



Risultati recenti sulle misure VBS e VBF dell'esperimento CMS

Incontri di Fisica delle Alte Energie
03/04/2024

Andrea Claudio Maria Bulla¹
per la collaborazione CMS

¹ Università e INFN di Cagliari



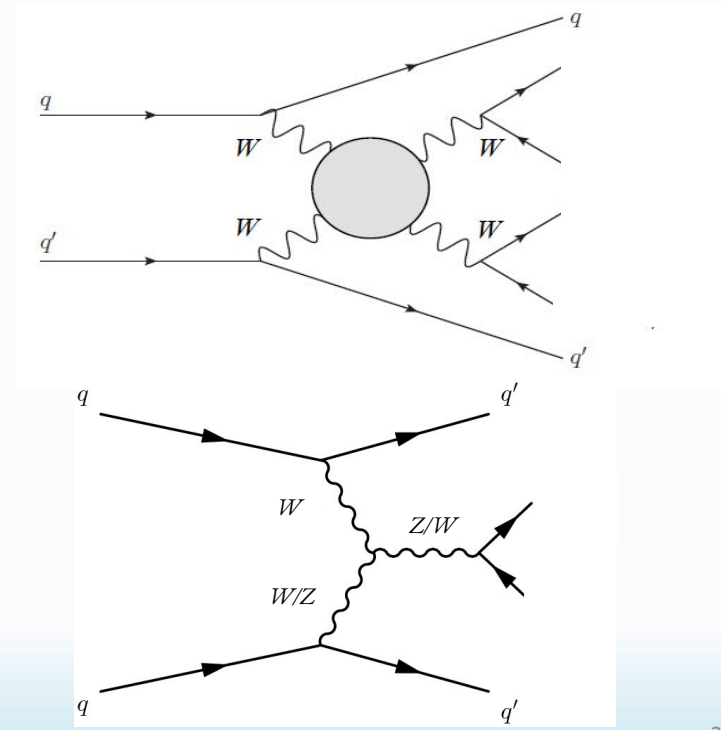
Introduzione



I processi Vector Boson Scattering (VBS) e Vector Boson Fusion (VBF) accadono quando i due quark iniziali irradiano bosoni vettori che interagiscono tra di loro

- Sono processi rari previsti dallo standard model (~ 10-1000 fb)
- Sono profondamente connessi con EWSSB
- Hanno una ricca fenomenologia (TGC, QGC, accoppiamento con h, EFT..)
- Sono caratterizzati da peculiarità cinematiche comuni ($|\Delta\eta_{jj}|, m_{jj}, Z_X$)

$$Z_X = \eta_X - \frac{1}{2}(\eta_{j_1} + \eta_{j_2})$$



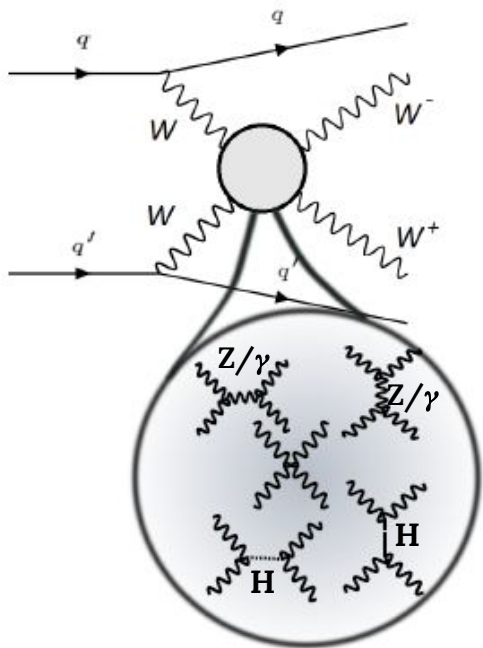


VBS \rightarrow W^+W^- leptonic



“First observation of the electroweak production of a leptonically decaying W^+W^- pair in association with two jets in $\sqrt{s} = 13$ TeV pp collisions”.

[Phys. Lett B 841 \(2023\) 137495](#)



Regione inclusiva

Misura: 99 ± 20 fb
LO predizione: 89 ± 5 fb

Osservabili
parton-level:
 $m_{qq} > 100$ GeV
 $p_T^i > 10$ GeV
 $i=l,q,v$

**Ottimo
accordo con
le predizioni
dello SM!**

Regione fiduciale

Misura: 10.2 ± 2.0 fb
LO predizione: 9.1 ± 0.6 fb

Objects	Requirements
	$e^+e^-, \mu^+\mu^-, e^+\mu^-, e^-\mu^+$
Leptons	$p_T^\ell = p_T^{\text{bare } \ell} + \sum_i p_T^{\gamma_i}$ if $\Delta R(\ell, \gamma_i) < 0.1$ $p_T^{\ell_1} > 25$ GeV, $p_T^{\ell_2} > 13$ GeV, $p_T^{\ell_3} < 10$ GeV $ \eta < 2.5$ $p_T^{\ell\ell} > 30$ GeV, $m_{\ell\ell} > 50$ GeV
Jets	$p_T^j > 30$ GeV, $ \eta < 4.7$ $\Delta R(j, \ell) > 0.4$ At least 2 jets, no b jets $m_{ij} > 300$ GeV, $\Delta\eta_{ij} > 2.5$
p_T^{miss}	$p_T^{\text{miss}} > 20$ GeV

