Probing EFT operators in top-Z associated production with machine learning.

Nicolas Tonon, Abideh Jafari, David Walter (DESY)

DESY – LHC Physics Discussion

20th September 2021

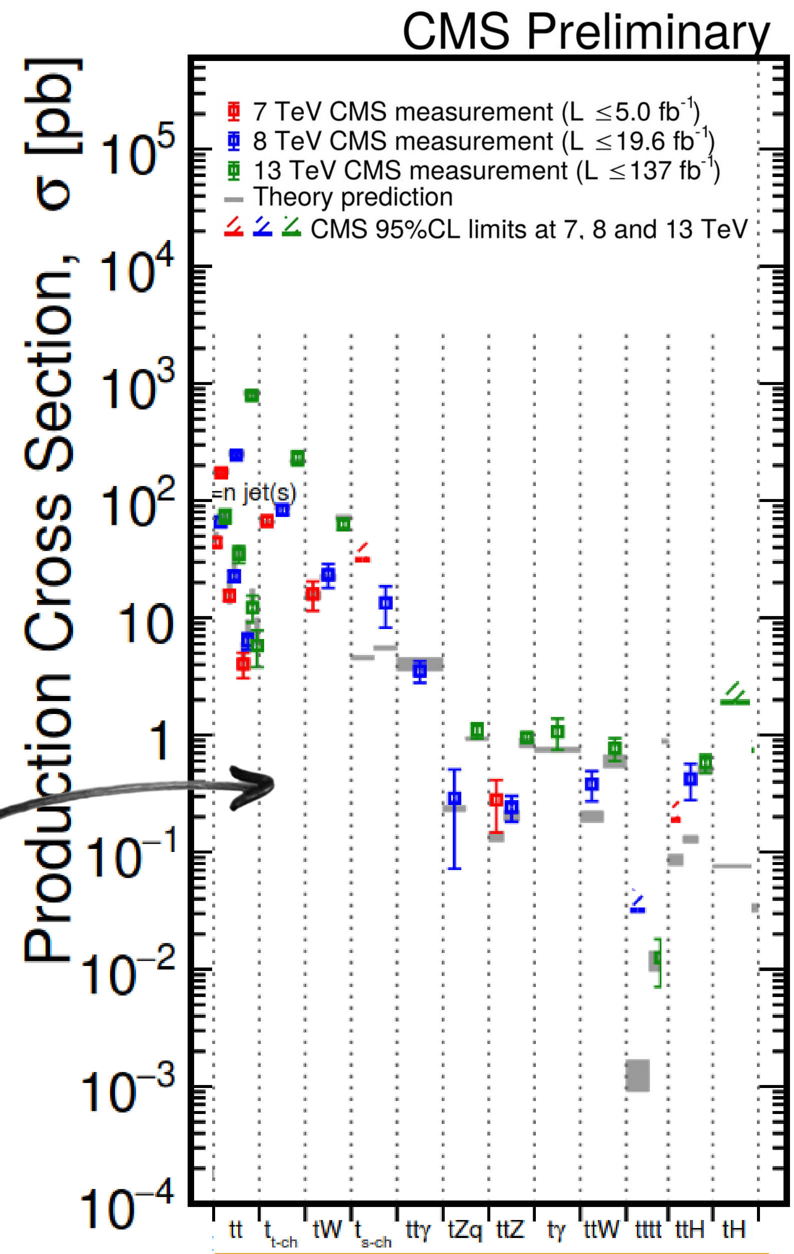
HELMHOLTZ RESEARCH FOR GRAND CHALLENGES

INTRODUCTION

- **No clear sign of new physics** (BSM) at LHC so far…
  - … and future accelerators in coming decade(s) will increase ∫L , not √s

- Many BSM theories predict sizeable **deviations of top quark's couplings** w.r.t. SM predictions

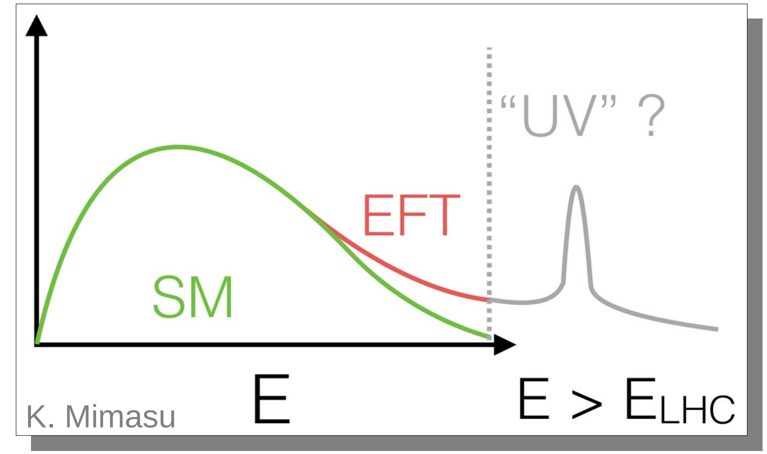
- Most of canonical top quark processes at LHC have reached **precision era** (systematics-limited)



Adapted from : http://cern.ch/go/pNj7

Motivates ambitious top physics programme to **reveal new physics** indirectly **through precision** measurements

EFFECTIVE FIELD THEORY



- **Effective Field Theory** (EFT) framework allows for systematic & model-independent interpretation of potential deviations in interactions between SM fields

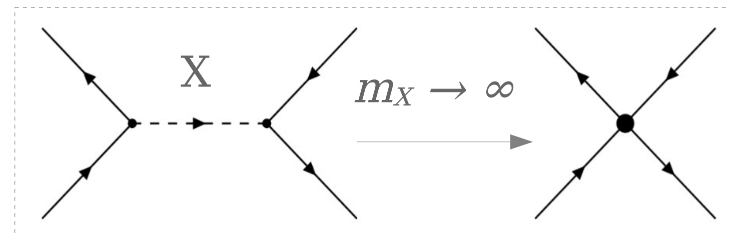
- Expand SM Lagrangian with higher-order operators :

Wilson coefficients ↔ interaction strengths

$$\mathcal{L}_{\mathrm{EFT}} = \mathcal{L}_{\mathrm{SM}} + \sum_{d,i} \frac{c_i^d}{\Lambda^{d-4}} \mathcal{O}_i^d$$

Higher-order operators

BSM energy scale ($\gg E_{LHC}$)



EFT example : Fermi theory of β-decay (W boson → 4-fermions interaction)

- BSM effects described by **finite number of EFT operators,** whose strengths are encoded into **Wilson coefficients** (WCs) ↔ **theory parameters**
- Predicts well-defined deviation **patterns**, **correlated** in different observables/processes

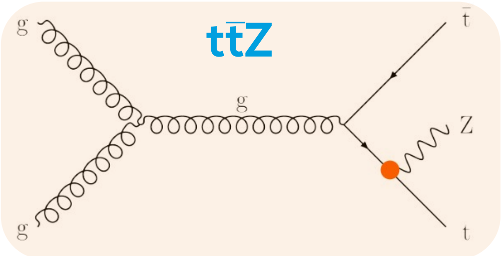
→ Well-motivated, global approach to maximize discovery potential of massive BSM states at (HL-)LHC

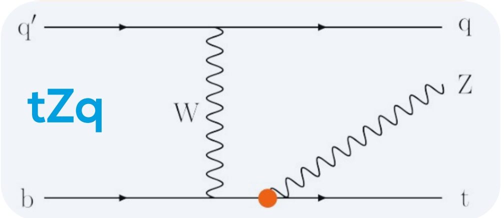
ANALYSIS OVERVIEW

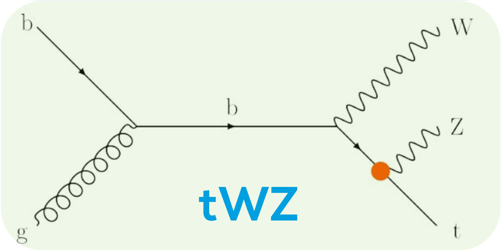
- *Existing EFT analyses* :
  - Reinterpretations (sub-optimal, assumptions)
  - Consider 1 single operator & signal process
  - 'Cut-and-count'
  - ...

- Target 3 top-Z associated production modes
  - Complementary, probe similar EFT operators
- Consider up to 5 operators simultaneously
- Analysis strategy entirely optimized to search for EFT effects
  - Simulated at detector-level (→ Direct measurement ≠ reinterpretation)

New Use novel machine-learning techniques to improve sensitivity to WCs







Run2 data (138 fb^{-1})

3ℓ & 4ℓ channels

Reference :

- *Probing effective field theory operators in the associated production of top quarks with a Z boson in multilepton final states at $\sqrt{s}$=13 TeV*, arXiv:2107.13896 (Submitted to JHEP)
  - *Learning to detect new top-quark interactions*, CERN Courier (Sep/Oct 2021)

SM+EFT SIMULATION

- At leading dimension-6, $\mathcal{O}$(60) non-redundant operators involving top quark !

