Measurements of the Higgs boson decaying into heavy flavored quarks at the CMS Experiment

Alejandro Gomez Espinosa

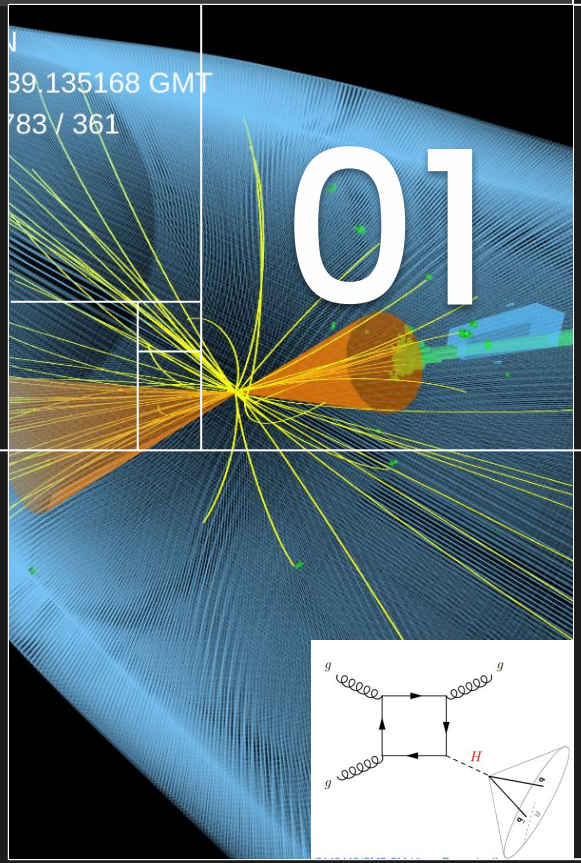
On behalf of the CMS Collaboration

SILAFAE 2022

ETH Zurich

Outlook

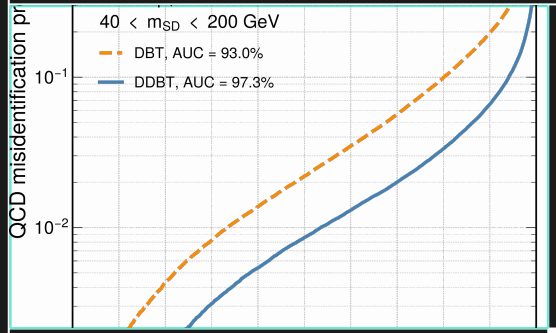
01	**Boosted ggHbb** High pt Higgs boson measurement with bottom pairs in final state	02	**Boosted ggHcc** High pt Higgs boson measurement with charm pairs in final state
03	**Inclusive VHcc** Boosted and resolved measurement of Higgs boson produced with W/Z decaying into charm pairs	04	**Inclusive VHbb** Boosted and resolved measurement of Higgs boson produced with W/Z decaying into bottom pairs



Boosted ggHbb

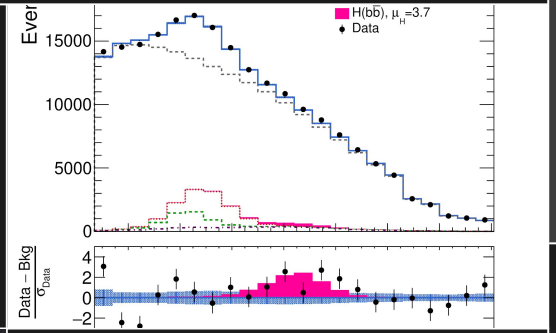
Inclusive search for highly boosted Higgs bosons decaying to a bottom quark-antiquark pair
JHEP 12 (2020) 085 (CMS-HIG-19-003)

Measurement using 137 fb^{-1} of data collected by the CMS Collaboration at 13 TeV. Using state-of-the-art object reconstruction and background estimation techniques.



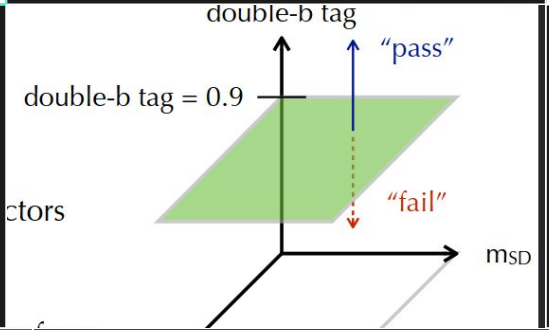
Bkg Estimation

Rhalphabet method profits of regions with higher stats to model bkg.