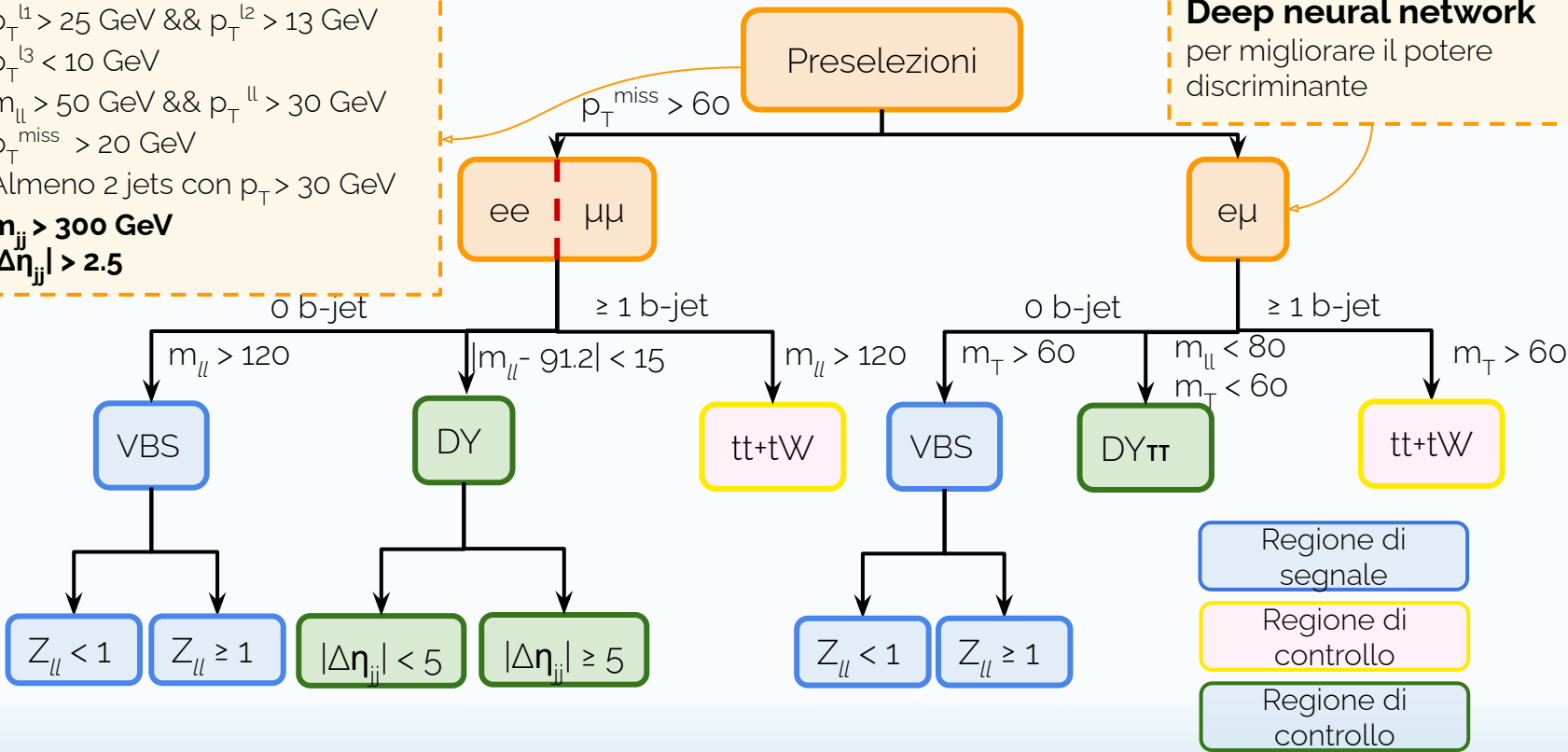


VBS \rightarrow W^+W^- leptonic



- $p_T^{l1} > 25$ GeV && $p_T^{l2} > 13$ GeV
- $p_T^{l3} < 10$ GeV
- $m_{ll} > 50$ GeV && $p_T^{ll} > 30$ GeV
- $p_T^{miss} > 20$ GeV
- Almeno 2 jets con $p_T > 30$ GeV
- **$m_{jj} > 300$ GeV**
- **$|\Delta\eta_{jj}| > 2.5$**

Deep neural network
per migliorare il potere discriminante



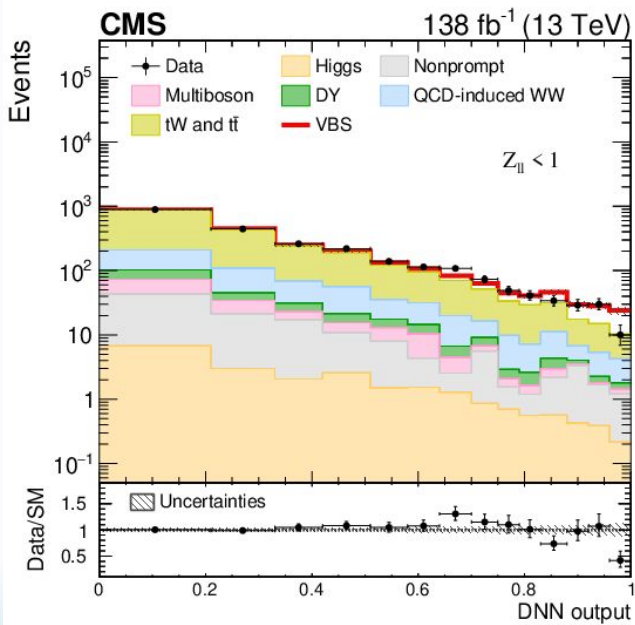


VBS \rightarrow W^+W^- leptonic

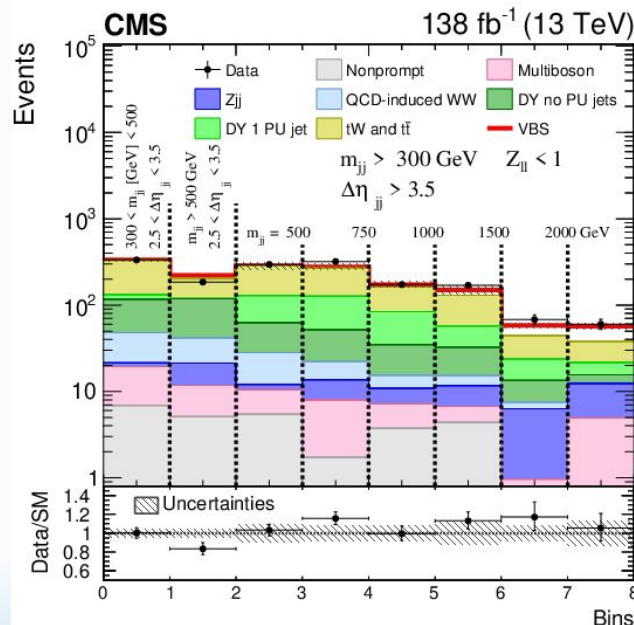


Combined binned maximum likelihood fit della distribuzione più discriminante con istogrammi di segnale e fondo

$$\mathcal{L}(\vec{n}|\mu, \theta) = p(\theta) \cdot \prod_{i=1}^N \text{Poisson}(n_i|\mu \cdot s_i(\theta) + b_i(\theta))$$



Tutte le categorie sono incluse simultaneamente nel fit





VBS \rightarrow W^+W^- leptónico



- Analisi statisticamente limitata \rightarrow Ci si aspetta che possa beneficiare di un maggior dataset con Run III (250 fb^{-1})
- Futuri sviluppi:
 - Studi di polarizzazione per investigare EWSSB e testare modelli di fisica BSM
 - Global fit di operatori EFT di dimensione 6 per avere una più completa comprensione del range di validità dello SM ([EFT dim6 study @ LHE](#))

	Uncertainty source	Impact
Incertezze teoriche	QCD-induced W^+W^- normalisation	5.3%
	$t\bar{t}$ QCD scale	5.1%
	QCD factorisation scale for VBS signal	5.0%
	$t\bar{t}$ normalisation	4.9%
Incertezze legate al b-tagging	b tagging	3.5%
	Prefiring corrections	3.3%
	DY normalisation	2.9%
	Jet energy scale + resolution	2.6%
	p_T^{miss} energy scale	2.4%
	QCD-induced W^+W^- QCD scale	2.1%
	Luminosity	2.1%
	Muon efficiency	2.0%
	Pileup	1.8%
	Electron efficiency	1.5%
	Underlying event	1.3%
Analisi dominata da incertezza statistica	Parton shower	1.0%
	Other	< 1%
	Total systematic uncertainty	13.1%
	Total statistical uncertainty	14.9%
	Total uncertainty	19.8%

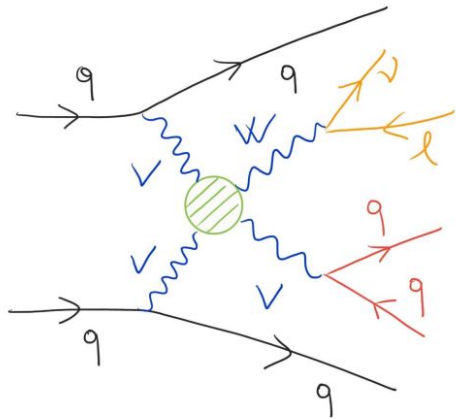


VBS \rightarrow WW semileptonico



“First evidence for WW/WZ vector boson scattering in the decay channel $lvqq$ produced in association with two jets in proton-proton collisions at $\sqrt{s} = 13$ TeV”.