- Focus on subset of five operators that…
  - Involve top quark and gauge bosons
  - Interfere with SM production of $t\bar{t}Z$ or tZq (~ Λ^{-2})
  - Conserve CP, lepton- and baryon-number
  - Ignore top chromomagnetic dipole moment,
    (probed with much better sensitivity in $t\bar{t}$)

$\mathcal{O}_{tZ}$, $\mathcal{O}_{tW}$, $\mathcal{O}^{-}_{\phi Q}$, $\mathcal{O}^{3}_{\phi Q}$ and $\mathcal{O}_{\phi t}$

parameter	$t\bar{t}$	single t	tW	tZ	t decay	$t\bar{t}Z$	$t\bar{t}W$
$C^{-}_{\phi Q}$	–	–	–	Λ^{-2}	–	Λ^{-2}	–
$C^{3}_{\phi Q}$	–	Λ^{-2}	Λ^{-2}	Λ^{-2}	Λ^{-2}	–	–
$C_{\phi t}$	–	–	–	Λ^{-2}	–	Λ^{-2}	–
$C_{\phi tb}$	–	Λ^{-4}	Λ^{-4}	Λ^{-4}	Λ^{-4}	–	–
C_{tZ}	–	–	–	Λ^{-2}	–	Λ^{-2}	–
C_{tW}	–	Λ^{-2}	Λ^{-2}	Λ^{-2}	Λ^{-2}	–	–
C_{bW}	–	Λ^{-4}	Λ^{-4}	Λ^{-4}	Λ^{-4}	–	–
C_{tG}	Λ^{-2}	$[\Lambda^{-2}]$	Λ^{-2}	–	$[\Lambda^{-2}]$	Λ^{-2}	Λ^{-2}

Brivio et al., arXiv:1910.03606

- SM+EFT signal samples simulated at leading-order with Madgraph
  - EFT effects included with *dim6top* model arXiv:1802.07237
  - Full detector simulation ↔ will constrain operators directly at detector-level
  - Validated against official CMS samples (under SM scenario)

EFT REWEIGHTING

- Cross sections & kinematics parameterized with Wilson coefficients

$$\mathcal{M} = \mathcal{M}_{SM} + \mathcal{M}_{EFT} = \mathcal{M}_{SM} + \sum_i c_i \cdot \mathcal{M}_i$$

$$\sigma \propto |\mathcal{M}^2|$$

Event weight

Coefficients to determine per event

$$w\left(\frac{\vec{c}}{\Lambda^2}\right) = \underbrace{s_0}_{\text{SM}} + \underbrace{\sum_j s_{1,j}\frac{c_j}{\Lambda^2}}_{\text{SM-EFT interference}} + \underbrace{\sum_j s_{2,j}\frac{c_j^2}{\Lambda^4}}_{\text{Pure-EFT (quadratic)}} + \underbrace{\sum_{j,k} s_{3,jk}\frac{c_j}{\Lambda^2}\frac{c_k}{\Lambda^2}}_{\text{EFT-EFT interference}}$$

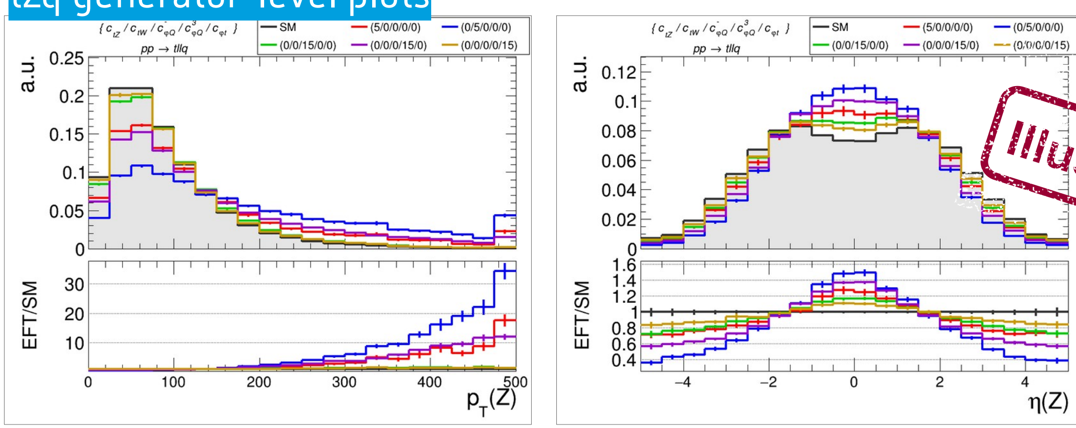
WC

1) **Reweight generator-level events** according to many different EFT scenarios

2) Perform **per-event quadratic fit** → Extract dependence of each event weight on each WC

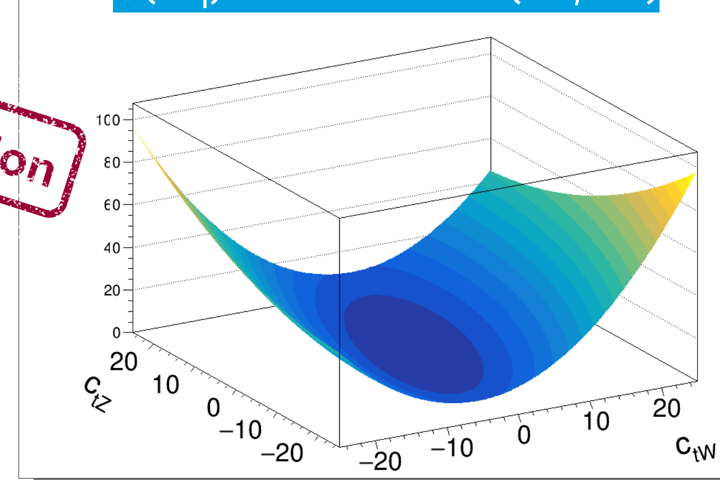
3) Parameterization → **Reweight any distribution according to any EFT scenario**

tZq generator-level plots



(C_{tW} = C_{tZ} = 5, and $C^-_{\phi Q}$ = $C^3_{\phi Q}$ = $C_{\phi t}$ = 15)

σ(tZq) as a function of (ctZ,ctW)



OBJECTS & BACKGROUNDS

Objects

- Leptons ($\ell = e, \mu$) :
  - Prompt lepton ID with dedicated BDT
- Jets :
  - AK4 (PF CHS) jets, $p_T > 25$ GeV, $|\eta| < 5$
- b jets :
  - DNN tagger (ε~75 % / mistag ~1 %)
  - $|\eta| < 2.4$ in 2016 (2.5 in 2017/18)

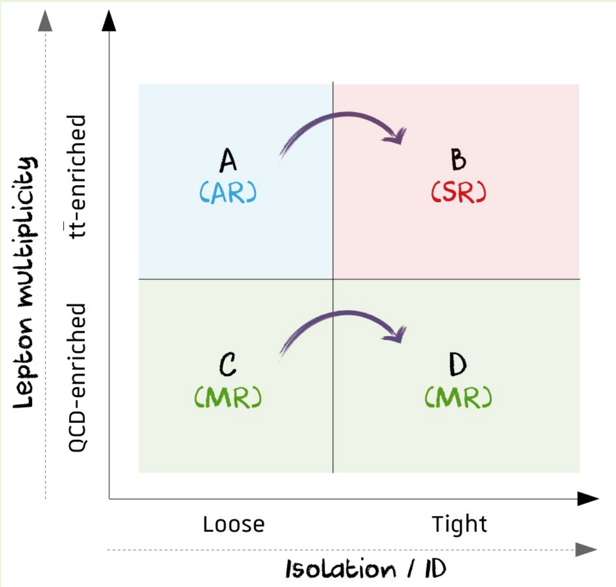
Event reco.

- Z boson candidate ↔ OSSF lepton pair within 15 GeV of Z peak (closest to m_Z)
- Leptonic top quark candidate (in 3ℓ) ↔ (closest to m_t) { 3^{rd} lepton + neutrino solution (from MET) + b-jet }

Baseline selection

- ≥ 3 isolated leptons with $p_T > 25 / 15 / 10$ GeV
- ≥ 1 Z boson candidate

- Prompt backgrounds estimated from simulation, validated in control regions
  - Grouped as : WZ / VV(V) / $t(\bar{t})X$ / $X\gamma$
- Nonprompt lepton backgrounds estimated from data with 'loose-to-tight method'
- Mis-identification probabilities binned in lepton (p_T, η)
  - QCD-enriched measurement region [C,D]
  - Dedicated application region [A] for each SR/CR
  - MC closure tests, good data/MC agreement

