Combination of Higgs Differential Production Cross Sections and Effective Field Theory Interpretation

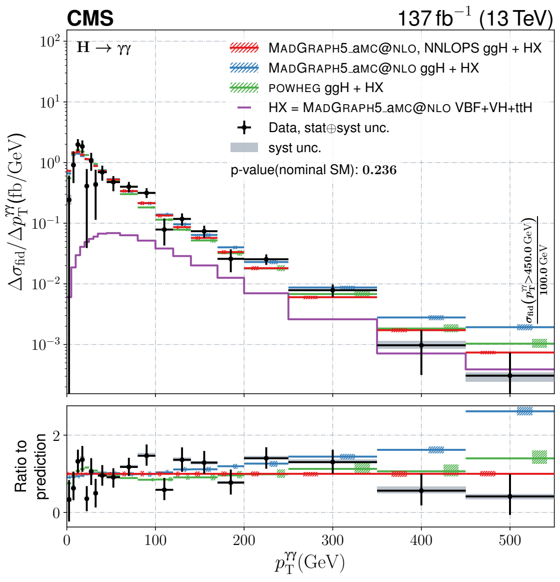
Massimiliano Galli

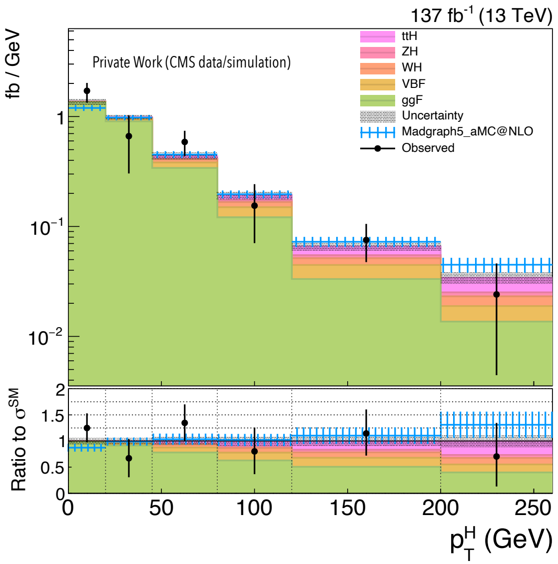
26th January 2023

CMS
ETH zürich

Introduction

- Since the discovery of the **Higgs boson** (2012), extensive effort has been put in measuring its **properties** and looking for possible **deviations** from the Standard Model predictions
- One of the main studies: **production cross sections**
- Interesting because they allow to:
  - test the SM
  - study **couplings** of the Higgs boson to other SM particles
- The **CMS experiment** has been measuring (and interpreting) cross sections in **many decay** channels since Run1, each time reducing the amount of uncertainty
- My job:
  - **combine** Run2 differential cross section measurements performed in multiple decay channels
  - provide Beyond Standard Model **interpretation**



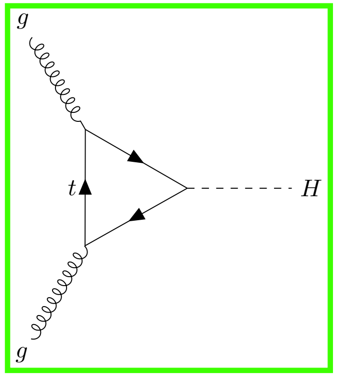


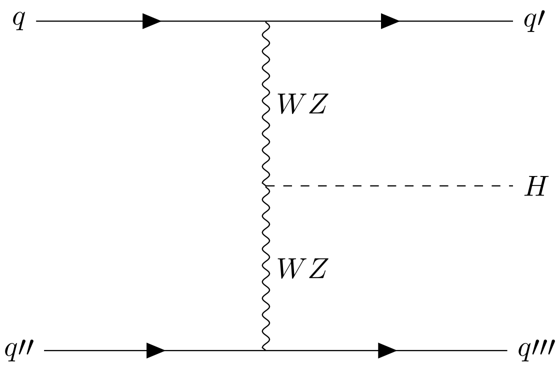
Differential Cross Sections

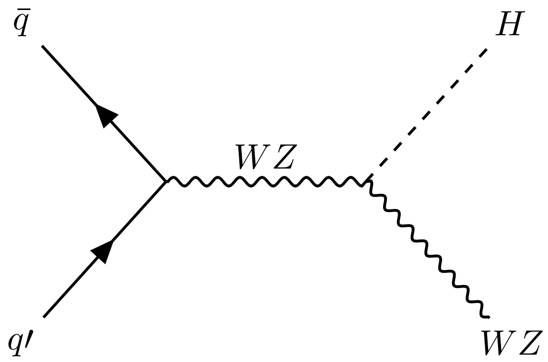
Measuring differential cross sections essentially means estimating the number of **signal events** in a **fiducial phase** space in **bins** of Higgs related kinematic observables

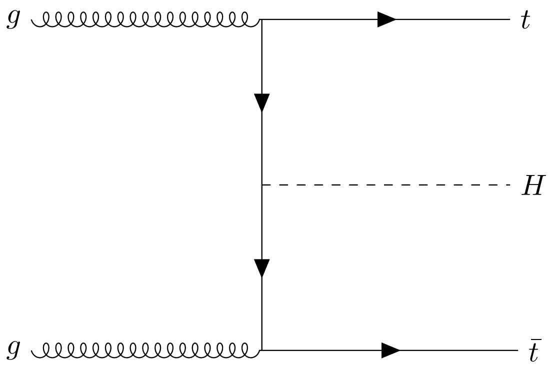
$$\sigma = \frac{N_S}{\epsilon A L}$$

- Compared to other ways of measuring cross sections, differentials have the advantage of being a **model independent** method (no a priori assumption when making the bins, contrary to e.g. STXS)
- On the other hand, this method makes the measurement only sensitive to **gluon fusion** production

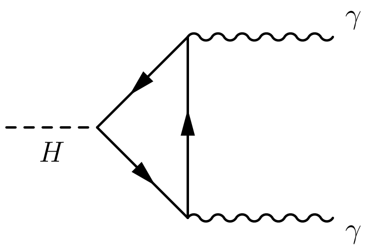




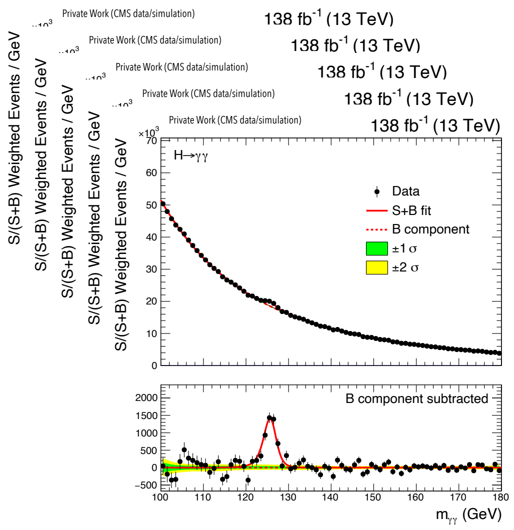




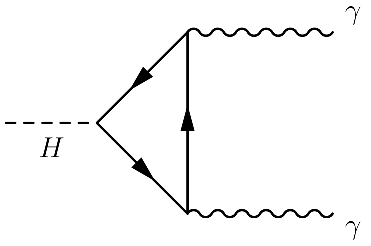
Measuring Differential Cross Sections



Measuring Differential Cross Sections



Data distributions of a certain variable (in this case invariant mass of the photon pair) + signal and background models in different bins and categories