[Phys. Lett. B 834 \(2022\) 137438](#)



σ_{EWK} only

Misura: 1.9 ± 0.5 pb

Predizione: 2.23 ± 0.10 pb

Significance: 4.4σ (5.1 exp)

Osservabili parton-level:

$$m_{qq} > 100 \text{ GeV}$$

$$p_T^i > 10 \text{ GeV}$$

$$i=l,q,\nu$$

**Ottimo
accordo con
le predizioni
dello SM!**

$\sigma_{EWK+QCD}$

Misura: 16.3 ± 3.2 pb

Predizione: 16.9 ± 2.5 pb

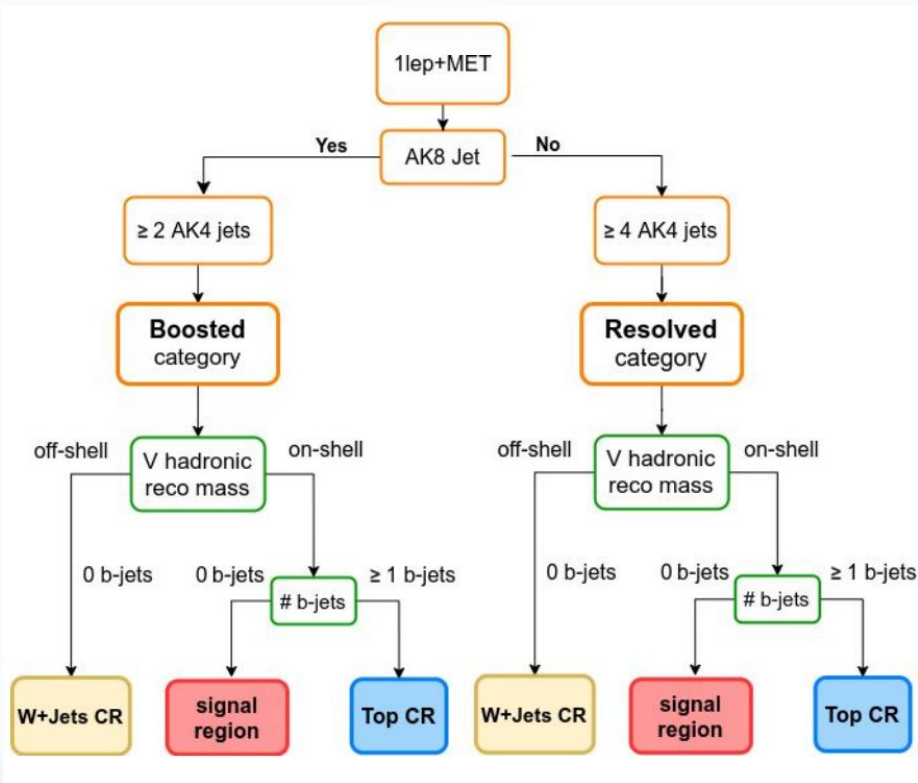
Osservabili parton-level:

$$m_{qq} > 100 \text{ GeV}$$

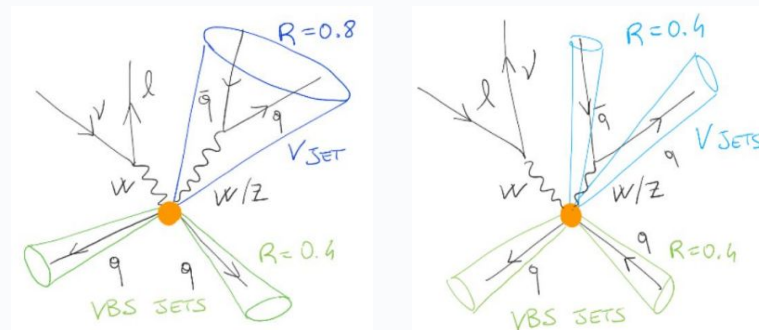
$$p_T^i > 10 \text{ GeV}$$

$$i=l,q,\nu$$

VBS \rightarrow WV semileptonico



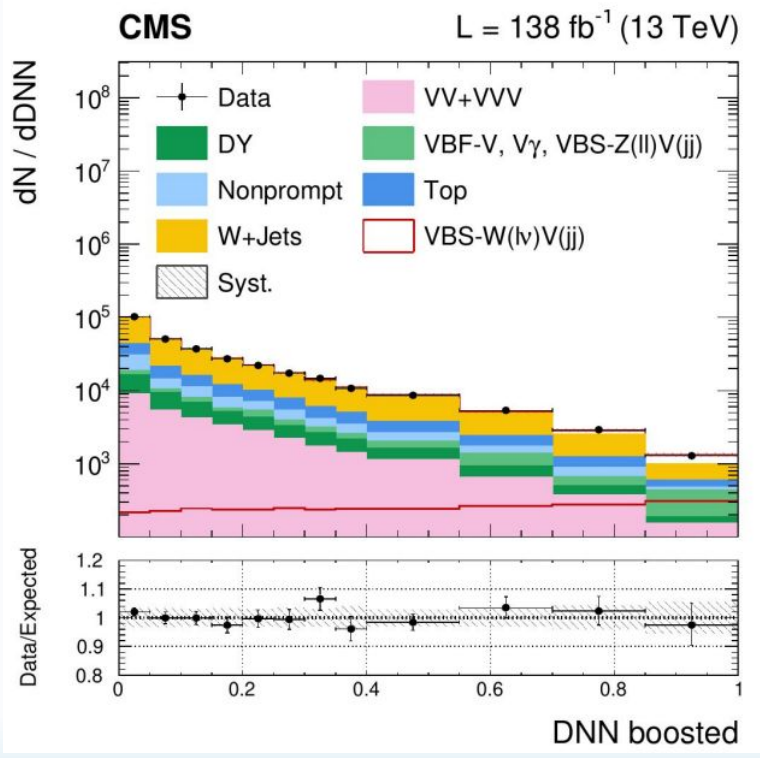
Regioni spittate a seconda dell'energia dei getti: boosted ($R=0.8$), resolved ($R=0.4$)



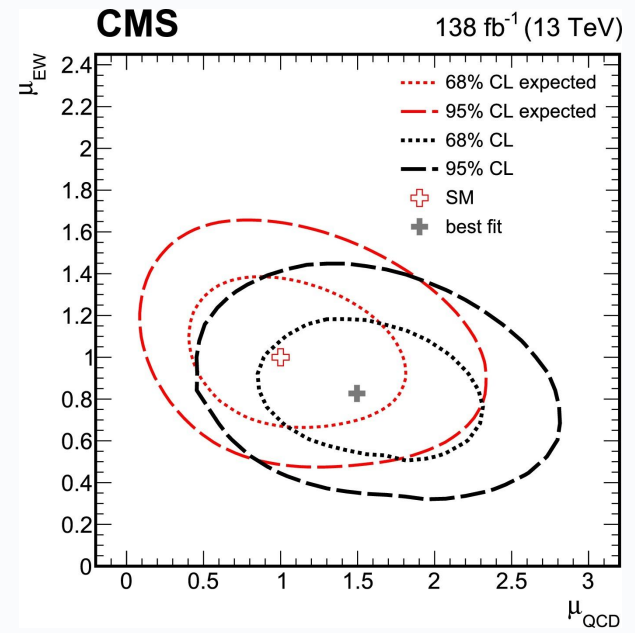
Anche in questo caso, nelle regioni di segnale si addestrano due DNN per aumentare il potere discriminante



VBS \rightarrow WV semileptonico



Tutte le categorie sono incluse simultaneamente nel fit





VBS \rightarrow WV semileptonico



Analisi dominata
da incertezza
statistica

- Anche questa analisi è statisticamente limitata \rightarrow Run III
- Futuri sviluppi:
 - Studi di polarizzazione per investigare EWSSB e testare modelli di fisica BSM
 - Global fit di operatori EFT di dimensione 6. Il canale è un ottimo candidato ed è avvantaggiato dalla maggiore statistica del decadimento adronico

Uncertainty source	$\Delta\mu_{EW}$
Statistical	0.12
Limited sample size	0.10
Normalization of backgrounds	0.08
Experimental	
b-tagging	0.05
Jet energy scale and resolution	0.04
Integrated luminosity	0.01
Lepton identification	0.01
Boosted V boson identification	0.01
Total	0.06
Theory	
Signal modeling	0.09
Background modeling	0.08
Total	0.12
Total	0.22