Improved tagger

A machine-learning based tagger improved the sensitivity of the search

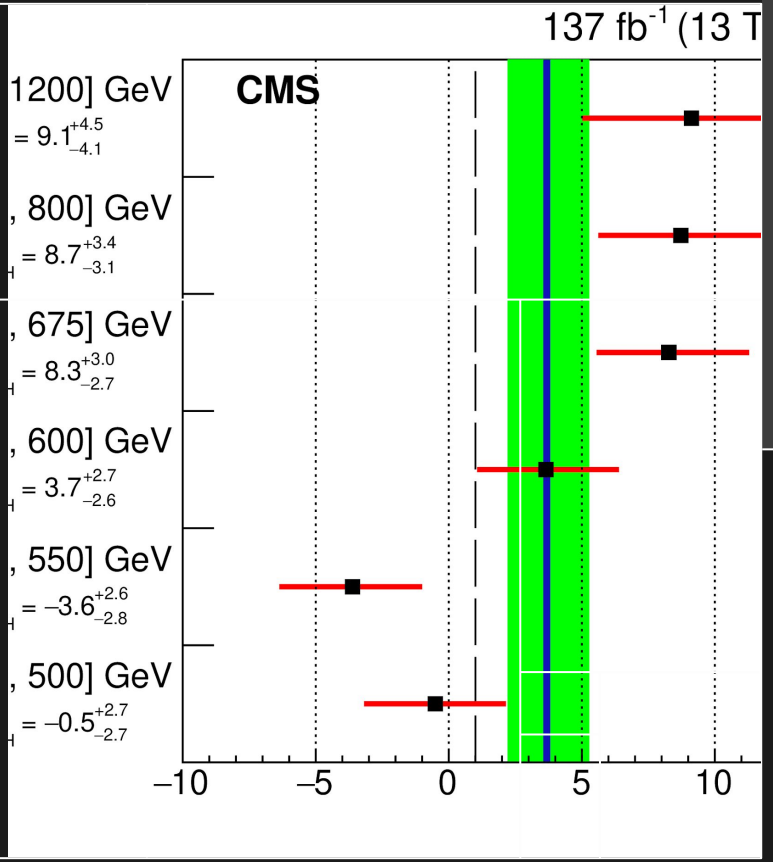


Jet Mass

Improved jet mass modelling with PUPPI+Softdrop

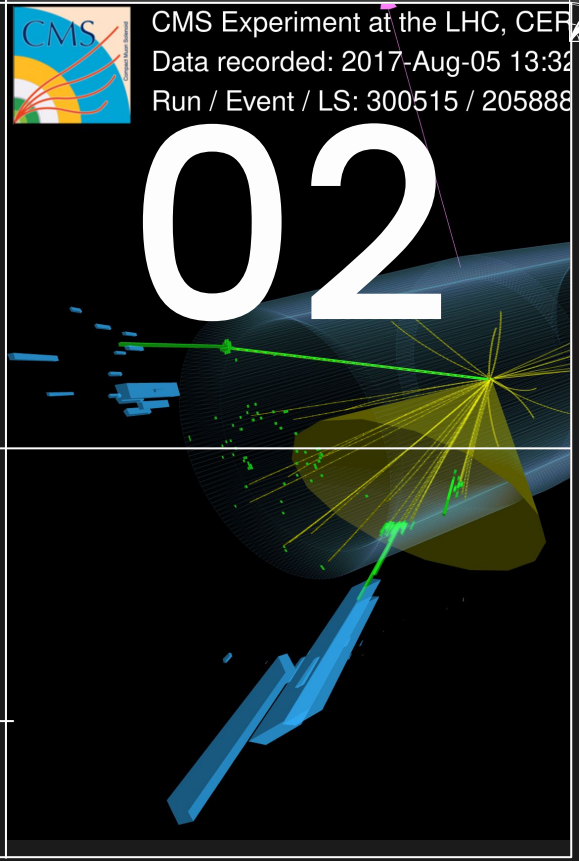
The observed significance of the ggHbb process is 2.5 standard deviations!

- The inclusive signal strength (μ) is measured to be **3.7 +1.6/-1.5**.
- It represents an improve of a factor of two with respect to previous results.
- Largest source of uncertainty comes from the modelling of the QCD background (stats) and V+jets (theory).



Boosted ggHcc

02

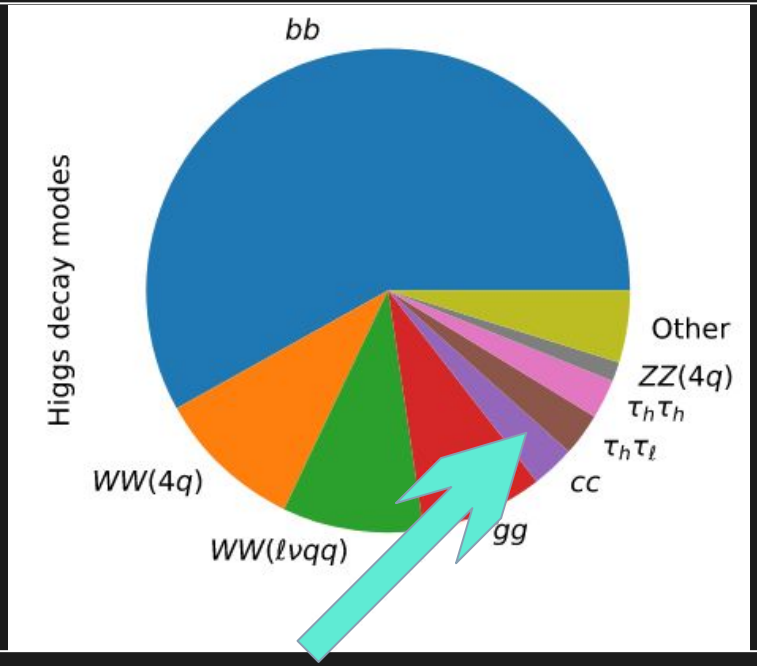


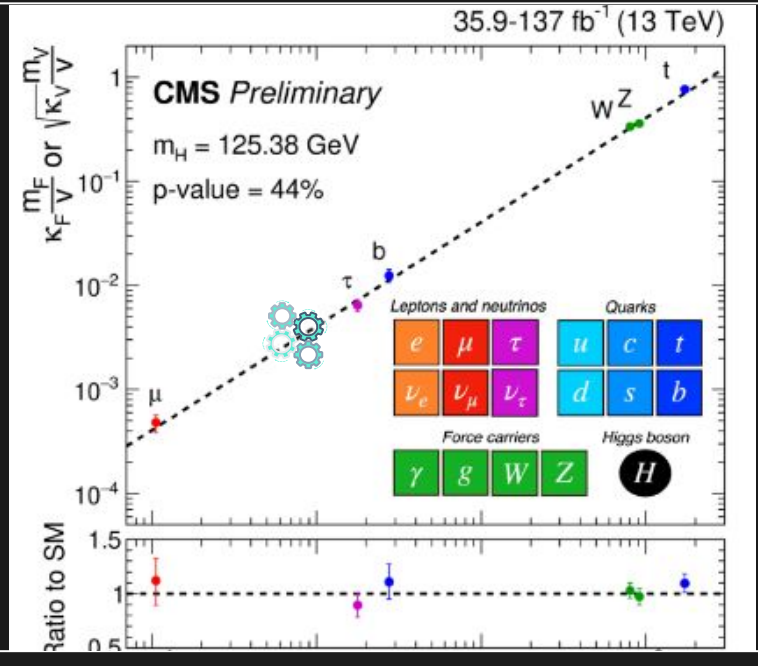
Inclusive search for a boosted Higgs boson and an observation of the Z boson decaying to charm quark pair (CMS-PAS-HIG-21-012)

