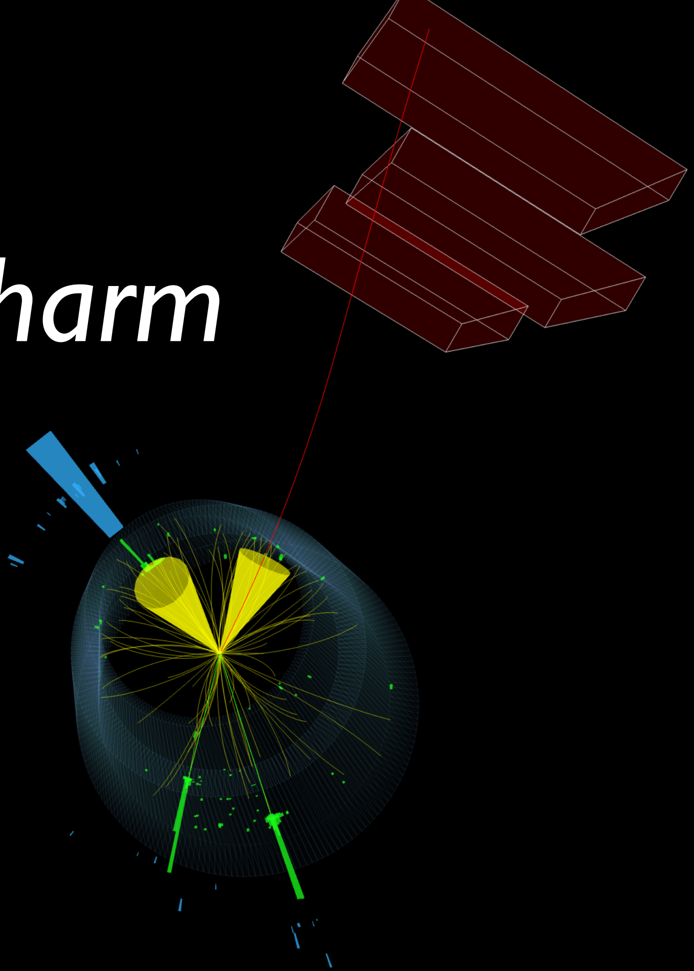
Constraining the Higgs-charm coupling at CMS

Luca Mastrolorenzo[1] and Huilin Qu[2]
on behalf of the CMS Collaboration

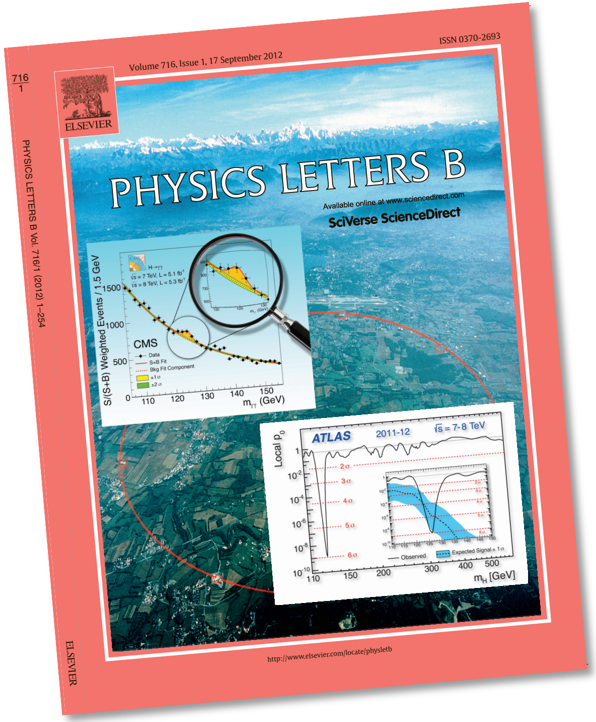
CERN LPCC EP-LHC Seminar
1 March 2022



[1] Rheinisch-Westfälische Technische Hochschule (RWTH), Aachen University
[2] CERN (European Organization for Nuclear Research)

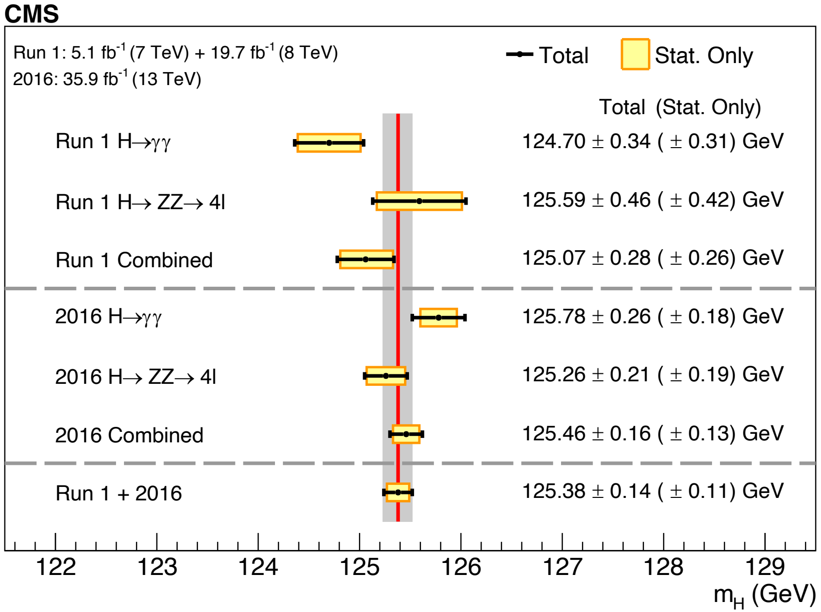
Introduction

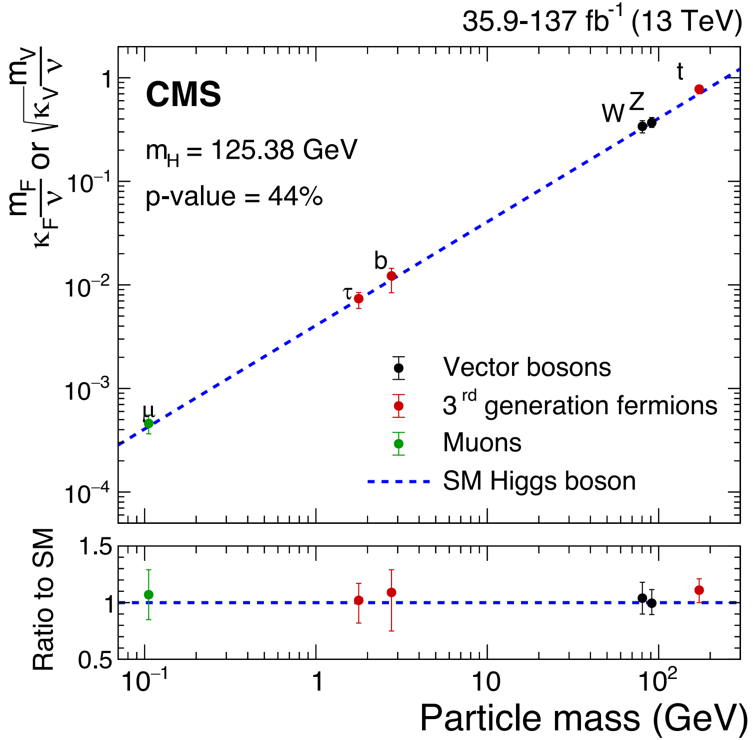
- Discovery of the Higgs boson in 2012: A new chapter of particle physics



Understanding the Higgs boson

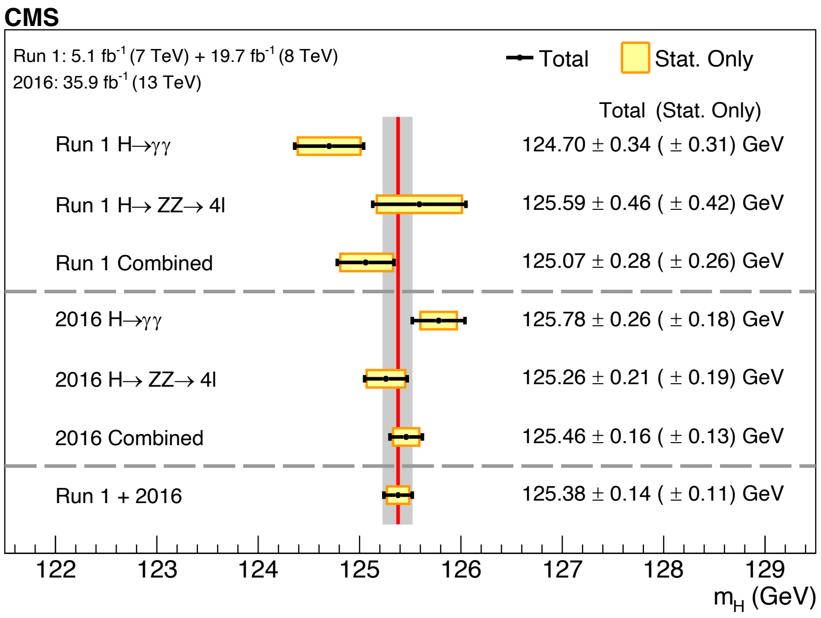
- Tremendous progress in our understanding of the Higgs boson in the past ten years

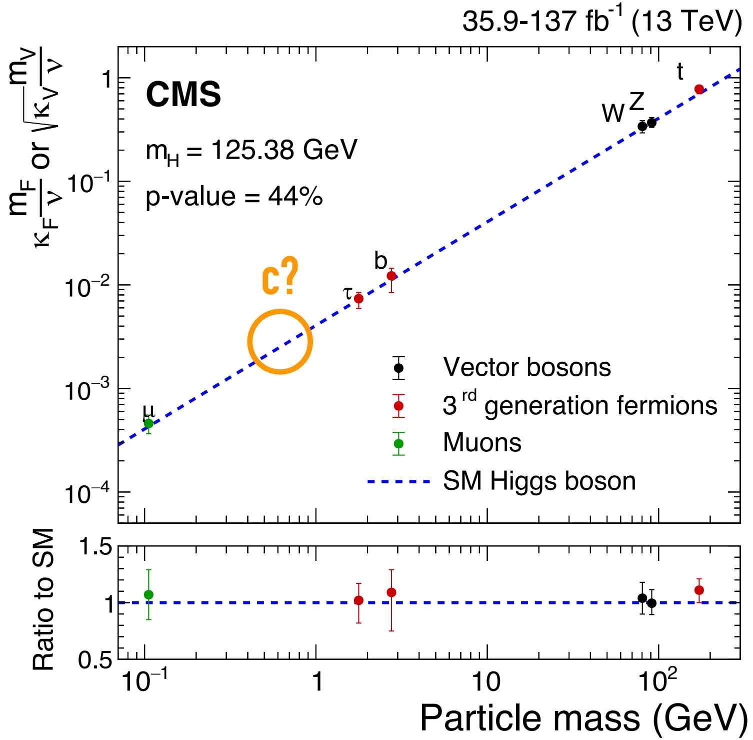




How charming is the Higgs boson?

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



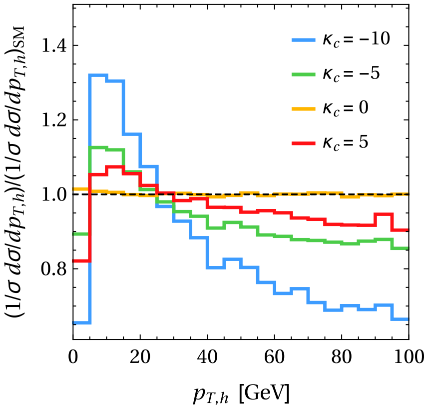


Probing the Higgs-charm coupling

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

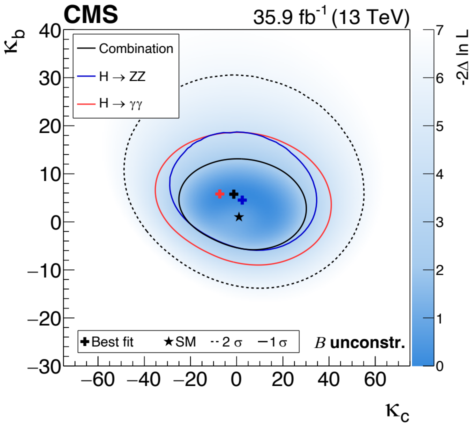
Indirect constraint from Higgs kinematics

Phys. Rev. Lett. 118, 121801



Variation of $p_T(H)$ shape as a function $\kappa_c = y_c/y_c^{SM}$

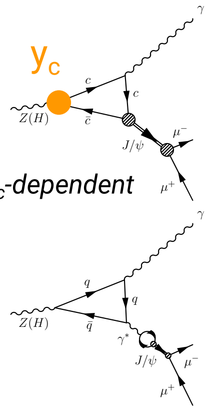
Phys. Lett. B 792 (2019) 369



$-33 < \kappa_c < 38$ (obs.)

$-31 < \kappa_c < 36$ (exp.)

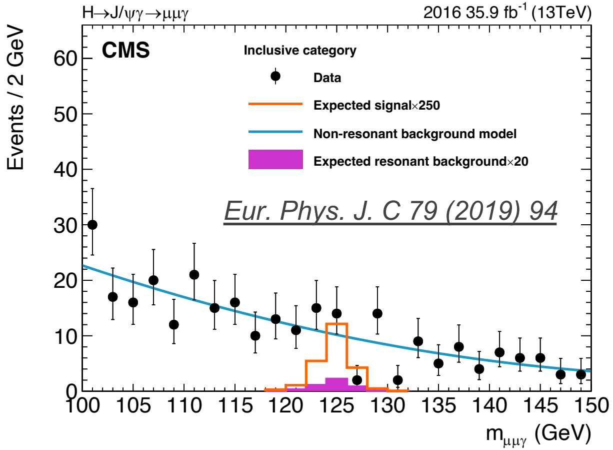
Search for exclusive H → J/Ψγ decays



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



Eur. Phys. J. C 79 (2019) 94

$\mathcal{B}$(H → J/Ψγ) < 220x SM(obs.)

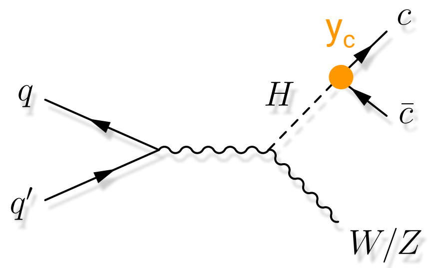
$\mathcal{B}$(H → J/Ψγ) < 170x SM(exp.)

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

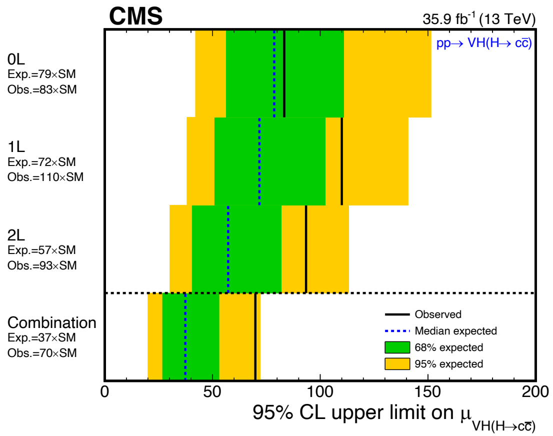
Corresponding ATLAS analyses: ATLAS-CONF-2022-002; Phys. Lett. B 786 (2018) 134.

Direct search for H → cc

- Search for $H \to cc$ decays: directly sensitive to y_c, but very challenging
  - small branching fraction (~3%) vs. large backgrounds (at a hadron collider)
  - charm quark identification is the key
- Exploit associated VH production (V = W, Z)
  - three channels: $Z \to \nu\nu$ (0L), $W \to \ell\nu$ (1L), $Z \to \ell\ell$ (2L) [ℓ = e, μ]
- Main backgrounds
  - V + jets, single and pair production of top quarks, dibosons
  - VH($H \to bb$): 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 tt contribution (0L & 1L)
- Previous result (36 fb^{-1}): [JHEP 03 (2020) 131]
- **Today: result with the full Run 2 data set (138 fb^{-1})**



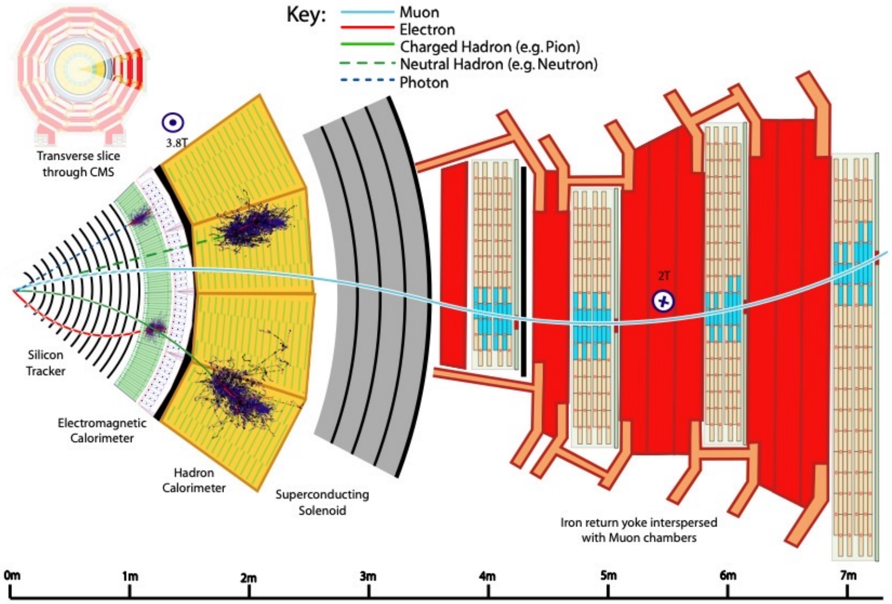
JHEP 03 (2020) 131



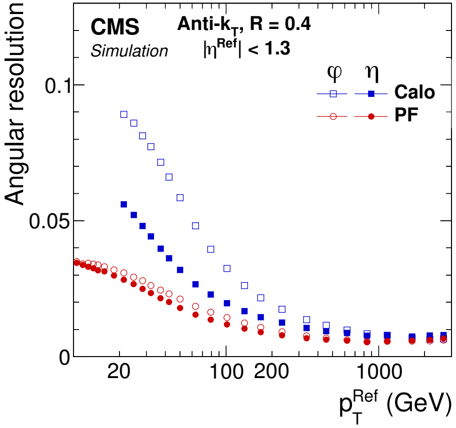
Corresponding ATLAS analysis: arXiv:2201.11428. See also recent LHC seminar by A. Chisholm.

