Finanziato dall'Unione europea NextGenerationEU
Ministero dell'Università e della Ricerca
Italiadomani PIANO NAZIONALE DI RIPRESA E RESILIENZA
ICSC Centro Nazionale di Ricerca in HPC, Big Data and Quantum Computing

Vector Boson Scattering with Machine Learning in Boosted Topologies at CMS

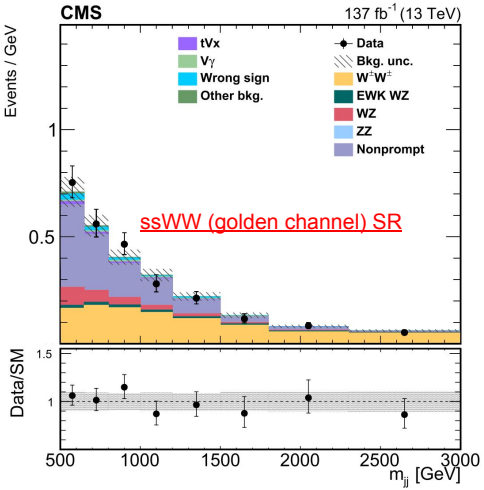
Luca Della Penna, Tommaso Tedeschi, David Butraigo Ceballos

unipg UNIVERSITÀ DEGLI STUDI DI PERUGIA
BOOST
INFN Istituto Nazionale di Fisica Nucleare

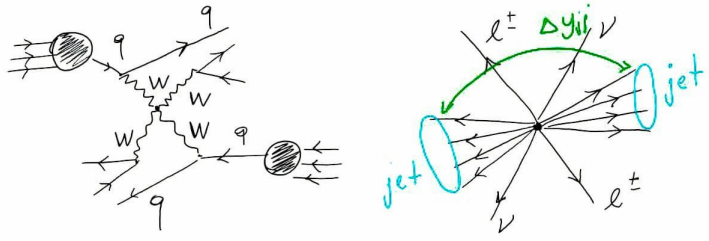
ICSC and Spoke2 - Where Are We Now?
Catania, 10-12 Dicembre 2024

Vector Boson Scattering

- scattering of W/Z gauge bosons, purely EW process of order $O(\alpha^6)$
  - Why?
    - Triple and quartic gauge couplings
    - Higgs and gauge bosons couplings (test of electroweak symmetry breaking)
    - Cancel the divergence of the cross-section at high energy
  - How?
    - Large rapidity gap between the two forward-backward jets in the detector
    - Large reconstructed dijet mass



Three types of final states:
-fully leptonic (electron and/or muons in the final states, clean but low statistics)
-semi-leptonic (one V in leptons and one V in quarks, balance)
-fully-hadronic (both V decay into a quark pair, high statistics and high background)

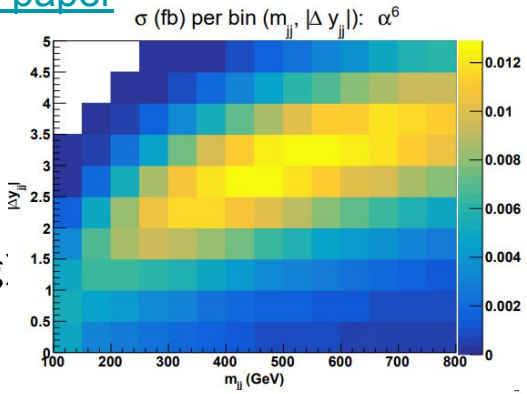


Vector Boson Scattering

- scattering of W/Z gauge bosons, purely EW process of order $O(\alpha^6)$
  - Why?
    - Triple and quartic gauge couplings
    - Higgs and gauge bosons couplings (test of electroweak symmetry breaking)
    - Cancel the divergence of the cross-section at high energy
  - How?
    - Large rapidity gap between the two forward-backward jets in the detector
    - Large reconstructed dijet mass

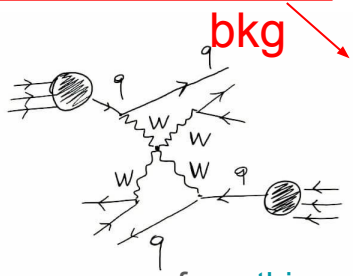
Three types of final states:
-fully leptonic (electron and/or muons in the final states, clean but low statistics)
-semi-leptonic (one V in leptons and one V in quarks, balance)
-fully-hadronic (both V decay into a quark pair, high statistics and high background)