Fit of signal strength: $\mu = \frac{\sigma}{\sigma_{SM}}$

$$\mathscr{L}(\overrightarrow{\mu} \,|\, \vec{\theta})$$

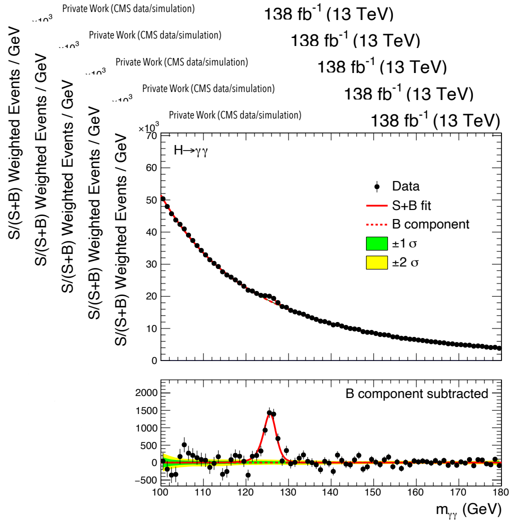
$$\overrightarrow{\mu} = (\mu_{0-5}, \mu_{5-10}, \ldots)$$

For one observable (e.g. p_T^H)

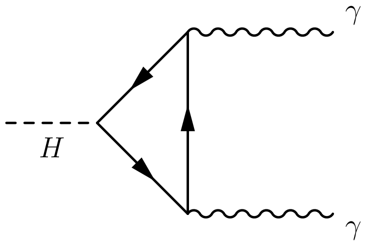
$\mathscr{L}$ is the **likelihood function,** which contains information about categorized data + signal and background probability distributions + nuisance parameters $\vec{\theta}$

CMS ETH zürich

Measuring Differential Cross Sections



Data distributions of a certain variable (in this case invariant mass of the photon pair) + signal and background models in different bins and categories



Fit of signal strength: $\mu = \frac{\sigma}{\sigma_{SM}}$

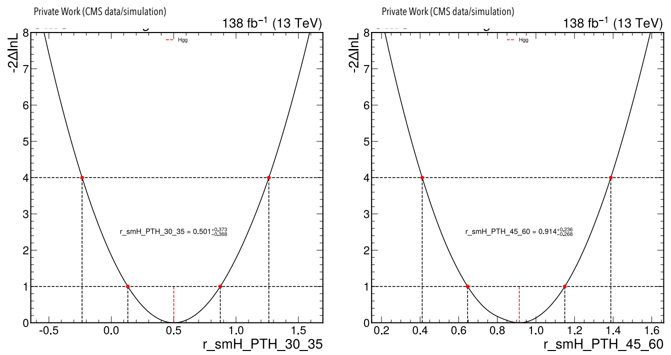
$$\mathscr{L}(\vec{\mu}\,|\,\vec{\theta})$$

$$\vec{\mu} = (\mu_{0-5}, \mu_{5-10}, \ldots)$$

For one observable (e.g. p_T^H)

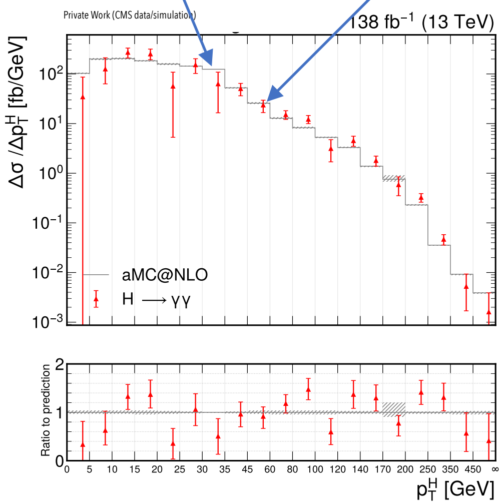
$\mathscr{L}$ is the **likelihood function,** which contains information about categorized data + signal and background probability distributions + nuisance parameters $\vec{\theta}$

Find values of $\vec{\mu}$ that maximize the likelihood (**minimize negative log likelihood**):



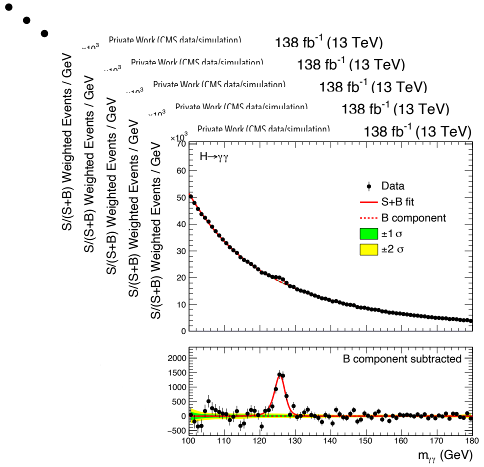
Best value for μ_{30-35}

Best value for μ_{45-60}

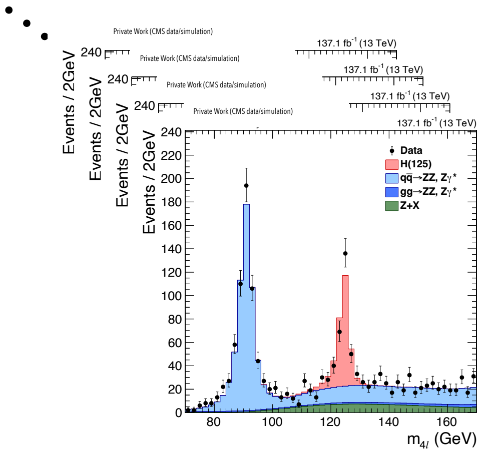


Build the observable spectrum: multiply SM prediction in each bin by the measured value of μ