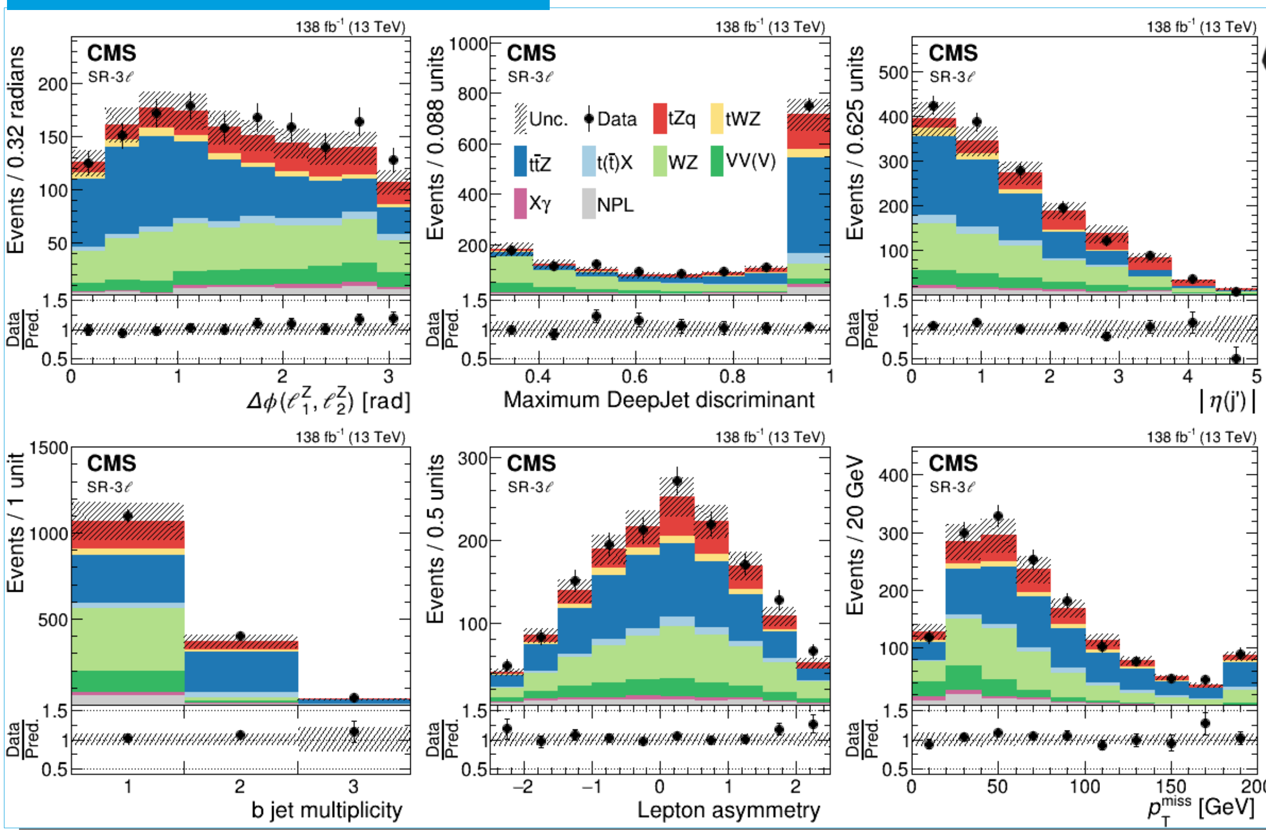


EVENT SELECTION

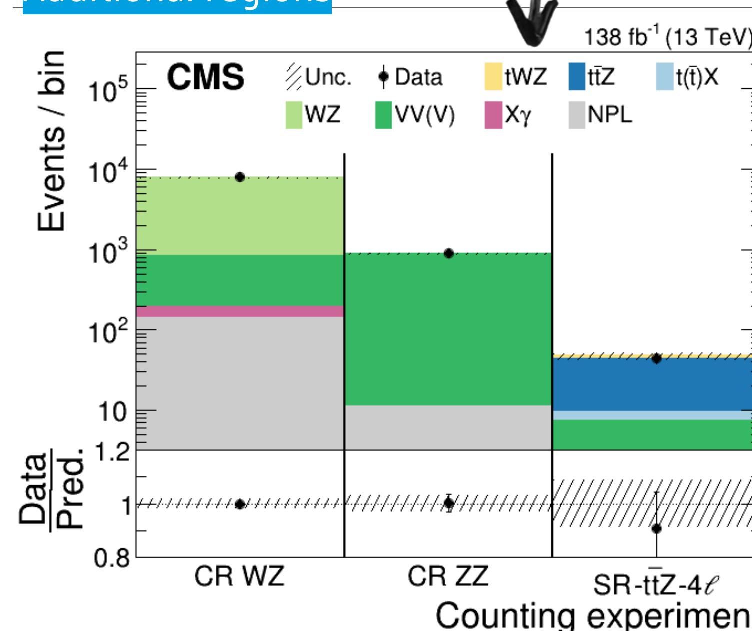
- Define 4 regions enriched in signals & main backgrounds

Selection requirement	SR-3ℓ	SR-$t\bar{t}$Z-4ℓ	WZ CR	ZZ CR
Lepton multiplicity	=3	=4	=3	=4
$m_{3\ell} - m_Z$	—	—	>15 GeV	—
Z boson candidates multiplicity	=1	=1	=1	=2
Jet multiplicity	≥2	≥2	—	—
b jet multiplicity	≥1	≥1	=0	—
p_T^{miss}	—	—	>50 GeV	—

Kinematic distributions in SR-3ℓ



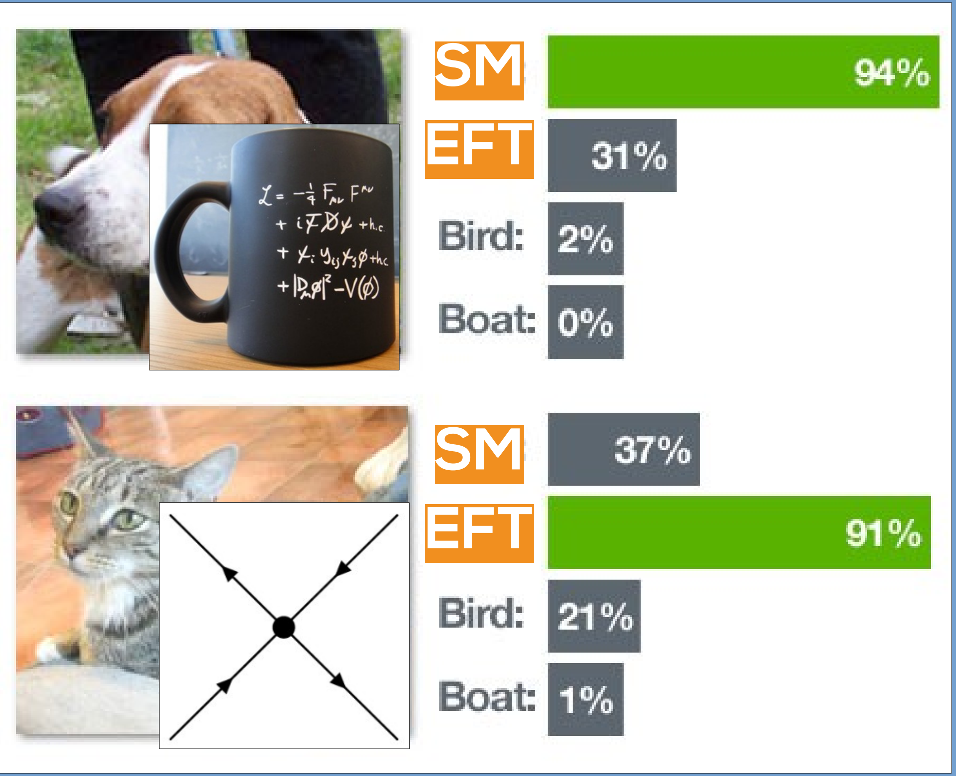
Additional regions



- SR-3ℓ drives the analysis' sensitivity
- Will be sub-divided based on MVA discriminant

- Additional regions included in fit
  → Control backgrounds, added sensitivity

Multivariate Analysis with Machine Learning



HOW CAN ML HELP ?

- *EFT operators usually …*
  - impact both cross section & kinematics
  - introduce new coupling structures leading to **subtle kinematics modifications**
  - **correlate** deviations in many different processes/observables (expected patterns)

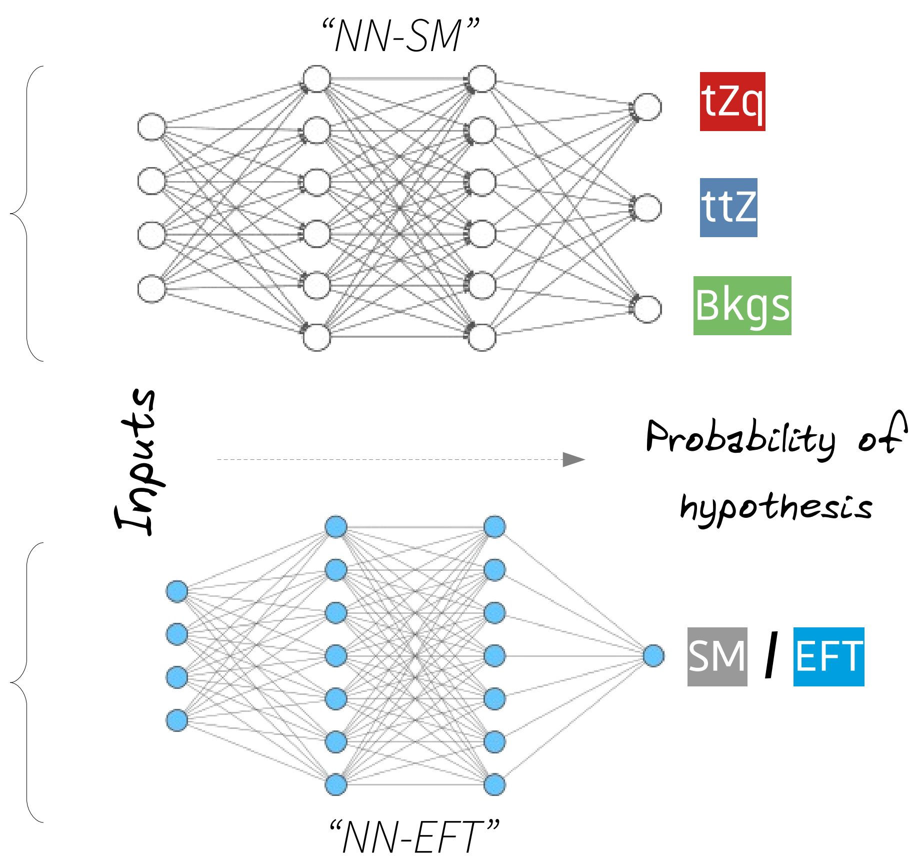
    → No single observable can constrain full parameter space

⇒ Perfect match for machine-learning (ML) techniques !

1) Separate physics processes
↔ Enrich SR with processes most sensitive to our operators

- Train neural networks (NN) in SR-3ℓ to :

2) Separate SM / EFT scenarios
↔ Learn characteristic features of new physics events

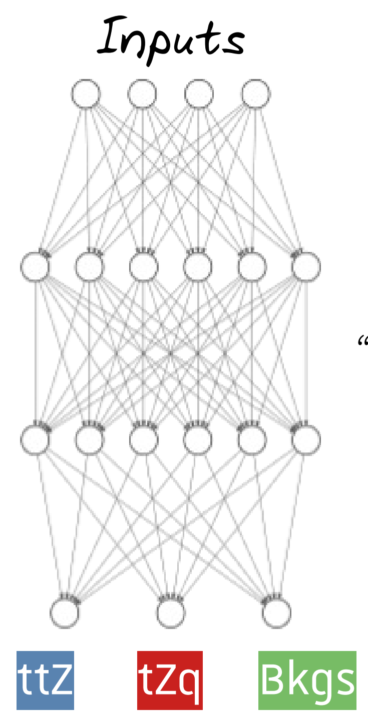


1) SM VS SM

- Multiclass classifier trained to separate tZq/$t\bar{t}Z$/others

ARCHITECTURE :

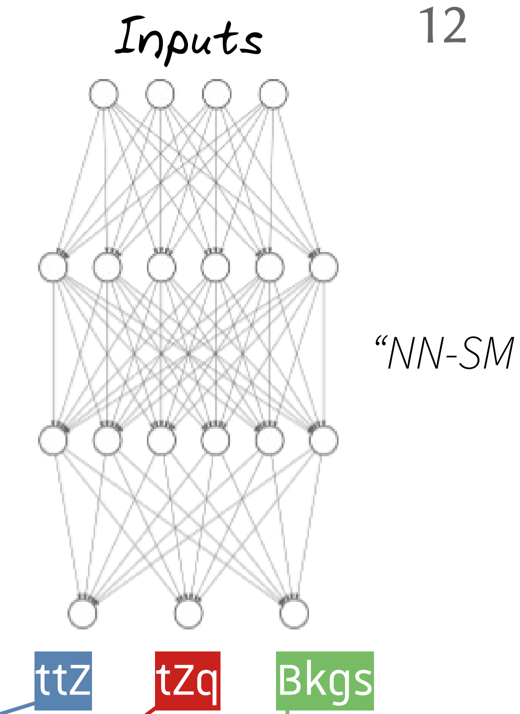
- 3 hidden layers, with 100 neurons each
- Activation function : reLu (softmax for output node)
- Optimizer : Adam
- Loss function : Cross-entropy
- Regularization techniques :
  - Dropout *(ignore random nodes during training)*
  - Batch normalization *(normalize input values at each layer)*
  - L2 regularization *(penalize large weights)*
- ✓ Data/MC validation of input distributions and 2D correlations

