Machine-learning VH-bb EFT effects in CMS

Suman Chatterjee

Deutsches Elektronen-Synchrotron (DESY), Hamburg

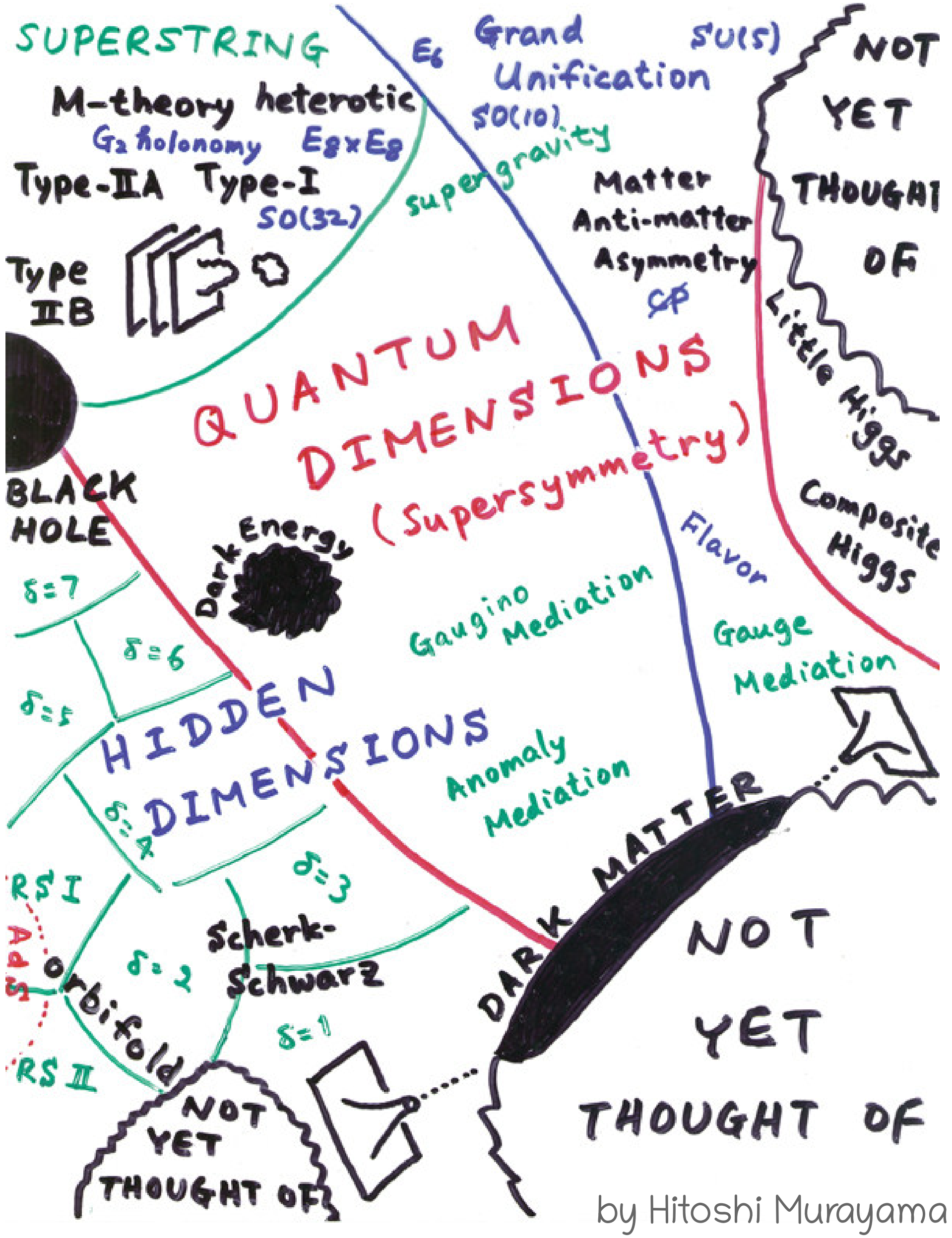
20/02/2025

COMETA workshop on

Polarized Perspectives: Tagging and Learning in the SM

Austrian Academy of Sciences, Vienna

Landscape of new physics



by Hitoshi Murayama

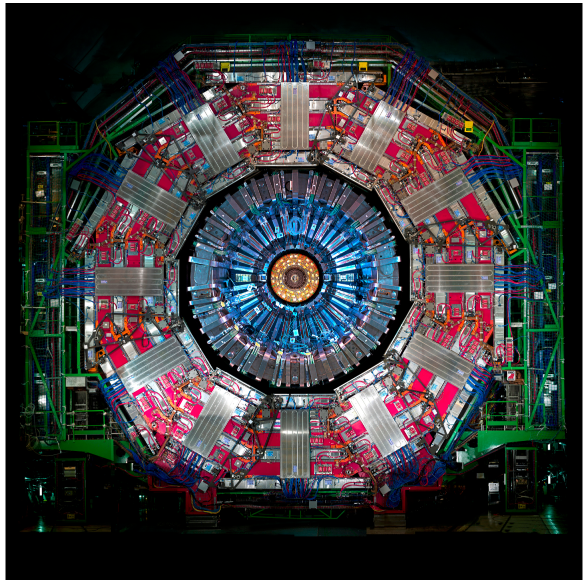
[image link]

Plethora of new ideas → solving one or many problems

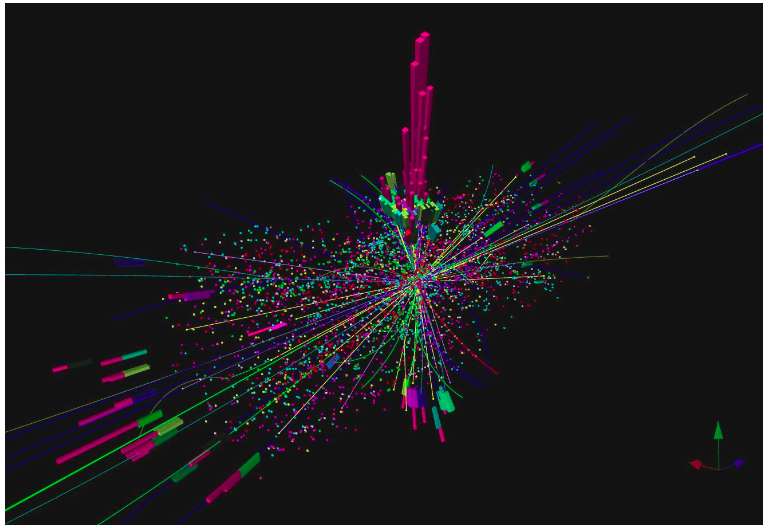
Formulated in terms of Lagrangians

← Parameters: couplings & energy scales

→ Predictions for new phenomena at experiments



Hidden Valley Z' → jets (simulation)

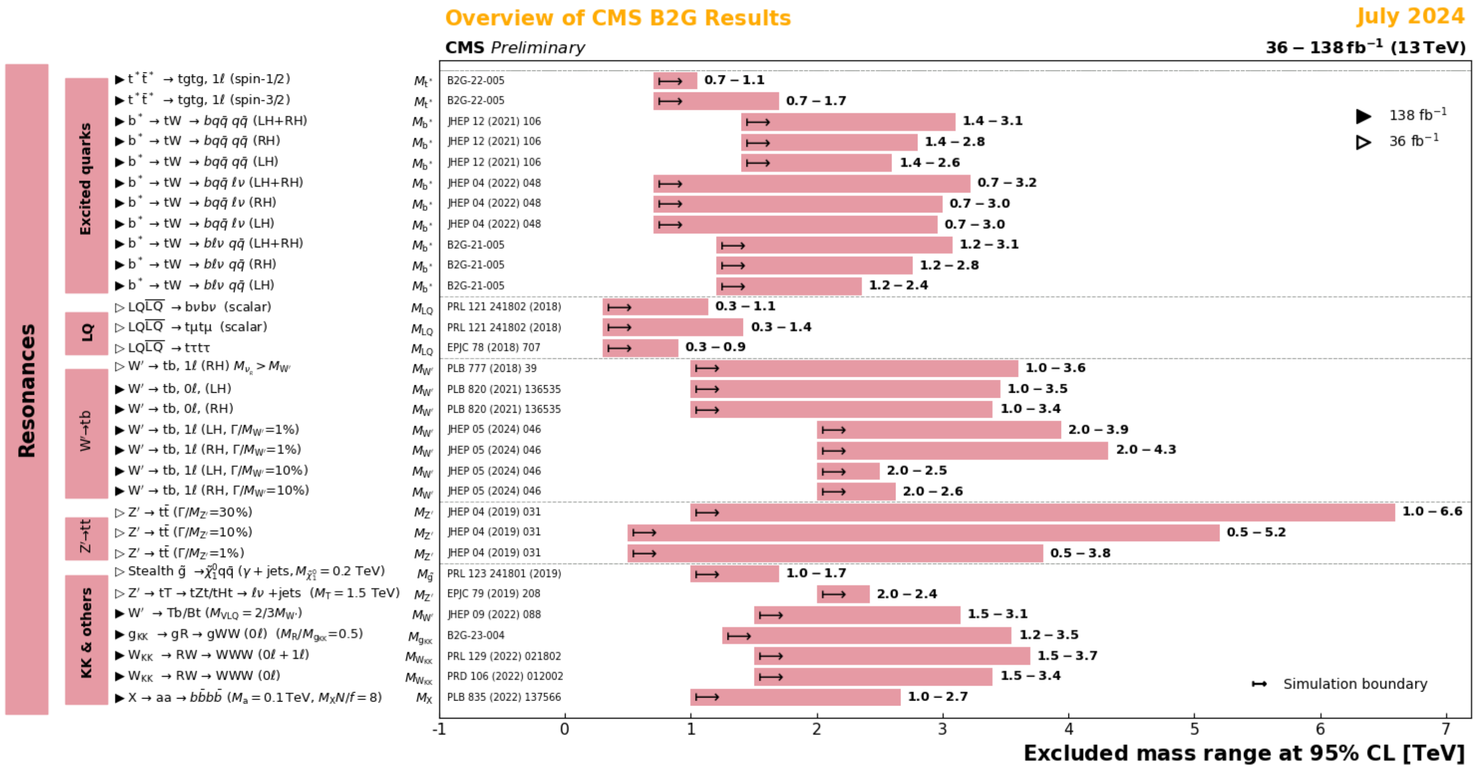


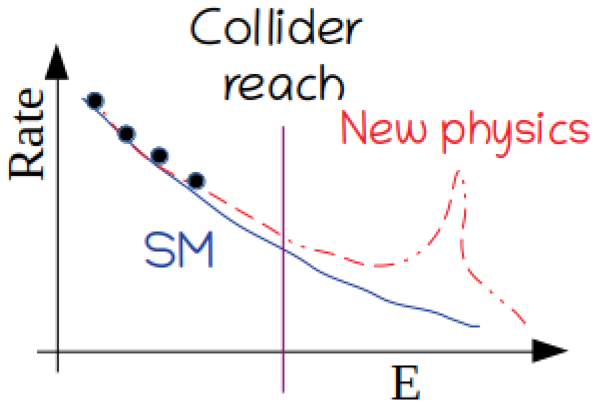
[image link]

Scale of new physics?

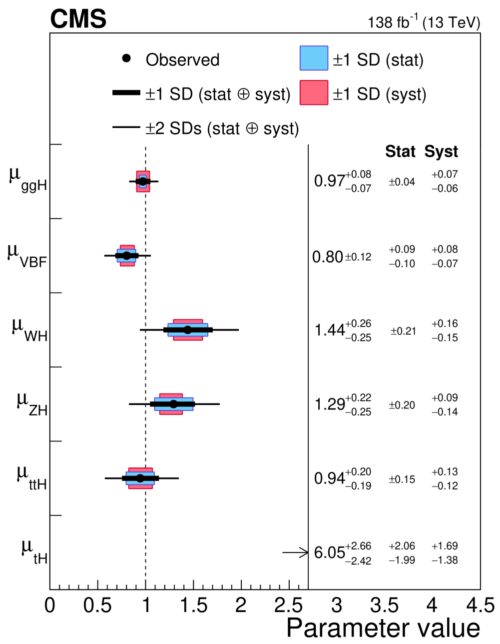
Newest fundamental particle discovered: Last missing piece in standard model (SM)

CMS B2G-Resonances summary plots





Nature 607 (2022) 60-68



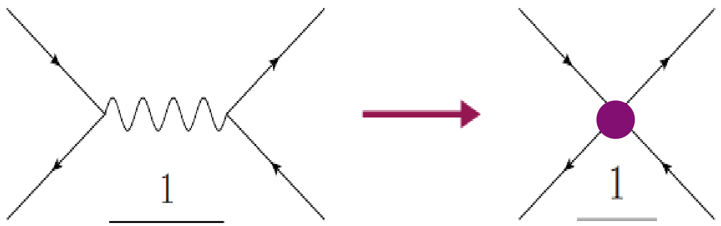
No smoking gun signature of a heavy resonance yet from LHC data

→ Hint for a separation of new physics scale & electroweak scale?

