



UNIVERSITÀ  
DEGLI STUDI  
FIRENZE



IFAE 2025

Incontri di Fisica delle Alte Energie

Cagliari, 9-11 Aprile 2025

# Misure differenziali di $H \rightarrow WW$ indipendenti dal modello

[CMS-PAS-HIG-24-004](#)

Benedetta Camaiani

Università degli Studi di Firenze e INFN Firenze

A nome della collaborazione CMS

# Struttura del talk



## 1. Misure indipendenti da modelli teorici

- a. Produzione del bosone di Higgs con 2 jet
- b. Dipendenza da modelli teorici nella strategia di analisi
- c. Rete neurale avversariale

## 2. Misura della sezione d'urto differenziale $H+2j$

- a. Osservabile di interesse
- b. Strategia di analisi
- c. Risultati

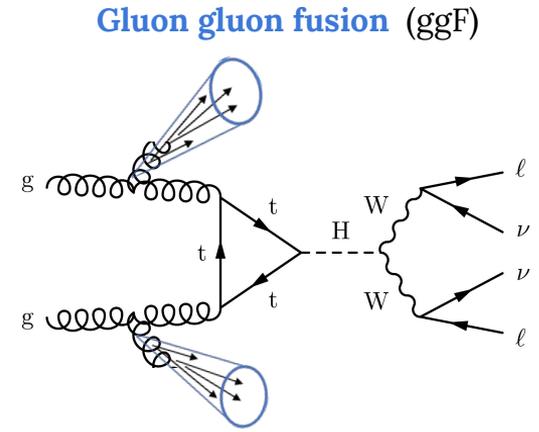
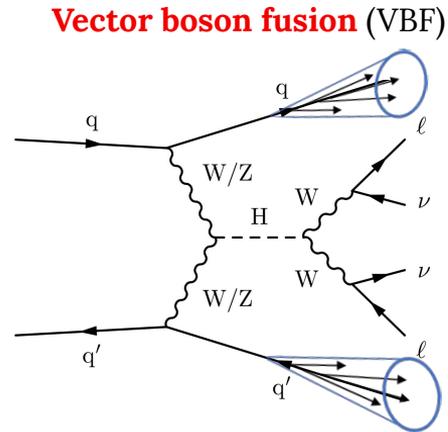
Parte 1

# Misure indipendenti da modelli teorici

# Introduzione



- **Misura differenziale** della produzione del bosone di **Higgs in associazione con 2 jet** nel canale  $H \rightarrow W^+W^- \rightarrow 2\ell 2\nu$  usando i dati raccolti dall'esperimento CMS durante il Run2 di LHC ( $L=138\text{fb}^{-1}$ ,  $\sqrt{s}=13\text{ TeV}$ )
- Osservabile di interesse legato alla separazione angolare tra i due jet



# Introduzione



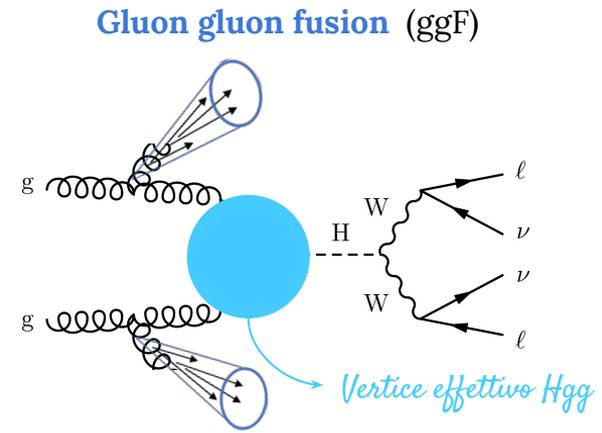
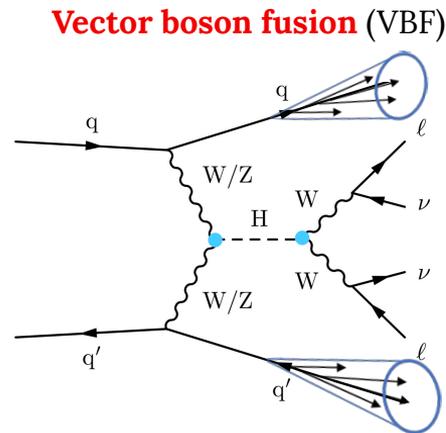
- **Misura differenziale** della produzione del bosone di **Higgs in associazione con 2 jet** nel canale  $H \rightarrow W^+W^- \rightarrow 2\ell 2\nu$  usando i dati raccolti dall'esperimento CMS durante il Run2 di LHC ( $L=138\text{fb}^{-1}$ ,  $\sqrt{s}=13\text{ TeV}$ )

- Osservabile di interesse legato alla separazione angolare tra i due jet

- Teorie BSM prevedono accoppiamenti anomali (**AC**) nel **vertice HVV/Hgg**

- Gli AC influenzano le distribuzioni cinematiche del sistema dei jet

→ **Misurare la sezione d'urto del segnale indipendentemente dal modello teorico che governa il vertice HVV/Hgg**

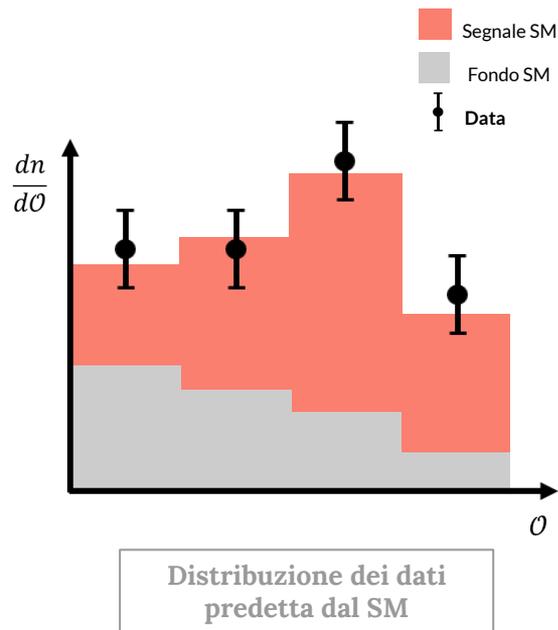


E' opportuno ridurre possibili bias che potrebbero favorire una specifica previsione teorica

# Procedura di fit



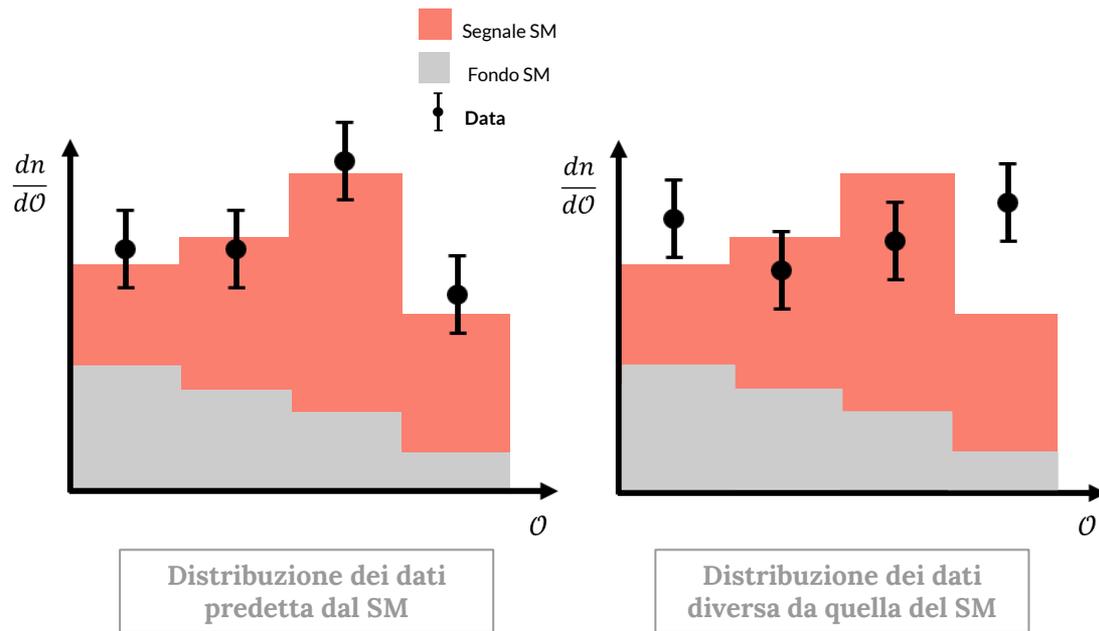
- La principale sorgente di dipendenza da modelli teorici risiede nella procedura di fit
- Si estrae la distribuzione di un osservabile da simulazioni Monte Carlo (MC)
  - Di solito **si assume l'ipotesi del SM per la simulazione**