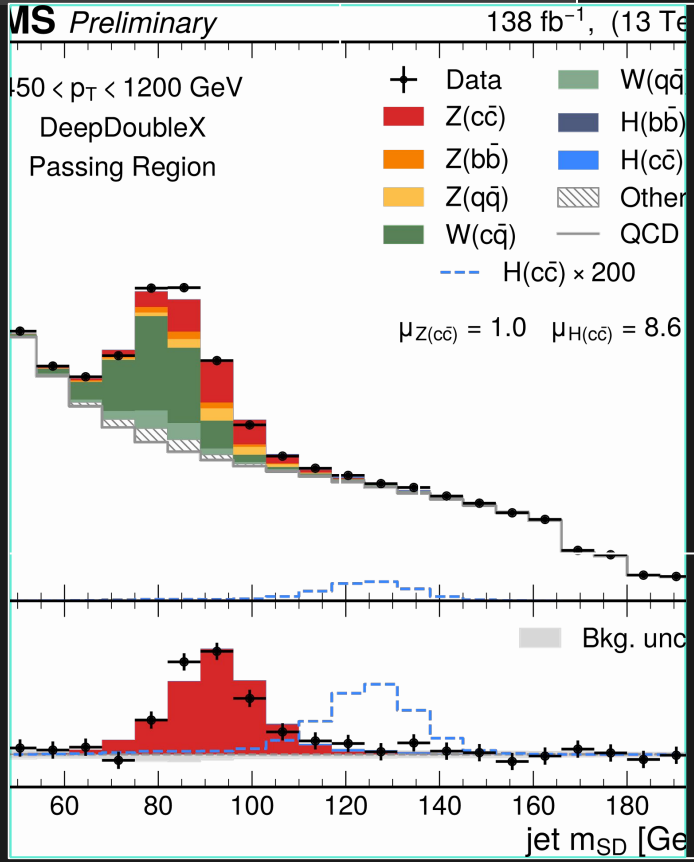
SILAFAE 2022

ETH Zurich - A.G.E.

Why H->cc is important?

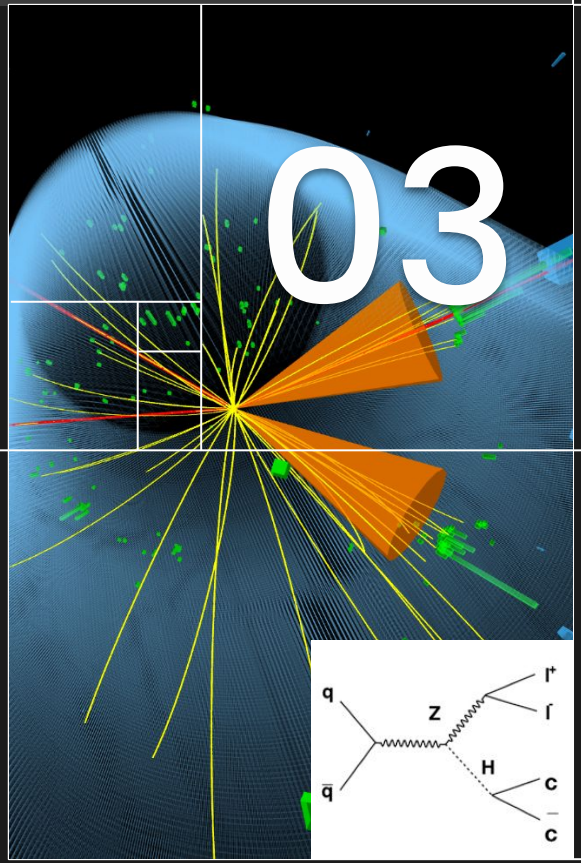






- Similar approach than ggHbb.
- **ML-based** Hcc tagger improves the sensitivity of the search.
- Rhalphabet data-driven method provides good QCD estimation.
- Residual QCD simulation is the highest source of uncertainty.
- Zcc process measured **well over 5 σ**.
- Hcc observed signal strength is 45 the SM prediction.

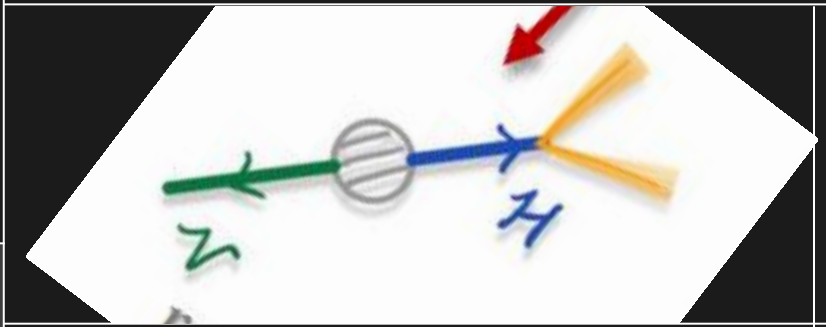
It is the first time that the Zcc process in Z+jets is observed in hadron colliders.



