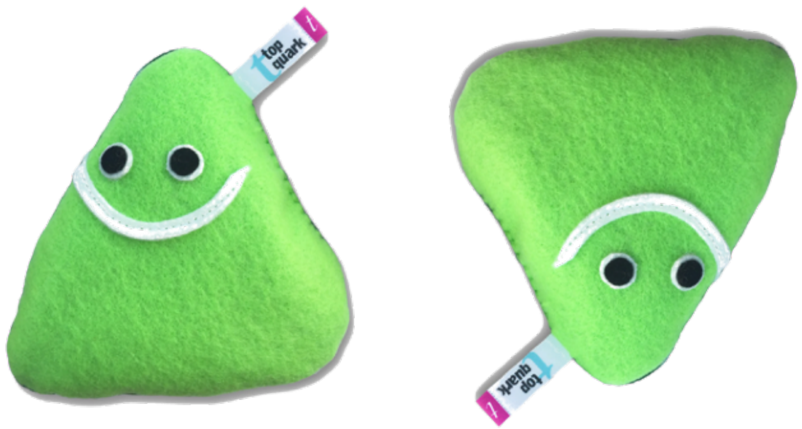


Equivariant neural networks for CP violation

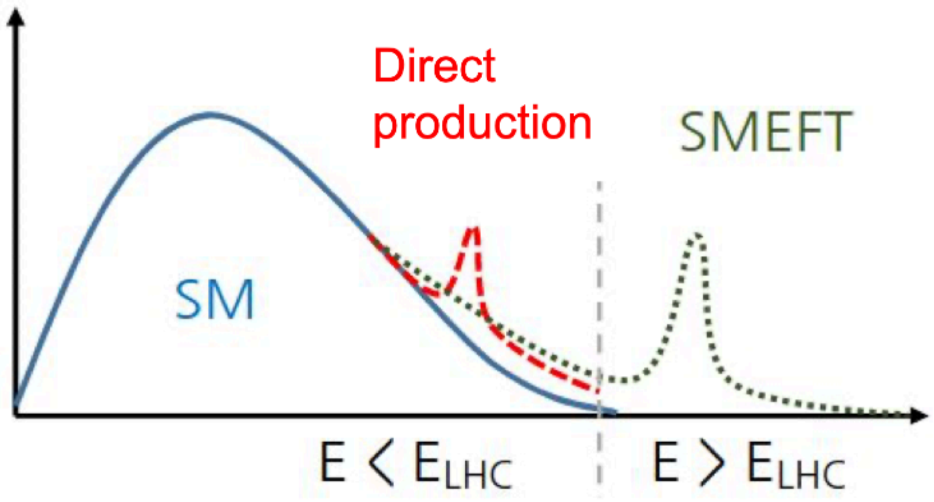
Sergio Sánchez Cruz in collaboration with Marina Kolosova, Giovanni Petrucciani, Clara Ramón, and Pietro Vischia