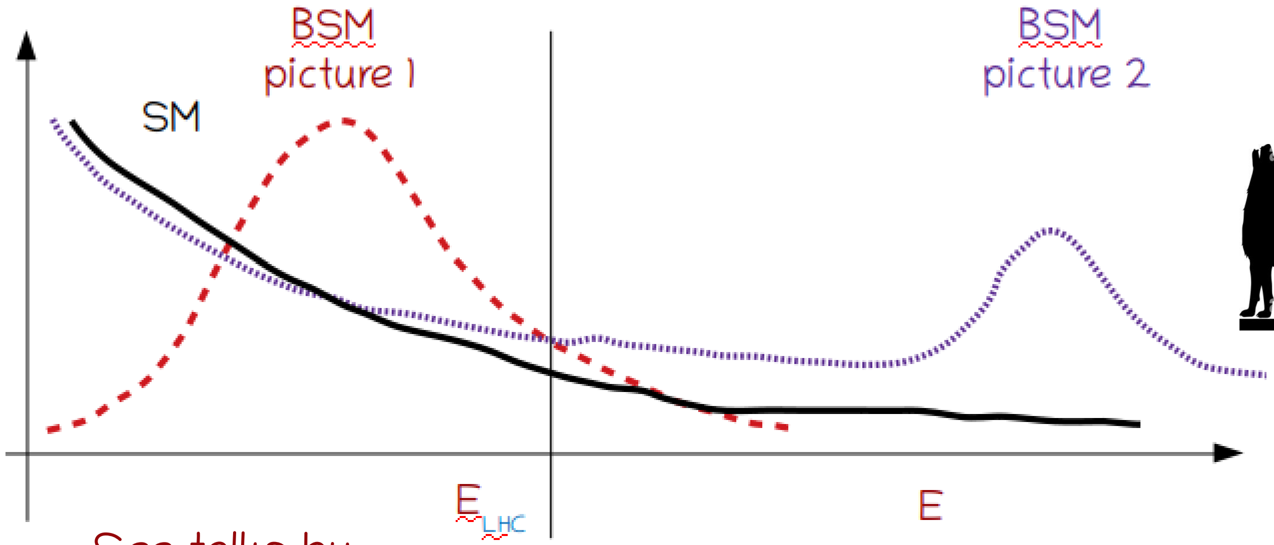
Large new physics scale consistent with inclusive measurements

Standard model effective field theory (SMEFT)

New (heavy) particles modify SM interactions



$$\frac{1}{p^2 - M^2} \rightarrow \frac{1}{-M^2}$$



See talks by

A. Malara

A. Agapitos

Discovery through precision measurements

Deviations from SM ← parameterized by effective field theory operators

$$\mathcal{L}_{\mathrm{SMEFT}} = \mathcal{L}_{\mathrm{SM}} + \cancel{\sum_i \frac{c_i^{(5)}}{\Lambda}\mathcal{O}_{5,i}} + \boxed{\sum_i \frac{c_i^{(6)}}{\Lambda^2}\mathcal{O}_{6,i}} + \cancel{\sum_i \frac{c_i^{(7)}}{\Lambda^3}\mathcal{O}_{7,i}} + \sum_i \frac{c_i^{(8)}}{\Lambda^4}\mathcal{O}_{8,i} + \dots .$$

Lepton number violation

Lepton & Baryon number violation

Assumptions:

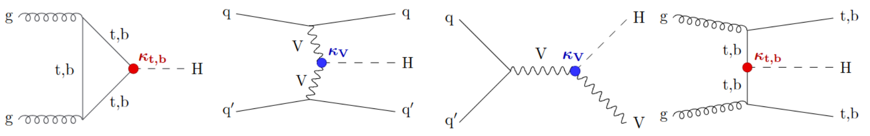
→ Particle content same as in SM

→ SM Gauge symmetries $SU(3)_C \times SU(2)_L \times U(1)_Y$ respected

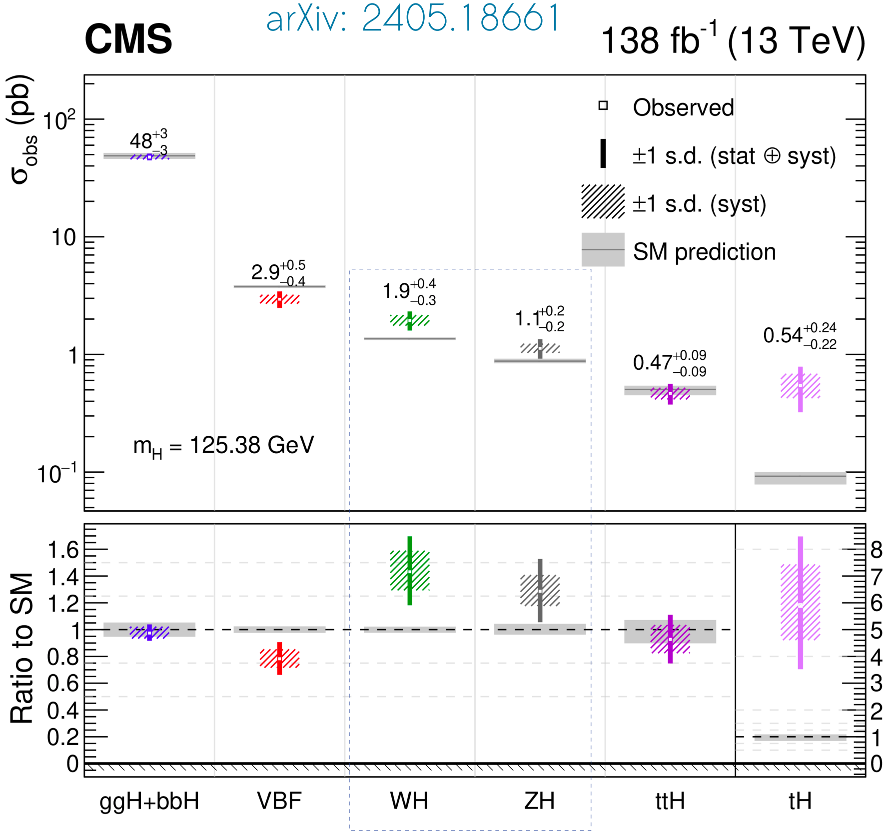
59 SMEFT operators @ dim=6 Grzadkowski, Iskrzyński, Misiak, Rosiek (2010)

Higgstrahlung: small but important

H production



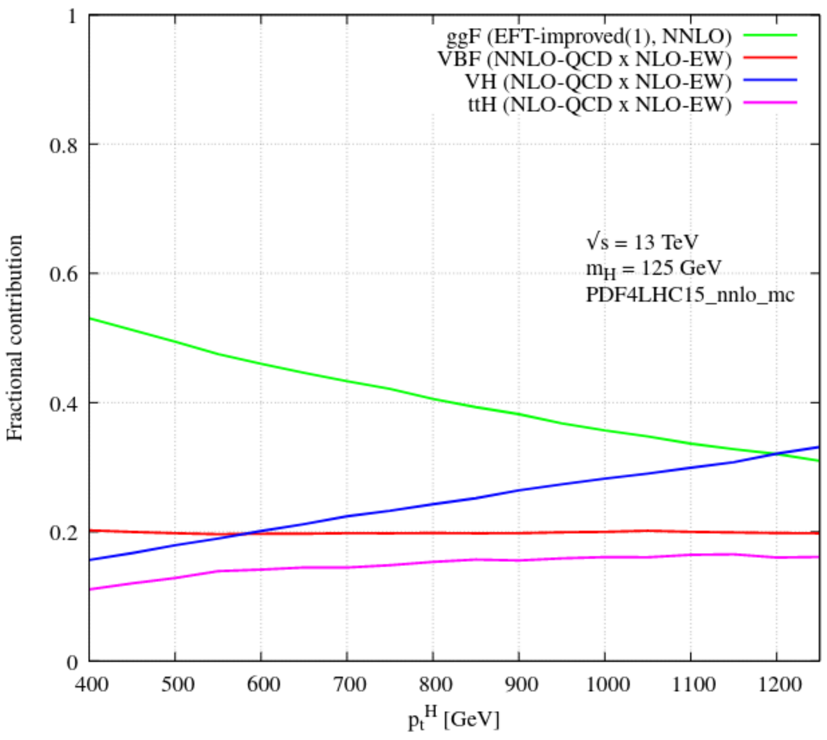
arXiv: 2405.18661



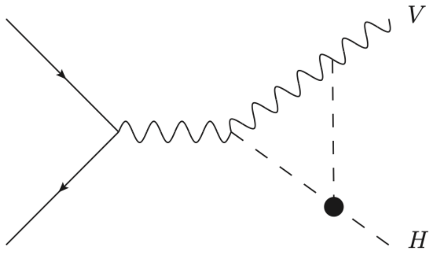
WH/ZH → small production cross section

Higgs-strahlung important at high energy

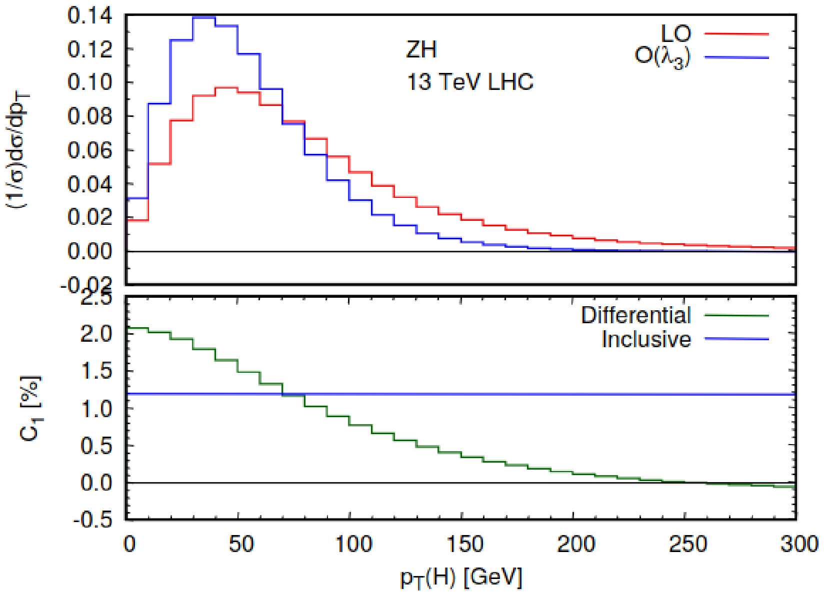
LHCHXSWG-2019-002



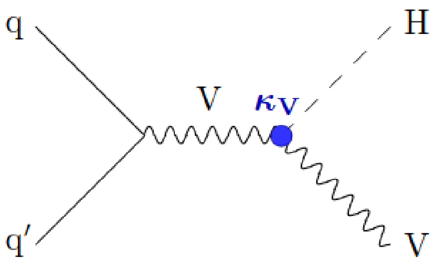
Sensitive to Higgs self-coupling



Maltoni, Pagani, Shivaji, Zhao (2017)