03 Inclusive VHcc

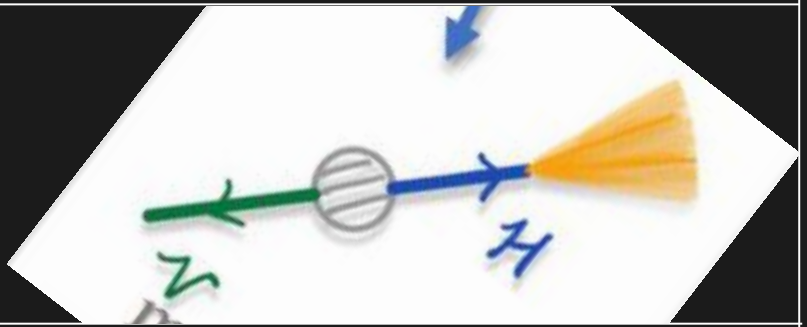
Search for Higgs boson decay to a charm quark-antiquark pair (CMS-HIG-21-008)

Inclusive search



Resolved analysis

Exploit the **low transverse momentum** of the Higgs decays. Reconstruction with **smaller** cone-size jets (AK4).



Merged analysis

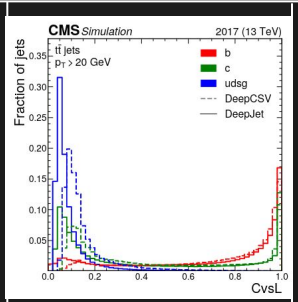
Exploit the **high transverse momentum** of the higgs decays. Reconstruction with **larger** cone-size jets (AK15).

Both analyses are further divided into 3 channels according to the V decay: 0, 1, 2 leptons.

Resolved VHcc

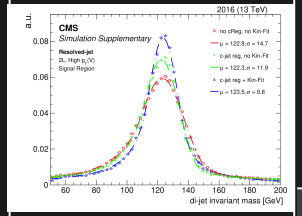
Charm tagging

Multiclassifier DNN (CNN+RNN) separating c-, b-, udsg-jets



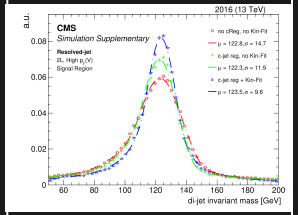
Kinematic fit

In 2 lepton channel, exploit kinematic constraints of collision



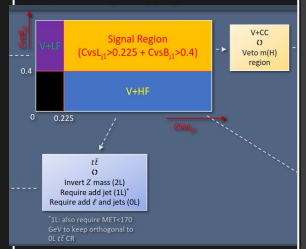
C-jet energy regression

DNN regression ro recover energy lost from neutrinos



Bkg estimation

Dedicated control regions by inverting c-tagging selection



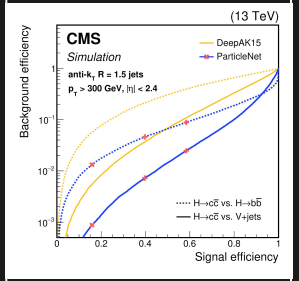
SILAFAE 2022

ETH Zurich - A.G.E.

Merged VHcc

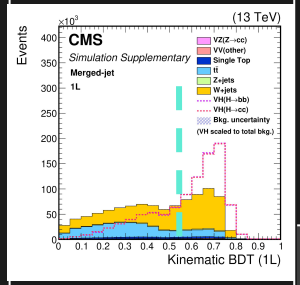
cc tagging

Graph neural network (GNN) to exploit substructure and flavor



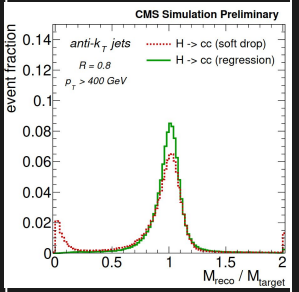
Kinematic BDT

Separate VH and bkg events. Output used to define signal and control regions



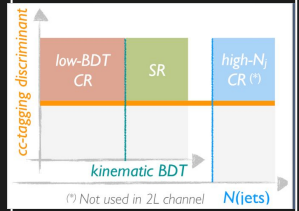
Mass regression

Similar GNN architecture, reduce mass dependency



Bkg estimation

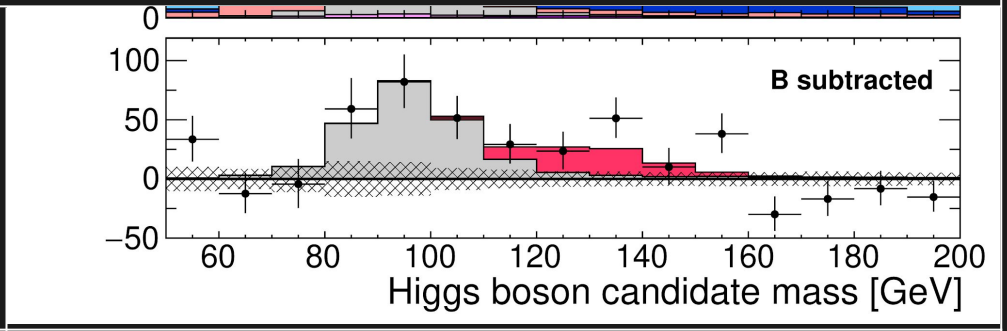
Dedicated control regions by inverting c-tagging selection



Most stringent constraint to date.

