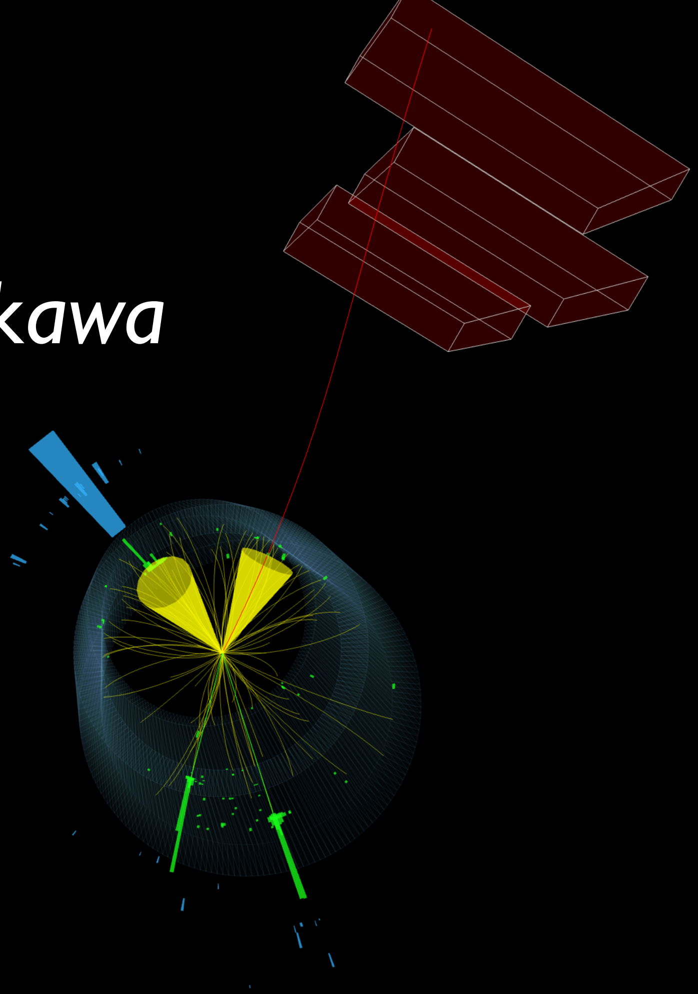
Constraining the Higgs-charm Yukawa coupling in the CMS experiment

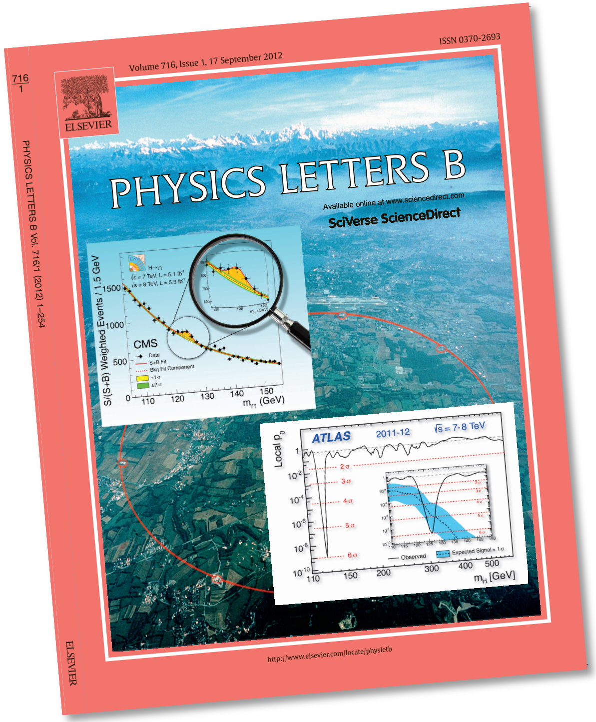


曲慧麟 (CERN)