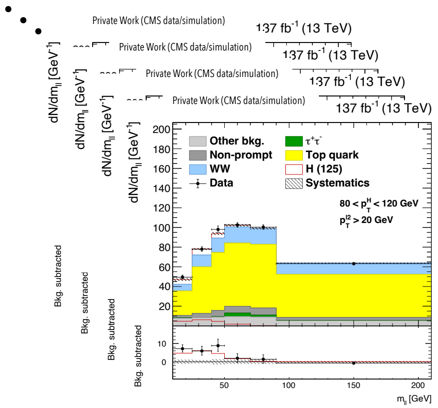
Combination



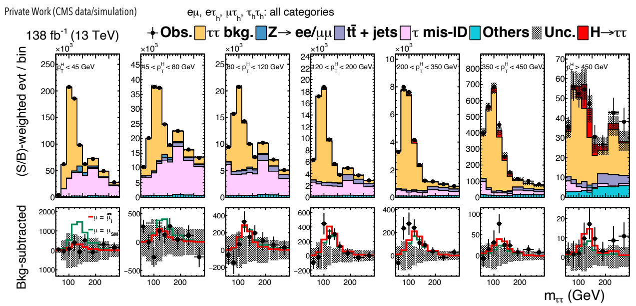
$H \to \gamma\gamma$



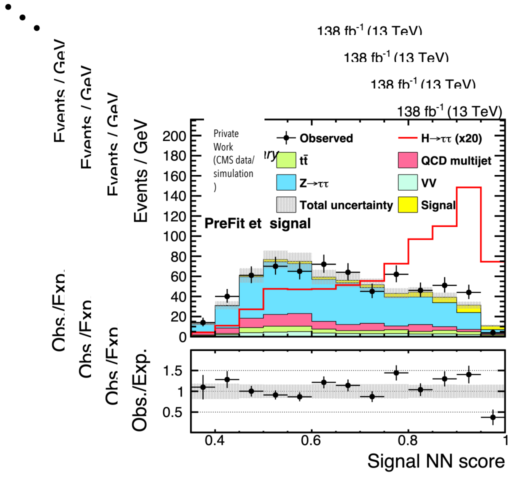
$H \to ZZ \to 4\ell$



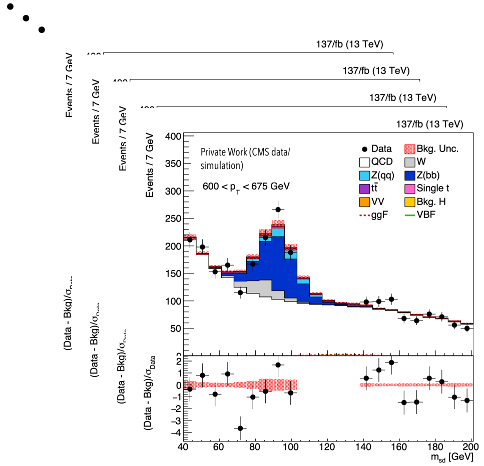
$H \to WW$



$H \to \tau\tau$

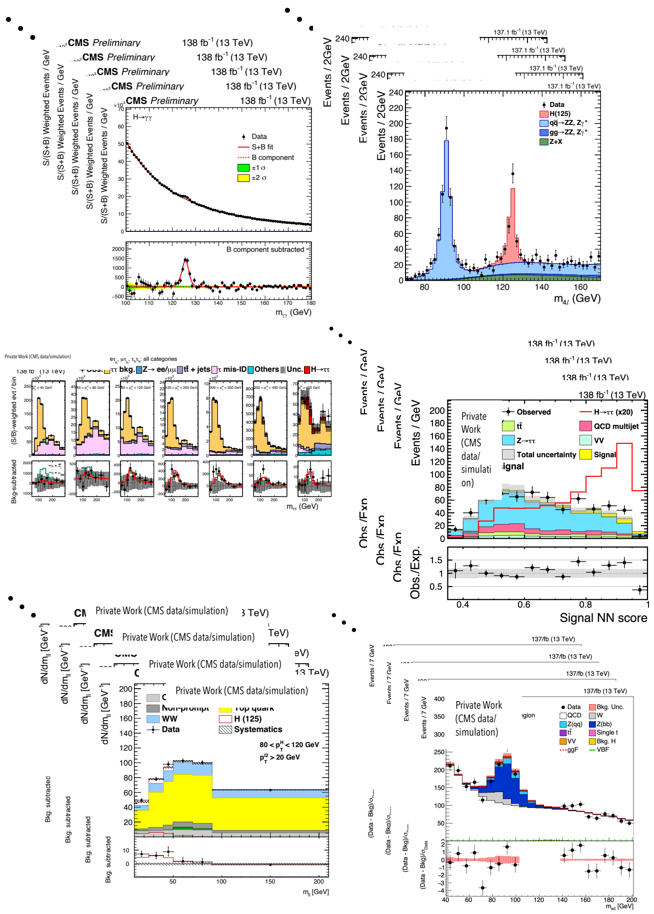


$H \to \tau\tau(\text{boosted})$



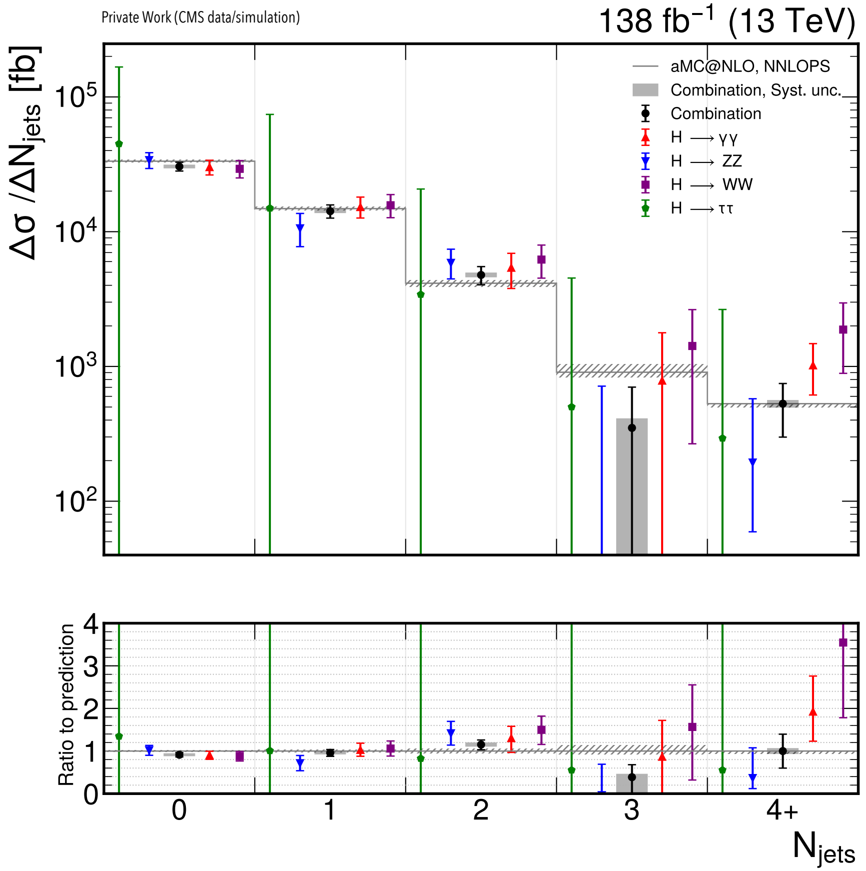
$H \to b\bar{b}$

Combination



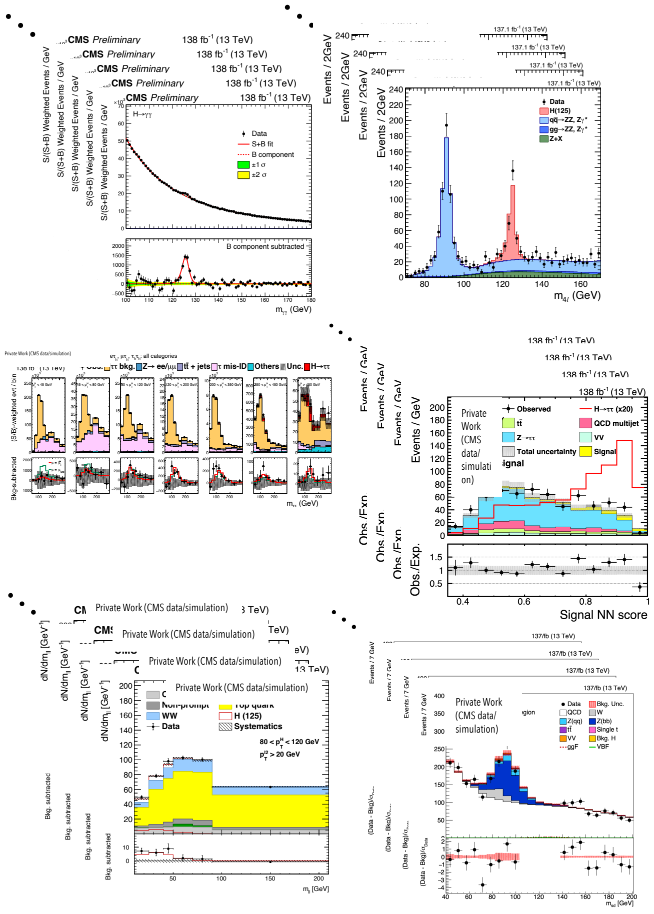
Multiply likelihoods for different decay channels (in one observable)

$$\mathscr{L}(\vec{\mu}\,|\,\vec{\theta}) = \prod_{i=1}^{n_{dc}} \mathscr{L}_i(\vec{\mu}\,|\,\vec{\theta})$$