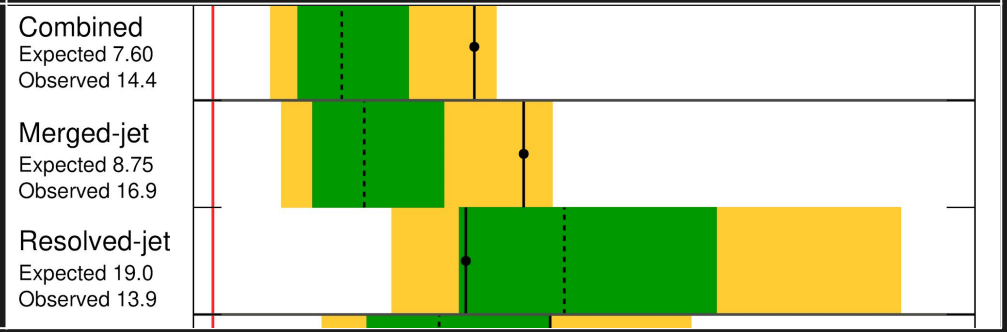
Z/Higgs mass

Zcc process in VZ events reach a **5.7 significance.**



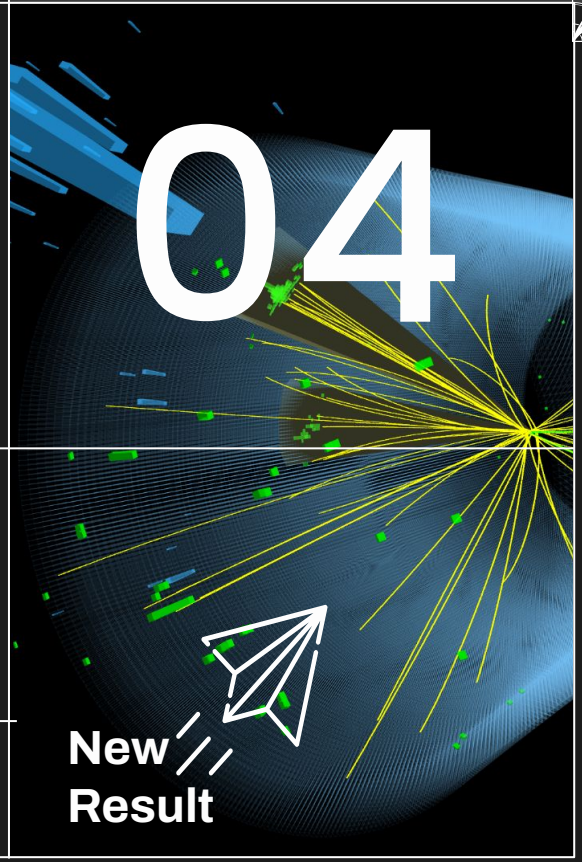
VHcc production

Observed limit: 0.94 pb corresponding to **14 times the SM prediction.**



Inclusive VHcc

04

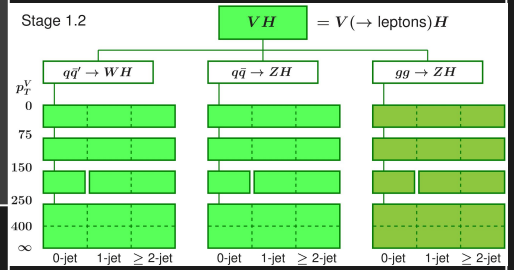


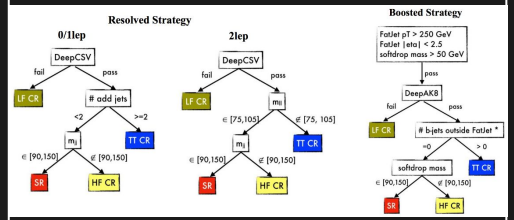
Simplified template cross section measurements of Higgs boson produced in association with vector bosons in the Hbb decay channel (CMS-PAS-HIG-20-001)

SILAFAE 2022

ETH Zurich - A.G.E.

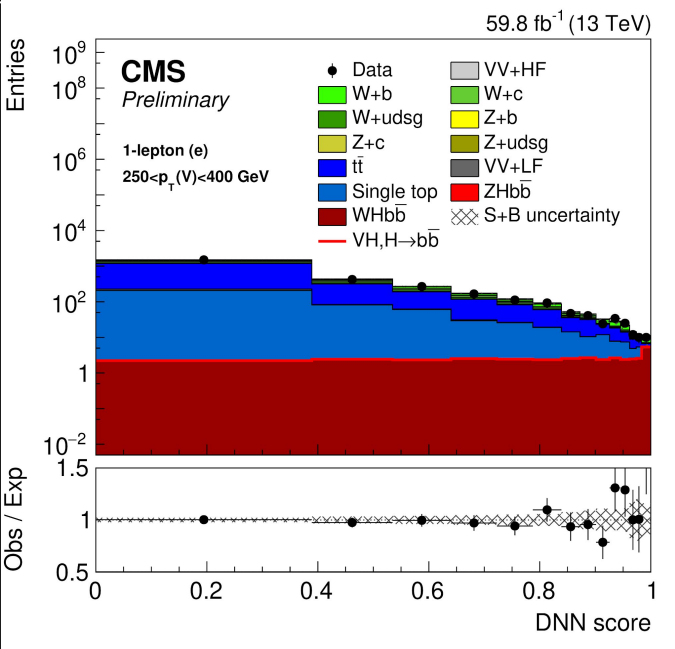
STXS measurement

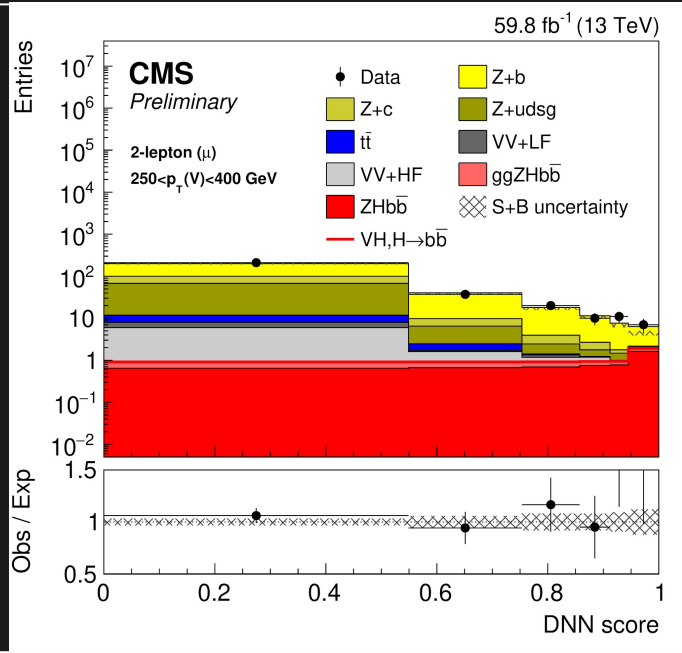




- Simplified Template Cross Section (**STXS**) analysis as a first step into a differential measurement.
- Analysis performed for **low and high** transverse momentum Higgs in association with a V boson.
- Selection targets 3 channels based on V decay: 0, 1, 2 leptons. Total of 30 signal and 51 control regions
- **DNN** and **BDT** techniques used to maximize separation of signal events and bkgs.
- Strategy: simultaneous max likelihood fit of signal and control regions in all regions.
- Analysis include a dedicated dijet invariant mass (Higgs mass) characterisation.