# Procedura di fit



- La principale sorgente di dipendenza da modelli teorici risiede nella procedura di fit
- Si estrae la distribuzione di un osservabile da simulazioni Monte Carlo (MC)
  - Di solito **si assume l'ipotesi del SM per la simulazione**
- Il modello fisico che descrive i dati è sconosciuto
  - Eseguire il fit utilizzando l'ipotesi del SM può introdurre un bias nei risultati



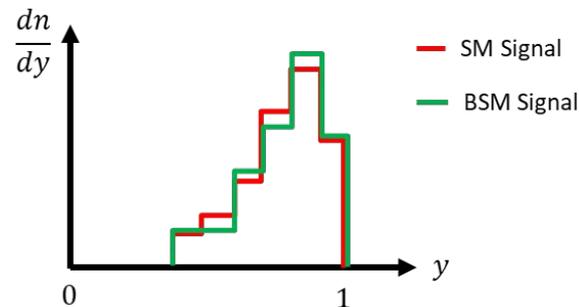
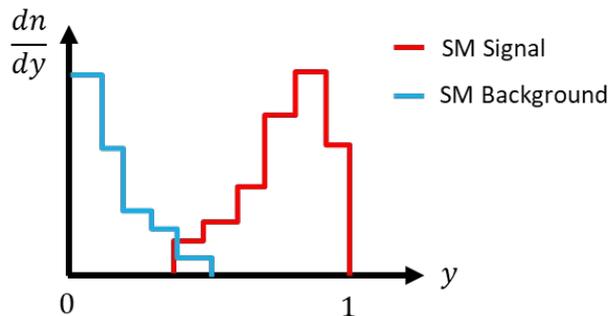
I risultati del fit hanno un bias verso il modello assunto per generare la simulazione MC

# Obiettivo



L'obiettivo principale è implementare una variabile di fit  $y$  agnostica verso le ipotesi di segnale:

- $y$  deve discriminare tra eventi di segnale e fondo
- $y$  non deve distinguere tra i modelli fisici degli eventi di segnale



**Domain Adaptation**

Rete neurale avversariale (ADNN)

# Rete neurale avversariale

Eur. Phys. J. C (2022) 82: 921



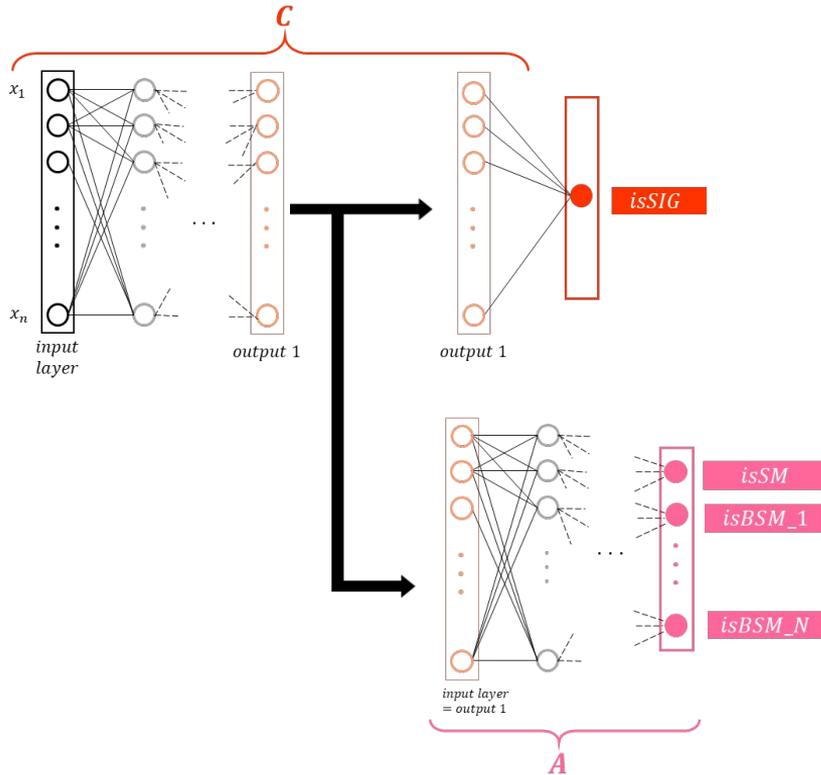
## Classificatore

- **Rete neurale binaria o multiclasse** che utilizza come input le variabili cinematiche misurabili di un evento
- **Addestrata su eventi di segnale e di fondo**
- Il campione di segnale include eventi provenienti da domini diversi, ovvero **modelli di segnale differenti**

## Avversario

- **Rete neurale multiclasse** addestrata **solo su eventi di segnale** (ipotesi SM + BSM)
- **Cerca di identificare il modello fisico degli eventi di segnale**, determinando il dominio a partire dal penultimo layer di C

$$Loss = Loss(C) - \underbrace{\alpha \cdot Loss(A)}_{\text{Termine di penalità}}$$



# Rete neurale avversariale

Eur. Phys. J. C (2022) 82: 921

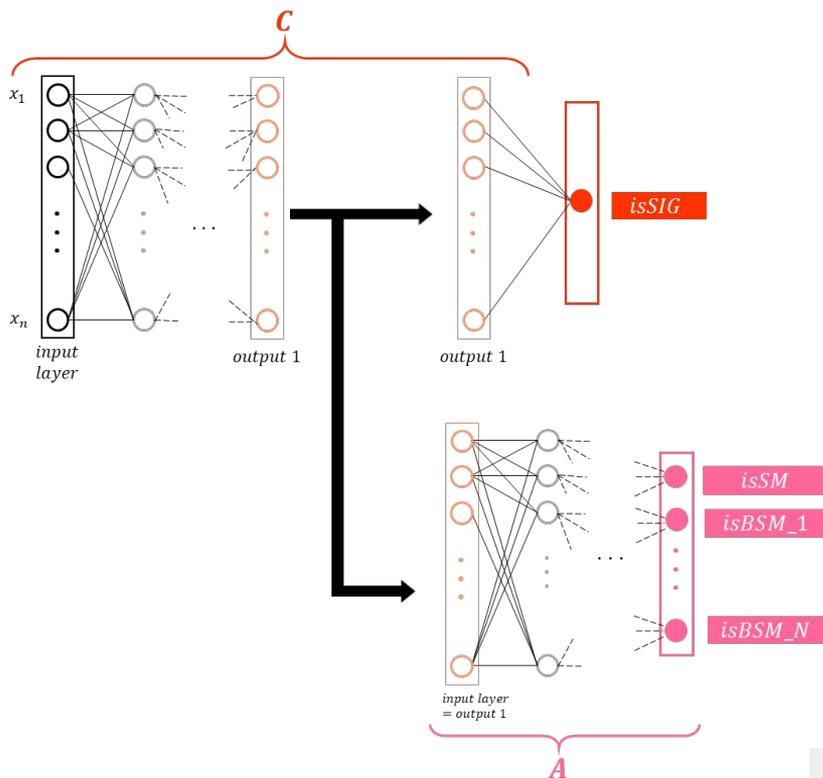


## Classificatore

- Rete neurale **binaria o multiclasse** che utilizza come input le variabili cinematiche misurabili di un evento
- Addestrata su eventi di segnale e di fondo
- Il campione di segnale include eventi provenienti da domini diversi, ovvero **modelli di segnale differenti**