Combination **reduces the uncertainty** on the measurements

Combination

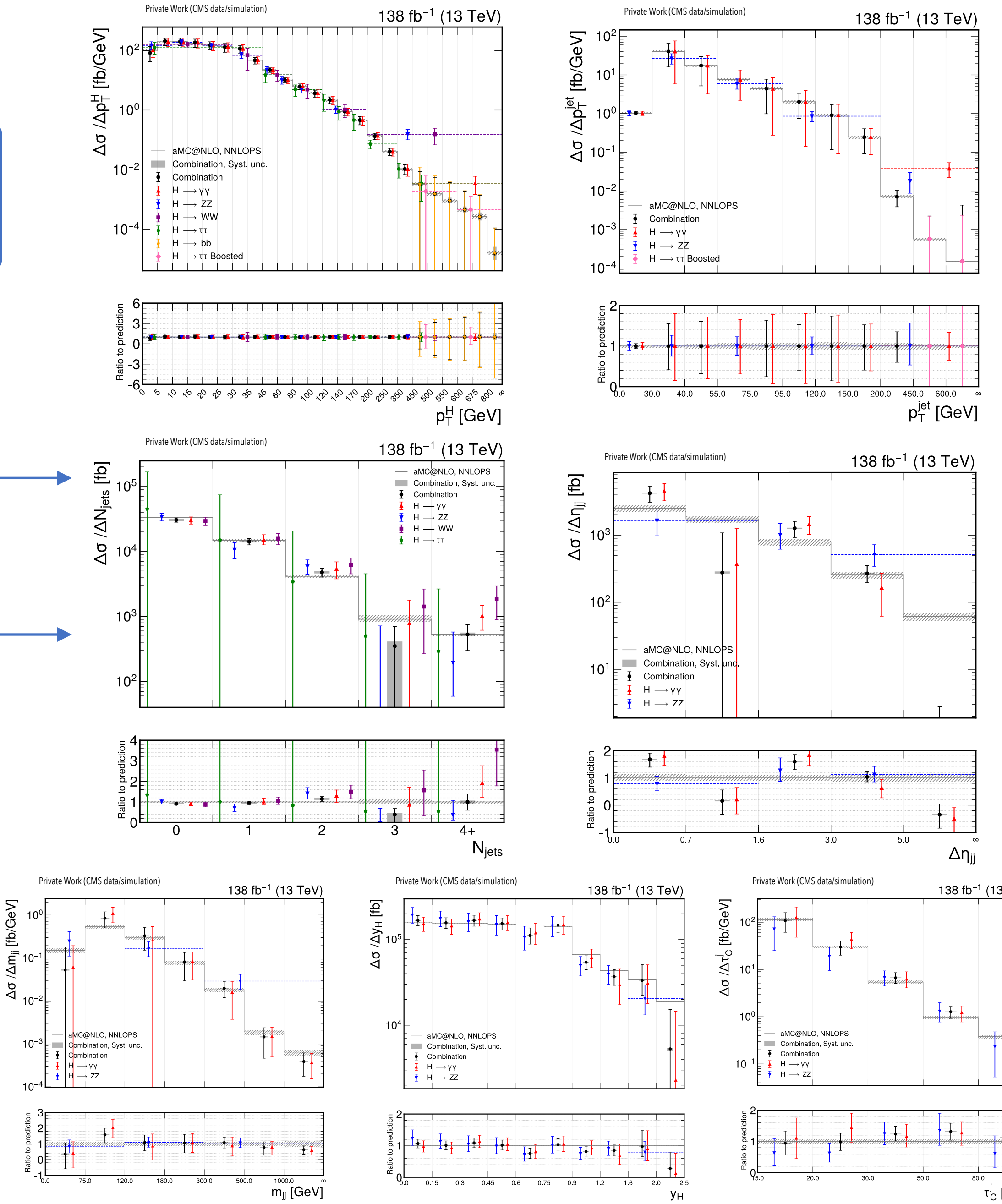


Same for many observables

$$\mathscr{L}(\vec{\mu}_{p_T^H}|\vec{\theta}) = \prod_{i=1}^{n_{dc}} \mathscr{L}_i(\vec{\mu}_{p_T^H}|\vec{\theta})$$

$$\mathscr{L}(\vec{\mu}_{Njets}|\vec{\theta}) = \prod_{i=1}^{n_{dc}} \mathscr{L}_i(\vec{\mu}_{Njets}|\vec{\theta})$$

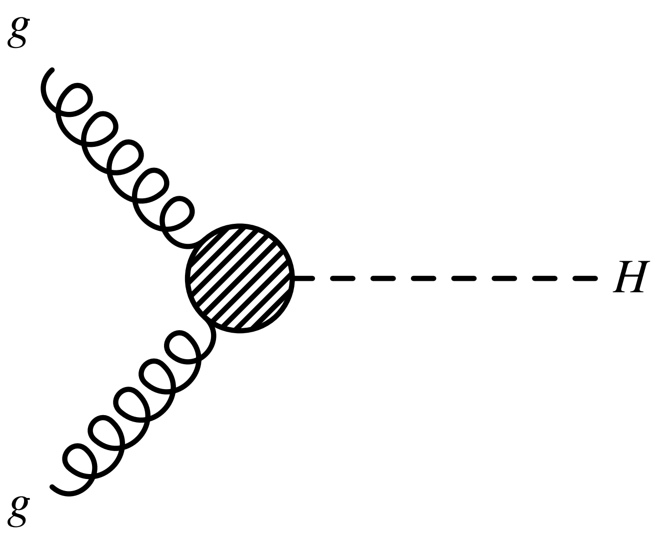
...