Selected distributions

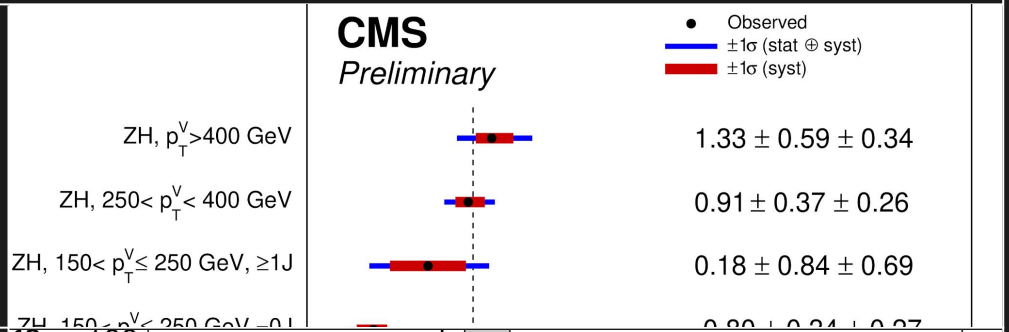




Observed significance of 3.3 σ

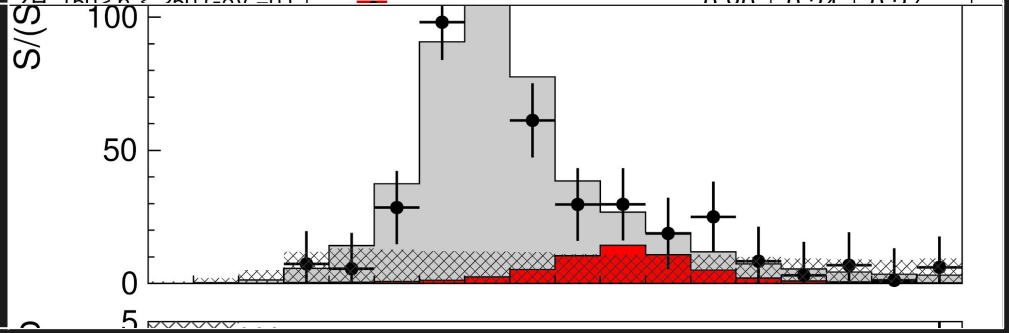
STXS signal strengths

Measured using 2016-2018 data



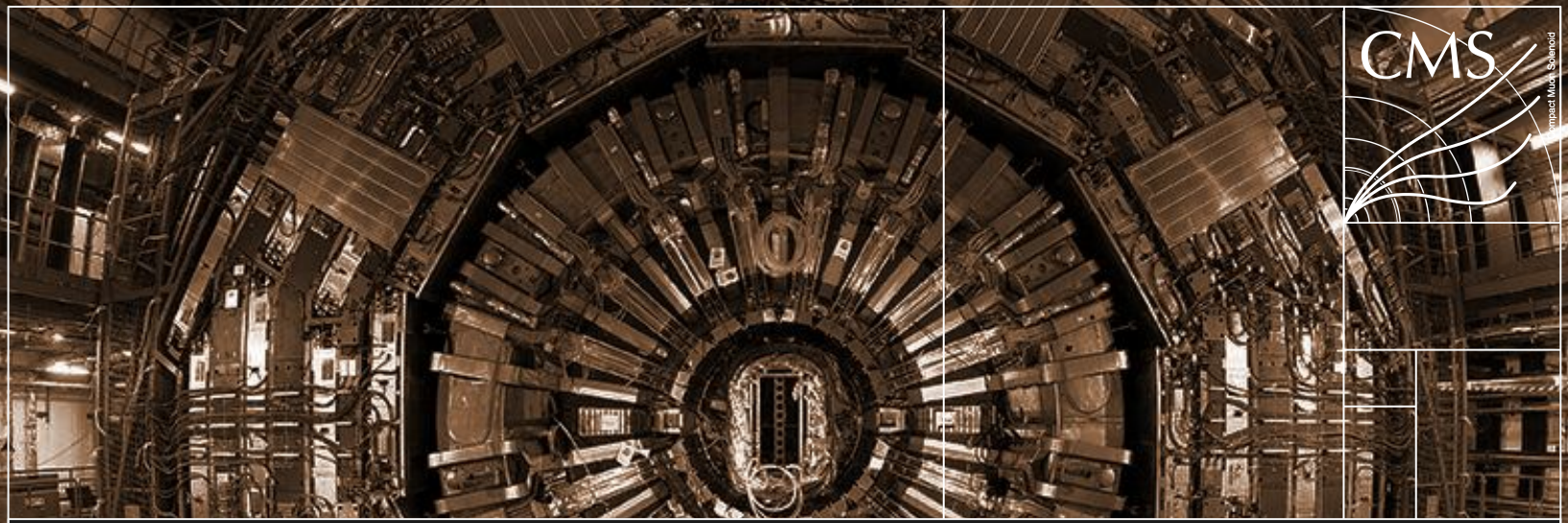
Mass distribution

Dijet invariant mass combining all channels and data.



