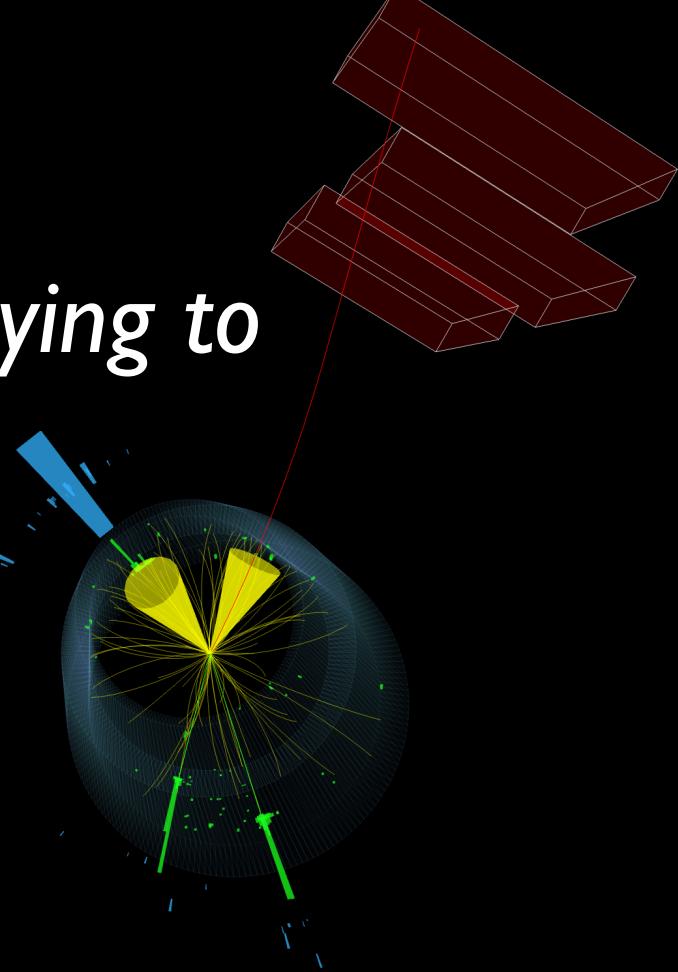


# *Search for Higgs boson decaying to charm quarks with CMS*

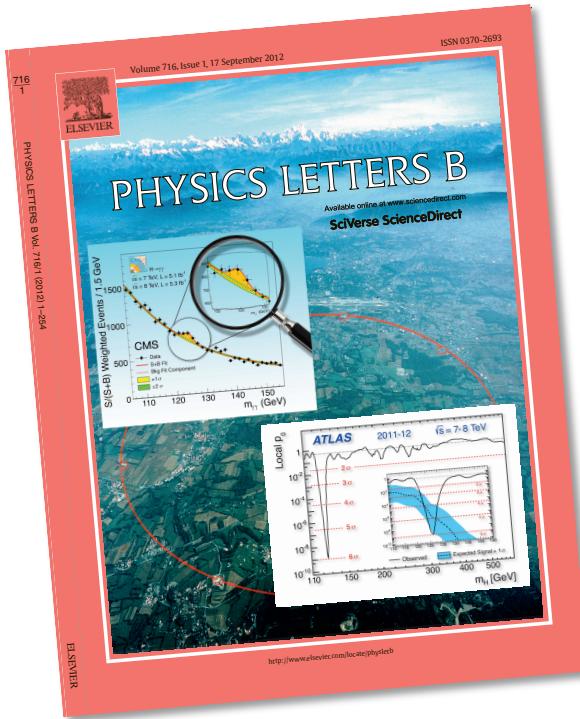
Huilin Qu (CERN)

*Experimental Particle and Astro-Particle Physics Seminar  
University of Zurich  
7 March 2022*



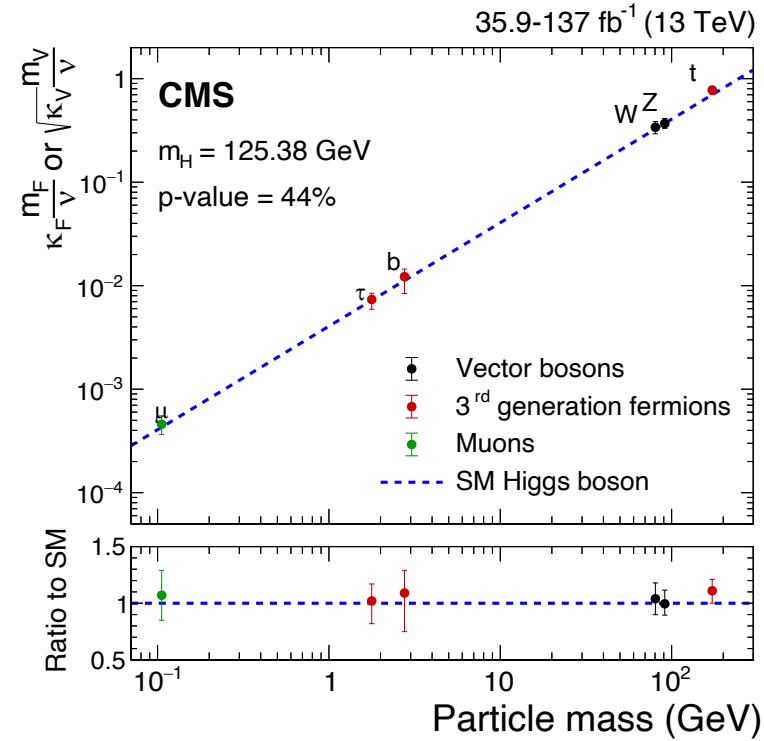
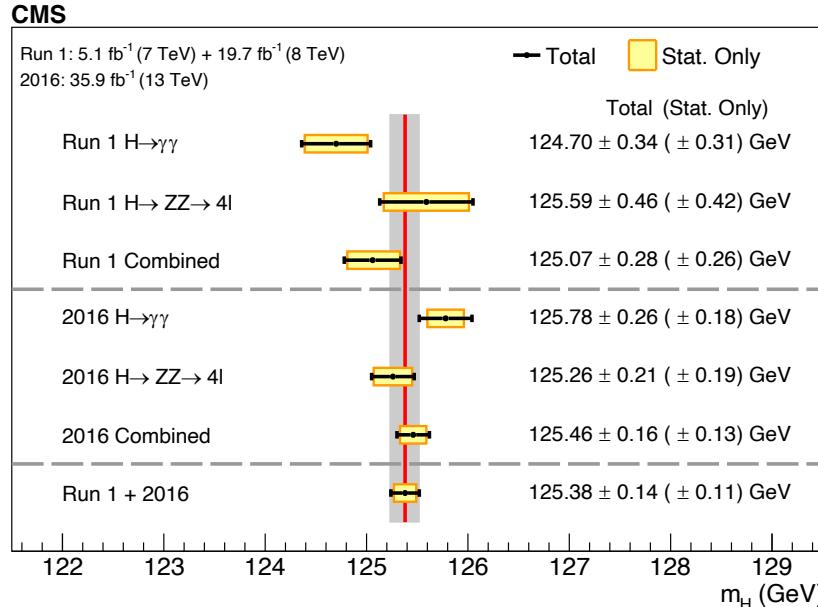
# Introduction

- Discovery of the Higgs boson in 2012: A new chapter of particle physics



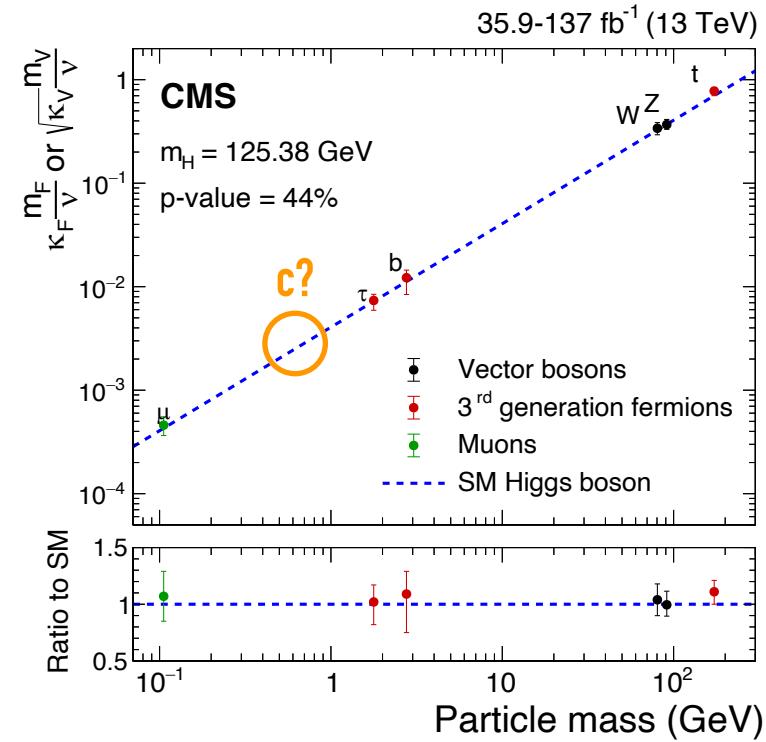
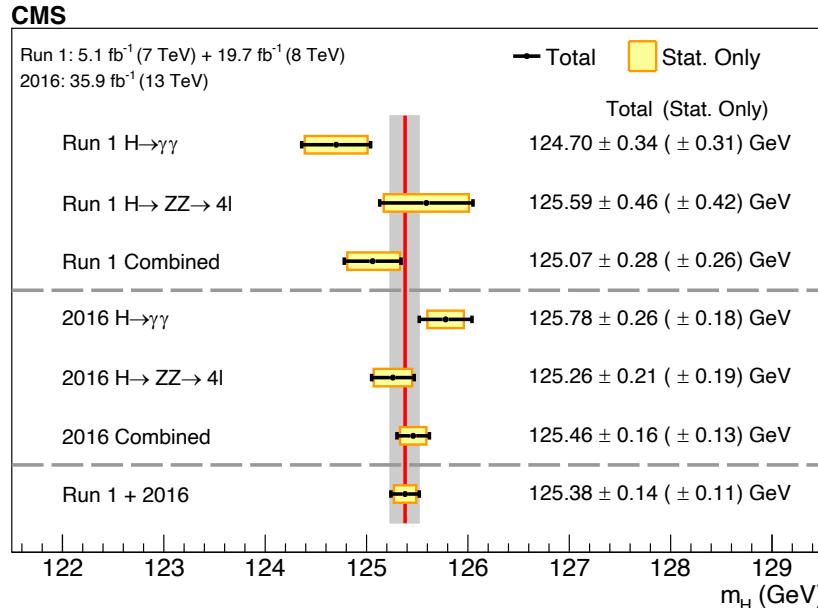
# Understanding the Higgs boson

- Tremendous progress in our understanding of the Higgs boson in the past ten years



# How charming is the Higgs boson?

- Tremendous progress in our understanding of the Higgs boson in the past ten years

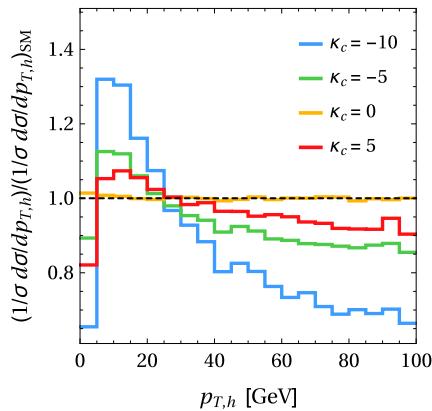


# Probing the Higgs-charm coupling

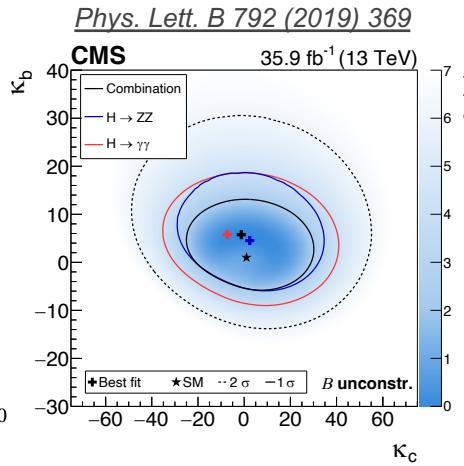
- Several methods explored by CMS to probe the Higgs-charm Yukawa coupling ( $y_c$ )

## Indirect constraint from Higgs kinematics

*Phys. Rev. Lett.* 118, 121801

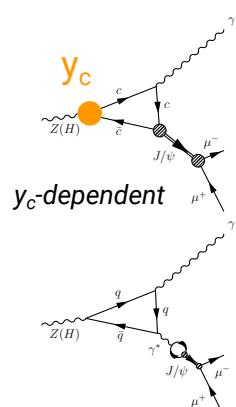


*Phys. Lett. B* 792 (2019) 369



## Search for exclusive $H \rightarrow J/\Psi\gamma$ decays

### $y_c$ -dependent



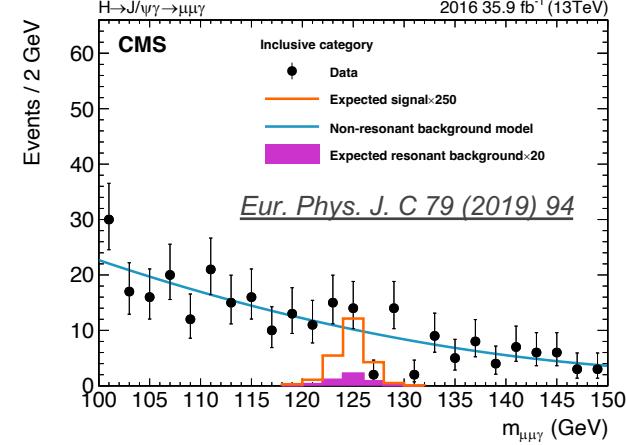
$y_c$ -independent  
(dominant contribution)

*Phys. Rev. D* 90 (2014) 11, 113010

*JHEP* 08 (2015) 012

*Phys. Rev. D* 95 (2017) 5, 054018

*Phys. Rev. D* 100 (2019) 5, 054038



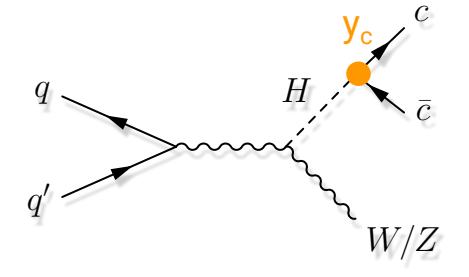
$\mathcal{B}(H \rightarrow J/\Psi\gamma) < 220 \times \text{SM}(\text{obs.})$

$\mathcal{B}(H \rightarrow J/\Psi\gamma) < 170 \times \text{SM}(\text{exp.})$

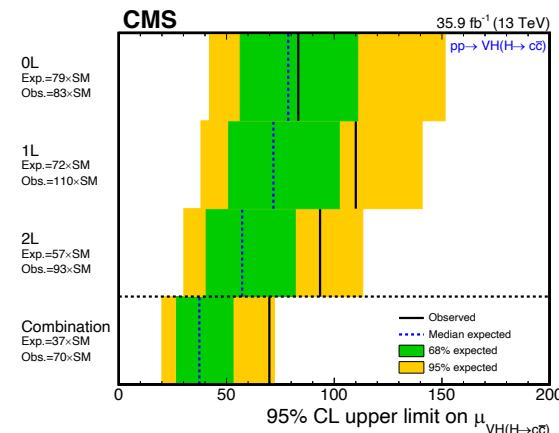
Roughly translates to  $\kappa_c < 0(100)$

# Direct search for $H \rightarrow cc$

- Search for  $H \rightarrow cc$  decays: directly sensitive to  $y_c$ , but very challenging
  - small branching fraction ( $\sim 3\%$ ) vs. large backgrounds (at a hadron collider)
  - charm quark identification is the key
- Exploit associated  $VH$  production ( $V = W, Z$ )
  - three channels:  $Z \rightarrow vv$  (0L),  $W \rightarrow \ell v$  (1L),  $Z \rightarrow \ell\ell$  (2L) [ $\ell = e, \mu$ ]
- Main backgrounds
  - $V + \text{jets}$ , single and pair production of top quarks, dibosons
  - $VH(H \rightarrow bb)$ : small but largely irreducible
- Baseline event selections
  - (high- $p_T$ ) vector boson recoiling against a Higgs boson candidate
  - veto events with high jet multiplicity to suppress  $t\bar{t}$  contribution (0L & 1L)
- Previous result ( $36 \text{ fb}^{-1}$ ): [\[JHEP 03 \(2020\) 131\]](#)
- Today: result with the full Run 2 data set ( $138 \text{ fb}^{-1}$ ) [CMS-PAS-HIG-21-008](#)



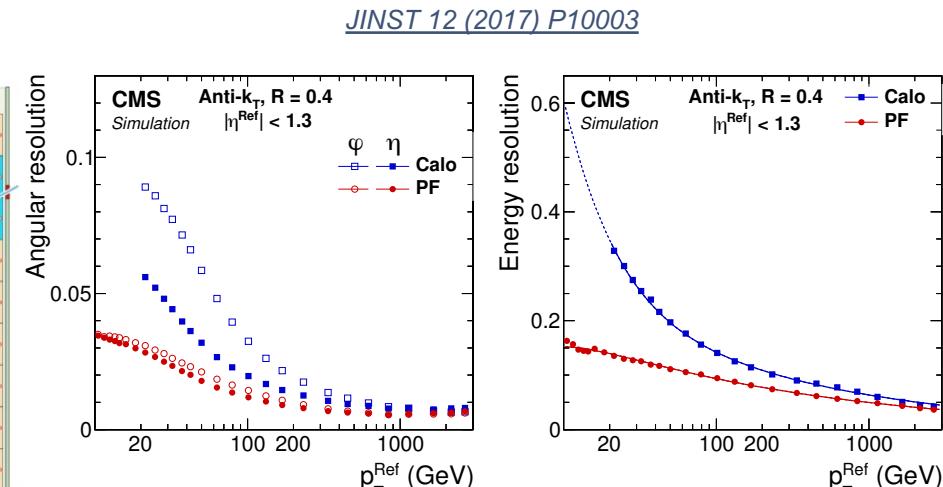
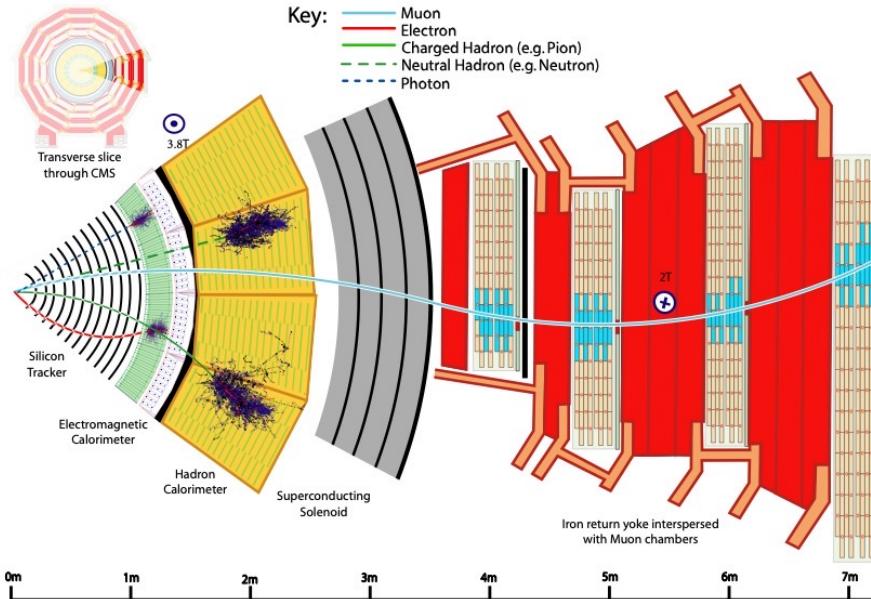
[JHEP 03 \(2020\) 131](#)