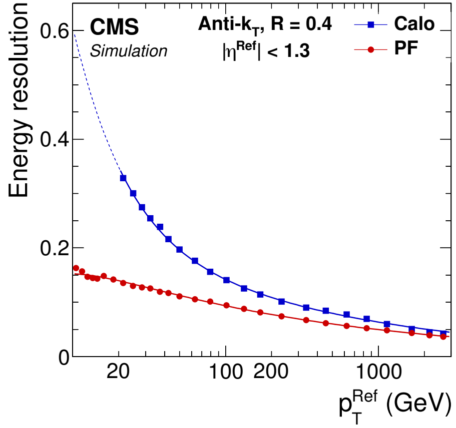
Particle-flow reconstruction

- 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



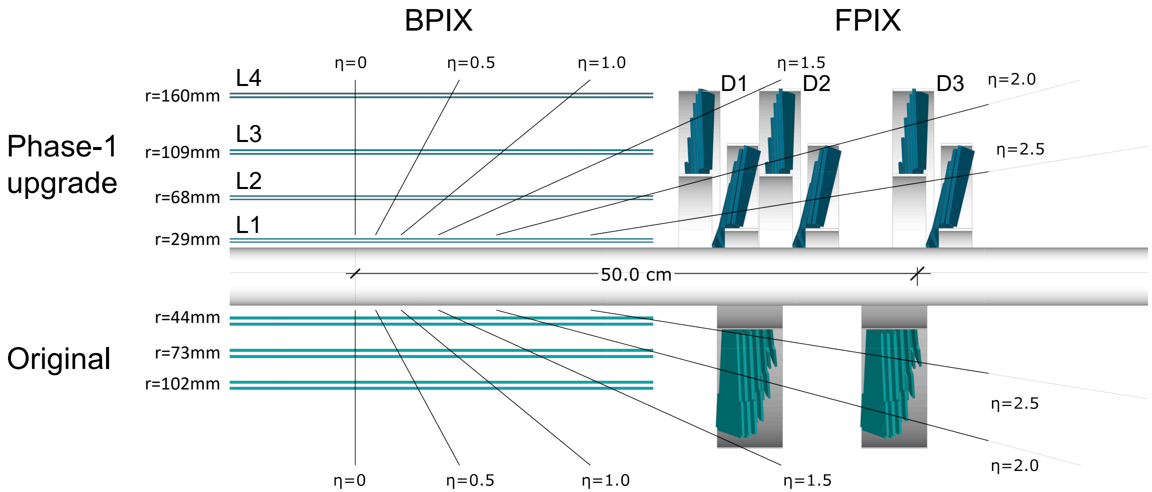
JINST 12 (2017) P10003



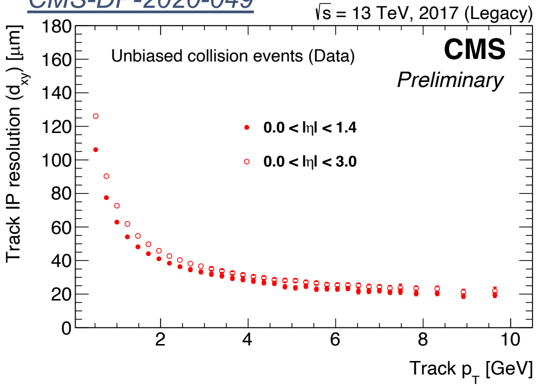


Phase-1 pixel detector upgrade

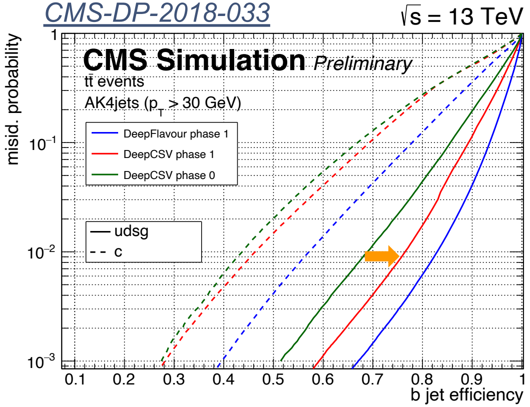
- New pixel detector installed during year-end stop 2016/2017



CMS-DP-2020-049



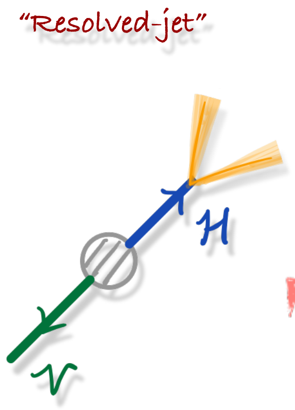
CMS-DP-2018-033

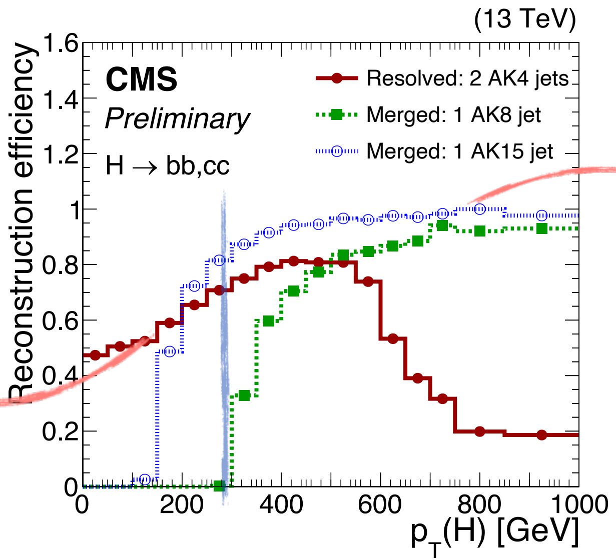


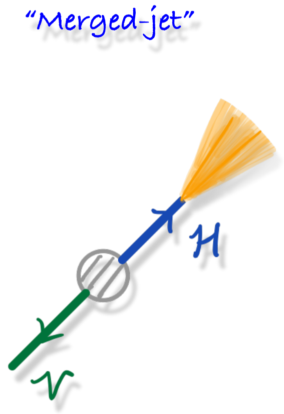
Improved tracking and flavour tagging performance in the 2017 — 2018 data set!

Analysis overview

- Two complementary approaches for Higgs boson candidate reconstruction







$\Delta R(c,\ c) \sim 2m(H)/p_T(H)$

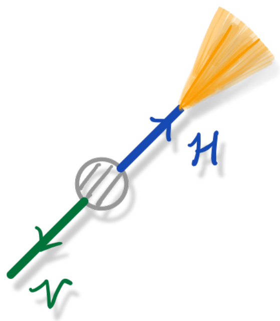
Resolved-jet topology

- reconstructs H → 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 → 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



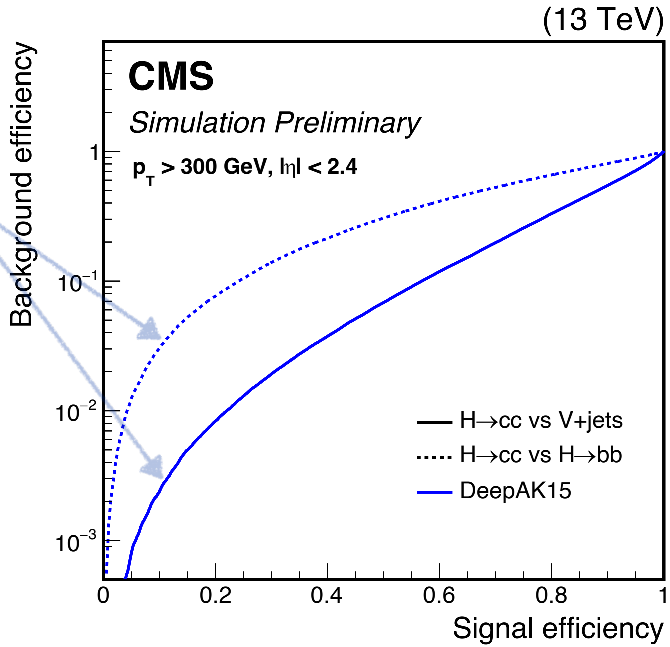
H → cc identification

- 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

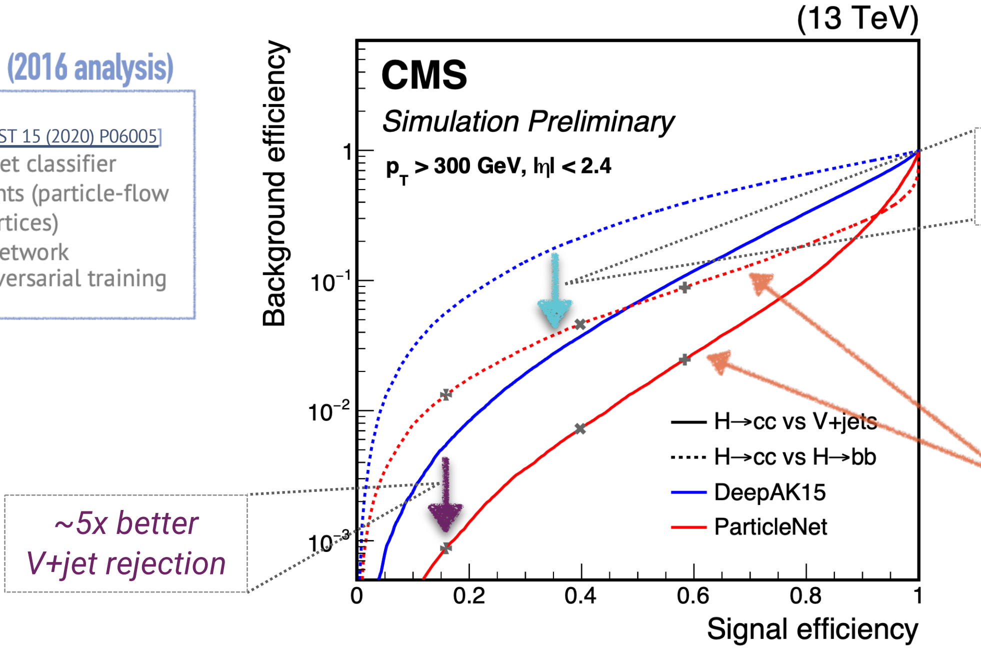


H → cc identification

- ❑ 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 → 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



CMS-PAS-HIG-21-008 (Run 2 analysis)

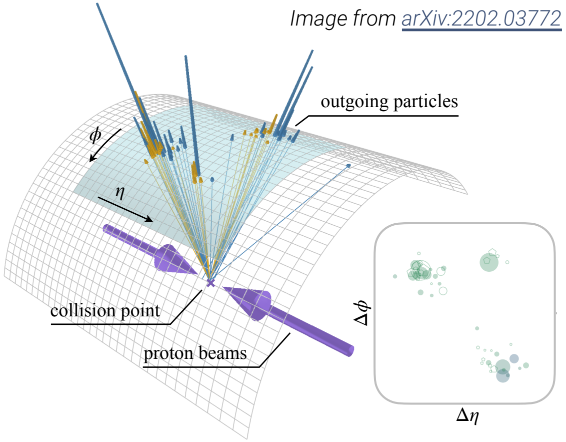
ParticleNet [CMS-DP-2020-002]
- same spirit as DeepAK8, but substantially improved:
- **graph neural network architecture**
- **novel mass decorrelation technique**

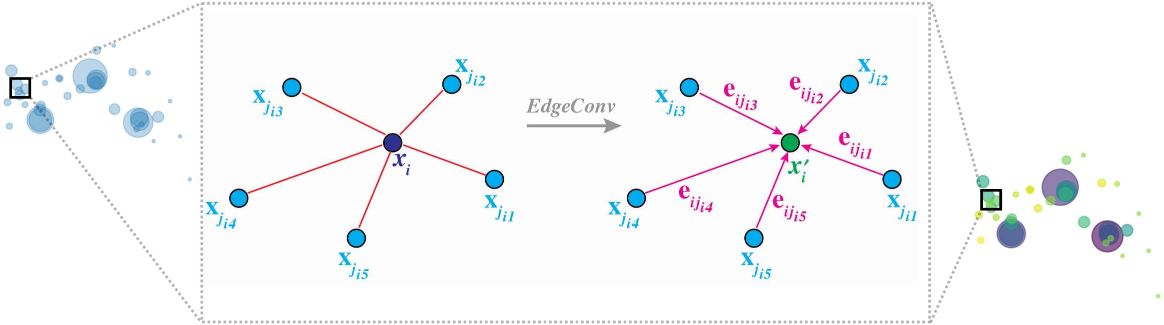
>2x improvement in the final sensitivity

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

Image from arXiv:2202.03772



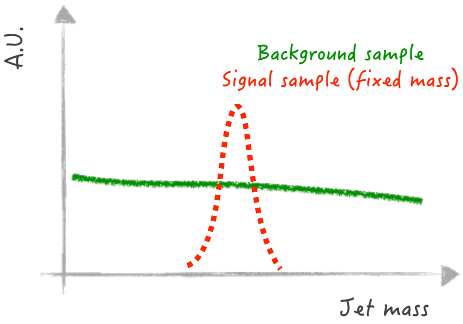


Performance on top quark tagging benchmark [*SciPost Phys. 7, 014 (2019)*]

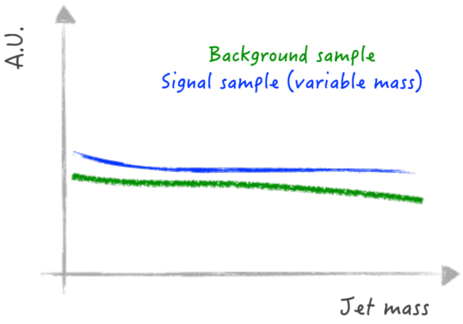
	$1/\varepsilon_b$ at $\varepsilon_s = 30\%$
ResNeXt-50	1147 ± 58
P-CNN	759 ± 24
PFN	888 ± 17
ParticleNet-Lite	1262 ± 49
ParticleNet	$\mathbf{1615} \pm \mathbf{93}$

Mass decorrelation

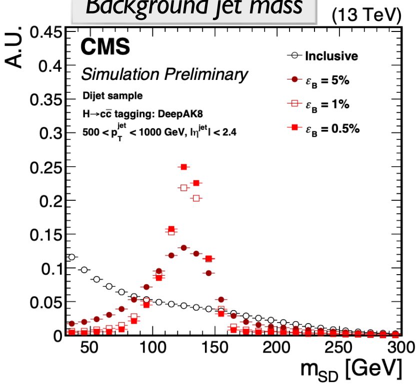
Plain training: no mass decorrelation



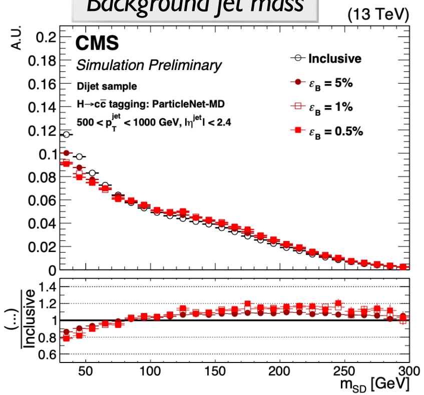
Mass-decorrelated training



Background jet mass



Background jet mass

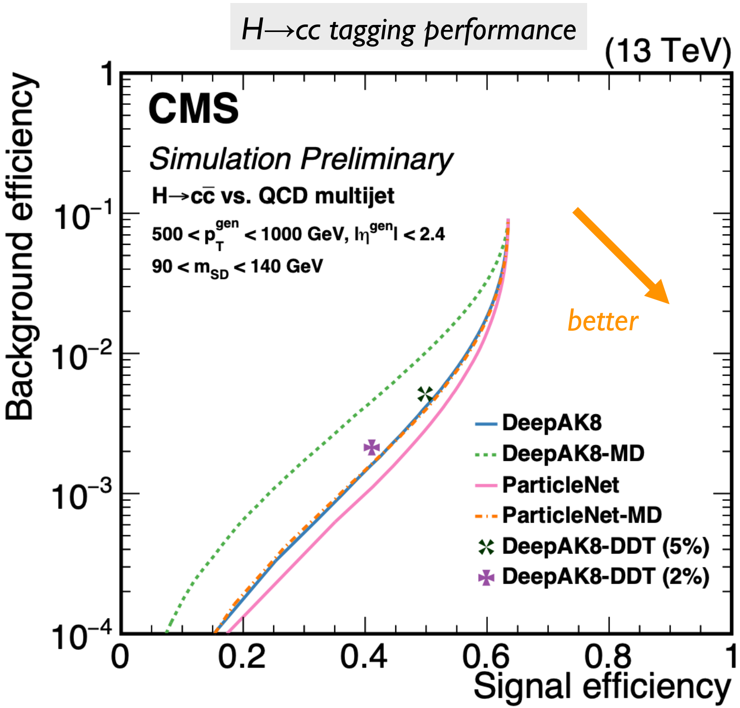


CMS-DP-2020-002

- "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**
      - $m(X) \in [15, 250]$ GeV
  - allows to easily reweight both signal and background to a ~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)

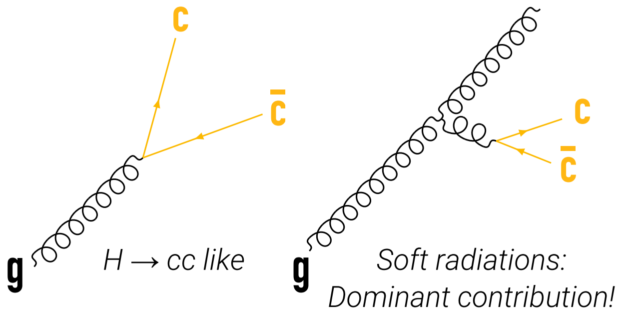
CMS-DP-2020-002



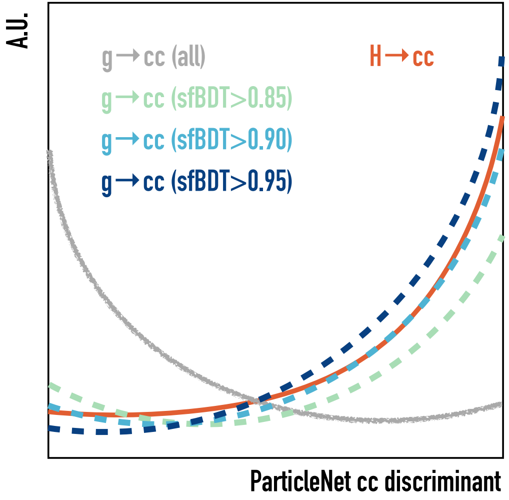
- ❑ "Mass sculpting": background jet mass shape becomes similar to signal after tagger selection
- ❑ New approach to prevent mass sculpting
  - using a special signal sample for training
    - hadronic decays of a spin-0 particle X
      - $X \to bb$, $X \to cc$, $X \to qq$
    - not a fixed mass, but a **flat mass spectrum**
      - $m(X) \in [15, 250]$ GeV
  - allows to easily reweight both signal and background to a ~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 → cc jets (or even Z → cc) in data
  - using g → cc in QCD multi-jet events as a proxy
- Difficulty: select a phase-space in g → cc that resembles H → 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%