Cross section measurement of V associated H production @CMS

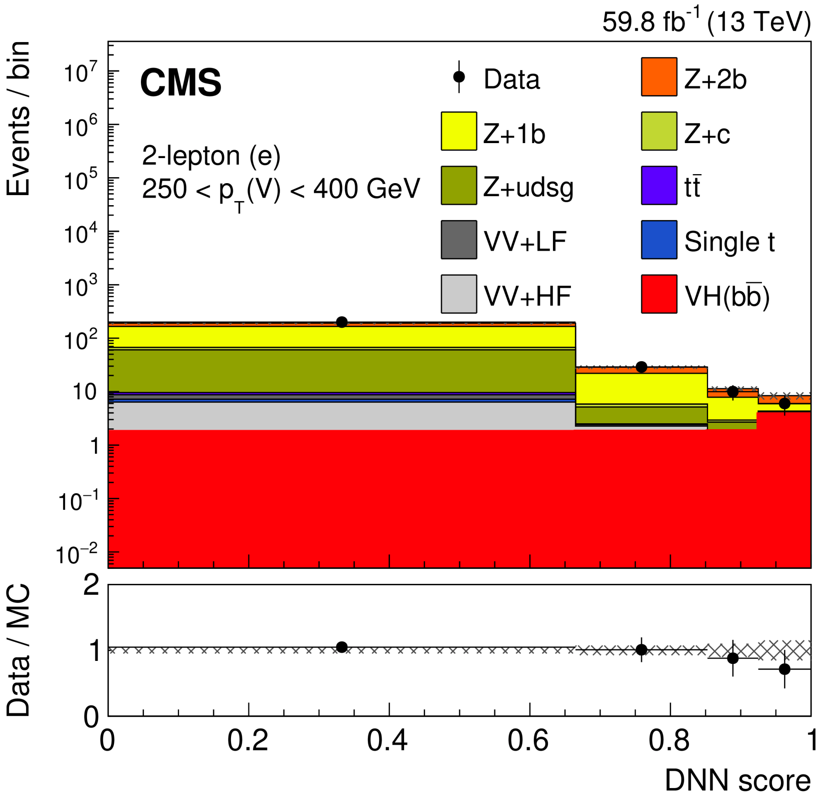


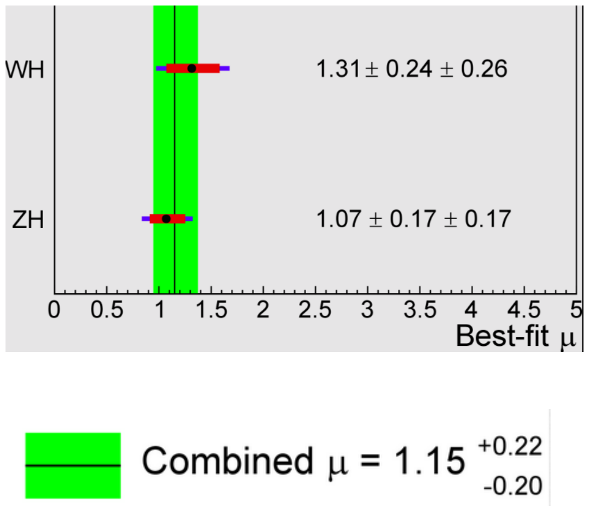
Phys.Rev.D 109 (2024) 9, 092011

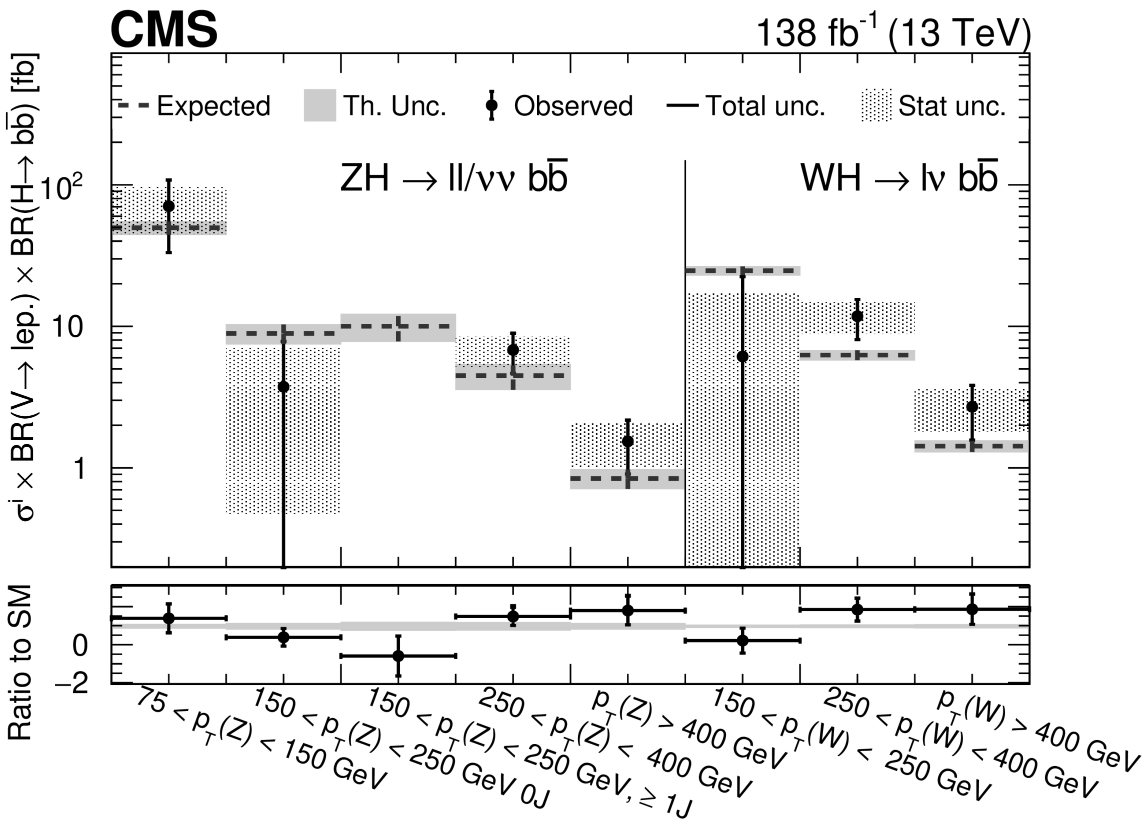
Small production cross section

- H→ bb decay ← largest branching ratio
- V→ leptons ← clean signature

Cross section reported in simplified template cross section framework

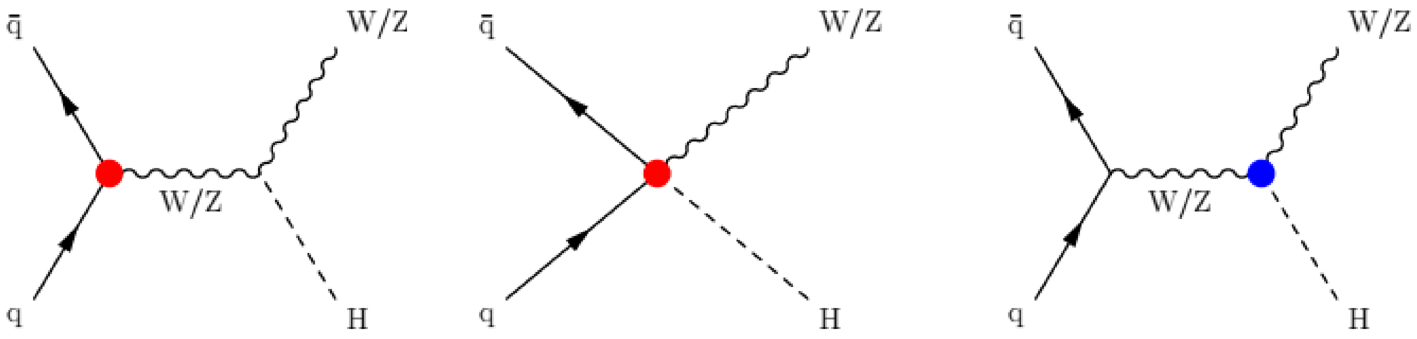






VH signal extracted using DNN score in signal-enriched regions

EFT effects in V associated H production



Rotation to mass eigenstate basis

Current operators

$\mathcal{O}_{\mathrm{Hq}}^{(1)}$	$i\mathrm{H}^{\dagger}\overleftrightarrow{D}_{\mu}\mathrm{H}\bar{\mathrm{q}}_{\mathrm{L}}\gamma^{\mu}\mathrm{q}_{\mathrm{L}}$	$c_{\mathrm{Hq}}^{(1)}$
$\mathcal{O}_{\mathrm{Hq}}^{(3)}$	$i\mathrm{H}^{\dagger}\sigma^{a}\overleftrightarrow{D}_{\mu}\mathrm{H}\bar{\mathrm{q}}_{\mathrm{L}}\sigma^{a}\gamma^{\mu}\mathrm{q}_{\mathrm{L}}$	$c_{\mathrm{Hq}}^{(3)}$
$\mathcal{O}_{\mathrm{Hu}}$	$i\mathrm{H}^{\dagger}\overleftrightarrow{D}_{\mu}\mathrm{H}\bar{\mathrm{u}}_{\mathrm{R}}\gamma^{\mu}\mathrm{u}_{\mathrm{R}}$	c_{Hu}
$\mathcal{O}_{\mathrm{Hd}}$	$i\mathrm{H}^{\dagger}\overleftrightarrow{D}_{\mu}\mathrm{H}\bar{\mathrm{d}}_{\mathrm{R}}\gamma^{\mu}\mathrm{d}_{\mathrm{R}}$	c_{Hd}

Gauge coupling operators

$\mathcal{O}_{\mathrm{HWB}}$	$\mathrm{H}^{\dagger}\sigma^{a}\mathrm{H}\mathrm{W}^{a}_{\mu\nu}\mathrm{B}^{\mu\nu}$	c_{HWB}
$\mathcal{O}_{\mathrm{HW}}$	$(\mathrm{H}^{\dagger}\mathrm{H})\,\mathrm{W}^{a}_{\mu\nu}\mathrm{W}^{a\mu\nu}$	c_{HW}
$\mathcal{O}_{\mathrm{HB}}$	$(\mathrm{H}^{\dagger}\mathrm{H})\,\mathrm{B}_{\mu\nu}\mathrm{B}^{\mu\nu}$	c_{HB}
$\mathcal{O}_{\mathrm{H\widetilde{W}B}}$	$\mathrm{H}^{\dagger}\sigma^{a}\mathrm{H}\mathrm{W}^{a}_{\mu\nu}\widetilde{\mathrm{B}}^{\mu\nu}$	$c_{\mathrm{H\widetilde{W}B}}$
$\mathcal{O}_{\mathrm{H\widetilde{W}}}$	$(\mathrm{H}^{\dagger}\mathrm{H})\,\mathrm{W}^{a}_{\mu\nu}\widetilde{\mathrm{W}}^{a\mu\nu}$	$c_{\mathrm{H\widetilde{W}}}$
$\mathcal{O}_{\mathrm{H\widetilde{B}}}$	$(\mathrm{H}^{\dagger}\mathrm{H})\,\mathrm{B}_{\mu\nu}\widetilde{\mathrm{B}}^{\mu\nu}$	$c_{\mathrm{H\widetilde{B}}}$

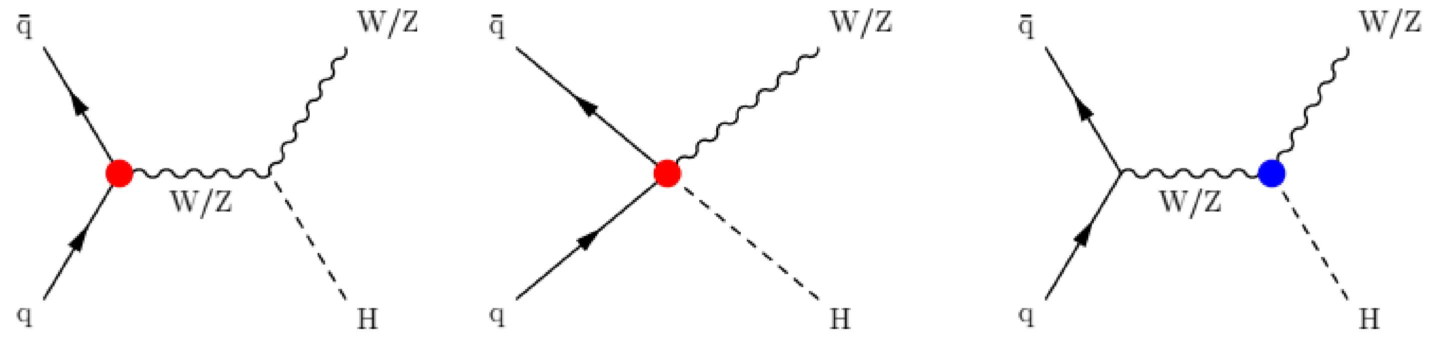
$$\boxed{g_2^{ZZ} = -2\frac{v^2}{\Lambda^2}\left(s_w^2 c_{\mathrm{HB}} + c_w^2 c_{\mathrm{HW}} + s_w c_w c_{\mathrm{HWB}}\right),}$$
$$g_2^{Z\gamma} = -2\frac{v^2}{\Lambda^2}\left(s_w c_w (c_{\mathrm{HW}} - c_{\mathrm{HB}}) + \frac{1}{2}(s_w^2 - c_w^2) c_{\mathrm{HWB}}\right),$$
$$g_2^{\gamma\gamma} = -2\frac{v^2}{\Lambda^2}\left(c_w^2 c_{\mathrm{HB}} + s_w^2 c_{\mathrm{HW}} - s_w c_w c_{\mathrm{HWB}}\right),$$

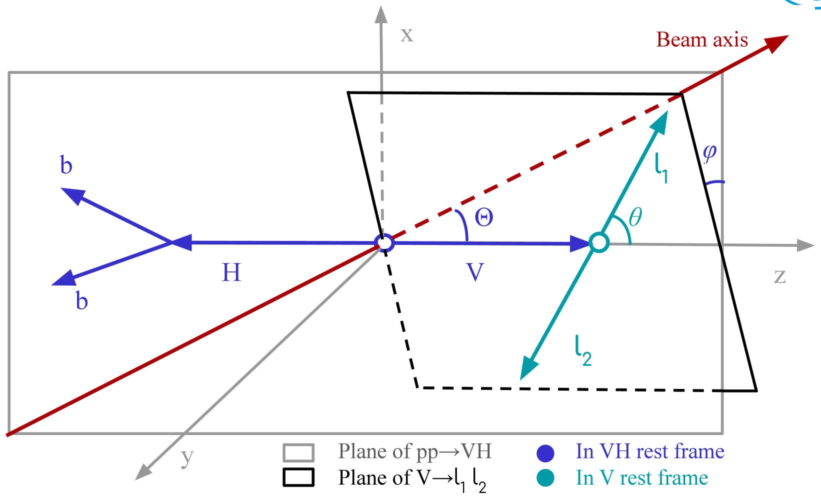
$$c_w = \cos\theta_{\mathrm{W}}$$

$$\boxed{g_4^{ZZ} = \tilde{g}_2^{ZZ} = -2\frac{v^2}{\Lambda^2}\left(s_w^2 c_{\mathrm{H\widetilde{B}}} + c_w^2 c_{\mathrm{H\widetilde{W}}} + s_w c_w c_{\mathrm{H\widetilde{W}B}}\right),}$$
$$g_4^{Z\gamma} = \tilde{g}_2^{Z\gamma} = -2\frac{v^2}{\Lambda^2}\left(s_w c_w (c_{\mathrm{H\widetilde{W}}} - c_{\mathrm{H\widetilde{B}}}) + \frac{1}{2}(s_w^2 - c_w^2) c_{\mathrm{H\widetilde{W}B}}\right),$$
$$g_4^{\gamma\gamma} = \tilde{g}_2^{\gamma\gamma} = -2\frac{v^2}{\Lambda^2}\left(c_w^2 c_{\mathrm{H\widetilde{B}}} + s_w^2 c_{\mathrm{H\widetilde{W}}} - s_w c_w c_{\mathrm{H\widetilde{W}B}}\right).$$

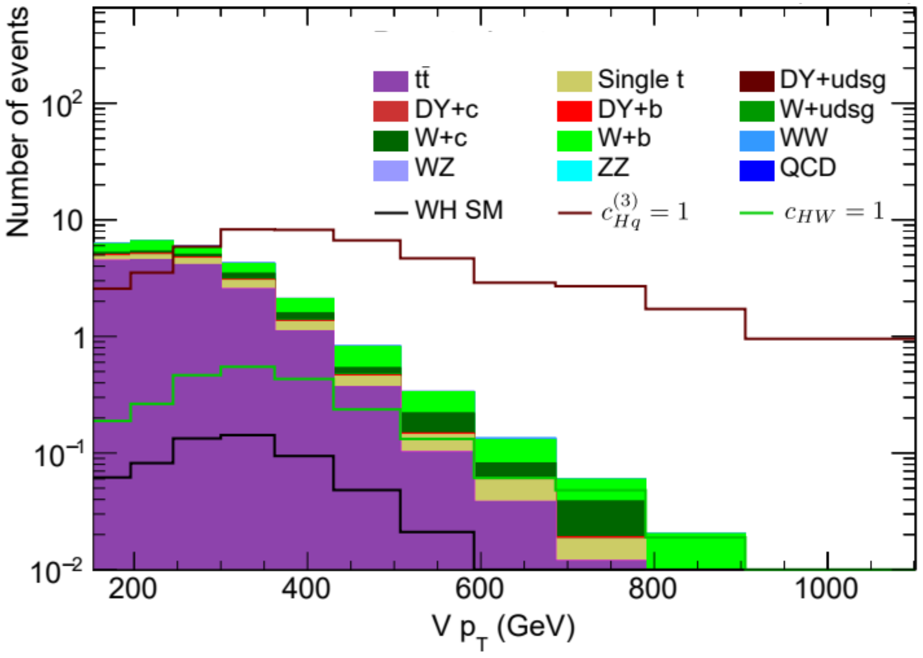
Coefficients targeted in measurement: $c_{\mathrm{Hq}}^{(1)},\ c_{\mathrm{Hq}}^{(3)},\ c_{\mathrm{Hu}},\ c_{\mathrm{Hd}},\ g_2^{ZZ},\ g_4^{ZZ}$

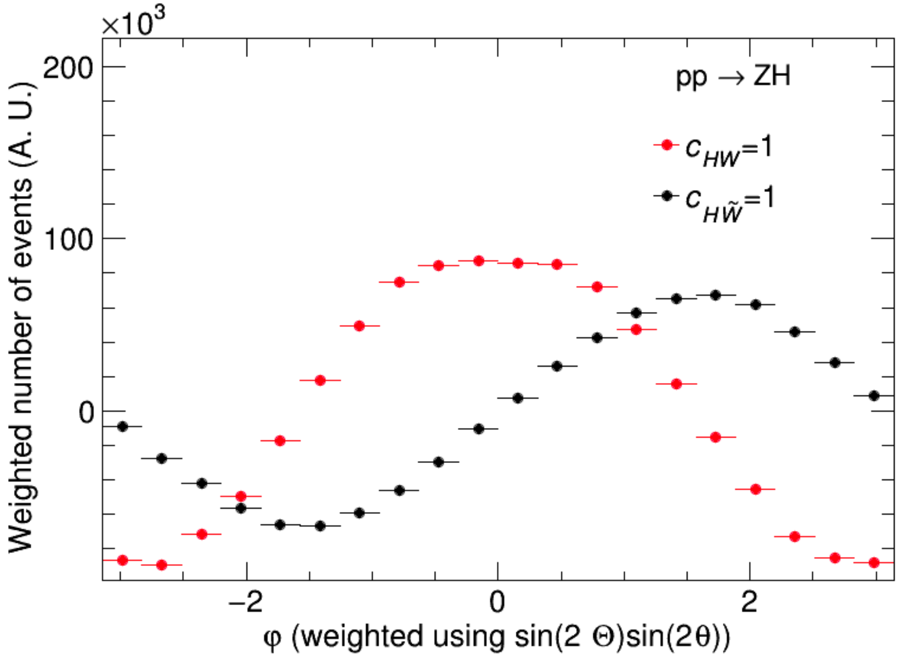
EFT effects in V associated H production





Use of angular variables followed from Banerjee, Gupta, Reiness, Seth, Spannowsky (2019)





EFT effects:

Changing energy spectra

+

Modifying angular correlation

Signal simulation strategy

Polynomial parameterization

$$\mathcal{M}_{\mathrm{SMEFT}} = \mathcal{M}_{\mathrm{SM}} + \sum_i \frac{c_i}{\Lambda^2}\mathcal{M}_{6,i}$$

$$\sigma \sim |\mathcal{M}_{\mathrm{SMEFT}}|^2$$
$$\sim |\mathcal{M}_{\mathrm{SM}}|^2 + \sum_i \frac{c_i}{\Lambda^2} 2\mathrm{Re}(\mathcal{M}^\dagger_{\mathrm{SM}}\mathcal{M}_{6,i}) + \sum_i \frac{c_i^2}{\Lambda^4}|\mathcal{M}_{6,i}|^2 + \sum_i\sum_{j,j>i} \frac{c_i c_j}{\Lambda^4}\mathcal{M}_{6,i}\mathcal{M}_{6,j}$$

σ is a quadratic function of coefficients !