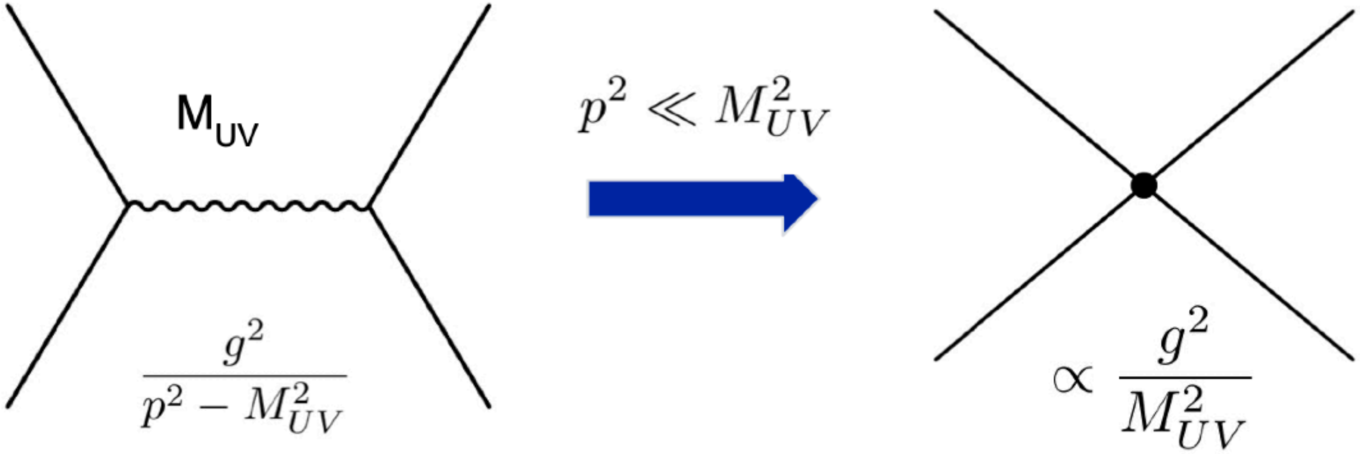
arXiv:2405.13524

10.06.2024

Introduction: SMEFT and CP-violation

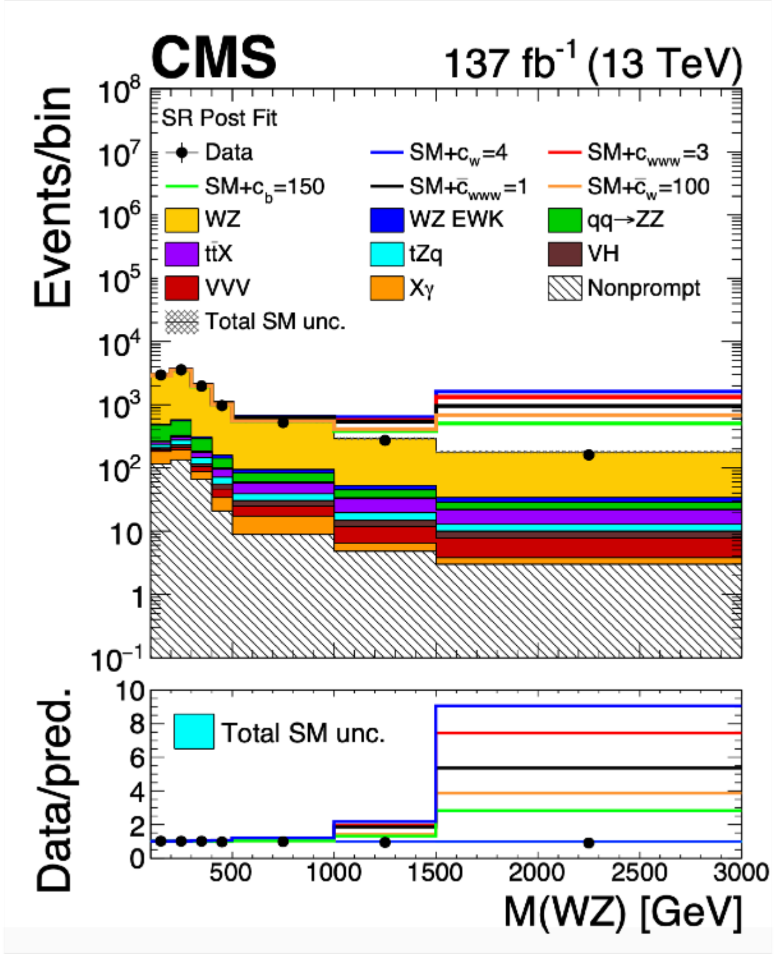
- SMEFT is an extension of the SM, adding contributions from high-mass BSM particles
- 1350 CP-even operators, 1149 CP-odd operators
- Plenty of CP-violation sources to study!





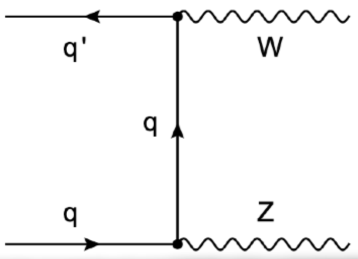
$$\mathcal{L} = \mathcal{L}_{SM} + \sum_i \frac{C_i}{\Lambda^2} O_i^{(6)}$$

EFT at observable level

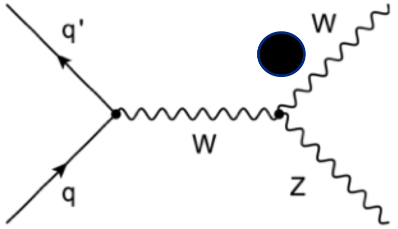


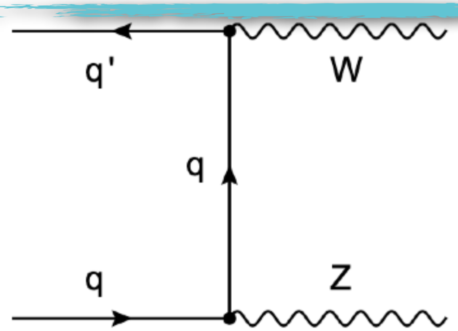
JHEP 07 (2022) 032

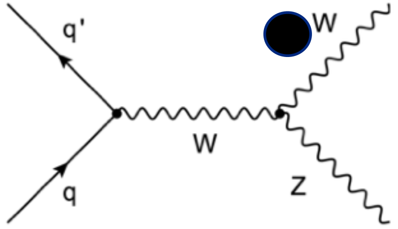
SM contribution

()^2

Pure BSM contribution

\+ ()^2

\+ 

X 

SM-BSM interference

CP-violating operators

SM contribution

Pure BSM contribution

- Mostly CP-invariant
- CP-invariant in e.g. the top/Higgs sectors

SM-BSM interference

- Particularly interesting from the phenomenological standpoint
- Odd under CP transformations

- CP-even observables (most of the LHC cross section measurement program) do not provide sensitivity to the interference
- CP-odd observables are robust against signal mismodeling / backgrounds

The algorithm

- The algorithm builds observables that are **equivariant with respect to the CP-symmetry**
  - **CP-invariant observables** are useful to discriminate among backgrounds for searches targeting the pure-BSM part
  - **CP-odd observables** are useful to get sensitivity to the interference term
  - Can be generalized to n_1 CP-invariant and n_2 CP-odd components
- A function f: D —> R is odd/even under CP transformation if f(CP(event)) = +/- f(event)
  - The function f(event) = g(event) +/- g(CP(event)) trivially satisfies that

- **The space of input features is fully general**
  - Can be the kinematics of a fixed set of particles or a particle set
- We take g to be a fully-connected neural network, could be any function

Training and cost function

- Method inspired in the SALLY method shown in 2401.10323
  - Equivariant networks can also be used with different cost functions

- Training the algorithm on weighted simulations $w(z) = w_{SM}(z) + c w_{int}(z) + c^2 w_{quad}(z)$
  - Function of parton-level kinematics
  - Can be used to compute the (non tractable) likelihood ratio

$$\frac{p(d, z|c_1)}{p(d, z|c = 0)} = \frac{w_{SM} + c w_{lin} + c^2 w_{quad}}{w_{SM}}$$

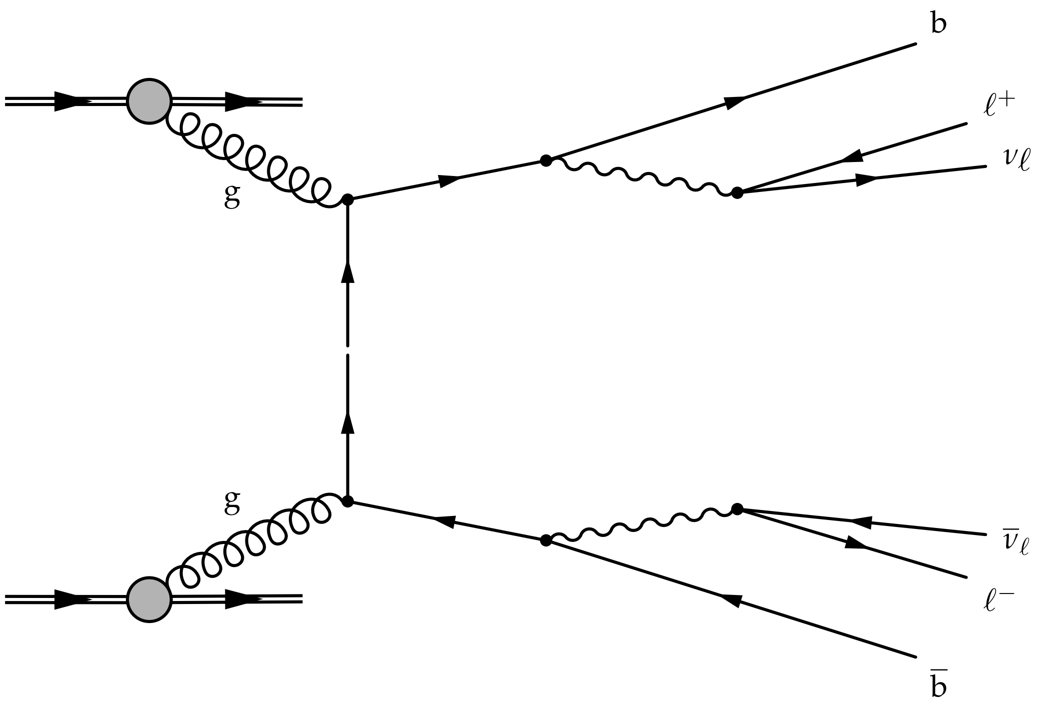
- We are interested in the likelihood score at the SM —> sufficient statistic for small values of c
  - Small values of c —> dominated by the interference
- Minimizing the loss function, we obtain a surrogate model of the score