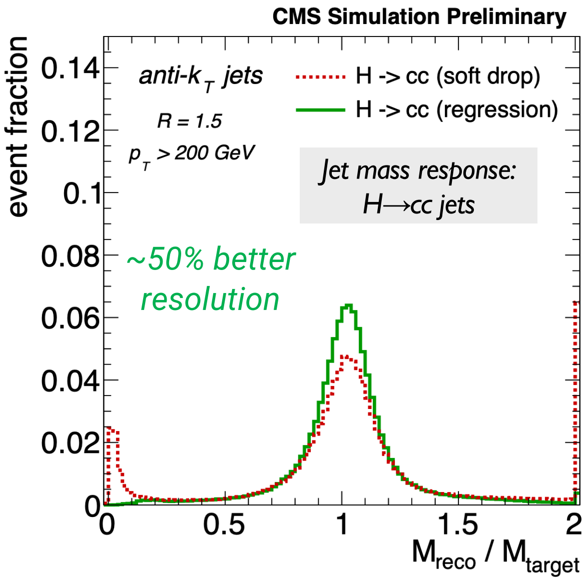


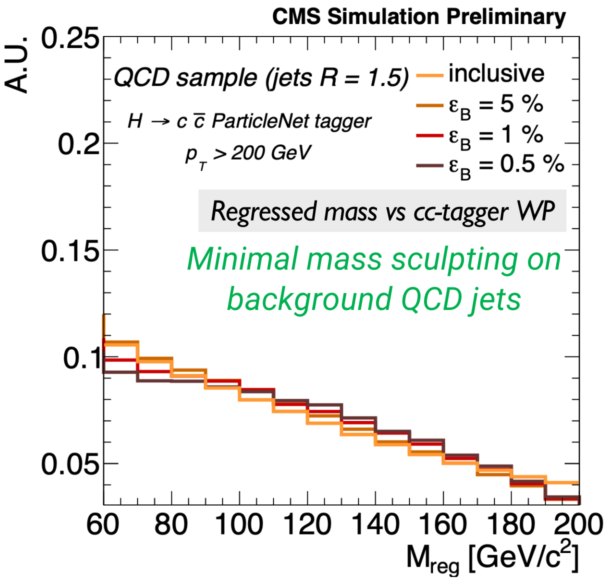
Effects of the BDT



Large-R jet mass regression

- Jet mass: one of the most powerful observable to distinguish signal and backgrounds — *CMS DP-2021/017*
- 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

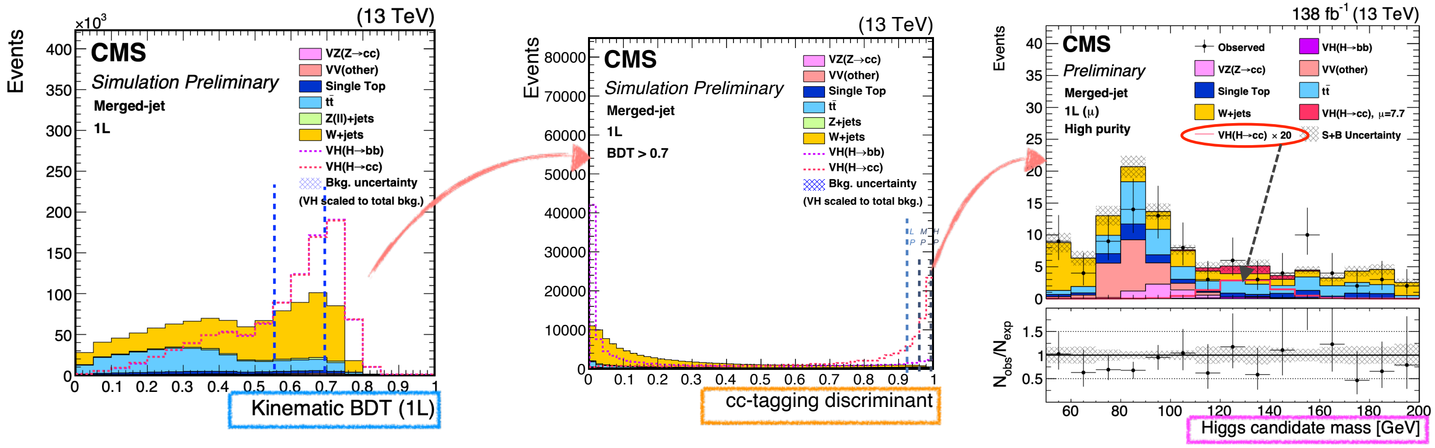




20 – 25% improvement in the final sensitivity

Analysis strategy

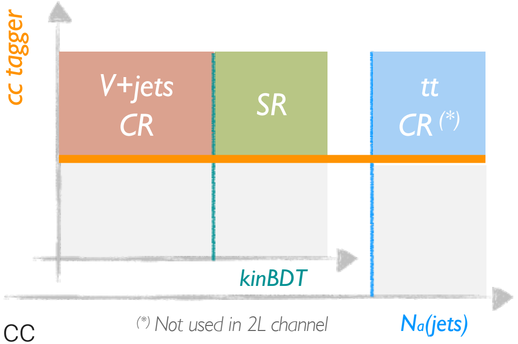
- Factorized approach for analysis design
  - event-level **kinematic BDT** developed in each channel to better suppress main backgrounds (V+jets, tt)
    - using only *event kinematics*, no intrinsic properties (e.g., mass/flavor) of the large-R jet
  - **ParticleNet cc-tagger** then used to define 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

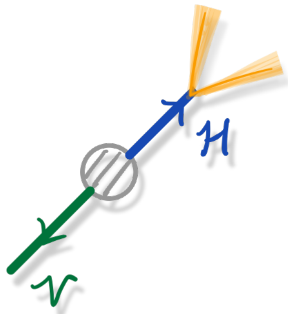
- Normalizations of main backgrounds estimated via dedicated data control regions (CRs)
  - V+jets CR: use the low kinematic BDT region
  - tt CR (0L & 1L): invert the cut on the number of additional small-R jets (i.e., $N_{aj} \geq 2$)
  - free-floating parameters scale the normalizations in CRs and 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 → cc) process (and VZ(Z → cc))
    - conservative uncertainty (2x/0.5x) for the misidentification of H(Z) → bb as H(Z) → cc

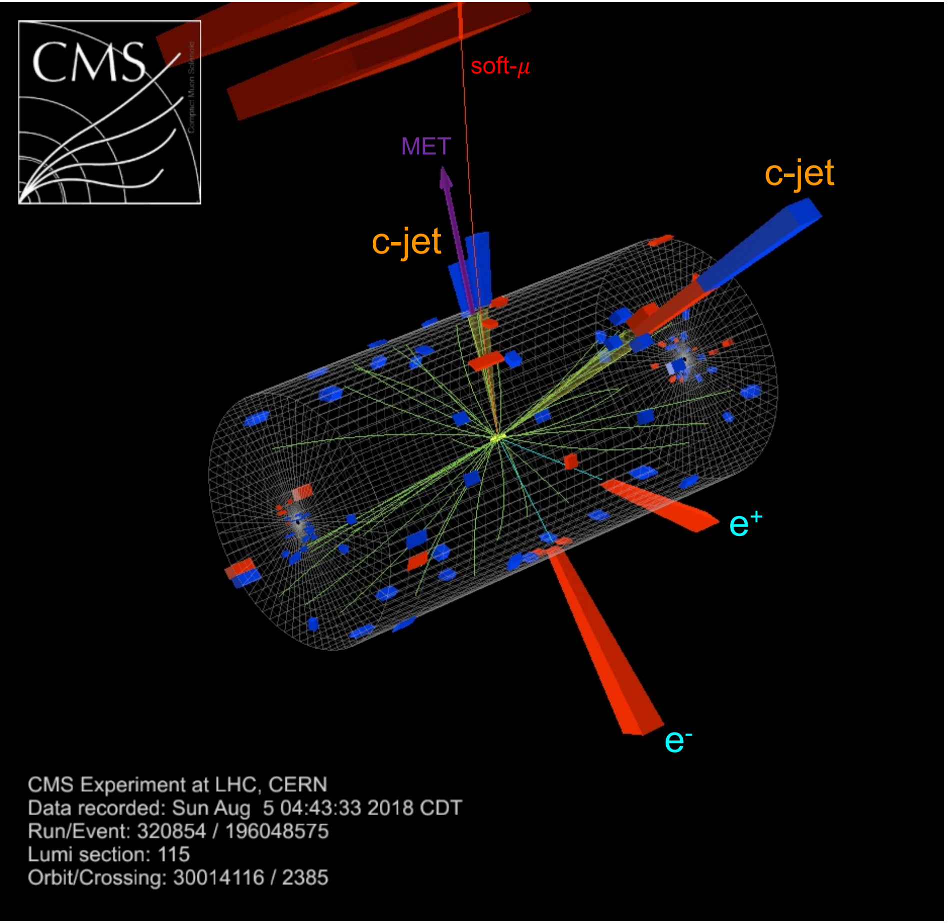


- Minor backgrounds (single top, dibosons, VH(H → bb)) estimated from simulation
  - dibosons: applying differential NNLO QCD + NLO EW corrections as a function of $p_T(V)$ [JHEP 2002 (2020) 087]

Resolved-jet topology



Overview of the resolved-jet topology

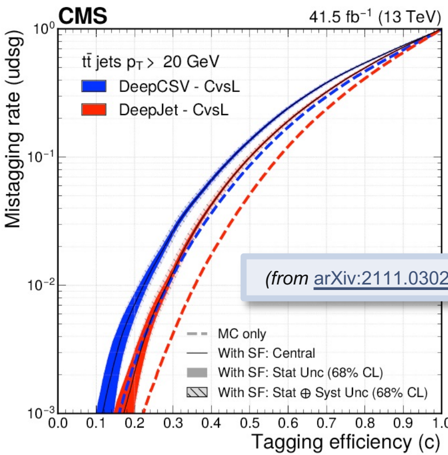


- Main challenges
  - Charm (c) tagging
  - QCD (reducible) and V+jets (irreducible) background
  - Relatively poor invariant mass resolution
- Higgs candidate reconstruction
  - Select two AK4-jets with the highest c-tagger discriminant score as Higgs jets
  - Dedicated c-jet energy regression for improved c-jet energy scale and resolution (eg. recovery of neutrino, unclustered hadrons, etc.) + Recover FSR-jets
  - Kinematic-fit (2L channels)
- Analysis strategy (three channels: 0L, 1L, 2L)
  - Control regions for background normalizations
  - BDT for final signal extraction

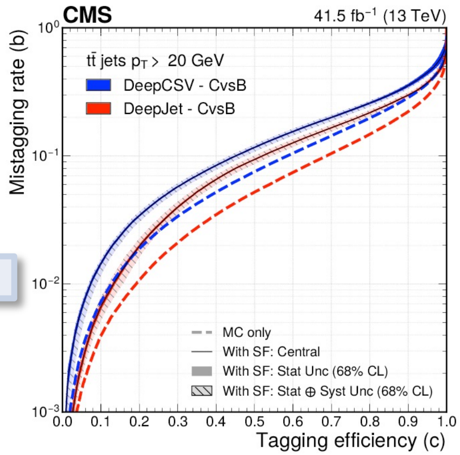
Charm-tagging in the resolved-jet topology

DeepJet algorithm as charm tagger

- C-jets have "intermediate" properties to b- and light-jets
  - It translates into the need to separate **c-jets** simultaneously from **light-jets** and **bottom jets**
- From DeepJet output score it is possible to build two c-jet taggers ➔ CvsL and CvsB
  - CvsL: it is optimized to differentiate charm-jets form light- or gluon-jets
  - CvsB: it is optimized to differentiate charm-jets from bottom-jets



(from arXiv:2111.03027)



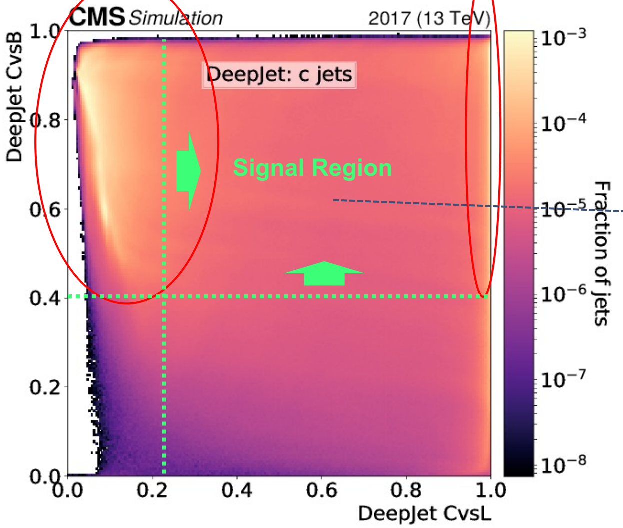
- Improvement vs DeepCSV (used in 2016 analysis)
  - Increase leading-jet c-tagging efficiency by ~30% for fixed b-jet and light-jet mis-tagging rate

Charm-tagging in the resolved-jet topology

- Definition of leading-jet working point
  - Studies of CvsB/CvsL jet score distributions in 2D plane
  - **CvsL>0.225, CvsB>0.4** ➔ c-jet identification efficiency of ~43% with a b-jet and light-jet mis-tagging rate respectively of ~15% and ~4% (depending on the year)

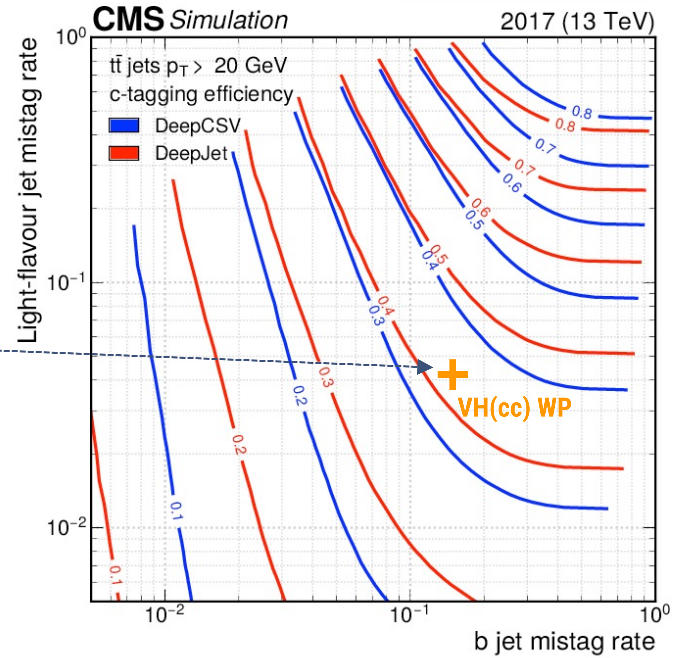
c-jets mis-identified as light ones

c-jets



(from arXiv:2111.03027)

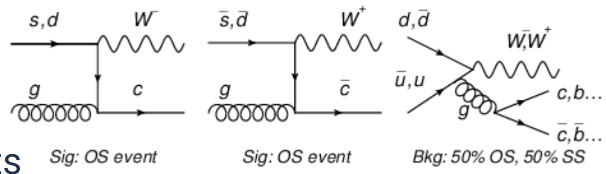
(from arXiv:2111.03027)

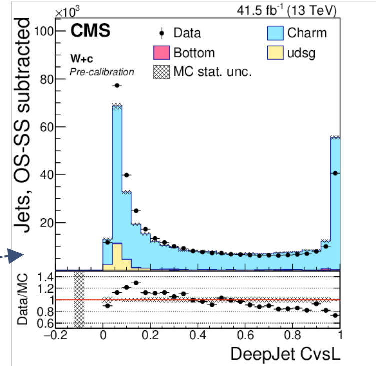


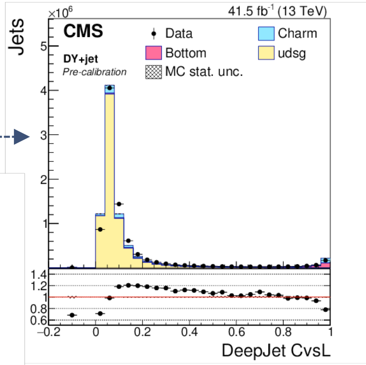
A new method to calibrate charm-taggers

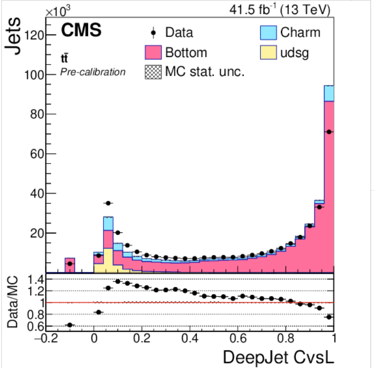
(from arXiv:2111.03027)

- Methodology
  - Iterative approach exploiting 3 distinct control regions, each enriched in b-jets, c-jets, or light-flavour jets
- Selecting an abundant and pure source of charm-jets
  - Target W production in association with charm quarks (W+c)
  - 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: semi-(di-)leptonic $t\bar{t}$+jets and DY(Z→$\mu\mu/ee$)+jets









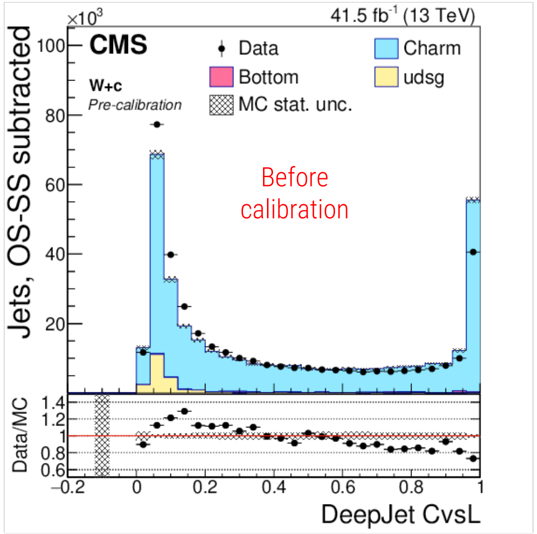
- First time that a calibration method to correct the 2D distribution of c-tagging discriminator shapes is presented

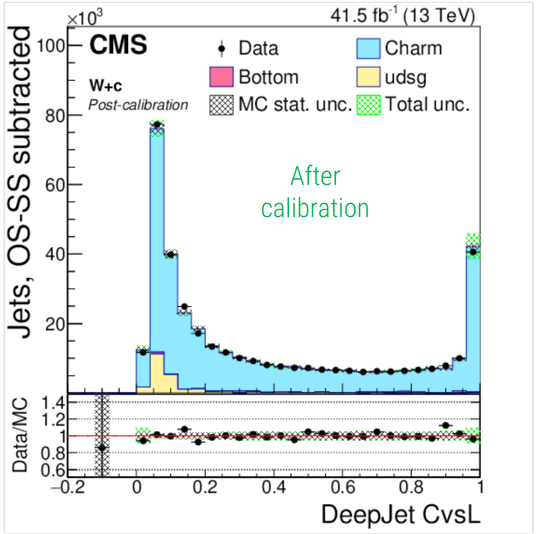
➔ arXiv:2111.03027 (**accepted by *JINST***)

A new method to calibrate charm-taggers

Application of the reshaping scale-factors

(from arXiv:2111.03027)



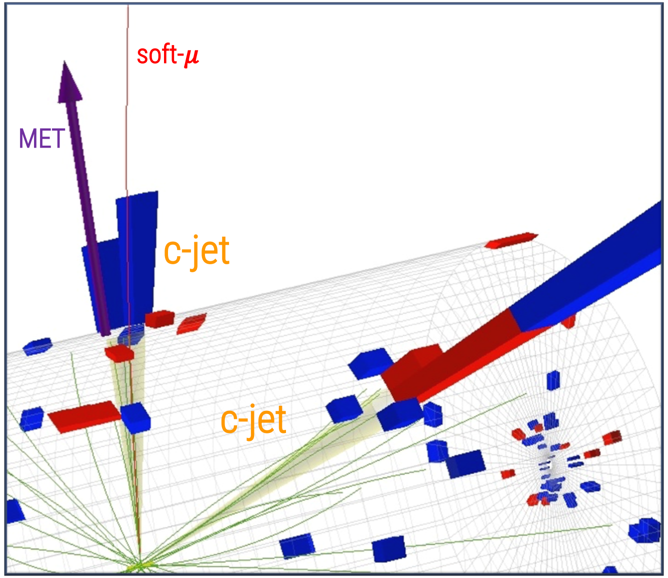


- Very good data/MC agreement after the calibration
  - Application through an event-by-event re-weighting: $w_i = \prod_{i=1}^{jets} sf_i(CvsL, CvsB)$