# of signal samples (for 'n' coefficients): N(n) =1 + n + n(n+1)/2 ← Sufficient

Possible to encode SMEFT prediction in event weights

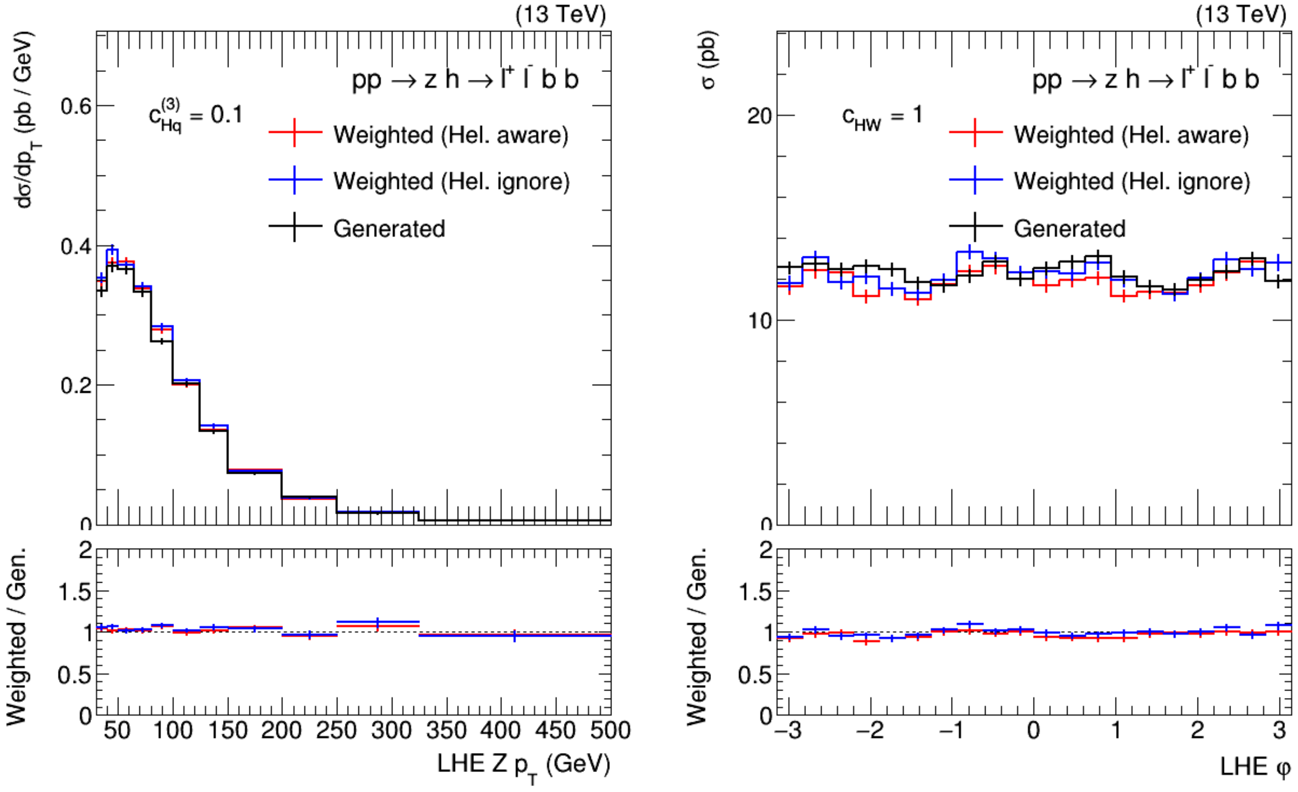
$$\omega = \frac{|\mathcal{M}_{SMEFT}(c = c_1)|^2}{|\mathcal{M}_{SMEFT}(c = c_0)|^2}$$

$$\omega(\boldsymbol{z}|\boldsymbol{c}) \sim \frac{d\sigma}{d\boldsymbol{z}} = \mathrm{Const}(\boldsymbol{z}) + \mathrm{LIN}_a(\boldsymbol{z})c_a + \frac{1}{2}\mathrm{QUAD}_{ab}(\boldsymbol{z})c_a c_b$$

Store N(n) weights per event

→ obtain EFT prediction for any coefficient value

SciPost Phys.Comm.Rep. (2024) 4



Observable: likelihood-free inference

Likelihood ratio trick in classification

$$L = \int dx \sum_{z\in 0,1} p(\boldsymbol{x}, z)\,(z - f(\boldsymbol{x}))^2$$

$0 \to \mathrm{SM},\ 1 \to \theta$

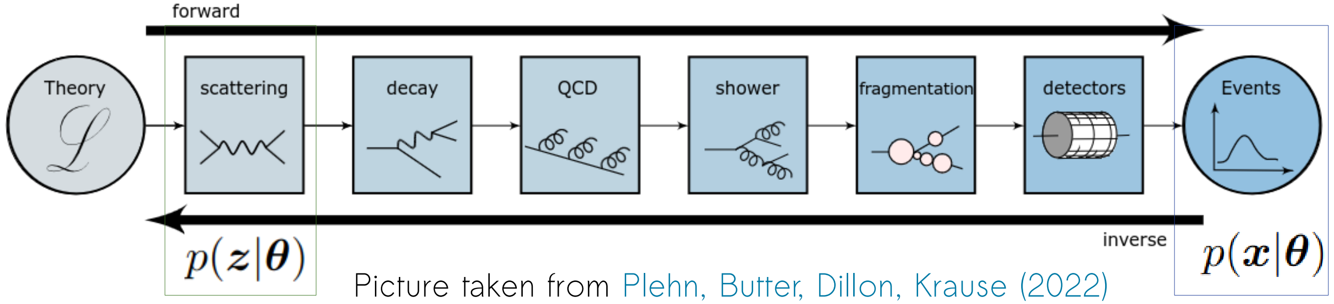
Minimizing L

$$f^*(\boldsymbol{x}) = \frac{p(\boldsymbol{x}|\mathrm{SM})}{p(\boldsymbol{x}|\mathrm{SM}) + p(\boldsymbol{x}|\theta)} = \frac{1}{1 + \frac{\sigma(\theta)}{\sigma(\mathrm{SM})} r(\boldsymbol{x})}$$

Optimal test statistic:
Likelihood ratio (LLR)

$$r(\boldsymbol{x}) = \frac{p(\boldsymbol{x}|\theta)}{p(\boldsymbol{x}|\mathrm{SM})} \to R(\boldsymbol{x}|\boldsymbol{\theta}_1, \boldsymbol{\theta}_0)$$

$\boldsymbol{x} \leftarrow$ observable



Picture taken from Plehn, Butter, Dillon, Krause (2022)

$$p(x|\boldsymbol{\theta}) = \int\int\int p(x|z_{\mathrm{Det}})p(z_{\mathrm{Det}}|z_{\mathrm{Had}})p(z_{\mathrm{Had}}|z_{\mathrm{PS}})p(z_{\mathrm{PS}}|z)dz_{\mathrm{Det}}dz_{\mathrm{Had}}dz_{\mathrm{PS}}\,p(z|\boldsymbol{\theta})dz$$

$$p(\boldsymbol{x}|\boldsymbol{\theta}) = \int p(\boldsymbol{x}|\boldsymbol{z})p(\boldsymbol{z}|\boldsymbol{\theta})dz$$

intractable
tractable

Inverse problem: Estimate θ from p(x)

LLR between two EFT hypotheses → learned by regressing on joint LLR

Brehmer, Cranmer, Louppe, Pavez (2018)

$$R(\boldsymbol{x}, \boldsymbol{z}|\boldsymbol{\theta}_1, \boldsymbol{\theta}_0) = \frac{p(\boldsymbol{x}, \boldsymbol{z}|\boldsymbol{\theta}_1)}{p(\boldsymbol{x}, \boldsymbol{z}|\boldsymbol{\theta}_0)} = \frac{p(\boldsymbol{x}|\boldsymbol{z})p(\boldsymbol{z}|\boldsymbol{\theta}_1)}{p(\boldsymbol{x}|\boldsymbol{z})p(\boldsymbol{z}|\boldsymbol{\theta}_0)} = \frac{p(\boldsymbol{z}|\boldsymbol{\theta}_1)}{p(\boldsymbol{z}|\boldsymbol{\theta}_0)}$$

$$L[\hat{F}] = \int dx dz\, p(\boldsymbol{x}, \boldsymbol{z}|\boldsymbol{\theta}_0)|F(\boldsymbol{x}, \boldsymbol{z}) - \hat{F}(x)|^2$$

$$F(\boldsymbol{x}, \boldsymbol{z}) = \frac{\sigma(\boldsymbol{\theta}_1)}{\sigma(\boldsymbol{\theta}_0)} R(\boldsymbol{x}, \boldsymbol{z}) = \frac{\sigma(\boldsymbol{\theta}_1)}{\sigma(\boldsymbol{\theta}_0)} \frac{p(\boldsymbol{x}, \boldsymbol{z}|\boldsymbol{\theta}_1)}{p(\boldsymbol{x}, \boldsymbol{z}|\boldsymbol{\theta}_0)}$$

$$\hat{R}(\boldsymbol{x}) = \frac{\int dz\, p(\boldsymbol{x}, \boldsymbol{z}|\boldsymbol{\theta}_0)\frac{\sigma(\boldsymbol{\theta}_1)}{\sigma(\boldsymbol{\theta}_0)}\frac{p(\boldsymbol{x}, \boldsymbol{z}|\boldsymbol{\theta}_1)}{p(\boldsymbol{x}, \boldsymbol{z}|\boldsymbol{\theta}_0)}}{\int dz\, p(\boldsymbol{x}, \boldsymbol{z}|\boldsymbol{\theta}_0)} = \frac{\sigma(\boldsymbol{\theta}_1)}{\sigma(\boldsymbol{\theta}_0)}\frac{p(\boldsymbol{x}|\boldsymbol{\theta}_1)}{p(\boldsymbol{x}|\boldsymbol{\theta}_0)} = R(\boldsymbol{x}|\boldsymbol{\theta}_1, \boldsymbol{\theta}_0)$$

Likelihood-free / simulation-based inference

Boosted information tree (BIT)

Using decision trees to regress on learn likelihood ratio terms
→ very fast!

SC, N. Frohner, L. Lechner, R. Schoefbeck, D. Schwarz (2021)

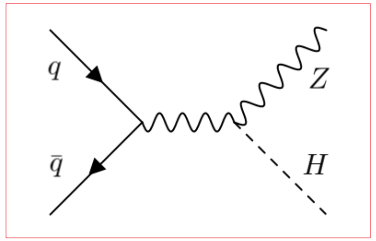
# of trainings = N + N + N(N-1)/2

$$R(\boldsymbol{x}|\boldsymbol{\theta}_1, \boldsymbol{\theta}_0) = 1 + (\theta_1 - \theta_0)_a R_a(\boldsymbol{x}) + \frac{1}{2}(\theta_1 - \theta_0)_a(\theta_1 - \theta_0)_b R_{a,b}(\boldsymbol{x})$$

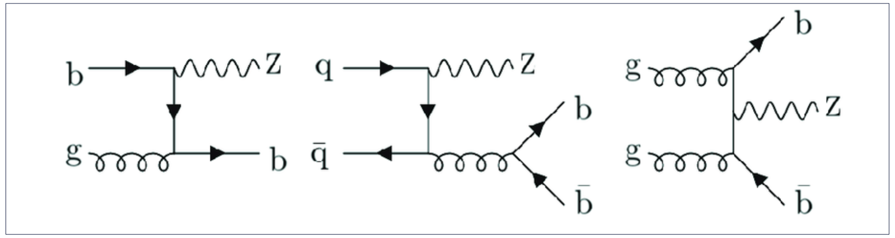
Learning SMEFT likelihood ratio

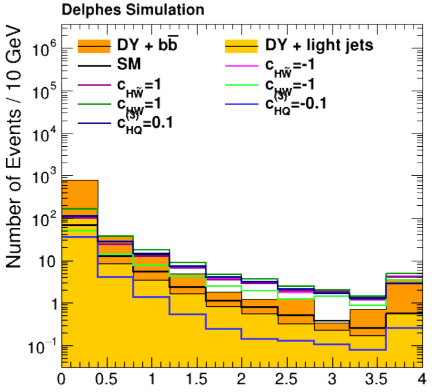
SC, S. Roshap, R. Schoefbeck, D. Schwarz (2022)