# Particle-flow reconstruction

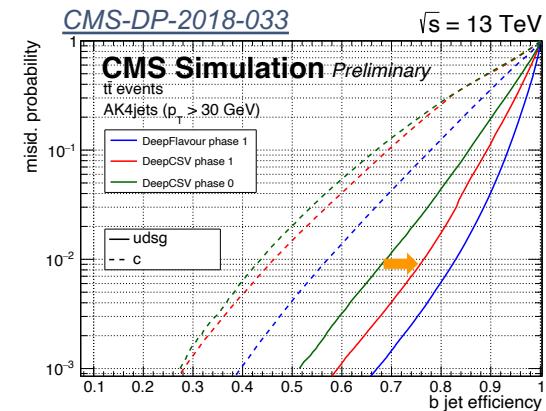
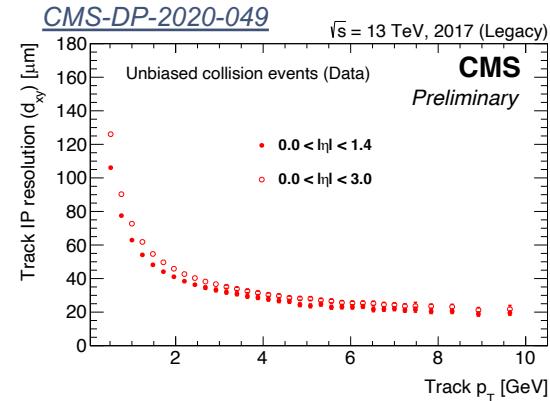
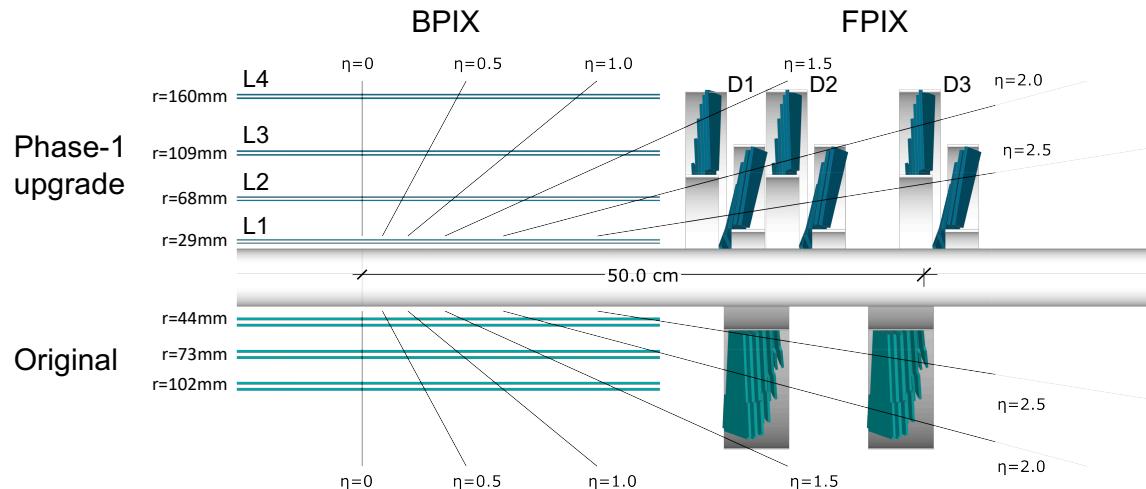
## □ Particle-flow (PF): powerful approach for jet reconstruction and flavor tagging

- excellent energy and angular resolutions
- each particle (PF candidate) contains a rich set of information from multiple sub-detectors – inputs to deep-learning



# Phase-1 pixel detector upgrade

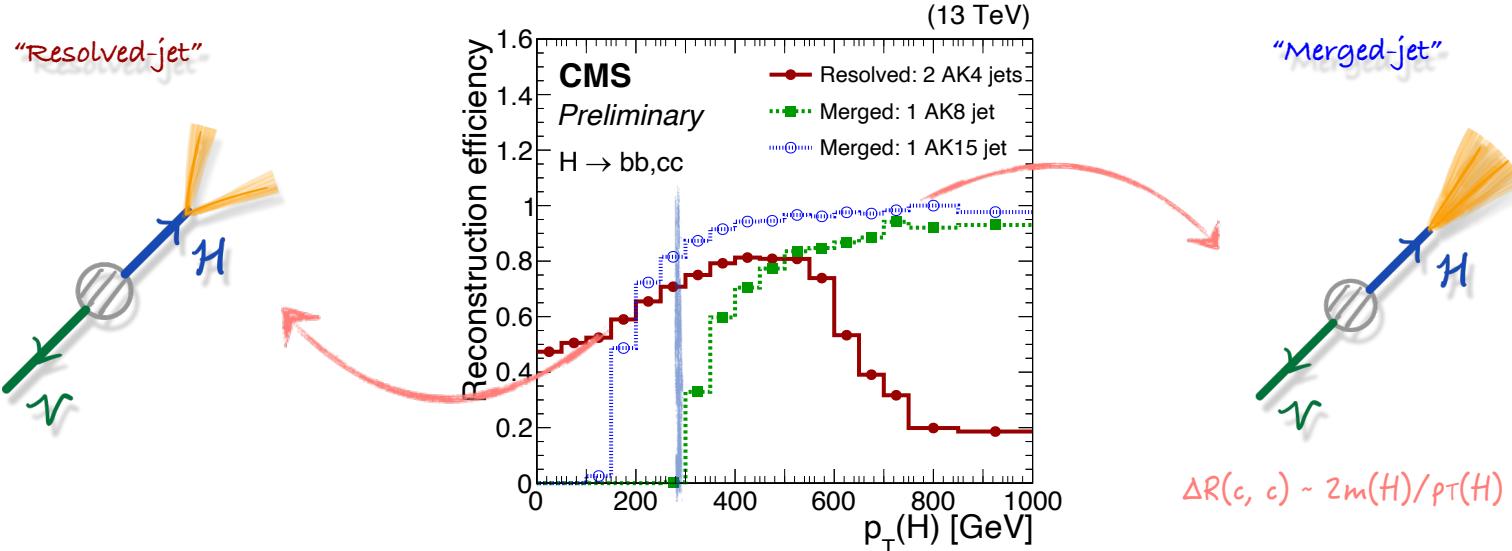
- New pixel detector installed during year-end stop 2016/2017



*Improved tracking and flavour tagging performance in the 2017 – 2018 data set!*

# Analysis overview

- Two complementary approaches for Higgs boson candidate reconstruction



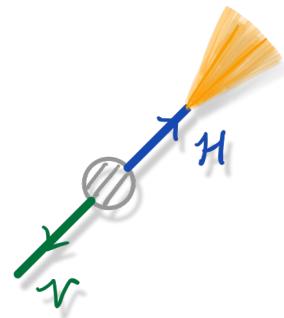
- Resolved-jet topology

- reconstructs  $H \rightarrow cc$  decay with two small-R jets ( $R=0.4$ , "AK4")
- probes the bulk (>95%) of the signal phase space

- Merged-jet topology

- reconstructs  $H \rightarrow cc$  decay with one large-R jets ( $R=1.5$ , "AK15")
- small signal acceptance (<5%) but higher purity
- better exploits the correlation between the two charm quarks

# Merged-jet topology



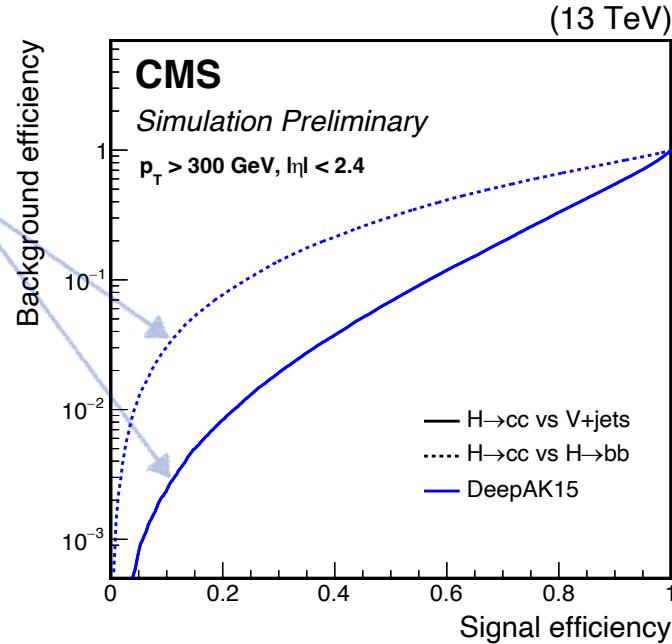
# $H \rightarrow cc$ identification

- Merged-jet topology: Higgs boson candidate reconstructed via a single large-R jet ( $p_T > 300$  GeV)

**JHEP 03 (2020) 131 (2016 analysis)**

**DeepAK8 (DeepAK15)** [[JINST 15 \(2020\) P06005](#)]

- multi-class DNN boosted jet classifier
- directly uses jet constituents (particle-flow candidates / secondary vertices)
- 1D convolutional neural network
- mass decorrelation via adversarial training



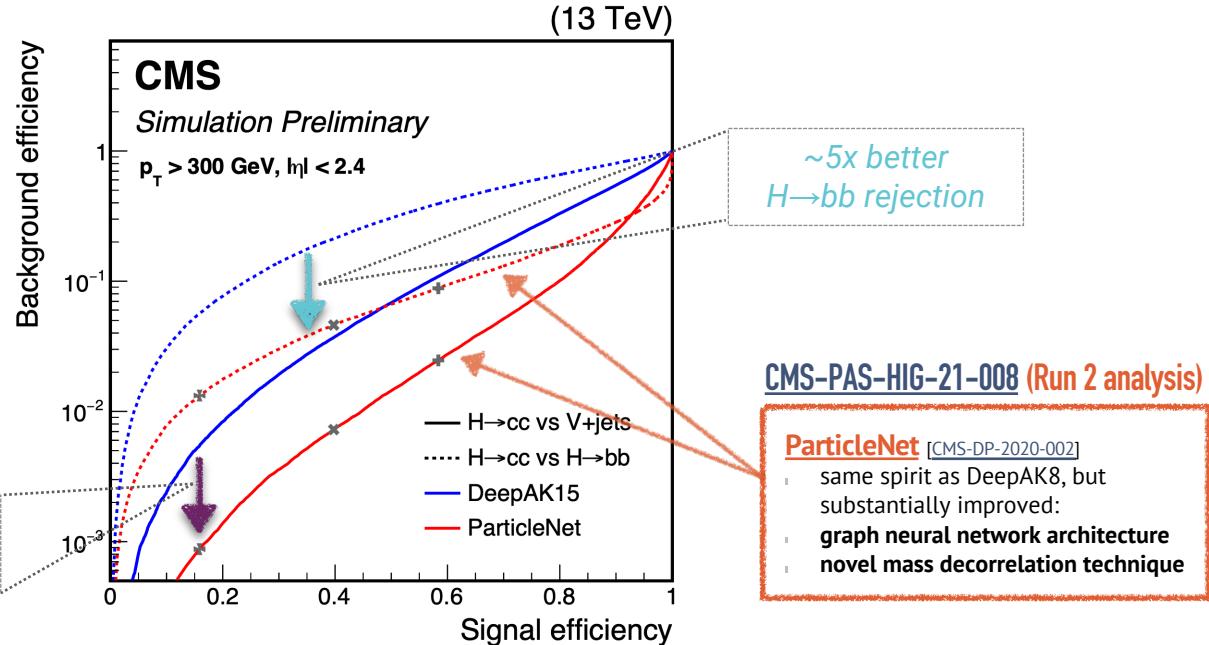
# $H \rightarrow cc$ identification

- Merged-jet topology: Higgs boson candidate reconstructed via a single large-R jet ( $p_T > 300$  GeV)
- A major improvement: **ParticleNet** tagger used to identify  $H \rightarrow cc$  decay

[JHEP 03 \(2020\) 131 \(2016 analysis\)](#)

**DeepAK8 (DeepAK15)** [[JINST 15 \(2020\) P06005](#)]

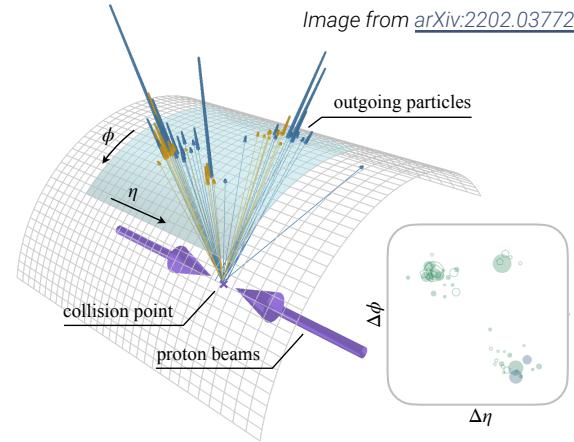
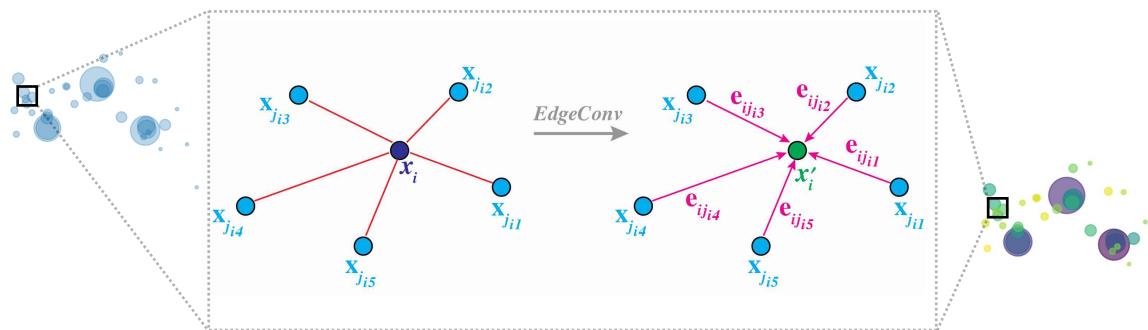
- multi-class DNN boosted jet classifier
- directly uses jet constituents (particle-flow candidates / secondary vertices)
- 1D convolutional neural network
- mass decorrelation via adversarial training



$>2x$  improvement in the final sensitivity

# ParticleNet architecture

- New jet representation: “particle cloud”
  - treating a jet as an unordered set of particles, distributed in the  $\eta - \varphi$  space
  
- ParticleNet [[Phys.Rev.D 101 \(2020\) 5, 056019](#)]
  - graph neural network architecture adapted from DGCNN [[arXiv:1801.07829](#)]
  - permutation-invariant architecture leads to significant performance improvement

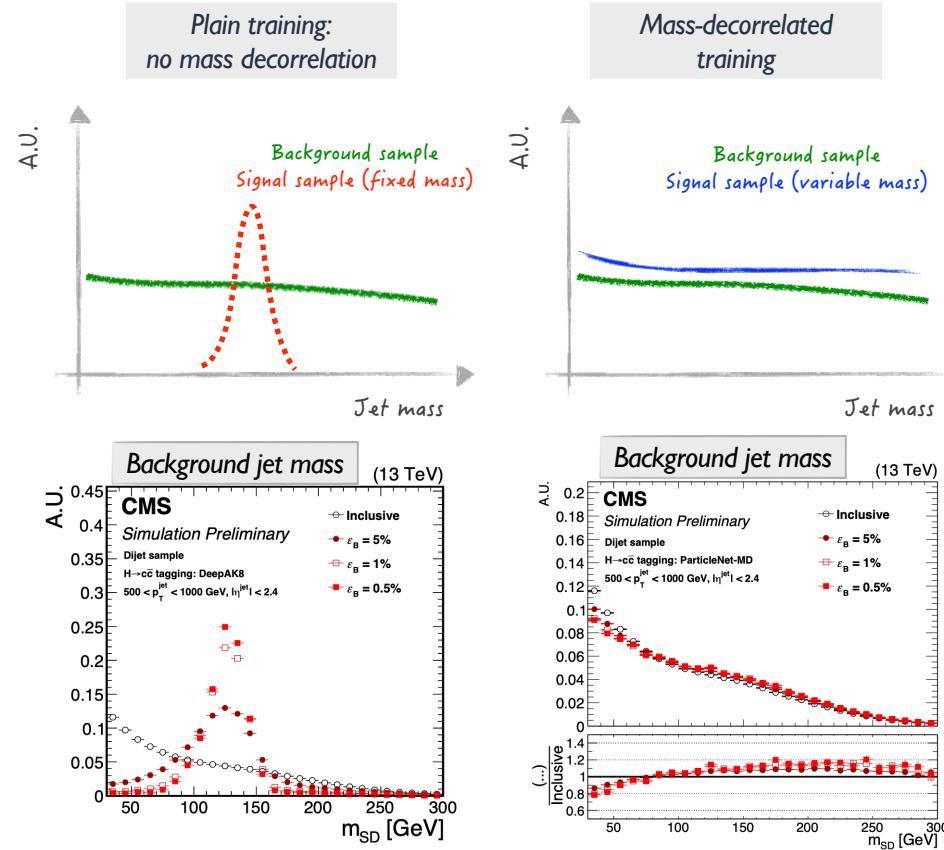


Performance on top quark tagging benchmark  
[[SciPost Phys. 7, 014 \(2019\)](#)]

| 1/ $\epsilon_b$ at $\epsilon_s = 30\%$ |                                 |
|--|---------------------------------|
| ResNeXt-50                             | $1147 \pm 58$                   |
| P-CNN                                  | $759 \pm 24$                    |
| PFN                                    | $888 \pm 17$                    |
| ParticleNet-Lite                       | $1262 \pm 49$                   |
| <b>ParticleNet</b>                     | <b><math>1615 \pm 93</math></b> |

# Mass decorrelation

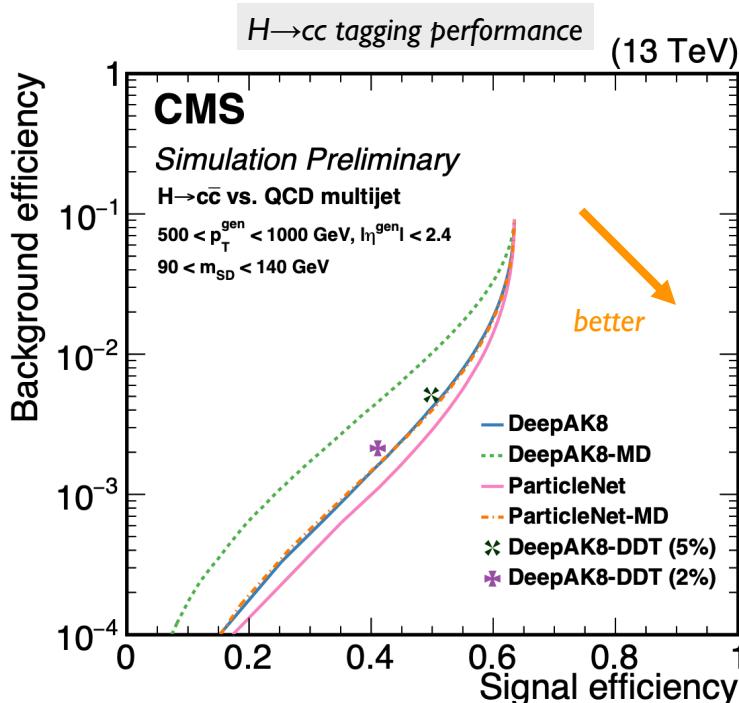
[CMS-DP-2020-002](#)



- “Mass sculpting”: background jet mass shape becomes similar to signal after tagger selection
- New approach to prevent mass sculpting
  - using a special signal sample for training
    - hadronic decays of a spin-0 particle X
      - $X \rightarrow bb$ ,  $X \rightarrow cc$ ,  $X \rightarrow qq$
    - not a fixed mass, but a **flat mass spectrum**
      - $m(X) \in [15, 250] \text{ GeV}$
  - allows to easily reweight both signal and background to a  $\sim$ flat 2D distribution in ( $p_T$ , mass) for the training
- Signal and background have the same (~flat) mass spectrum, thus no sculpting will develop in the training

# Mass decorrelation (II)

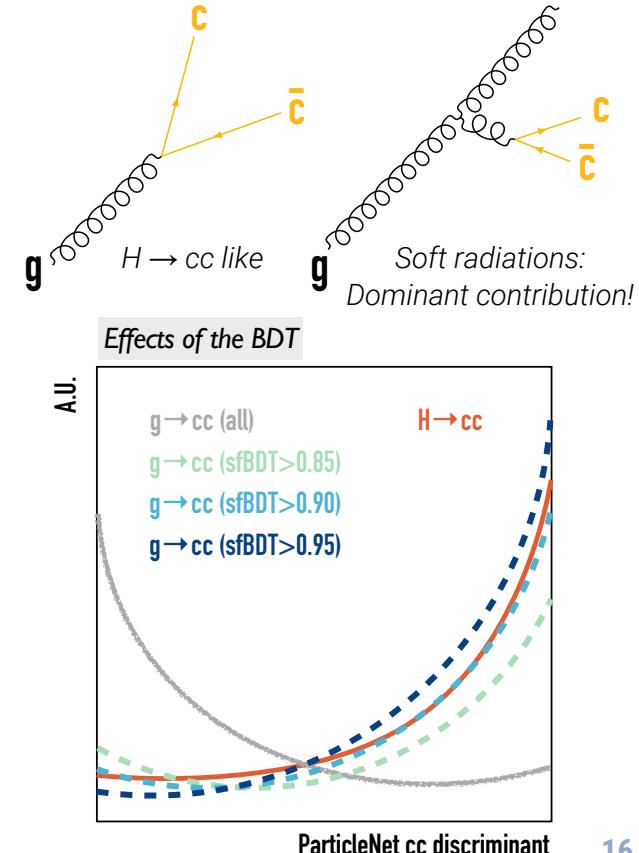
CMS-DP-2020-002



- “Mass sculpting”: background jet mass shape becomes similar to signal after tagger selection
- New approach to prevent mass sculpting
  - using a special signal sample for training
    - hadronic decays of a spin-0 particle X
      - $X \rightarrow bb$ ,  $X \rightarrow cc$ ,  $X \rightarrow qq$
    - not a fixed mass, but a **flat mass spectrum**
    - $m(X) \in [15, 250] \text{ GeV}$
  - allows to easily reweight both signal and background to a  $\sim$ flat 2D distribution in ( $p_T$ , mass) for the training- Performance loss due to mass decorrelation greatly reduced compared to the previous approach (DeepAK8-MD, based on “adversarial training”)