A dedicated charm-jet energy regression

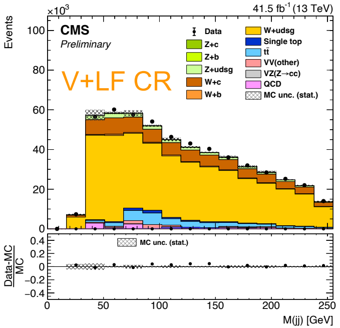
Goal: improve c-jet energy scale and resolution

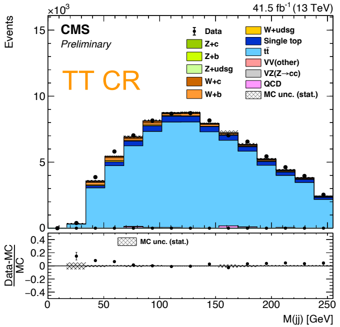
- Inspired by b-jet energy regression [arXiv:1912.06046]
  - Jet energy measurements not always accurate:
    - neutrinos, hadrons outside jet radius, etc. Effect enhanced in c-jets and b-jets
  - Dedicated algorithm to determine c-jet energy scale and resolution
  - A DNN algorithm pioneered for the observation of the H→bb decay
- Regression performed using DNN architecture:
  - Trained using c-jets collected from W→cq decays in $t\bar{t}$+jets MC events
  - Target is represented by $p_T(\text{gen})/p_T(\text{reco})$
- Input features
  - Total of 43 input variables as input to the network
  - Jets: kinematics, energy fraction, leading+soft-lepton tracks, pile-up, secondary vertices
  - Jet energy shapes (e.g. energy fraction, etc), jet constituents, $p_T(\text{jet})/p_T(\text{lepton})$

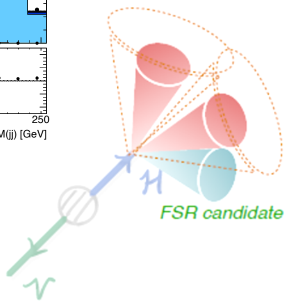


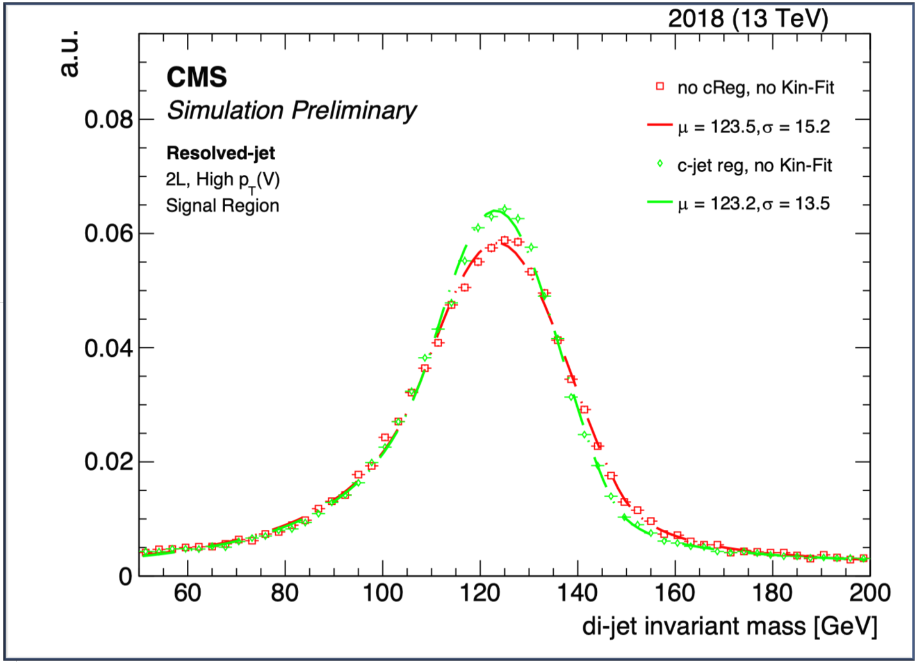
A dedicated charm-jet energy regression

- **~15% improvement in mass resolution**
  - Depending on the jet p_T
- Validated in VH(H→cc) control regions





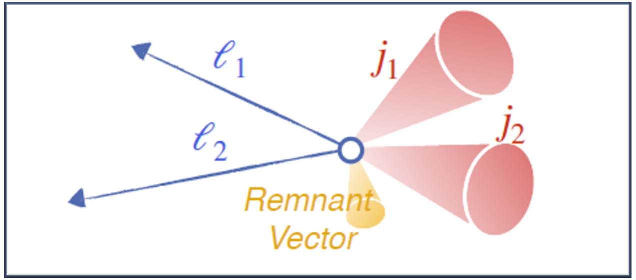


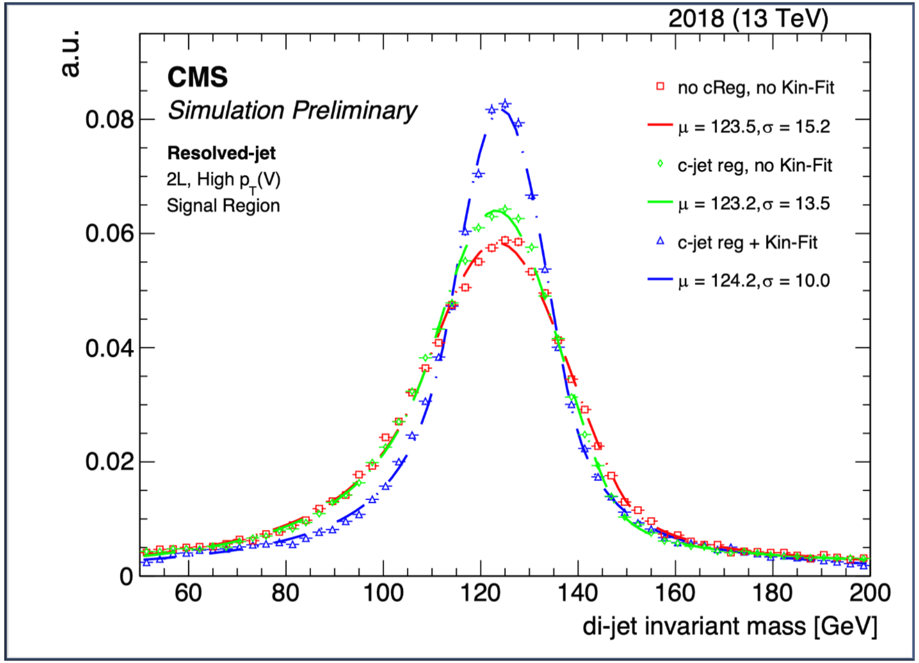


- FSR recovery
  - Further improve di-jet invariant mass resolution
  - Jets with $p_T < 20$ GeV, $|\eta| < 3$, and within $\Delta R < 0.8$ of Higgs jets are included in Higgs 4-momentum

Kinematic-fit in the 2L channels

- No intrinsic missing energy in $Z(\ell\ell)H(cc)$ process
- Improve jet p_T measurement through a kinematic fit:
  - Constrain di-lepton system to the Z boson mass
  - Balance the $\ell\ell$+cc+jets system in the (p_x, p_y) plane
  - Allow MET to adjust within the experimental resolution

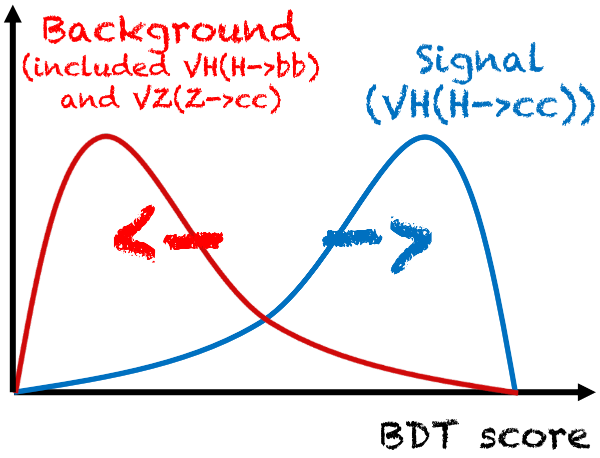




- Up to **~30% improvement in Higgs mass resolution**

Signal extraction – BDT training in SRs

- BDT trained to separate signal from background samples
  - Use combination of event kinematic observables, Higgs and vector boson properties, particle flavor variables (tagger information), and kinematic-fit variables (only in 2L channels)
- 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 distributions used in SR for the final fit



Analysis categories and background estimation

- ❑ 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) + $t\bar{t}$ +jets
  - Freely floating in each channel/year

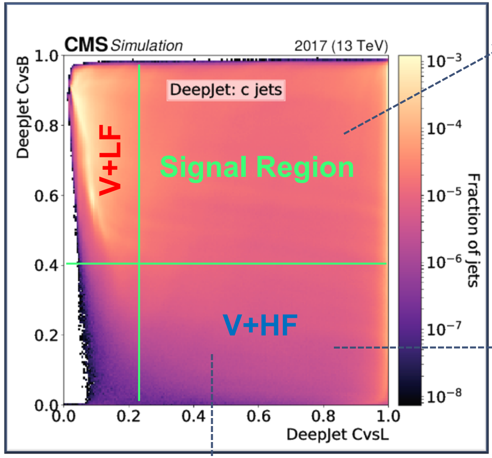
❑ Selections optimized for the different decay of the vector boson considered

❑ Definition of 4 analysis categories

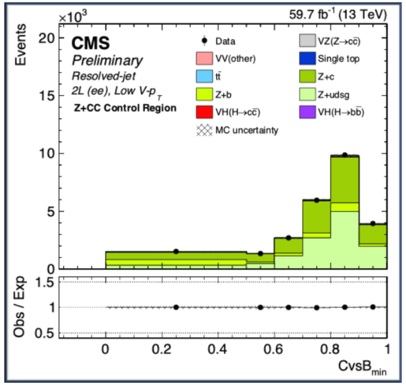
- ❑ **0L**: $p_T(Z)>170$ GeV
- ❑ **1L**: $p_T(W)>100$ GeV
- ❑ **2L Low-p_T**: 60 GeV $<p_T(Z)<150$ GeV
- ❑ **2L High-p_T**: $p_T(Z)>150$ GeV

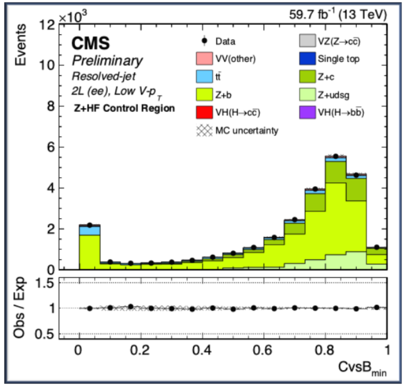
❑ All the categories have TT, LF, HF and CC CRs (1L has not HF) + 1 SR

❑ Simultaneous fit to BDT in SR and tagger shapes in CRs



V+CC
⟲
Veto m(H) region



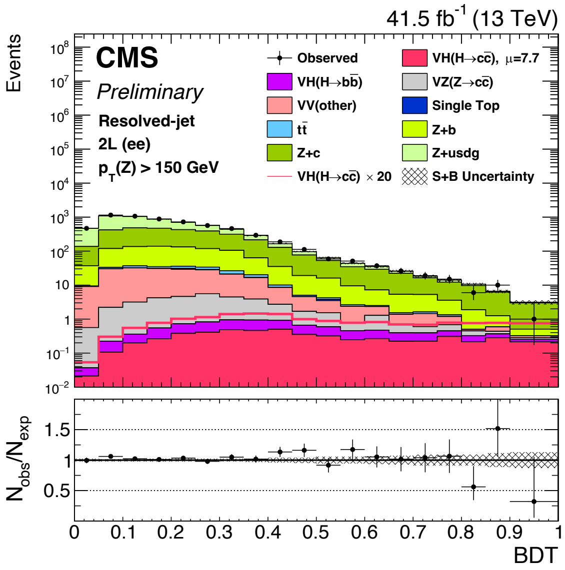


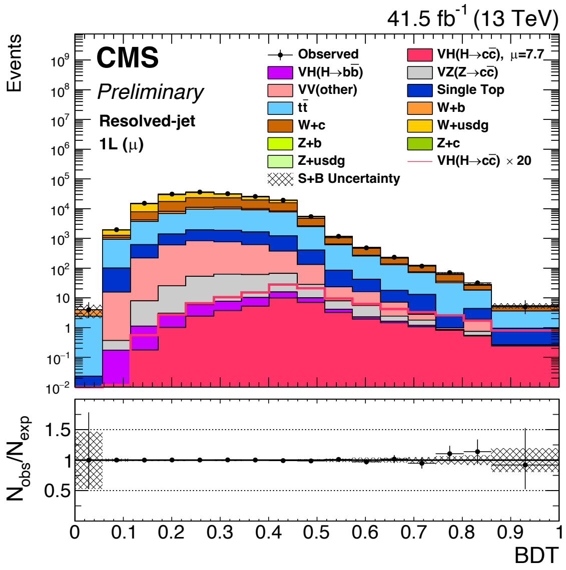
$t\bar{t}$
⟲
Invert Z mass (2L)
Require add jet (1L)*
Require add ℓ and jets (0L)

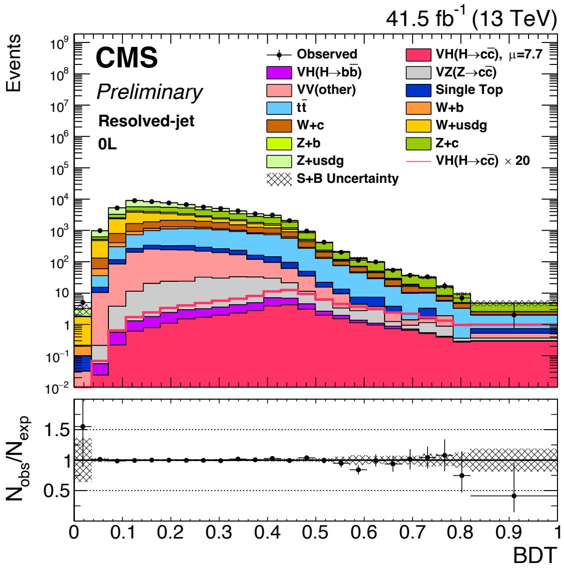
*1L: also require MET<170 GeV to keep orthogonal to 0L $t\bar{t}$ CR

Postfit plots – Signal regions

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





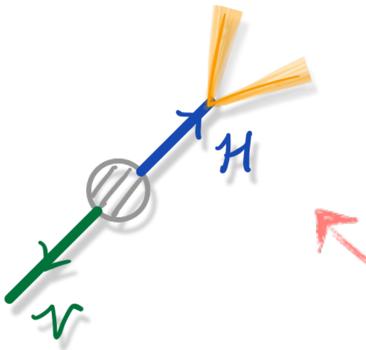


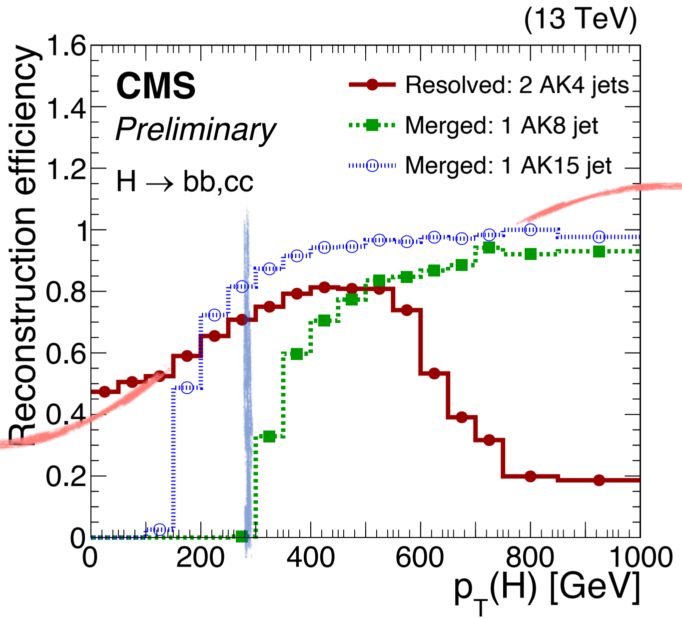
Results

Combination of the two topologies

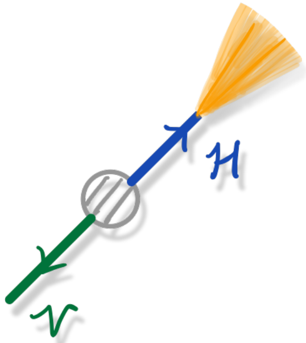
- The two topologies are made orthogonal via presence of AK15 jet with p_T > 300 GeV
  - p_T threshold chosen to maximize expected sensitivity

"Resolved-jet"





"Merged-jet"



$\Delta R(c, c) \sim 2m(H)/p_T(H)$

Uncertainties

- All correlated between topologies, except:
  - Background normalization SFs for V+jets and $t\bar{t}$
  - c-tagging efficiencies

- Main uncertainties
  - Limited statistics of data
  - Statistical uncertainties of V+jets samples
  - Charm tagging efficiencies

Uncertainty source	$\Delta\mu / (\Delta\mu)_{\mathrm{tot}}$
Statistical	**85%**
Background normalizations	37%
Experimental	**48%**
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%
Theory	**22%**
Backgrounds	17%
Signal	15%

VZ(Z→cc) results

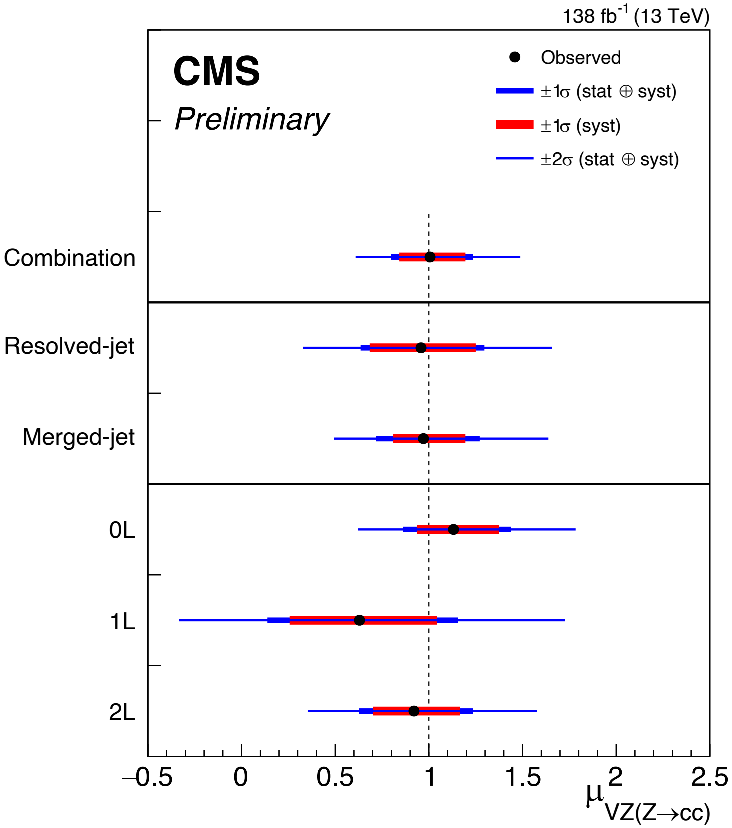
- Analysis validated by looking for VZ(Z→cc) process
  - Same analysis procedure, but extracting VZ(Z→cc) signal in the final fit
  - Resolved-jet: retrained BDTs with VZ(Z→cc) as signal
  - VH(H→cc) fixed to SM expectation