Signal



Background

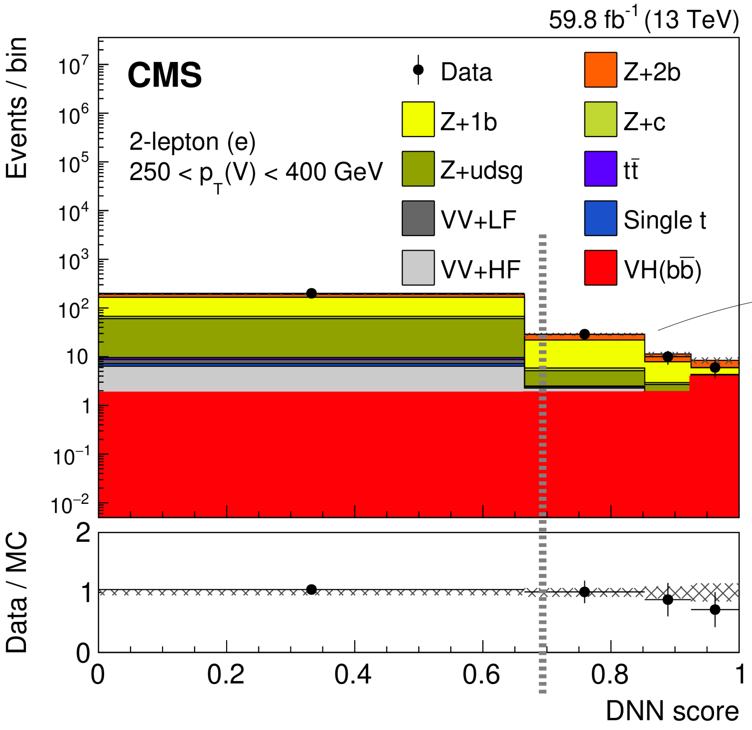


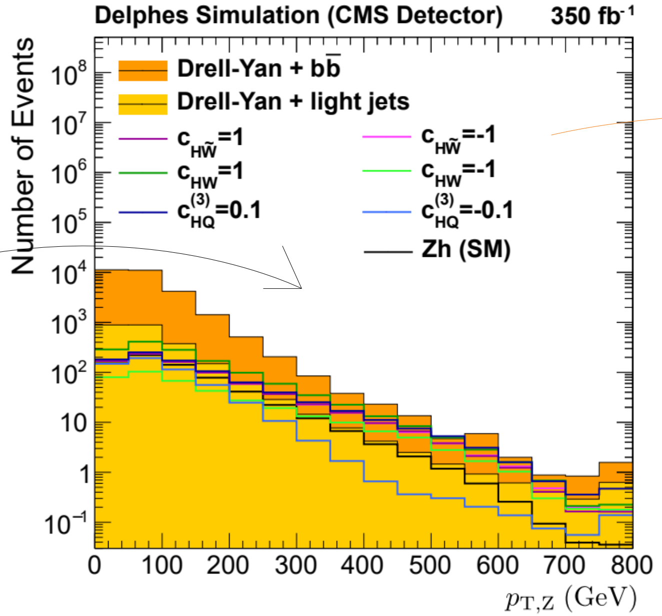


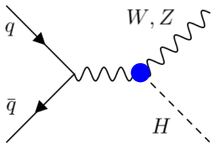
BIT approach

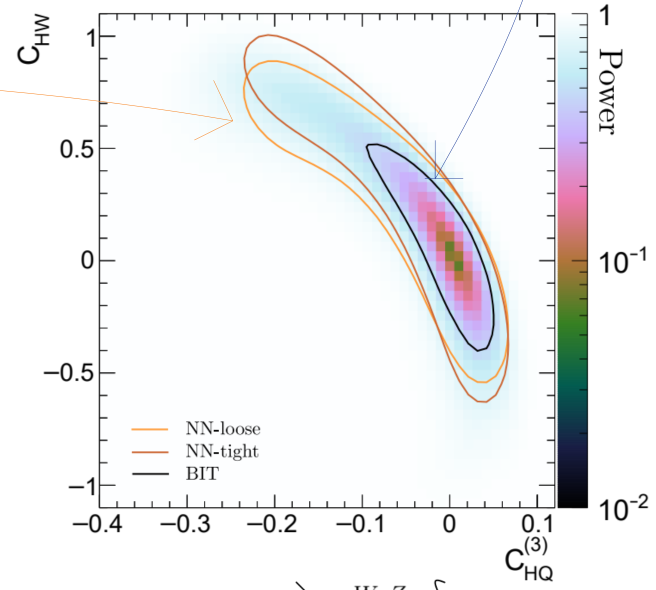
$$R(\boldsymbol{x}|\boldsymbol{\theta}_1,\boldsymbol{\theta}_0) = 1 + (\theta_1-\theta_0)_a R_a(\boldsymbol{x}) + \frac{1}{2}(\theta_1-\theta_0)_a(\theta_1-\theta_0)_b R_{a,b}(\boldsymbol{x})$$

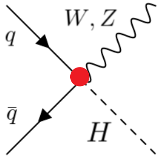
Traditional approach







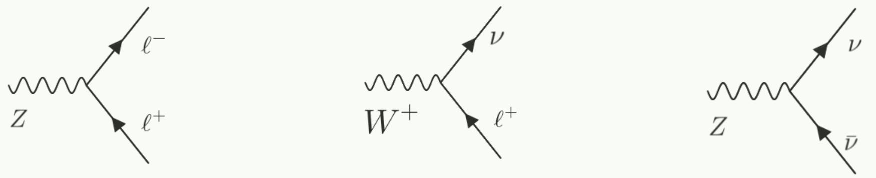




Event selection & categorization

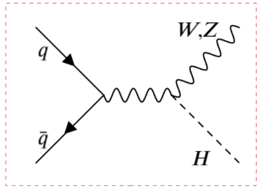
Final state

2-lepton 1-lepton 0-lepton



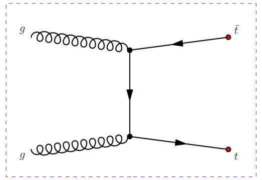
Regions

Signal region (SR)

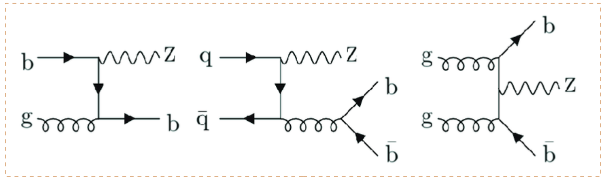


Control regions (CRs)

$t\bar{t}$



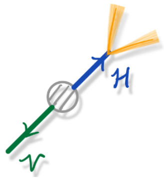
V+ heavy-flavor jets

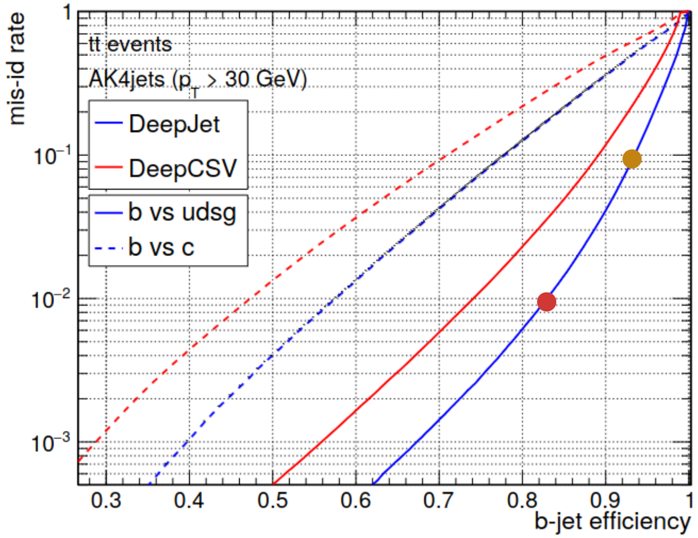


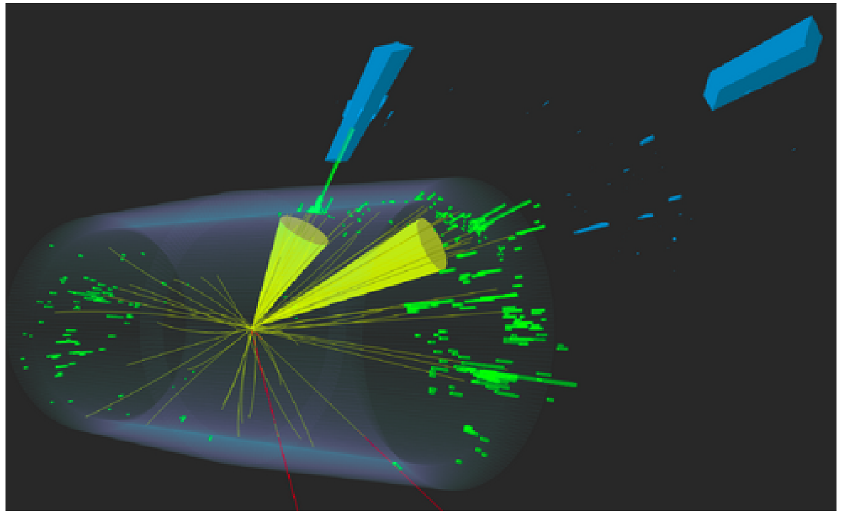
Used to measure background

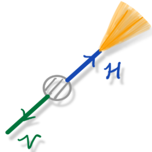
Bols, Kieseler, Verzetti, Stoye, Stakia (2020)

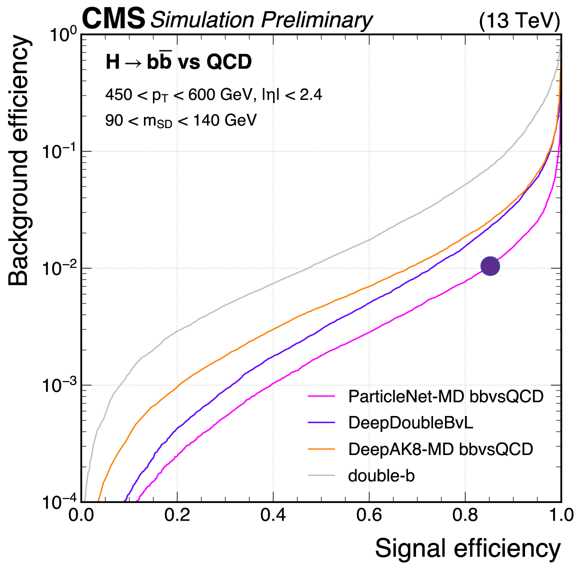
CMS-PAS-BTV-22-001











BIT training & template optimization

arXiv: 2411.16907

<u>Objective:</u>

SMEFT effects *vs*
- SM VH
- Background

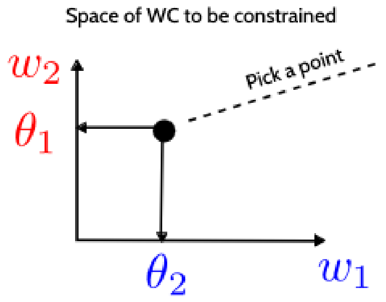
<u>Inputs:</u>

- Object kinematics
- Jet tagging scores
- Angular variables

<u>Template optimization</u>

Optimal performance guaranteed
if binning is chosen at each point in EFT space

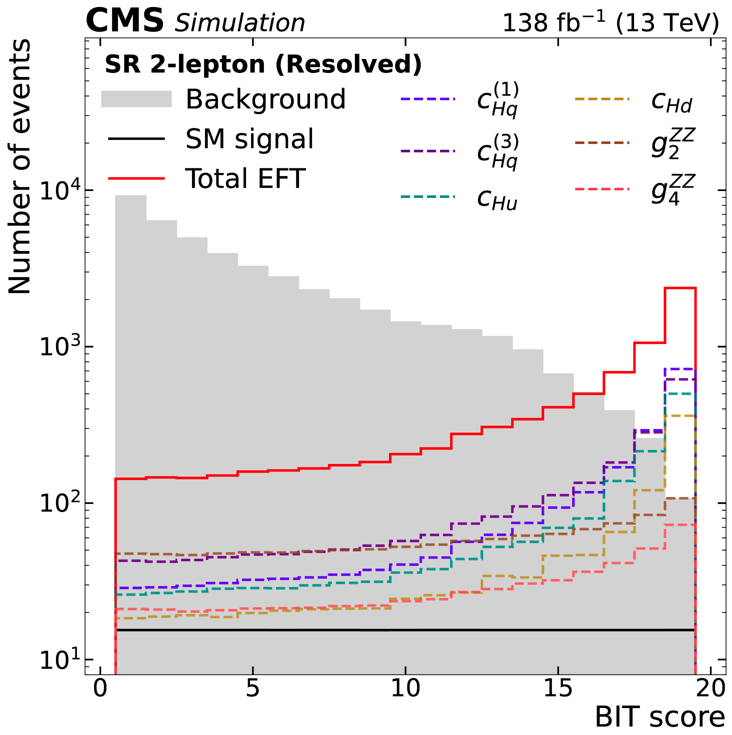
very challenging

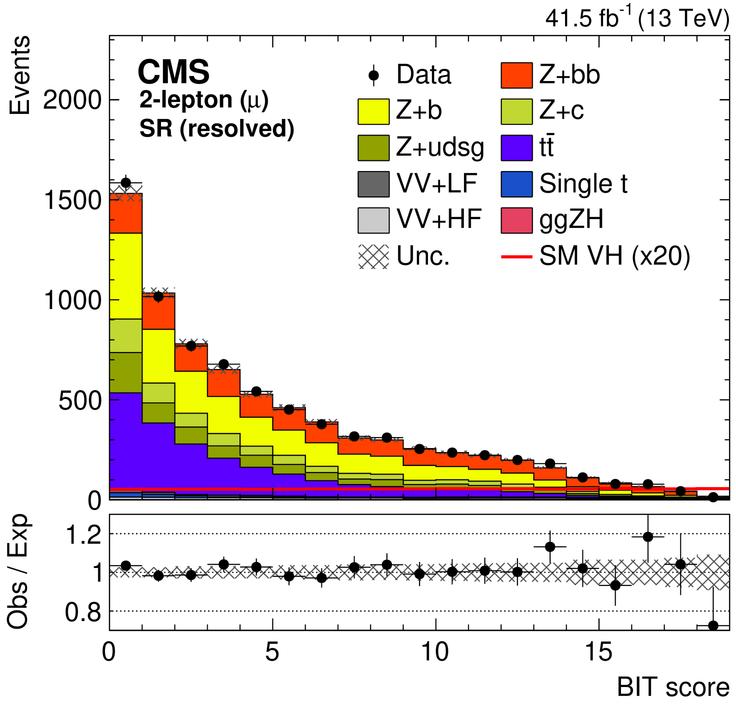


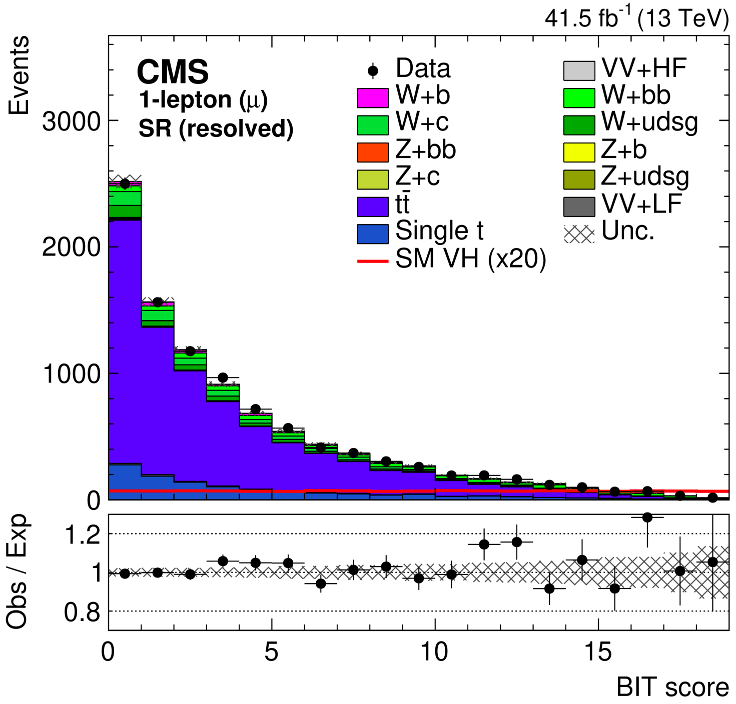
No EFT effects considered in background

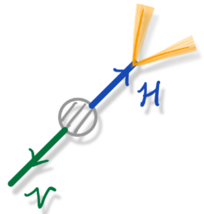
Template is chosen minimizing $\prod$ 2σ intervals (profiled)