# Calibration of the cc-tagger

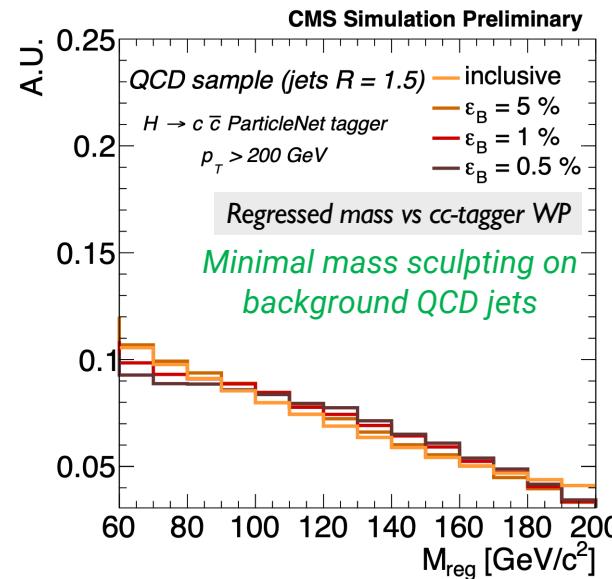
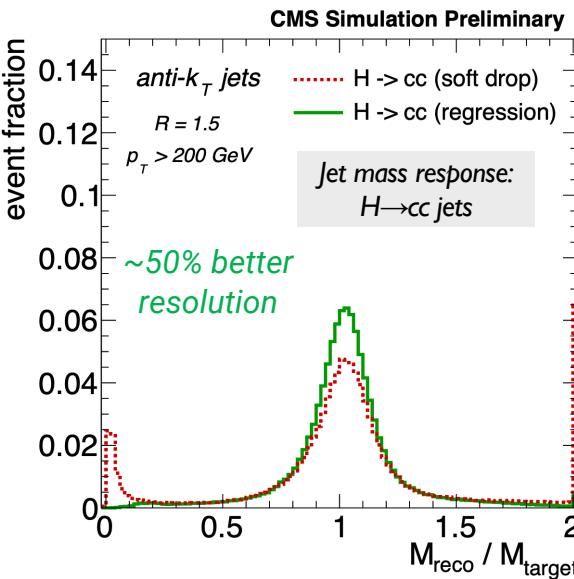
- Need to measure ParticleNet cc-tagging efficiency in data
  - no pure sample of  $H \rightarrow cc$  jets (or even  $Z \rightarrow cc$ ) in data
  - using  $g \rightarrow cc$  in QCD multi-jet events as a proxy
- Difficulty: select a phase-space in  $g \rightarrow cc$  that resembles  $H \rightarrow cc$ 
  - solution: a **dedicated BDT** developed to distinguish **hard 2-prong splittings** (i.e., high quark contribution to the jet momentum) from **soft cc radiations** (i.e., high gluon contribution to the jet momentum)
  - also allows to adjust the similarity between proxy and signal jets
    - by varying the sfBDT cut – treated as a systematic uncertainty
- Perform a fit to the secondary vertex mass shapes in the “passing” and “failing” regions simultaneously to extract the scale factors
  - three templates: cc (+ single c), bb (+ single b), light flavor jets
- Derived cc-tagging scale factors typically 0.9–1.3
  - corresponding uncertainties are 20–30%



# Large-R jet mass regression

- Jet mass: one of the most powerful observable to distinguish signal and backgrounds
- New ParticleNet-based regression algorithm to improve the large-R jet mass reconstruction
  - training setup similar to the ParticleNet tagger; the regression target:
    - signal ( $X \rightarrow bb/cc/qq$ ): generated particle mass of  $X$  [flat spectrum in 15 – 250 GeV]
    - background (QCD) jets: soft drop mass of the particle-level jet

CMS DP-2021/017



20 – 25% improvement  
in the final sensitivity

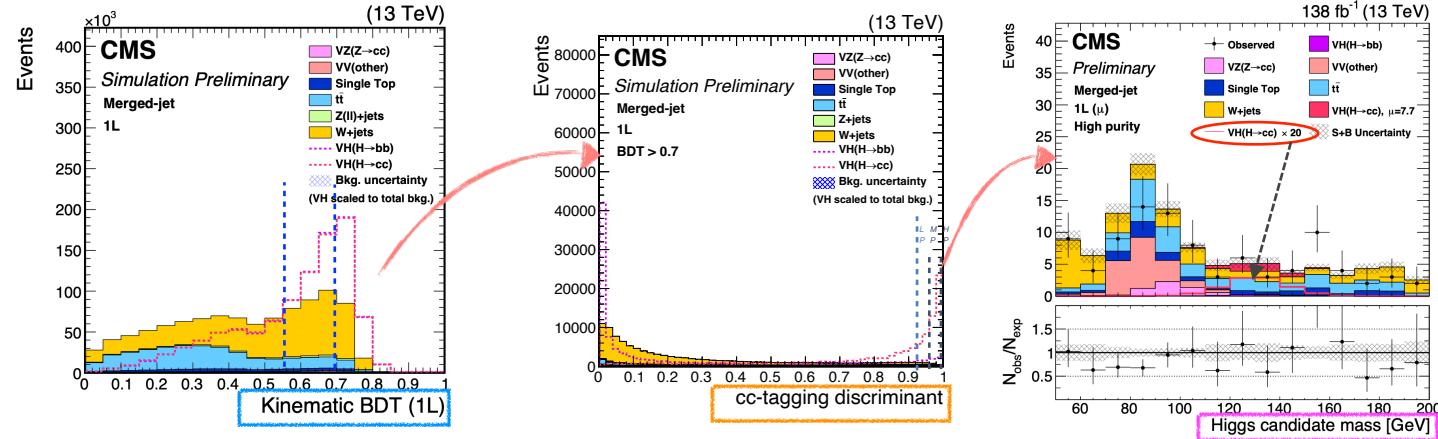
# Analysis strategy

## Factorized approach for analysis design

- event-level **kinematic BDT** developed in each channel to better suppress main backgrounds ( $V+jets$ ,  $t\bar{t}$ )
  - using only event *kinematics*, no intrinsic properties (e.g., mass/flavor) of the large-R jet
- ParticleNet cc-tagger** then used to define 3 cc-flavor enriched regions and reject light/bb-flavor jets
- finally: fit to the **ParticleNet-regressed large-R jet mass** shape for signal extraction

## Kinematic BDT, ParticleNet cc-tagger and regressed jet mass largely independent of each other

- allowing for a simple and robust strategy for background estimation and signal extraction



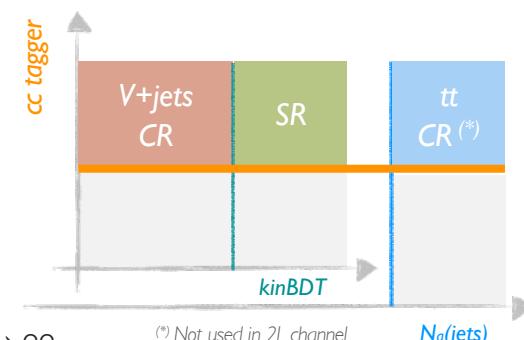
# Background estimation

## □ Normalizations of main backgrounds estimated via dedicated data control regions (CRs)

- V+jets CR: use the low kinematic BDT region
- tt CR (0L & 1L): invert the cut on the number of additional small-R jets (i.e.,  $N_{aj} \geq 2$ )
- free-floating parameters scale the normalizations in CRs and signal regions (SRs) simultaneously

## □ CRs designed to have similar jet flavor composition as the SR

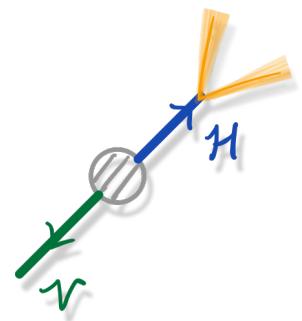
- flavor-independent kinematic BDT + same cc-tagging requirement in CRs as in SR
- allows to correct cc-tagging efficiency for backgrounds directly from data
- cc-tagging SFs only needed for the signal VH( $H \rightarrow cc$ ) process (and VZ( $Z \rightarrow cc$ ))
  - conservative uncertainty ( $2x/0.5x$ ) for the misidentification of  $H(Z) \rightarrow bb$  as  $H(Z) \rightarrow cc$



## □ Minor backgrounds (single top, dibosons, VH( $H \rightarrow bb$ )) estimated from simulation

- dibosons: applying differential NNLO QCD + NLO EW corrections as a function of  $p_T(V)$  [[JHEP 002 \(2020\) 087](#)]

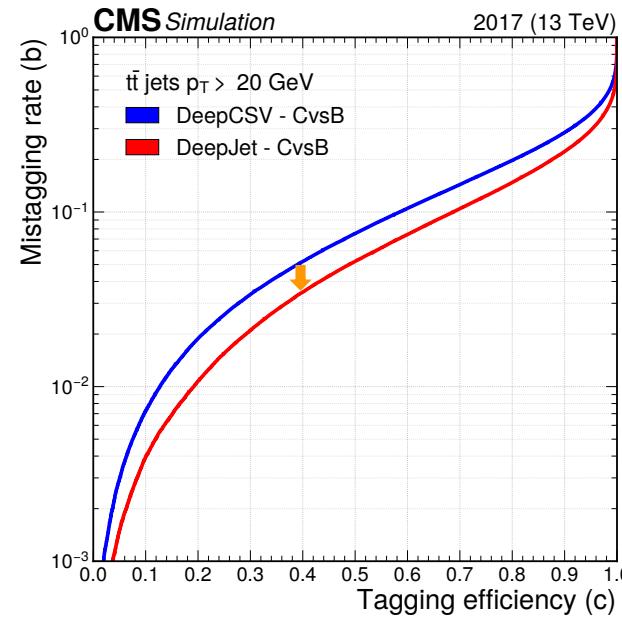
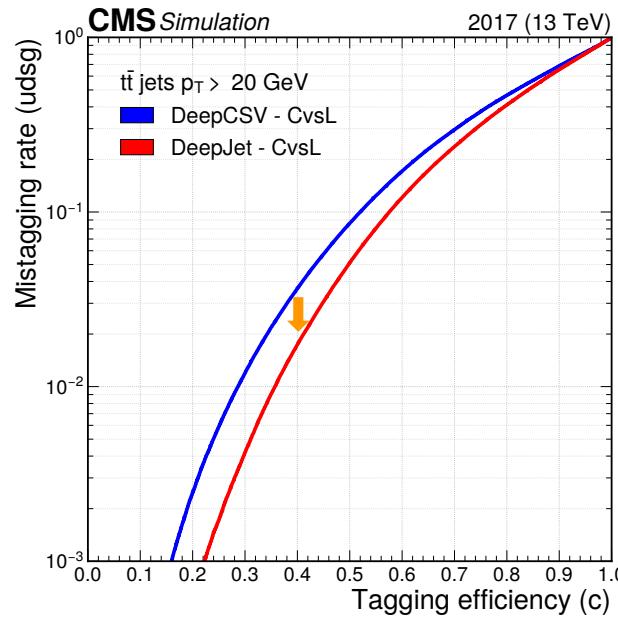
# Resolved-jet topology



# Charm quark identification

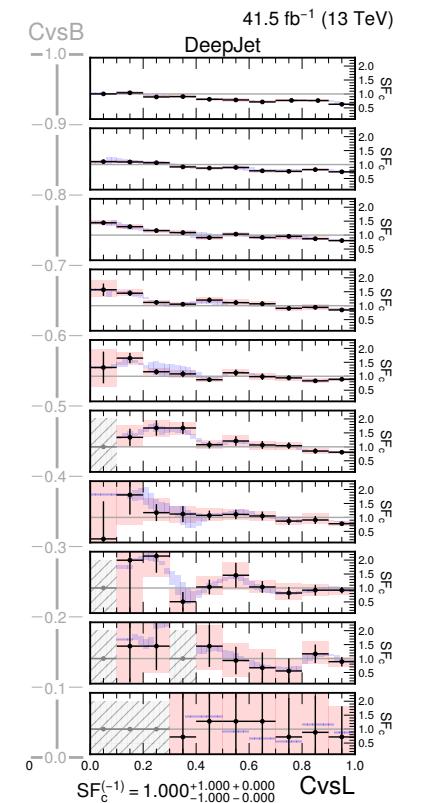
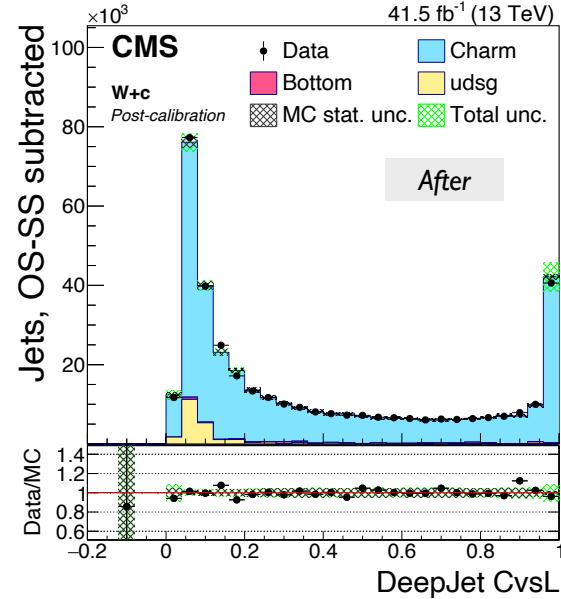
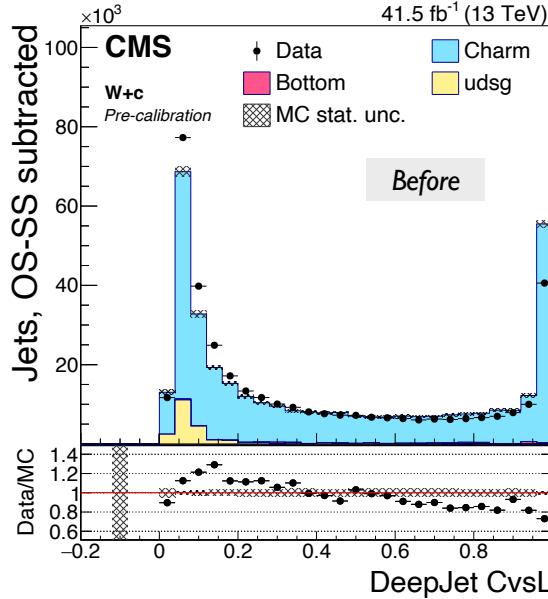
- Resolved-jet topology: Higgs boson candidate reconstructed with two small-R jets
- Charm quark jet identification: **DeepJet** algorithm
  - ~2x (~40%) improvement in light (b) jet rejection at 40% c jet efficiency compared to DeepCSV

[arXiv:2111.03027](https://arxiv.org/abs/2111.03027)



# Charm tagging calibration

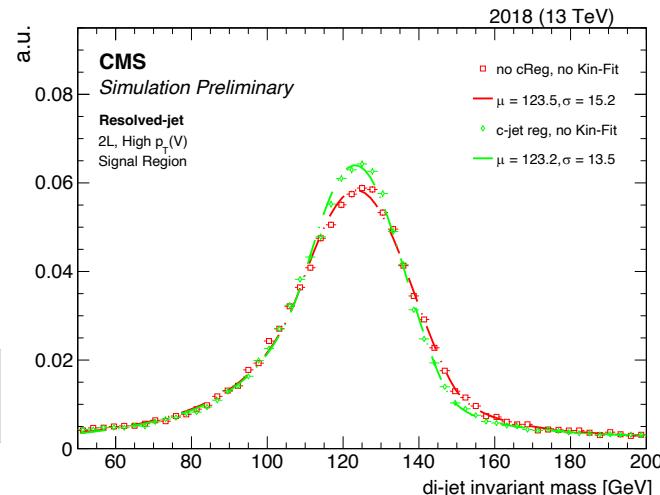
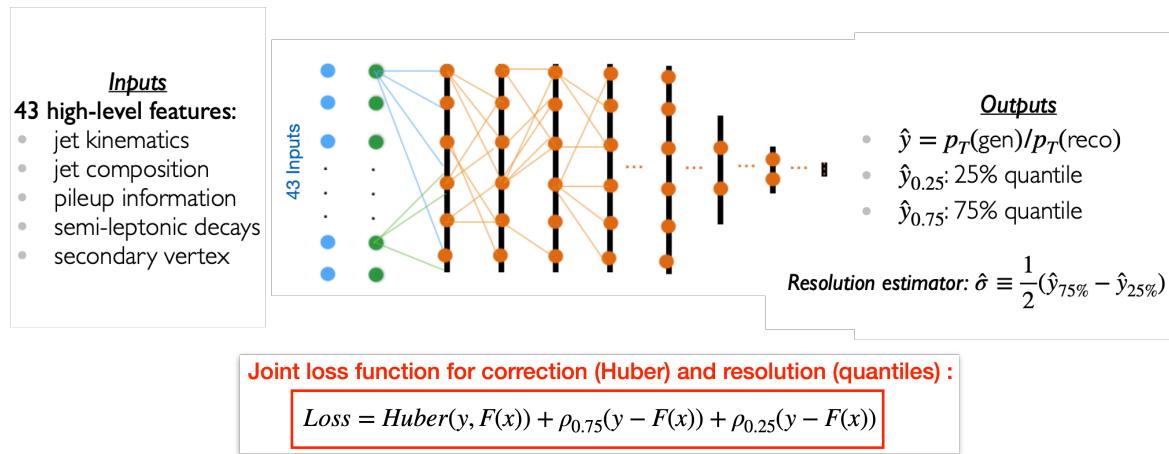
- Novel calibration method to correct the entire distributions of the c-tagging discriminants [arXiv:2111.03027](https://arxiv.org/abs/2111.03027)
  - per-jet SFs derived as a function of (CvsL, CvsB | truth flavor)
- SFs derived with an iterative approach using three samples
  - $Z(\ell\ell) + \text{jets}$  (light jet enriched);  $W + c$  (c-jet enriched);  $t\bar{t}$  (b-jet enriched)



# Charm jet energy regression

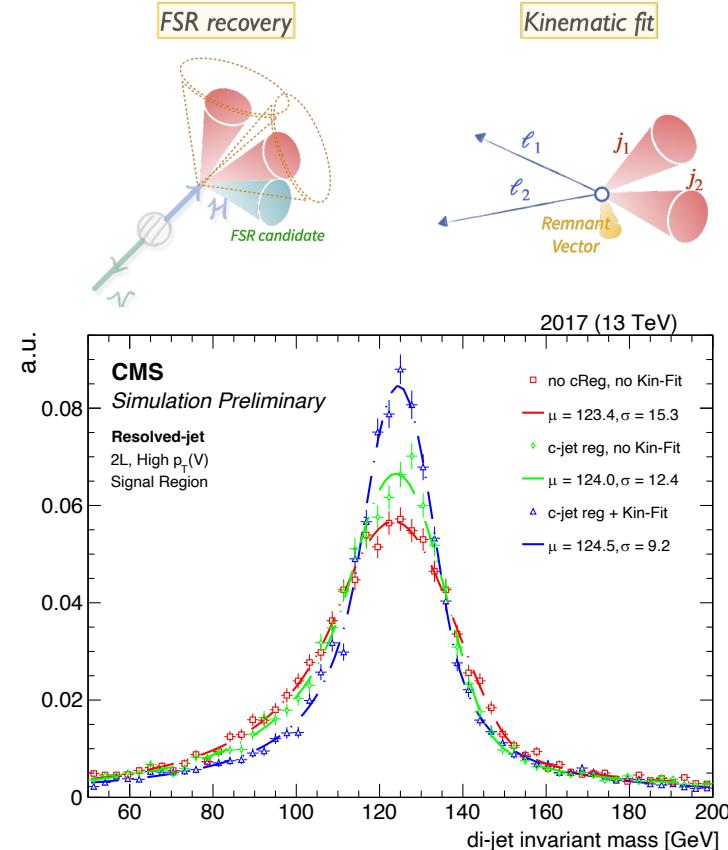
## □ Dedicated jet energy regression algorithm developed to improve the c-jet energy scale and resolution

- based on the b jet energy regression [[Comput. Softw. Big Sci. 4 \(2020\) 10](#)] used in several CMS H $\rightarrow$ bb analyses
- re-trained for c jets instead of b jets
  - c jets collected from W $\rightarrow$ c $\bar{c}$  decay in tt MC events
- provides simultaneous estimation of the c jet energy and its resolution
  - both used as inputs to the signal extraction BDT