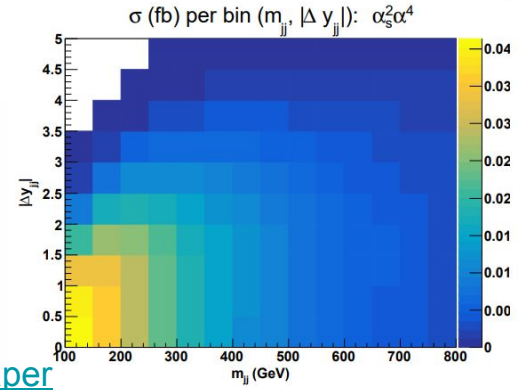
from this paper



signal

bkg

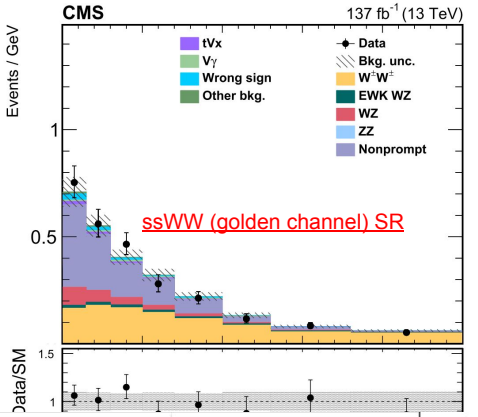


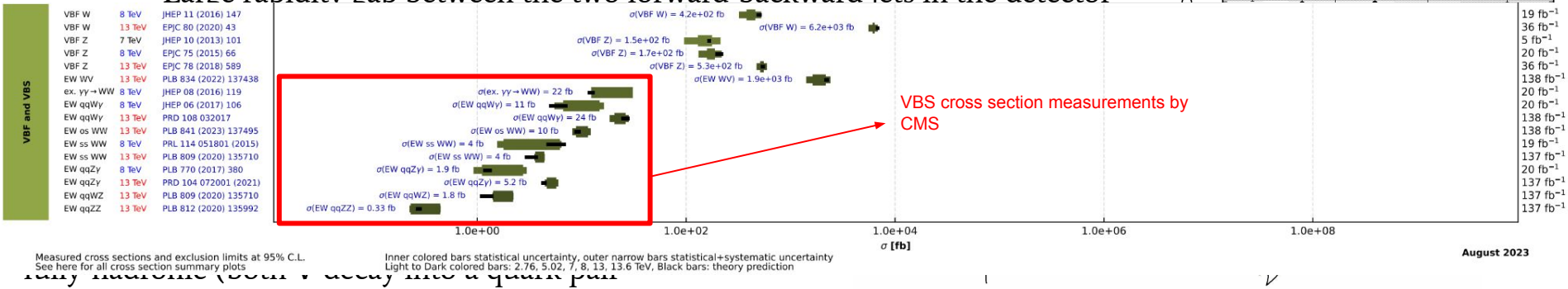


from this paper

Vector Boson Scattering

- scattering of W/Z gauge bosons, purely EW process of order $O(\alpha^6)$
  - Why?
    - Triple and quartic gauge couplings
    - Higgs and gauge bosons couplings (test of electroweak symmetry breaking)
    - Cancel the divergence on the cross-section at high energy
  - How?
    - Large rapidity gap between the two forward-backward jets in the detector



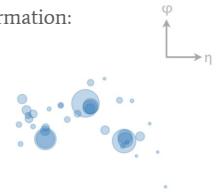


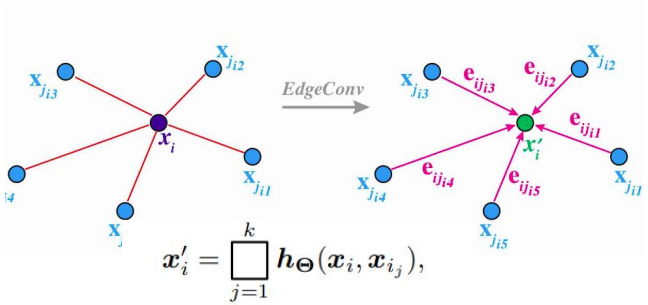
VBS cross section measurements by CMS

What we have: ParticleNet

- ParticleNet is a DGNN machine learning techniques used in CMS to extract informations from particles features.
- Inputs: unordered set of particles in space (point cloud) with their informations (rapidity and azimuth angle, momentum, charge...) and it is based on the EdgeConv operation to find correlation among the points on a graph.
- Outputs: classification score to tag X->qq, X->bb, X->cc jets against QCD jets.