## Avversario

- Rete neurale **multiclasse** addestrata **solo su eventi di segnale** (ipotesi SM + BSM)
- Cerca di **identificare il modello fisico degli eventi di segnale**, determinando il dominio a partire dal penultimo layer di C



La ADNN viene usata nella misura della sezione d'urto differenziale H+2jet per mitigare la dipendenza da modelli teorici dei risultati

Parte 2

Misura differenziale della sezione d'urto  
differenziale  $H+2j$

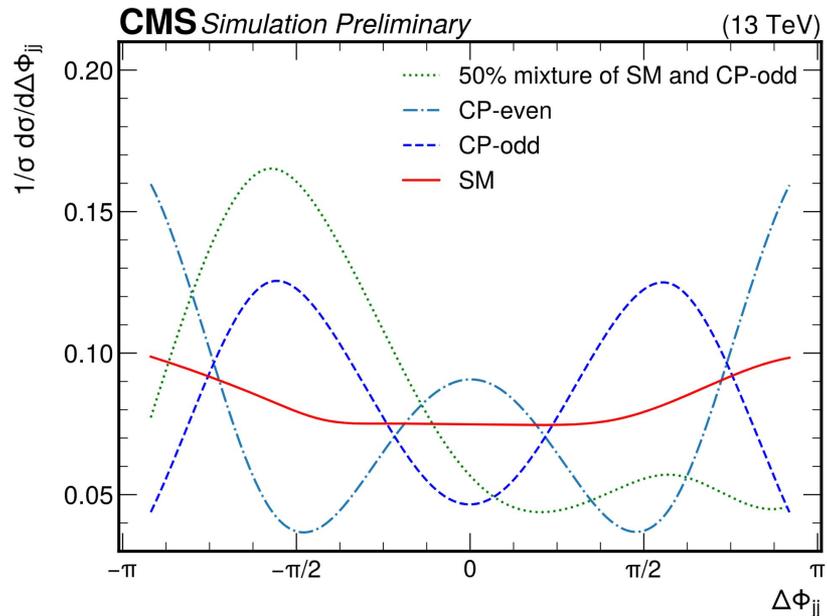
# Osservabile di interesse



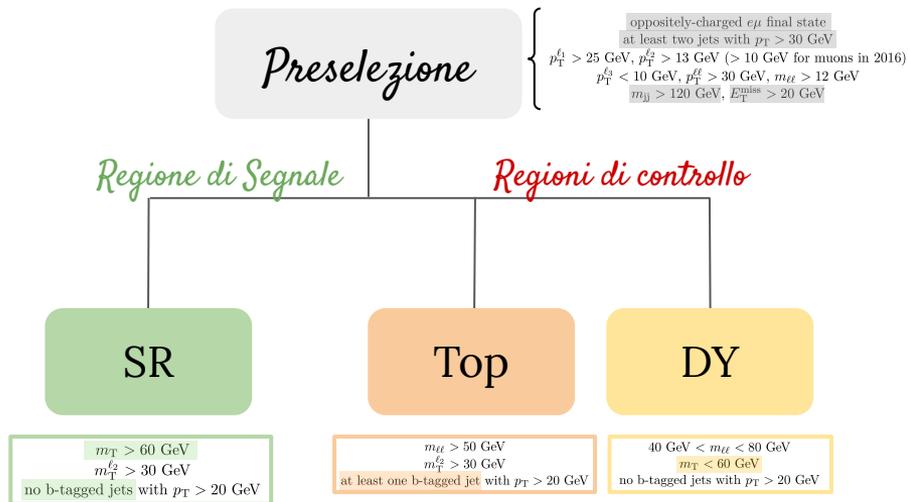
Differenza (con segno) in angolo azimutale  $\Delta\Phi_{jj}$  tra i due jet più energetici nello stato finale

$$\Delta\Phi_{jj} = \phi_{jk} - \phi_{jl} \quad \text{con} \quad \eta_{jk} > \eta_{jl}$$

- Distribuzione simmetrica secondo il Modello Standard
- Sensibile a contributi **CP-dispari e CP-pari in HVV/Hgg**
- **Asimmetria per i contributi CP-misti**



# Strategia di analisi



- **Preselezione:** 2 leptoni carichi ( $e\mu$ ),  $MET > 20$  GeV, 2 jet con  $p_T < 30$  GeV e  $m_{jj} > 120$  GeV
- **Regione di segnale (SR):** veto per jet provenienti da quark b (**b-jet veto**) e  $m_T > 60$  GeV
- **Fondi principali** a valle della preselezione: **top quark**,  $Z/\gamma^* \rightarrow \tau\tau$ ,  $W^+W^-$
- **2 regioni di controllo (CR) ortogonali alla SR** per costringere la normalizzazione dei principali fondi:
  - **Top CR per fondo da quark top** (almeno 1 b-jet)
  - **DY CR per fondo  $Z/\gamma^* \rightarrow \tau\tau$**  ( $m_T < 60$  GeV)

- Ogni regione è divisa in bin di  $\Delta\Phi_{jj}$ :

Bin 0	Bin 1	Bin 2	Bin 3
$-\pi < \Delta\Phi_{jj} \leq -\frac{\pi}{2}$	$-\frac{\pi}{2} < \Delta\Phi_{jj} \leq 0$	$0 < \Delta\Phi_{jj} \leq \frac{\pi}{2}$	$\frac{\pi}{2} < \Delta\Phi_{jj} \leq \pi$