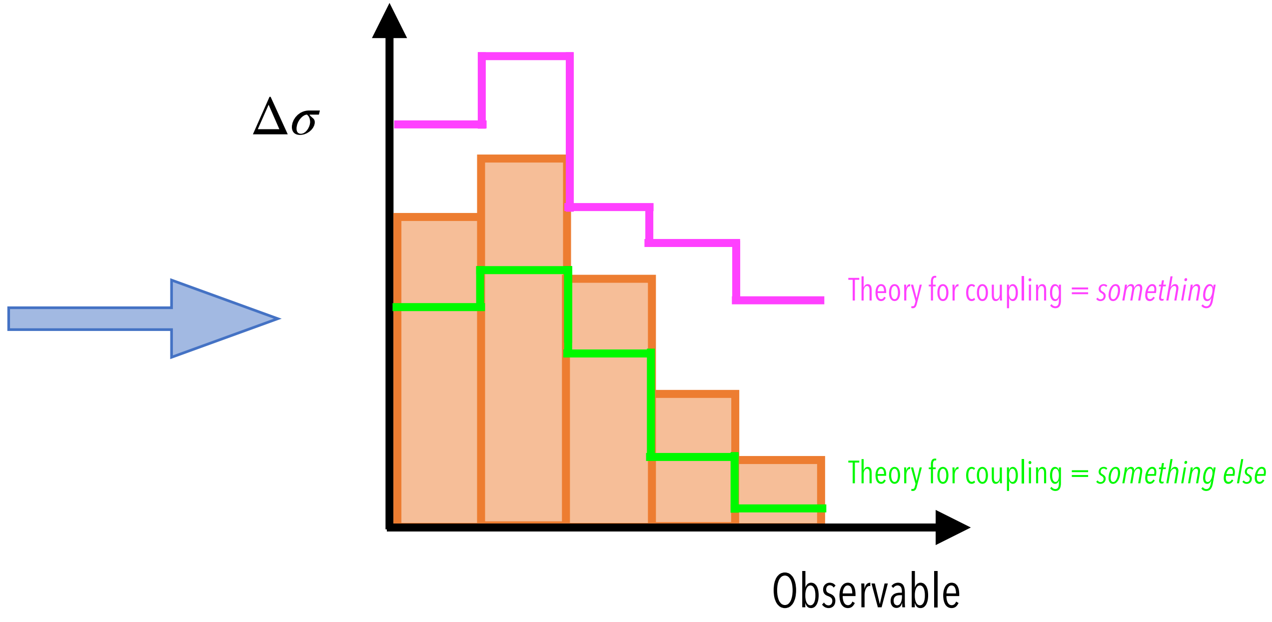


Expected results using Asimov datasets

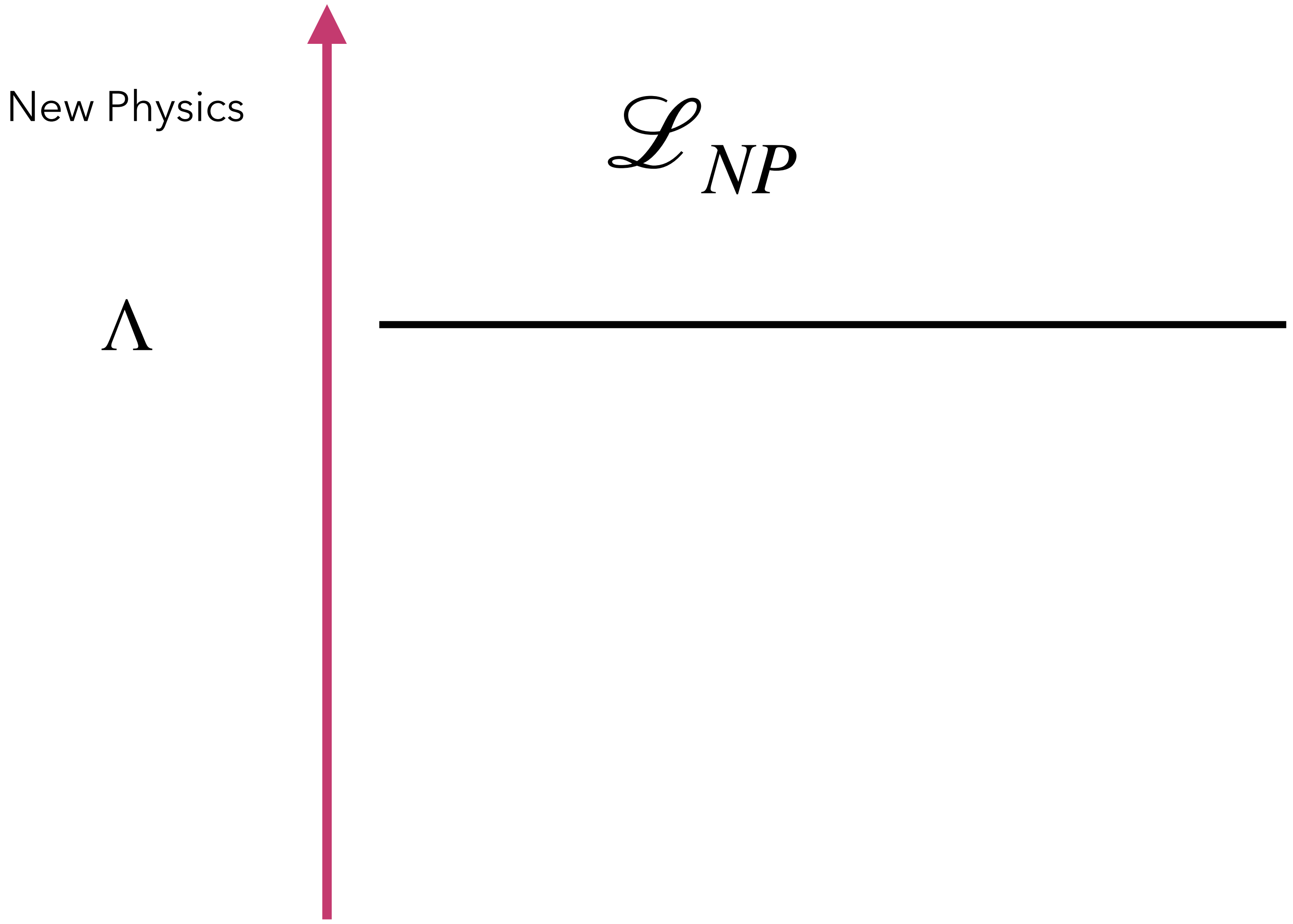
What else?

- Providing differential cross sections measurements is interesting, but… **that's not all**!
- The shape of the distribution that we obtain can be tested against its Standard Model expectation
- Relatively small coupling variations can lead to **significant shape distortions**
- **Higgs transverse momentum** is one of the most affected





SMEFT: Standard Model Effective Field Theories



- Assume existence of **unknown theory** beyond energy scale Λ

SMEFT: Standard Model Effective Field Theories

New Physics

$$\mathscr{L}_{NP}$$

Λ

Standard Model

$$\mathscr{L}_{\mathrm{SMEFT}} = \mathscr{L}_{\mathrm{SM}} + \sum_i \frac{c_j^{(6)}}{\Lambda^2} O_j^{(6)} + \mathcal{O}(\Lambda^{-4})$$

- Assume existence of **unknown theory** beyond energy scale Λ
- Effects manifest at lower energies through **new effective interactions between SM fields**
- **Wilson coefficients** c associated to operators in the Lagrangian
- Stop at dimension 6
- 59 independent coefficients
- Values of WC different from 0 = **hint of discrepancy from SM**

SMEFT: Standard Model Effective Field Theories

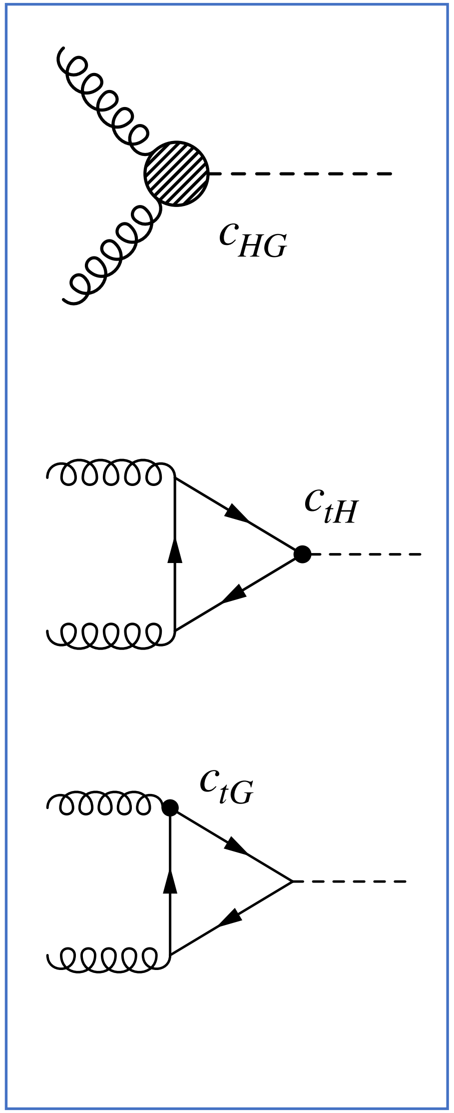
New Physics

$$\mathscr{L}_{NP}$$

Λ

Standard Model

$$\mathscr{L}_{\text{SMEFT}} = \mathscr{L}_{\text{SM}} + \sum_i \frac{c_j^{(6)}}{\Lambda^2} O_j^{(6)} + \mathcal{O}(\Lambda^{-4})$$