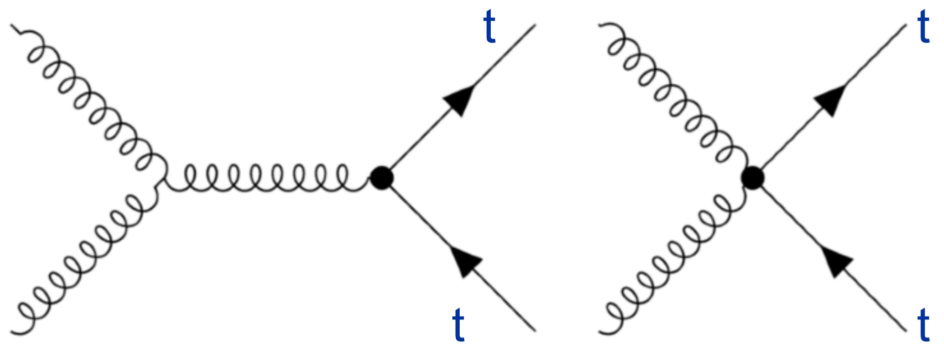
$$L = w_{SM} \left(f(d) - \frac{w_{int}(z)}{w_{SM}(d)} \right)^2$$

Use case: ttbar production

- Targeting ttbar production in the dileptonic final state
- Potentially affected by the chromoelectric dipole moment operator (CP violating)

$$g_s \frac{v}{\sqrt{2}} (\bar{t}\sigma^{\mu\nu}\gamma_5 T^A t) G^A_{\mu\nu}$$



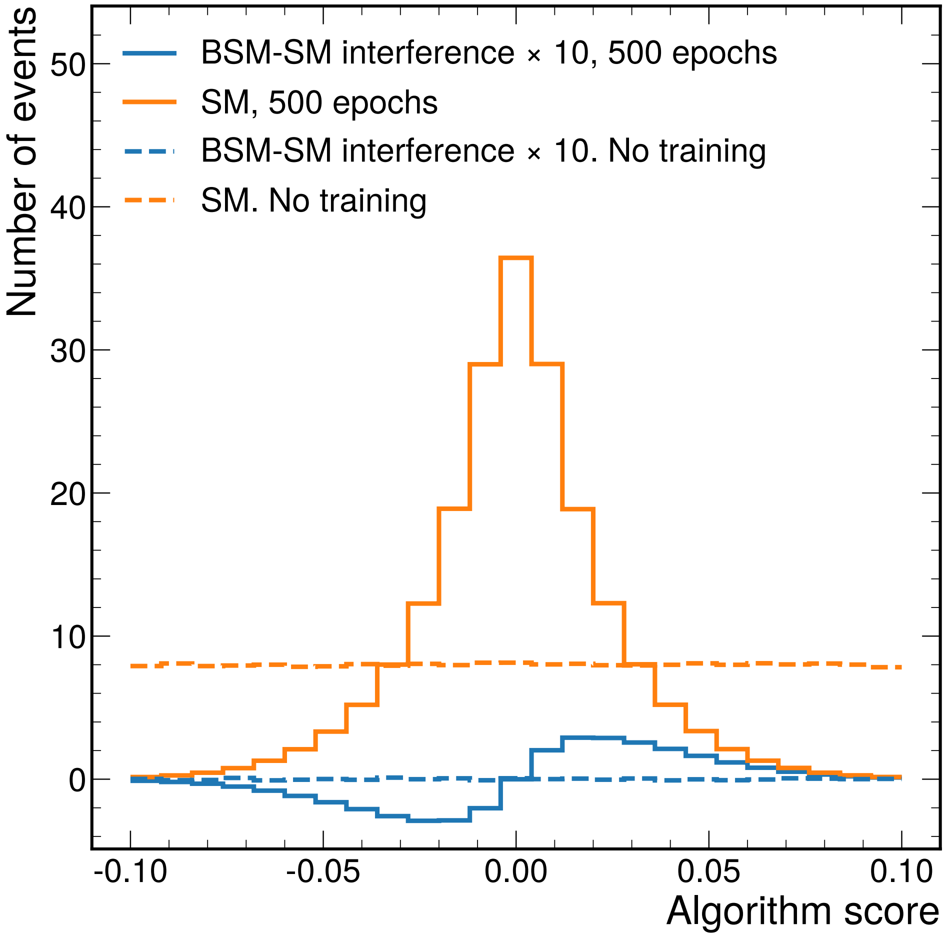


$$\overrightarrow{p_\ell^+}, \overrightarrow{p_\ell^-}, \overrightarrow{p_{j_1}}, \overrightarrow{p_{j_2}}, \overrightarrow{p_\mathrm{T}^\mathrm{miss}} \xrightarrow{\mathrm{CP}} -\overrightarrow{p_\ell^-}, -\overrightarrow{p_\ell^+}, -\overrightarrow{p_{j_2}}, -\overrightarrow{p_{j_1}}, -\overrightarrow{p_\mathrm{T}^\mathrm{miss}}$$