# Risultati del training

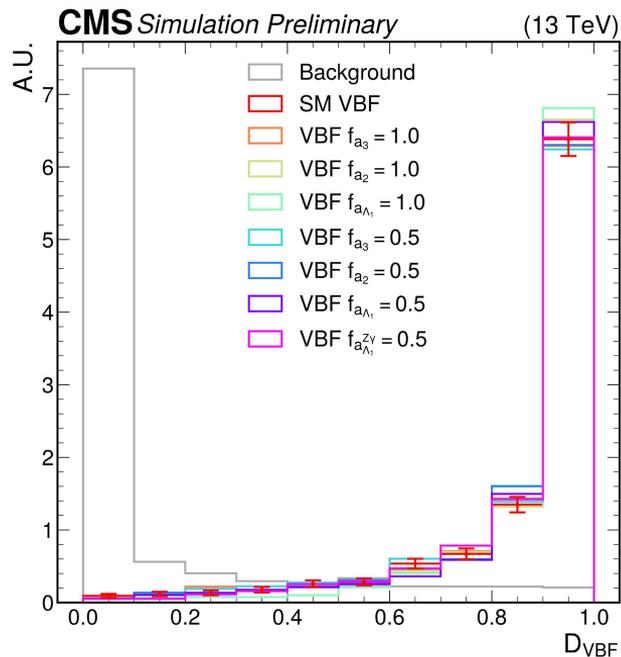


- Sono state **addestrate due ADNN nella SR inclusiva**, utilizzando eventi MC

→ Fortemente discriminanti tra segnale e fondo

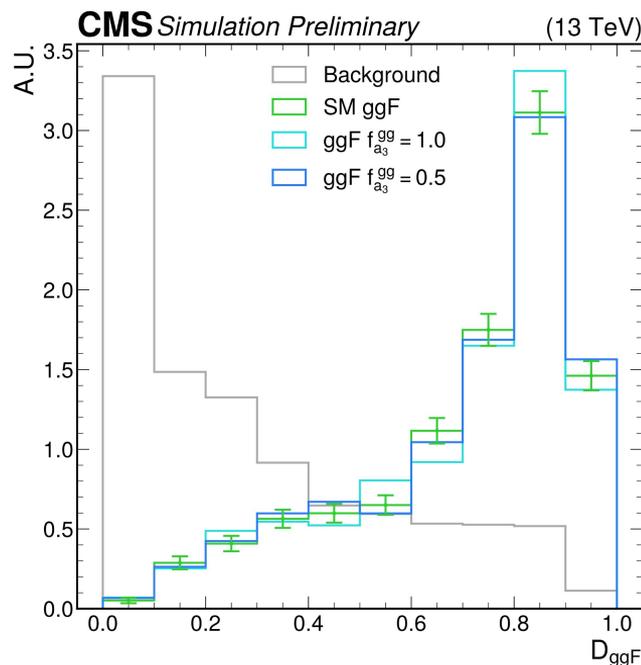
→ Le forme su eventi SM e BSM sono in buon accordo tra loro, **il classificatore non è in grado di riconoscere i modelli fisici degli eventi di segnale**

## VBF-ADNN



Accuracy 91%

## GGH-ADNN

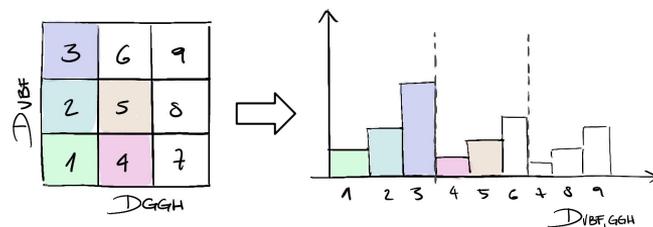


Accuracy 79%

# Estrazione del segnale



- **Maximum likelihood fit** della distribuzione di output dell'ADNN, eseguito simultaneamente in tutti i bin della SR
  - variabile 1D:  $D_{\text{VBF}}$  (output di VBF-ADNN)
  - variabile 2D:  $D_{\text{VBF,ggF}}$  composta da  $D_{\text{VBF}}$  e  $D_{\text{ggF}}$  (output di GGH-ADNN)

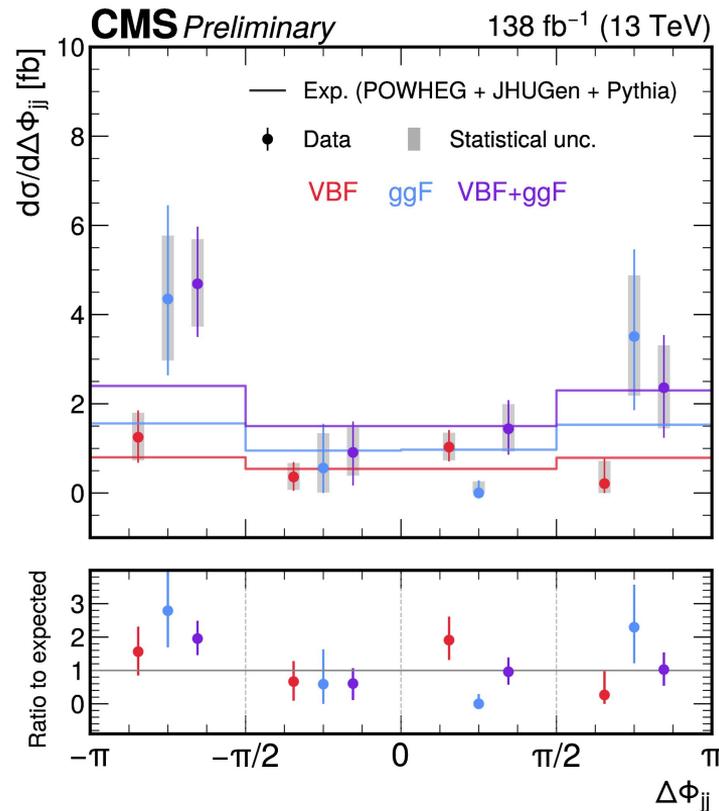


- Sezioni d'urto misurate in un **volume fiduciale**
  - Campioni di segnale suddivisi in bin di  $\Delta\Phi_{jj}$  con lo stesso schema delle categorie di analisi
  - Procedura di unfolding integrata nel fit
- Il **numero di eventi** in ciascun bin di  $\Delta\Phi_{jj}$  delle CR top e DY viene incluso nel fit per misurare la normalizzazione del fondo corrispondente

# Sezioni d'urto differenziali



- Segnale estratto in tre diverse configurazioni
  - Sezione d'urto H+2jet ( $D_{\text{VBF,ggF}}$ )
  - Misura simultanea di sezione d'urto VBF e ggF ( $D_{\text{VBF,ggF}}$ )
  - Sezione d'urto VBF ( $D_{\text{VBF}}$ )
- Risultati dominati dall'incertezza statistica
- Dipendenza dal modello ridotta di un fattore 30-70% in base al bin di  $\Delta\Phi_{jj}$  rispetto all'uso di una DNN semplice



# Interpretazione SMEFT



- Le misure differenziali sono utilizzate per costringere l'estensione del SM che introduce interazioni anomale tra il bosone di Higgs e le particelle del SM utilizzando un approccio EFT
- Lagrangiana SMEFT:

$$\mathcal{L}_{\text{SMEFT}} = \mathcal{L}_{\text{SM}} + \mathcal{L}_{\text{EFT}} = \mathcal{L}_{\text{SM}} + \sum_{i=5}^{\infty} \sum_j^{N_i} \frac{c_j^{(i)}}{\Lambda^{i-4}} \mathcal{O}_j^{(i)}$$