Effects on Cross Section

Wilson coefficients affect the number of signal events that we detect

→

Parametrize signal strength as $\mu = \mu(\vec{c})$

→

Rewrite the likelihood function as $\mathscr{L}(\overrightarrow{\mu}(\vec{c})\,|\,\vec{\theta})$ and find the values of $\vec{c}$ that maximize it

Effects on Cross Section

Wilson coefficients affect the number of signal events that we detect → Parametrize signal strength as $\mu = \mu(\vec{c})$ → Rewrite the likelihood function as $\mathscr{L}(\vec{\mu}(\vec{c})\,|\,\vec{\theta})$ and find the values of $\vec{c}$ that maximize it

In narrow width approximation:

$\sigma(gg \to H \to X) = \sigma(gg \to H) \cdot \mathrm{BR}(H \to X) = \sigma(gg \to H) \cdot \frac{\Gamma^{H\to X}}{\Gamma^{\mathrm{tot}}}$

Production

Decay width channel X

$$\mu_i^X(c_j) = \frac{\sigma_{SMEFT}(gg \to H)}{\sigma_{SM}(gg \to H)} \cdot \frac{\Gamma^{H\to X}_{SMEFT}/\Gamma^{H\to X}_{SM}}{\Gamma^{H}_{SMEFT}/\Gamma^{H}_{SM}}$$

Total decay width

$$\frac{\sigma_{SMEFT}(gg \to H)}{\sigma_{SM}(gg \to H)} = 1 + \sum_j A_j^{gg\to H} c_j + \sum_{jk} B_{jk}^{gg\to H} c_j c_k$$

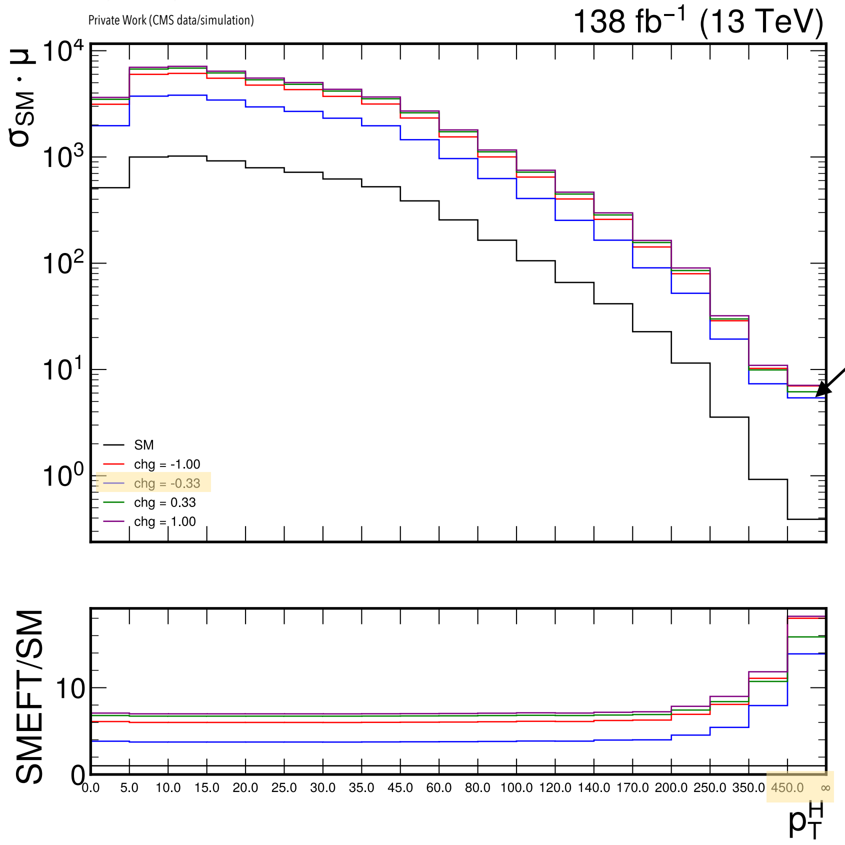
$$\frac{\Gamma^{H\to X}_{SMEFT}}{\Gamma^{H\to X}_{SM}} = 1 + \sum_j A_j^{H\to X} c_j + \sum_{jk} B_{jk}^{H\to X} c_j c_k$$

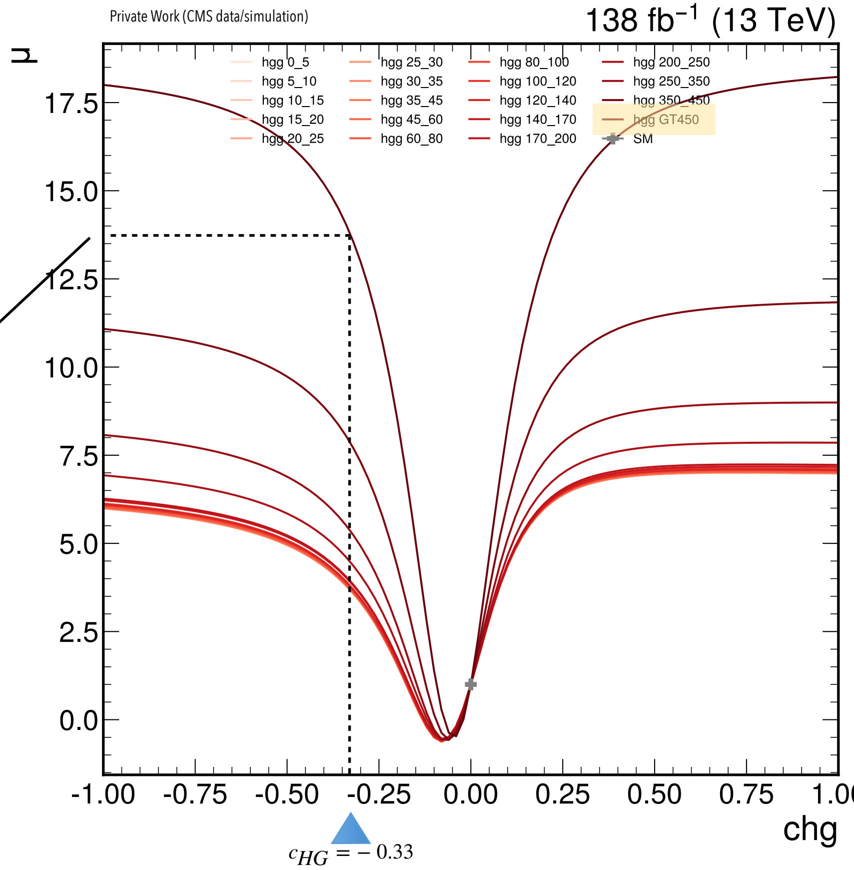
$$\frac{\Gamma^{H}_{SMEFT}}{\Gamma^{H}_{SM}} = 1 + \sum_j A_j^{H} c_j + \sum_{jk} B_{jk}^{H} c_j c_k$$

Equations can be obtained with simulation

Effects on Cross Section

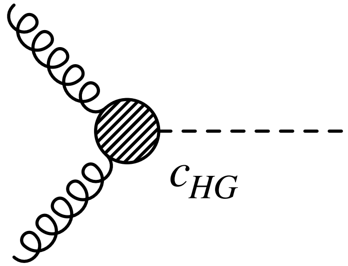
$$\mu_i^X(c_j) = (1 + \sum_j A_j^{gg\to H} c_j + \sum_{jk} B_{jk}^{gg\to H} c_j c_k) \cdot \frac{(1 + \sum_j A_j^{H\to X} c_j + \sum_{jk} B_{jk}^{H\to X} c_j c_k)}{(1 + \sum_j A_j^{tot} c_j + \sum_{jk} B_{jk}^{tot} c_j c_k)}$$