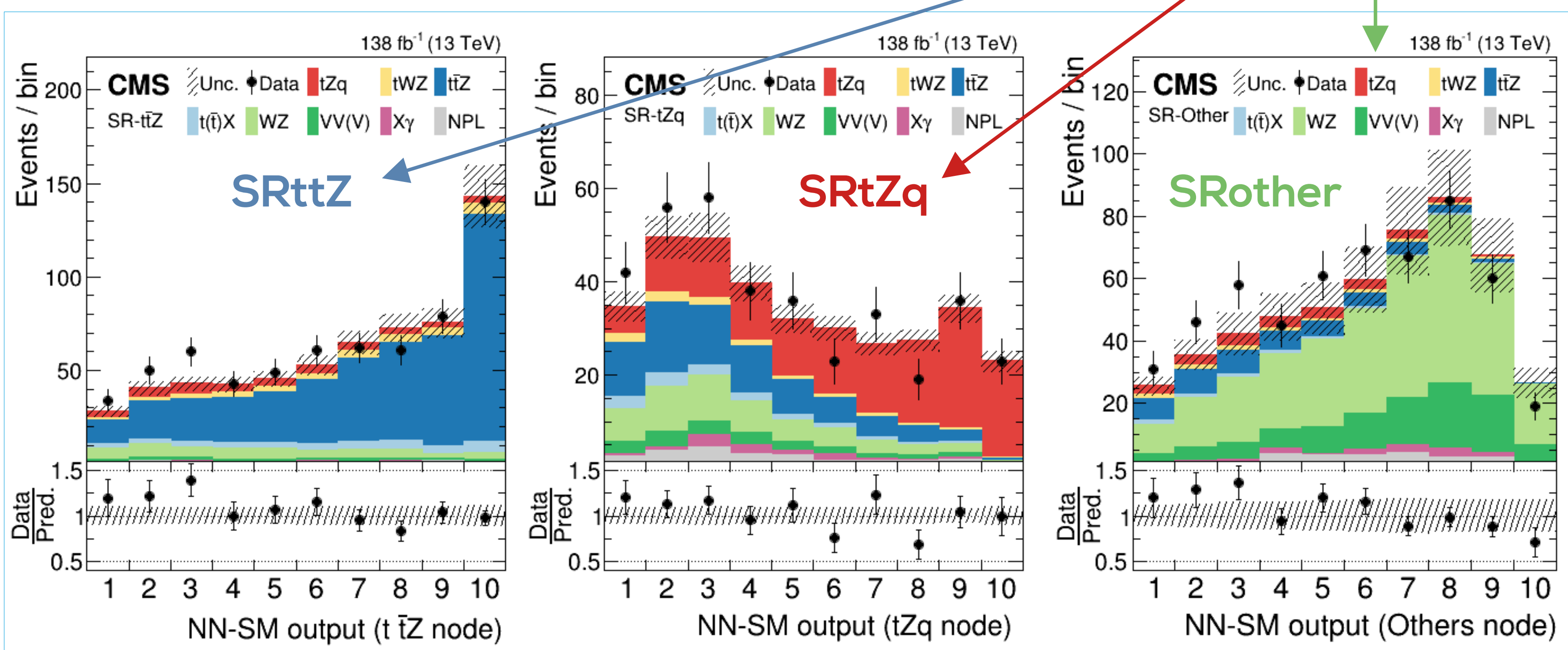


“NN-SM”

1) SM VS SM

- Multiclass classifier trained to separate tZq/t$\bar{t}$Z/others
- Assign events to 3 orthogonal subregions, based on their maximum node value
  → Obtain categories very pure in targeted processes !

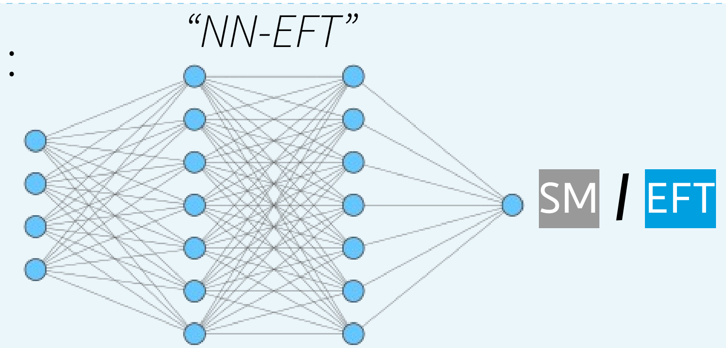




Prefit NN-SM distributions in SR-3ℓ

2) SM VS EFT

- Binary classifiers trained to separate events sampled under :
  - a) SM hypothesis
  - *or*
  - b) many different non-zero values of EFT operator(s)



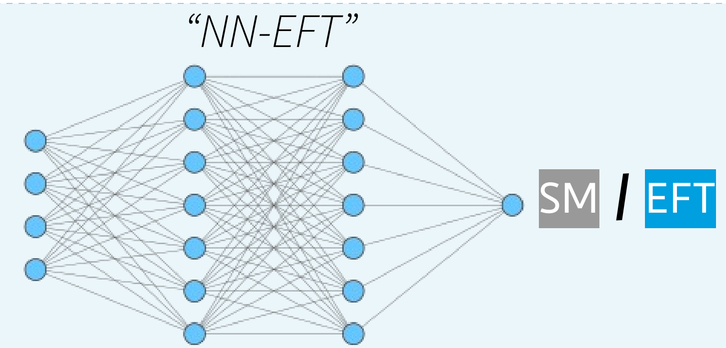
- Learn non-trivial patterns due to new physics out of high-dimensional data
  - → Construct optimized observables describing the compatibility of an event with [SM] or [SM+EFT]

- SM/EFT interference included in training for first time in LHC analysis
  - Difficulty : shapes of kinematic distributions depend on WC values
    - → Train under many different scenarios, NNs learn to interpolate between WC values

- Works with **state-of-the-art physics generators & full detector simulation**
  - Can handle higher-order simulation (but negative weights may reduce training stat.)
  - Fast inference time
  - Optimal in limit of infinite training statistics

2) SM VS EFT

- Binary classifiers trained to separate events sampled for :
  - a) SM hypothesis
  - *or*
  - b) many different non-zero values of EFT operator(s)

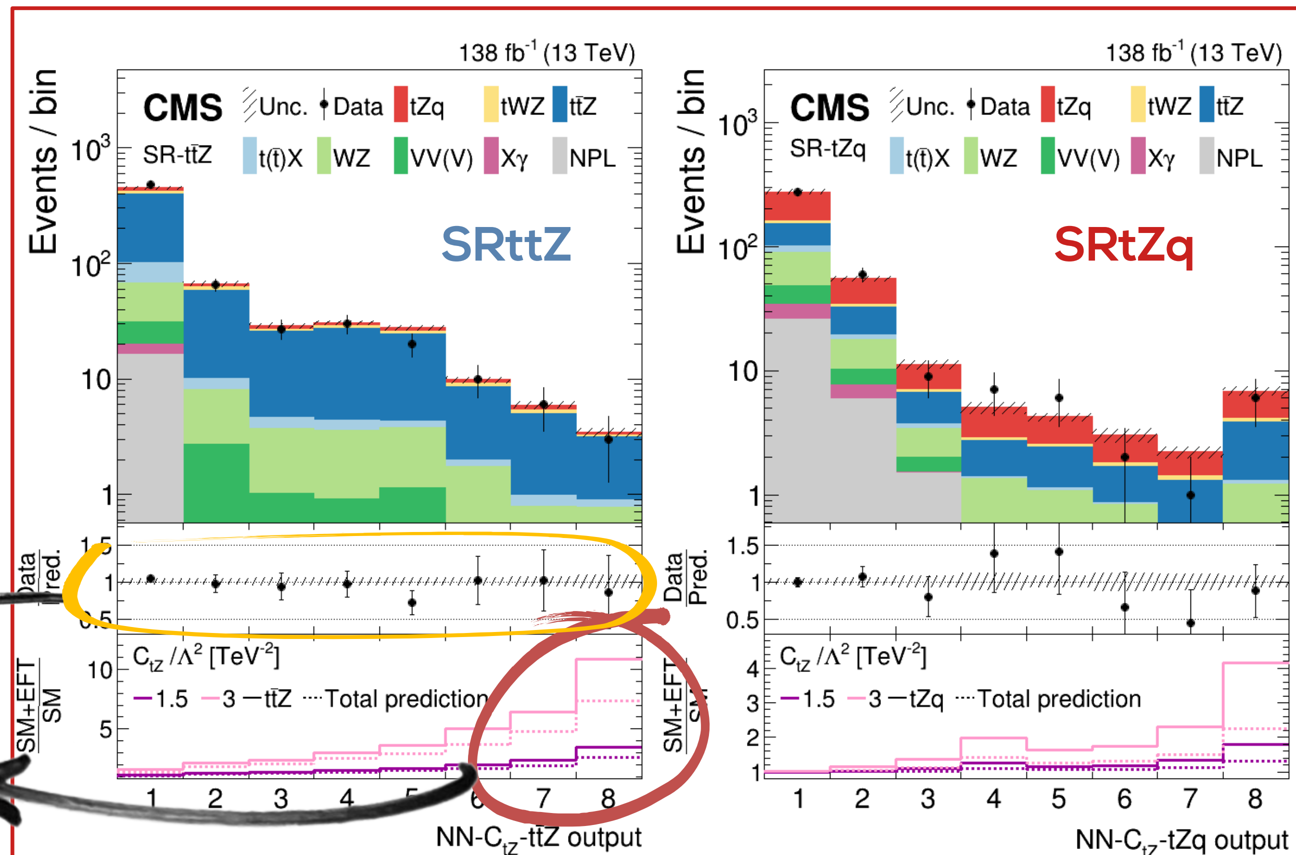


- Learn non-trivial patterns of deviations due to new physics from high-dimensional data
  - → Construct optimized observables describing the compatibility of an event with [SM] or [SM+EFT]