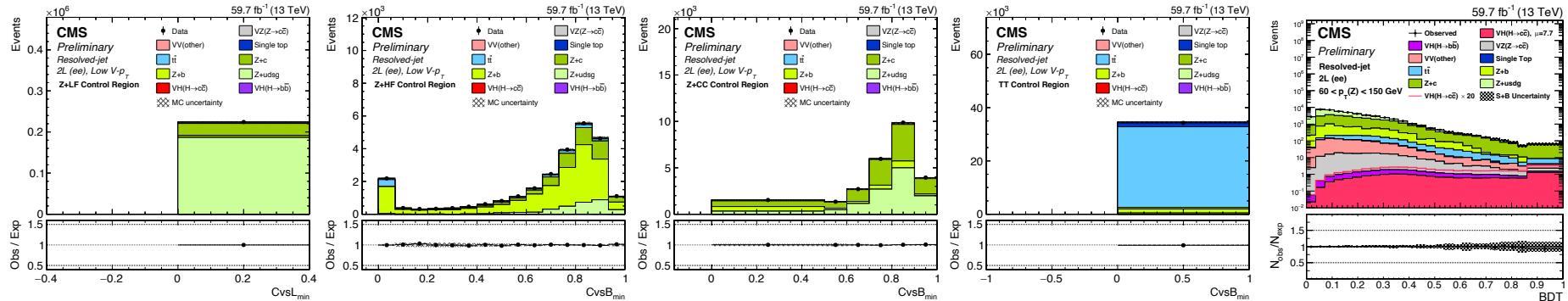
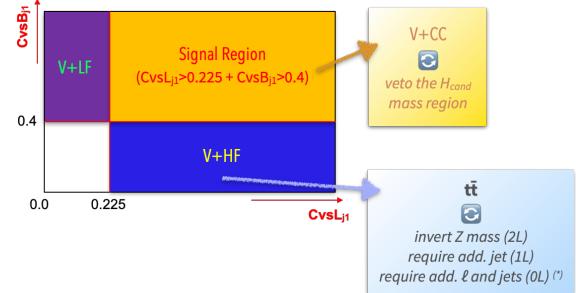
# Higgs boson candidate reconstruction

- Higgs boson candidate reconstructed using the two small-R jets with highest CvsL scores
- To improve the Higgs candidate mass resolution:
  - recovery of final state radiations (FSR)
    - additional jets within  $\Delta R < 0.8$  from either of the two selected jets are included in the calculation of the Higgs candidate's 4-momentum
    - improves the Higgs mass resolution by a few percent
  - new DNN-based c-jet energy regression
    - ~20% improvement in Higgs candidate mass resolution
  - improved kinematic fit in the 2L channel
    - better reconstruction of the Higgs candidate's 4-momentum using constraints from the  $Z \rightarrow \ell\ell$  system
    - up to 30% improvement in Higgs candidate mass resolution



# Signal extraction strategy

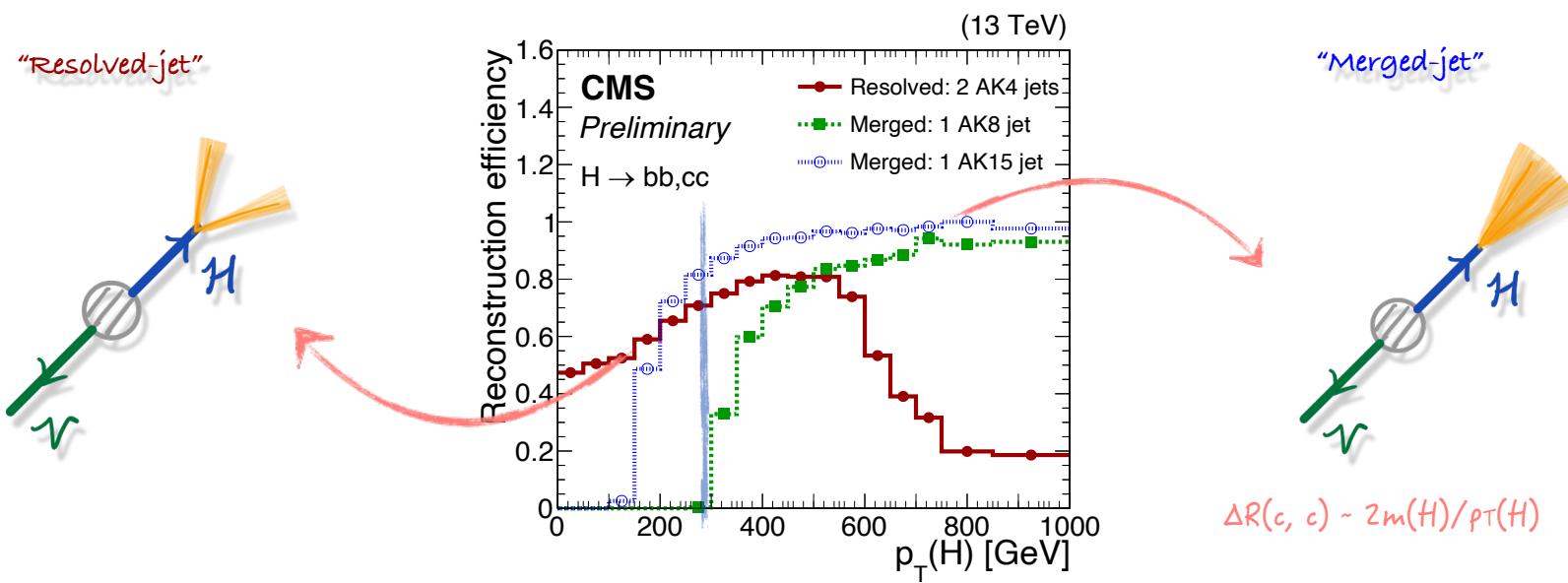
- Event-level BDT trained in each channel to maximize the signal vs background separation
  - inputs: event kinematics, Higgs candidate properties, c-tagging discriminants
    - + kinematic fit variables in 2L
- Background estimation
  - dedicated CRs to constrain the normalizations of main backgrounds ( $V + \text{jets}$ ,  $t\bar{t}$ )
    - $V + \text{jets}$  split based on flavor:  $V + b$ ,  $V + c$ ,  $V + \text{udsg}$
- Simultaneous fit of SRs (BDT shapes) and CRs (c-tagging discriminants) for signal extraction



# Results

# Combination of the two topologies

- The two topologies are made orthogonal via the presence of large-R jet with  $p_T > 300$  GeV
  - $p_T$  threshold chosen to maximize expected sensitivity



# Uncertainties

## □ Systematic uncertainties correlated between topologies, except:

- background normalizations for V+jets and tt
- charm quark identification efficiencies

*Relative contributions to the total uncertainty on  $\mu$*

| Uncertainty source                 | $\Delta\mu / (\Delta\mu)_{\text{tot}}$ |
|------------------------------------|--|
| <b>Statistical</b>                 | <b>85%</b>                             |
| Background normalizations          | 37%                                    |
| <b>Experimental</b>                | <b>48%</b>                             |
| Sizes of the simulated samples     | 37%                                    |
| Charm identification efficiencies  | 23%                                    |
| Jet energy scale and resolution    | 15%                                    |
| Simulation modeling                | 11%                                    |
| Luminosity                         | 6%                                     |
| Lepton identification efficiencies | 4%                                     |
| <b>Theory</b>                      | <b>22%</b>                             |
| Backgrounds                        | 17%                                    |
| Signal                             | 15%                                    |

## □ Main uncertainties

- limited statistics of the data set
- size of simulated samples (especially NLO V+jets)
- charm quark identification efficiencies

# VZ( $Z \rightarrow cc$ ) results

## □ The full analysis procedure is validated by measuring the VZ( $Z \rightarrow cc$ ) process

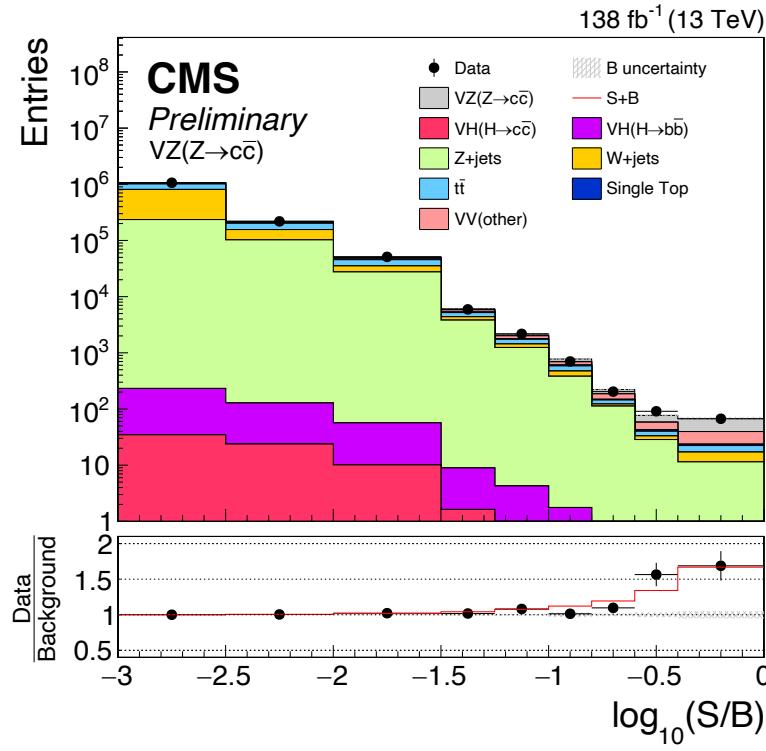
- resolved-jet topology:
  - BDT re-trained using VZ( $Z \rightarrow cc$ ) as signal
  - fit to the BDT shapes to extract the signal
- merged-jet topology:
  - no change to the analysis procedure
  - fit to the large-R jet mass shapes to extract the signal

# VZ( $Z \rightarrow cc$ ) results

- The full analysis procedure is validated by measuring the VZ( $Z \rightarrow cc$ ) process

[CMS-PAS-HIG-21-008](#)

Search for  $H \rightarrow cc$  with CMS – March 7, 2022 - Huilin Qu (CERN)



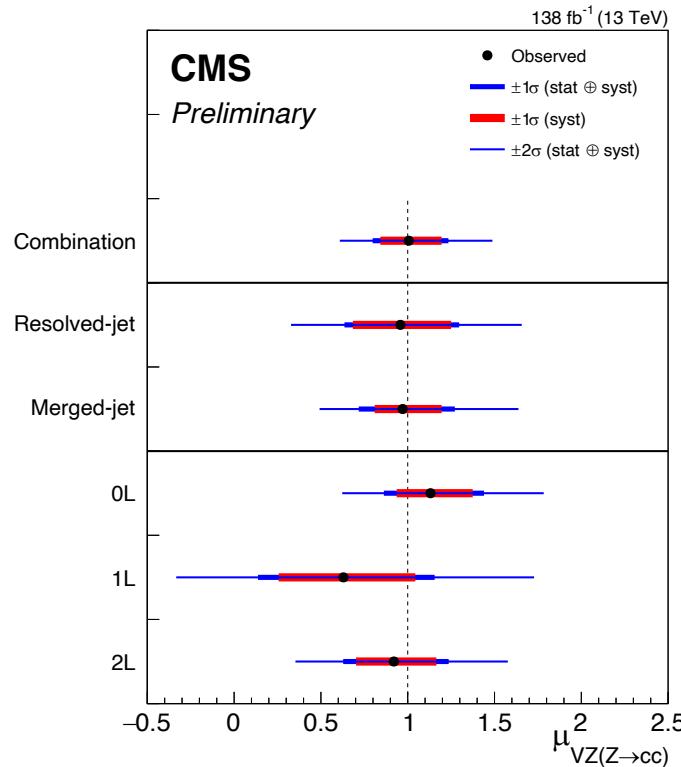
Observed significance for VZ( $Z \rightarrow cc$ ): **5.7 $\sigma$**   
- expected significance: 5.9 $\sigma$

**First observation of  $Z \rightarrow cc$  at a hadron collider!**

# VZ( $Z \rightarrow cc$ ) results

- The full analysis procedure is validated by measuring the VZ( $Z \rightarrow cc$ ) process

[CMS-PAS-HIG-21-008](#)



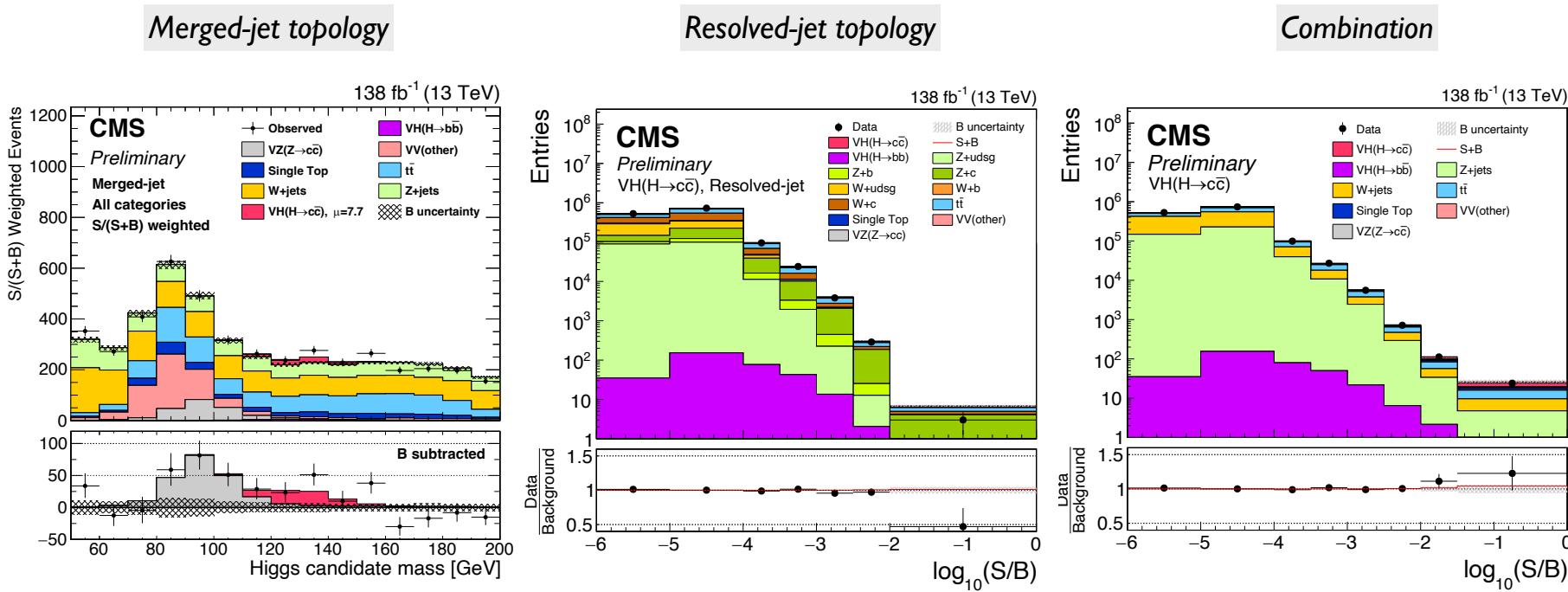
Best-fit signal strength:  $\mu_{VZ(Z \rightarrow cc)} = 1.01^{+0.23}_{-0.21}$

- very good agreement with SM expectation
- consistent results between topologies/channels

# VH( $H \rightarrow cc$ ) results

- Post-fit distributions in the two topologies and the combination

[CMS-PAS-HIG-21-008](#)



# VH( $H \rightarrow cc$ ) results

- Upper limits on the VH( $H \rightarrow cc$ ) signal strength at 95% CL:

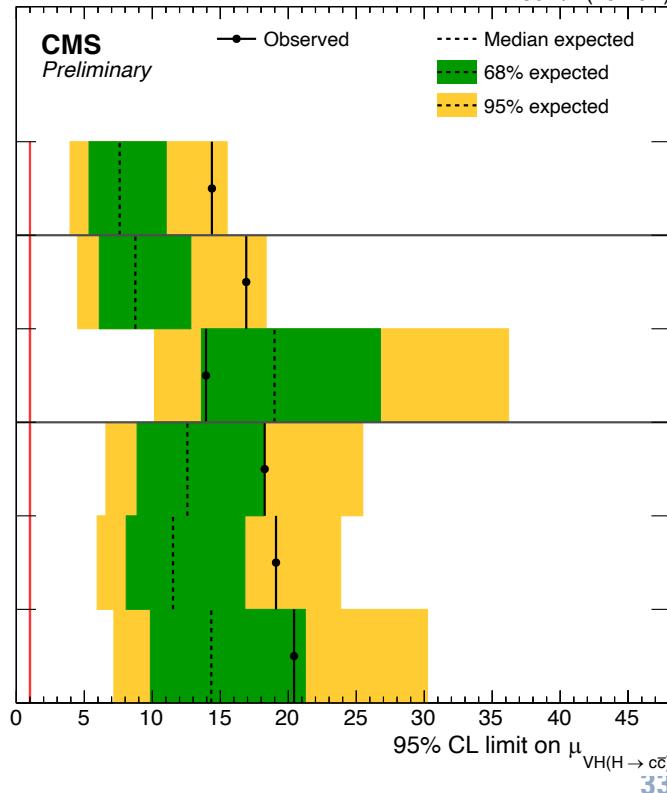
- $\mu_{VH(H \rightarrow cc)} < 14$  (7.6) observed (expected)
- substantially stronger than ATLAS full Run 2 result
  - $\mu_{VH(H \rightarrow cc)} < 26$  (31) obs. (exp.) [[arXiv:2201.11428](https://arxiv.org/abs/2201.11428)]

- Best fit signal strength

- $\mu_{VH(H \rightarrow cc)} = 7.7^{+3.8}_{-3.5}$
- consistent with the SM prediction within  $2\sigma$

[CMS-PAS-HIG-21-008](#)

138  $\text{fb}^{-1}$  (13 TeV)



# VH( $H \rightarrow cc$ ) results

- ❑ Results used to set a constraint on the charm quark Yukawa coupling modifier  $\kappa_c = y_c / y_c^{\text{SM}}$

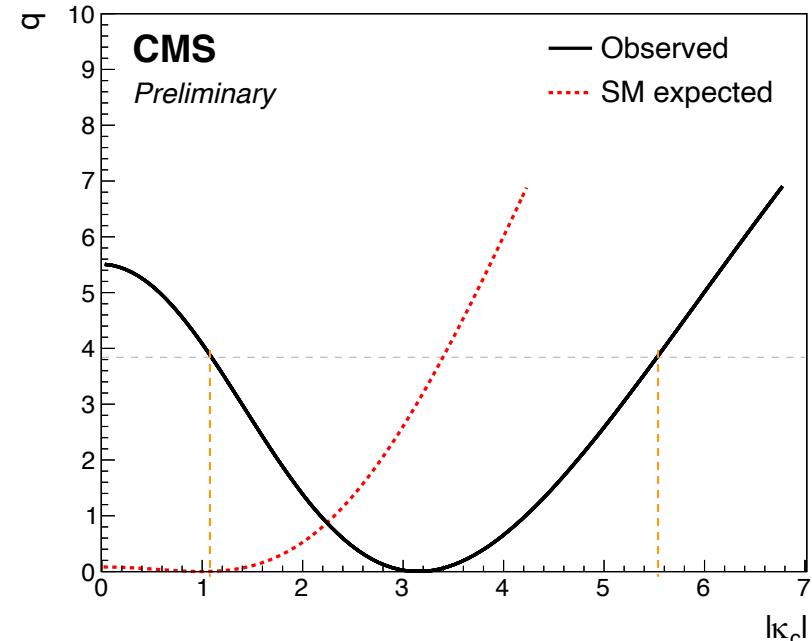
- for simplicity, only considering effects on  $B(H \rightarrow cc\bar{c}\bar{c})$  and fixing all other couplings at SM values:

$$\mu_{VH(H \rightarrow cc\bar{c}\bar{c})} = \frac{\kappa_c^2}{1 + \mathcal{B}_{\text{SM}}(H \rightarrow cc\bar{c}\bar{c}) \times (\kappa_c^2 - 1)}.$$

- ❑ The 95% CL interval on  $\kappa_c$ :

- observed:  $1.1 < |\kappa_c| < 5.5$
- expected:  $|\kappa_c| < 3.4$
- **most stringent constraint on  $\kappa_c$  to date!**
- comparable to the previous projection for HL-LHC w/  
 $3000 \text{ fb}^{-1}$ :  $|\kappa_c| < 3.0$  [ATL-PHYS-PUB-2021-039]

[CMS-PAS-HIG-21-008](#)