- Observed (expected) signal strength for VZ(Z→cc):

$$\mu_{VZ(Z\to cc)} = 1.01^{+0.23}_{-0.21}(1.00^{+0.22}_{-0.20})$$

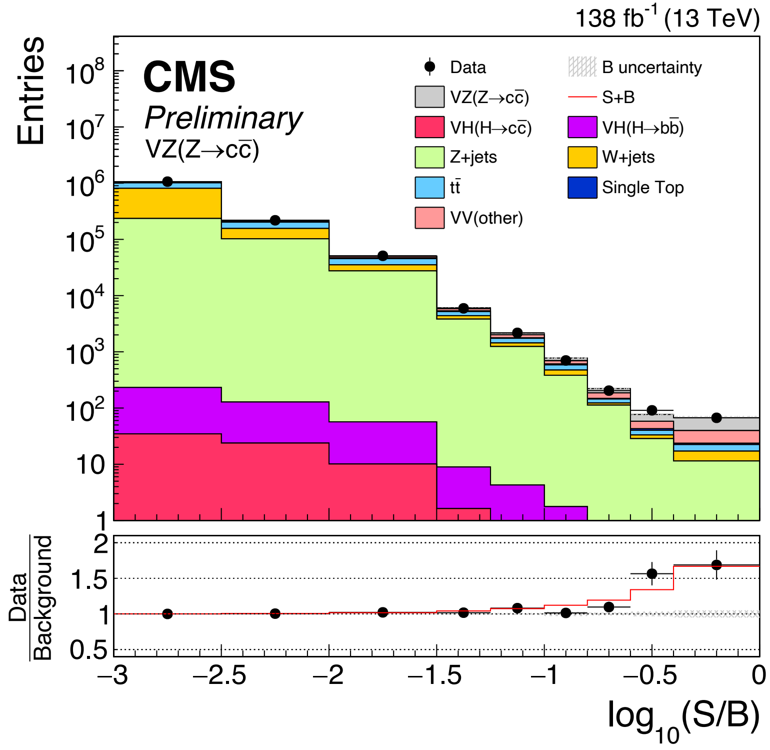
with a significance of 5.7σ (5.9σ)

- **First observation of Z→cc at hadron collider!**



VZ(Z→cc) results

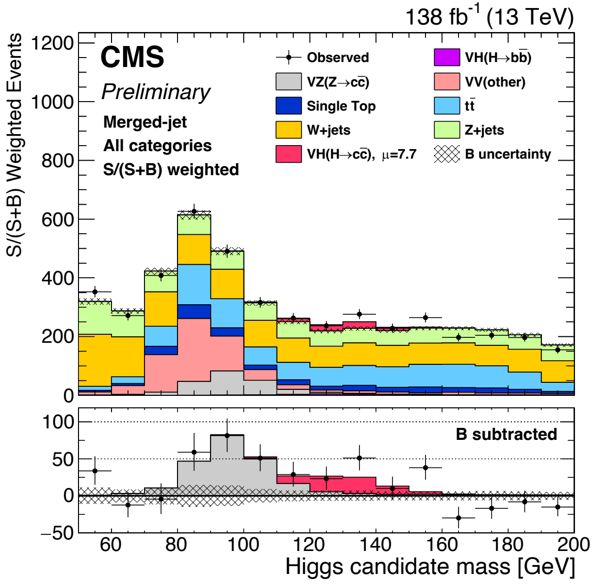
- Observing the excess: distribution of events ordered by $\log_{10}(S/B)$



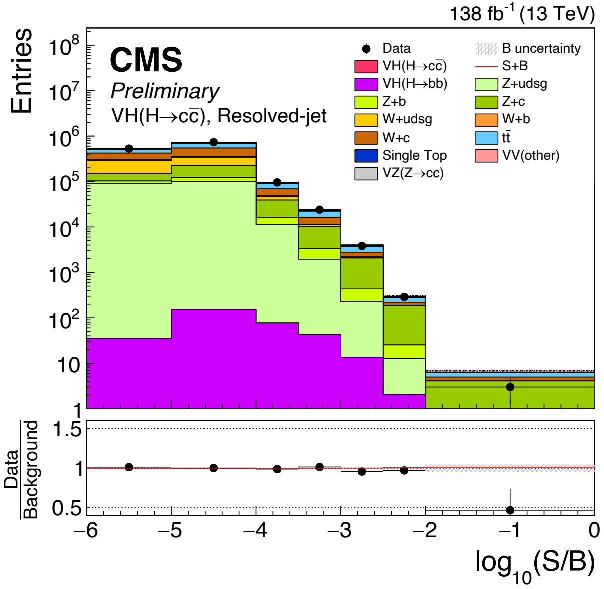
VH(H→cc) results

- ❑ Merged-jet topology: distribution of the Higgs boson candidate mass
- ❑ Resolved-jet topology and the combination: ordering the events by $\log_{10}(S/B)$

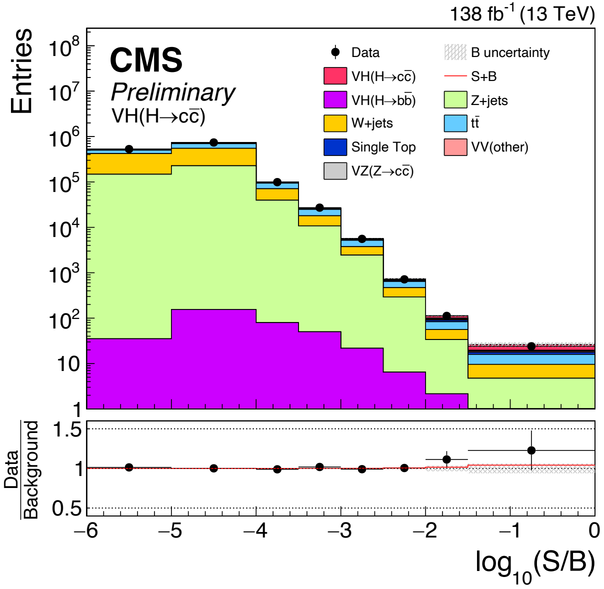
Merged-jet



Resolved-jet

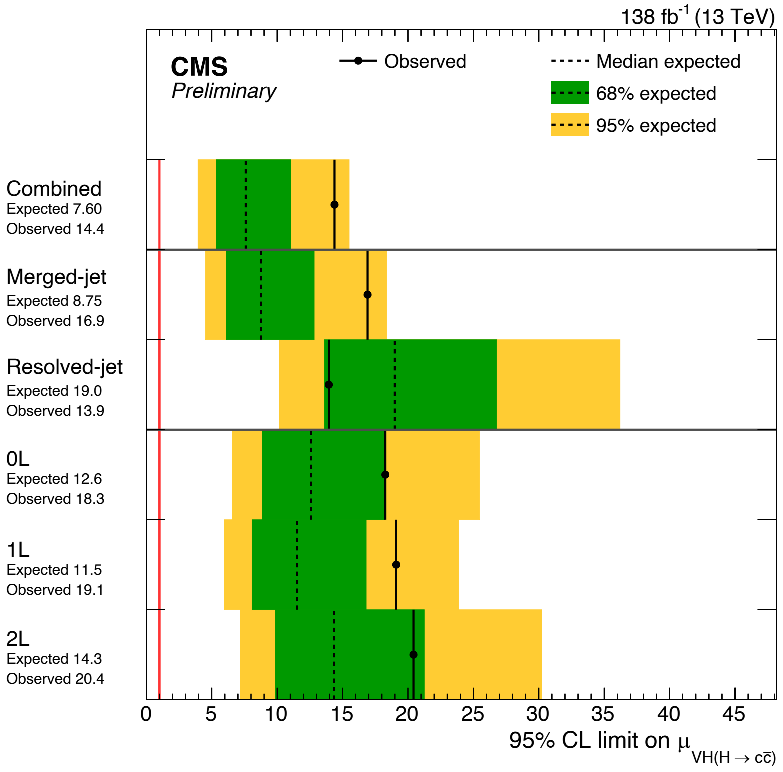


Merged + Resolved



VH(H→cc) results

- Observed (expected) upper limit on VH(H→cc) signal strength at 95% CL: $\mu_{VH(H\to cc)} < 14\ (7.6^{+3.4}_{-2.3})$
  - **Strongest limits on VH(H→cc) process to date!**
  - ATLAS Full Run 2 result: $\mu_{VH(H\to cc)} < 26\ (31)$ [arXiv:2201.11428]
- Best fit signal strength: $\mu_{VH(H\to cc)} = 7.7^{+3.8}_{-3.5}$
  - Consistent with the SM prediction within 2σ
- Obs. (Exp.) upper limits from each topology:
  - **Resolved-jet** topology: **14(19) × SM**
  - **Merged-jet** topology: **17(8.8) × SM**



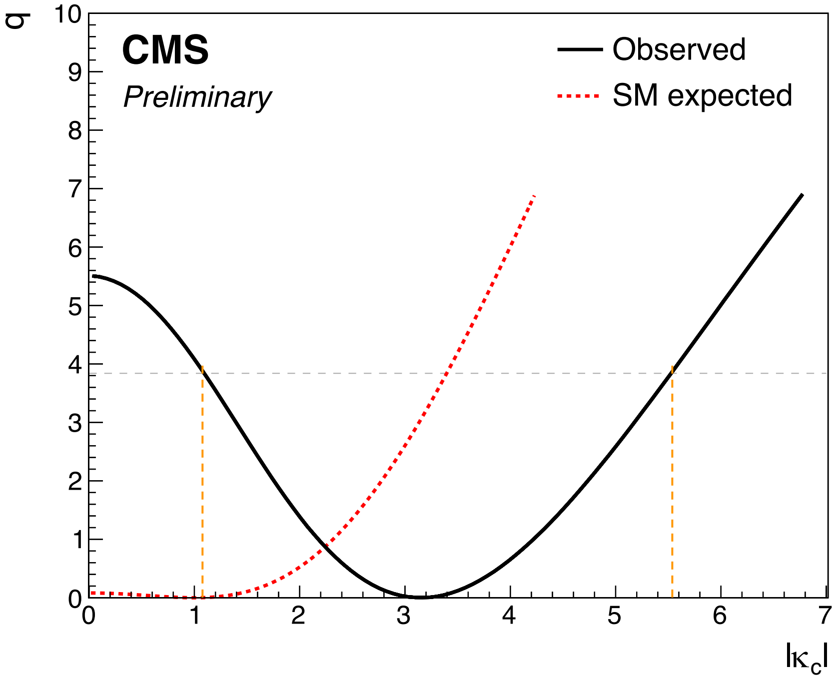
VH(H→cc) results

Results used to place new constraints on κ_c

- Only considering effects on $\mathcal{B}(\mathrm{H} \to \mathrm{cc})$ and fixing all other couplings to their SM values

$$\mu_{VH(H\to cc)} = \frac{\kappa_c^2}{1+\mathcal{B}_{SM}(H\to cc)\times(\kappa_c^2-1)}$$

- The 95% CL intervals obtained with likelihood scans
  - observed: $1.1 < |\kappa_c| < 5.5$
  - expected: $|\kappa_c| < 3.4$
- **Strongest constraints on $|\kappa_c|$ to date**
  - Outperforming indirect measurements of $|\kappa_c| \lesssim 6.2$: PRD 92 (2015) 033016
  - Comparable to the previous projection for HL-LHC [ATL-PHYS-PUB-2021-039]



Conclusions

- New results of the CMS search for the VH(H→cc) process are presented
  - Benefit from the full Run 2 dataset
  - Substantial improvements in charm tagging performance
  - Major upgrades of analysis techniques, such as jet energy/mass regression, kinematic fits, etc.
- Analysis validated by measuring VZ(Z→cc) signal strength: $\mu_{VZ(Z \to cc)} = 1.01^{+0.23}_{-0.21}$
  - Significance of 5.7σ (5.9σ) ➔ First observation of Z→cc at a Hadron Collider!
- Upper limits on VH(H→cc): $\mu_{VH(H \to cc)} < 14$ (7.6 exp.)
  - Almost 5x increase in expected sensitivity compared to analysis using 2016 data
  - Constraints on Higgs-charm coupling: $1.1 < |\kappa_c| < 5.5$ ($|\kappa_c| < 3.4$ exp.) – Most stringent to date!

Prospects

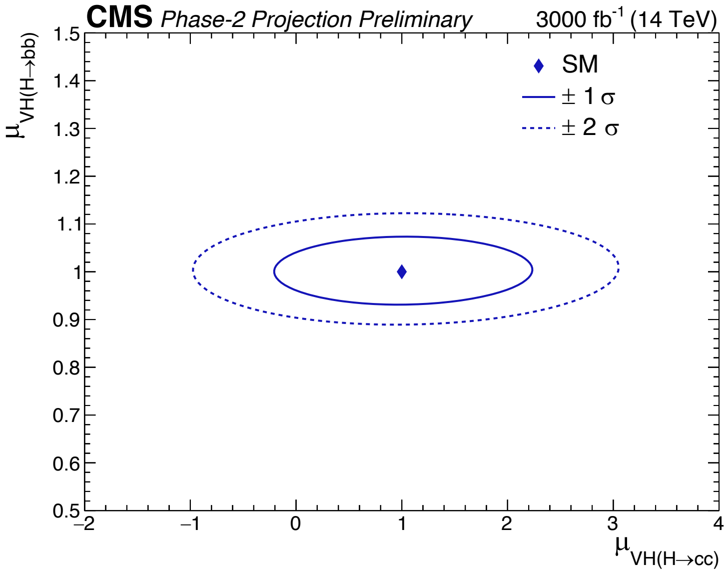
Projection at HL-LHC: Setup

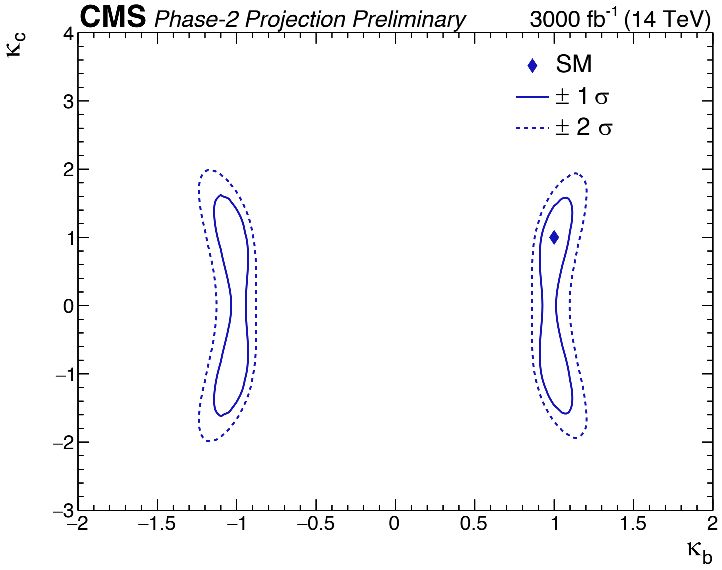
- Extrapolation of the merged-jet analysis to HL-LHC with 3000 fb^{-1} data
- Modifications to the Run 2 analysis to allow for a simultaneous constraint on $H \rightarrow bb$ and $H \rightarrow cc$
  - addition of 3 categories enriched in $H \rightarrow bb$ decays, selected with the ParticleNet bb-tagging 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 ~3% and ~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 ~100% (Run 2) to 20% in the projection

Projection at HL-LHC

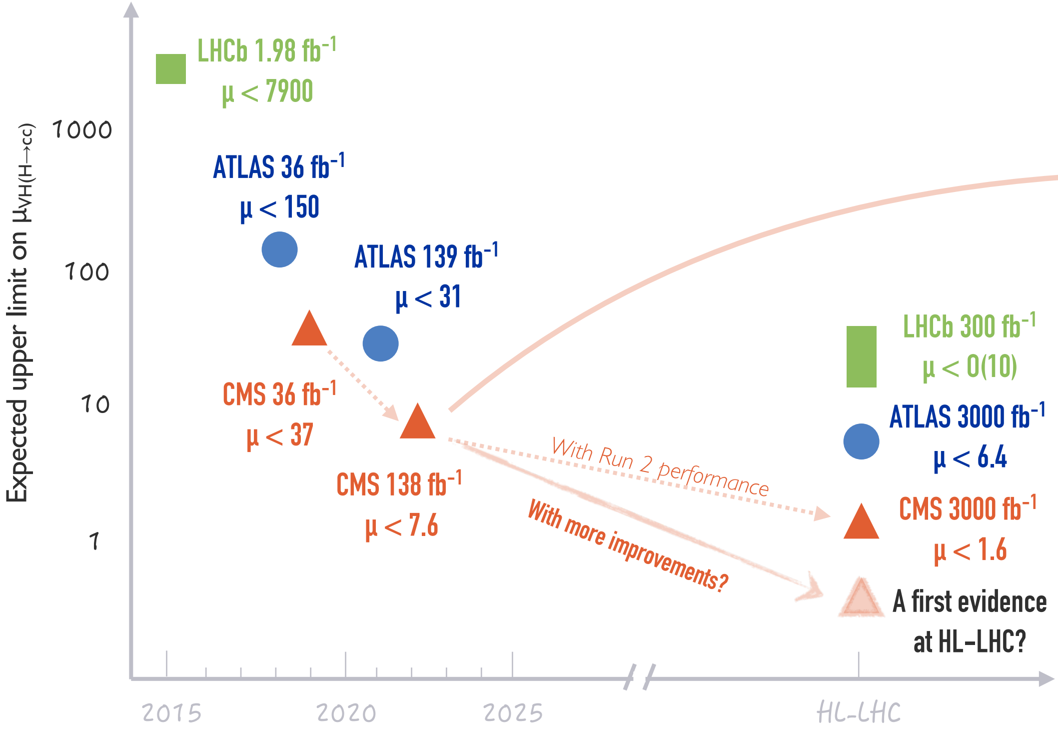
- Simultaneous extraction of the H → bb and H → cc signal strengths
  - $\mu_{VH(H \to bb)} = 1.00 \pm 0.03\ (\text{stat.}) \pm 0.04\ (\text{syst.}) = 1.00 \pm 0.05\ (\text{total})$
  - $\mu_{VH(H \to cc)} = 1.0 \pm 0.6\ (\text{stat.}) \pm 0.5\ (\text{syst.}) = 1.0 \pm 0.8\ (\text{total})$

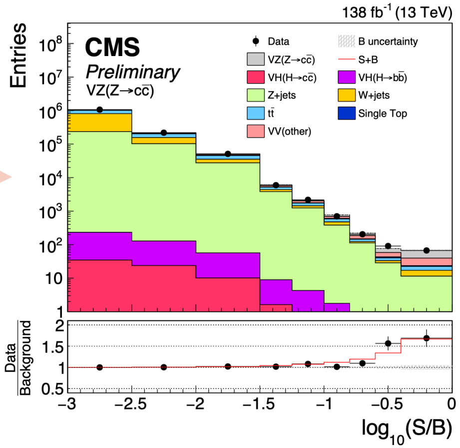




Expected sensitivity approaches the SM value for the Higgs-charm coupling.