Primary information:
2D (η, φ)

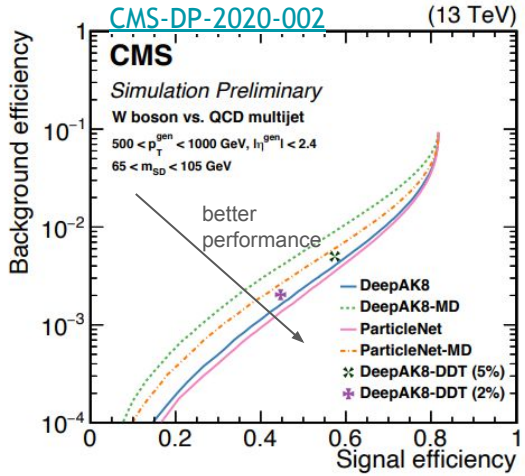




ParticleNet outputs are available in the "standard" set of variables in CMS (NanoAOD format)

ParticleNet variables for R=0.8 jets

FatJet_particleNet_XbbVsQCD	Float_t
FatJet_particleNet_XccVsQCD	Float_t
FatJet_particleNet_XggVsQCD	Float_t
FatJet_particleNet_XqqVsQCD	Float_t



What we have: ParticleNet

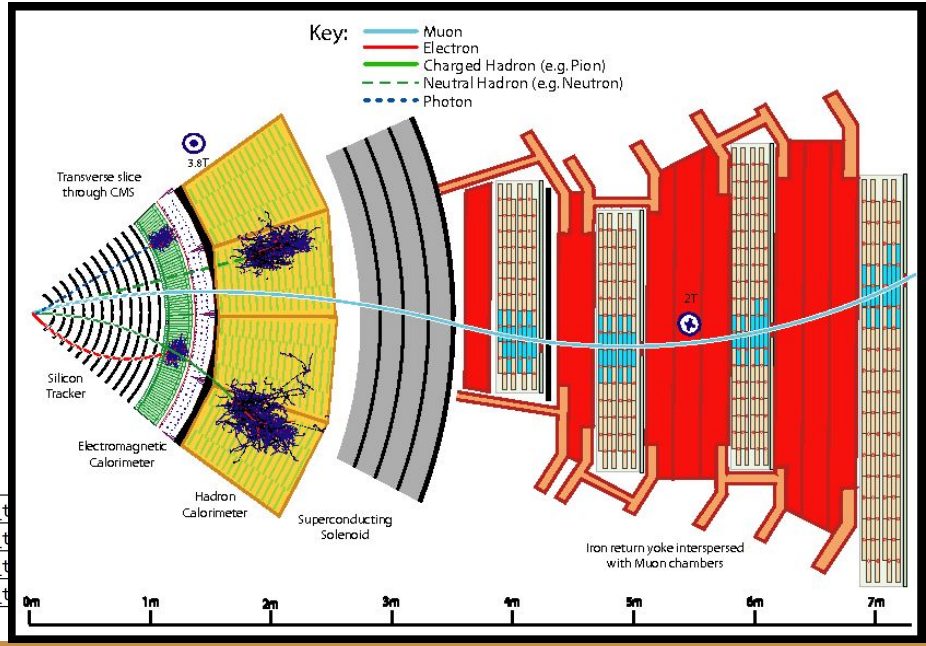
- ParticleNet is a DGNN machine learning techniques used in CMS to extract informations from particles features.
- Inputs: unordered set of particles in space (point cloud) with their informations (rapidity and azimuth angle, momentum, charge...) and it is based on the EdgeConv operation to find correlation among the points on a graph.
- Outputs: classification score to tag X->qq, X->bb, X->cc jets against QCD jets.