← Find the point in EFT space









BIT training & template optimization

arXiv: 2411.16907

Objective:

SMEFT effects *vs*
- SM VH
- Background

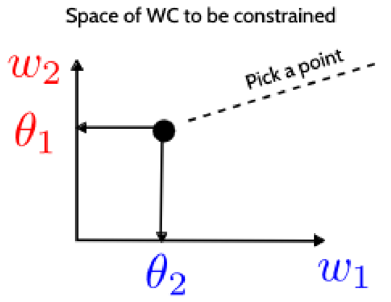
Inputs:

- Object kinematics
- Jet tagging scores
- Angular variables

Template optimization

Optimal performance guaranteed
if binning is chosen at each point in EFT space

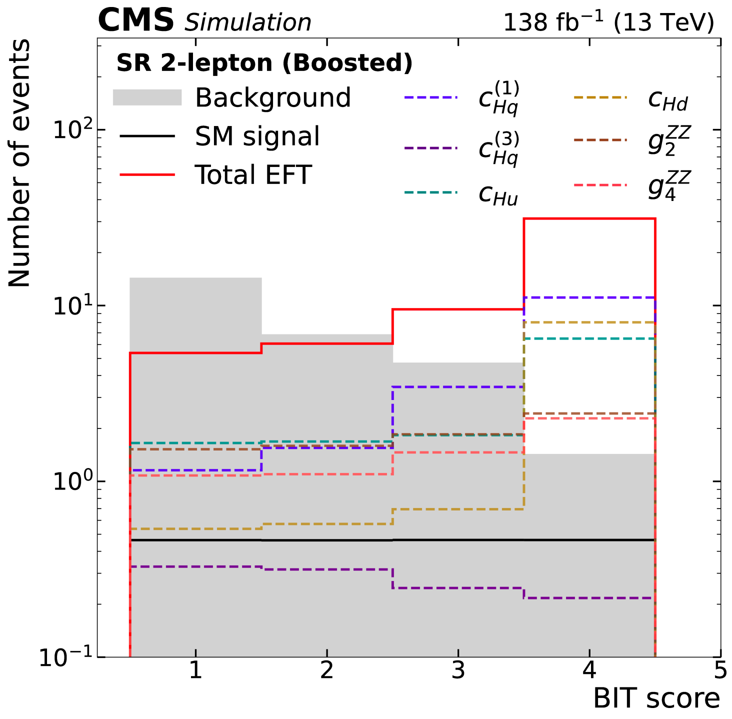
very challenging

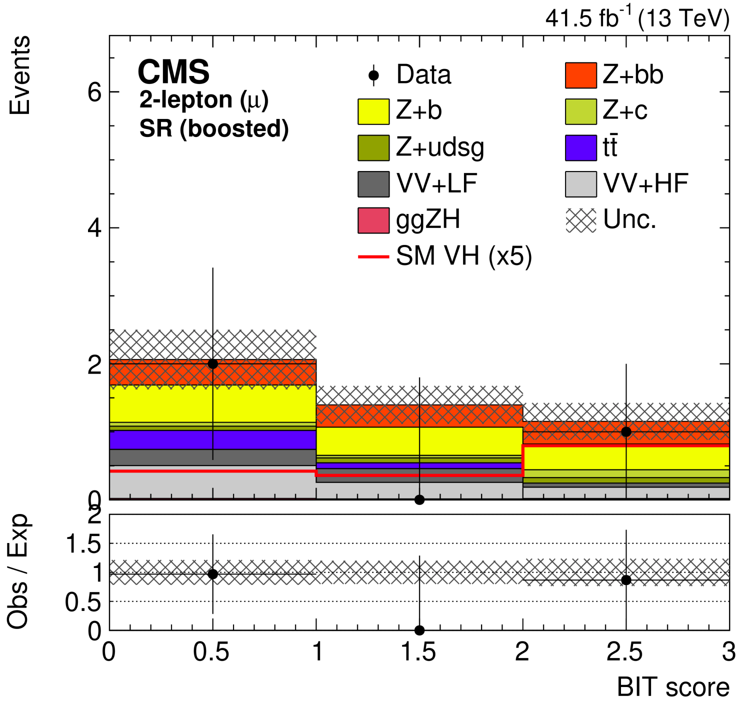


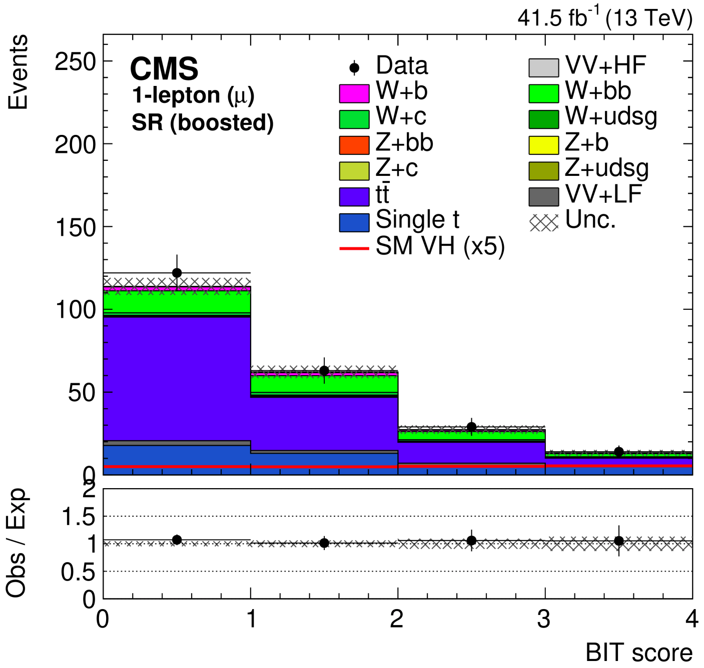
No EFT effects considered in background

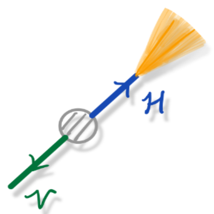
Template is chosen minimizing ∏ 2σ intervals (profiled)

← Find the point in EFT space









EFT interpretation: Wilson coefficients

With linear or full quadratic expansion

$$\sigma \sim |\mathcal{M}_{\mathrm{SMEFT}}|^2$$
$$\sim \boxed{|\mathcal{M}_{\mathrm{SM}}|^2 + \sum_i \frac{c_i}{\Lambda^2} 2\mathrm{Re}(\mathcal{M}_{\mathrm{SM}}^\dagger \mathcal{M}_{6,i})} + \sum_i \frac{c_i^2}{\Lambda^4}|\mathcal{M}_{6,i}|^2 + \sum_i \sum_{j,j>i} \frac{c_i c_j}{\Lambda^4}\mathcal{M}_{6,i}\mathcal{M}_{6,j}$$

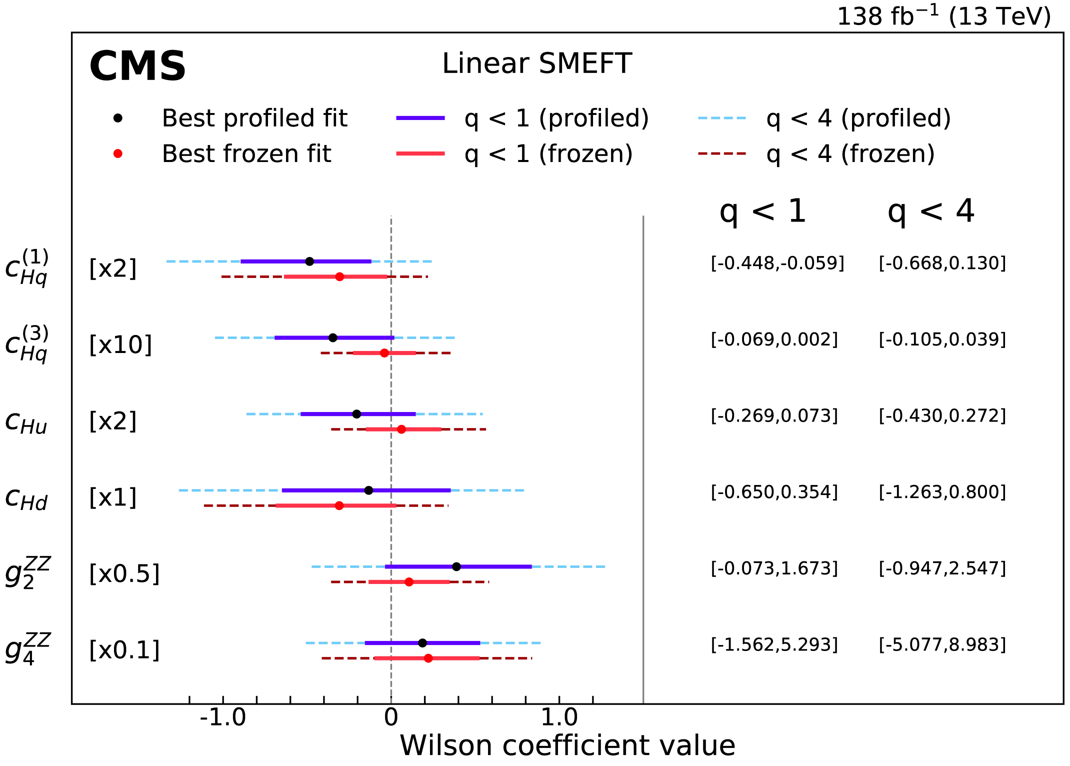
Λ = 1 TeV

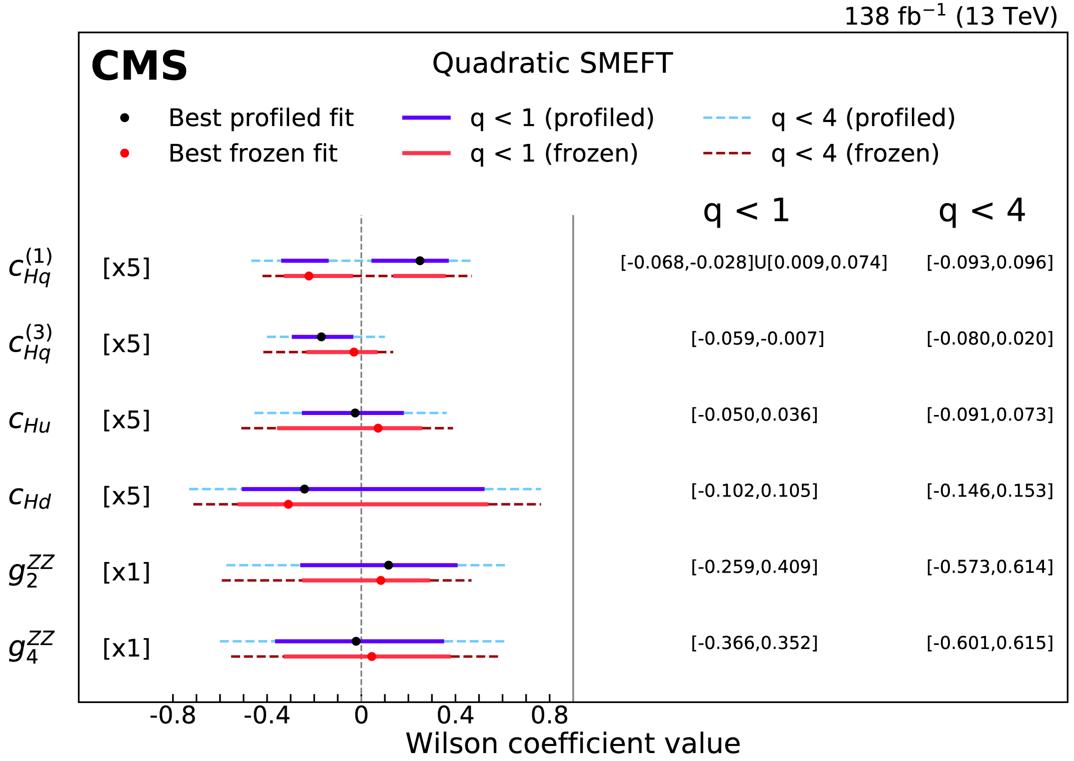
Compatibility with SM while varying all coefficients together

arXiv: 2411.16907

p-value = 73%

p-value = 84%





Quadratic component dominates sensitivity for most coefficients

EFT interpretation: energy scale

With linear or full quadratic expansion

$$\sigma \sim |\mathcal{M}_{\text{SMEFT}}|^2$$

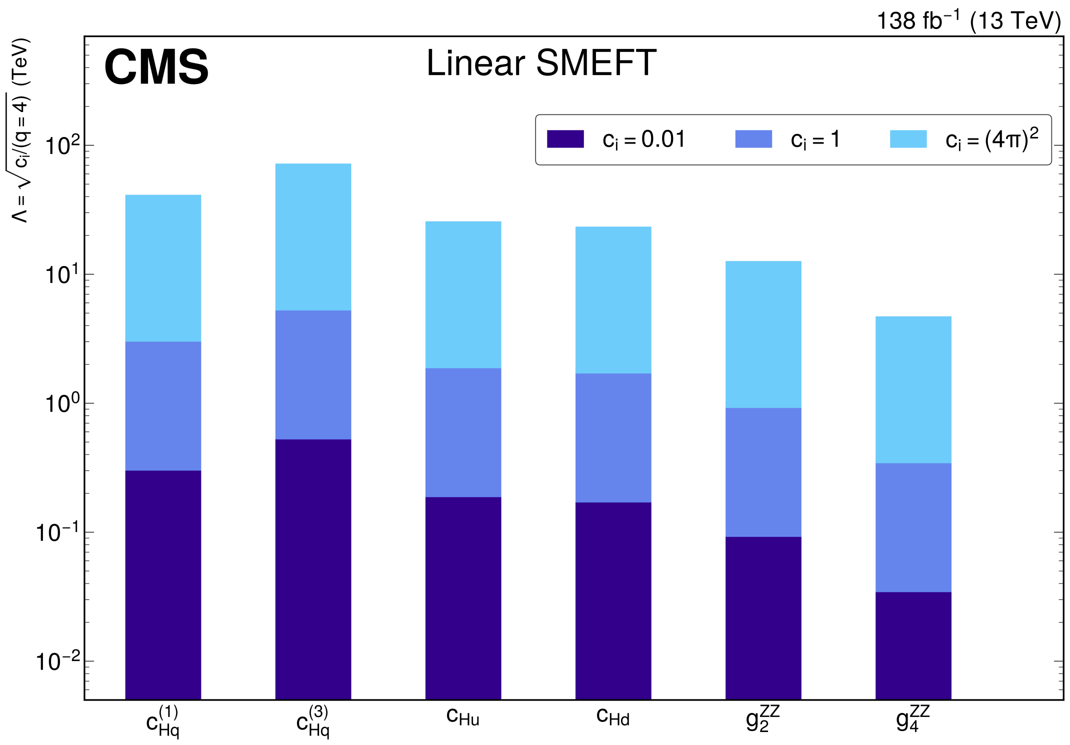
$$\sim \boxed{|\mathcal{M}_{\text{SM}}|^2 + \sum_i \frac{c_i}{\Lambda^2} 2\text{Re}(\mathcal{M}^\dagger_{\text{SM}}\mathcal{M}_{6,i})} + \sum_i \frac{c_i^2}{\Lambda^4}|\mathcal{M}_{6,i}|^2 + \sum_i \sum_{j,j>i} \frac{c_i c_j}{\Lambda^4}\mathcal{M}_{6,i}\mathcal{M}_{6,j}$$

