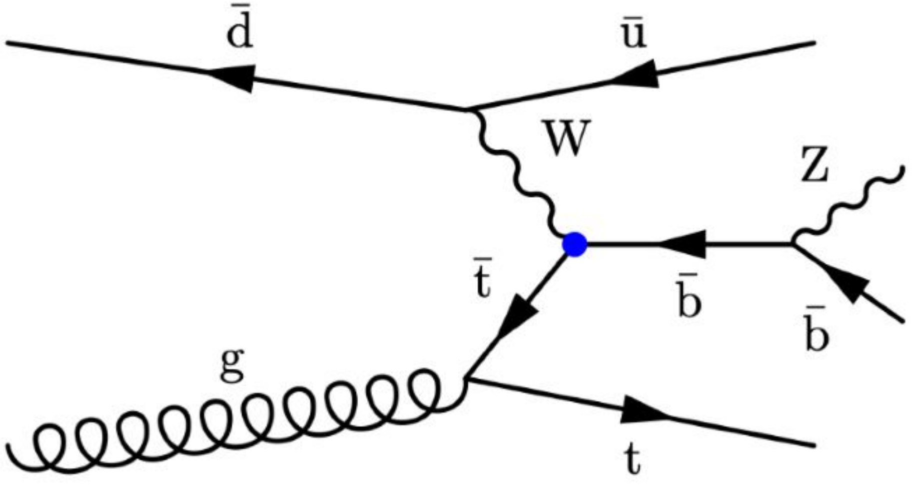


Phys. Rev. D 110, 096023 (2024)

CP violation in top+Z production

- Exploring CP-violation interactions affecting **ttZ** or **tZq production**
- Using 13 TeV and 13.6 TeV data; events with 3 leptons, jets and b-tagged jets





- Studying **two EFT operators**: c_{tW}^{I}, c_{tZ}^{I} $\quad (\bar{q}_i\sigma^{\mu\nu}\tau^I u_j)\,\tilde{\varphi}W^I_{\mu\nu} \quad (\bar{q}_i\sigma^{\mu\nu}u_j)\,\tilde{\varphi}B_{\mu\nu}$

Equivariant network

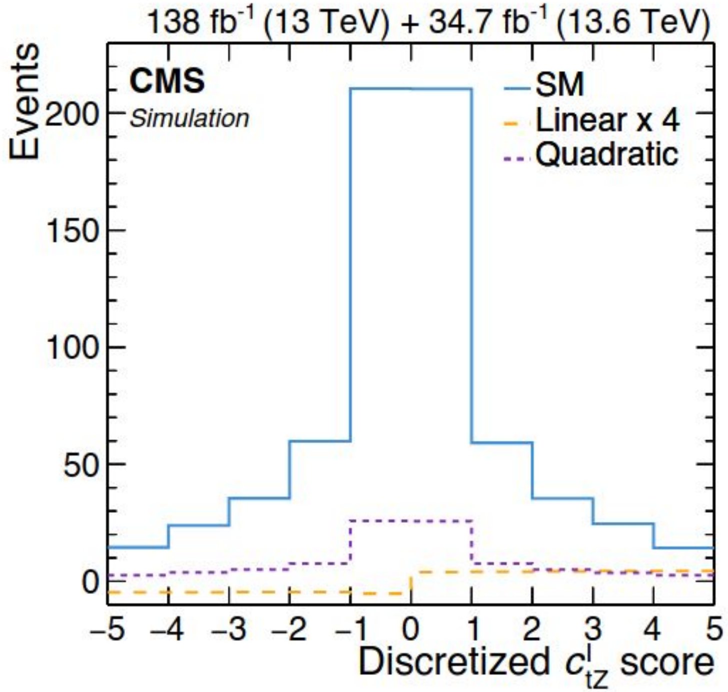
- Building a neural network per operator to learn the score
- Loss function using the SALLY method

$$L = w_{SM}(\text{parton kin.})\left(f(\text{observables}) - \frac{l_i(\text{parton kin.})}{w_{SM}(\text{parton kin.})}\right)^2$$

- Equivariance imposed on the network architecture:

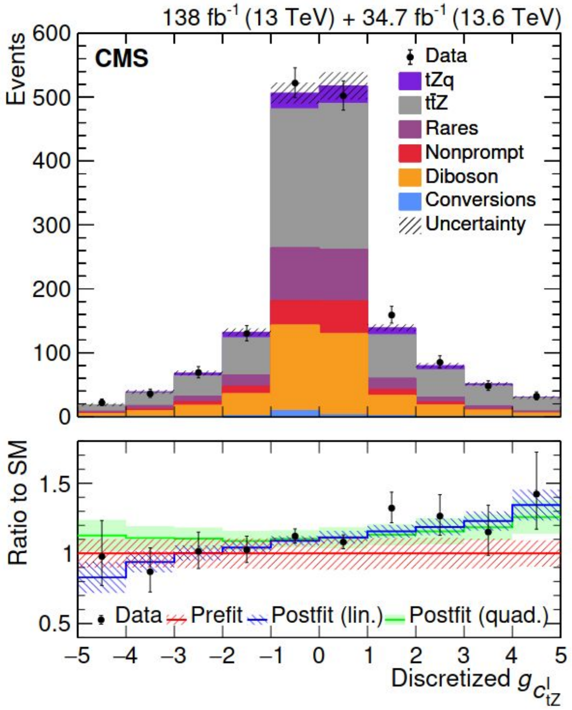
	Input variables							
x	$\vec{p}_{\ell Z+}$	$\vec{p}_{\ell Z-}$	$\vec{p}_{\ell W}$	$\vec{p}_{j_i}$	$Q_{\ell W}$	$\vec{p}_T^{miss}$	bscore_i	era
$CP(x)$	$-\vec{p}_{\ell Z-}$	$-\vec{p}_{\ell Z+}$	$-\vec{p}_{\ell W}$	$-\vec{p}_{j_i}$	$-Q_{\ell W}$	$-\vec{p}_T^{miss}$	bscore_i	era

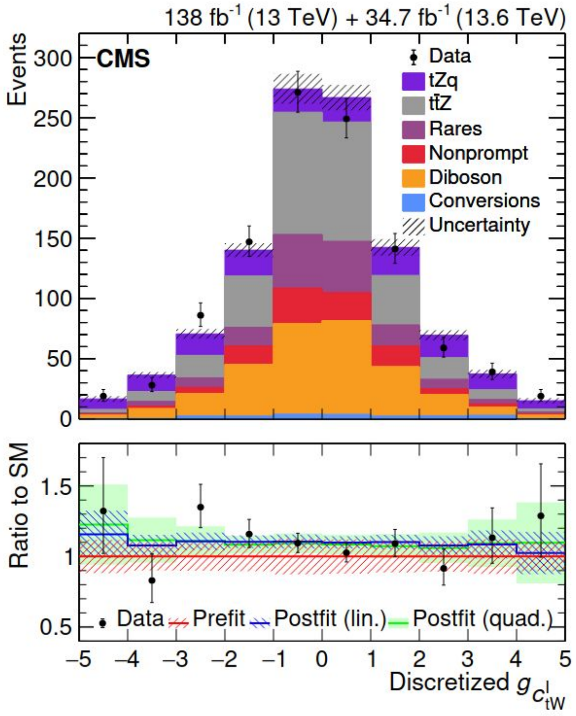
- Background included in the training, taking "SM-like" values

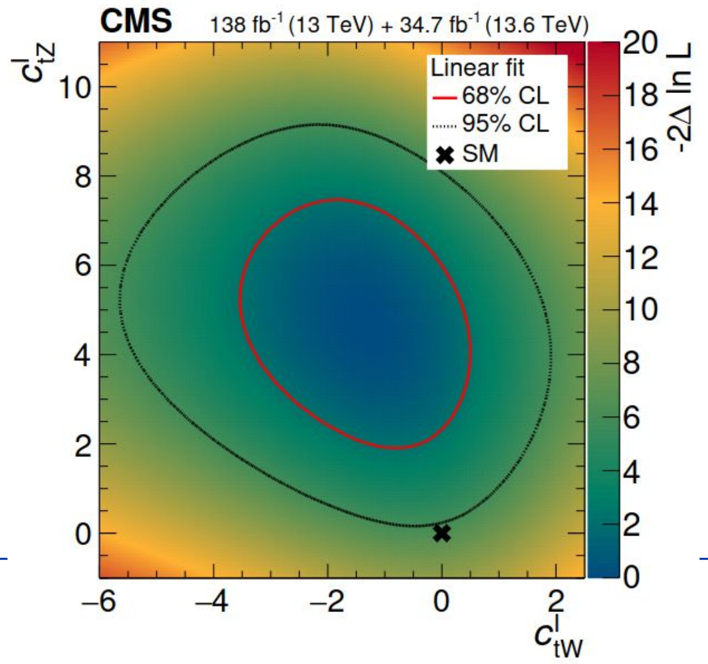


Results

- Resulting scores are
  - symmetric for SM backgrounds
  - optimal observables for the linear contribution of ctZI and ctWI
- Fairly good agreement with the SM
  - Slight (2sigma) asymmetry in ctZI
  - Discrepancy translated to the limits







Conclusions

- Shown two recent EFT analyses in CMS using SBI techniques
- Optimal observables are constructed to optimize the sensitivity
- One case incorporates CP-equivariant → more interpretable model

- Work ongoing to incorporate systematics uncertainties to the trainings to improve the sensitivity