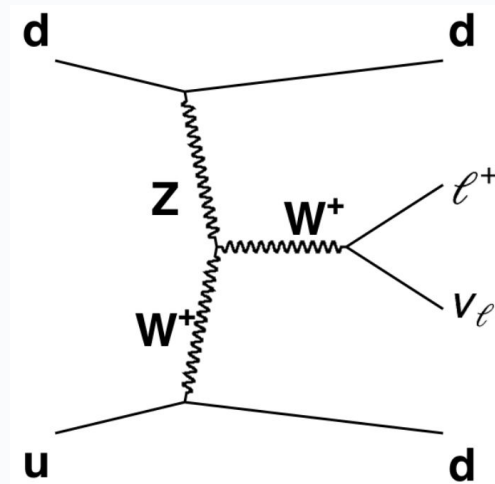
VBF \rightarrow W leptonico



- Misura inclusiva e fiduciale di sezione d'urto per la produzione EWK di W utilizzando tutto il dataset di Run II nel canale leptonico
- Misura differenziale di sezione d'urto in funzione di alcune variabili cinematiche chiave:

$$m_{jj}, \Delta\phi_{jj}, \Delta\eta_{jj}, pT_j, \eta_j, \eta_l, \phi_l, pT_l, mT(W), pT(W), ..$$

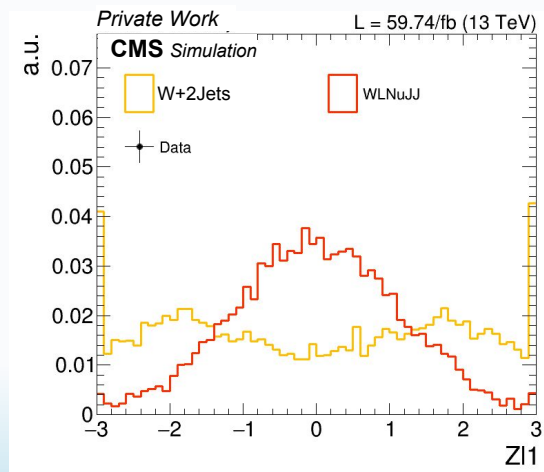
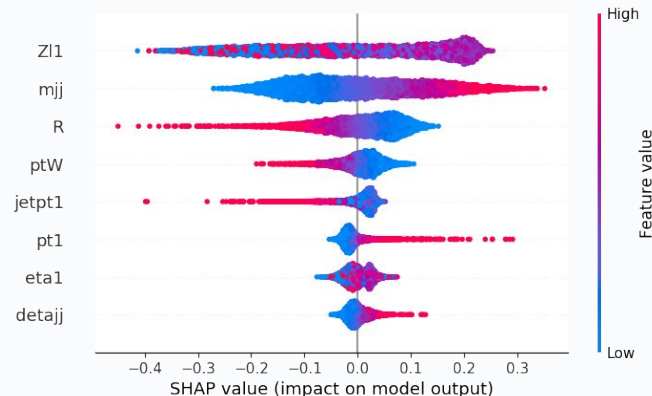
- Ricerche indirette di nuova fisica tramite un approccio EFT con operatori di dimensione 6:





To disentangle signal from **W+2Jets** background in SR:

- **1 universal model:** implemented for both the lepton flavors
- **Training samples:** (70%-30%) dataset from SRs. 60k training, 20k splitted in test & validation. 24% enriched in signal
- **Architecture:** 3 hidden layers with (64,64,64) neurons
- **Set of training variables:** 8 physical-motivated variables



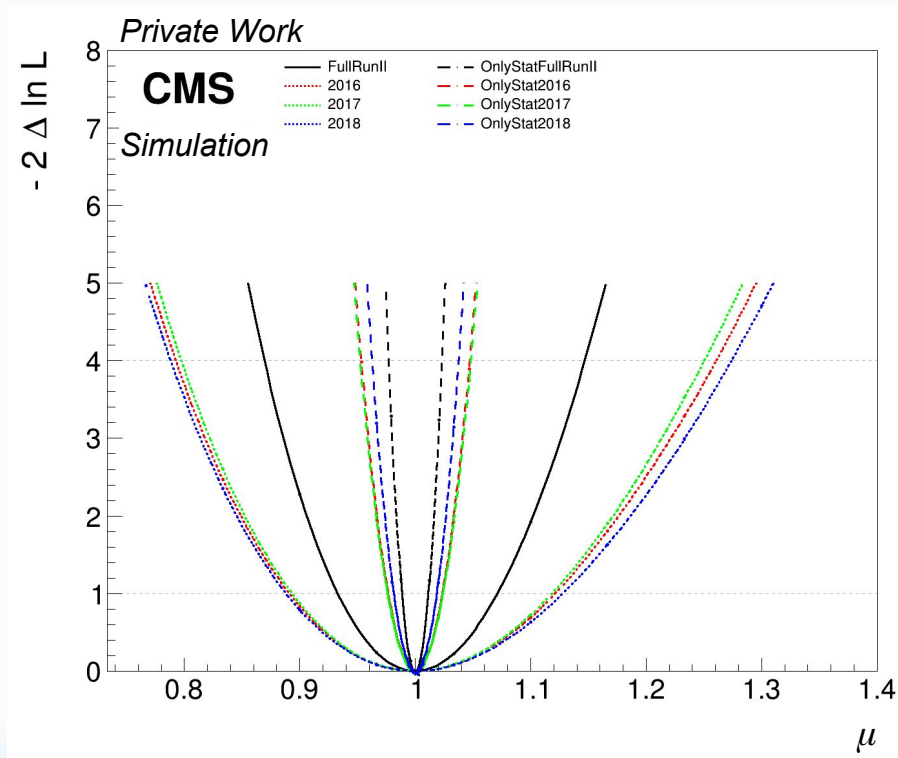


VBF \rightarrow W leptonico



$$\frac{\sigma_{exp}}{\sigma_{exp}} = 1.0_{-0.013}^{+0.013} (stat)_{-0.077}^{+0.083} (syst)$$

- Grazie ad una maggiore statistica ed a tecniche ML più avanzate, è atteso un miglioramento rispetto all'analisi che sfrutta solo i dati del 2016
- Tutti e tre gli anni ed entrambi i canali leptonici vengono combinati nel fit
- L'analisi è dominata dalle incertezze sistematiche \rightarrow il maggior contributo proviene dalla QCD scale variation del segnale
- Valutazione di un sample NLO di segnale per ridurre il suo impatto in corso

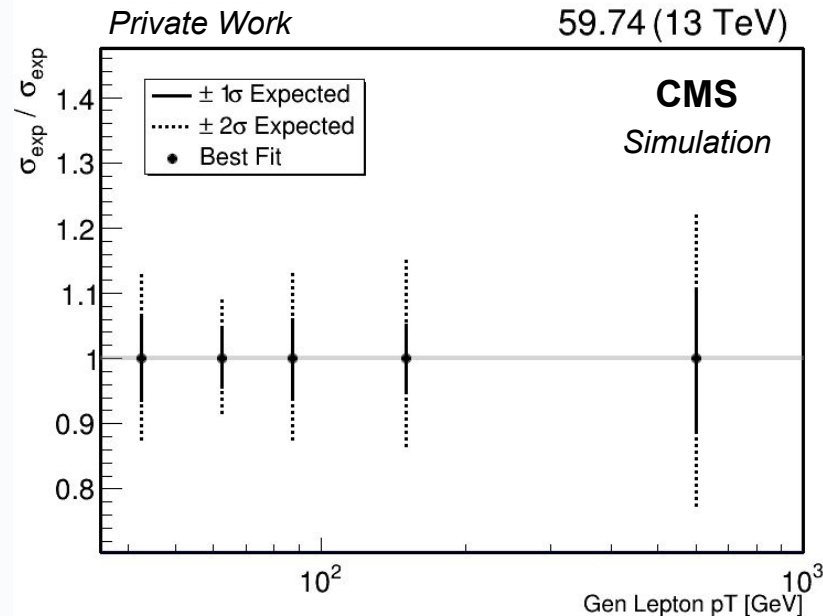




VBF \rightarrow W leptonico



- Goal: misura **differenziale** di cross-section in gen bins (in questo caso, p_T del leptone)
- Per **massimizzare** la **purezza** del segnale, viene sfruttata la DNN \rightarrow **2D** fit del **DNN output** e della corrispondente **variabile reco**
- Panifichiamo di fittare anche su variabili legate ai jets (p_{Tj} , η_j , ϕ_j), al sistema di-jet (m_{jj} , $\Delta\phi_{jj}$, $\Delta\eta_{jj}$), al leptone (p_{Tl} , η_l) e alla W (m_T , p_T)