- Separate trainings for tZq & $t\bar{t}Z$
- Separate trainings for individual operators & all 5 simultaneously

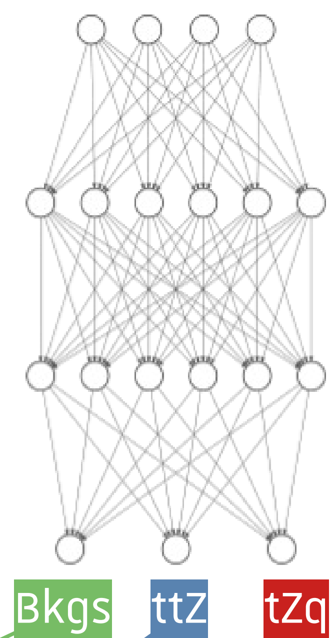
Good data/MC agreement

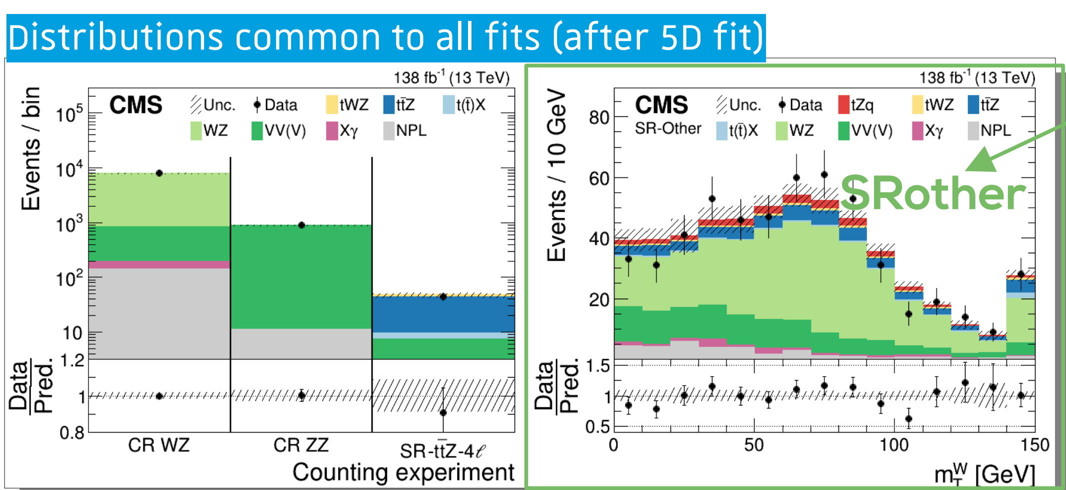
Last bins highly sensitive to EFT scenarios

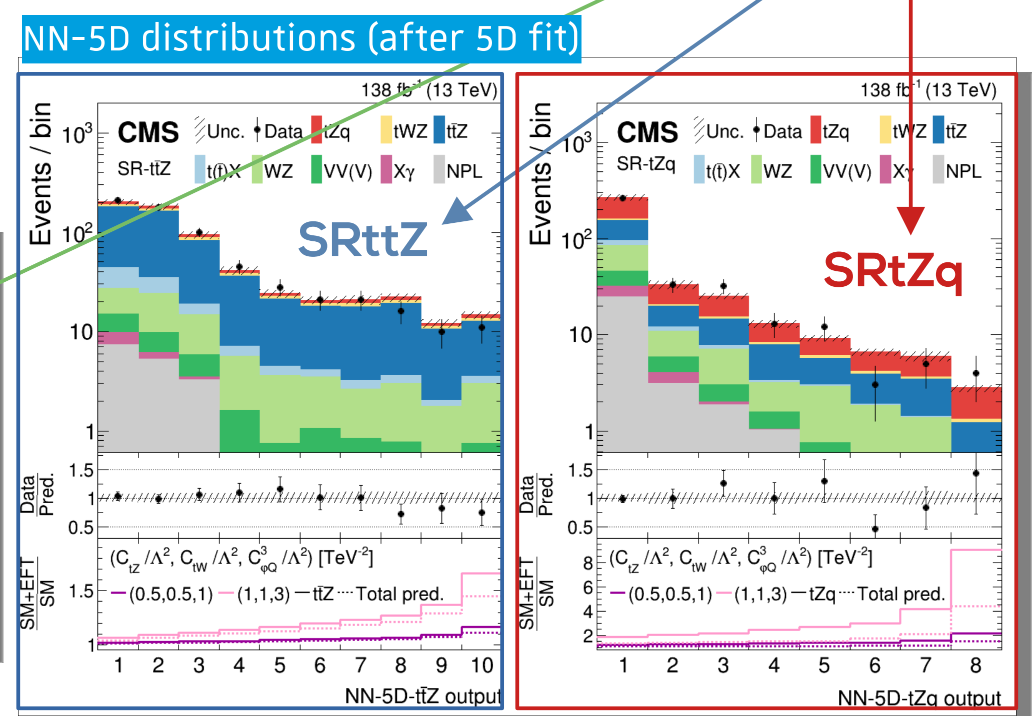


SIGNAL EXTRACTION

- Simultaneous template fits in [6 regions x 3 years]
  - NN distributions used in SRtZq & SRttZ are fit-dependent





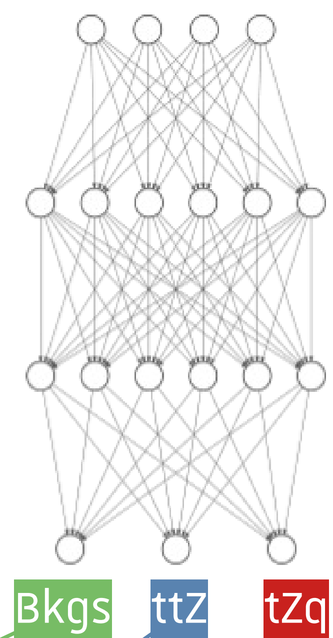


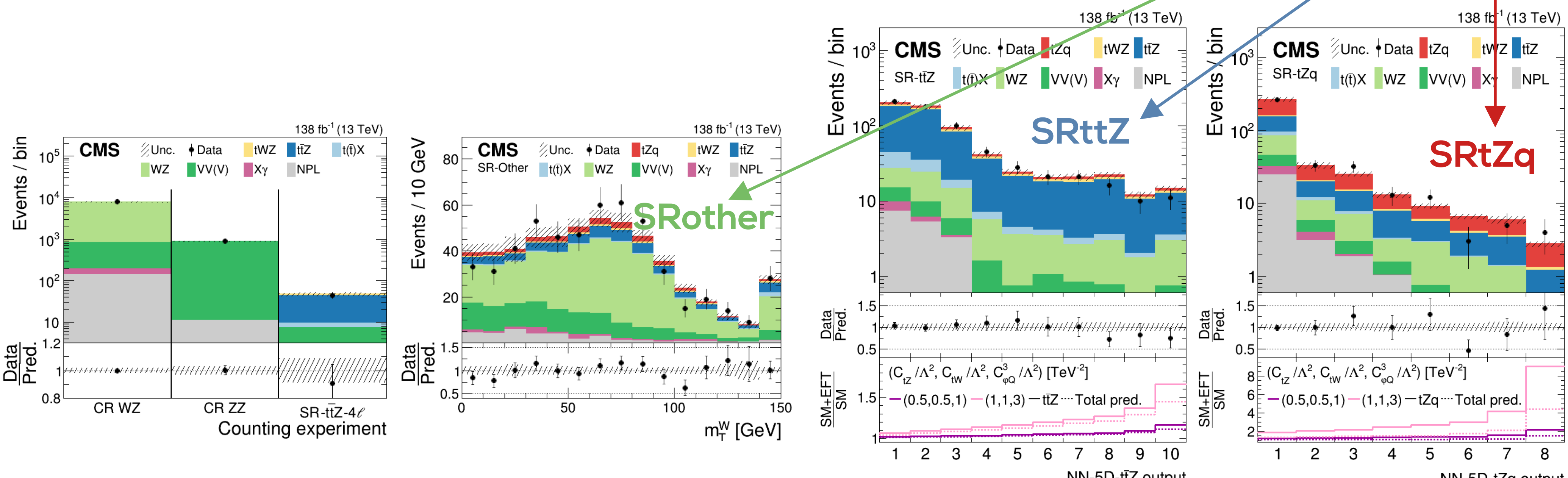
SIGNAL EXTRACTION

WC/Λ^2 $[\mathrm{TeV}^{-2}]$	95% CL confidence intervals			
	Other WCs fixed to SM		5D fit	
	Expected	Observed	Expected	Observed
c_{tZ}	$[-0.97, 0.96]$	$[-0.76, 0.71]$	$[-1.24, 1.17]$	$[-0.85, 0.76]$
c_{tW}	$[-0.76, 0.74]$	$[-0.52, 0.52]$	$[-0.96, 0.93]$	$[-0.69, 0.70]$
$c^3_{\varphi Q}$	$[-1.39, 1.25]$	$[-1.10, 1.41]$	$[-1.91, 1.36]$	$[-1.26, 1.43]$
$c^-_{\varphi Q}$	$[-2.86, 2.33]$	$[-3.00, 2.29]$	$[-6.06, 14.09]$	$[-7.09, 14.76]$
$c_{\varphi t}$	$[-3.70, 3.71]$	$[-21.65, -14.61] \cup [-2.06, 2.69]$	$[-16.18, 10.46]$	$[-19.15, 10.34]$