# **Prospects: HL-LHC**

# Projection at HL-LHC: Setup

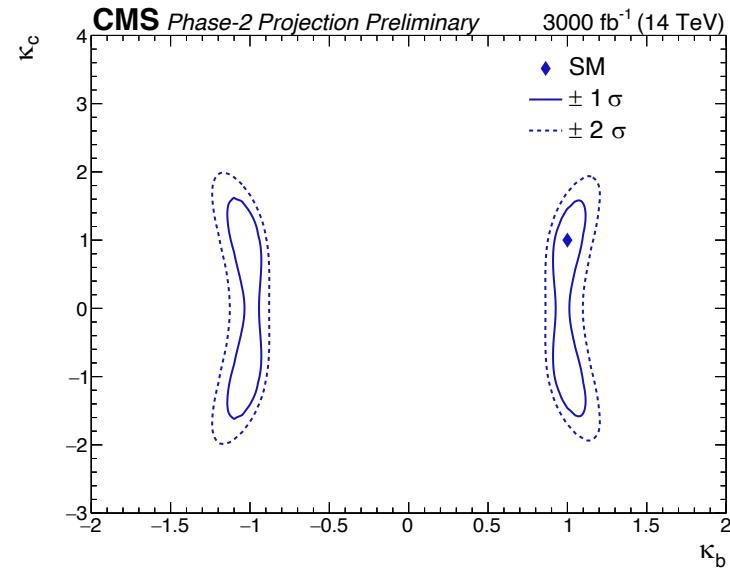
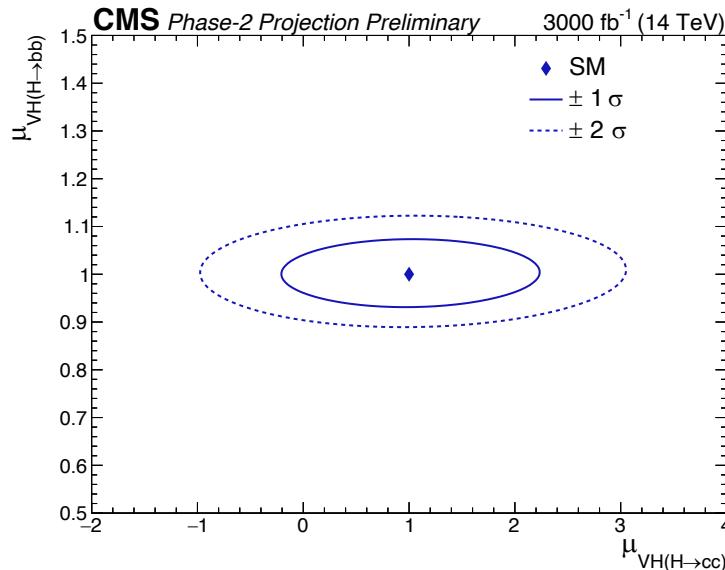
- ❑ Extrapolation of the merged-jet analysis to HL-LHC with  $3000 \text{ fb}^{-1}$  data
- ❑ Modifications to the Run 2 analysis to allow for a simultaneous constraint on  $H \rightarrow bb$  and  $H \rightarrow cc$ 
  - addition of 3 categories enriched in  $H \rightarrow bb$  decays, selected with the ParticleNet bb-tagging discriminant
    - very small (1-2%) overlap of bb and cc categories – events assigned to a unique category
  - large-R jet  $p_T$  threshold lowered from 300 GeV to 200 GeV – increasing signal acceptance
- ❑ Systematic uncertainties adjusted according to the Yellow Report [[CERN-2019-007](#)]
  - theoretical uncertainties: reduced by half
  - most experimental uncertainties: scaled down with  $\sqrt{\mathcal{L}}$ 
    - bb and cc tagging efficiencies: constrained by  $VZ(Z \rightarrow bb)$  and  $VZ(Z \rightarrow cc)$  events to  $\sim 3\%$  and  $\sim 5\%$
    - misidentification of  $H \rightarrow bb$  as  $H \rightarrow cc$ : a prominent uncertainty on  $H \rightarrow cc$  measurement at HL-LHC
      - assumed to be reduced from  $\sim 100\%$  (Run 2) to 20% in the projection

# Projection at HL-LHC

## □ Simultaneous extraction of the $H \rightarrow bb$ and $H \rightarrow cc$ signal strengths

[CMS-PAS-HIG-21-008](#)

- $\mu_{VH(H \rightarrow bb)} = 1.00 \pm 0.03 \text{ (stat.)} \pm 0.04 \text{ (syst.)} = 1.00 \pm 0.05 \text{ (total)}$
- $\mu_{VH(H \rightarrow cc)} = 1.0 \pm 0.6 \text{ (stat.)} \pm 0.5 \text{ (syst.)} = 1.0 \pm 0.8 \text{ (total)}$



**Expected sensitivity approaches the SM value for the Higgs-charm coupling.**

# **Prospects: ML for jet tagging**

# Beyond ParticleNet

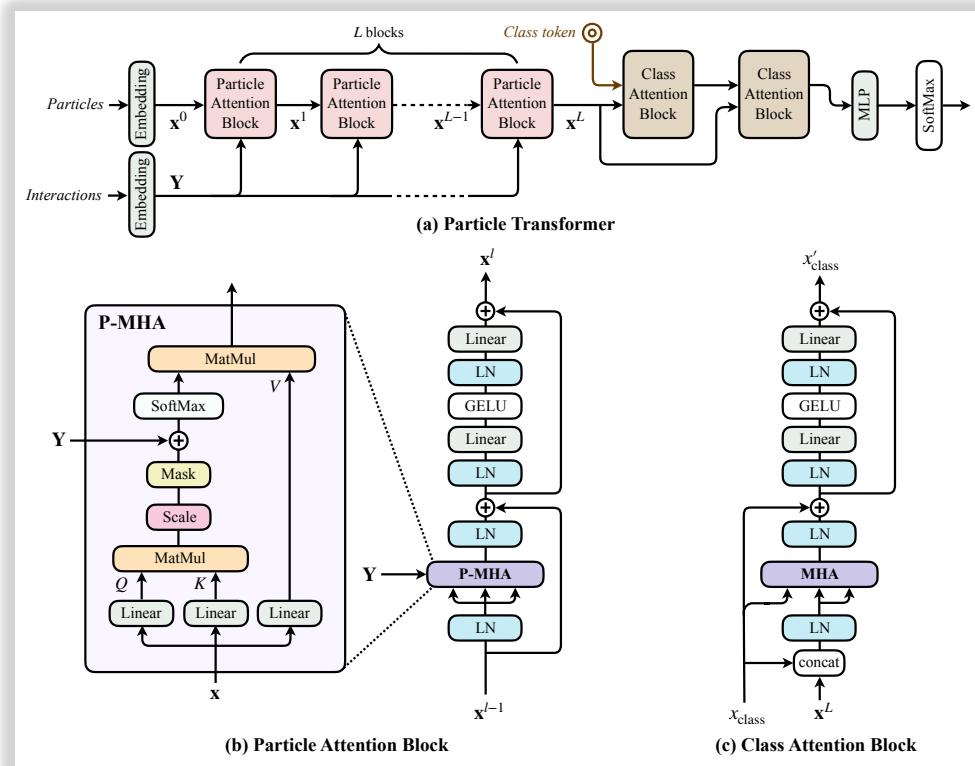
- Particle Transformer (ParT): a new Transformer-based architecture for jet tagging

H. Qu, C. Li, S. Qian,  
[arXiv:2202.03772](https://arxiv.org/abs/2202.03772)

## Interaction features

$$\begin{aligned}\Delta &= \sqrt{(y_a - y_b)^2 + (\phi_a - \phi_b)^2}, \\ k_T &= \min(p_{T,a}, p_{T,b})\Delta, \\ z &= \min(p_{T,a}, p_{T,b})/(p_{T,a} + p_{T,b}), \\ m^2 &= (E_a + E_b)^2 - \|\mathbf{p}_a + \mathbf{p}_b\|^2,\end{aligned}$$

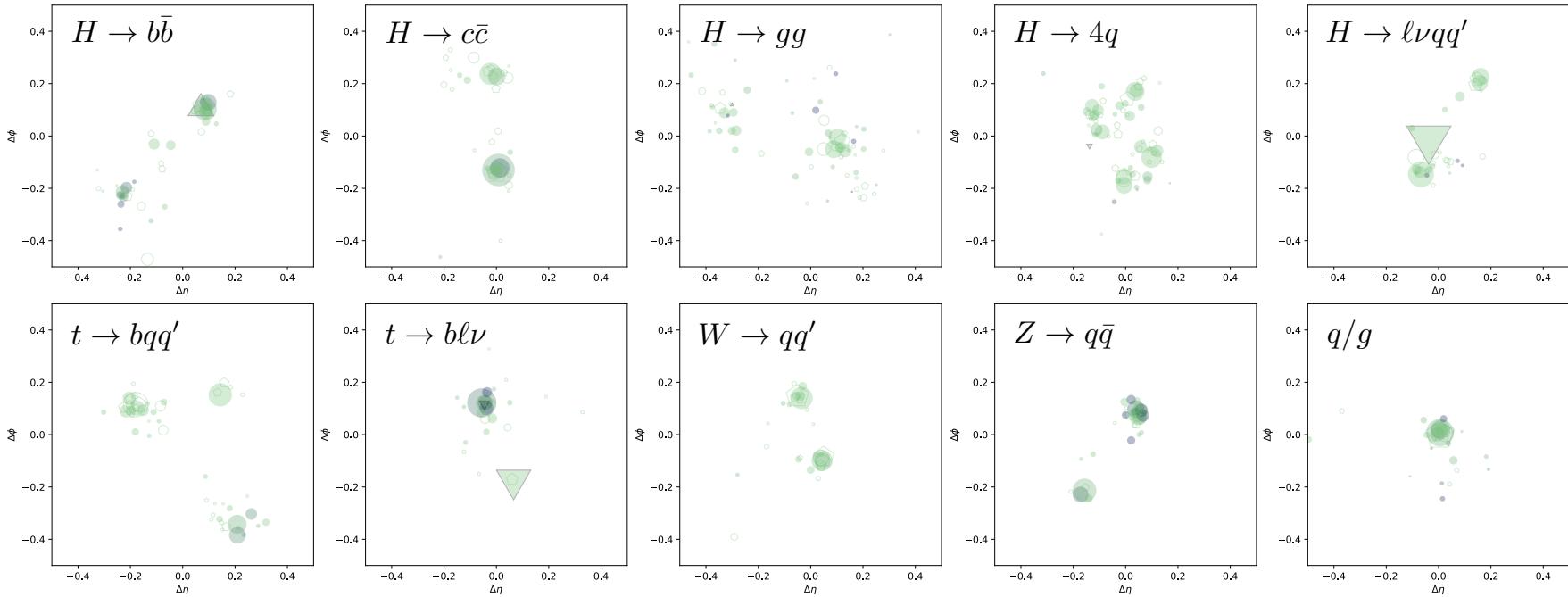
Motivated by LundNet  
[F. Dreyer and H. Qu,  
[JHEP 03 \(2021\) 052](https://doi.org/10.1007/JHEP03(2021)052)]



# Beyond ParticleNet: JetClass dataset

- A new large-scale public jet dataset: 100M jets in 10 classes
  - MadGraph + Pythia (Herwig) + Delphes

H. Qu, C. Li, S. Qian,  
[arXiv:2202.03772](https://arxiv.org/abs/2202.03772)



# Beyond ParticleNet: Particle Transformer

## □ Performance of Particle Transformer (ParT) on the JetClass dataset

H. Qu, C. Li, S. Qian,  
[arXiv:2202.03772](https://arxiv.org/abs/2202.03772)

*Performance comparison on the JetClass dataset*

|              | All classes  |               | $H \rightarrow b\bar{b}$ | $H \rightarrow c\bar{c}$ | $H \rightarrow gg$ | $H \rightarrow 4q$ | $H \rightarrow \ell\nu qq'$ | $t \rightarrow bq\bar{q}'$ | $t \rightarrow b\ell\nu$ | $W \rightarrow qq'$ | $Z \rightarrow q\bar{q}$ |
|--------------|--------------|---------------|--------------------------|--------------------------|--------------------|--------------------|-----------------------------|----------------------------|--------------------------|---------------------|--------------------------|
|              | Accuracy     | AUC           | Rej <sub>50%</sub>       | Rej <sub>50%</sub>       | Rej <sub>50%</sub> | Rej <sub>50%</sub> | Rej <sub>99%</sub>          | Rej <sub>50%</sub>         | Rej <sub>99.5%</sub>     | Rej <sub>50%</sub>  | Rej <sub>50%</sub>       |
| 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                      |
| <b>ParT</b>  | <b>0.861</b> | <b>0.9877</b> | <b>10638</b>             | <b>4149</b>              | <b>123</b>         | <b>1864</b>        | <b>5479</b>                 | <b>32787</b>               | <b>15873</b>             | <b>543</b>          | <b>402</b>               |
| ParT (plain) | 0.849        | 0.9859        | 9569                     | 2911                     | 112                | 1185               | 3868                        | 17699                      | 12987                    | 384                 | 311                      |

*Fine-tuning result on Top-tagging Benchmark [SciPost Phys. 7 (2019) 014]*

|                         | Accuracy     | AUC           | Rej <sub>50%</sub>             | Rej <sub>30%</sub>               |
|-------------------------|--------------|---------------|--------------------------------|----------------------------------|
| P-CNN                   | 0.930        | 0.9803        | $201 \pm 4$                    | $759 \pm 24$                     |
| PFN                     | —            | 0.9819        | $247 \pm 3$                    | $888 \pm 17$                     |
| ParticleNet             | 0.940        | 0.9858        | $397 \pm 7$                    | $1615 \pm 93$                    |
| JEDI-net (w/ $\sum O$ ) | 0.930        | 0.9807        | —                              | 774.6                            |
| PCT                     | 0.940        | 0.9855        | $392 \pm 7$                    | $1533 \pm 101$                   |
| LGN                     | 0.929        | 0.964         | —                              | $435 \pm 95$                     |
| rPCN                    | —            | 0.9845        | $364 \pm 9$                    | $1642 \pm 93$                    |
| ParT                    | 0.940        | 0.9858        | $413 \pm 16$                   | $1602 \pm 81$                    |
| <b>ParT-f.t.</b>        | <b>0.944</b> | <b>0.9877</b> | <b><math>691 \pm 15</math></b> | <b><math>2766 \pm 130</math></b> |

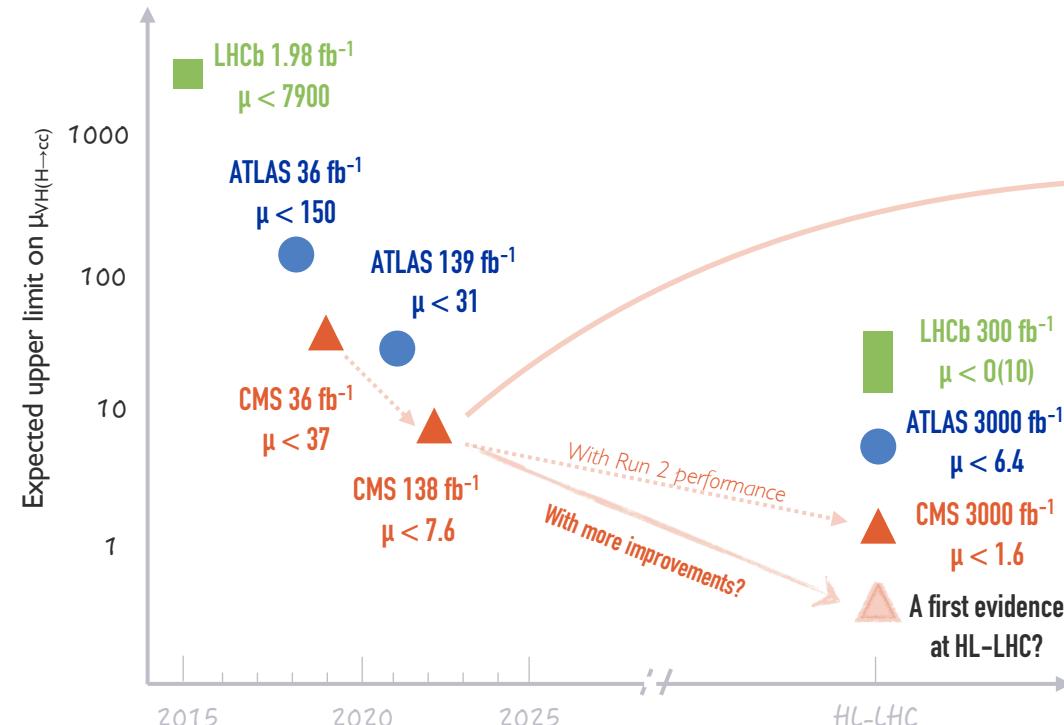
*Model complexity*

|              | Accuracy     | # params | FLOPs  |
|--------------|--------------|----------|--------|
| PFN          | 0.772        | 86.1 k   | 4.62 M |
| P-CNN        | 0.809        | 354 k    | 15.5 M |
| ParticleNet  | 0.844        | 370 k    | 540 M  |
| <b>ParT</b>  | <b>0.861</b> | 2.14 M   | 340 M  |
| ParT (plain) | 0.849        | 2.13 M   | 260 M  |

*Significant performance improvement.  
 Similar computational cost.*

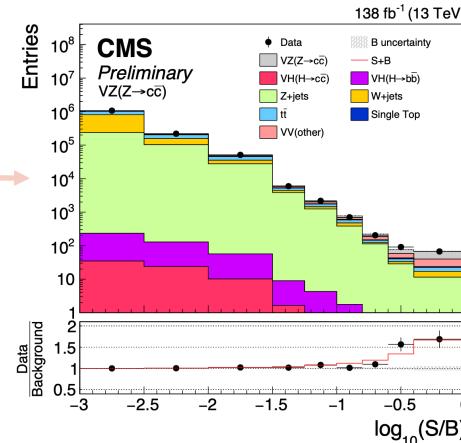
# **Summary & Outlook**

# A charming journey



$\text{From } \mathcal{O}(1000) \text{ to } \mathcal{O}(100) \text{ to } \mathcal{O}(10) \text{ in } \sim 5 \text{ years.}$

A combined effort and creativity from instrumentation,  
physics objects and analysis techniques!



First observation of  $Z \rightarrow cc$  at a hadron collider!  
Opening a new era for future explorations.

- More channels:  $t\bar{t}H(cc)$ , VBF  $H(cc)$ , indirect constraints, etc.
- Improvements in advanced analysis techniques (e.g., Deep Learning) and instrumentation (e.g., tracker)
- Reduction of systematic uncertainties:  $c$ -tagging, event modeling, theoretical uncertainties, ...

A charming journey ahead!