A charming journey





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

- More channels: ttH(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, . . .

From $\mathcal{O}(1000)$ to $\mathcal{O}(100)$ to $\mathcal{O}(10)$ in ~5 years.
A combined effort and creativity from instrumentation, physics objects and analysis techniques!

A charming journey ahead!

Backups

$H \to cc$ searches at the LHC

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

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

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

Baseline event selections

Merged-jet topology

Variable	0L	1L	2L
p_T^ℓ	—	(>25,>30)	>20
Lepton isolation	—	(<0.06, —)	(<0.25, —)
$N_{a\ell}$	=0	=0	—
$M(\ell\ell)$	—	—	75–105
$N^{aj}_{small\text{-}R}$	<2	<2	<3
p_T^{miss}	>200	>60	—
$p_T(V)$	>200	>150	>150
$p_T(H_{cand})$	>300	>300	>300
$m\ (H_{cand})$	50–200	50–200	50–200
$\Delta\phi(V, H_{cand})$	>2.5	>2.5	>2.5
$\Delta\phi(\vec{p}_T^{miss}, j)$	>0.5	—	—
$\Delta\phi(\vec{p}_T^{miss}, \ell)$	—	<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^ℓ	—	(>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
$N^{aj}_{small\text{-}R}$	—	<2	—	—
p_T^{miss}	> 170	—	—	—
p_T^{miss} significance	—	>4	—	—
$p_T(V)$	>170	>100	60–150	>150
$p_T(H_{cand})$	>120	>100	—	—
$m\ (H_{cand})$	<250	<250	<250	<250
$\Delta\phi(V, H_{cand})$	>2.0	>2.5	>2.5	>2.5
$\Delta\phi(\vec{p}_T^{miss}, j)$	>0.5	—	—	—
$\Delta\phi(\vec{p}_T^{miss}, \ell)$	—	<2.0	—	—

Uncertainties

- Breakdown of the uncertainties in each topology

Merged-jet topology

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

Uncertainty source	$\Delta\mu/\ (\Delta\mu)_{\mathrm{tot}}$
Statistical	**88%**
Background normalizations	39%
Experimental	**40%**
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%
Theory	**25%**
Backgrounds	21%
Signal	14%

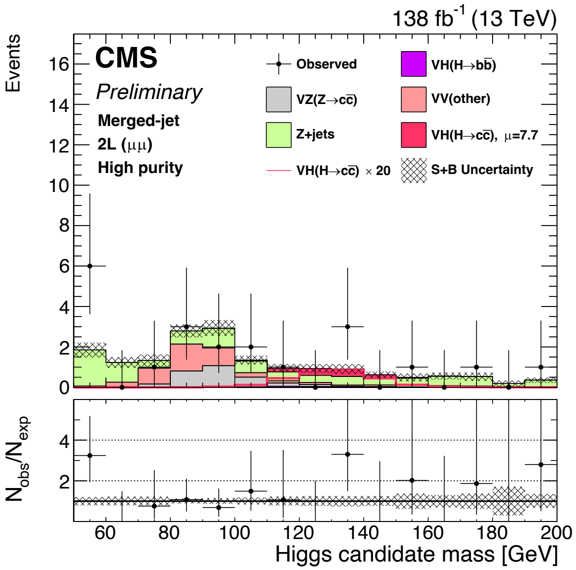
Resolved-jet topology

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

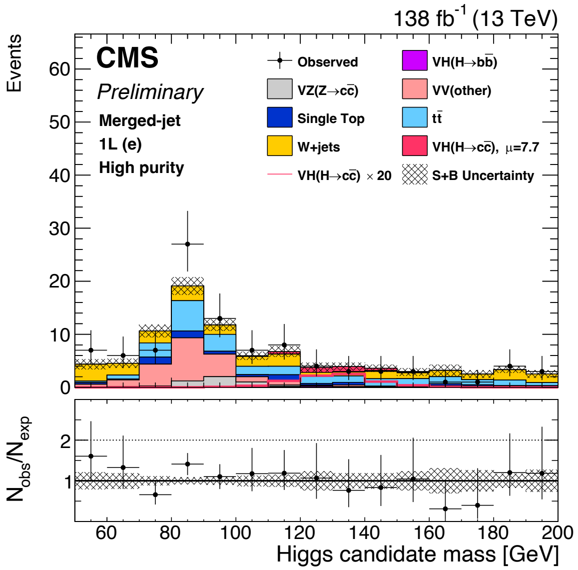
Uncertainty source	$\Delta\mu/\ (\Delta\mu)_{\mathrm{tot}}$
Statistical	**66%**
Background normalizations	28%
Experimental	**72%**
Sizes of the simulated samples	59%
Charm identification efficiencies	27%
Jet energy scale and resolution	17%
Simulation modeling	20%
Luminosity	13%
Lepton identification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

Merged-jet topology: signal regions

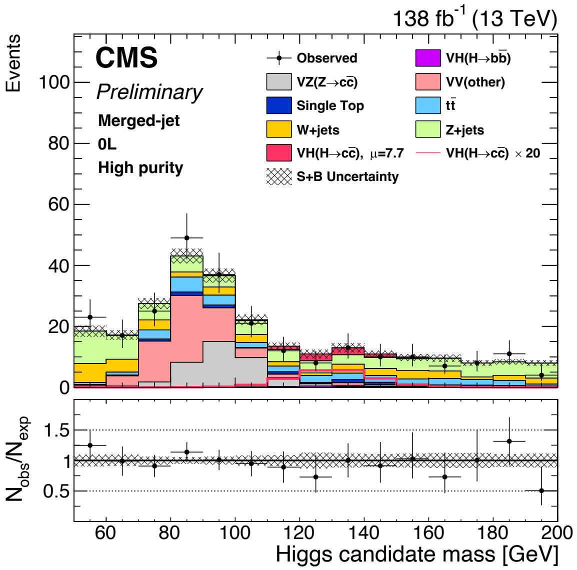
2L(μμ), high cc-purity



1L(e), high cc-purity



0L, high cc-purity

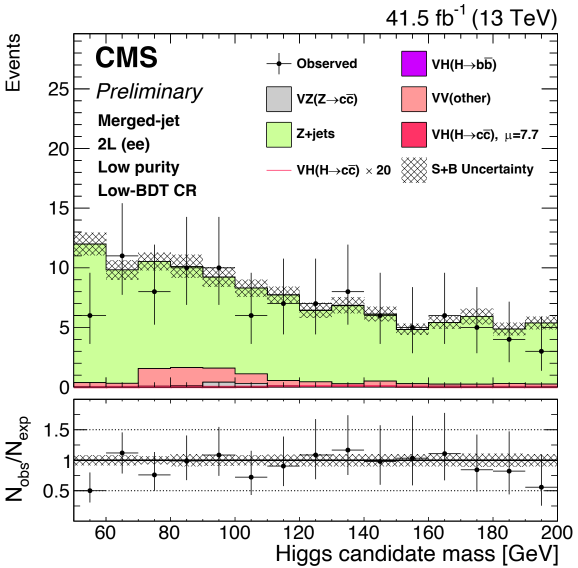


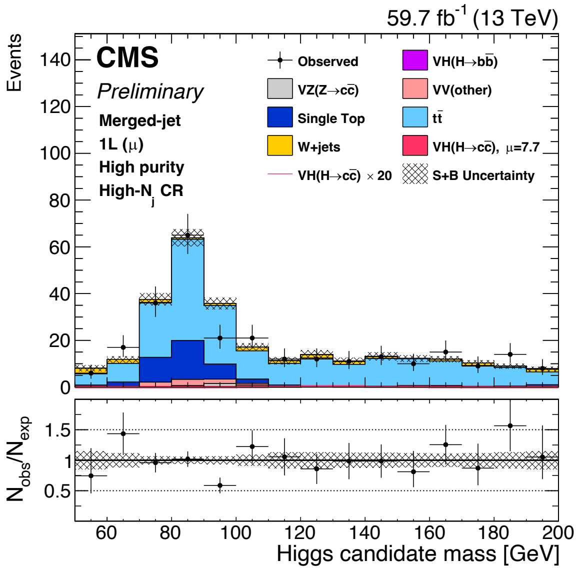
Merged-jet topology: control regions

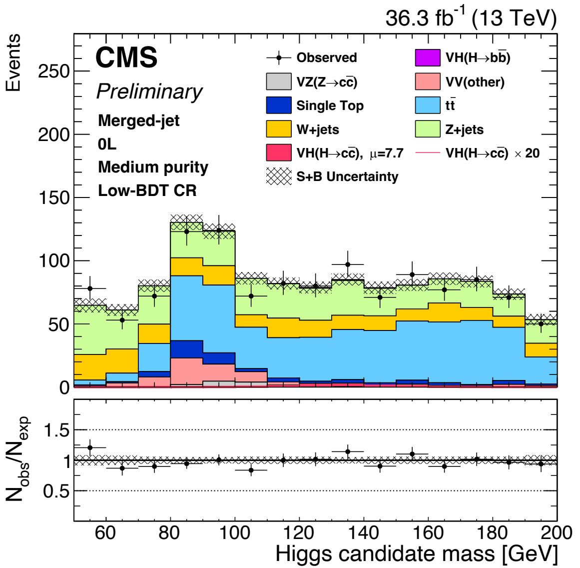
2L(ee), V+jets CR, low cc-purity

1L(μ), tt CR, high cc-purity

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

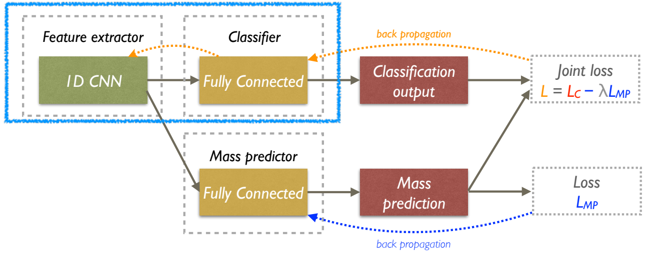




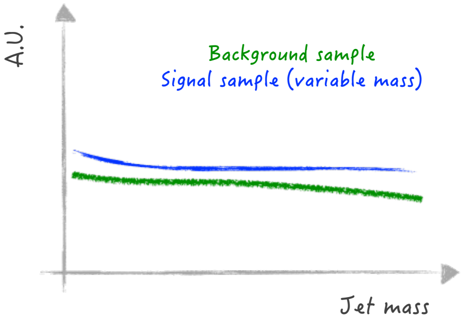


Comparison of mass decorrelation methods

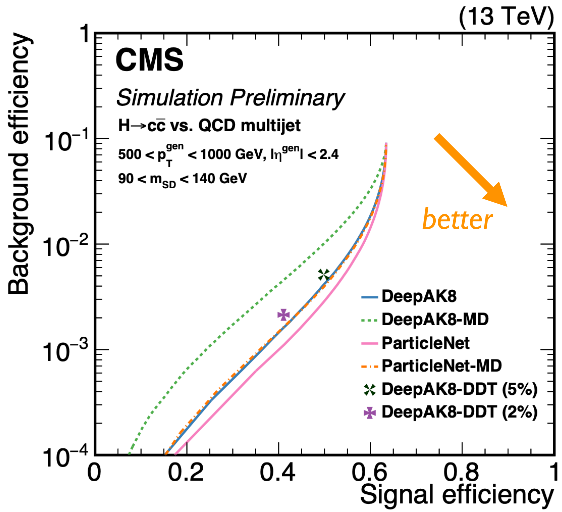
DeepAK8-MD



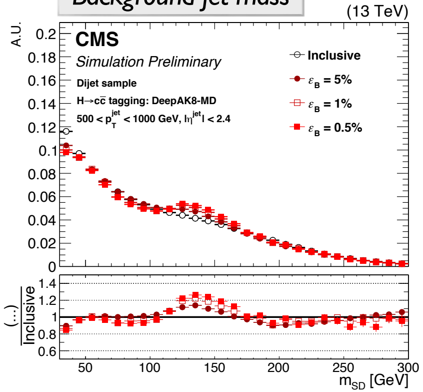
ParticleNet-MD



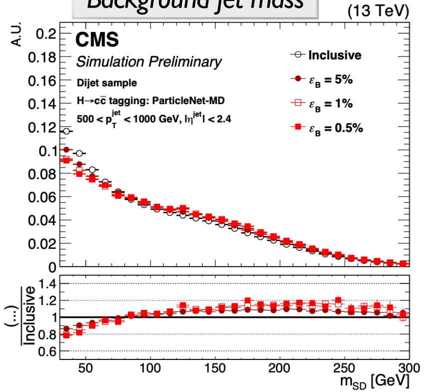
H→cc tagging performance



Background jet mass



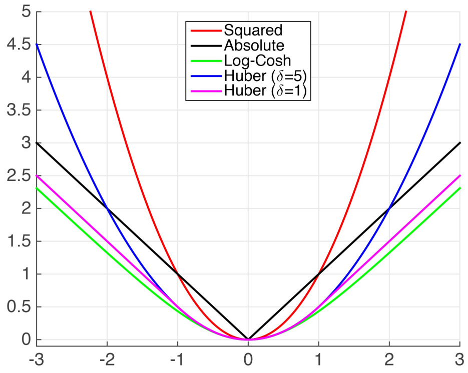
Background jet mass



Large-R jet mass regression

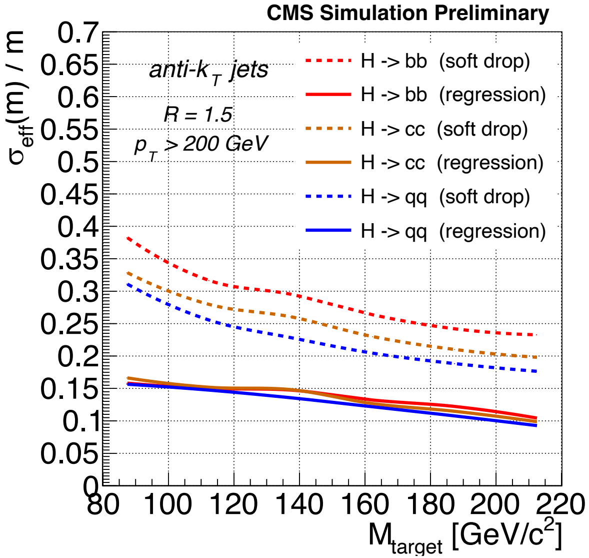
Loss function: LogCosh

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

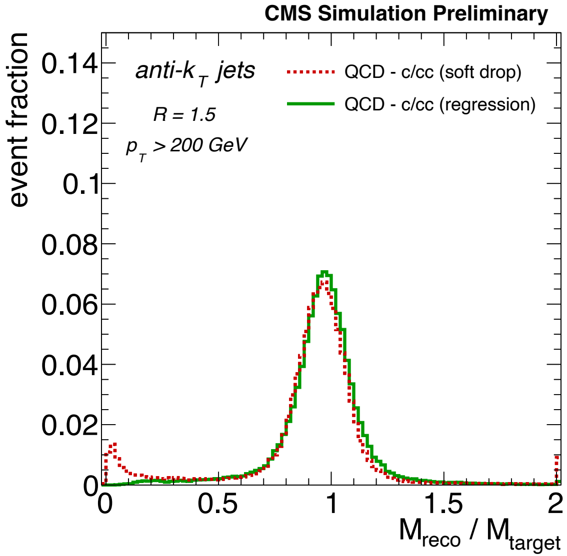


https://www.cs.cornell.edu/courses/cs4780/2015fa/web/lecturenotes/lecturenote10.html

Signal jet mass resolution

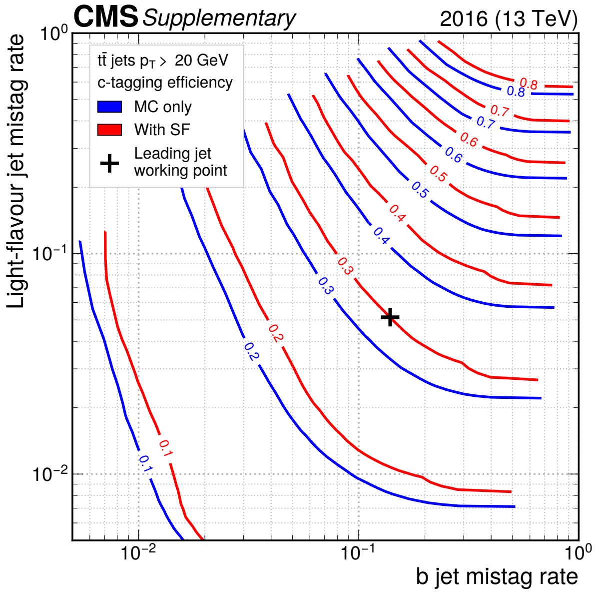


Background jet mass response

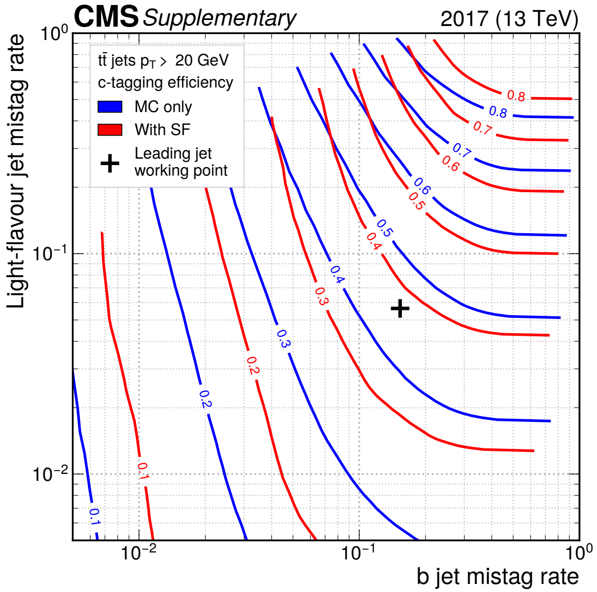


C-tagger ROC curves

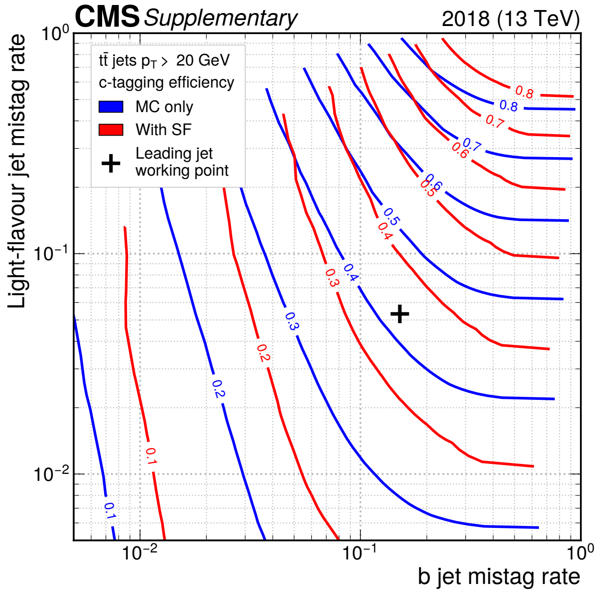
2016



2017



2018

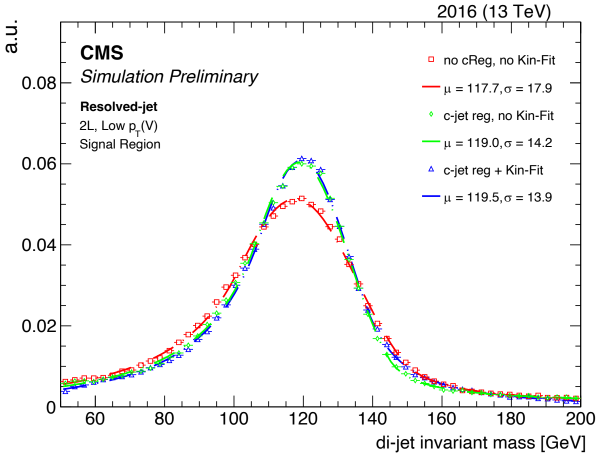


- CMS c-tagging WP: ~40% (c), ~16% (b), ~4% (light)
- ATLAS c-tagging WP [arXiv:2201.11428]: 27% (c), 8% (b), 1.6% (light)