USTC Seminar
2022.05.13

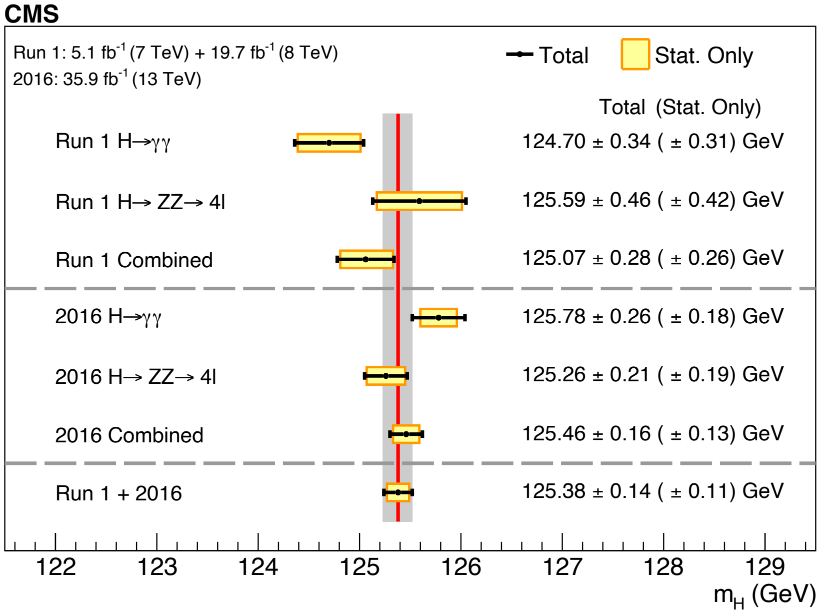
Introduction

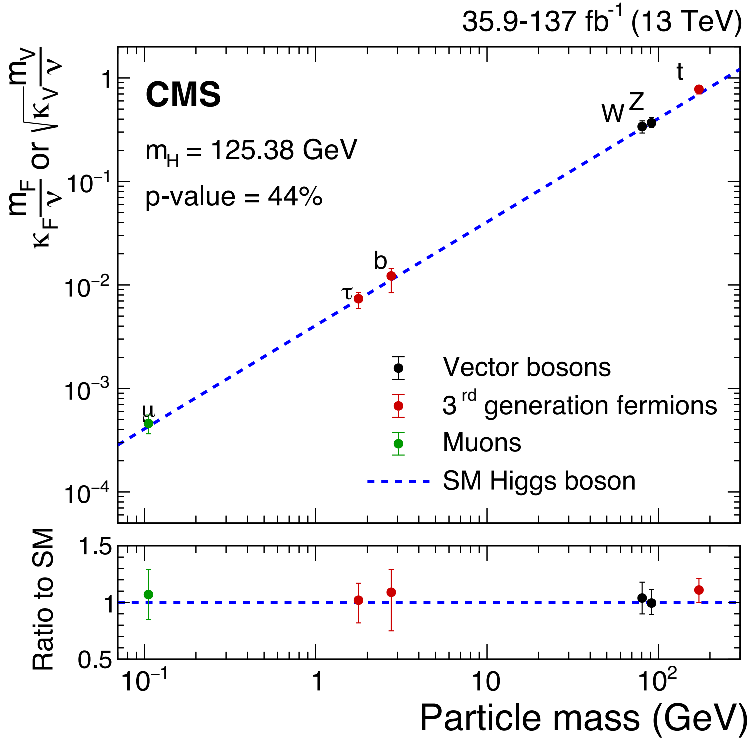
- Discovery of the Higgs boson in 2012: A new chapter of particle physics



Understanding the Higgs boson

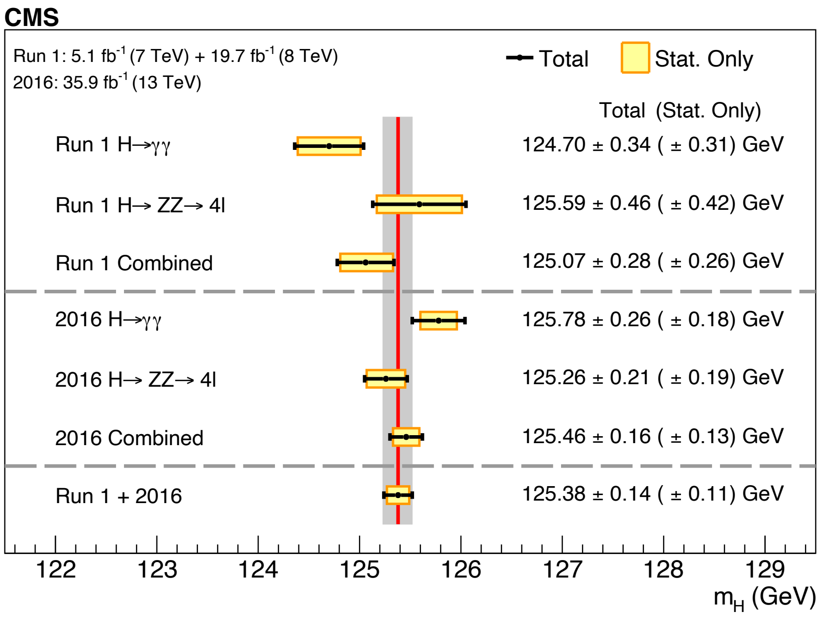
- Tremendous progress in our understanding of the Higgs boson in the past ten years

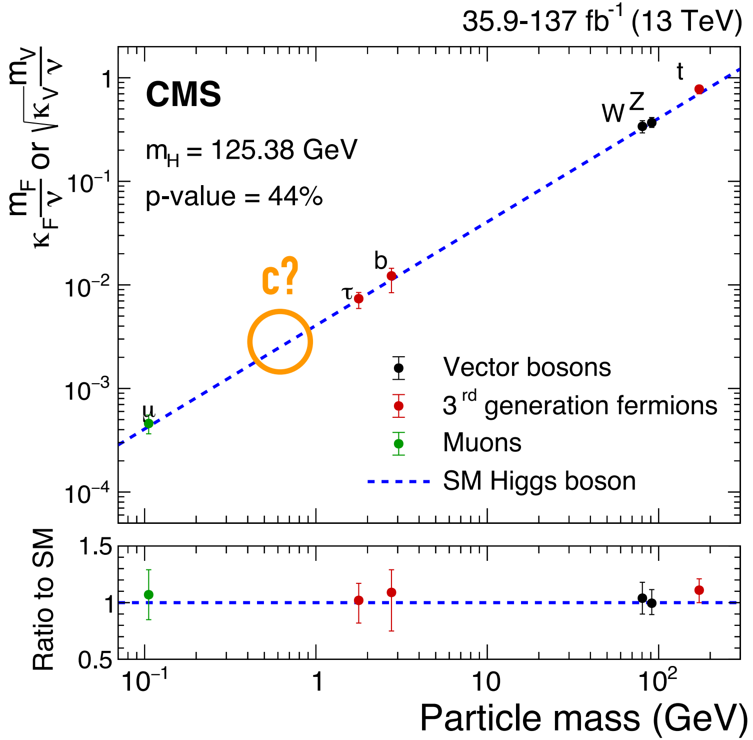




How charming is the Higgs boson?