ParticleNet outputs are available in the "standard" set of variables in CMS (NanoAOD format)

ParticleNet variables for R=0.8 jets

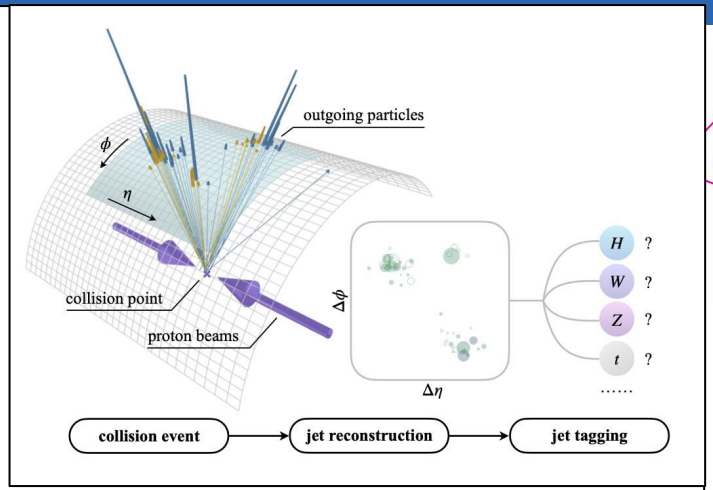
FatJet_particleNet_XbbVsQCD	Float_t
FatJet_particleNet_XccVsQCD	Float_t
FatJet_particleNet_XggVsQCD	Float_t
FatJet_particleNet_XqqVsQCD	Float_t

At the CMS experiment:
Particle informations = ParticleFlow
(putting together informations from the different subdetectors)



A new tool: ParT

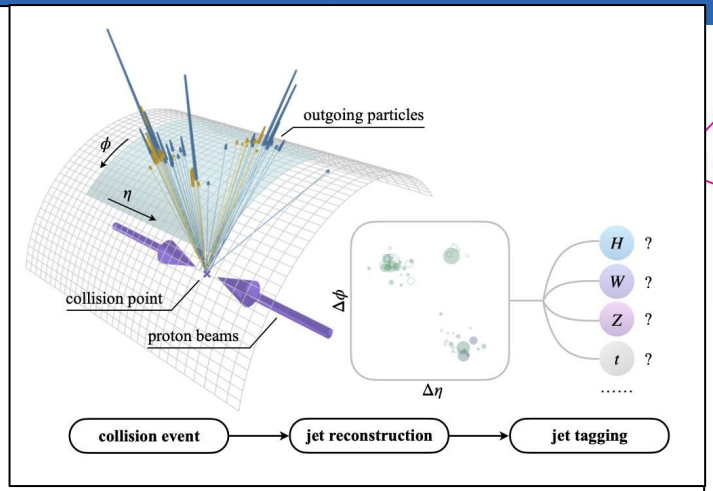
CMS is going to introduce new ML variables for jets using ParT
-It shows better performance, particularly in background rejection
-Inputs are again particle features but now introducing "interaction" (C') among them (N, C, C')
-Trained on the JETCLASS dataset that contains 10 "types" of jets (100M jet event in total)

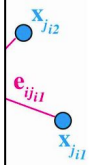


Finanziato dall'Unione europea NextGenerationEU
Ministero dell'Università e della Ricerca
Italiadomani PIANO NAZIONALE DI RIPRESA E RESILIENZA
ICSC Centro Nazionale di Ricerca in HPC, Big Data and Quantum Computing

Part A new tool: ParT

- CMS is going to introduce new ML variables for jets using ParT
- -It shows better performance, particularly in background rejection
- -Inputs are again particle features but now introducing "interaction" (C') among them (N, C, C')
- -Trained on the JETCLASS dataset that contains 10 "types" of jets (100M jet event in total)