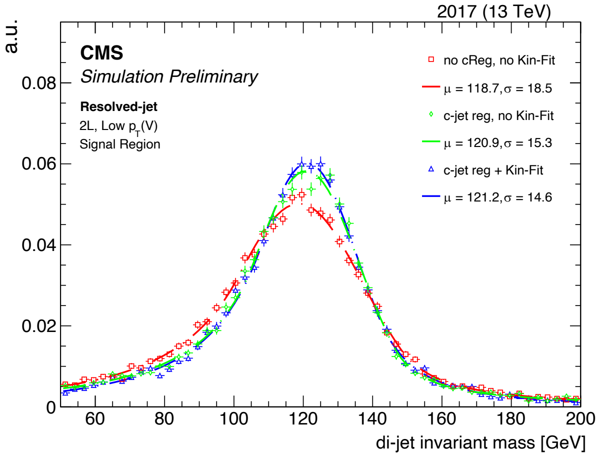
C-jet energy regression and kinematic fit

- 2-lepton Low-p_T(V) category – 60 GeV < p_T(V) < 150 GeV

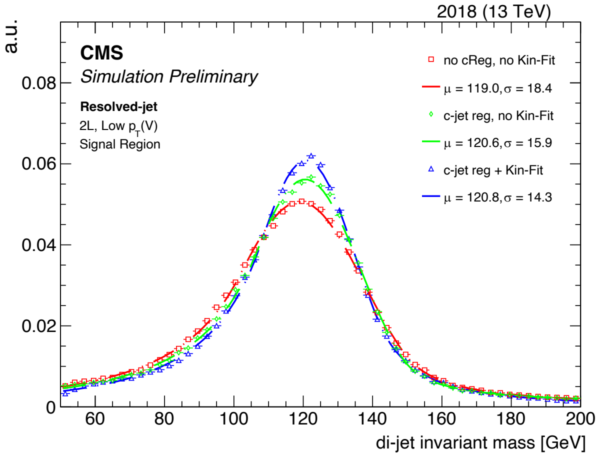
2016



2017



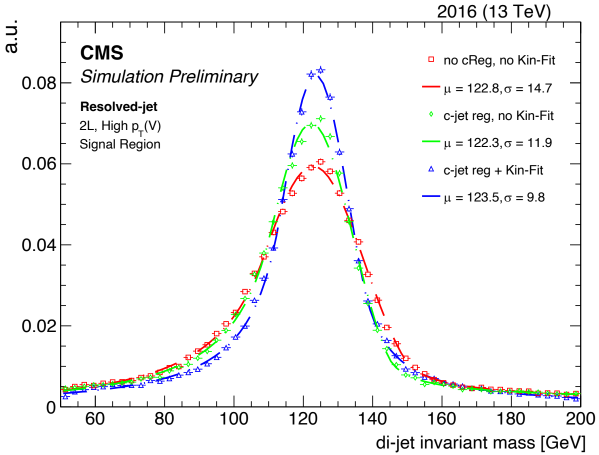
2018



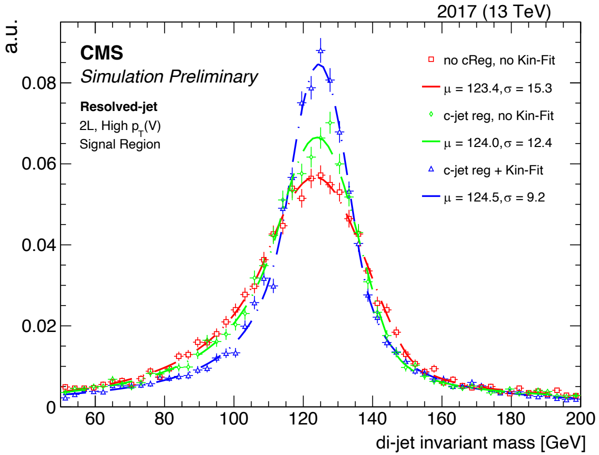
C-jet energy regression and kinematic fit

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

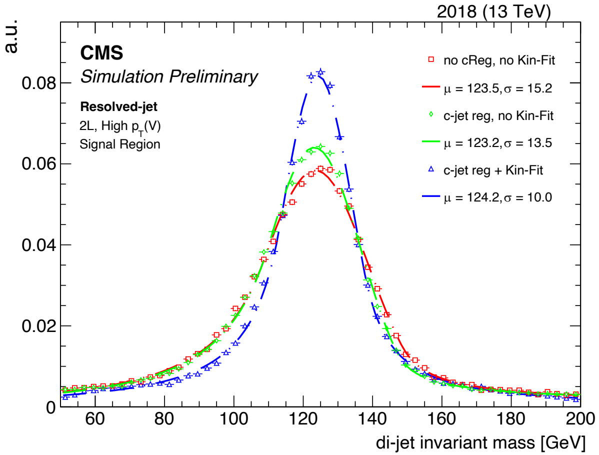
2016



2017



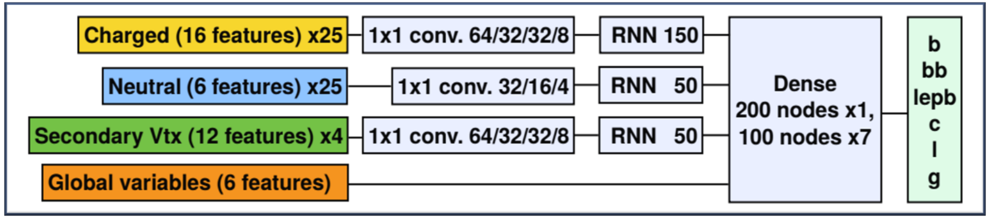
2018



Charm-tagging in the "resolved-jet" topology

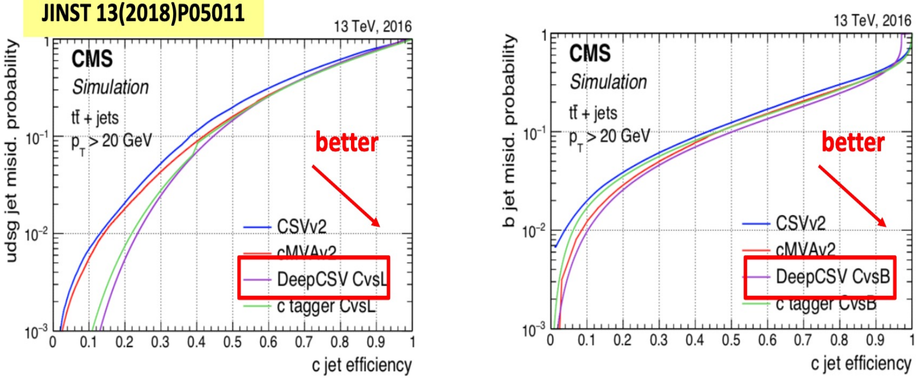
DeepJet algorithm – the cornerstone of the VH(cc) resolved-jet topology analysis

- Multiclassifier Deep Neural Network
  - Optimized for AK4-jets
  - Returns the probability for a given jet to be originated by a b-, c- or light-quark
- DNN architecture:
  - Separate 1D CNNs to process three low-level feature classes
    - For each class, concatenate multiple CNNs with decreasing dimensions
    - Compress the features to lower dimensional space
  - RNNs (LSTM type) applied after CNNs
    - Better handles the variable length sequence (PF candidates/SV)
  - Fully connected layer to connect all channels
- Input variables:
  - Properties of PF-candidates
  - Global jet features
  - Secondary vertices
- Output:
  - 6 raw scores



DeepJet architecture (from 2008.10519)

JINST 13(2018)P05011



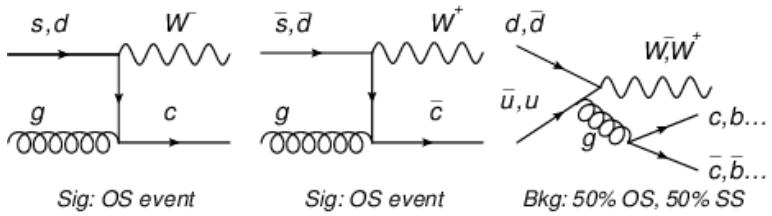
- DeepCSV: predecessor of DeepJet
- Used in the CMS VH(cc) analysis with 2016 data [JHEP **2020**,131]

A new method to calibrate charm-taggers

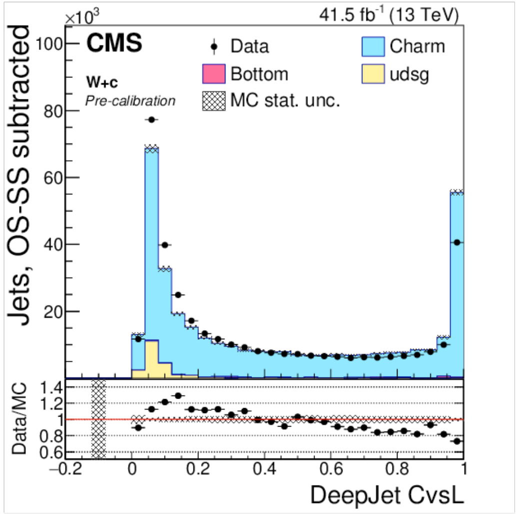
DeepJet algorithm calibration

- 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 (**accepted by *JINST***)

(from arXiv:2111.03027)

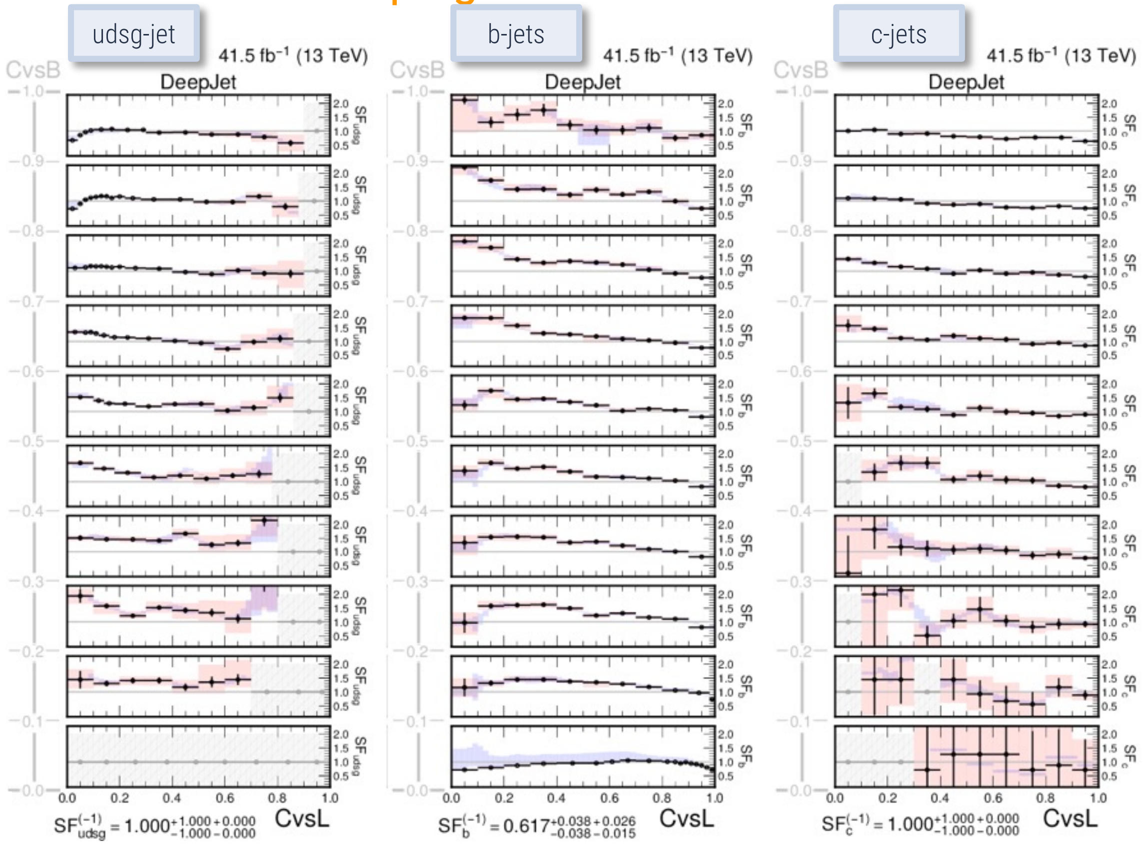


- Search for an abundant and pure source of charm-jets
  - Target W production in association with charm quarks
    - The relevant events involve a leptonically decaying W boson and a c-jet
    - These c-jets are identified using the semileptonic decay of the charmed hadrons, which produces a soft muon 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→$\mu\mu/ee$)+jets processes are considered



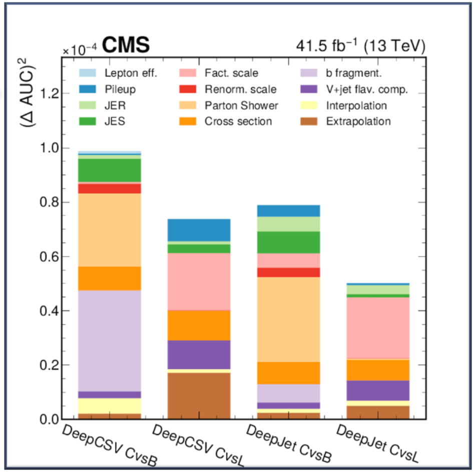
A new method to calibrate charm-taggers

Extraction of reshaping data-to-simulation scale factors



- SFs as a function of CvsL in bins of CvsB
  - Fixed bin width along CvsB and an adaptive binning scheme along CvsL (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

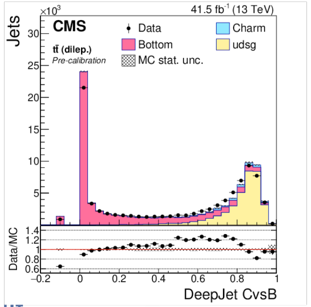
Validate robustness of the SFs derivation

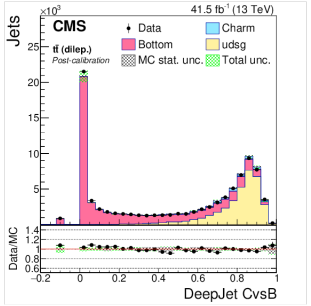
(from arXiv:2111.03027)

- ❑ Check possible bias due to the soft-μ-in-jet selection
  - SFs are derived without soft-μ 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

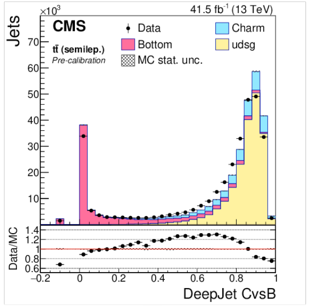
All the checks shown no bias in the SFs derivation

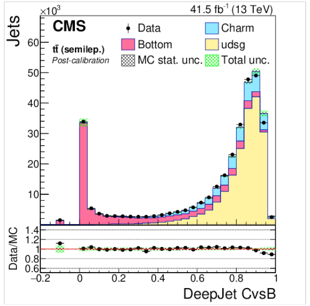
Only di-lep. tt
No soft-μ-in-jet



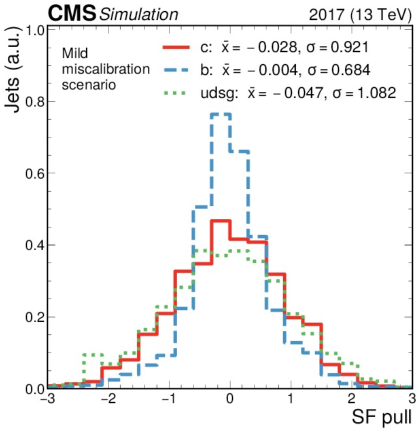


Only semi-lep. tt
No soft-μ-in-jet





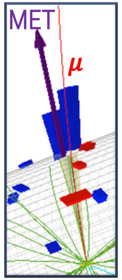
Bias test – pull distribution



A dedicated charm-jet energy regression

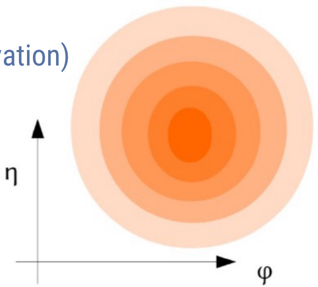
Goal: improve *c*-jet energy scale and resolution

- Inspired by b-jet energy regression [arXiv:1912.06046]
  - Jet energy measurements not always accurate:
    - loss of neutrinos, hadrons outside jet radius. Effect enhanced in c-jets and b-jets
  - Dedicated algorithm to determine c-jet energy scale and resolution
  - Algorithm pioneered for the observation of the H→bb decay mode

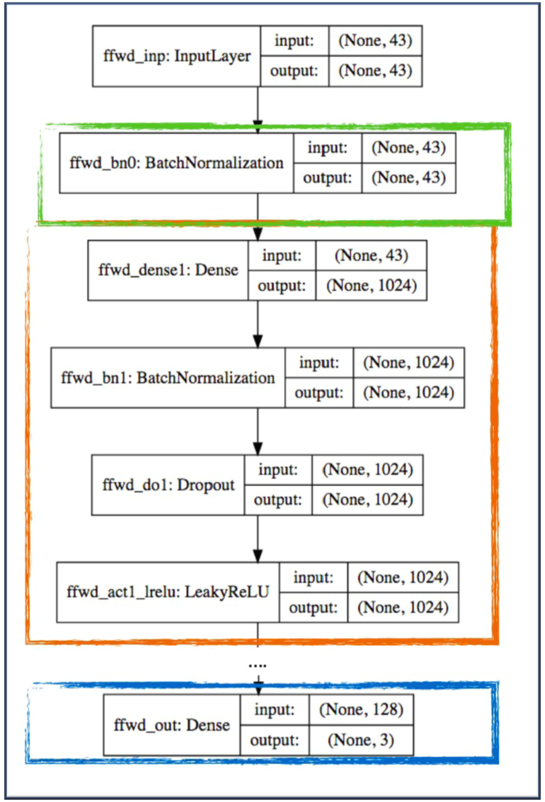


Charm-jet

- Regression performed using DNN architecture:
  - Feed-forward fully connected Deep NN (neurons with Leaky ReLu activation)
    - 6 hidden layers + batch normalization + dropout
  - Trained using c-jets collected from $W \rightarrow cq$ decays in $t\bar{t}$ MC events
  - Target is represented by $p_T(gen)/p_T(reco)$