Conclusions

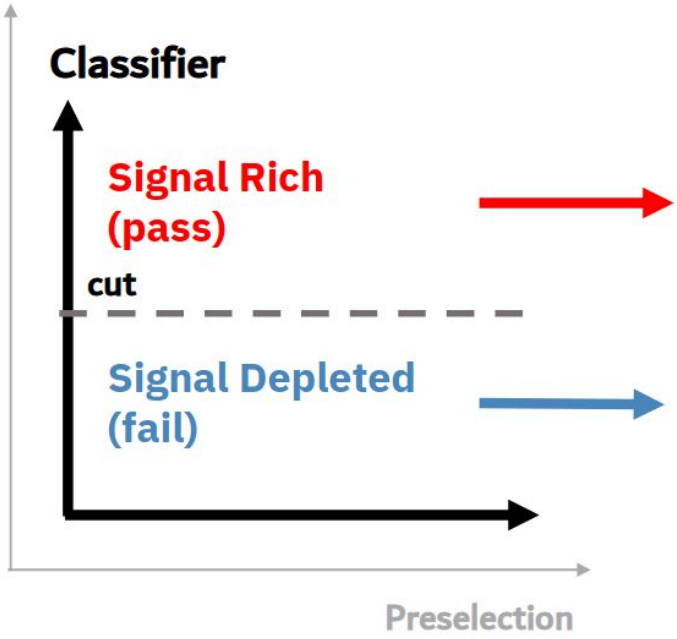
- **CMS Collaboration keeps pushing the boundaries** in object reconstruction and analysis techniques to study rare processes, like Higgs decaying into charm quarks.
- In this talk, an overview of the state-of-the-art analyses targeting final states with **Higgs decaying into heavy flavor quarks** is presented.
- In ggHcc, we provided the **first measurement of Zcc** in Z+jets events at hadron colliders.
- In VHcc, we reach the **most stringent constraint** to date in the production of VHcc.
- In VHbb, we start to provide a **more detailed characterization** of the VHbb process in several transverse momentum regions.



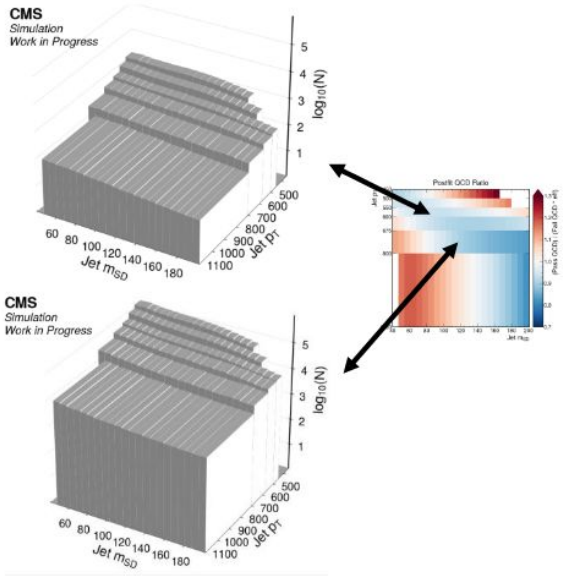
Stay tuned to more CMS results.