*coefficienti di Wilson*

Valori dei WC diversi da 0  $\Rightarrow$  possibile discrepanza dal SM

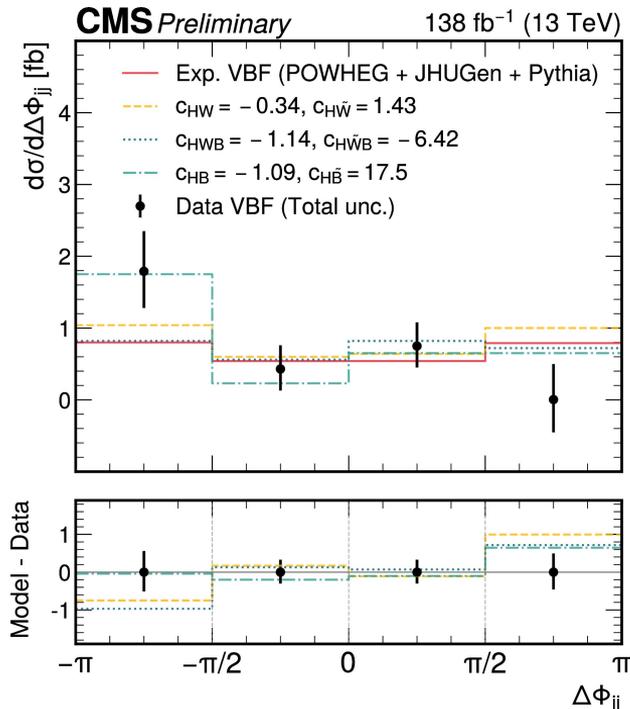
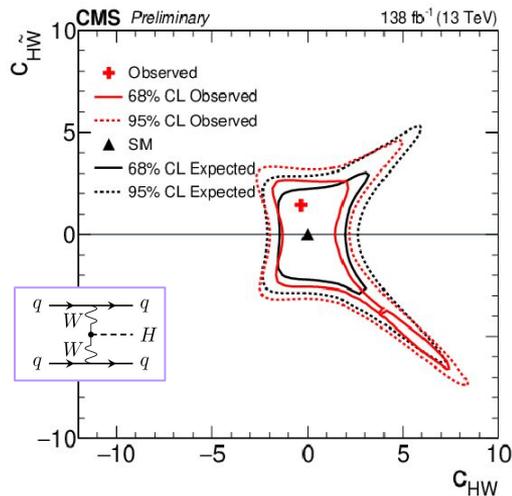
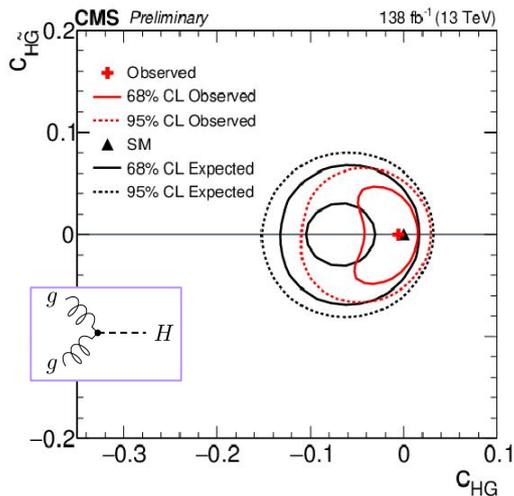
- L'obiettivo è studiare un sottoinsieme di operatori  $i=6$  e vincolare i corrispondenti coefficienti di Wilson nella base di Warsaw

Class	Operator	Wilson Coefficient
$\mathcal{L}_6^{(4)} - X^2 H^2$ (CP-even)	$H^\dagger H W_{\mu\nu}^i W^{i\mu\nu}$	$c_{\text{HW}}$
	$H^\dagger H W_{\mu\nu}^i B^{i\mu\nu}$	$c_{\text{HWB}}$
	$H^\dagger H B_{\mu\nu} B^{\mu\nu}$	$c_{\text{HB}}$
	$H^\dagger H G_{\mu\nu}^a G^{a\mu\nu}$	$c_{\text{HG}}$
$\mathcal{L}_6^{(4)} - X^2 H^2$ (CP-odd)	$H^\dagger H \tilde{W}_{\mu\nu}^i W^{i\mu\nu}$	$c_{\text{H}\tilde{W}}$
	$H^\dagger H \tilde{W}_{\mu\nu}^i B^{i\mu\nu}$	$c_{\text{H}\tilde{W}B}$
	$H^\dagger H \tilde{B}_{\mu\nu} B^{\mu\nu}$	$c_{\text{H}\tilde{B}}$
	$H^\dagger H \tilde{G}_{\mu\nu}^a G^{a\mu\nu}$	$c_{\text{H}\tilde{G}}$
$\mathcal{L}_6^{(3)} - H^4 D^2$	$(H^\dagger H) \square (H^\dagger H)$	$c_{\text{H}\square}$
	$(D^\mu H^\dagger H) (H^\dagger D_\mu H)$	$c_{\text{HD}}$

# Limiti sui coefficienti di Wilson



- Vincoli più stringenti ottenuti per la misura della sezione d'urto VBF sotto l'assunzione del coefficiente (CP-pari)  $c_{HW}$ , e per la misura della sezione d'urto ggF, sensibile al coefficiente (CP-pari)  $c_{HG}$

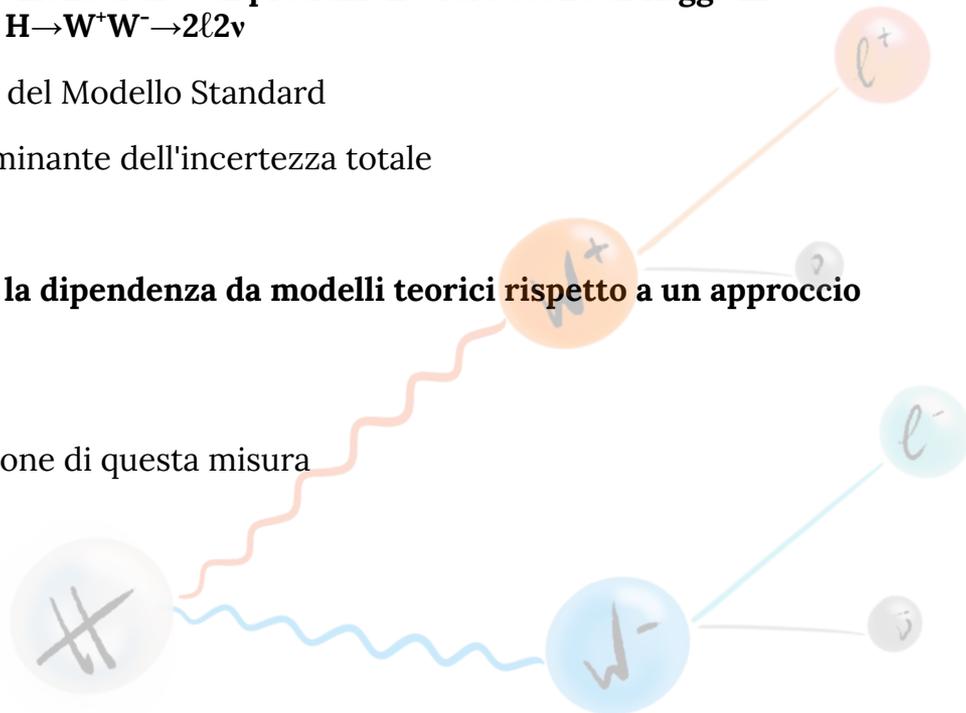


- Sezione d'urto differenziale ricalcolata utilizzando i valori misurati dei WC

# Conclusioni



- E' stata presentata la **misura differenziale della sezione d'urto di produzione del bosone di Higgs in associazione con 2 jet** nel canale di decadimento  $H \rightarrow W^+W^- \rightarrow 2\ell 2\nu$ 
  - I risultati sono in accordo con le previsioni del Modello Standard
  - L'incertezza statistica è la componente dominante dell'incertezza totale
- L'uso di una **rete neurale avversariale ha ridotto la dipendenza da modelli teorici rispetto a un approccio standard**
- Il Run 3 di LHC consentirà di migliorare la precisione di questa misura



*grazie per l'attenzione*

CMS-PAS-HIG-24-004

Back up

# Accoppiamenti anomali in HVV/Hgg

Ampiezza di scattering del bosone di Higgs ( $H$ ) con due bosoni di gauge ( $V_1 V_2$ ) con spin-1

$$A(HVV) \sim \left[ a_1^{VV} + \frac{k_1^{VV} q_{V_1}^2 + k_2^{VV} q_{V_2}^2}{(\Lambda_1^{VV})^2} \right] m_{V_1}^2 \epsilon_{V_1}^* \epsilon_{V_2}^* + a_2^{VV} f_{\mu\nu}^{*(1)} f^{*(2)\mu\nu} + a_3^{VV} f_{\mu\nu}^{*(1)} \tilde{f}^{*(2)\mu\nu}$$