VBF → W leptonico



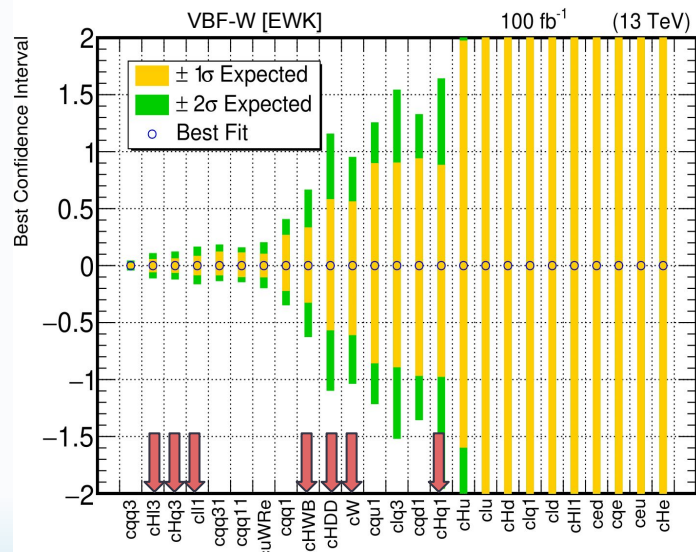
Solo 1 Coefficiente di Wilson è lasciato libero, tutti gli altri sono fissati al loro valore SM

$$N \propto |\mathcal{A}_{EFTSM}|^2 = \boxed{|\mathcal{A}_{SM}|^2} + \boxed{2c_i \text{Re}(\mathcal{A}_{SM} \mathcal{A}_i^*)} + \boxed{c_i^2 |\mathcal{A}_i|^2}$$

SM
Lineare
Quad

Per **ciascuna** variabile cinematica @ LHE, viene effettuato un fit **1D** e il risultato migliore viene plottato

Gli operatori bosonici sembrano essere **competitivi** rispetto ad altre analisi. Verranno effettuati ulteriori studi @ reco level

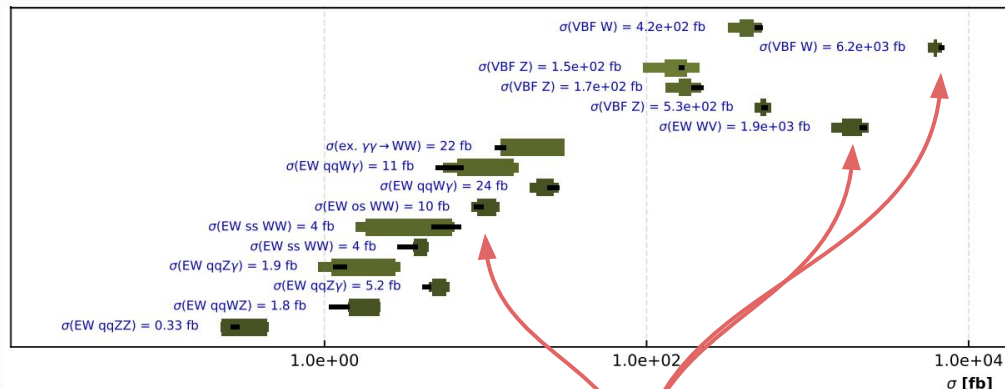




Conclusioni



- Presentati alcuni risultati tra i più recenti prodotti (e ongoing) da CMS per analisi VBS e VBF
- Grande partecipazione da parte della comunità italiana nelle analisi presentate
- Sforzo comune e collettivo per una nuova interpretazione di risultati SM, attraverso EFT
- Futura combinazione dei risultati EFT per costringere meglio il più ampio spettro possibile di operatori comuni



!!TALK DI OGGI!!



Backup



VBS \rightarrow W^+W^- leptonic



ttbar - tW:

$$\sigma_{tt} \sim 1 \text{ nb}$$

Maggior fondo dell'analisi.

Strategia: b jets veto.

QCD WW:

Interazione forte tra i quark dello stato iniziale.

Stesso stato finale ma caratterizzato da una cinematica differente.

Strategia: selezioni VBS.

Drell-Yan:

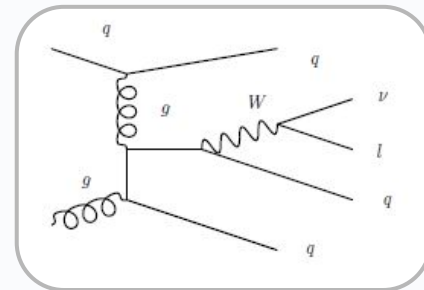
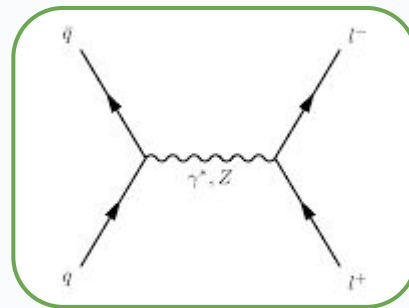
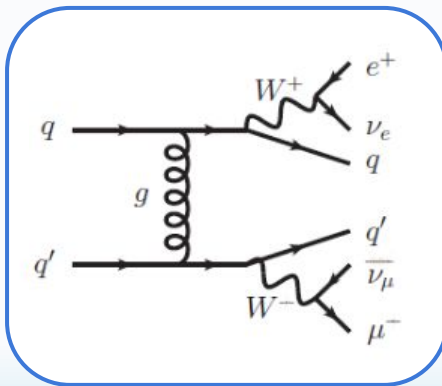
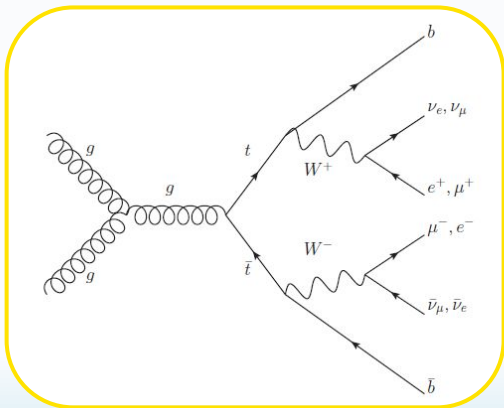
Per lo più presente nel canale SF.

Strategia: selezione sulla massa invariante e sul pT del sistema dileptoni, selezioni più dure su p_T^{miss} .

Non Prompt:

Per lo più $W^+ \text{Jet}$:

con un jet mis-identificato come leptone (fake lepton).





VBS \rightarrow W^+W^- leptónico



Per separare il segnale dai fondi **top** e **QCD-WW**:

- Solo per different flavour (DF) final state ($e\mu$);
- **2 modelli** implementati, stessa architettura e stesse variabili:
 - $Z_{ll} < 1$ phase space;
 - $Z_{ll} \geq 1$ phase space.
- **Training samples** (80%-20%) dataset separati dai test dataset (analysis datasets)
- **Architettura**: 5 hidden layers con (128,128,64,64,64) neuroni
- **Set di variabili di training**:
9 variabili con motivazioni fisiche

Variable	Description
m_{jj}	Invariant mass of the two VBS jets pair
$\Delta\eta_{jj}$	Pseudorapidity gap between the two VBS jets
p_{Tj_1}	p_T of the highest- p_T jet
p_{Tj_2}	p_T of the second highest p_T jet
$p_{T\ell\ell}$	p_T of the lepton pair
$\Delta\phi_{\ell\ell}$	Azimuthal angle between the two leptons
Z_{ℓ_1}	Zeppenfeld variable of the highest- p_T lepton
Z_{ℓ_2}	Zeppenfeld variable of the second highest p_T lepton
m_{TW_1}	Transverse mass of the $(p_{T\ell_1}, p_T^{\text{miss}})$ system