Use case: ttbar production

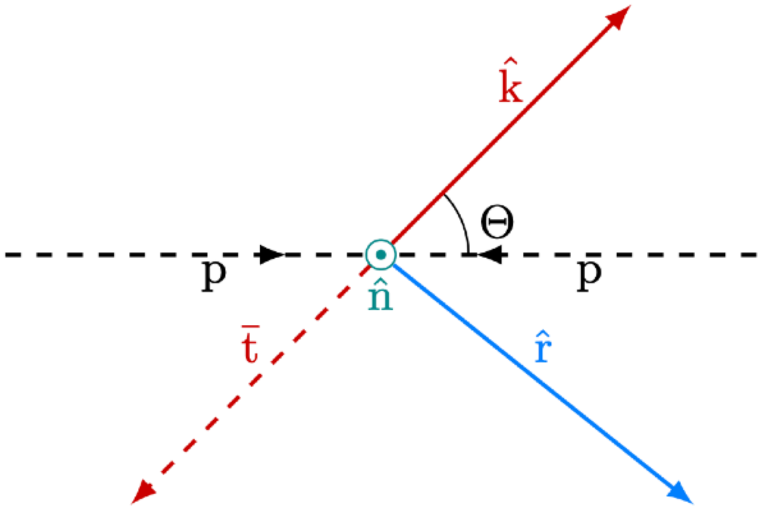
- Score after the training is a CP-odd observable
  - Symmetric for the SM contribution
  - **Any SM-like mismodeling / background will be symmetric by construction**
  - Interference contributes constructively for positive values and negatively for negative values

- Equivariance respected even during (or before) training
- **Observable is robust even if the training has not converged**



Use case: ttbar production

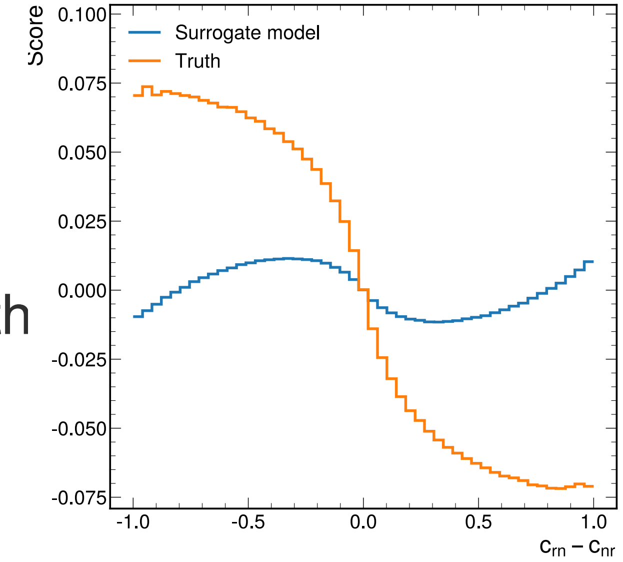
- The algorithm provides a surrogate model of the score
  - Comparing against the true model (from parton level quantities)
- 1508.05271 proposes two observables, relying on the reconstruction of the ttbar system, based on angles between leptons and axes

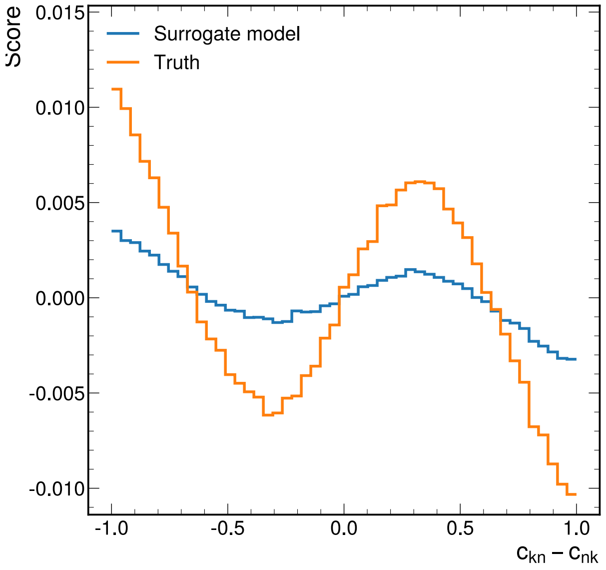


c_{rn}-c_{rn} = $\cos(l^{+}{}_{r})\cos(l^{-}{}_{n})-\cos(l^{+}{}_{n})\cos(l^{-}{}_{r})$

c_{kn}-c_{nk} = $\cos(l^{+}{}_{k})\cos(l^{-}{}_{n})-\cos(l^{+}{}_{n})\cos(l^{-}{}_{k})$

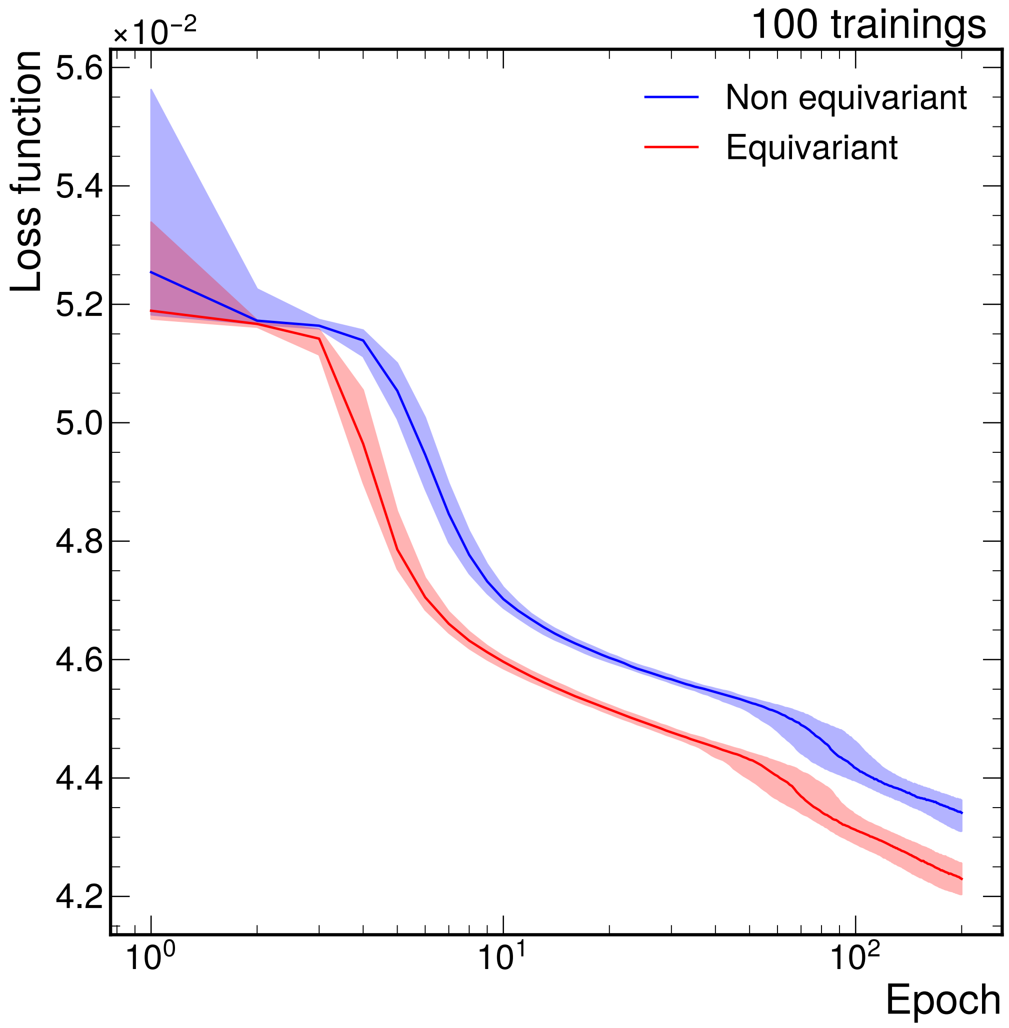
- Partially learning them
  - Limited by the possibility of reconstructing th system