## Accoppiamenti HVV

- $a_1^{VV} = 1$  accoppiamento del MS  $J^{CP}=0^{++}$
- $L_1$  CP-pari AC
- $a_2^{VV}$  CP-pari AC
- $a_3^{VV}$  CP-dispari AC

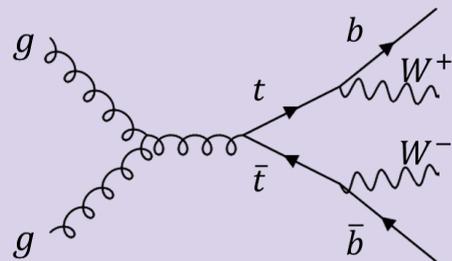
## Accoppiamento Hgg

- $a_2^{VV}$  SM loop
- $a_3^{VV}$  CP-dispari AC

# Principali eventi di fondo

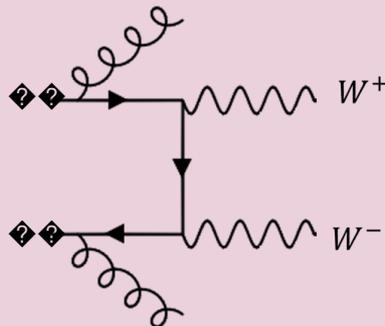


## Produzione di quark top



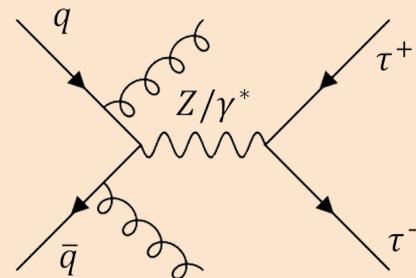
- Sezione d'urto grande
- Algoritmi di b-tagging per identificare i b-jets
- Normalizzazione misurata sui dati

## $W^+ W^-$ non risonante



- + 2 jets da ISR
- No correlazione di spin
- Bosoni  $W$  on-shell
- Cinematica diversa ( $m_{\ell\ell}$  più grande del segnale)

## Drell-Yan $\tau^+ \tau^-$

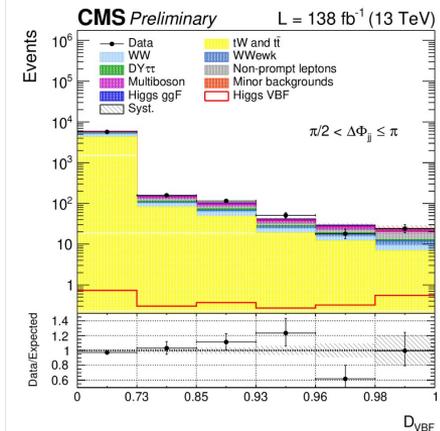
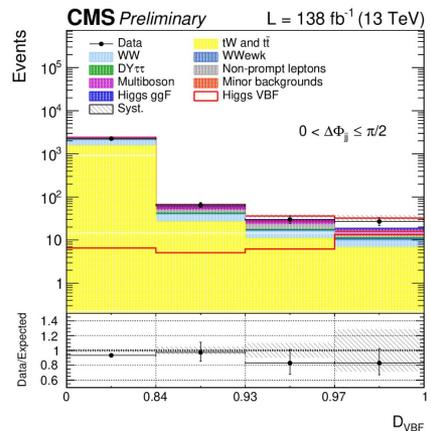
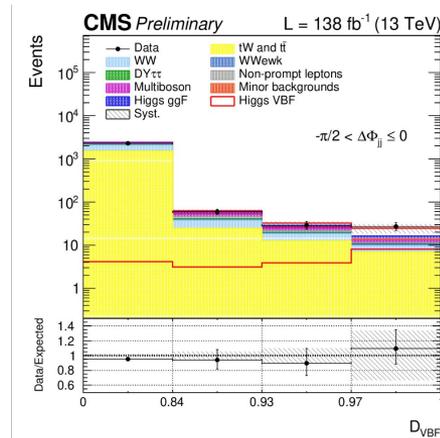
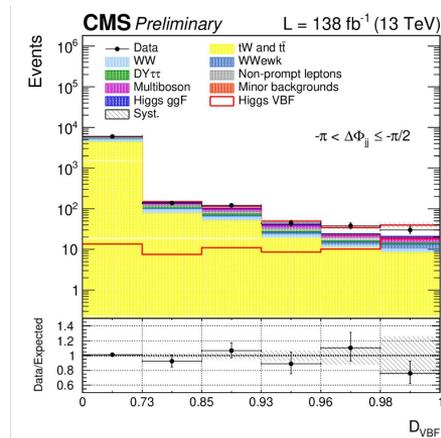


- + 2 jet da ISR
- $\tau \rightarrow e\nu\nu^-$  o  $\tau \rightarrow \mu\nu\nu^-$
- Ridotto con  $m_T > 60$  GeV
- Normalizzazione misurata sui dati

# Regione di segnale: $D_{\text{VBF}}$



- **Distribuzioni del discriminante  $D_{\text{VBF}}$  nei bin di  $\Delta\Phi_{jj}$  della regione di segnale (scala logaritmica)**

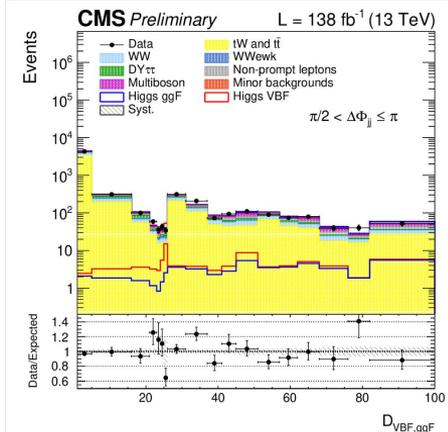
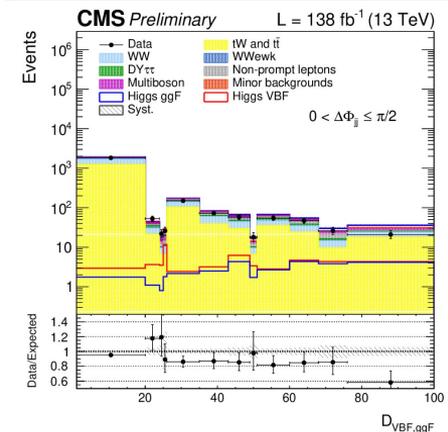
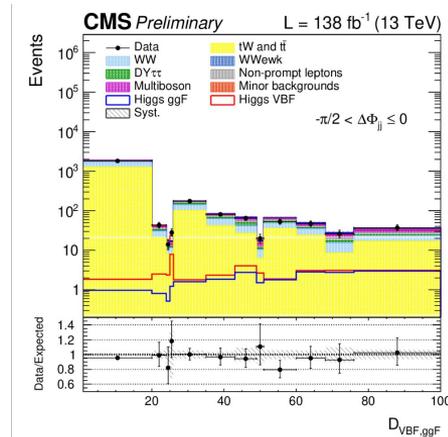
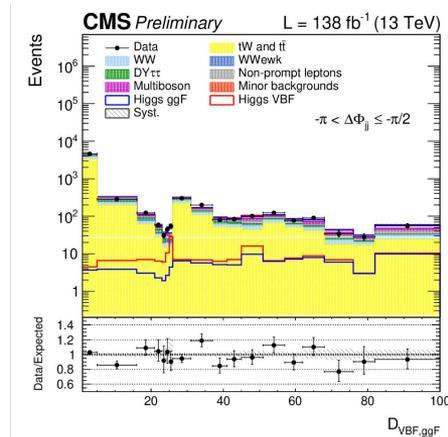
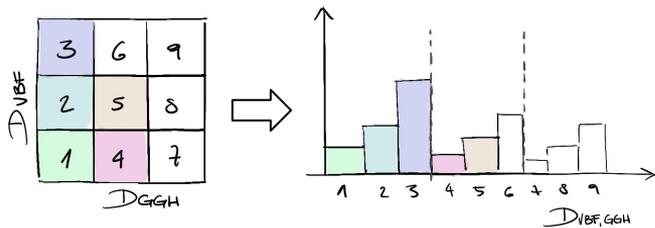