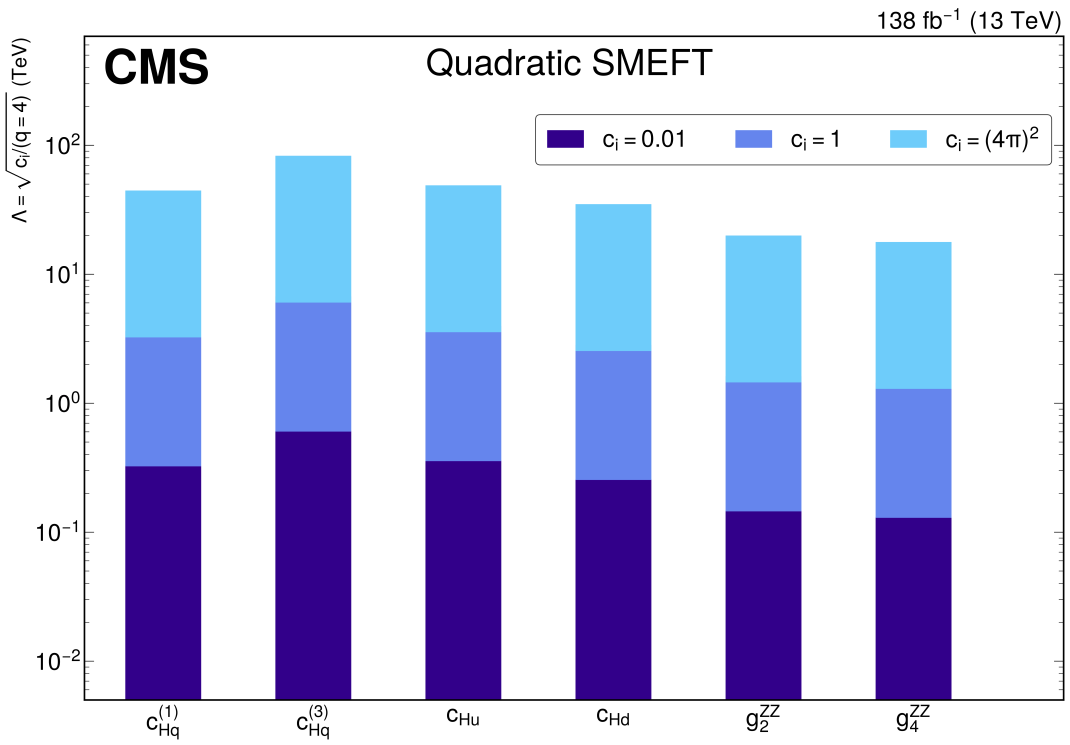
Lower bound on Λ

Three BSM scenarios:

- weakly coupled model (c=0.01)
- no assumption (c=1)
- strongly-coupling perturbativity limit (c=$16\pi^2$)

arXiv: 2411.16907





Sensitivity: ~ few TeV for c=1

EFT interpretation: Wilson coefficients in 2D

With
full quadratic
expansion

$$\sigma \sim |\mathcal{M}_{\mathrm{SMEFT}}|^2$$

$$\boxed{\sim |\mathcal{M}_{\mathrm{SM}}|^2 + \sum_i \frac{c_i}{\Lambda^2} 2\mathrm{Re}(\mathcal{M}_{\mathrm{SM}}^{\dagger}\mathcal{M}_{6,i}) + \sum_i \frac{c_i^2}{\Lambda^4}|\mathcal{M}_{6,i}|^2 + \sum_i \sum_{j,j>i} \frac{c_i c_j}{\Lambda^4}\mathcal{M}_{6,i}\mathcal{M}_{6,j}}$$

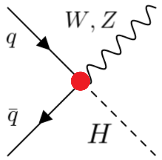
Λ = 1 TeV

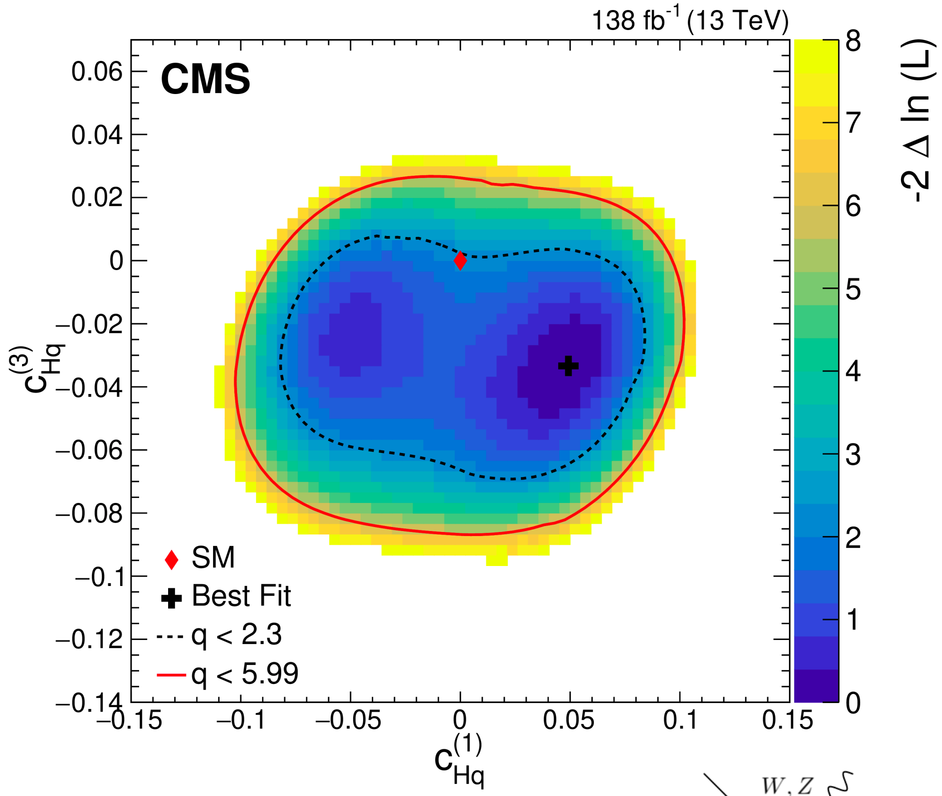
Other coefficients profiled

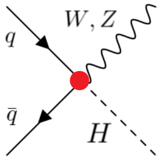
Vector coupling
vs.
Vector coupling

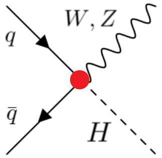
Other coefficients fixed at 0

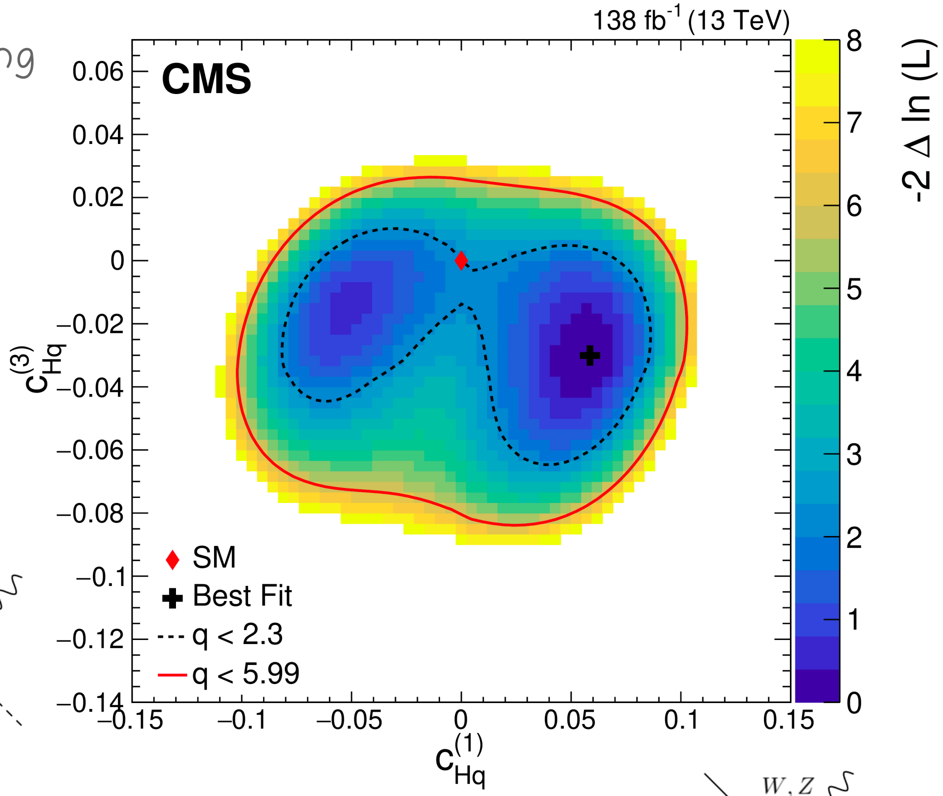
arXiv: 2411.16907

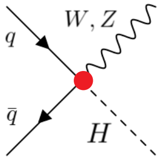












EFT interpretation: Wilson coefficients in 2D

With
full quadratic
expansion

$$\sigma \sim |\mathcal{M}_{\mathrm{SMEFT}}|^2$$

$$\boxed{\sim |\mathcal{M}_{\mathrm{SM}}|^2 + \sum_i \frac{c_i}{\Lambda^2} 2\mathrm{Re}(\mathcal{M}_{\mathrm{SM}}^{\dagger}\mathcal{M}_{6,i}) + \sum_i \frac{c_i^2}{\Lambda^4}|\mathcal{M}_{6,i}|^2 + \sum_i \sum_{j,j>i} \frac{c_i c_j}{\Lambda^4}\mathcal{M}_{6,i}\mathcal{M}_{6,j}}$$

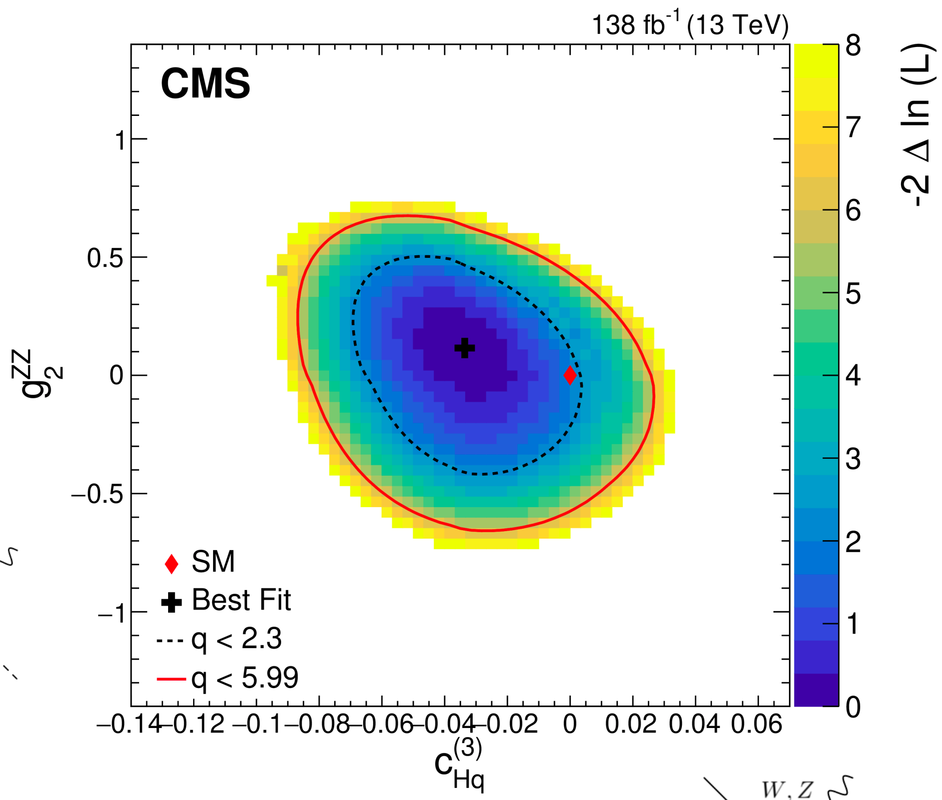
Λ = 1 TeV

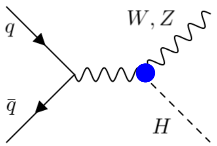
Other coefficients profiled

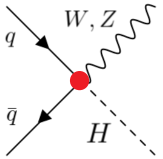
Vector coupling
vs.
Gauge coupling

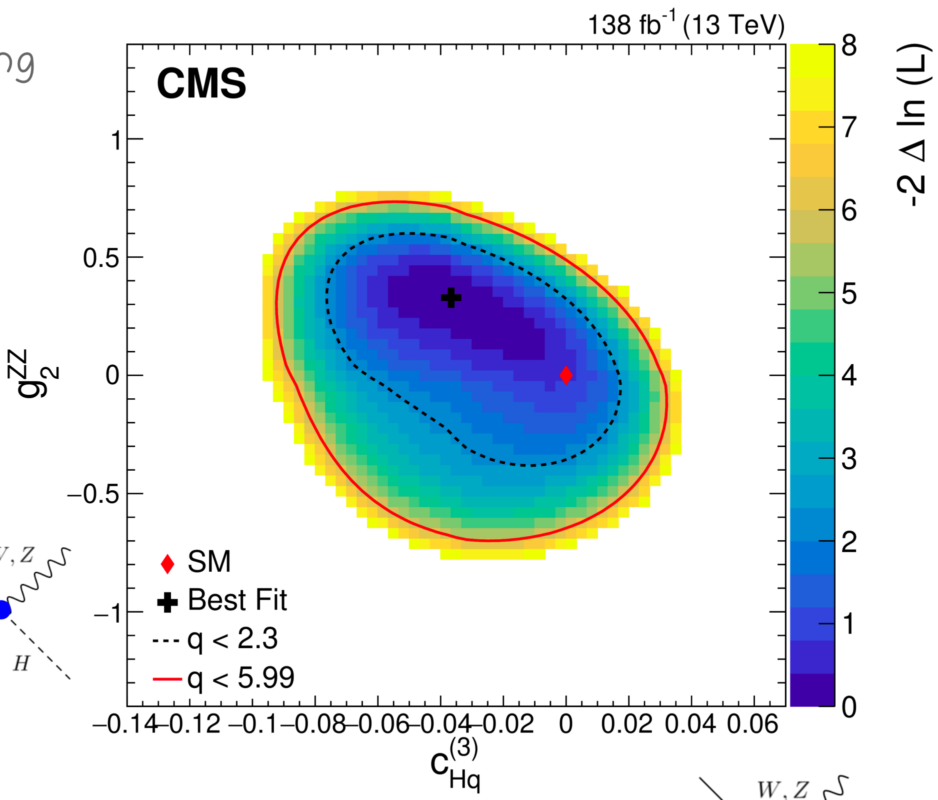
Other coefficients fixed at 0

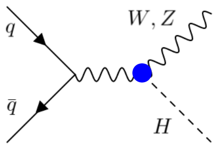
arXiv: 2411.16907

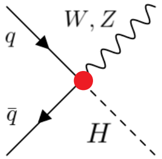












Highlights

In K. Cranmer's DESY Colloquium

ATLAS CONF Note

ATLAS-CONF-2024-015

28th October 2024

An implementation of Neural Simulation-Based Inference for Parameter Estimation in ATLAS