Use case: ttbar production (II)

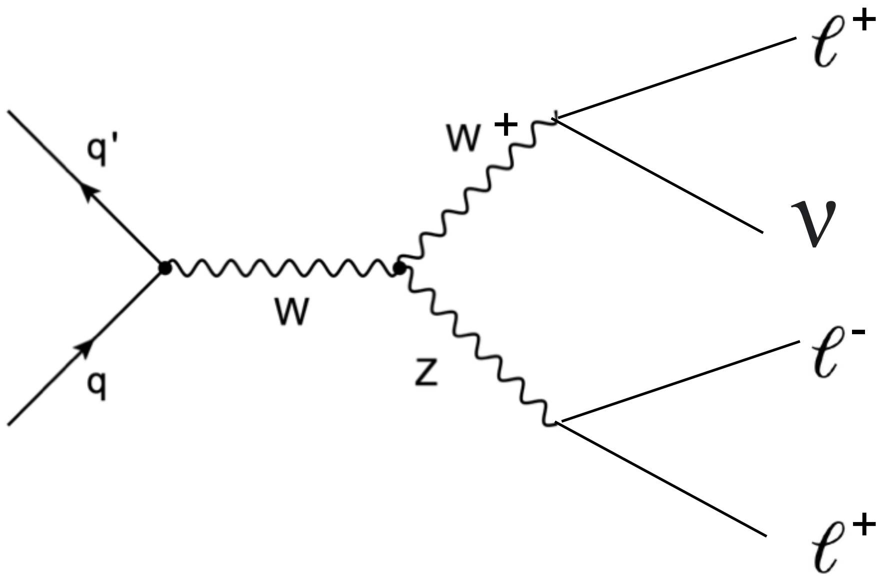
- Imposing equivariance as inductive bias improves the convergence of the model
- Training 100 instances of equivariant and non-equivariant model
- Smaller variance in the first steps of the training
- Overall, between 40 and 300% less iterations needed to achieve the same loss function

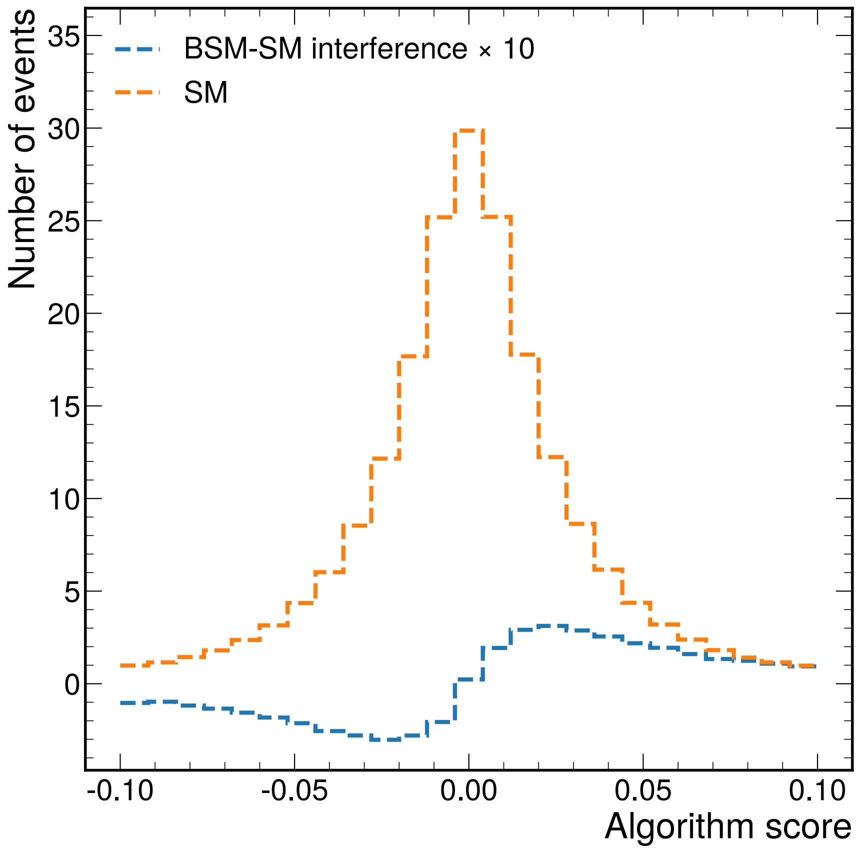


Use case: WZ production

- Targeting WZ production in the three lepton final state
- Potentially affected by the cWtilde operator