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



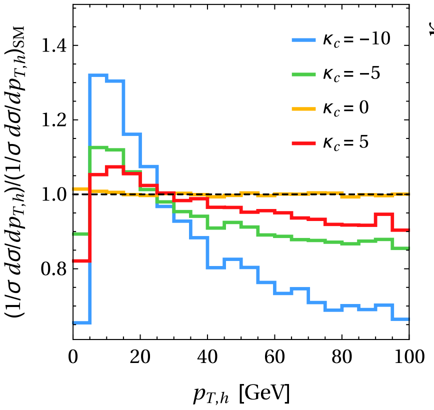


Probing the Higgs-charm coupling

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

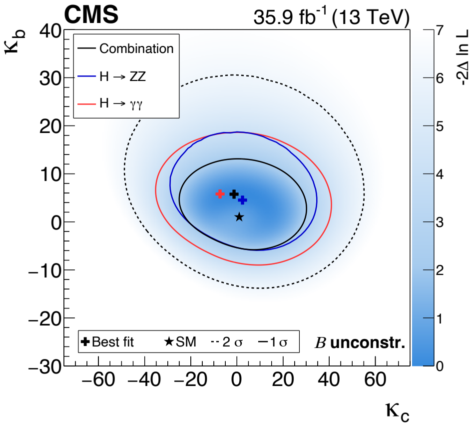
Indirect constraint from Higgs kinematics

Phys. Rev. Lett. 118, 121801



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

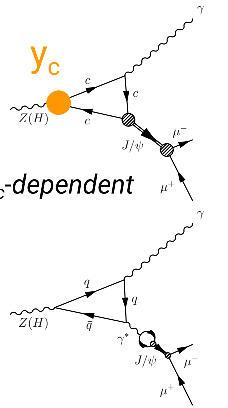
Phys. Lett. B 792 (2019) 369



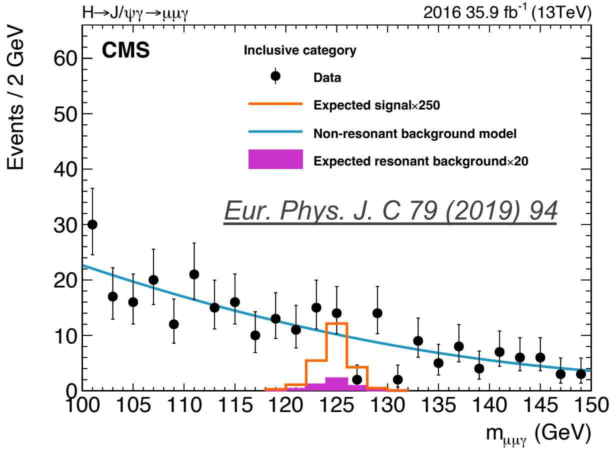
-33 < κ_c < 38 (obs.)

-31 < κ_c < 36 (exp.)

Search for exclusive H → J/Ψγ decays



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.)

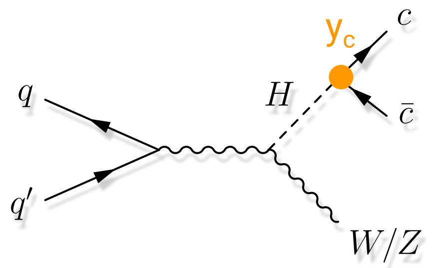
$\mathcal{B}$(H → J/Ψγ) < 170x SM(exp.)

Roughly translates to κ_c < O(100)

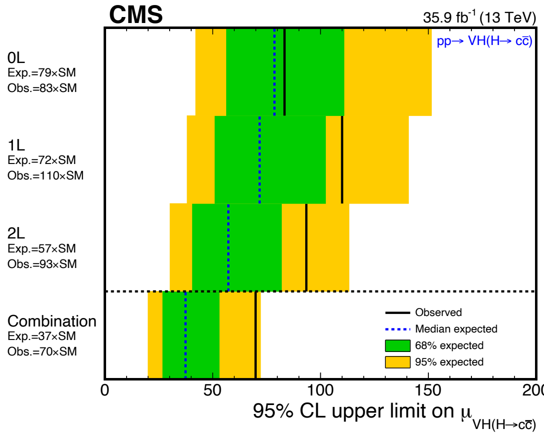
Corresponding ATLAS analyses: ATLAS-CONF-2022-002; Phys. Lett. B 786 (2018) 134.