# Regione di segnale: $D_{VBF,ggF}$



- **Distribuzioni del discriminante bidimensionale  $D_{VBF,ggF}$  nei bin di  $\Delta\Phi_{jj}$  della regione di segnale** (scala logaritmica)
  - Distribuzione trasposta in un dimensione

Il contenuto di ciascun bin della distribuzione 2D viene trasposto in un istogramma 1D, in cui le etichette dei bin sull'asse x corrispondono alla numerazione dei bin della mappa 2D



# Incertezze di misura

- Contributi delle diverse sorgenti di incertezza, espressi come percentuale dell'incertezza totale ( $\Delta\sigma_i/\Delta\sigma_{\text{tot}} \times 100$ ).
- Per errori asimmetrici, viene riportata l'incertezza maggiore tra quella verso l'alto e quella verso il basso.
- La componente sistematica include tutte le sorgenti, ad eccezione della normalizzazione del fondo, che fa parte della componente statistica.

Uncertainty source	$-\pi < \Delta\Phi_{ij} \leq -\frac{\pi}{2}$	$-\frac{\pi}{2} < \Delta\Phi_{ij} \leq 0$	$0 < \Delta\Phi_{ij} \leq \frac{\pi}{2}$	$\frac{\pi}{2} < \Delta\Phi_{ij} \leq \pi$
Theory (sig.)	18%	16%	16%	17%
Theory (backg.)	30%	17%	16%	27%
Integrated luminosity	16%	17%	15%	14%
b tagging	16%	14%	15%	13%
Jets	15%	15%	13%	14%
Non-prompt	14%	15%	15%	17%
PileUp	15%	15%	13%	16%
Leptons	17%	15%	13%	15%
MET	14%	15%	13%	14%
Backg. norm.	14%	15%	13%	14%
MC stat.	31%	31%	32%	40%
Stat. uncertainty	91%	93%	93%	86%
Syst. uncertainty	46%	38%	39%	52%

# Volume fiduciale

- Gli stessi tagli dell'SR vengono applicati alle quantità definite a livello gen, ad eccezione dei requisiti MET e  $m_T^{\ell 2}$
- I leptoni sono “dressed”, ovvero gli impulsi dei fotoni irradiati all'interno di un cono di  $\Delta R < 0,1$  attorno alla direzione del leptone vengono sommati all'impulso del leptone
- I campioni di segnale vengono suddivisi in contenitori  $\Delta\Phi_{jj}$ , utilizzando lo stesso schema di suddivisione delle categorie di analisi

---

Fiducial requirements
oppositely-charged $e\mu$ (not from $\tau$ decay) final state, at least two jets with $p_T > 30$ GeV, $p_T^{\ell 1} > 25$ GeV, $p_T^{\ell 2} > 13$ GeV, $p_T^{\ell 3} < 10$ GeV, $p_T^{\ell\ell} > 30$ GeV, $m_{\ell\ell} > 12$ GeV, $m_{jj} > 120$ GeV, $m_T > 60$ GeV, $ \eta_{j_1}  < 4.7$ , $ \eta_{j_2}  < 4.7$ , $ \eta_{\ell_1}  < 2.5$ , $ \eta_{\ell_2}  < 2.5$

---

# Likelihood

- Il segnale viene estratto tramite un **maximum likelihood template fit**
- Gli eventi seguono una distribuzione poissoniana a media  $(s_i + b_i)$ 
  - Il **numero di eventi** nelle CR top e DY viene incluso nel fit per misurare la normalizzazione del fondo corrispondente
- Parametri di interesse:
  - **signal strength modifier per ogni componente gen-level del segnale**

$$\mu_j = \frac{\sigma_j}{\sigma_j^{exp}}$$

*Incertezze sistematiche*

$$L(\boldsymbol{\mu}, \boldsymbol{\nu}) = \prod_{i=0}^N \mathcal{P}(N_{i,obs} ; s_i(\boldsymbol{\mu}, \boldsymbol{\nu}) + b_i(\boldsymbol{\nu})) \cdot \mathcal{N}(\boldsymbol{\nu})$$

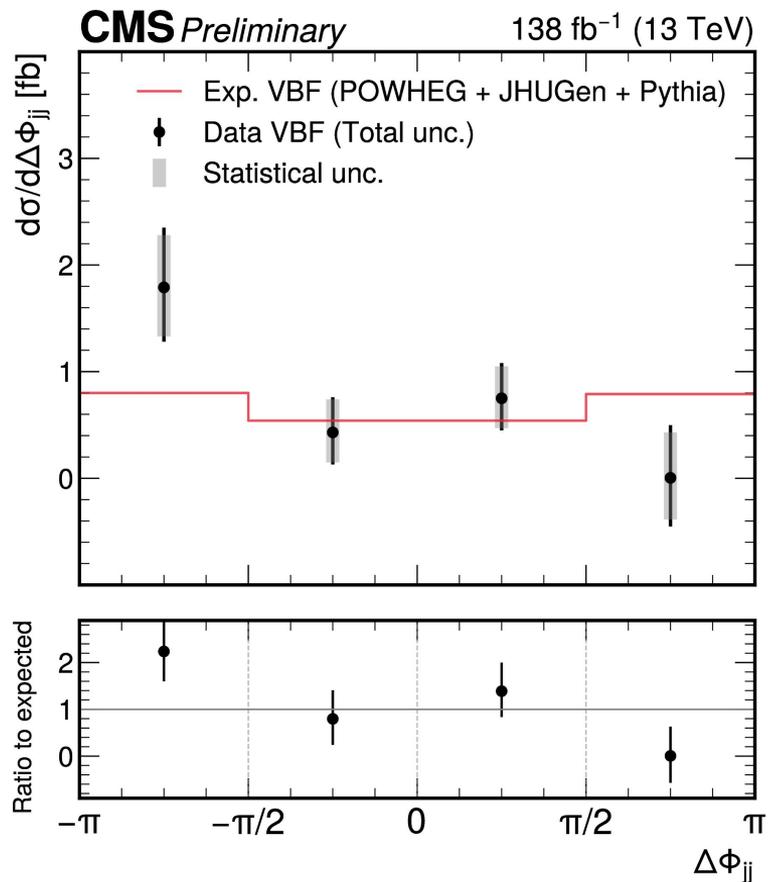
$$L(\boldsymbol{\mu}, \boldsymbol{\nu}) = \prod_{i=0}^{N_{bins}^{reco}} \mathcal{P}\left(N_{i,obs} ; \underbrace{\sum_{j=0}^{N_{bins}^{gen}} (R_{ij}^{fid} \mu_j s_j^{fid} + R_{ij}^{non-fid} \mu_j s_j^{non-fid})}_{s_i} + b_i\right) \cdot \mathcal{N}(\boldsymbol{\nu})$$