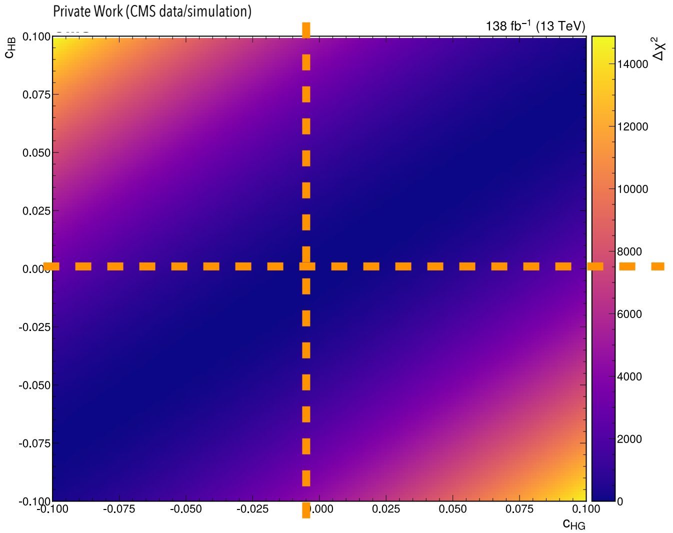
To get the shape for some value (e.g. -0.33) of a coefficient (e.g. chg):

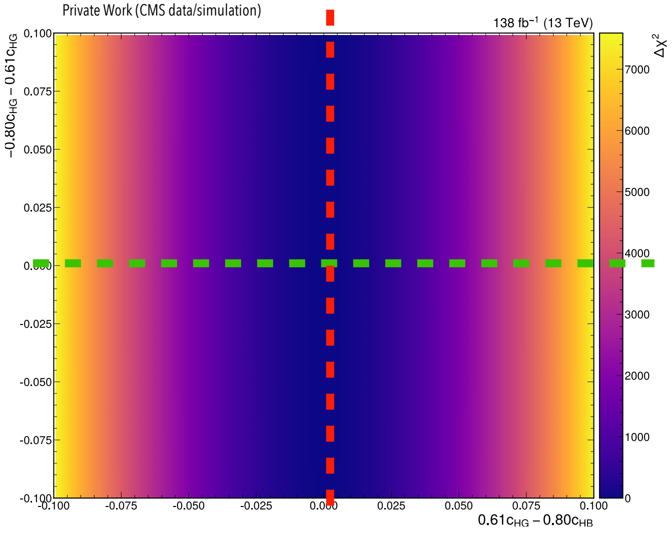
- get μ for each bin by interpolating the functions on the right
- multiply the SM cross section in each bin by the corresponding value of μ



Basis Rotation

- Unfortunately fitting together many Wilson coefficients is not straightforward, so a Principal Component Analysis (**PCA**) procedure was developed to **find constrained directions** in the parameter space
- Practically speaking we **rotate the basis** and fit linear combinations of the original WCs





$$c_0, \ldots, c_n$$

$$EV_0 = f_0(c_0, \ldots, c_n)$$
$$\ldots$$
$$EV_n = f_n(c_0, \ldots, c_n)$$

Results

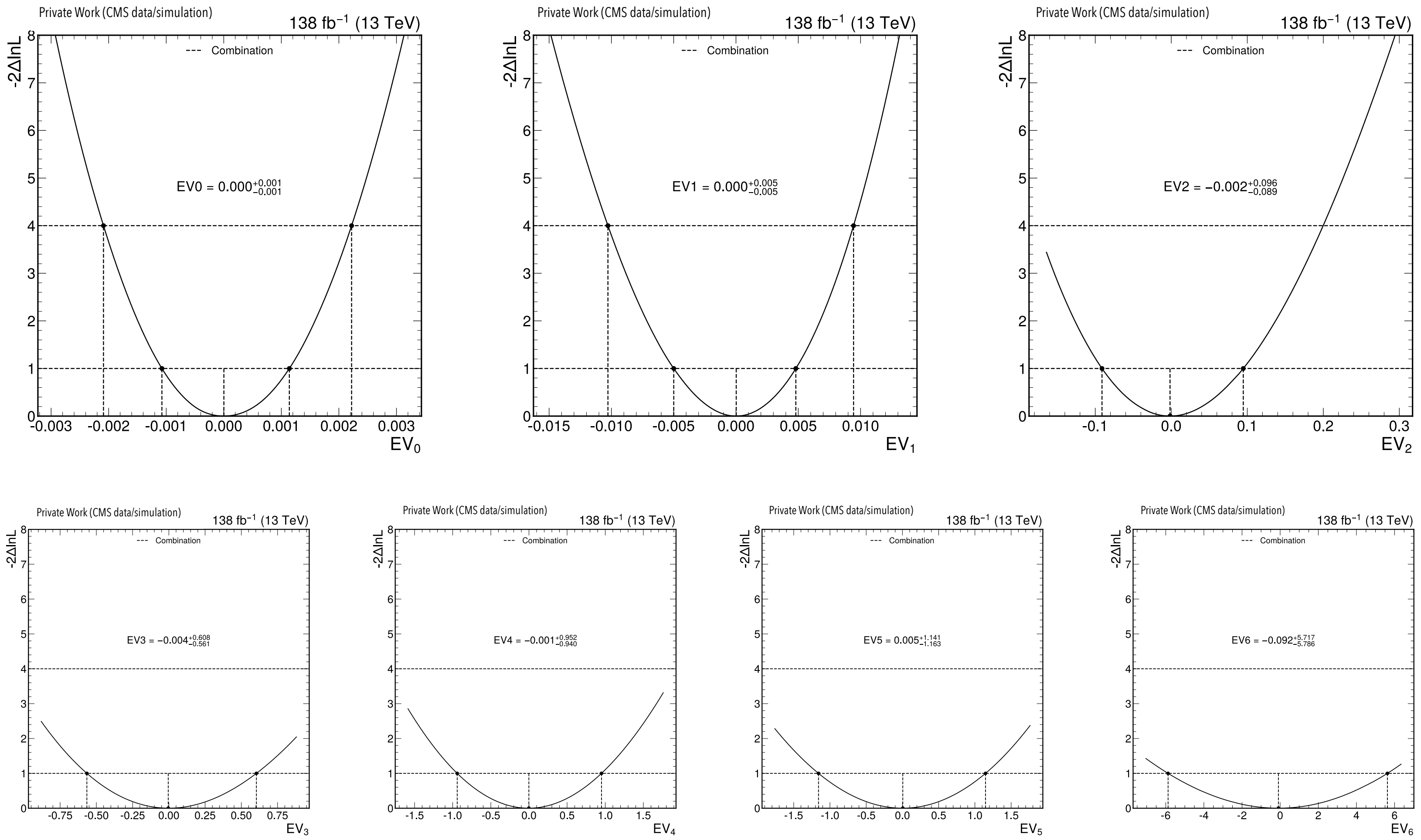
Write

$$\mathscr{L}(\overrightarrow{\mu}_{p_T^H}(\overrightarrow{EV})\,|\,\vec{\theta})$$

with

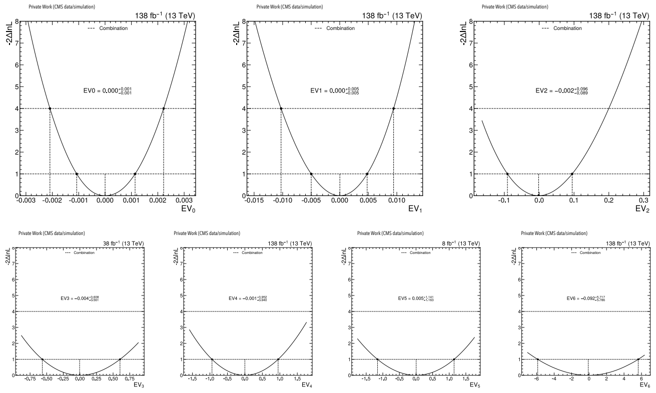
$$\overrightarrow{\mu}_{p_T^H} = (\mu^{\gamma\gamma}_{0-5}, \ldots, \mu^{ZZ}_{0-10}, \ldots, \mu^{WW}_{0-30}, \ldots)$$

and find values of $\overrightarrow{EV}$ that maximize $\mathscr{L}$



Expected results using Asimov datasets

How is it Useful?