Direct search for H → cc

- Search for H → 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 → νν (0L), W → ℓν (1L), Z → ℓℓ (2L) [ℓ = e, μ]
- Main backgrounds
  - V + jets, single and pair production of top quarks, dibosons
  - VH(H → 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})** *arXiv:2205.05550 (submitted to PRL)*



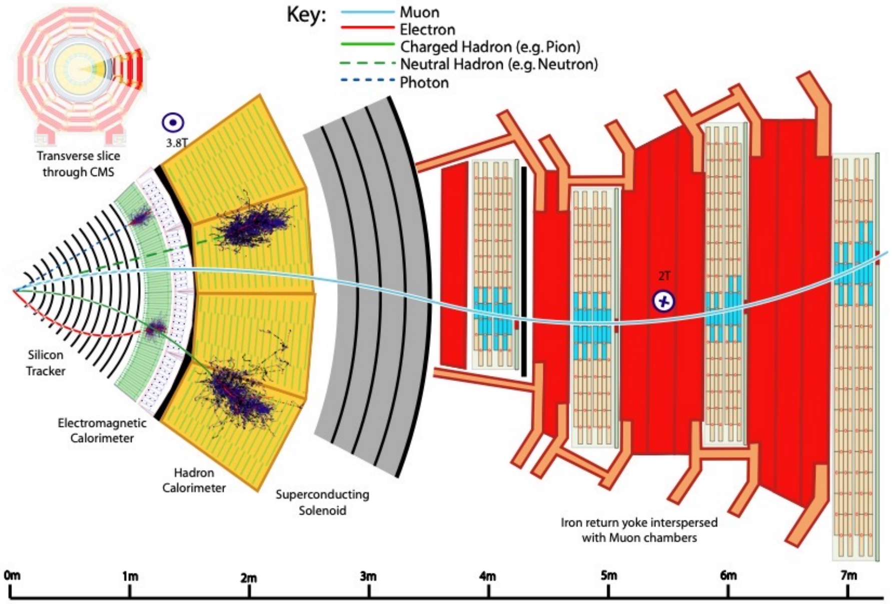
JHEP 03 (2020) 131



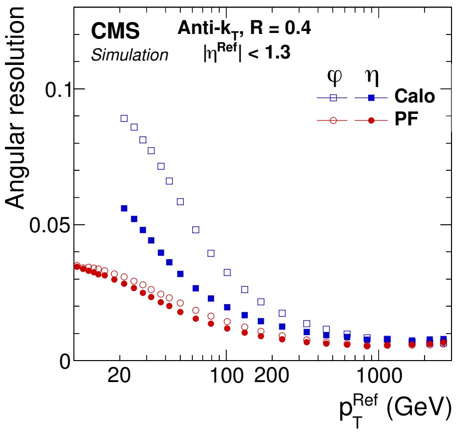
Corresponding ATLAS analysis: arXiv:2201.11428. See also recent LHC seminar by A. Chisholm.