# **Backups**

# $H \rightarrow cc$ searches at the LHC

## □ ATLAS:

- [[Phys. Rev. Lett. 120 \(2018\) 211802](#)] ( $36 \text{ fb}^{-1}$ )
- [[arXiv:2201.11428](#)] ( $139 \text{ fb}^{-1}$ )
- [[ATL-PHYS-PUB-2021-039](#)] (HL-LHC projection,  $3000 \text{ fb}^{-1}$ )

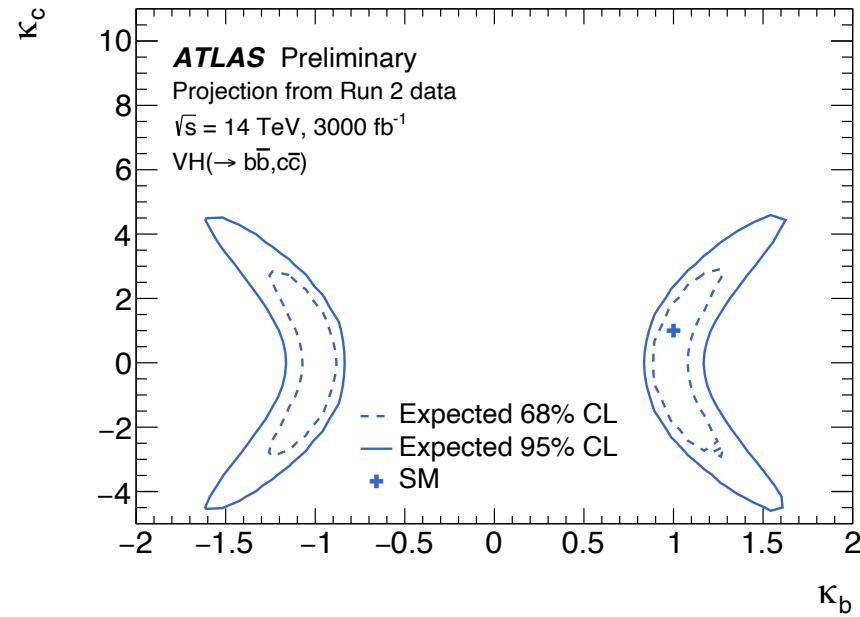
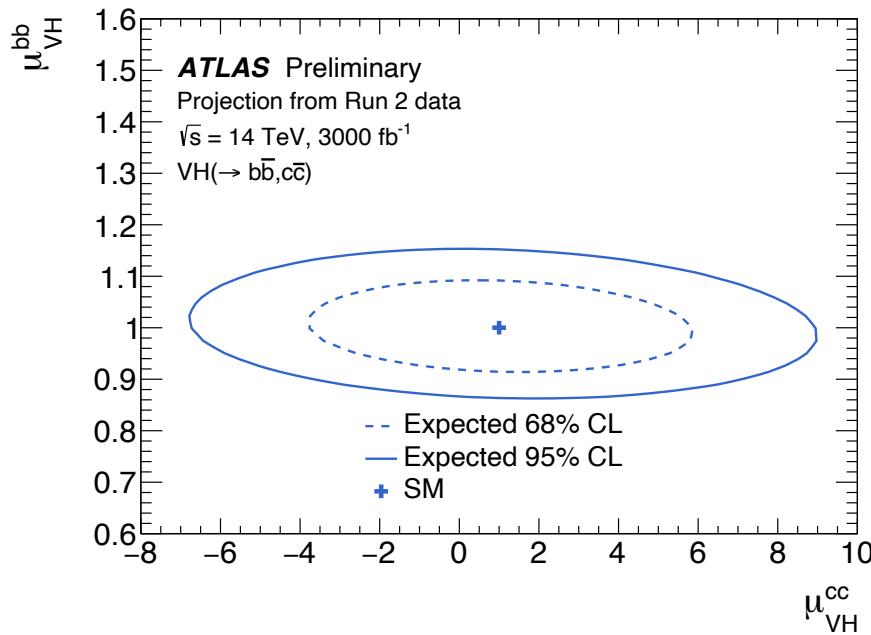
## □ CMS:

- [[JHEP 03 \(2020\) 131](#)] ( $36 \text{ fb}^{-1}$ )
- [[CMS-PAS-HIG-21-008](#)] ( $138 \text{ fb}^{-1}$ ; HL-LHC projection,  $3000 \text{ fb}^{-1}$ )

## □ LHCb:

- [[LHCb-CONF-2016-006](#)] ( $1.98 \text{ fb}^{-1}$ )
- [[LHCb-PUB-2018-009](#)] (HL-LHC projection,  $300 \text{ fb}^{-1}$ )

# ATLAS HL-LHC projection for $H \rightarrow cc$



# Baseline event selections

Merged-jet topology

| Variable                                    | 0L        | 1L               | 2L             |
|---|-----------|------------------|----------------|
| $p_T^\ell$                                  | —         | ( $>25, >30$ )   | $>20$          |
| Lepton isolation                            | —         | ( $<0.06, —$ )   | ( $<0.25, —$ ) |
| $N_{a\ell}$                                 | =0        | =0               | —              |
| $M(\ell\ell)$                               | —         | —                | 75–105         |
| $N_{\text{small-}R}^{\text{aj}}$            | <2        | <2               | <3             |
| $p_T^{\text{miss}}$                         | $>200$    | $>60$            | —              |
| $p_T(V)$                                    | $>200$    | $>150$           | $>150$         |
| $p_T(H_{\text{cand}})$                      | $>300$    | $>300$           | $>300$         |
| $m(H_{\text{cand}})$                        | 50–200    | 50–200           | 50–200         |
| $\Delta\phi(V, H_{\text{cand}})$            | $>2.5$    | $>2.5$           | $>2.5$         |
| $\Delta\phi(\vec{p}_T^{\text{miss}}, j)$    | $>0.5$    | —                | —              |
| $\Delta\phi(\vec{p}_T^{\text{miss}}, \ell)$ | —         | $<1.5$           | —              |
| Kinematic BDT                               | $>0.55$   | 0.55–0.7, $>0.7$ | $>0.55$        |
| c $\bar{c}$ discriminant                    |           |                  |                |
| High purity                                 | $>0.99$   | $>0.99$          | $>0.99$        |
| Medium purity                               | 0.96–0.99 | 0.96–0.99        | 0.96–0.99      |
| Low purity                                  | 0.90–0.96 | 0.90–0.96        | 0.90–0.96      |

Resolved-jet topology

| Variable                                    | 0L       | 1L             | 2L low- $p_T(V)$ | 2L high- $p_T(V)$ |
|---|----------|----------------|------------------|-------------------|
| $p_T^\ell$                                  | —        | ( $>25, >30$ ) | $>20$            | $>20$             |
| Lepton isolation                            | —        | ( $<0.06, —$ ) | ( $<0.25, —$ )   | ( $<0.25, —$ )    |
| $N_{a\ell}$                                 | =0       | =0             | —                | —                 |
| $M(\ell\ell)$                               | —        | —              | 75–105           | 75–105            |
| $p_T(j_1)$                                  | $>60$    | $>25$          | $>20$            | $>20$             |
| $p_T(j_2)$                                  | $>35$    | $>25$          | $>20$            | $>20$             |
| $CvsL(j_1)$                                 | $>0.225$ | $>0.225$       | $>0.225$         | $>0.225$          |
| $CvsB(j_2)$                                 | $>0.4$   | $>0.4$         | $>0.4$           | $>0.4$            |
| $N_{\text{small-}R}^{\text{aj}}$            | —        | <2             | —                | —                 |
| $p_T^{\text{miss}}$                         | $> 170$  | —              | —                | —                 |
| $p_T^{\text{miss}}$ significance            | —        | >4             | —                | —                 |
| $p_T(V)$                                    | $>170$   | $>100$         | 60–150           | $>150$            |
| $p_T(H_{\text{cand}})$                      | $>120$   | $>100$         | —                | —                 |
| $m(H_{\text{cand}})$                        | $<250$   | $<250$         | $<250$           | $<250$            |
| $\Delta\phi(V, H_{\text{cand}})$            | $>2.0$   | $>2.5$         | $>2.5$           | $>2.5$            |
| $\Delta\phi(\vec{p}_T^{\text{miss}}, j)$    | $>0.5$   | —              | —                | —                 |
| $\Delta\phi(\vec{p}_T^{\text{miss}}, \ell)$ | —        | $<2.0$         | —                | —                 |

# Uncertainties

## □ Breakdown of the uncertainties in each topology

### Merged-jet topology

Table 3: The relative contributions to the total uncertainty on  $\mu_{VH(H \rightarrow c\bar{c})}$  in the merged-jet analysis, with a best fit value  $\mu_{VH(H \rightarrow c\bar{c})} = 8.7^{+4.6}_{-4.0}$ .

| Uncertainty source                 | $\Delta\mu / (\Delta\mu)_{\text{tot}}$ |
|------------------------------------|--|
| <b>Statistical</b>                 | <b>88%</b>                             |
| Background normalizations          | 39%                                    |
| <b>Experimental</b>                | <b>40%</b>                             |
| Sizes of the simulated samples     | 24%                                    |
| Charm identification efficiencies  | 26%                                    |
| Jet energy scale and resolution    | 15%                                    |
| Simulation modeling                | 1%                                     |
| Luminosity                         | 5%                                     |
| Lepton identification efficiencies | 2%                                     |
| <b>Theory</b>                      | <b>25%</b>                             |
| Backgrounds                        | 21%                                    |
| Signal                             | 14%                                    |

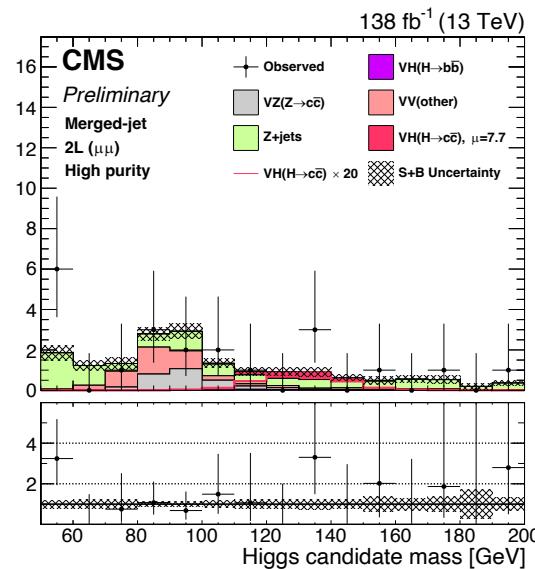
### Resolved-jet topology

Table 4: The relative contributions to the total uncertainty on  $\mu_{VH(H \rightarrow c\bar{c})}$  in the resolved-jet analysis, with a best fit value  $\mu_{VH(H \rightarrow c\bar{c})} = -9.5 \pm 9.6$ .

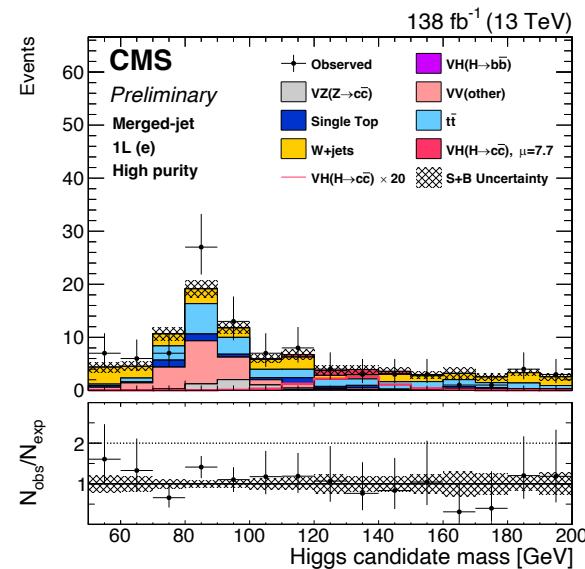
| Uncertainty source                 | $\Delta\mu / (\Delta\mu)_{\text{tot}}$ |
|------------------------------------|--|
| <b>Statistical</b>                 | <b>66%</b>                             |
| Background normalizations          | 28%                                    |
| <b>Experimental</b>                | <b>72%</b>                             |
| Sizes of the simulated samples     | 59%                                    |
| Charm identification efficiencies  | 27%                                    |
| Jet energy scale and resolution    | 17%                                    |
| Simulation modeling                | 20%                                    |
| Luminosity                         | 13%                                    |
| Lepton identification efficiencies | 10%                                    |
| <b>Theory</b>                      | <b>22%</b>                             |
| Backgrounds                        | 21%                                    |
| Signal                             | 7%                                     |

# Merged-jet topology: signal regions

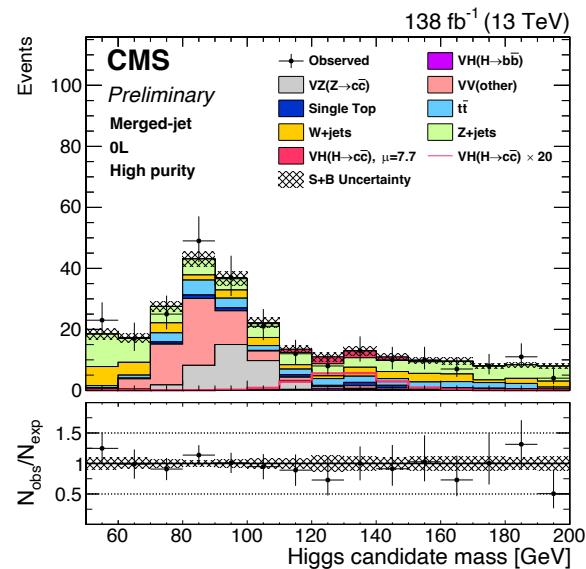
2L( $\mu\mu$ ), high cc-purity



1L(e), high cc-purity

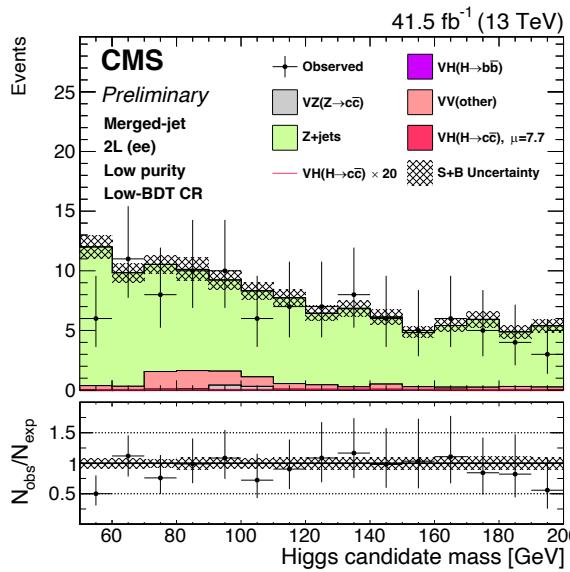


0L, high cc-purity

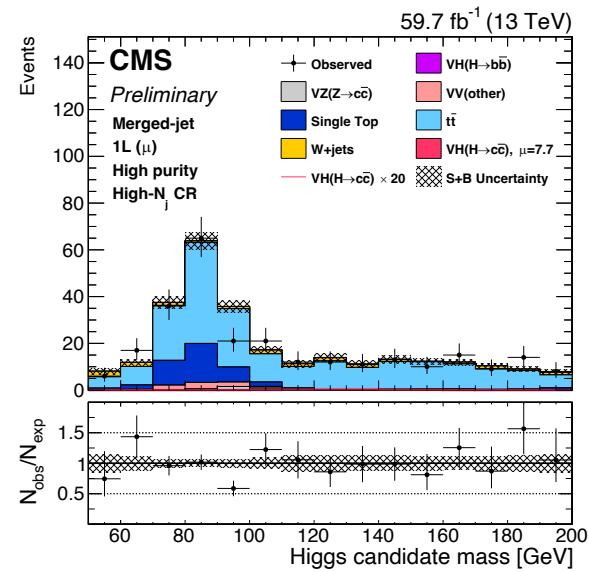


# Merged-jet topology: control regions

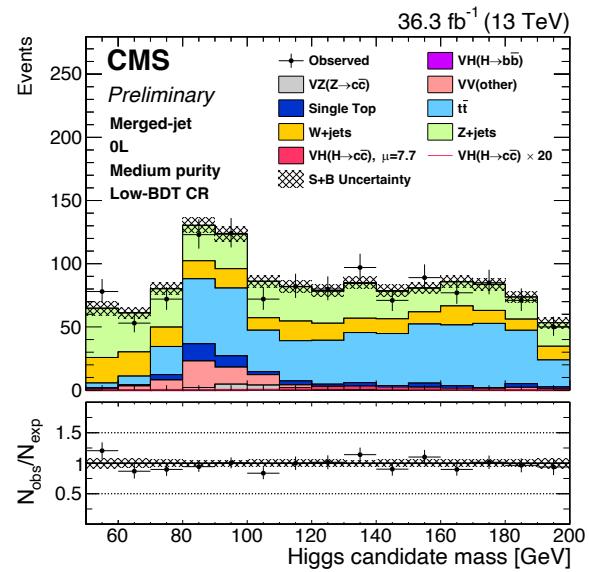
2L(ee), V+jets CR, low cc-purity



1L( $\mu$ ), tt CR, high cc-purity

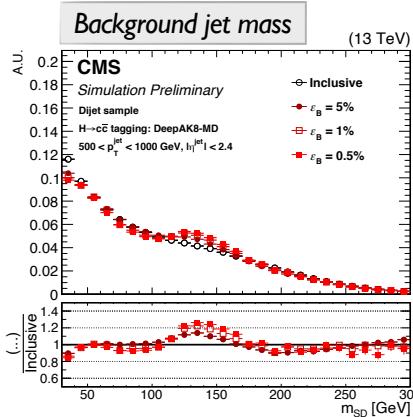
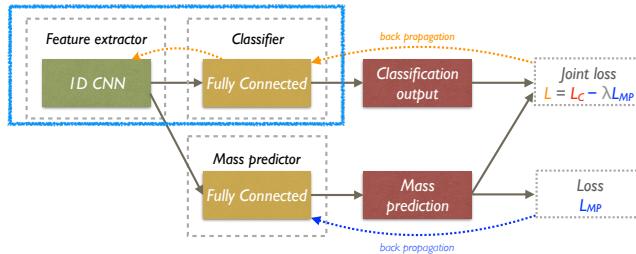


0L, V+jets CR, medium cc-purity

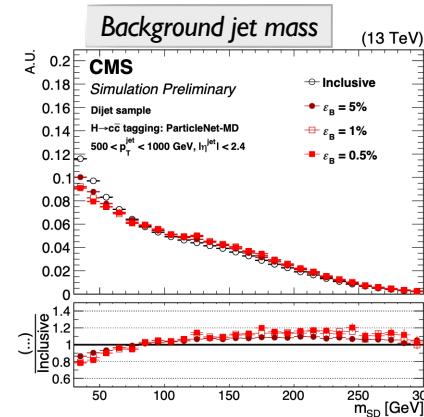
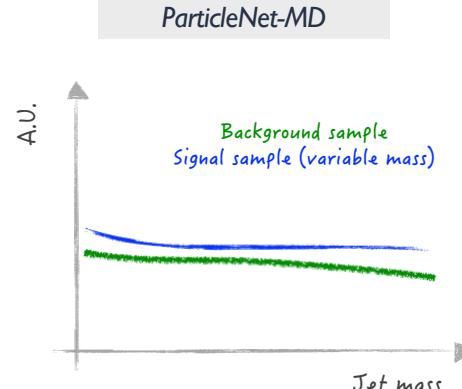


# Comparison of mass decorrelation methods

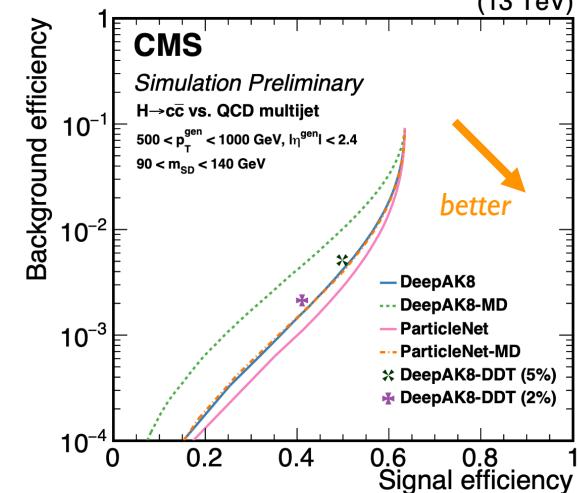
DeepAK8-MD



ParticleNet-MD



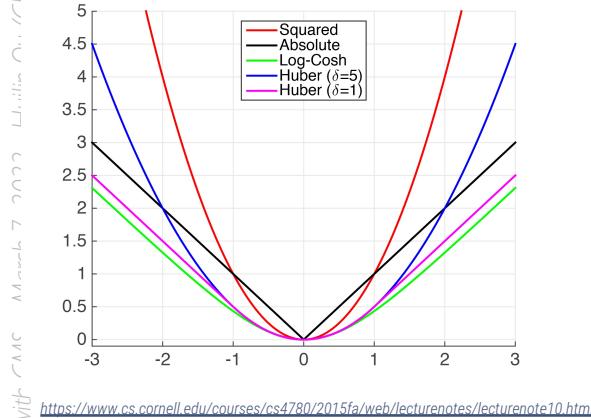
$H \rightarrow cc$  tagging performance



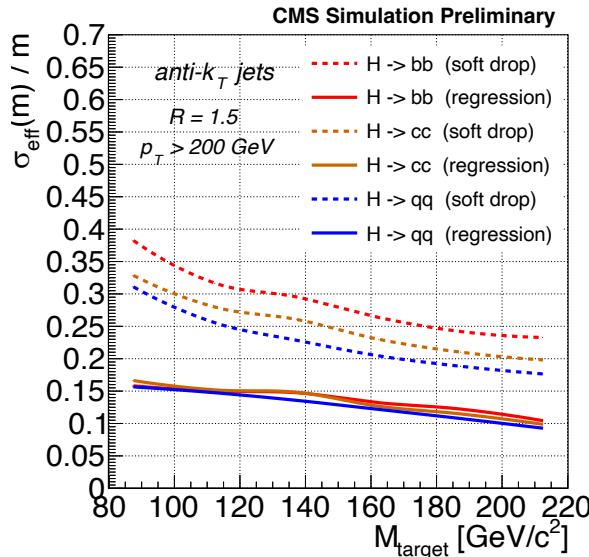
# Large-R jet mass regression

Loss function: LogCosh

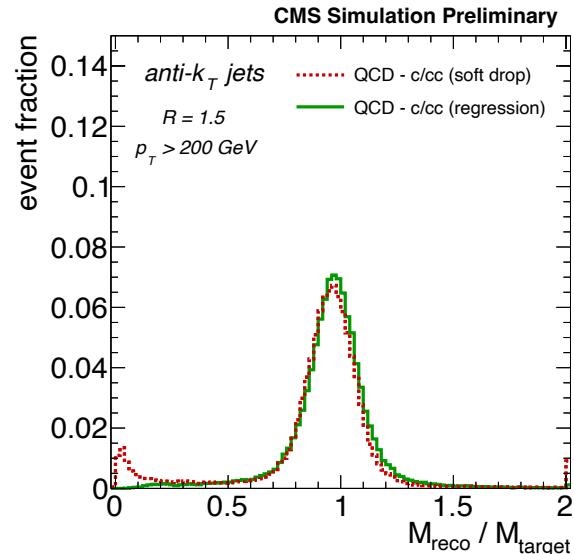
$$L(y, y^p) = \sum_{i=1}^n \log(\cosh(y_i^p - y_i))$$