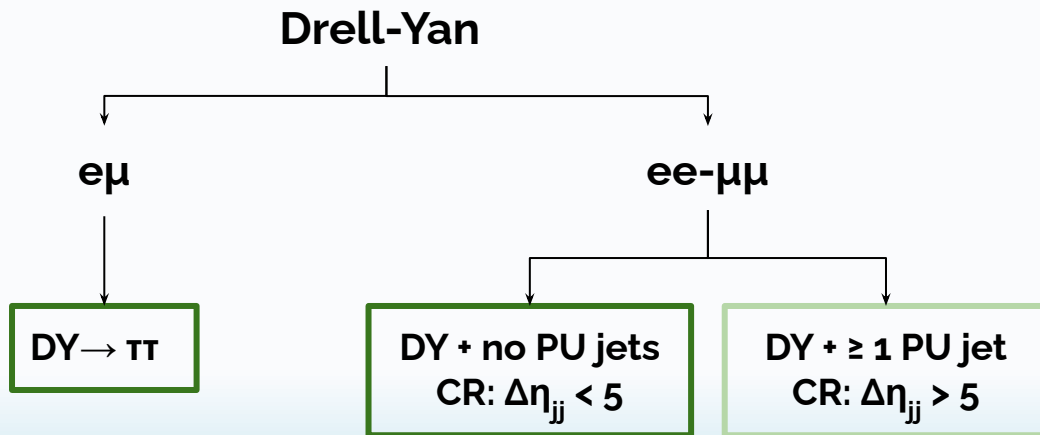


VBS \rightarrow W^+W^- leptonic



- In DY CRs sono chiaramente visibili 2 contributi differenti e molto sensibili rispetto alla distribuzione di $\Delta\eta_{jj}$:
 - Gli eventi "Hard" di DY popolano la regione a basso $\Delta\eta_{jj}$
 - Gli eventi DY in cui c'è almeno 1 PU jet piccano attorno a $\Delta\eta_{jj} \sim 5$
- 3 contributi totali in regioni di controllo differenti
- La loro normalizzazione è lasciata libera nel fit ed è maggiormente costretta dalle CRs dedicate $|\Delta\eta_{jj}| \geq 5$

PU jet = Reco jet
non-matched a nessun
 $p_T > 25$ GeV gen jet

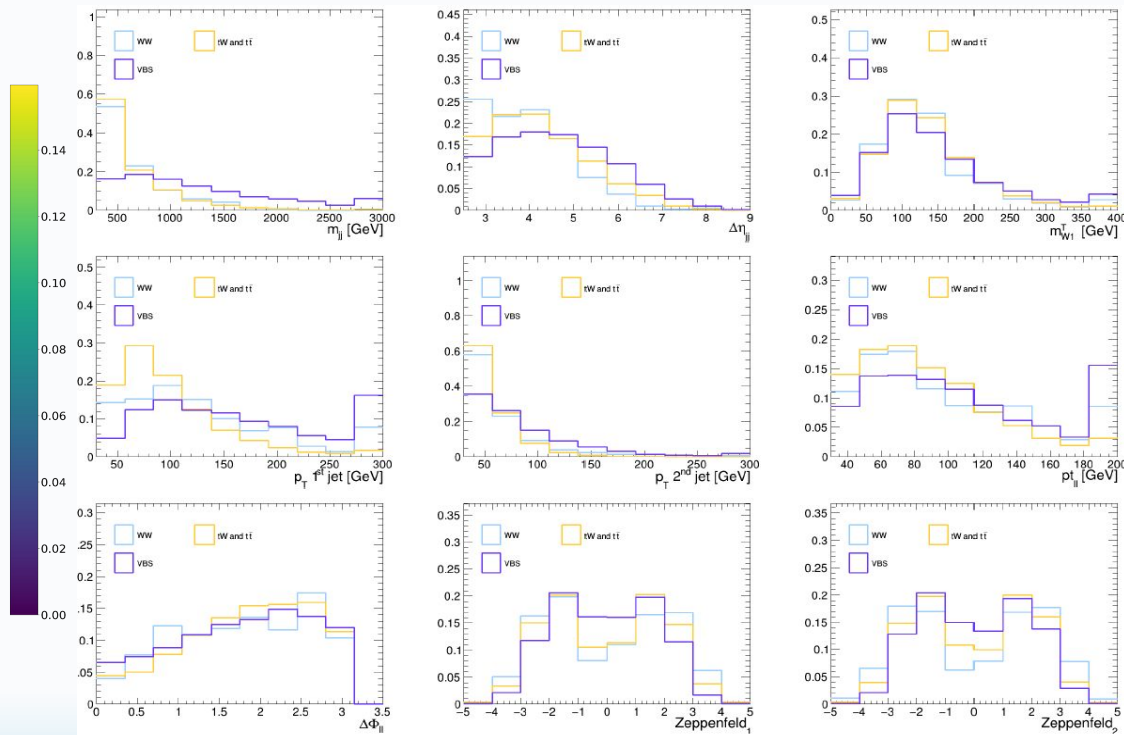
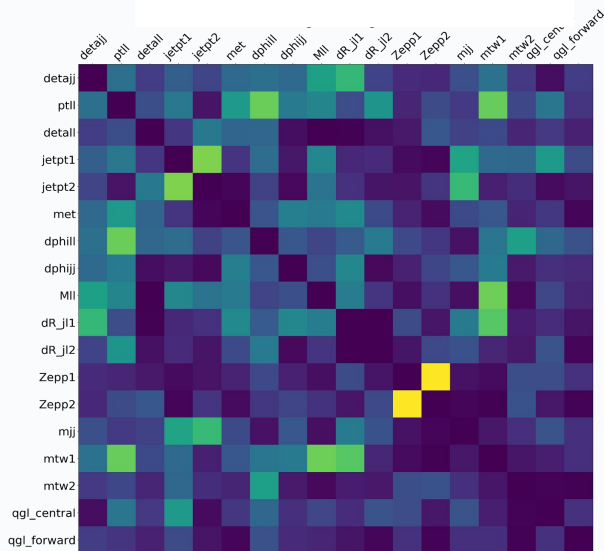




VBS \rightarrow W^+W^- leptónico



Background-signal matrice di correlazione



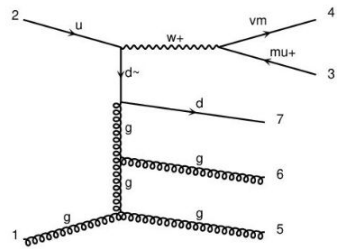


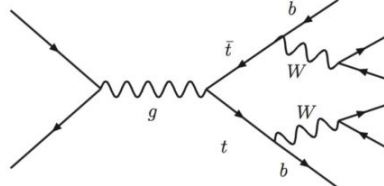
diagram 4 QCD=3, QED=2

W+ jets

Order $\alpha_{EWK}^2 \alpha_S^4$

$\sigma_{tot} \sim 61500 \text{ pb}$

Maggior fondo dell'analisi.
Misurato in regioni di controllo dedicate, tramite approccio data-driven

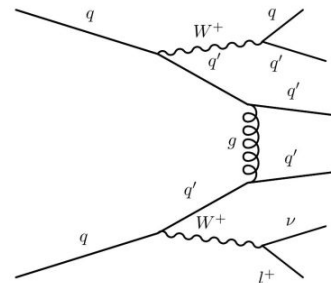


QCD $t\bar{t}$

Order $\alpha_{EWK}^4 \alpha_S^2$

$\sigma_{tot} \sim 450 \text{ pb}$

Ridotto applicando un b-quark veto nella regione di segnale e misurato in dedicate regioni di controllo