The ATLAS Collaboration

Neural Simulation-Based Inference (NSBI) is a powerful class of machine learning (ML)-based methods for statistical inference that naturally handles high-dimensional parameter estimation without the need to bin data into low-dimensional summary histograms. Such methods are promising for a range of measurements, including at the Large Hadron Collider (LHC), where no single observable may be optimal to scan over the entire theoretical phase space under consideration, or where binning data into histograms could result in a loss of sensitivity. This work develops an NSBI framework for statistical inference, using neural networks to estimate probability density ratios, which enables the application of NSBI to a full-scale LHC analysis. It incorporates a large number of systematic uncertainties, quantifies the uncertainty coming from finite training statistics, develops a method to construct confidence intervals, and demonstrates a series of intermediate diagnostic checks that can be performed to validate the robustness of the method. As an example, the power and feasibility of the method are demonstrated on simulated data for a simplified version of an off-shell Higgs boson couplings measurement in the four-leptons final states. This NSBI framework is an extension of the standard statistical framework used by LHC experiments and can benefit a large number of physics analyses.

ATLAS-CONF-2024-015
28 October 2024

Available on the CERN CDS information server **CMS PAS HIG-23-016**

CMS Physics Analysis Summary

Contact: cms-pag-conveners-higgs@cern.ch 2024/09/20

Constraints on standard model effective field theory effects with Higgs bosons produced in association with W or Z bosons in the $H \to b\bar{b}$ decay channel in proton-proton collisions at $\sqrt{s} = 13$ TeV

The CMS Collaboration

Abstract

A standard model effective field theory (SMEFT) analysis with dimension-six operators is performed in the Higgsstrahlung process, where the Higgs boson is produced in association with a W or Z boson, in proton-proton collisions at a center-of-mass energy of 13 TeV. The final states where the W or Z boson decay leptonically and the Higgs boson decays to a pair of bottom quarks are considered. The analyzed data were collected by the CMS experiment between 2016 and 2018 and correspond to an integrated luminosity of 138 fb^{-1}. An approach targeted to optimize simultaneously the sensitivity to Wilson coefficients of multiple SMEFT operators is employed. The observed results are consistent with the predictions of the standard model.

ATLAS used simulation-based inference for off-shell H production

Neural network-based approach: arXiv: 2412.01548

Summary

- Detailed effective field theory analysis in Higgs production in association with W/Z boson with full Run 2 data

  - *Follow-up & complementary to cross section measurements*

- Probed effects of both vector and gauge coupling operators

- First application of likelihood-free/simulation-based inference at colliders

  - *Developed outside collaboration & technology available for any EFT analysis*

  --- *Road towards fully unbinned EFT analysis using optimal observables* [see Schöfbeck (2024)]

- Results reported as constraints on Wilson coefficients

  - *in 1D with other coefficients profiled or set to SM values*

  - *in 2D with other coefficients profiled or set to SM values*

  ← *findings consistent with SM*

- Best SMEFT sensitivity reported in *VH(bb) channel* till date

Documentation

arXiv: 2411.16907

(Accepted in JHEP)

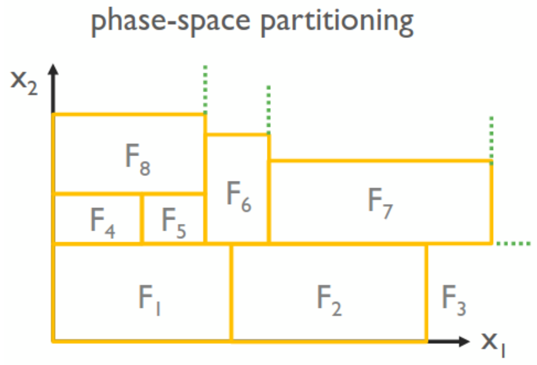
Extra Material

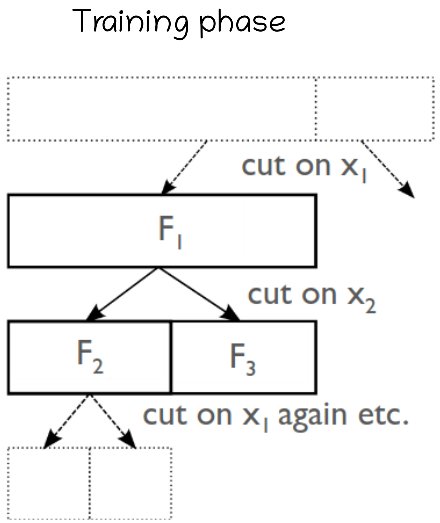
Learning SMEFT likelihood with decision trees (1)

Tree prediction $\hat{F}(\mathbf{x}) = \sum_{j\in\mathcal{J}} \mathbb{1}_{\alpha_j}(\mathbf{x}) F_j$

Phase space partitioning — Prediction

x: Feature vectors
j: Terminal nodes
α_j : Requirements on x for node j
F_j : Prediction for node j





Minimization of loss function w.r.t. α_j and F_j

Linear term in SMEFT expansion

$$\mathrm{MSE}[\hat{F}_a] = \sum_{j\in\mathcal{J}}\sum_{i\in j} w_i \left|\frac{w_{i,a}}{w_i} - F_j\right|^2 = \sum_{i=1}^{N_{\mathrm{sim}}} \frac{w_{i,a}^2}{w_i} - 2\sum_{j\in\mathcal{J}} F_j \sum_{i\in j} w_{i,a} + \sum_{j\in\mathcal{J}} F_j^2 \sum_{i\in j} w_i.$$

Integral replaced by summation

$$\lambda_j = \sum_{i\in j}\omega_i$$

$$\frac{\partial}{\partial F_j}\mathrm{MSE}[\hat{F}_a] = 0 \longrightarrow \quad F_j = \frac{\sum_{i\in j} w_{i,a}}{\sum_{i\in j} w_i} = \left.\frac{\partial_a \lambda_j}{\lambda_j}\right|_{\boldsymbol{\theta}=\boldsymbol{\theta}_0} = \partial_a \log\lambda_j\Big|_{\boldsymbol{\theta}=\boldsymbol{\theta}_0}$$

$$\mathrm{MSE}[\hat{F}_a] = -\sum_{j\in\mathcal{J}} \frac{\left(\sum_{i\in j} w_{i,a}\right)^2}{\sum_{i\in j} w_i} = -\sum_{j\in\mathcal{J}} \left.\frac{(\partial_a\lambda_j)^2}{\lambda_j}\right|_{\boldsymbol{\theta}=\boldsymbol{\theta}_0} = -\sum_{j\in\mathcal{J}} I^{(\lambda_j)}$$

Fisher information for measurement of θ

Fisher information = Variance of score (= derivative of log-likelihood)

Quadratic term in SMEFT expansion

$$F_j = \frac{\sum_{i\in j} w_{i,ab}}{\sum_{i\in j} w_i} = \left.\frac{\partial_a\partial_b\lambda_j}{\lambda_j}\right|_{\boldsymbol{\theta}=\boldsymbol{\theta}_0}.$$

$$\mathrm{MSE}[\hat{F}_{ab}] = -\sum_{j\in\mathcal{J}} \frac{\left(\sum_{i\in j} w_{i,ab}\right)^2}{\sum_{i\in j} w_i} = -\sum_{j\in\mathcal{J}} \left.\frac{(\partial_a\partial_b\lambda_j)^2}{\lambda_j}\right|_{\boldsymbol{\theta}=\boldsymbol{\theta}_0}$$

Cramér-Rao bound: $\mathrm{Variance}(\hat{\theta}) \geq \frac{1}{I(\theta)}$ wiki

Node-split criterion maximizes Fisher information $\rightarrow$ Optimal in precision

$$\lambda_j = b_j + \theta s_j \qquad \rho_j = \frac{\theta s_j}{b_j + \theta s_j}$$

$$L = -\sum_j \frac{(\partial_\theta \lambda_j)^2}{\lambda_j} = -\sum_j \frac{s_j^2}{b_j + \theta s_j} = -\frac{1}{\theta^2}\sum_j \lambda_j \rho^2 \equiv \sum_j \lambda_j \rho(1-\rho)$$

Starting point of SC, N. Frohner, L. Lechner, R. Schoefbeck, D. Schwarz (2021)

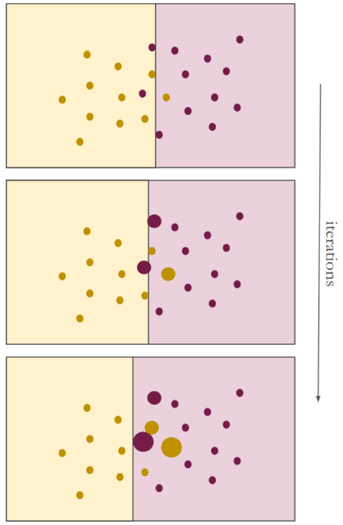
Gini index implemented in TMVA for classification

Learning SMEFT likelihood with decision trees (2)

Boosting: Provides a strong learner by iteratively training an ensemble of weak learners to pseudo-residuals of previous iteration

$$\hat{F}^b(\boldsymbol{x}) = \hat{f}^b(\boldsymbol{x}) + \eta \hat{F}^{b-1}(\boldsymbol{x})$$

Minimize loss function loss w.r.t. *f(x)* ← Goes on till a pre-defined number B



$$\mathrm{MSE}[\hat{F}_a] = \sum_{j\in\mathcal{J}}\sum_{i\in j} w_i \left|\frac{w_{i,a}}{w_i} - \hat{f}^b(\boldsymbol{x}) - \eta\hat{F}^{b-1}(\boldsymbol{x})\right|^2 = \sum_{j\in\mathcal{J}}\sum_{i\in j} w_i \left|\frac{w_{i,a} - \eta\hat{F}^{b-1}(\boldsymbol{x})w_i}{w_i} - \hat{f}^b(\boldsymbol{x})\right|^2$$

Weak learner needs to fit w - ηF ← Target needs to be updated in each iteration

Final outcome of algorithm $\hat{R}(\boldsymbol{x}|\boldsymbol{\theta},\boldsymbol{\theta}_0) = 1 + (\theta-\theta_0)_a \hat{F}_a^{(B)}(\boldsymbol{x}) + \frac{1}{2}(\theta-\theta_0)_a(\theta-\theta_0)_b \hat{F}_{ab}^{(B)}(\boldsymbol{x})$ Boosted information tree (BIT)

SC, S. Roshap, R. Schoefbeck, D. Schwarz (2022)

Separate training for each linear (‘a’) & quadratic terms (‘ab’) → Total # of trainings = n + n(n+1)/2

LLR to achieve

$$q(\mathcal{D}(N_{\mathrm{obs}},\boldsymbol{x})|\boldsymbol{\theta}_1,\boldsymbol{\theta}_0) = -2\left[\mathcal{L}(\sigma(\boldsymbol{\theta}_1)-\sigma(\boldsymbol{\theta}_0)) - \sum_{i=1}^{N_{\mathrm{obs}}} \ln R(\boldsymbol{x}|\boldsymbol{\theta}_1,\boldsymbol{\theta}_0)\right]$$

= (in large sample limit)

LLR obtained

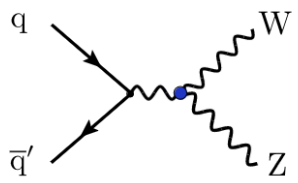
$$\hat{q}(\mathcal{D}(N_{\mathrm{obs}},\boldsymbol{x})|\boldsymbol{\theta}_1,\boldsymbol{\theta}_0) = -2\left[\mathcal{L}(\sigma(\boldsymbol{\theta}_1)-\sigma(\boldsymbol{\theta}_0)) - \sum_{i=1}^{N_{\mathrm{obs}}} \ln \hat{R}(\boldsymbol{x}|\boldsymbol{\theta}_1,\boldsymbol{\theta}_0)\right]$$

Simulation strategy for SMEFT

Possible to encode SMEFT prediction in event weights $\omega = \frac{|\mathcal{M}_{SMEFT}(c = c_1)|^2}{|\mathcal{M}_{SMEFT}(c = c_0)|^2}$

$$\omega(\boldsymbol{z}|\boldsymbol{c}) \sim \frac{d\sigma}{dz} = \mathrm{Const}(\boldsymbol{z}) + \mathrm{LIN}_a(\boldsymbol{z})c_a + \frac{1}{2}\mathrm{QUAD}_{ab}(\boldsymbol{z})c_a c_b$$

Store N(n) weights per event → obtain EFT prediction for any coefficient value



→ EFT operator changes helicity configuration

SM: +-/-+ EFT: ++/--

Helicity-aware reweighting $\omega = \frac{|\mathcal{M}(z_p, h|c_1)|^2}{|\mathcal{M}(z_p, h|c_0)|^2}$

Needs careful choice of reference point

Helicity-ignorant reweighting $\omega = \frac{\sum_h |\mathcal{M}(z_p, h|c_1)|^2}{\sum_h |\mathcal{M}(z_p, h|c_0)|^2}$

Robust option

SC & others (2024)