Signal jet mass resolution

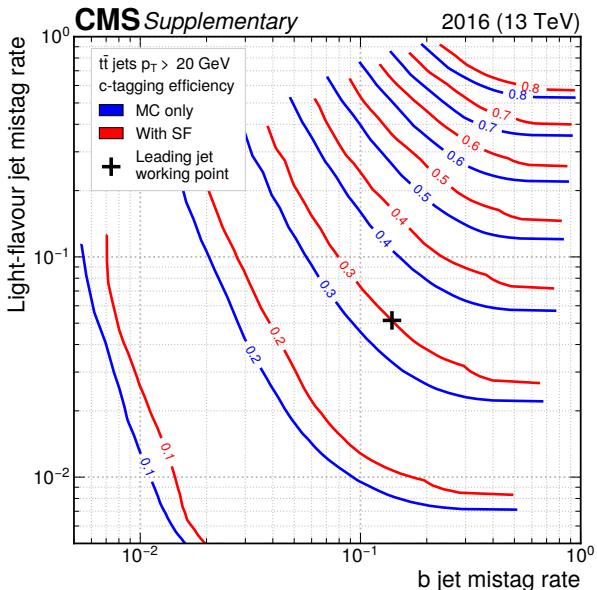


Background jet mass response

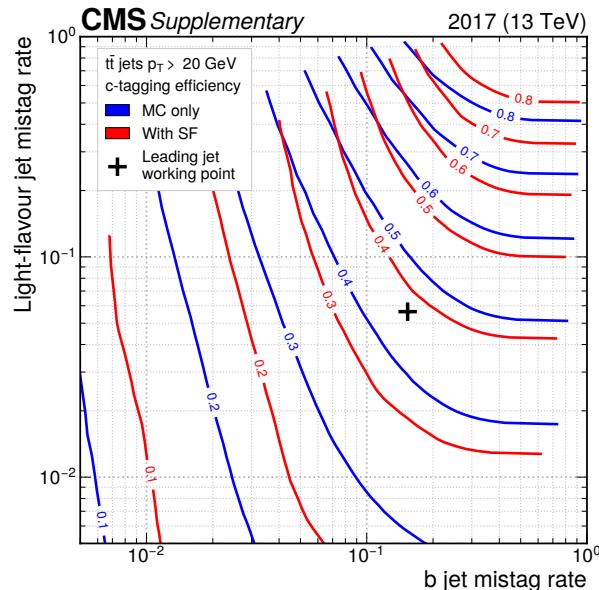


# C-tagger ROC curves

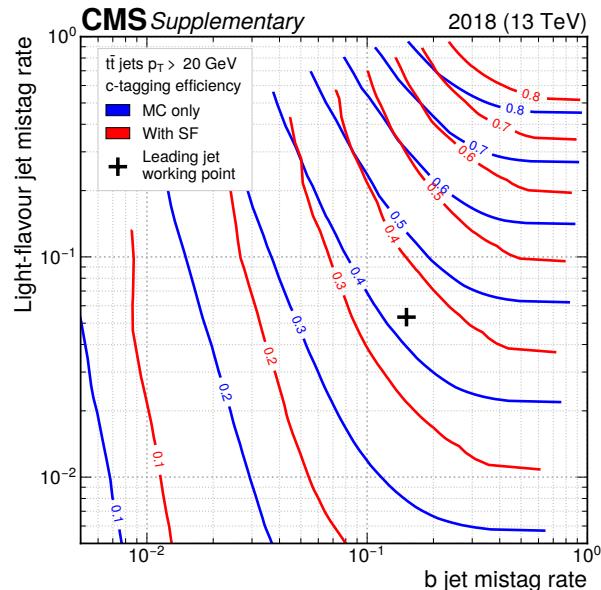
2016



2017



2018



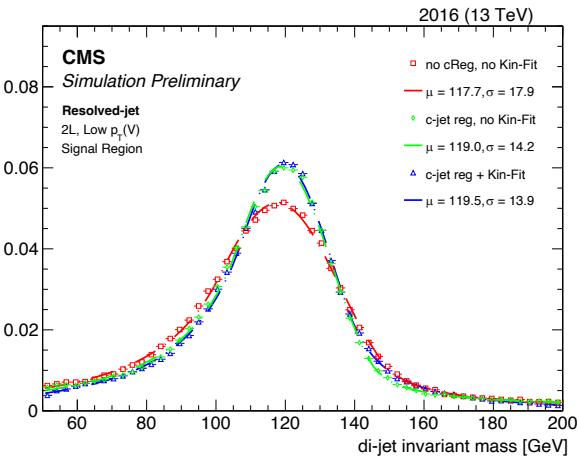
- CMS c-tagging WP: ~40% (c), ~16% (b), ~4% (light)
- ATLAS c-tagging WP [arXiv:2201.11428]: 27% (c), 8% (b), 1.6% (light)

# C-jet energy regression and kinematic fit

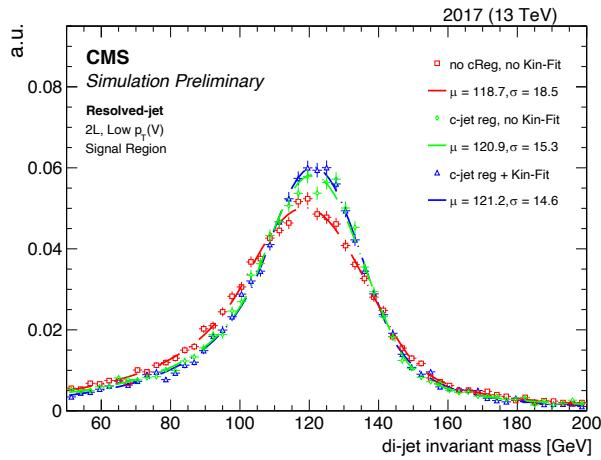
- 2-lepton Low- $p_T(V)$  category –  $60 \text{ GeV} < p_T(V) < 150 \text{ GeV}$

(CERN)  
Search for  $f$

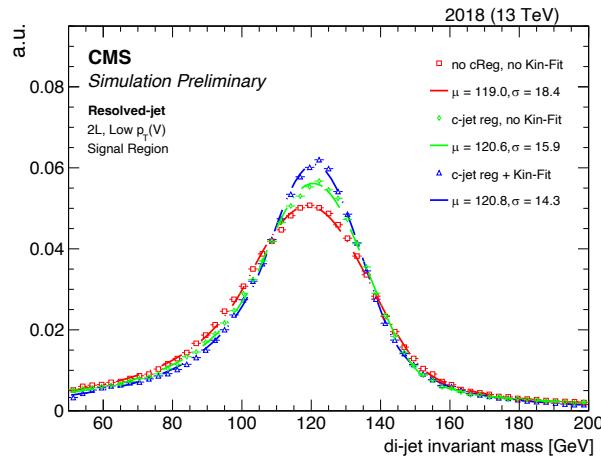
2016



2017



2018

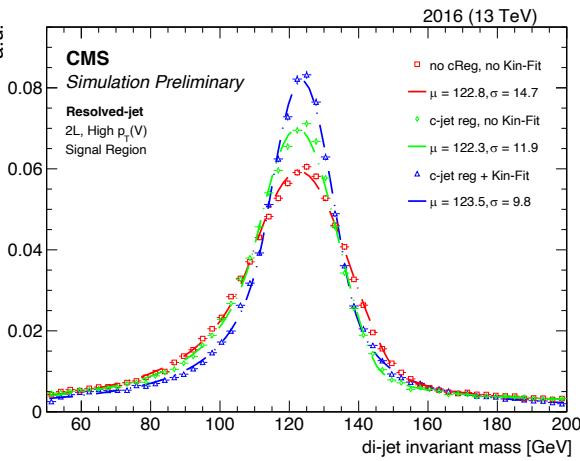


# C-jet energy regression and kinematic fit

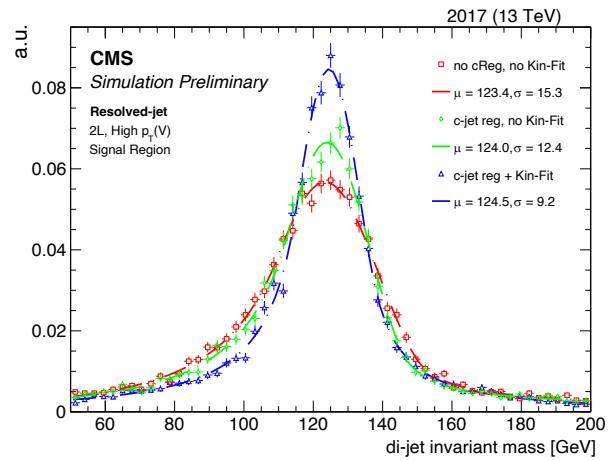
- 2-lepton High- $p_T(V)$  category –  $p_T(V) > 150$  GeV

Search for  $f$

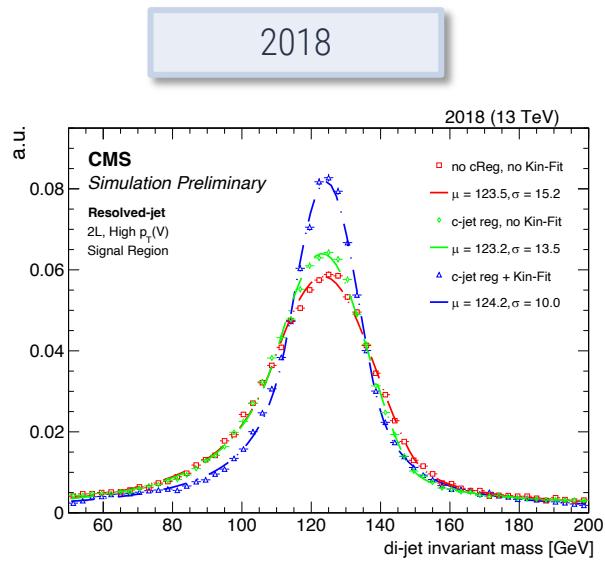
2016



2017



2018



# A new method to calibrate charm-taggers

## DeepJet algorithm calibration

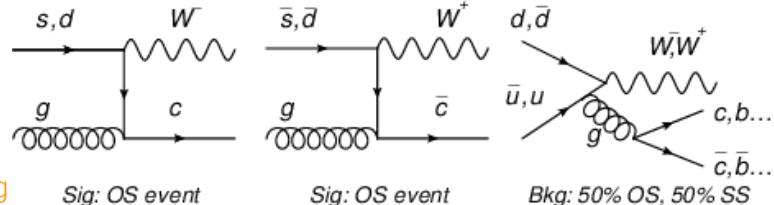
(from arXiv:2111.03027)

### Methodology

- Iterative approach exploiting three distinct control regions that are enriched with either b-jets, c-jets, or light-flavour and gluon jets
- First time that a **calibration** method to correct the 2D distribution of **c-tagging discriminator shapes** is presented ➔ [arXiv:2111.03027](https://arxiv.org/abs/2111.03027) (accepted by JINST)

### Search for an abundant and pure source of charm-jets

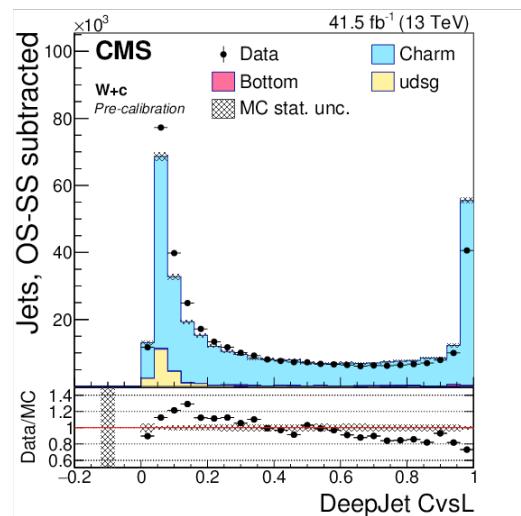
- Target W production in association with charm quarks
  - The relevant events involve a leptonically decaying W boson and a c-jet
  - These c-jets are identified using the semileptonic decay of the charmed hadrons, which produces a soft muon within the jet
- Major background has 50% chance to have SS or OS final states ➔ performing an OS-SS subtraction reduces considerably the W+gluon process
- To enrich in b-jets and light-jets, the semi-(di-)leptonic  $t\bar{t}$ +jets and DY( $Z \rightarrow \mu\mu/\text{ee}$ )+jets processes are considered



Sig: OS event

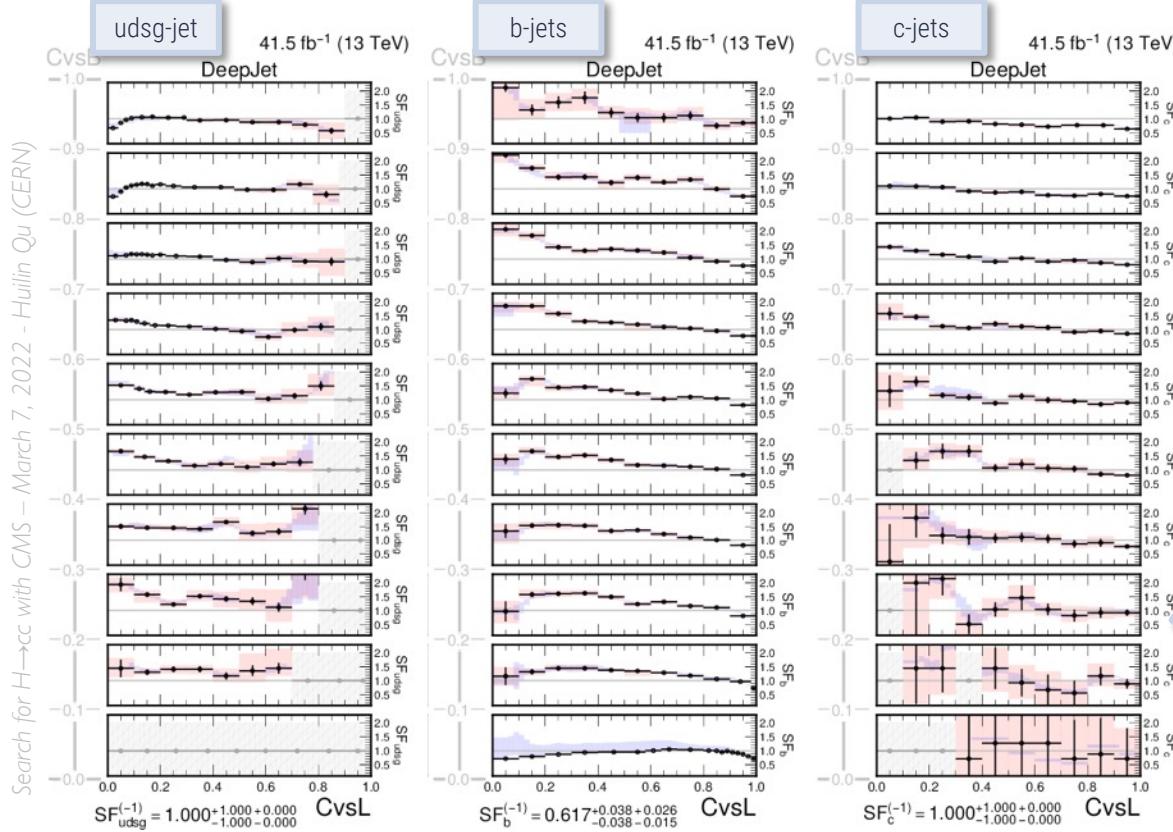
Sig: OS event

Bkg: 50% OS, 50% SS



# A new method to calibrate charm-taggers

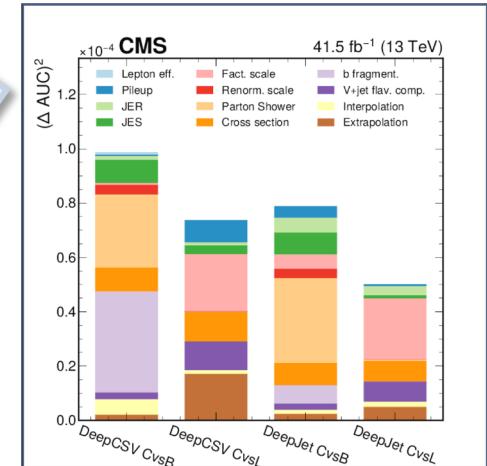
## Extraction of reshaping data-to-simulation scale factors



□ SFs as a function of  $C_{\text{vsL}}$  in bins of  $C_{\text{vsB}}$

- Fixed bin width along  $C_{\text{vsB}}$  and an adaptive binning scheme along  $C_{\text{vsL}}$  (stat. depending)
- Total uncertainties (red envelopes) relatively small in the region of interest of the analysis
- Total uncertainties breakdown
  - Overall smaller than DeepCSV

(from arXiv:2111.03027)



# A new method to calibrate charm-taggers

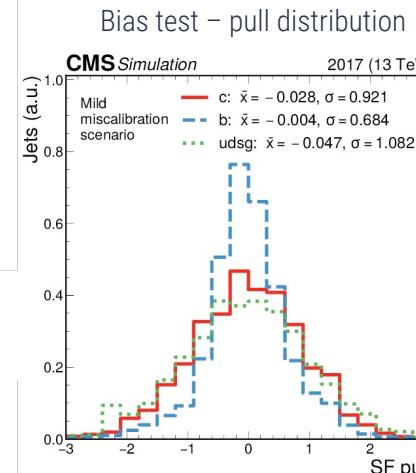
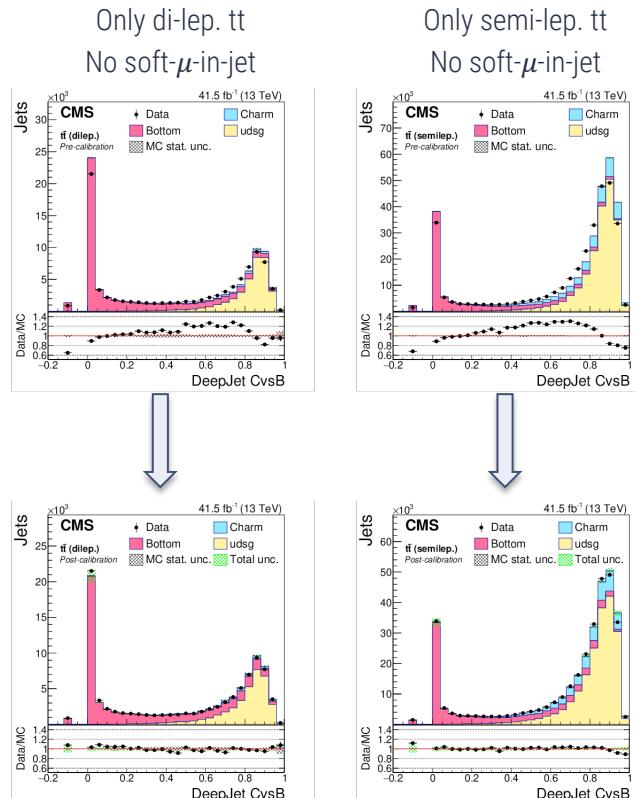
## Validate robustness of the SFs derivation

(from arXiv:2111.03027)

- ❑ Check possible bias due to the soft- $\mu$ -in-jet selection
  - SFs are derived without soft- $\mu$  selection
- ❑ Check possible bias between semileptonic or di-leptonic tt final states
  - SFs are derived also for the two separate processes independently
- ❑ Check possible bias due in the fit:
  - Inject *artificial* SFs to calculate the pulls between the fit result and the injected one