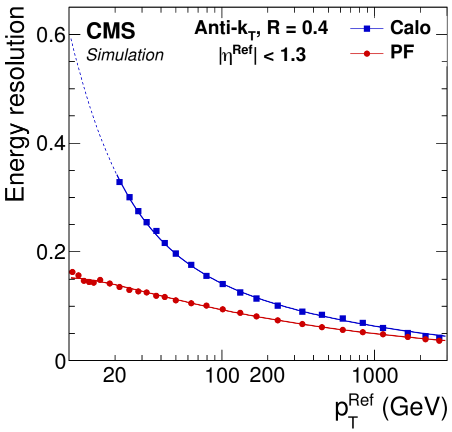
Particle-flow reconstruction

- 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



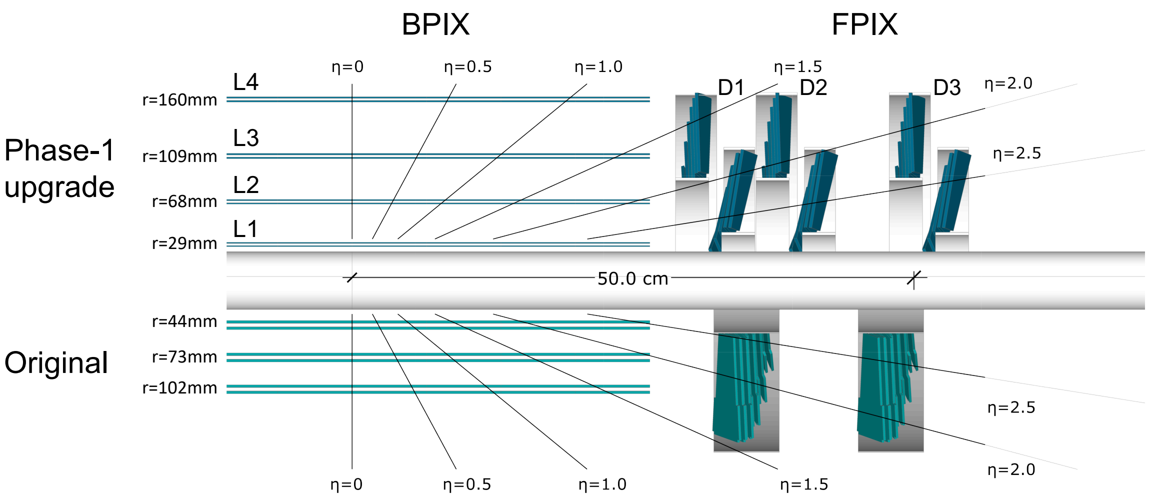
JINST 12 (2017) P10003



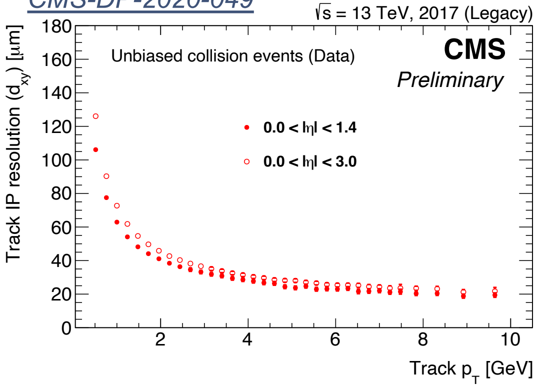


Phase-1 pixel detector upgrade

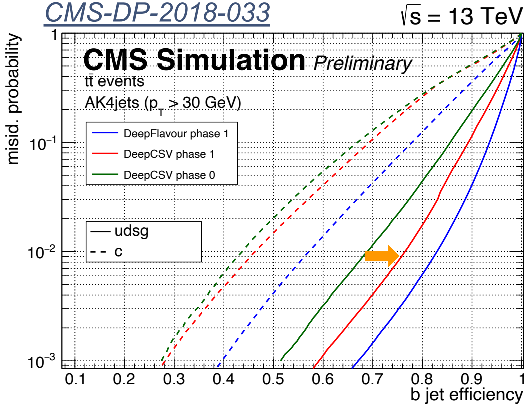
- New pixel detector installed during year-end stop 2016/2017



CMS-DP-2020-049



CMS-DP-2018-033

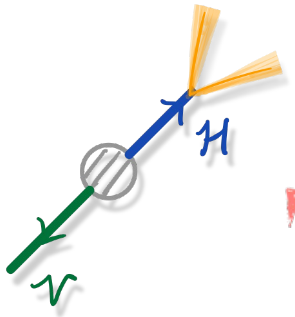


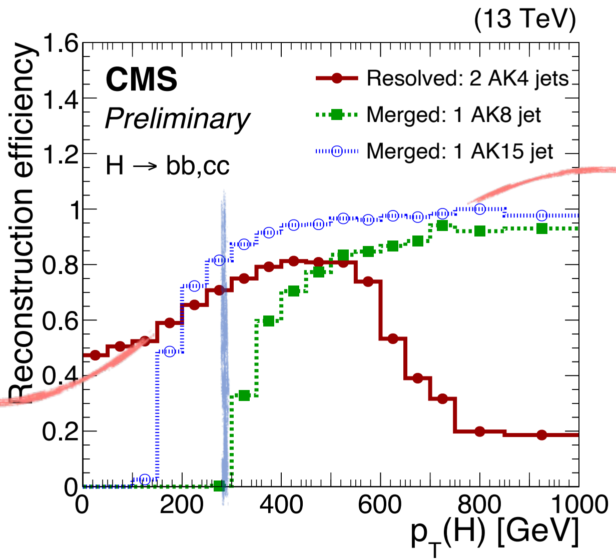
Improved tracking and flavour tagging performance in the 2017 — 2018 data set!

Analysis overview

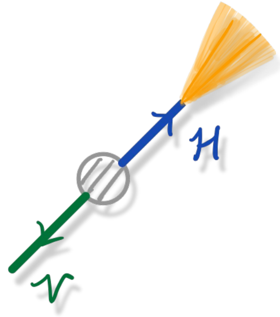
- Two complementary approaches for Higgs boson candidate reconstruction

“Resolved-jet”





“Merged-jet”