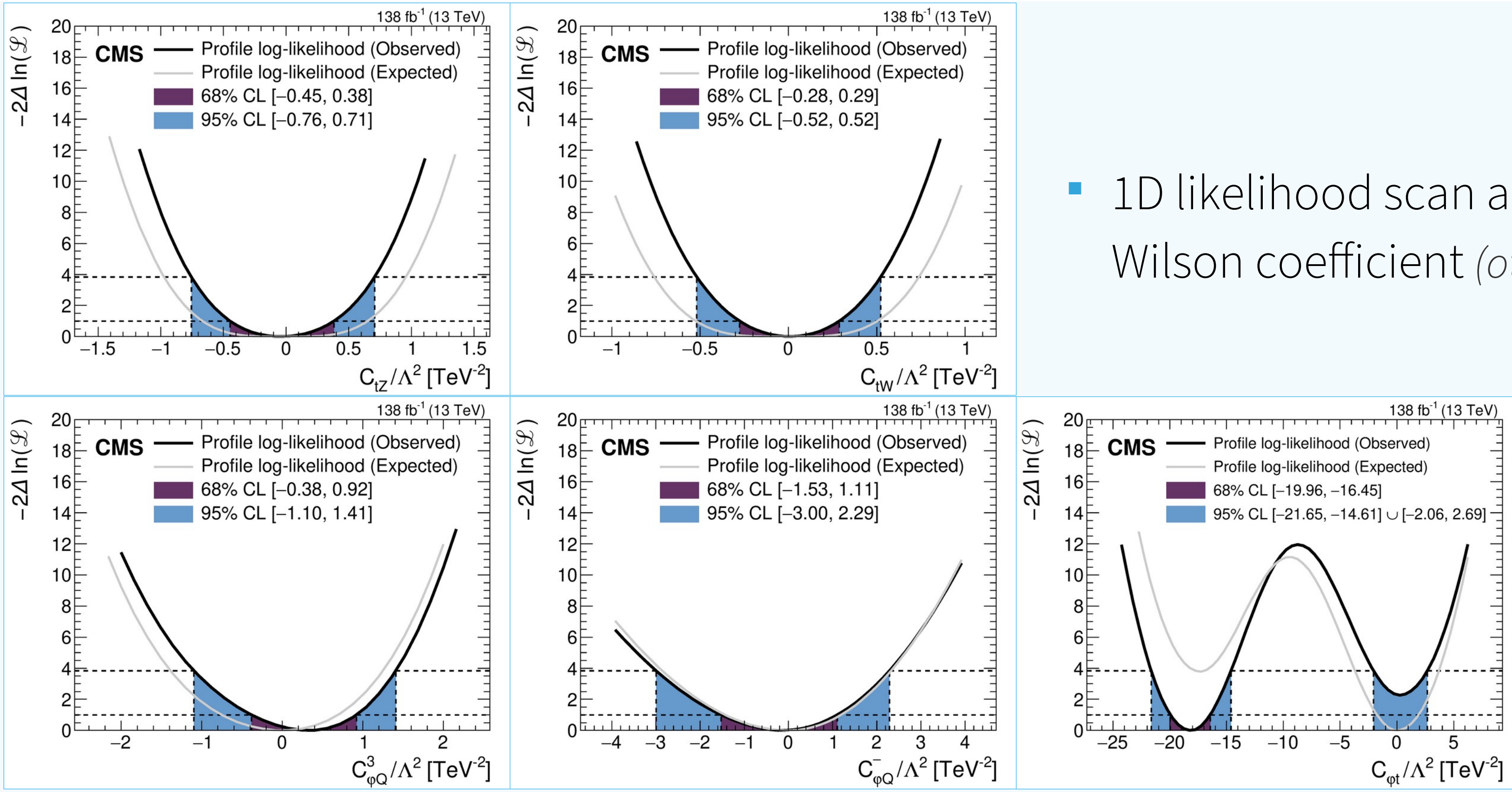
- Extract 1D & 5D confidence intervals @ 95 % CL for each Wilson coefficient
  - All intervals contain the SM predictions
- Shape information reduces widths of confidence intervals by 20 up to 70 %
  - → Significant sensitivity gains using optimal observables

"NN-SM"



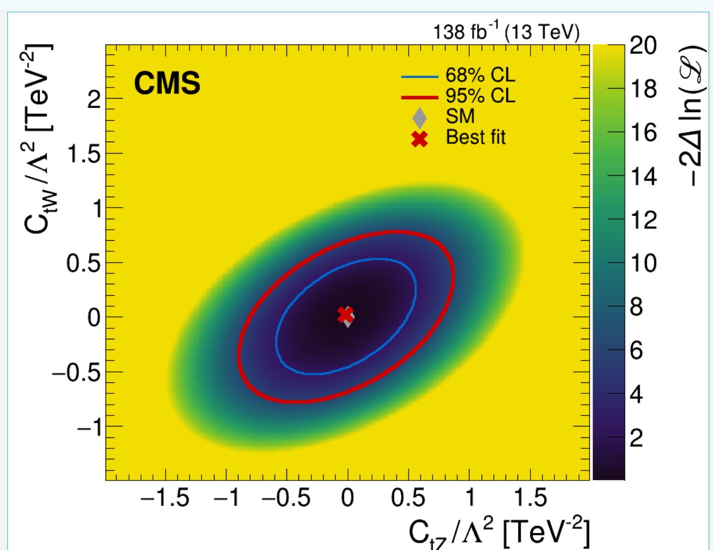


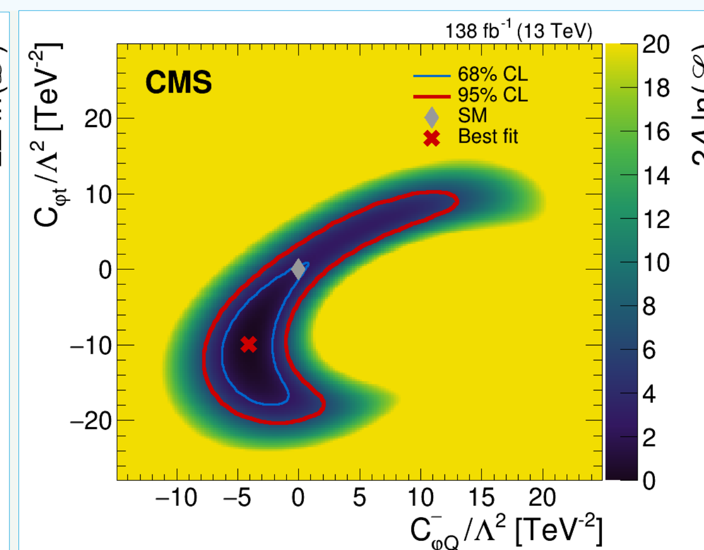
LIKELIHOOD SCANS



- 1D likelihood scan as a function of each Wilson coefficient *(other WCs fixed to 0)*

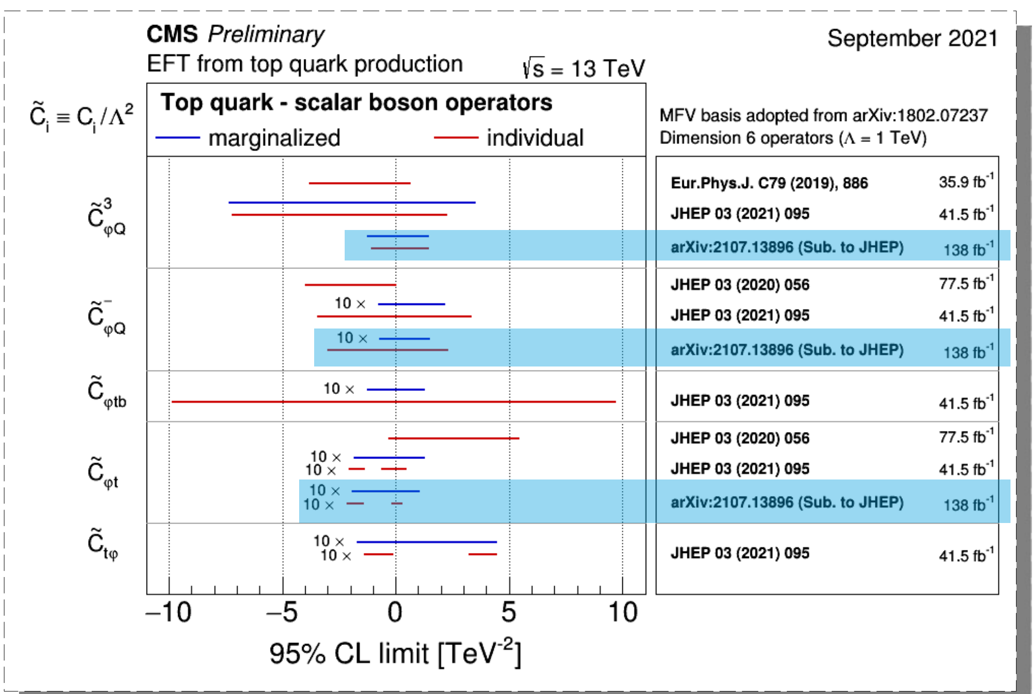
- 2D likelihood scans illustrate correlations of pairs of WCs *(other WCs fixed to 0)*

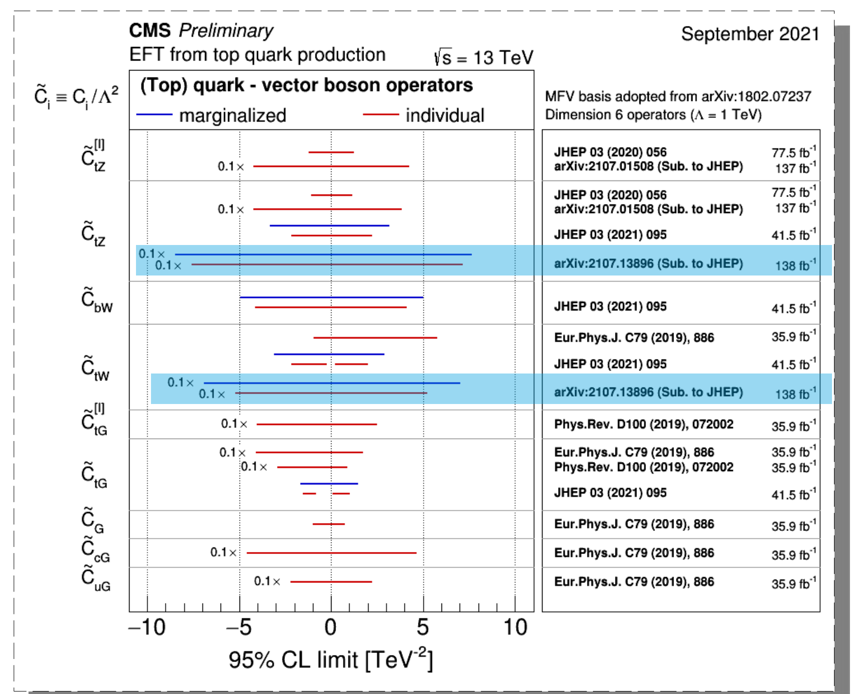




SUMMARY

- Constrain top-electroweak EFT operators with $t(\bar{t})Z$ production in multilepton channels
  - EFT effects parameterized at detector-level via event weights
- Rely on novel ML techniques to enhance sensitivity to EFT
  - Significant sensitivity gains from shape information
  - Not the ‘new default’, but complementary tool with great potential – active research area !
- Obtain **best direct limits to date** from multilepton final states on several Wilson coefficients
  - All 95 % CL confidence intervals contain the SM predictions