All the checks shown no bias in the SFs derivation



# Signal extraction – BDT training in SRs

| Variable                                   | Description   | 0L | 1L | 2L |
|--|---|----|----|----|
| $m(H)$                                     | H mass  | ✓  | ✓  | ✓  |
| $p_T(H)$                                   | H transverse momentum   | —  | ✓  | ✓  |
| $p_T(V)$                                   | vector boson transverse momentum  | —  | ✓  | ✓  |
| $m_T(V)$                                   | vector boson transverse mass  | —  | ✓  | —  |
| $p_T^{\text{miss}}$                        | missing transverse momentum   | ✓  | ✓  | —  |
| $p_T(V)/p_T(H)$                            | ratio between vector boson and H transverse momenta   | ✓  | ✓  | ✓  |
| $Cvsl_{\max}$                              | $Cvsl$ value of the leading $Cvsl$ jet  | ✓  | ✓  | ✓  |
| $CvsB_{\max}$                              | $CvsB$ value of the leading $Cvsl$ jet  | ✓  | ✓  | ✓  |
| $Cvsl_{\min}$                              | $Cvsl$ value of the subleading $Cvsl$ jet   | ✓  | ✓  | ✓  |
| $CvsB_{\min}$                              | $CvsB$ value of the subleading $Cvsl$ jet   | ✓  | ✓  | ✓  |
| $p_{T\max}$                                | $p_T$ of the leading $Cvsl$ jet   | ✓  | ✓  | ✓  |
| $p_{T\min}$                                | $p_T$ of the subleading $Cvsl$ jet  | ✓  | ✓  | ✓  |
| $\Delta\phi(V, H)$                         | azimuthal angle between vector boson and H  | ✓  | ✓  | ✓  |
| $\Delta R(j_1, j_2)$                       | $\Delta R$ between leading and subleading $Cvsl$ jets   | —  | ✓  | ✓  |
| $\Delta\phi(j_1, j_2)$                     | azimuthal angle between leading and subleading $Cvsl$ jets                                      | ✓  | ✓  | —  |
| $\Delta\eta(j_1, j_2)$                     | difference in pseudorapidity between leading and subleading $Cvsl$ jets                         | ✓  | ✓  | ✓  |
| $\Delta\phi(\ell_1, \ell_2)$               | azimuthal angle between leading and subleading $p_T$ leptons                                    | —  | —  | ✓  |
| $\Delta\eta(\ell_1, \ell_2)$               | difference in pseudorapidity between leading and subleading $p_T$ leptons                       | —  | —  | ✓  |
| $\Delta\phi(\ell_1, j_1)$                  | azimuthal angle between leading $p_T$ lepton and leading $Cvsl$ jet                             | —  | ✓  | —  |
| $\Delta\phi(\ell_2, j_1)$                  | azimuthal angle between subleading $p_T$ lepton and leading $Cvsl$ jet                          | —  | —  | ✓  |
| $\Delta\phi(\ell_2, j_2)$                  | azimuthal angle between subleading $p_T$ lepton and subleading $Cvsl$ jet                       | —  | —  | ✓  |
| $\Delta\phi(\ell_1, p_T^{\text{miss}})$    | azimuthal angle between leading $p_T$ lepton and missing transverse momentum                    | ✓  | —  | —  |
| $\Delta\eta(\ell_1, t)$                    | difference in pseudorapidity between leading $p_T$ lepton and b-tagged jet from top quark decay | —  | ✓  | —  |
| $\Delta\phi(\ell_1, t)$                    | azimuthal angle between leading $p_T$ lepton and b-tagged jet from top quark decay              | —  | ✓  | —  |
| $\Delta R(\ell_1, t)$                      | $\Delta R$ between leading $p_T$ lepton and b-tagged jet from top quark decay                   | —  | ✓  | —  |
| $Cvsl_t$                                   | $Cvsl$ value of the b-tagged jet from top quark decay   | —  | ✓  | —  |
| $CvsB_t$                                   | $CvsB$ value of the b-tagged jet from top quark decay   | —  | ✓  | —  |
| $P(b+bb)_t$                                | Deepjet prob( $b+bb$ ) value of the b-tagged jet from top quark decay                           | —  | ✓  | —  |
| $m(t)$                                     | Reconstructed top quark mass  | —  | ✓  | —  |
| $N_{\text{small-}R}^{ij}$                  | Number of small- $R$ additional jets after the FSR subtraction                                  | —  | ✓  | —  |
| $\sigma_{CReg}(j_1)$                       | leading $p_T$ jet resolution from c-jet energy regression                                       | ✓  | ✓  | ✓  |
| $\sigma_{CReg}(j_2)$                       | subleading $p_T$ jet resolution from c-jet energy regression                                    | ✓  | ✓  | ✓  |
| $\Delta\eta(V, H) \parallel \text{kinfit}$ | difference in pseudorapidity between vector boson and H, after kinematic-fit                    | —  | —  | ✓  |
| $\Delta\phi(V, H) \parallel \text{kinfit}$ | azimuthal angle between vector boson and H, after kinematic-fit                                 | —  | —  | ✓  |
| $m(H) \parallel \text{kinfit}$             | H mass after kinematic-fit  | —  | —  | ✓  |
| $p_T(H) \parallel \text{kinfit}$           | H transverse momentum after kinematic-fit   | —  | —  | ✓  |
| $p_{T\max} \parallel \text{kinfit}$        | $p_T$ of the leading $Cvsl$ jet after kinematic-fit   | —  | —  | ✓  |
| $p_{T\min} \parallel \text{kinfit}$        | $p_T$ of the subleading $Cvsl$ jet after kinematic-fit  | —  | —  | ✓  |
| $p_T(V)/p_T(H) \parallel \text{kinfit}$    | ratio between vector boson and H transverse momenta after kinematic-fit                         | —  | —  | ✓  |
| $\sigma(H) \parallel \text{kinfit}$        | H invariant mass resolution from kinematic fit  | —  | —  | ✓  |

Higgs and vector boson properties

c-tagging score

event kinematics

Kinfit Variables  
(2L only)

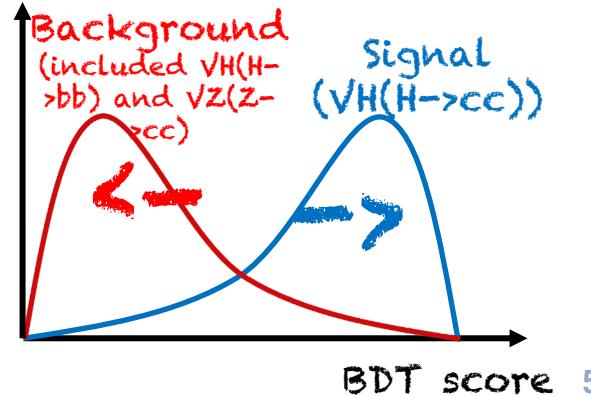
- BDT trained to separate signal from background samples

- Use combination of kinematic observables and particle flavor variables (tagger informations)

- Separate BDTs trained for each channel and data taking year

- Separate BDTs trained for high- and low- $p_T(V)$  2L
- Variables used dependent on channel

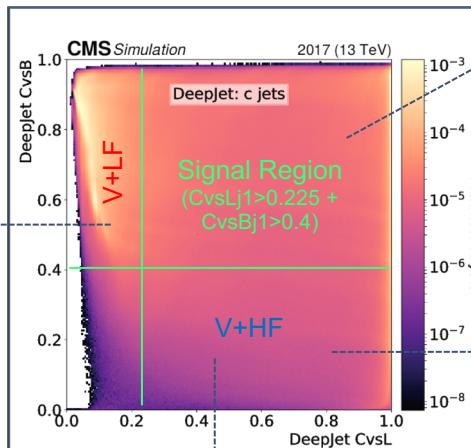
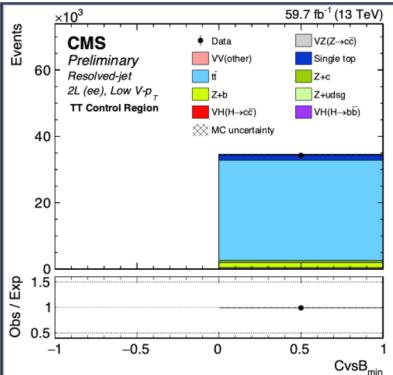
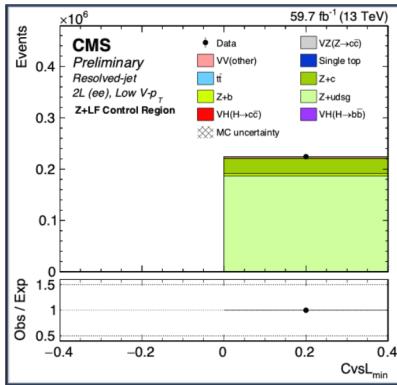
- Reshaped BDT distribution used in SR during final fit



# Background estimation – Resolved-jet

- Accurate modeling of jet flavor in V+Jet background is vital for proper signal extraction

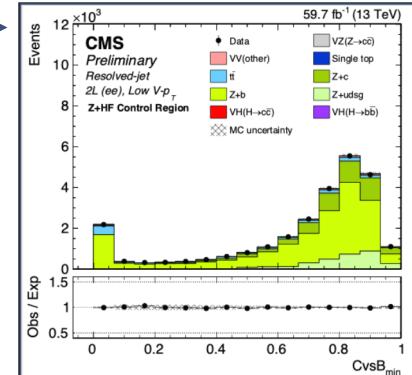
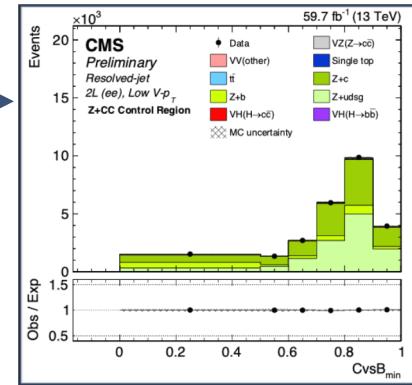
- Separate rate parameters for **V+c**, **V+b**, and **V+light** processes (no W+b)
- Additional rate parameter for  $t\bar{t}$  background



$t\bar{t}$   
Invert  $Z$  mass (2L)  
Require add jet (1L)\*  
Require add  $\ell$  and jets (0L)

\*1L: also require MET<170 GeV to keep orthogonal to 0L  $t\bar{t}$  CR

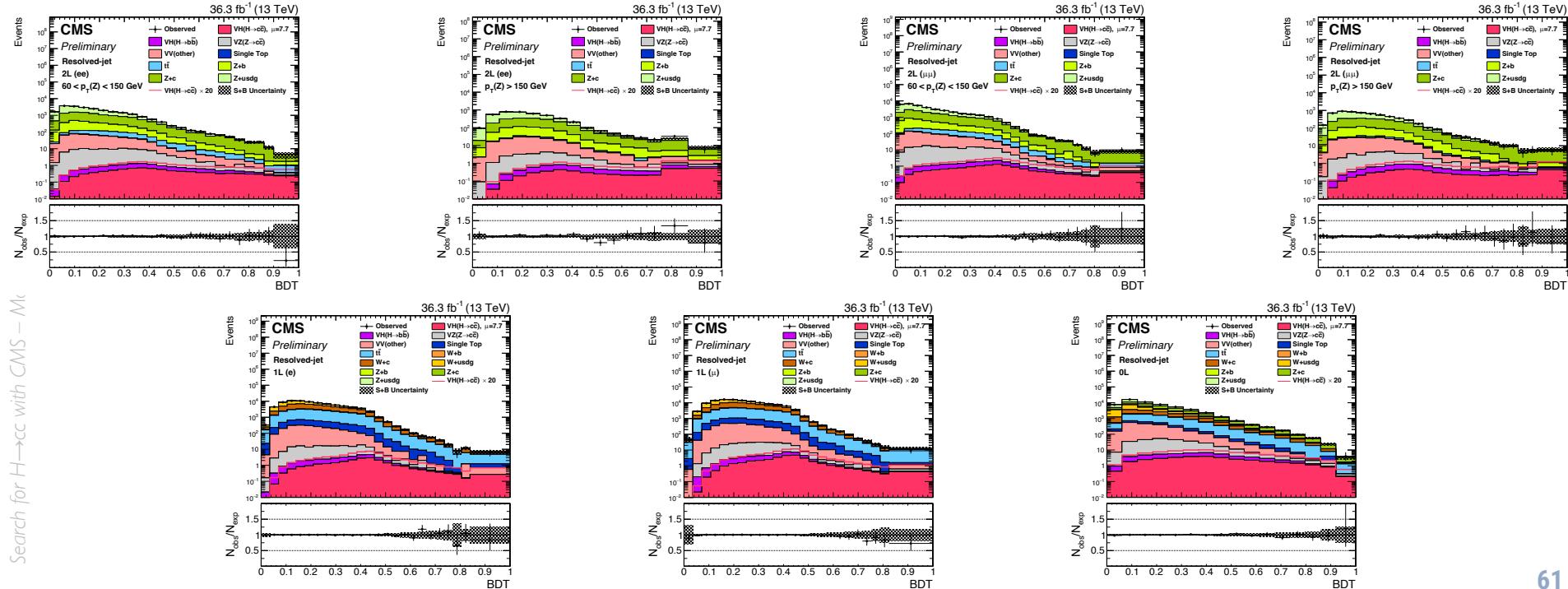
V+CC  
Veto m(H)  
region



# Postfit plots – Signal regions - 2016

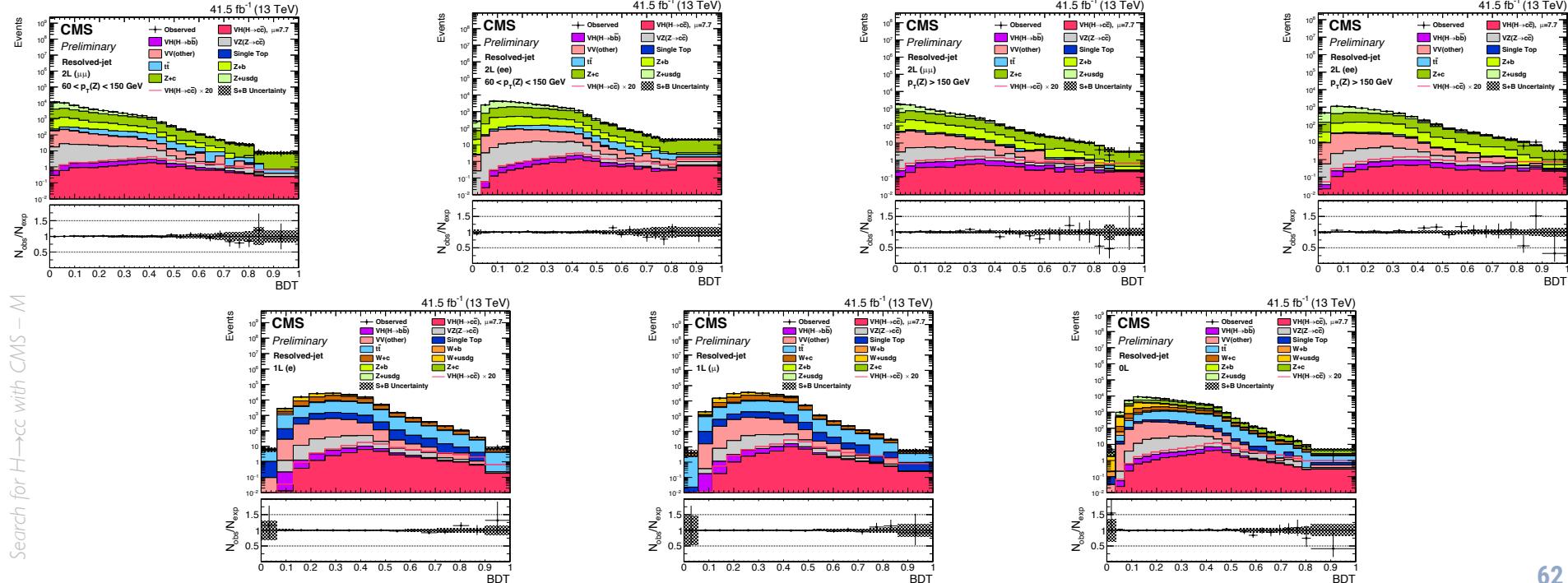
□ Postfit distribution of the BDT discriminant obtained with the 2016 data

■ 7 Signal regions in each year: 2L( $e/\mu$ ) Low- $p_T(V)$  and –High- $p_T(V)$ , 1L( $e/\mu$ ) and 0L



# Postfit plots – Signal regions - 2017

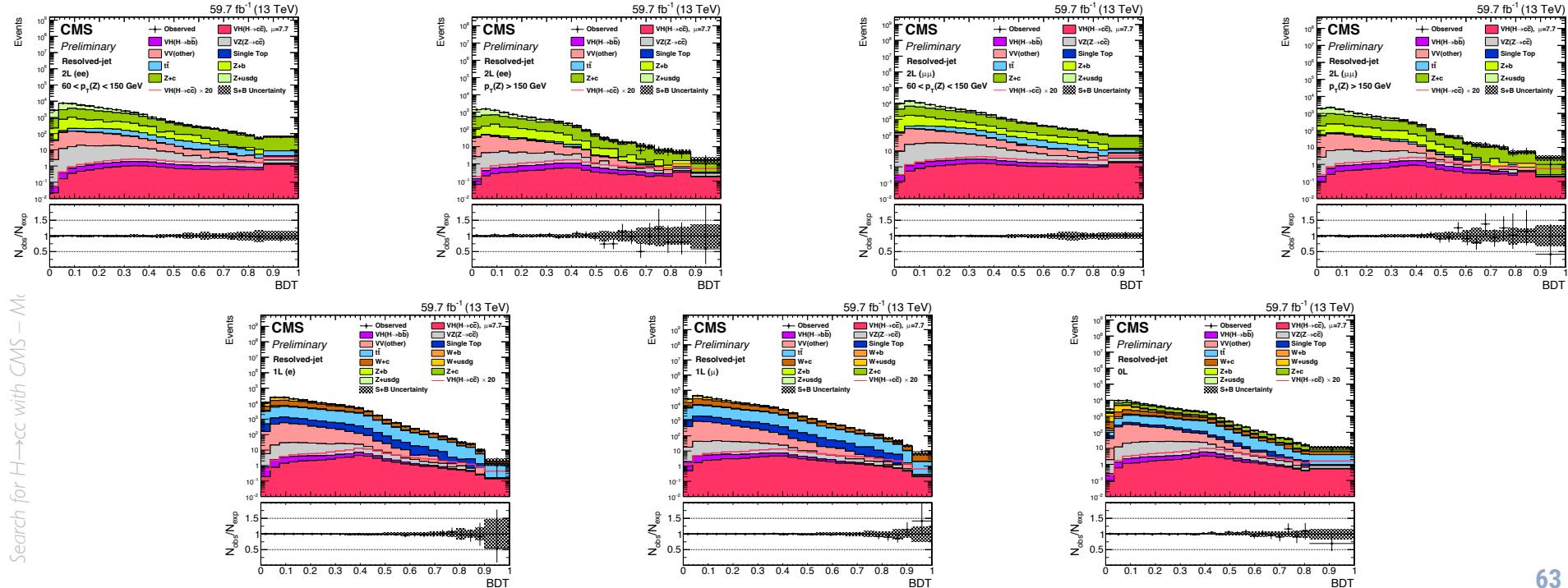
- Postfit distribution of the BDT discriminant obtained with the 2017 data
  - 7 Signal regions in each year: 2L( $e/\mu$ ) Low- $p_T(V)$  and –High- $p_T(V)$ , 1L( $e/\mu$ ) and 0L



# Postfit plots – Signal regions - 2018

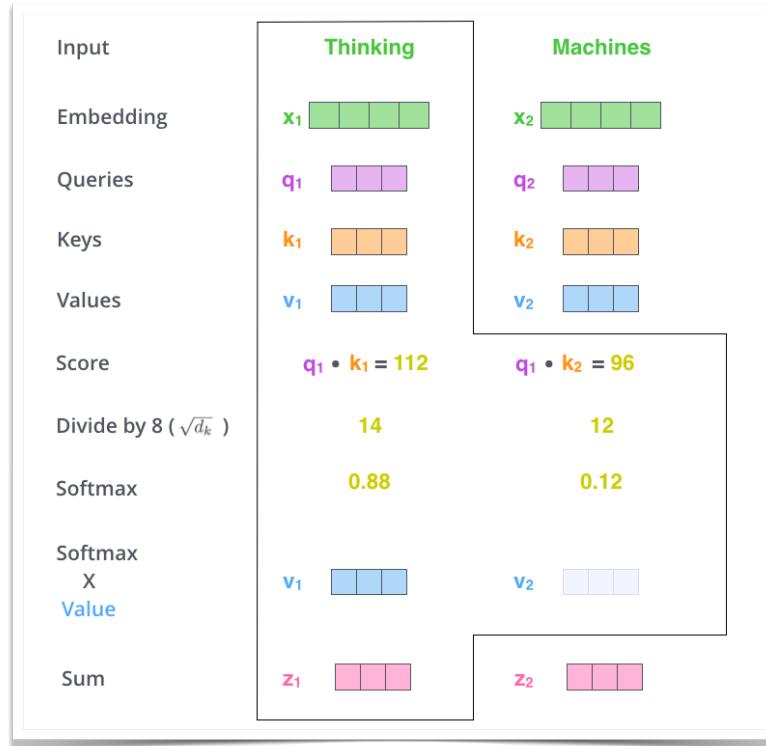
□ Postfit distribution of the BDT discriminant obtained with the 2018 data

■ 7 Signal regions in each year: 2L( $e\mu/\mu\mu$ ) Low- $p_T(V)$  and –High- $p_T(V)$ , 1L( $e/\mu\mu$ ) and 0L



# Transformer 101

Vaswani, Shazeer, Parmar, Uszkoreit, Jones, Gomez, Kaiser, Polosukhin, arXiv:1706.03762



$$\begin{aligned}
 X &\times W^Q = Q \\
 X &\times W^K = K \\
 X &\times W^V = V \\
 \text{softmax}\left(\frac{Q \times K^T}{\sqrt{d_k}}\right) V &= Z
 \end{aligned}$$