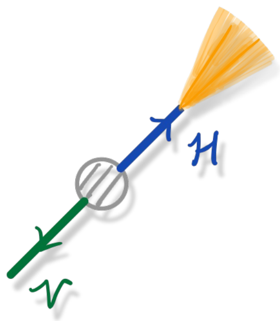


$\Delta R(c, c) \sim 2m(H)/p_T(H)$

- **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



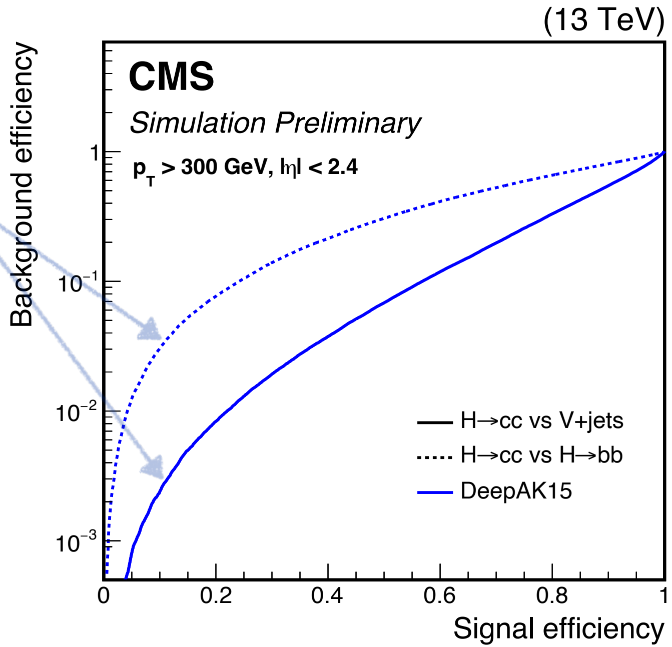
H → cc identification

- 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



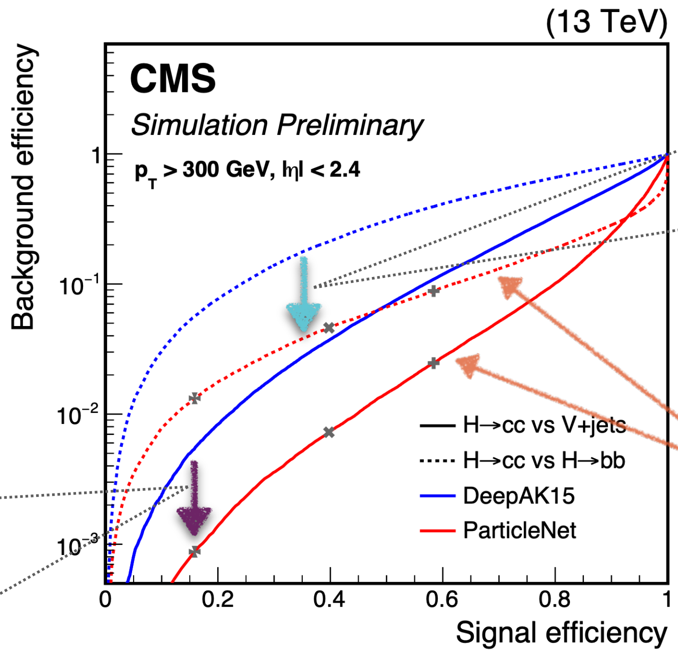
H → cc identification

- ❑ 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



~5x better
H→bb rejection

~5x better
V+jet rejection

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

ParticleNet [CMS-DP-2020-002]

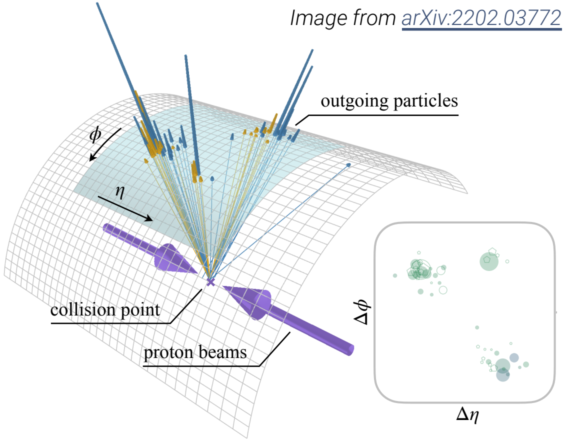
- same spirit as DeepAK8, but substantially improved:
- **graph neural network architecture**
- **novel mass decorrelation technique**

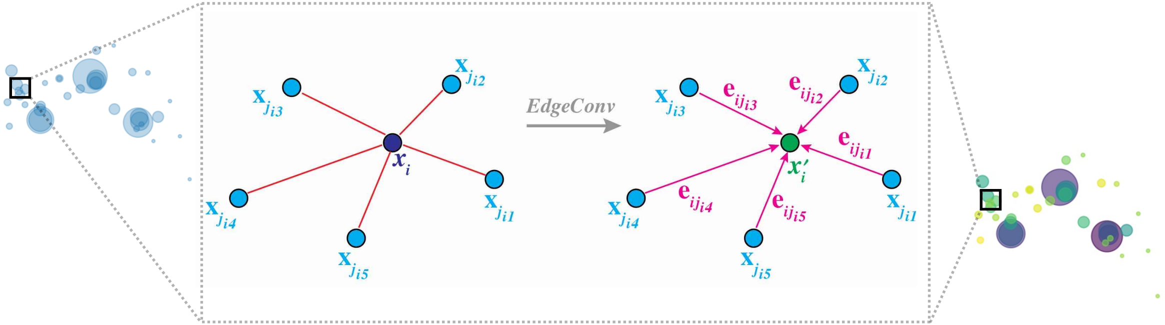
>2x improvement in the final sensitivity

ParticleNet architecture

- New jet representation: "particle cloud"
  - treating a jet as an unordered set of particles, distributed in the η − φ space
- ParticleNet [H. Qu and L. Gouskos, 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



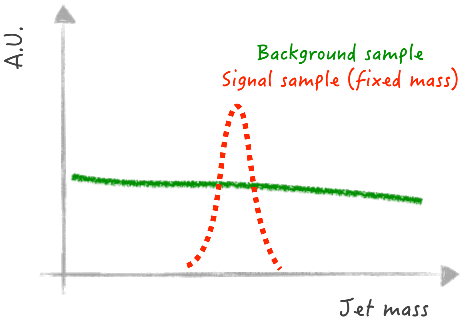


Performance on top quark tagging benchmark
[SciPost Phys. 7, 014 (2019)]

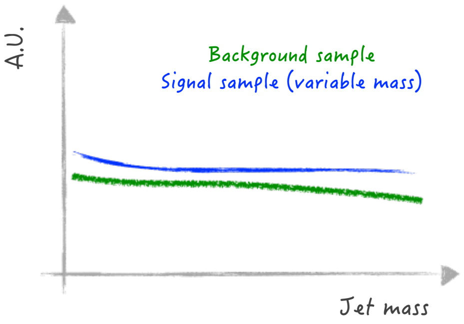
	$1/\varepsilon_b$ at $\varepsilon_s = 30\%$
ResNeXt-50	1147 ± 58
P-CNN	759 ± 24
PFN	888 ± 17
ParticleNet-Lite	1262 ± 49
ParticleNet	**1615 ± 93**