BACKUP

SYSTEMATIC UNCERTAINTIES

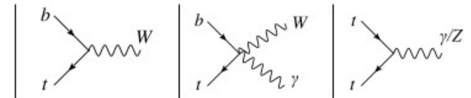
	Source	Type	Correlation
Experimental	Integrated luminosity	Yield	Partial
	Trigger efficiency	Yield	—
	Pileup	Both	✓
	Lepton identification and isolation	Both	✓
	b tagging	Both	Partial
	Jet energy scale	Both	Partial
	Jet energy resolution	Both	—
	Missing transverse momentum	Both	—
	L1 ECAL inefficiency	Both	✓
Theoretical	PDF	Both	✓
	α_S	Both	✓
	ME scales μ_R, μ_F	Both	✓
	Signal SM cross sections	Yield	✓
	ISR and FSR	Both	✓
	Additional radiation	Shape	✓
Backgrounds	WZ normalization	Yield	✓
	VV(V) normalization	Yield	✓
	$t(\bar{t})X$ normalization	Yield	✓
	$X\gamma$ normalization	Yield	✓
	NPL normalization	Yield	✓
	NPL misidentification probabilities	Both	✓

Impacts of nuisance groups on each WC

Source	c_{tZ}	c_{tW}	t	$c^{-}_{\varphi Q}$	$c_{\varphi t}$
tZq normalization	<0.1	<0.1	1.2	0.1	0.8
$t\bar{t}Z$ normalization	0.6	<0.1	0.4	37	38
tWZ normalization	0.1	0.1	<0.1	0.7	2.1
Background normalizations	<0.1	<0.1	6.9	3.6	6.8
NPL background estimation	1.4	0.2	5.6	0.3	3.8
Jet energy scale	<0.1	<0.1	0.8	0.7	2.3
Jet energy resolution	<0.1	<0.1	<0.1	<0.1	1.4
p_T^{miss}	<0.1	<0.1	<0.1	<0.1	0.2
b tagging	<0.1	<0.1	0.9	2.0	0.3
Other (experimental)	<0.1	<0.1	1.6	0.8	0.6
Lepton identification and isolation	0.4	0.4	1.2	2.2	0.8
Theory	2.1	1.1	0.4	0.9	0.9

Stat. limited **Syst. limited**

EFT OPERATORS

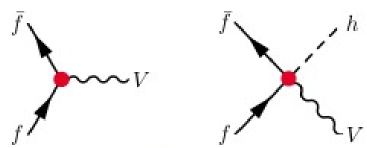


C_{uW}^{33}	$\otimes$	$\otimes$	$\otimes$

- $\mathcal{O}_{tZ}$ and $\mathcal{O}_{tW} \leftrightarrow$ Electroweak dipole operators
- $\mathcal{O}^3_{\phi Q} \leftrightarrow$ Left-handed SU(2) triplet current operator
- $\mathcal{O}^-_{\phi Q} \leftrightarrow$ Induces anomalous neutral-current interactions
- $\mathcal{O}_{\phi t} \leftrightarrow$ Right-handed neutral current operator (idem)

Operator	WC	Mapping to Warsaw-basis coefficients
$\mathcal{O}_{\mathrm{tZ}}$	c_{tZ}	$\mathrm{Re}\{-s_W c_{\mathrm{uB}}^{(33)} + c_W c_{\mathrm{uW}}^{(33)}\}$
$\mathcal{O}_{\mathrm{tW}}$	c_{tW}	$\mathrm{Re}\{c_{\mathrm{uW}}^{(33)}\}$
$\mathcal{O}^3_{\varphi\mathrm{Q}}$	$c^3_{\varphi\mathrm{Q}}$	$c_{\varphi q}^{3(33)}$
$\mathcal{O}^-_{\varphi\mathrm{Q}}$	$c^-_{\varphi\mathrm{Q}}$	$c_{\varphi q}^{1(33)} - c_{\varphi q}^{3(33)}$
$\mathcal{O}_{\varphi\mathrm{t}}$	$c_{\varphi\mathrm{t}}$	$c_{\varphi\mathrm{u}}^{(33)}$

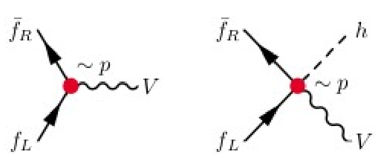
currents $i(\varphi^\dagger \overleftrightarrow{D}^\mu \varphi)(\bar{Q}\gamma^\mu Q)$



$C_{\phi f}$

- Shift SM $f\bar{f}V$ couplings
- $f\bar{f}Vh$ contact interactions

dipole $(\bar{q}\,\sigma_{\mu\nu}\,t\,\tilde{\varphi})V^{\mu\nu}$



C_{tV}

- Chirality flipping $f\bar{f}V$ couplings
- $f\bar{f}V(V)h$ contact interactions
- W, B & G fields

parameter	$t\bar{t}$	single t	tW	tZ	t decay	$t\bar{t}Z$	$t\bar{t}W$
$C_{Qq}^{1,8}$	Λ^{-2}	–	–	–	–	Λ^{-2}	Λ^{-2}
$C_{Qq}^{3,8}$	Λ^{-2}	Λ^{-4} $[\Lambda^{-2}]$	–	Λ^{-4} $[\Lambda^{-2}]$	Λ^{-4} $[\Lambda^{-2}]$	Λ^{-2}	Λ^{-2}
C_{tu}^8, C_{td}^8	Λ^{-2}	–	–	–	–	Λ^{-2}	–
$C_{Qq}^{1,1}$	Λ^{-4} $[\Lambda^{-2}]$	–	–	–	–	Λ^{-4} $[\Lambda^{-2}]$	Λ^{-4} $[\Lambda^{-2}]$
$C_{Qq}^{3,1}$	Λ^{-4} $[\Lambda^{-2}]$	Λ^{-2}	–	Λ^{-2}	Λ^{-2}	Λ^{-4} $[\Lambda^{-2}]$	Λ^{-4} $[\Lambda^{-2}]$
C_{tu}^1, C_{td}^1	Λ^{-4} $[\Lambda^{-2}]$	–	–	–	–	Λ^{-4} $[\Lambda^{-2}]$	–
C_{Qu}^8, C_{Qd}^8	Λ^{-2}	–	–	–	–	Λ^{-2}	–
C_{tq}^8	Λ^{-2}	–	–	–	–	Λ^{-2}	Λ^{-2}
C_{Qu}^1, C_{Qd}^1	Λ^{-4} $[\Lambda^{-2}]$	–	–	–	–	Λ^{-4} $[\Lambda^{-2}]$	–
C_{tq}^1	Λ^{-4} $[\Lambda^{-2}]$	–	–	–	–	Λ^{-4} $[\Lambda^{-2}]$	Λ^{-4} $[\Lambda^{-2}]$
$C_{\phi Q}^-$	–	–	–	Λ^{-2}	–	Λ^{-2}	–
$C_{\phi Q}^3$	–	Λ^{-2}	Λ^{-2}	Λ^{-2}	Λ^{-2}	–	–
$C_{\phi t}$	–	–	–	Λ^{-2}	–	Λ^{-2}	–
$C_{\phi tb}$	–	Λ^{-4}	Λ^{-4}	Λ^{-4}	Λ^{-4}	–	–
C_{tZ}	–	–	–	Λ^{-2}	–	Λ^{-2}	–
C_{tW}	–	Λ^{-2}	Λ^{-2}	Λ^{-2}	Λ^{-2}	–	–
C_{bW}	–	Λ^{-4}	Λ^{-4}	Λ^{-4}	Λ^{-4}	–	–
C_{tG}	Λ^{-2}	$[\Lambda^{-2}]$	Λ^{-2}	–	$[\Lambda^{-2}]$	Λ^{-2}	Λ^{-2}

Brivio et al., arXiv:1910.03606

MVA INPUT FEATURES

- Choice of input variables informed by generator-level studies, correlations, relative rankings, and modeling

Variable	NN-SM	NN-c_{tZ}-tZq	NN-c_{tZ}-t$\bar{t}$Z	NN-c_{tW}-tZq	NN-c_{tW}-t$\bar{t}$Z	NN-$c^3_{\varphi Q}$-tZq	NN-$c^3_{\varphi Q}$-t$\bar{t}$Z	NN-5D-tZq	NN-5D-t$\bar{t}$Z
p_T^Z	—	✓	✓	✓	✓	✓	✓	✓	✓
$\eta(Z)$	✓	✓	✓	—	—	✓	—	—	✓
$\Delta\phi(\ell_1^Z \ell_2^Z)$	✓	✓	✓	✓	✓	✓	✓	✓	✓
$p_T(t)$	✓	✓	✓	—	✓	✓	—	✓	✓
$\eta(t)$	—	✓	✓	✓	✓	✓	—	—	✓
$m(t, Z)$	—	—	—	—	—	—	—	—	—
$\lvert\eta(j')\rvert$	✓	—	—	—	—	—	—	✓	—
$p_T(j')$	✓	✓	—	✓	—	—	—	—	—
$\Delta R(b, \ell_t)$	—	✓	—	✓	—	—	—	—	—
$\Delta R(j', \ell_t)$	✓	—	—	—	—	—	—	—	—
$\Delta R(t, Z)$	—	✓	✓	✓	—	✓	—	—	✓
$\Delta\eta(Z, j')$	—	✓	—	—	—	—	—	✓	—
ΔR between t and the closest lepton	—	✓	—	✓	—	—	—	—	—
ΔR between j' and the closest lepton	—	—	—	—	—	—	—	✓	—
$m_{3\ell}$	✓	—	—	—	✓	—	✓	—	✓
m_T^W	✓	✓	✓	—	—	—	—	—	✓
p_T^{miss}	✓	—	—	—	—	—	—	—	—
Lepton asymmetry	✓	—	—	✓	✓	—	—	✓	—
$\cos\theta_Z^\star$	—	—	✓	—	—	✓	—	—	✓
Max. p_T among jet pairs	—	—	—	—	—	—	✓	—	✓
Max. DeepJet discriminant	✓	—	—	—	—	—	—	—	—
b jet multiplicity	✓	—	—	—	—	—	—	—	—
Three-momenta of the three leading leptons	✓	—	—	—	—	—	—	—	—
Three-momenta of the three leading jets	✓	—	—	—	—	—	—	—	—
DeepJet discriminants of the three leading jets	✓	—	—	—	—	—	—	—	—
Number of variables	33	11	8	8	6	7	4	7	10