Jet rings in dR :
(0 → 0.05 → 0.1 → 0.2 → 0.3 → 0.4)



- Input features
  - Total of 43 input variables in input to the network
  - Jets: kinematics, energy fraction, leading+soft-lepton tracks, pile-up, secondary vertexes
  - Jet energy shapes (e.g. energy fraction, etc), jet constituents, $p_T(jet)/p_T(lepton)$

Signal extraction – BDT training in SRs

Variable	Description	0L	1L	2L
$m(\mathrm{H})$	H mass	✓	✓	✓
$p_T(\mathrm{H})$	H transverse momentum	—	✓	✓
$p_T(\mathrm{V})$	vector boson transverse momentum	—	✓	✓
$m_T(\mathrm{V})$	vector boson transverse mass	—	✓	—
p_T^{miss}	missing transverse momentum	✓	✓	—
$p_T(\mathrm{V})/p_T(\mathrm{H})$	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_{Tmax}	p_T of the leading $CvsL$ jet	✓	✓	✓
p_{Tmin}	p_T of the subleading $CvsL$ jet	✓	✓	✓
$\Delta\phi(\mathrm{V,H})$	azimuthal angle between vector boson and H	✓	✓	✓
$\Delta R(j_1,j_2)$	ΔR 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	—	—	✓
$\Delta\phi(\ell_1,p_T^{miss})$	azimuthal angle between leading p_T lepton and missing transverse momentum	—	✓	—
$\Delta\eta(\ell_1,\mathrm{t})$	difference in pseudorapidity between leading p_T lepton and b-tagged jet from top quark decay	—	✓	—
$\Delta\phi(\ell_1,\mathrm{t})$	azimuthal angle between leading p_T lepton and b-tagged jet from top quark decay	—	✓	—
$\Delta R(\ell_1,\mathrm{t})$	ΔR 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	—	✓	—
$m(\mathrm{t})$	Reconstructed top quark mass	—	✓	—
$N^{aj}_{small\text{-}R}$	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	✓	✓	✓
$\Delta\eta(\mathrm{V,H})\|_{kinfit}$	difference in pseudorapidity between vector boson and H, after kinematic-fit	—	—	✓
$\Delta\phi(\mathrm{V,H})\|_{kinfit}$	azimuthal angle between vector boson and H, after kinematic-fit	—	—	✓
$m(\mathrm{H})\|_{kinfit}$	H mass after kinematic-fit	—	—	✓
$p_T(\mathrm{H})\|_{kinfit}$	H transverse momentum after kinematic-fit	—	—	✓
$p_{Tmax}\|_{kinfit}$	p_T of the leading $CvsL$ jet after kinematic-fit	—	—	✓
$p_{Tmin}\|_{kinfit}$	p_T of the subleading $CvsL$ jet after kinematic-fit	—	—	✓
$p_T(\mathrm{V})/p_T(\mathrm{H})\|_{kinfit}$	ratio between vector boson and H transverse momenta after kinematic-fit	—	—	✓
$\sigma(\mathrm{H})\|_{kinfit}$	H invariant mass resolution from kinematic fit	—	—	✓

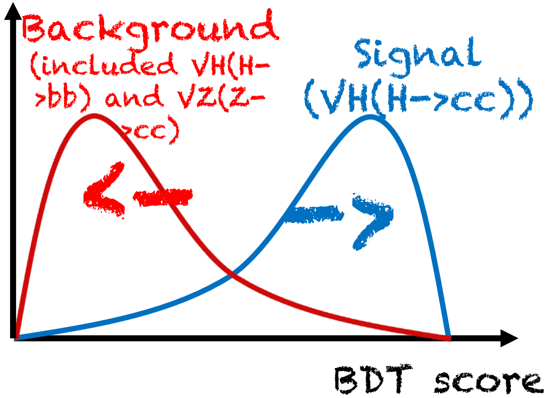
Higgs and vector boson properties

c-tagging score

event kinematics

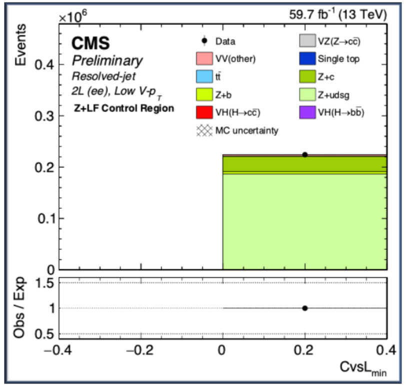
Kinfit Variables (2L only)

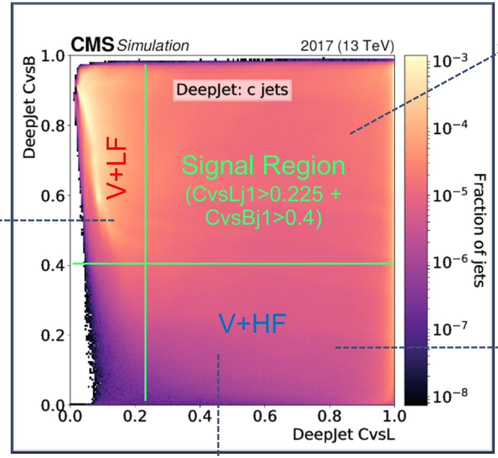
- ❑ BDT trained to separate signal from background samples
  - Use combination of kinematic observables and particle flavor variables (tagger informations)
- ❑ Separate BDTs trained for each channel and data taking year
  - Separate BDTs trained for high- and low-p_T(V) 2L
  - Variables used dependent on channel
- ❑ Reshaped BDT distribution used in SR during final fit



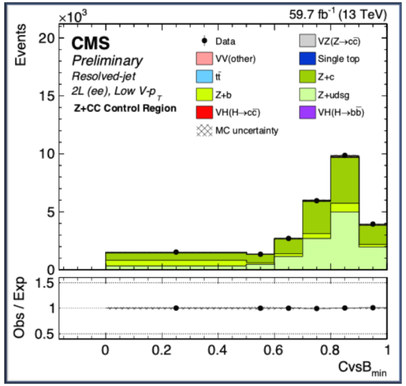
Background estimation – Resolved-jet

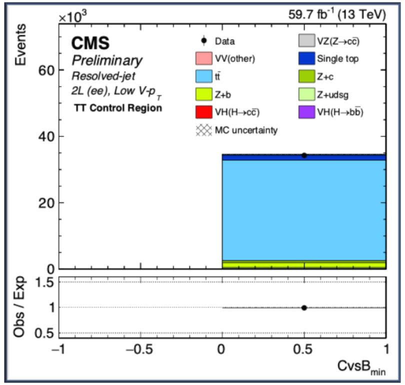
- Accurate modeling of jet flavor in V+Jet background is vital for proper signal extraction
  - Separate rate parameters for V+c, V+b, and V+light processes (no W+b)
  - Additional rate parameter for $t\bar{t}$ background



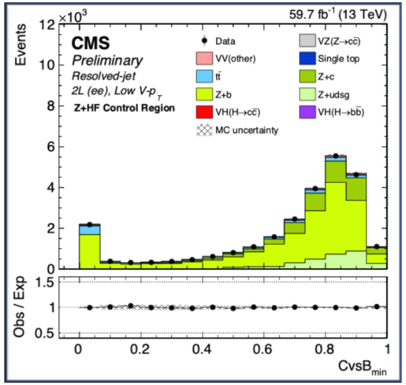


V+CC
⇕
Veto m(H) region





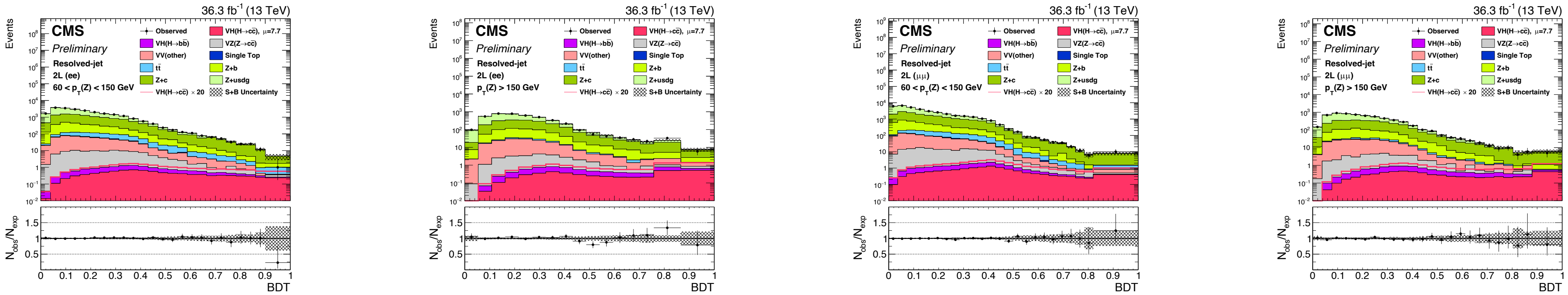
$t\bar{t}$
⇕
Invert Z mass (2L)
Require add jet (1L)*
Require add ℓ and jets (0L)

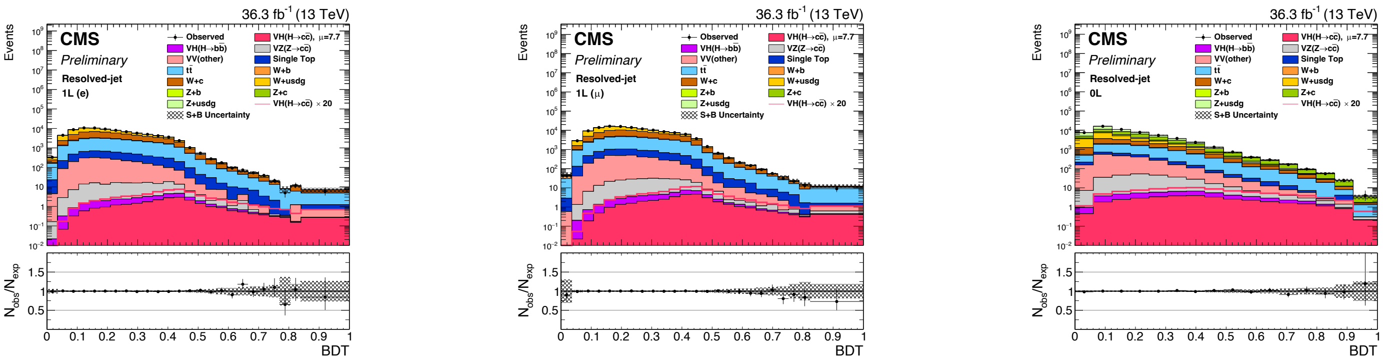


*1L: also require MET<170 GeV to keep orthogonal to 0L $t\bar{t}$ CR

Postfit plots – Signal regions - 2016

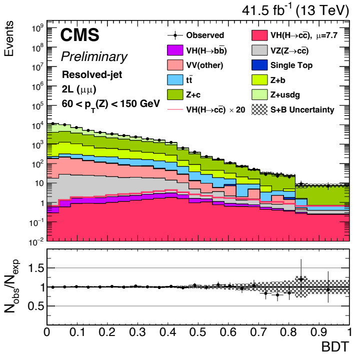
- Postfit distribution of the BDT discriminant obtained with the 2016 data
  - 7 Signal regions in each year: 2L(ee/μ) Low-p_T(V) and –High-p_T(V), 1L(e/μ) and 0L

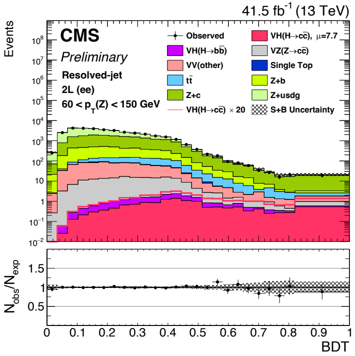


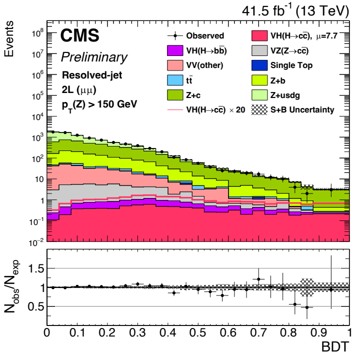


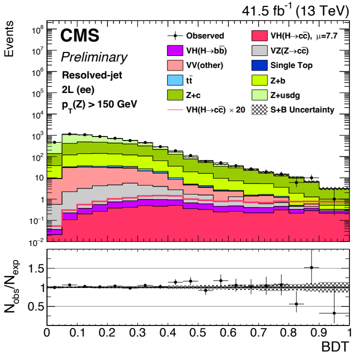
Postfit plots – Signal regions - 2017

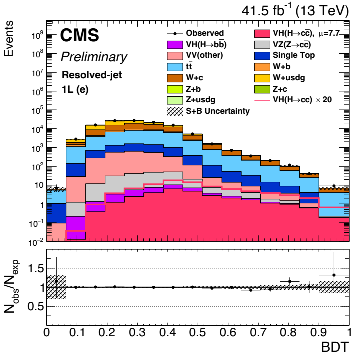
- Postfit distribution of the BDT discriminant obtained with the 2017 data
  - 7 Signal regions in each year: 2L(ee/μ) Low-p_T(V) and –High-p_T(V), 1L(e/μ) and 0L

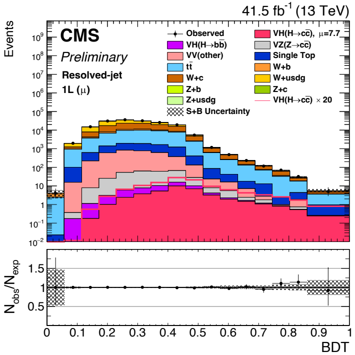


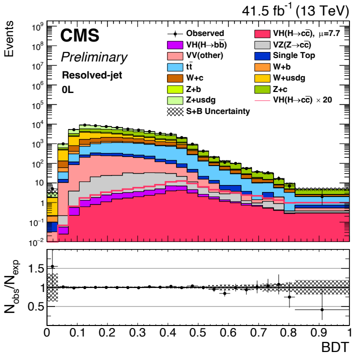






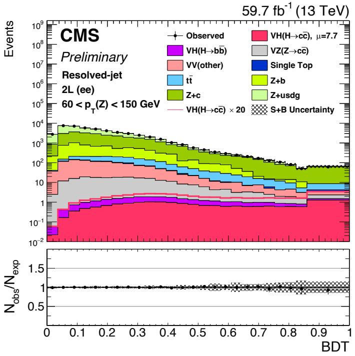


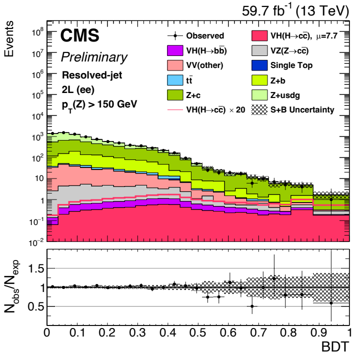


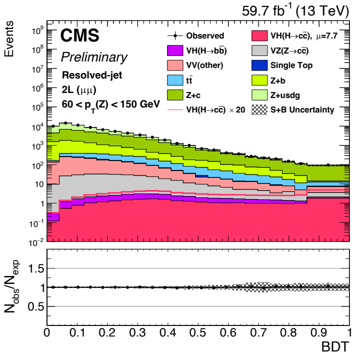


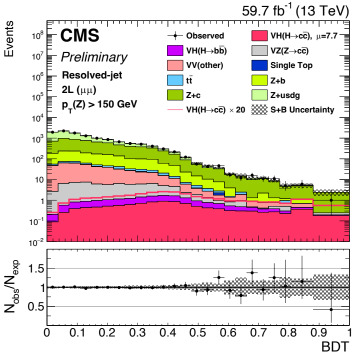
Postfit plots – Signal regions - 2018

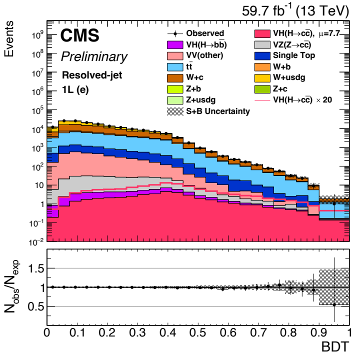
- Postfit distribution of the BDT discriminant obtained with the 2018 data
  - 7 Signal regions in each year: 2L(ee/μ) Low-p_T(V) and –High-p_T(V), 1L(e/μ) and 0L

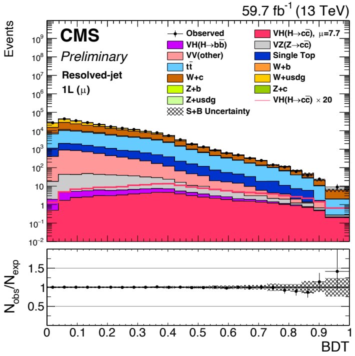


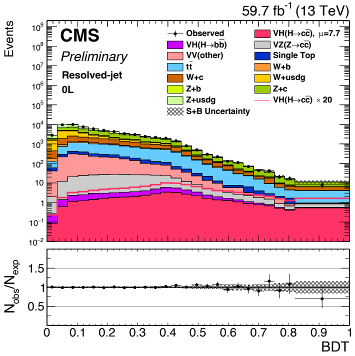








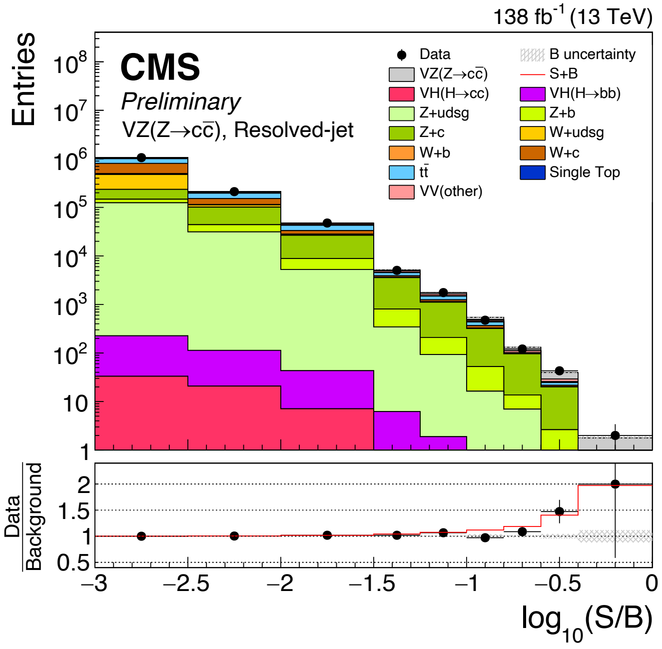




Resolved-jet topology - results

- Resolved-jet – all categories: ordering the events by $\log_{10}(S/B)$

VZ(cc)



VH(cc)