Mass decorrelation

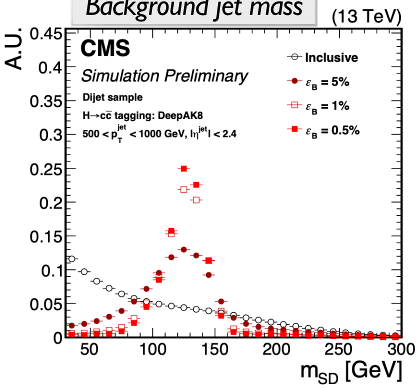
Plain training: no mass decorrelation



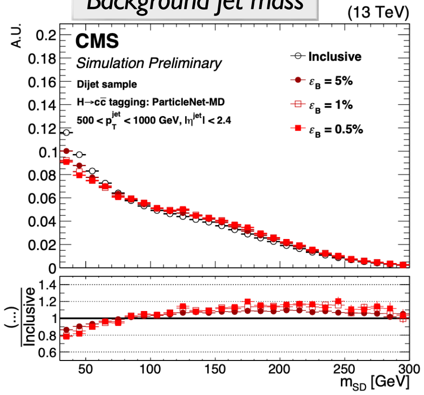
Mass-decorrelated training



Background jet mass



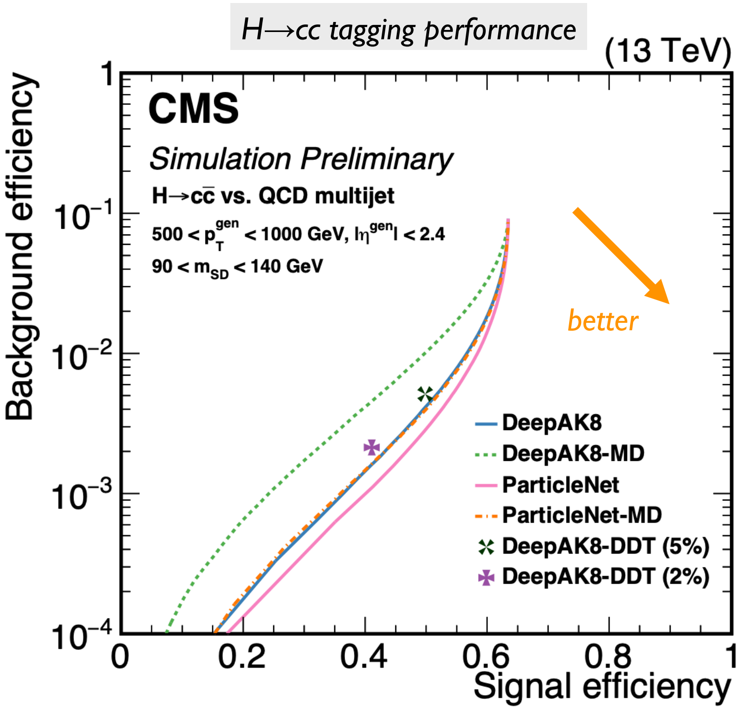
Background jet mass



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)



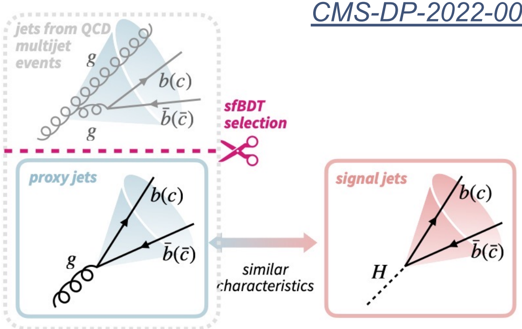
CMS-DP-2020-002

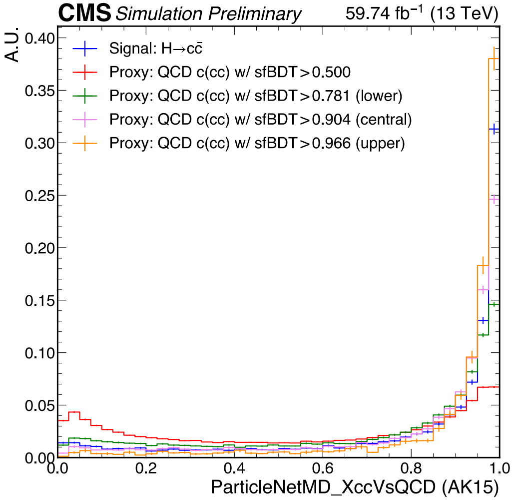
- ❑ "Mass sculpting": background jet mass shape becomes similar to signal after tagger selection
- ❑ New approach to prevent mass sculpting
  - using a special signal sample for training
    - hadronic decays of a spin-0 particle X
      - $X \to bb$, $X \to cc$, $X \to qq$
    - not a fixed mass, but 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%