$$\epsilon^{ijk}\widetilde{W}_{\mu}^{i\nu}W_{\nu}^{j\rho}W_{\rho}^{k\mu}$$





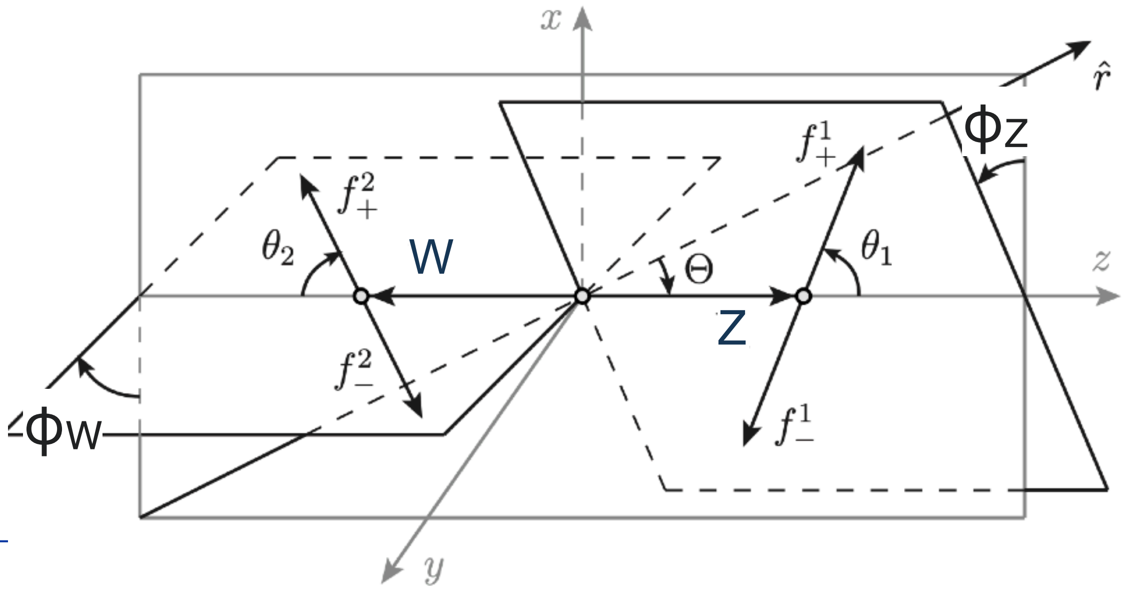
$$\overrightarrow{p^{Z}_{\ell+}}, \overrightarrow{p^{Z}_{\ell-}}, \overrightarrow{p^{W}_{\ell}}, Q^{W}, \overrightarrow{p^{miss}_{T}}$$

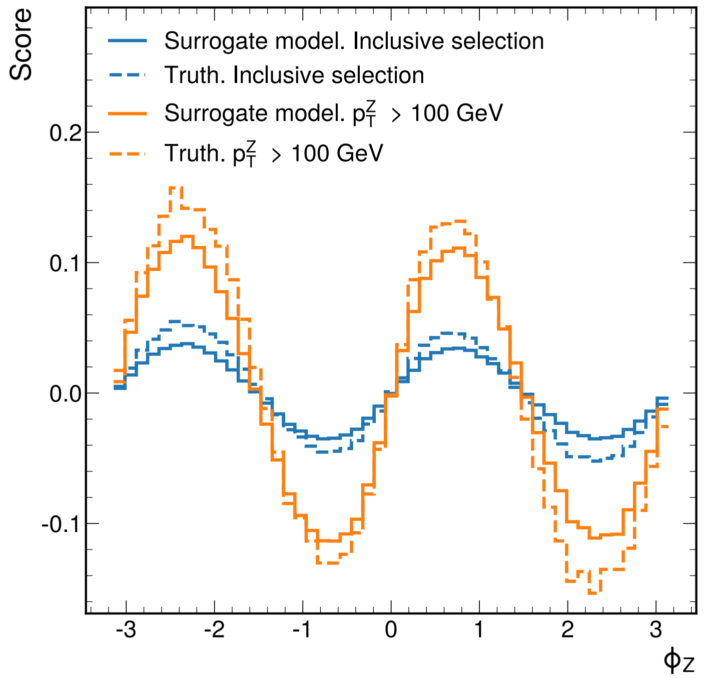
CP

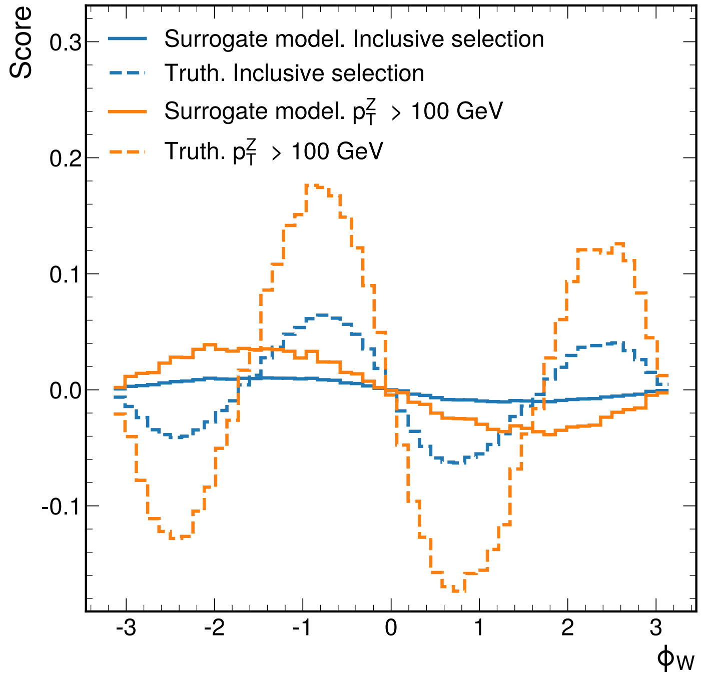
$$-\overrightarrow{p^{Z}_{\ell-}}, -\overrightarrow{p^{Z}_{\ell+}}, -\overrightarrow{p^{W}_{\ell}}, -Q^{W}, -\overrightarrow{p^{miss}_{T}}$$

Use case: WZ production (II)

- Checking if the model has learned dedicated observables based on spin correlation
- Modulation introduced in ϕ_Z is well captured
- Mostly insensitive to ϕ_W due to ambiguity in the W decay reconstruction
- **More sensitive than dedicated observables, can capture energy growth**