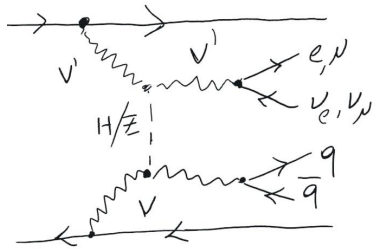


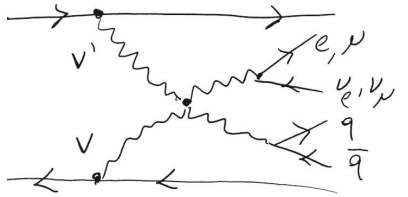
background rejection

	All classes		$H \to b\bar{b}$	$H \to c\bar{c}$	$H \to gg$	$H \to 4q$	$H \to \ell\nu qq'$	$t \to bqq'$	$t \to b\ell\nu$	$W \to qq'$	$Z \to q\bar{q}$
	Accuracy	AUC	$\text{Rej}_{50\%}$	$\text{Rej}_{50\%}$	$\text{Rej}_{50\%}$	$\text{Rej}_{50\%}$	$\text{Rej}_{99\%}$	$\text{Rej}_{50\%}$	$\text{Rej}_{99.5\%}$	$\text{Rej}_{50\%}$	$\text{Rej}_{50\%}$
PFN	0.772	0.9714	2924	841	75	198	265	797	721	189	159
P-CNN	0.809	0.9789	4890	1276	88	474	947	2907	2304	241	204
ParticleNet	0.844	0.9849	7634	2475	104	954	3339	10526	11173	347	283
ParT	**0.861**	**0.9877**	**10638**	**4149**	**123**	**1864**	**5479**	**32787**	**15873**	**543**	**402**

What we want to do?

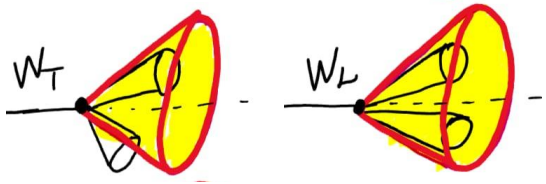
- Polarized cross-section in VBS processes in semi-leptonic and fully-hadronic final state...using ML techniques
- We need to extract informations from boosted W/Z jets, retrieving properties of the subjects to reconstruct the polarization of the initial bosons.
- Polarization accessible through the angular distribution of the final state particles. Different variables have been proposed for this task that take the momentum unbalance of the final state in the case of transverse and longitudinal polarization (p_Θ, z_j , z_g and others)
- Mass-decorrelated tagger using ParticleNet are already available in CMS





Process	$\sigma\,\mathcal{B}$ (fb)	Theoretical prediction (fb)
$W_L^\pm W_L^\pm$	$0.24^{+0.40}_{-0.37}$	0.28 ± 0.03
$W_X^\pm W_T^\pm$	$3.25^{+0.50}_{-0.48}$	3.32 ± 0.37
$W_L^\pm W_X^\pm$	$1.40^{+0.60}_{-0.57}$	1.71 ± 0.19
$W_T^\pm W_T^\pm$	$2.03^{+0.51}_{-0.50}$	1.89 ± 0.21

polarized cross section for ssWW in fully leptonic final state measured by CMS (this paper)



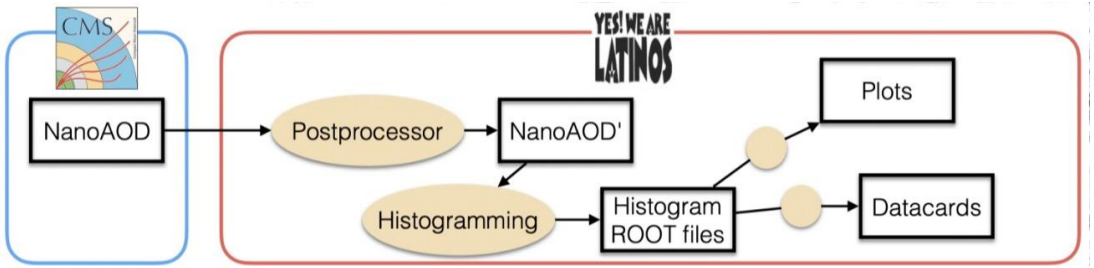
$$\mathcal{M}_\pm \propto \frac{1 \mp \cos\theta_*}{2},$$

$$\mathcal{M}_0 \propto -\frac{\sin\theta_*}{\sqrt{2}}.$$

How?

mkShapesRDF (Latino) + ParticleNet/ParT variables

- mkShapesRDF is a software for post-processing and shape analysis in CMS
  - Post-Processing: selection and filtering variables to produce a modified "standard" format (NanoAOD')
  - We included new modules on the framework to filtering FatJets (jets with R=0.8) + corrections on the jet variables (mass, transverse momentum ...) due to detector effects
- All these can be used on the CMS INFN Analysis Facility (see Tommaso's talk of yesterday)



Person Power (INFN and University of Perugia):

- David Butraigo Ceballos (1st year PhD student)
- Luca Della Penna (2nd year PhD student)
- Tommaso Tedeschi (Post-Doc)

Thanks for your attention!