NN-EFT – EVENT SAMPLING

Basic algorithm

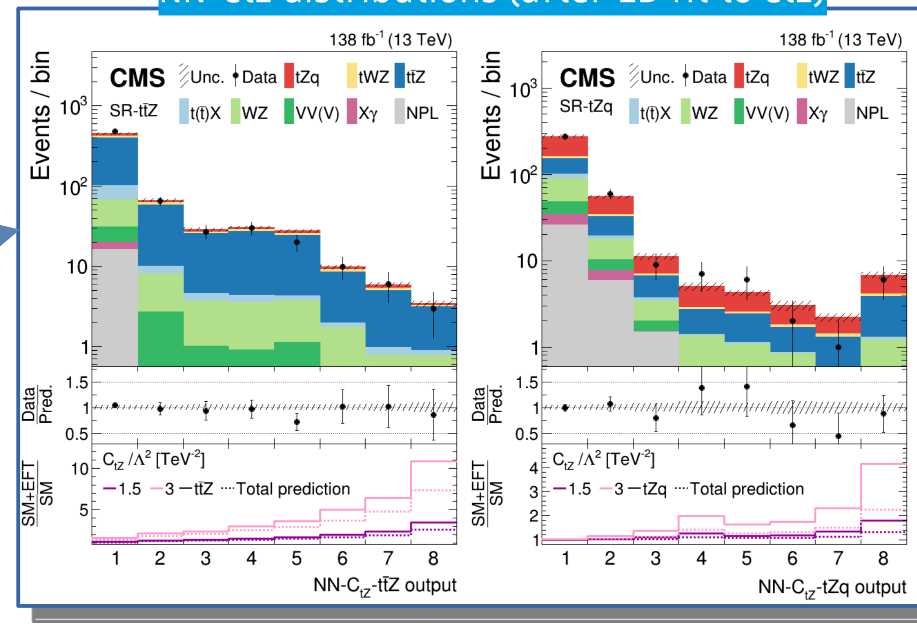
- Choose N hypotheses θ corresponding to different EFT scenarios (ex : ctZ ∈ [-5, -4, …, 4, 5])
- For each hypothesis θ :
  - Sample N_e events according to **θ** : set class label to **0**
  - Sample N_e events according to **SM** : set class label to **1**

➡ Network learns how to separate SM/EFT, for different Wilson coefficients

- In practice, the data set is *unweighted* ↔ sum of all event weights normalized to unity, so that each event weight can be interpreted as a probability for an event under a given hypothesis
  - Can then sample a subset of events of any size, representative of a given hypothesis
  - Must allow for replacement (↔ the same event may be selected multiple times)
- The range of WC values over which training events are sampled is a hyperparameter
  - Generally chosen a few times larger than existing constraints (→ avoid bias on previous results, and easier for NN to learn features at larger WC values)

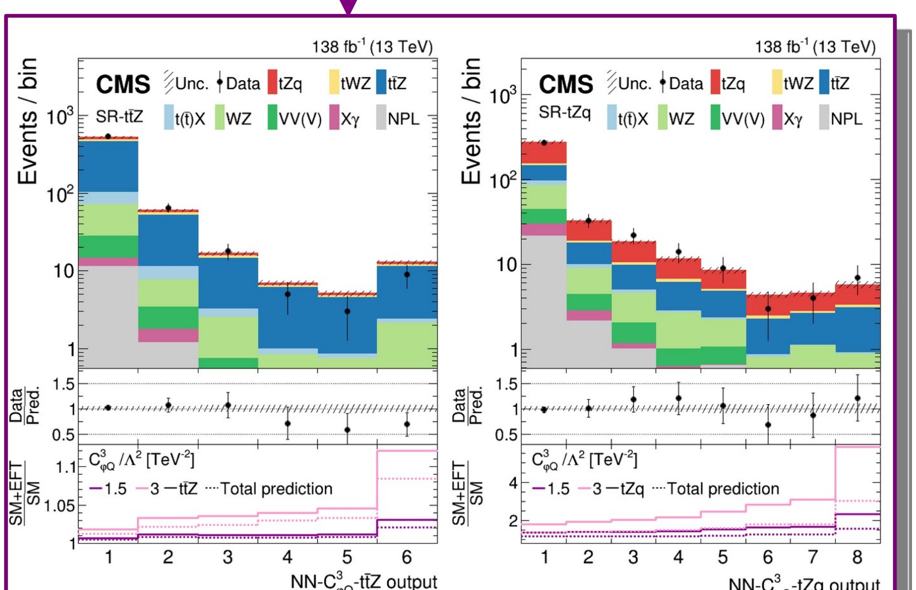
SIGNAL EXTRACTION

NN-ctz distributions (after 1D fit to ctz)

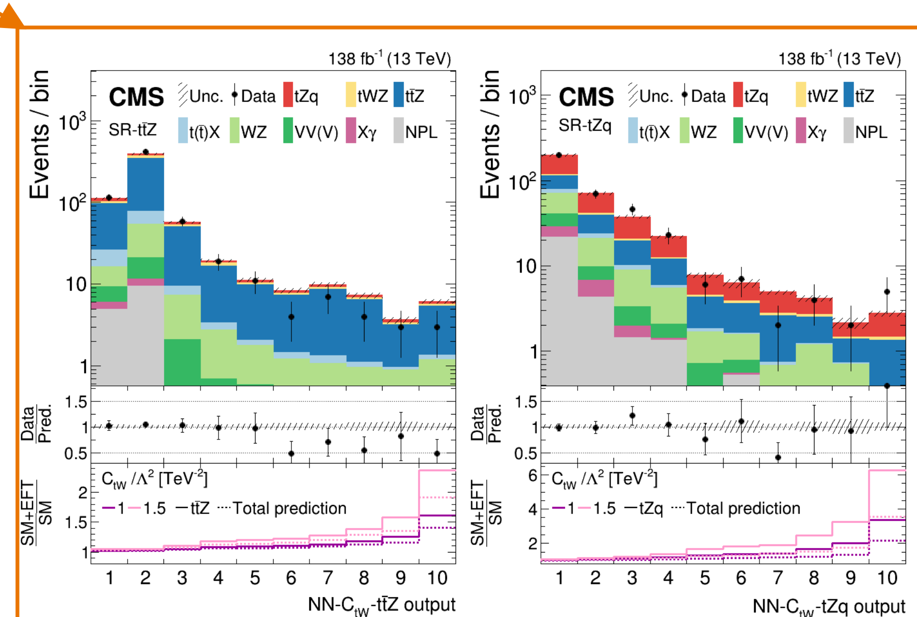


Observables used in the different regions and fits

Fit configuration	Region					
	SR-tZq	SR-$t\bar{t}$Z	SR-Others	SR-$t\bar{t}$Z-4ℓ	CR WZ	CR ZZ
1D c_{tZ}	NN-c_{tZ}-tZq	NN-c_{tZ}-$t\bar{t}$Z				
1D c_{tW}	NN-c_{tW}-tZq	NN-c_{tW}-$t\bar{t}$Z				
1D $c^3_{\varphi Q}$	NN-$c^3_{\varphi Q}$-tZq	NN-$c^3_{\varphi Q}$-$t\bar{t}$Z	m_T^W	Counting experiments		
1D $c^-_{\varphi Q}$	NN-SM (tZq node)	NN-SM ($t\bar{t}$Z node)				
1D $c_{\varphi t}$	NN-SM (tZq node)	NN-SM ($t\bar{t}$Z node)				
2D and 5D	NN-5D-tZq	NN-5D-$t\bar{t}$Z				



NN-cpq3 distributions (after 1D fit to cpq3)



NN-ctw distributions (after 1D fit to ctw)

(SOME) PROMISING DIRECTIONS

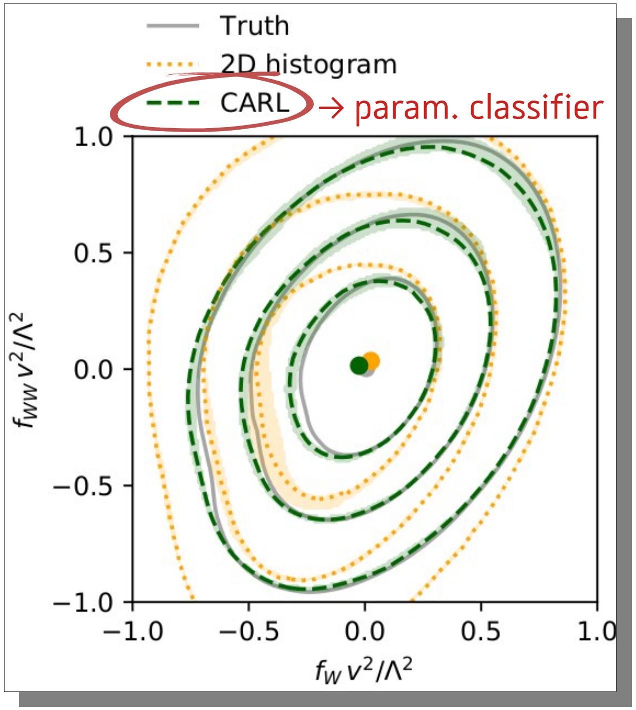
- *Successful proof-of-concept studies by Jonas Rübenach (link)*

- *Likelihood-free inference* : regress on true likelihood ratio
  - Set limits directly from NN response
  - Outperforms limited summary statistics
  - Only deal with few theory systematics for now

- *Augmented data* : maximize use of generator info. in training
  - Better performance with less training data

- Morphing-aware architectures
  - Impose EFT structure on ML architecture
  - Model SM / interference / EFT components with individual estimators (→ Sum)

- … and more !



10.1103/PhysRevD.98.052004