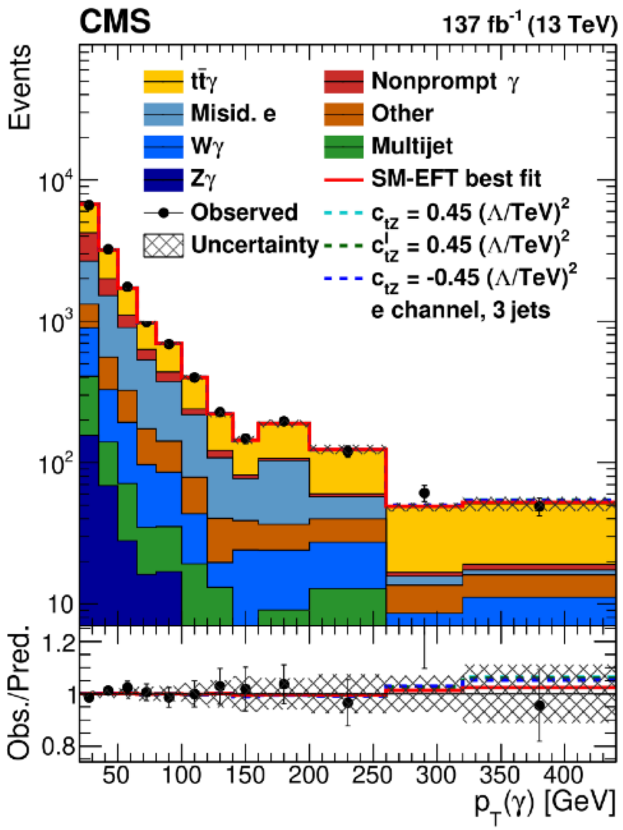


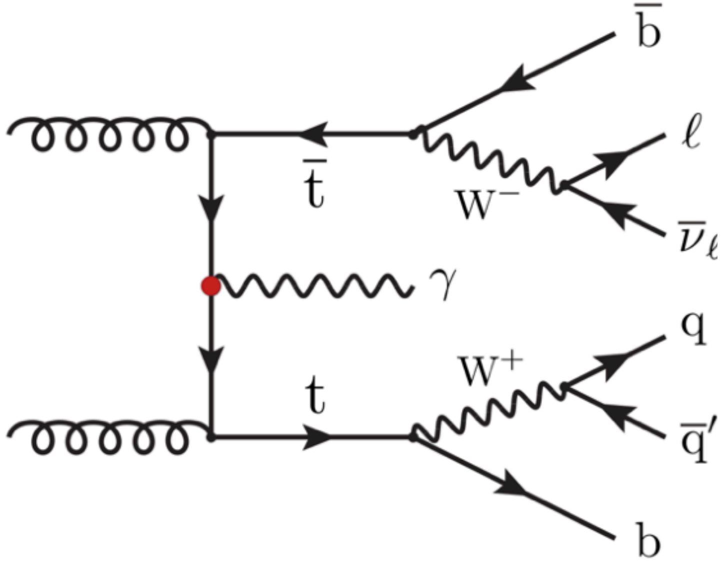
arXiv:1708.07823

Use case: ttγ production

- Checking ttγ in the single lepton channel, affecting c_{tZ}^{l}
- Operator related to $(\bar{Q}\sigma^{\mu\nu}t)\tilde{H}B_{\mu\nu}$

- Often looked for using the pT of the photon
  - Not a CP-violating observable





$$\overrightarrow{p_\gamma}, \overrightarrow{p_\ell}, Q_\ell, \overrightarrow{p_{b_1}}, \overrightarrow{p_{b_2}}, \overrightarrow{p_{j_1}}, \overrightarrow{p_{j_2}}$$