QCD diboson

Order $\alpha_{EWK}^4 \alpha_S^2$

$\sigma_{tot} \sim 15 \text{ pb}$

Difficile da distinguere dal segnale VBS.
Stimato dal MC

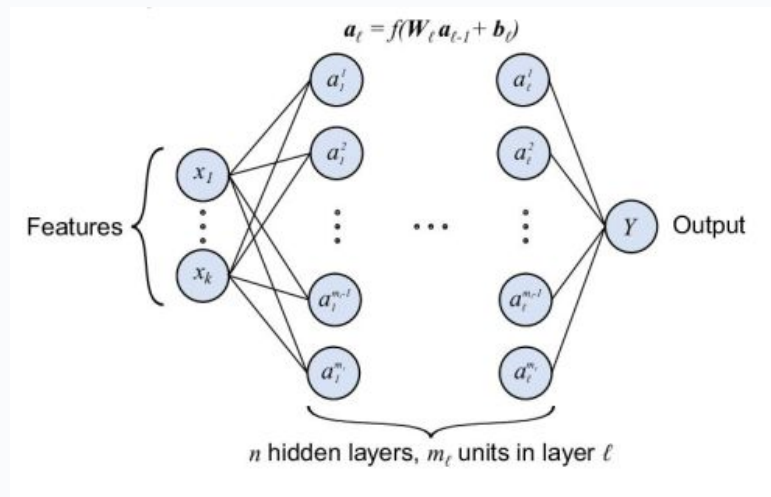


VBS \rightarrow WV semileptonico



Per separare il segnale da tutti i fondi:

- Solo per different flavour (DF) final state ($e\mu$);
- 2 modelli implementati:
 - Per categorie boosted e resolved
 - Dataset ottenuto da tutti e tre gli anni
- Training samples (80%-20%) dataset separati dai test dataset (analysis datasets)



Boosted

- DNN fully connected (64-32-32-32)
- 13 variabili di input
- 416k eventi di fondo, 50k di segnale

Resolved

- DNN fully connected (64-64-64-64)
- 16 variabili di input
- 1.7 M eventi di fondo, 220k di segnale

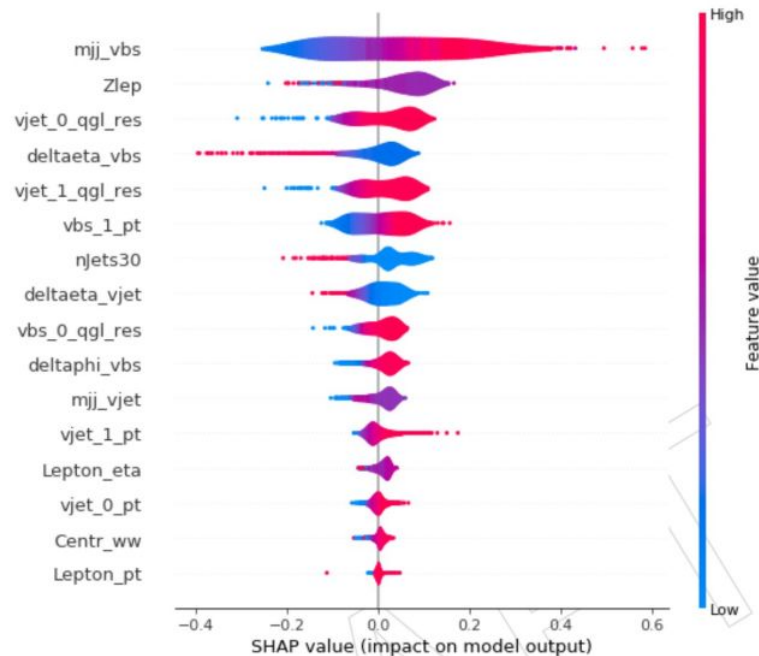


VBS \rightarrow WV semileptonico



Variable	Resolved	Boosted	SHAP ranking	
			Resolved	Boosted
Lepton pseudorapidity	✓	✓	13	12
Lepton transverse momentum	✓	✓	16	10
Zeppenfeld variable for the lepton	✓	✓	2	2
Number of jets with $p_T > 30$ GeV	✓	✓	7	3
Leading VBS tag jet p_T	-	✓	-	11
Trailing VBS tag jet p_T	✓	✓	6	6
Pseudorapidity interval $\Delta\eta_{jj}^{VBS}$ between tag jets	✓	✓	4	4
Quark/gluon discriminator of leading VBS tag jet	✓	✓	9	7
Azimuthal angle distance between VBS tag jets	✓	-	10	-
Invariant mass of the VBS tag jets pair	✓	✓	1	1
p_T of the leading V_{had} jet	✓	-	14	-
p_T of the trailing V_{had} jet	✓	-	12	-
Pseudorapidity difference between V_{had} jets	✓	-	8	-
Quark/gluon discriminator of the leading V_{had} jet	✓	-	3	-
Quark/gluon discriminator of the trailing V_{had} jet	✓	-	5	-
p_T of the AK8 V_{had} jet candidate	-	✓	-	8
Invariant mass of V_{had}	✓	✓	11	5
Zeppenfeld variable for V_{had}	-	✓	-	9
Centrality	✓	✓	15	13

SHAP interpretation





VBF \rightarrow W leptonic

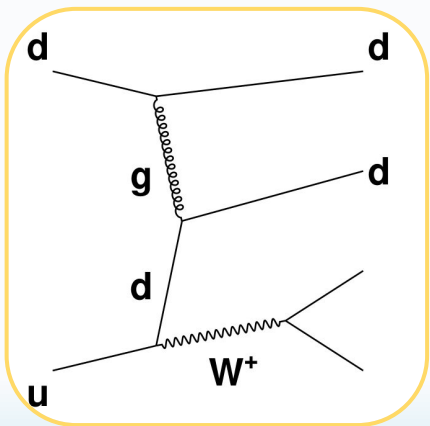


W+Jets:

Interazione forte tra quarks dello stato iniziale.

Stesso stato finale ma con una cinematica diversa.

Strategia: selezioni VBF.

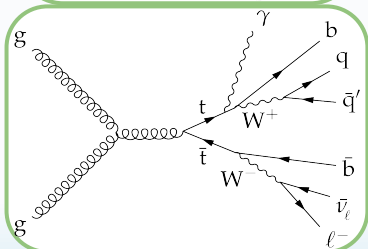
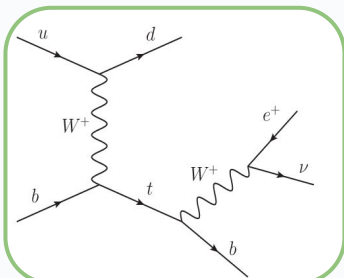


Single (double) top

Single: s & t channel

Double: un W che decade leptonicamente

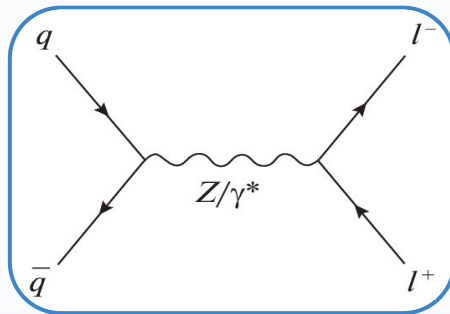
Strategia: b jets veto.



Drell-Yan:

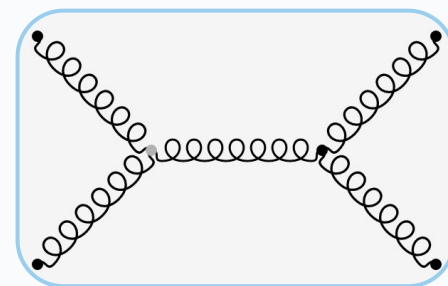
Un leptone non ricostruito (o fuori dall'accettazione).