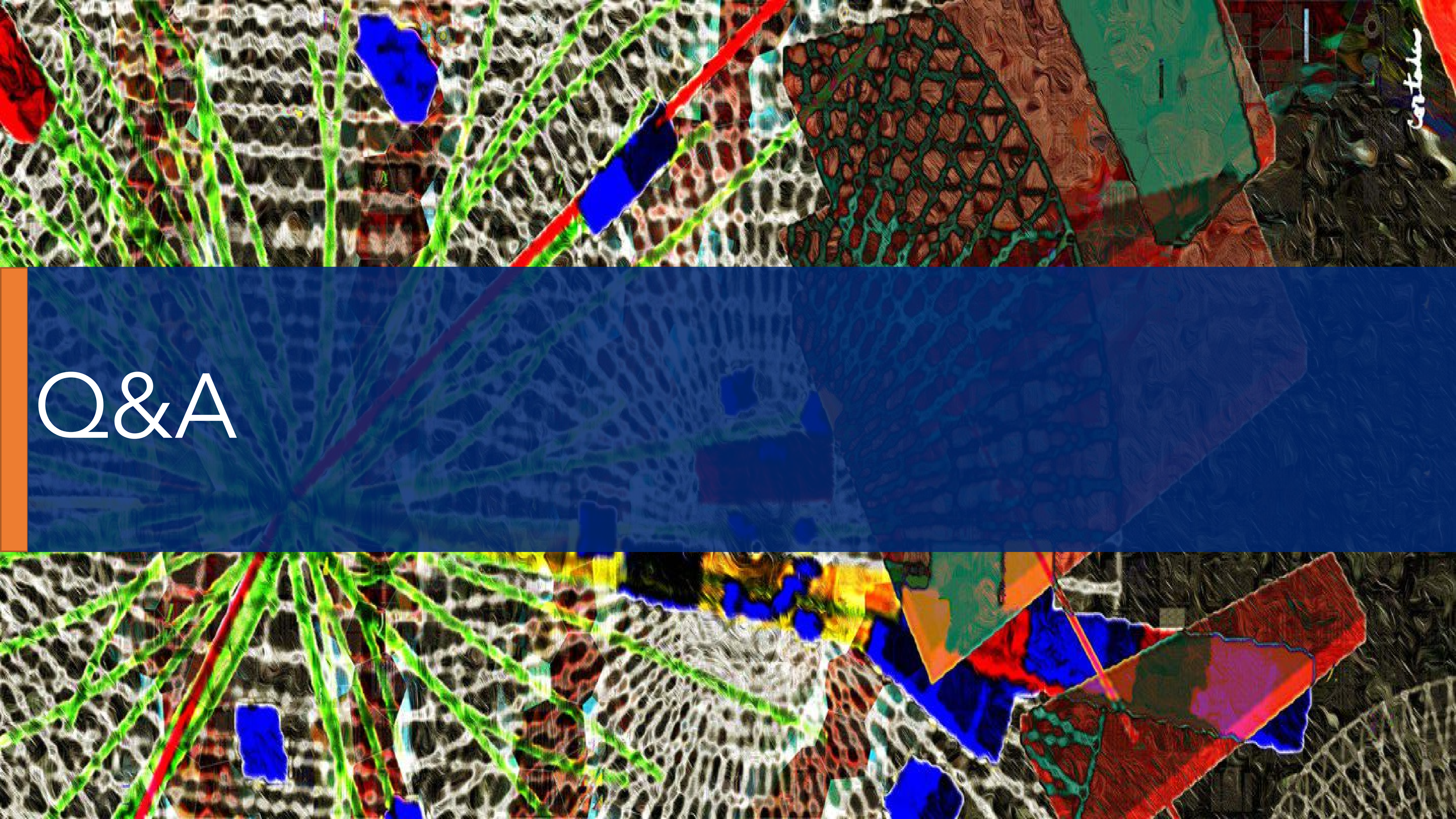


Confidence levels on linear combinations of coefficients

$$EV_0 = f_0(c_0, \ldots, c_n)$$
$$\ldots$$
$$EV_n = f_n(c_0, \ldots, c_n)$$

Definitions of fitted linear combinations

Theorists can use these information in **new and improved studies** on BSM models



Q&A

Backup

EFT2Obs

1

Event generation with MadGraph (SMEFT@NLO for gluon fusion predictions)

↓

Shower with Pythia8

↓

Rivet routine to get histograms for the observables we are interested in

Repeat $2N + (N^2 - N)/2$ weights + SM with N number of Wilson coefficients

2 Transform following:

Label	c_1	c_2	Weight	Transformed weight
W_1	0	0	1	1
W_2	$0.5D_1$	0	$1 + 0.5D_1A_1 + 0.25{D_1}^2B_{11}$	D_1A_1
W_3	D_1	0	$1 + D_1A_1 + {D_1}^2B_{11}$	${D_1}^2B_{11}$
W_4	0	$0.5D_2$	$1 + 0.5D_2A_2 + 0.25{D_2}^2B_{22}$	D_2A_2
W_5	0	D_2	$1 + D_2A_2 + {D_2}^2B_{22}$	${D_2}^2B_{22}$
W_6	D_1	D_2	$1 + D_1A_1 + D_2A_2 + D_1^2B_{11} + {D_2}^2B_{22} + D_1D_2B_{12}$	$D_1D_2B_{12}$

3 Get coefficients of the equations:

$$A_j = \frac{\overline{A_j}}{\bar{S}D_j}$$

$$B_{jj} = \frac{\overline{B_{jj}}}{\bar{S}D_j^2}$$

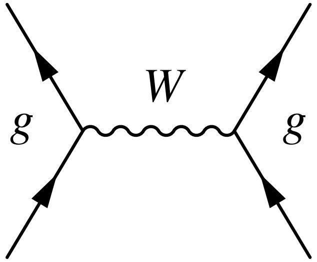
$$B_{jk} = \frac{\overline{B_{jk}}}{\bar{S}D_jD_k}$$

With:

- $\bar{S}$ ratio between sum of weights and number of events in bin i of SM histogram
- $\overline{A_j}$ same for bin i of histogram with transformed weight D_jA_j
- $\overline{B_{jj}}$ same for bin i of histogram with transformed weight $D_j^2B_{jj}$
- $\overline{B_{jk}}$ same for bin i of histogram with transformed weight $D_jD_kB_{jk}$

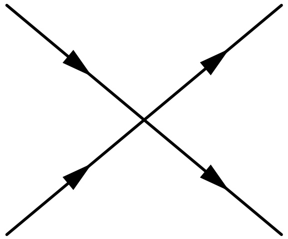
Weak Interactions as EFTs

Standard model



$$\frac{g^2}{p^2 - m_W^2}$$

Λ



$$G_F \propto -\frac{g^2}{m_W^2}$$

$E_{\beta-\mathrm{decay}}$

Effects on Cross Section

$$\mu_i^X(c_j) = 1 + \sum_j (A_{ij}^{gg\to H} + A_j^{H\to X} - A_j^{\text{tot}})c_j$$