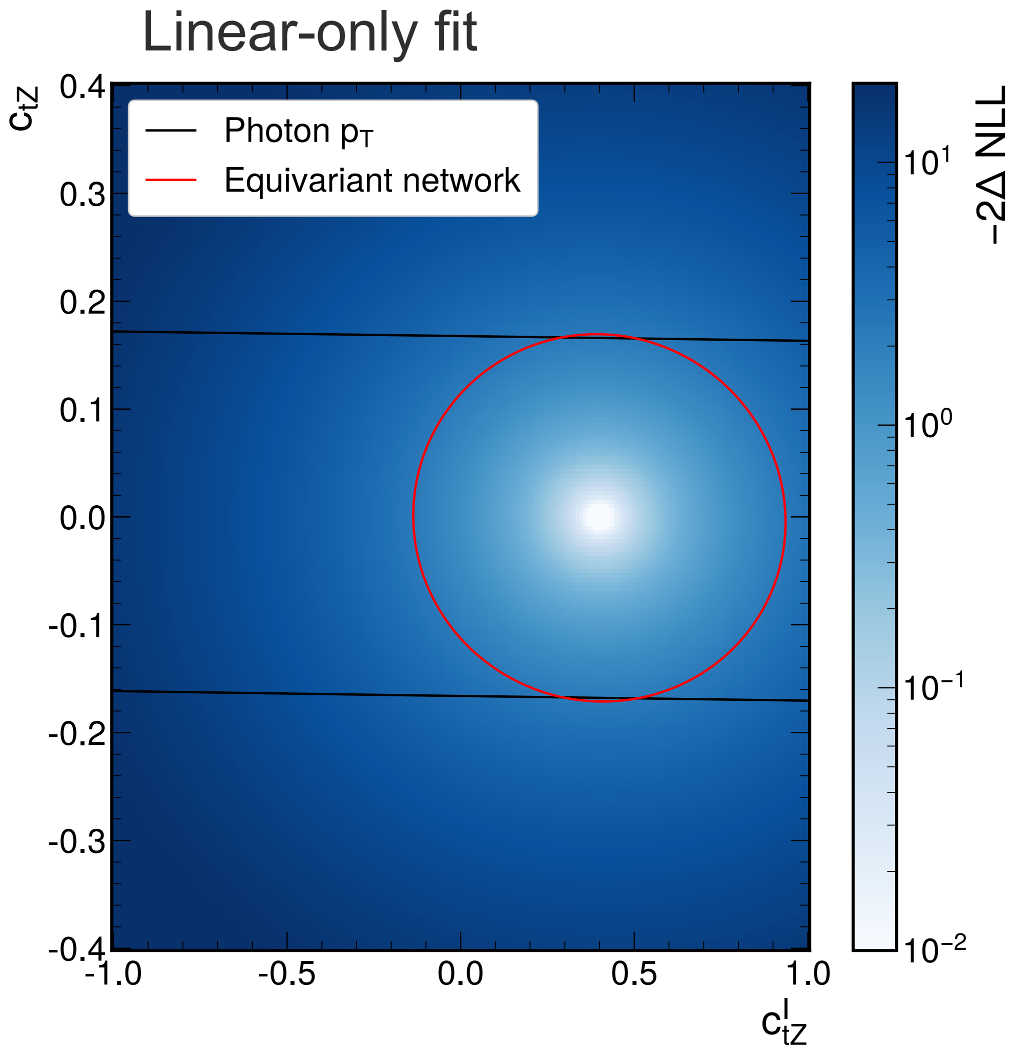
CP

$$-\overrightarrow{p_\gamma}, -\overrightarrow{p_\ell}, -Q_\ell, -\overrightarrow{p_{b_2}}, -\overrightarrow{p_{b_1}}, -\overrightarrow{p_{j_2}}, -\overrightarrow{p_{j_1}}$$

Use case: $t\bar{t}\gamma$ production

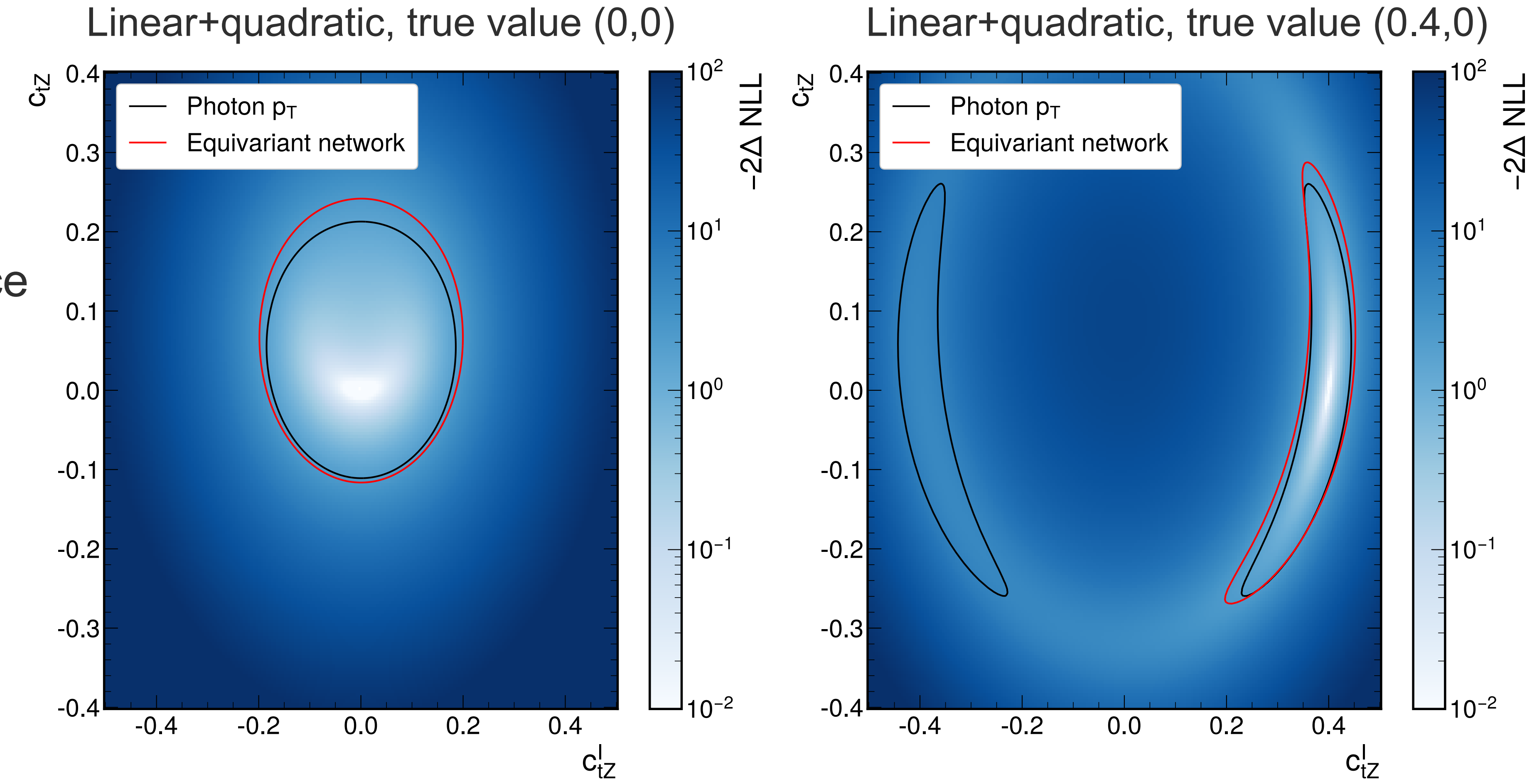
- Comparing counting analyses binning on the score of the equivariant network and photon p_T
- Setting expected limits on c_{tZ}^{I} and c_{tZ}
- Using Poissonian likelihood
  - **When considering CP-odd observables systematic uncertainties are suppressed**

- When considering linear contributions only, the **equivariant network provides sensitivity to c_{tZ}^{I}**
  - **Photon p_T is completely blind to this operator**
  - Both observables give similar sensitivity to c_{tZ}



Use case: $t\bar{t}\gamma$ production

- Assuming the SM, both observables give similar sensitivity
  - Equivariant network targets the interference
- In BSM cases, the equivariant network can disentangle positive and negative c_{tZ}^{I} values



- **Overall, the equivariant network is more powerful, even if it has not been trained to capture quadratic effects**

Conclusions

- Showcased the properties of equivariant neural networks to search for CP-violation
- Using equivariance as an inductive bias, we obtain robust CP-odd observables
  - Robust observables, regardless of the convergence of the training
  - Better numerical convergence properties than non-equivariant algorithms
- Produced optimal CP-odd observables for ttbar, WZ and ttγ production
  - Observables developed in WZ and ttγ improve the existing state-of-the-art observables
  - Highly relevant for analysis in the top, Higgs and electroweak sectors targetting CP violation, potential for improving any such analysis

- Possible trivial extensions
  - Many of the physics we look at are CP-even —> improving convergence by considering CP-invariant networks?