*$\mu$  introduced for each generator-level signal*

# Configurazioni di fit

Fit	Misura	POI	Variabile	Motivazione
1	VBF e ggF misurati insieme	VBF e ggF scalati usando lo stesso POI ( $\mu_{i,\text{VBF+ggF}}$ con $i=0,\dots,3$ )	$D_{\text{VBF,ggF}}$	Sezione d'urto H+2jets
2	VBF e ggF come processi indipendenti	VBF e ggF con diversi POIs ( $\mu_{i,\text{VBF}}$ e $\mu_{i,\text{ggF}}$ with $i=0,\dots,3$ )	$D_{\text{VBF,ggF}}$	Misura indipendente dal modello
3	Sezione d'urto VBF	Solo VBF ( $\mu_{i,\text{VBF}}$ con $i=0,\dots,3$ )	$D_{\text{VBF}}$	Permette i limiti più stringenti sugli accoppiamenti SMEFT

# VBF differential cross section



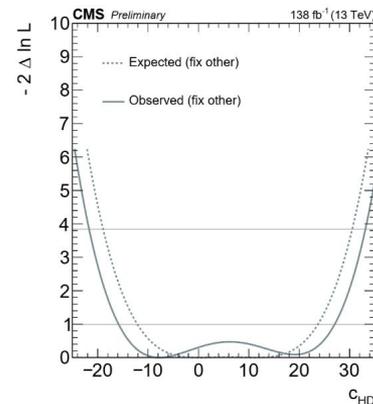
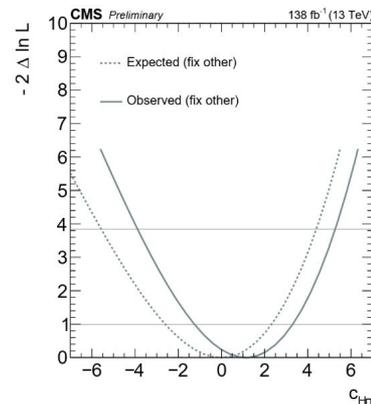
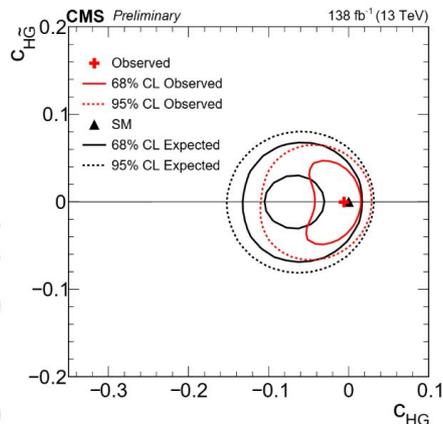
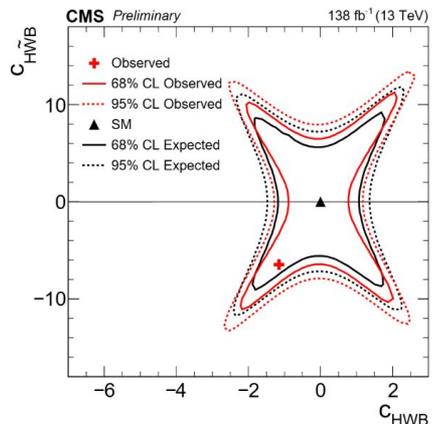
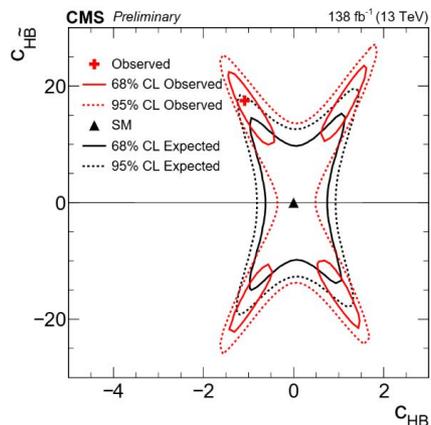
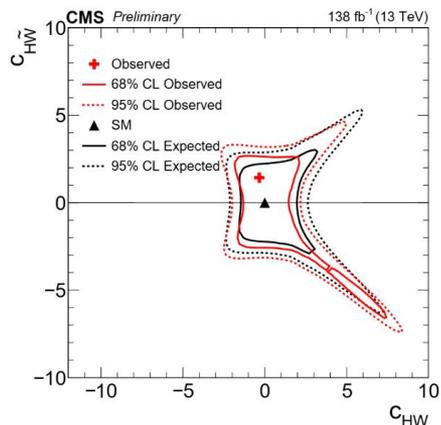
# Asimmetria

- Misura della asimmetria della distribuzione di  $\Delta\Phi_{jj}$

$$A = \frac{N(\Delta\Phi_{jj} > 0) - N(\Delta\Phi_{jj} \leq 0)}{N(\Delta\Phi_{jj} > 0) + N(\Delta\Phi_{jj} \leq 0)}$$

$$A = 0.4 \pm 0.3 \text{ (tot. unc.)}$$

# Coefficienti di Wilson



# Coefficienti di Wilson

Wilson Coefficients		Observed	Expected	Significance
$c_{H\Box}$	Best fit	1.14	0.00	0.5
	68% CL	[-1.26,3.31]	[-2.60,2.33]	
	95% CL	[-3.88,5.27]	[-5.56,4.39]	
$c_{HD}$	Best fit	-7.76	0.00	0.0
	68% CL	[-15.8,27.0]	[-12.1,23.7]	
	95% CL	[-21.7,33.1]	[-19.0,30.5]	
$c_{HW}$	Best fit	-0.34	0.00	0.0
	68% CL	[-1.46,1.15]	[-1.17,1.76]	
	95% CL	[-2.35,3.57]	[-2.07,6.50]	
$c_{HWB}$	Best fit	-1.14	0.00	0.5
	68% CL	[-1.92,1.82]	[-1.47,1.50]	
	95% CL	[-2.44,2.34]	[-2.06,2.10]	
$c_{HB}$	Best fit	-1.09	0.00	1.7
	68% CL	[-1.35,-0.72] $\cup$ [0.92,1.51]	[-0.80,0.96]	
	95% CL	[-1.56,1.75]	[-1.13,1.27]	
$c_{HG}$	Best fit	-0.01	0.00	0.0
	68% CL	[-0.03,0.01]	[-0.10,0.02]	
	95% CL	[-0.08,0.02]	[-0.14,0.03]	
$c_{H\bar{W}}$	Best fit	1.43	0.00	0.7
	68% CL	[-5.95,-5.22] $\cup$ [-2.21,2.32]	[-1.87,1.91]	
	95% CL	[-7.12,3.00]	[-5.60,4.53]	
$c_{H\bar{W}B}$	Best fit	-6.42	0.00	0.9
	68% CL	[-9.84,8.90]	[-7.40,7.52]	
	95% CL	[-12.3,11.4]	[-10.51,10.62]	
$c_{H\bar{B}}$	Best fit	17.5	0.00	2.3
	68% CL	[-20.6,-14.5] $\cup$ [13.0,21.1]	[-12.2,11.9]	
	95% CL	[-24.5,-5.9] $\cup$ [6.0,24.1]	[-17.3,16.6]	
$c_{H\hat{G}}$	Best fit	-0.000	0.00	0.0
	68% CL	[-0.035,0.034]	[-0.062,0.061]	
	95% CL	[-0.058,0.057]	[-0.075,0.073]	

# Sezioni d'urto reinterpretate