Strategia: selezioni VBF



Nonprompt:

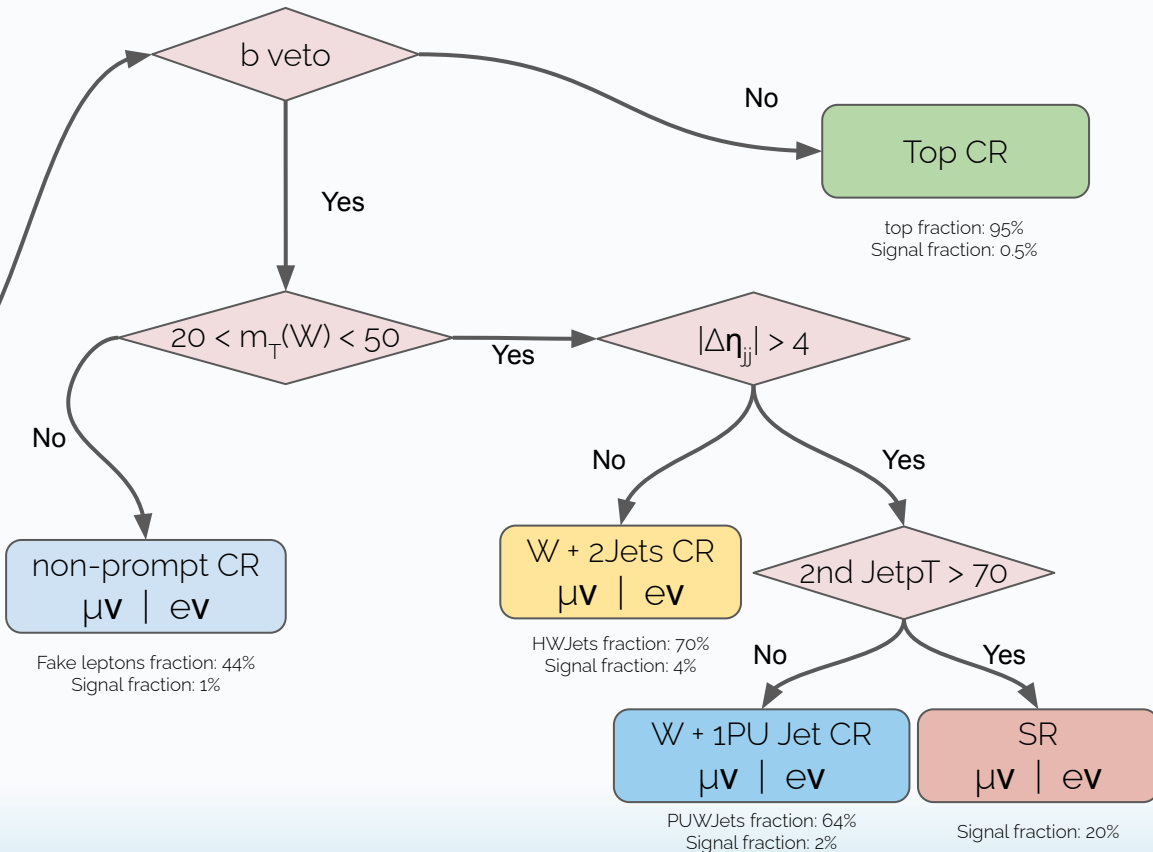
Per lo più multijet con 1 jet misidentificato come leptone (fake).





VBF \rightarrow W leptonico

- Preselezioni comuni:
- $p_T^{l1} > 25$ GeV
ma consistenti con le selezioni HLT
 - $|\eta_{l1}| < 2$
 - $p_T^{l2} < 15$ GeV (veto)
 - Almeno 2 jets con $p_T > 30$ GeV
and $|\eta_j| < 4.7$
 - $p_T^{j1} > 40$ GeV
 - $m_{jj} > 400$ GeV**
 - $|\Delta\eta_{jj}| > 2$**





VBF \rightarrow W leptonic

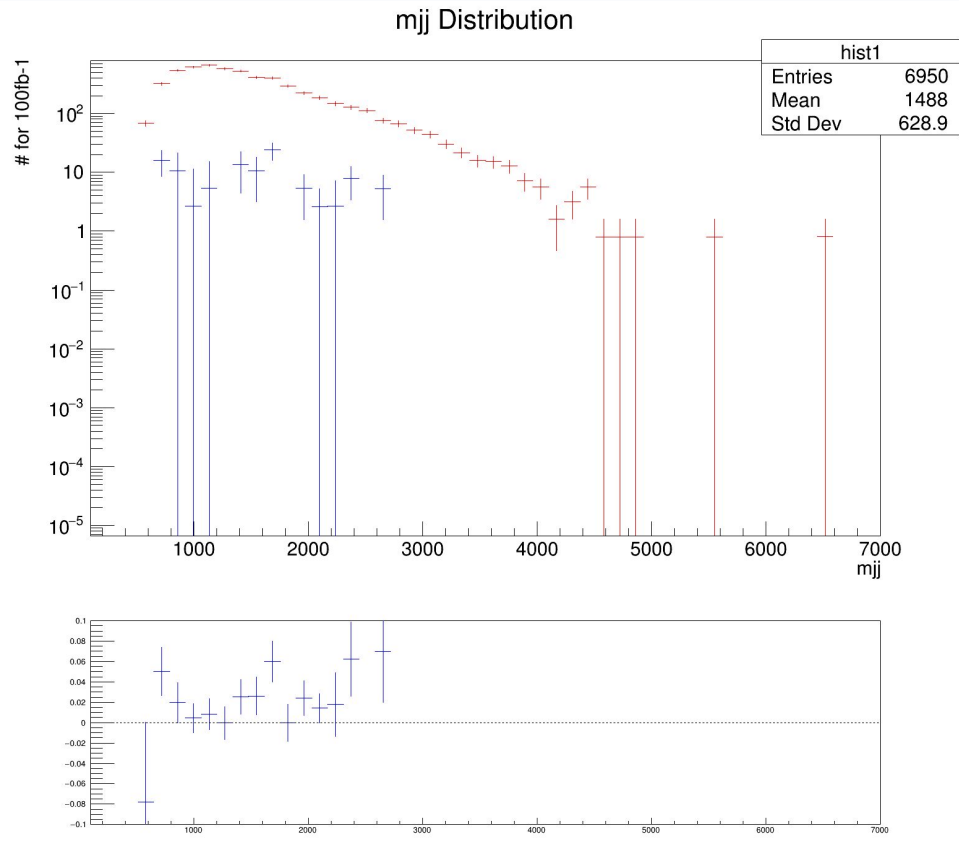


L'interferenza tra il segnale puramente EWK signal e il fondo mixed QCD è stata investigata.

In SR @LHE è presente un effetto del $\sim 4\%$. Attualmente stiamo svolgendo un'ulteriore indagine a livello @reco.

Il sample di interferenza è stato generato con MadGraph5_aMC@NLO con la seguente sintassi:

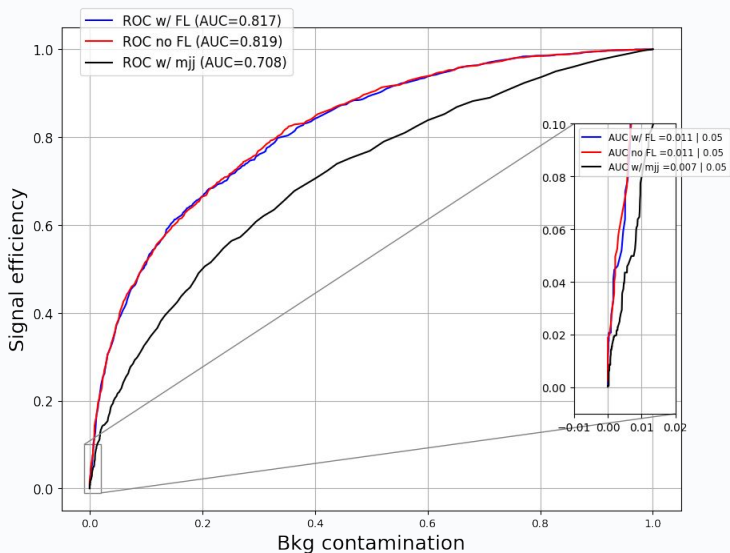
```
generate p p > l+ vl j j / t t~ h QED^2==6 QCD^2==2
add process p p > l- vl~ j j / t t~ h QED^2==6 QCD^2==2
```





VBF \rightarrow W leptonic

ROC comparison DNN vs m_{jj}



Normalized DNN output