Rhalphabet method

Signal Region Definition



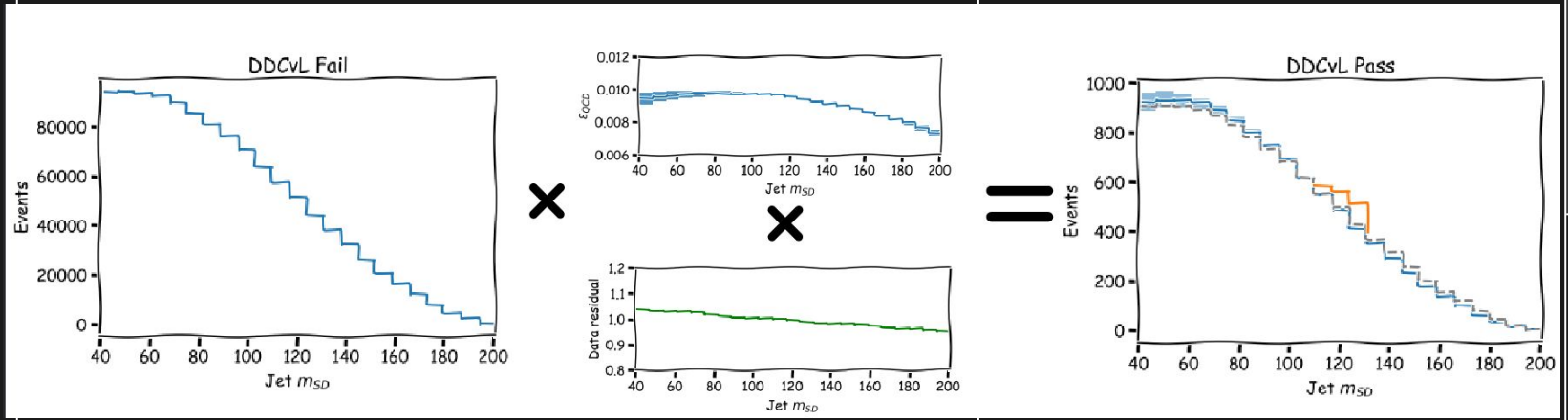
Simultaneous Fit



Constrain QCD **from data *in-situ***

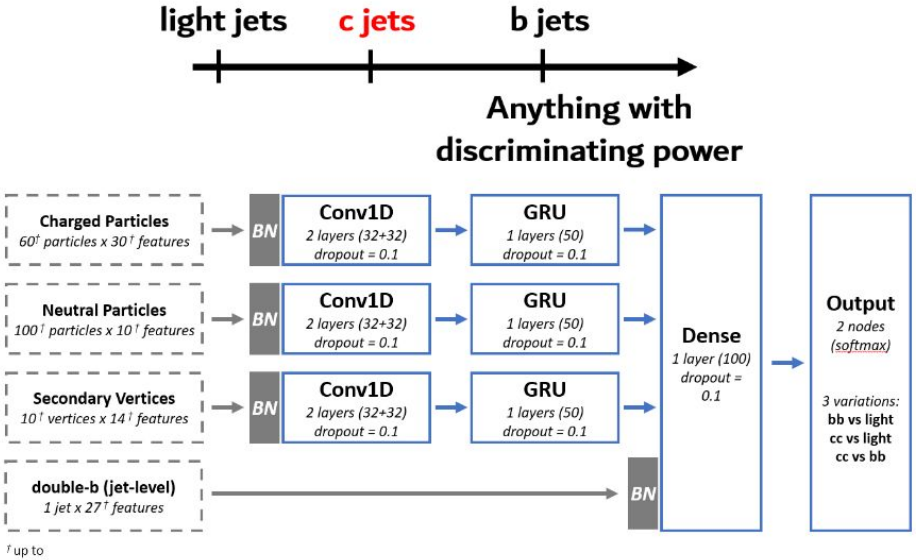
Fit jet soft-drop mass in bins of p_T

Rhalphabet method



DeepDoubleX – jet identification

- bb jet identification 2x better than last gen
- cc jet identification meaningfully performant



DeepJet – jet identification

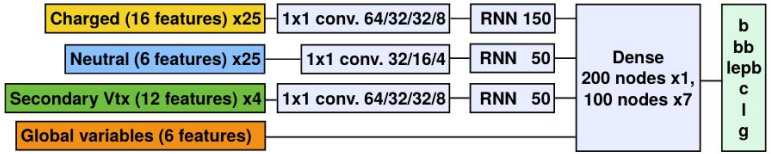


Figure 1. An illustration of the DeepJet architecture. Three seperate branches are used to process charged candidates, neutral candidates and secondary vertices. The algorithm makes use of 1x1 convolutional layers to perform automatic feature engineering for each class of jet constituents. The three RNN (LSTM) layers combine the information for each sequence of constituents. Finally the full jet information is combined using fully connected layers.

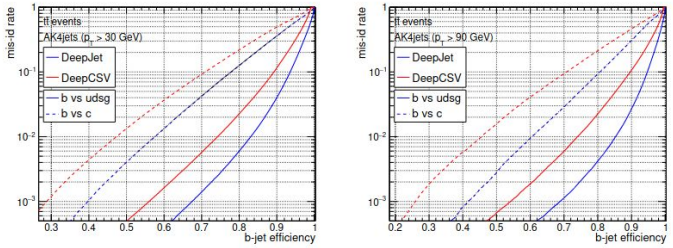
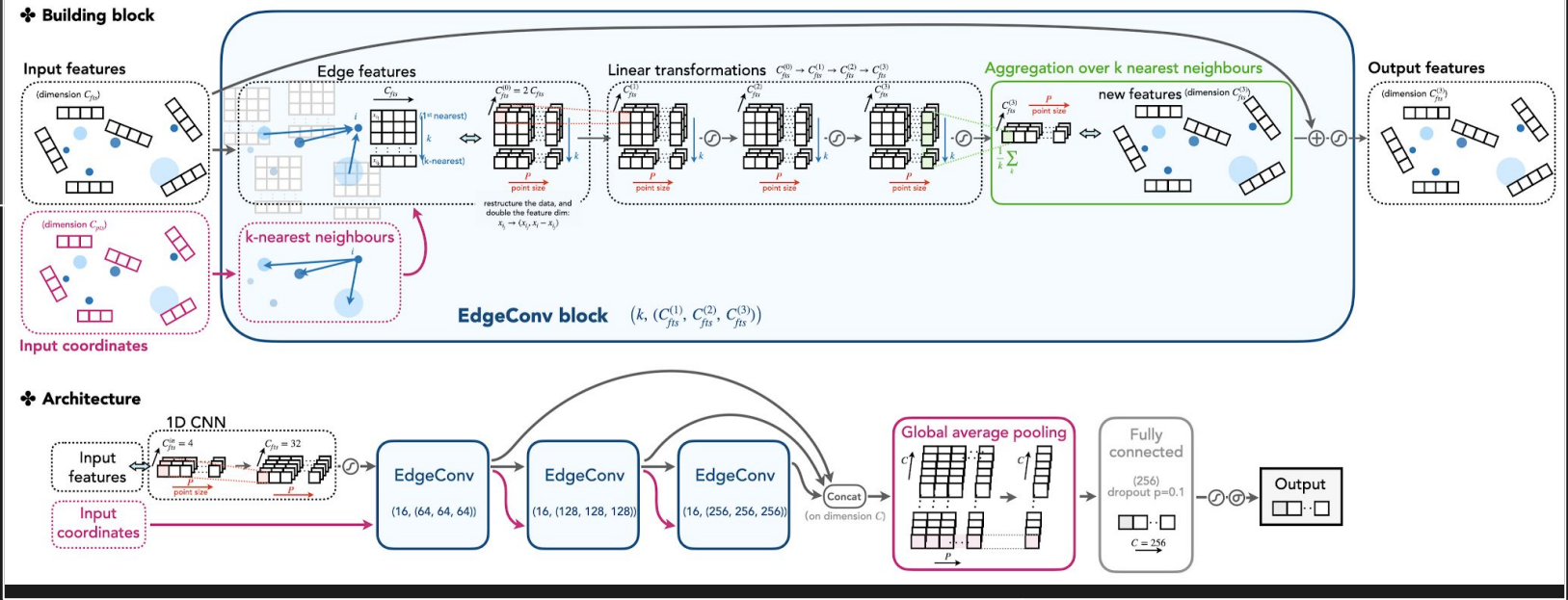


Figure 2. Performance of the DeepJet and DeepCSV b-tagging algorithms on $t\bar{t}$ events with both top quarks decaying hadronically. The jets are required to have $p_T > 30$ GeV (left) and $p_T > 90$ GeV (right). The performance is shown for both b vs. c classification (dashed lines), and b vs. light (solid lines).

ParticleNet – jet identification



SILAFAE 2022

ETH Zurich - A.G.E.