CMS-DP-2022-005

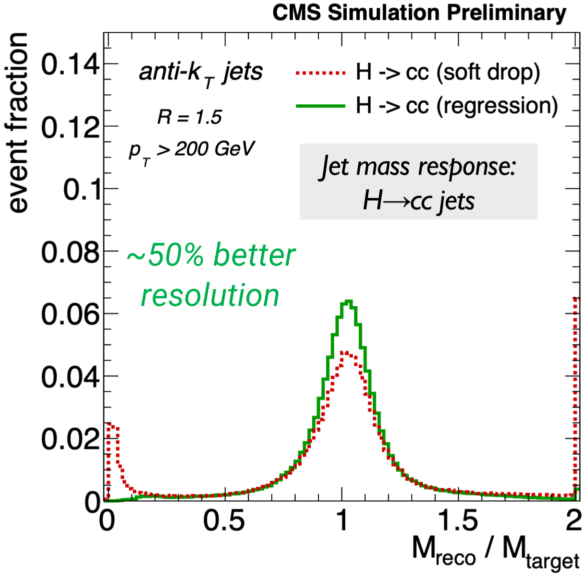


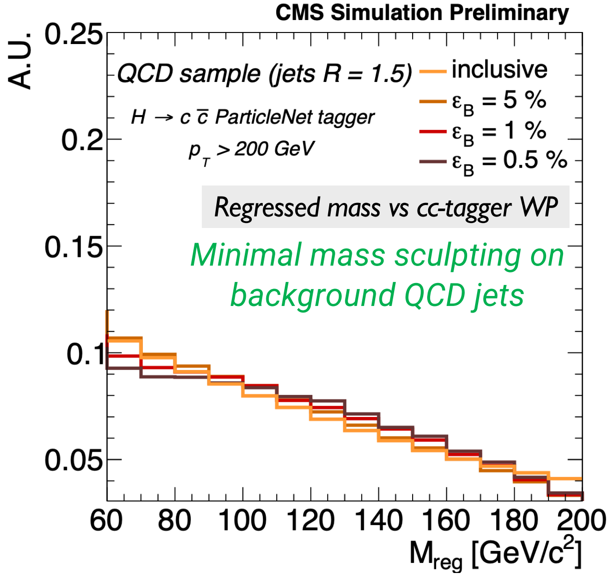


Large-R jet mass regression

CMS DP-2021/017

- 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 → 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

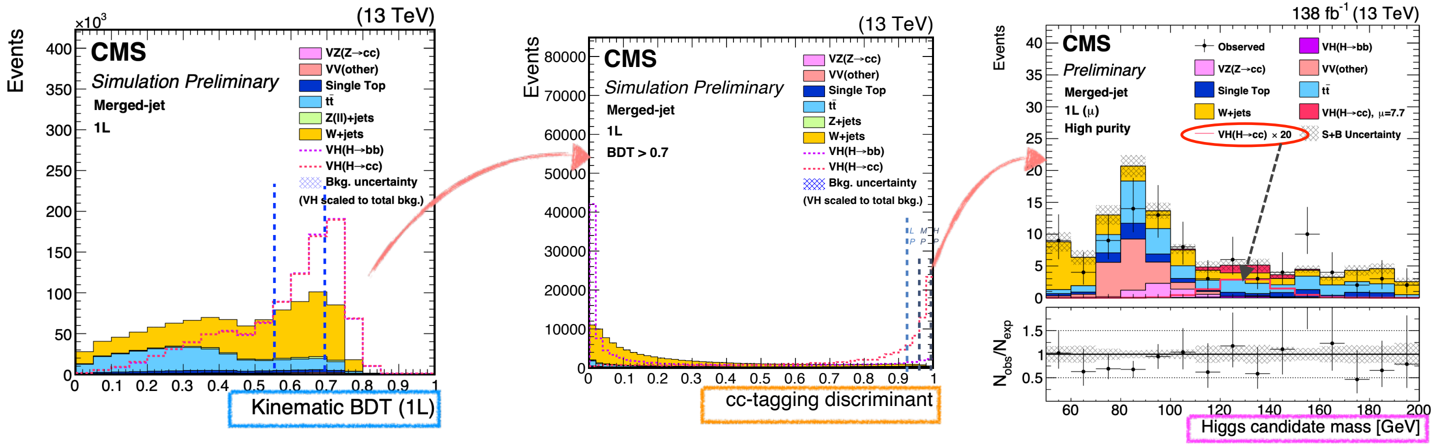




20 – 25% improvement in the final sensitivity

Analysis strategy

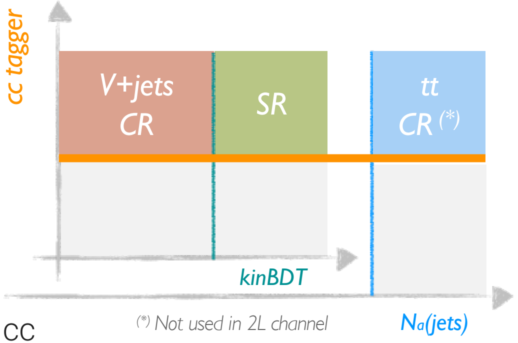
- Factorized approach for analysis design
  - event-level **kinematic BDT** developed in each channel to better suppress main backgrounds (V+jets, tt)
    - using only *event kinematics*, no intrinsic properties (e.g., mass/flavor) of the large-R jet
  - **ParticleNet cc-tagger** then used to define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

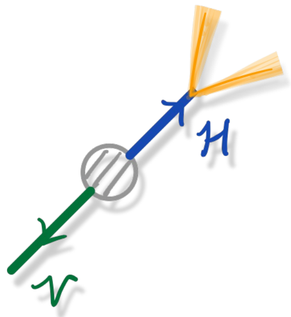
- Normalizations of main backgrounds estimated via dedicated data control regions (CRs)
  - V+jets CR: use the low kinematic BDT region
  - tt CR (0L & 1L): invert the cut on the number of additional small-R jets (i.e., $N_{aj} \geq 2$)
  - free-floating parameters scale the normalizations in CRs and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

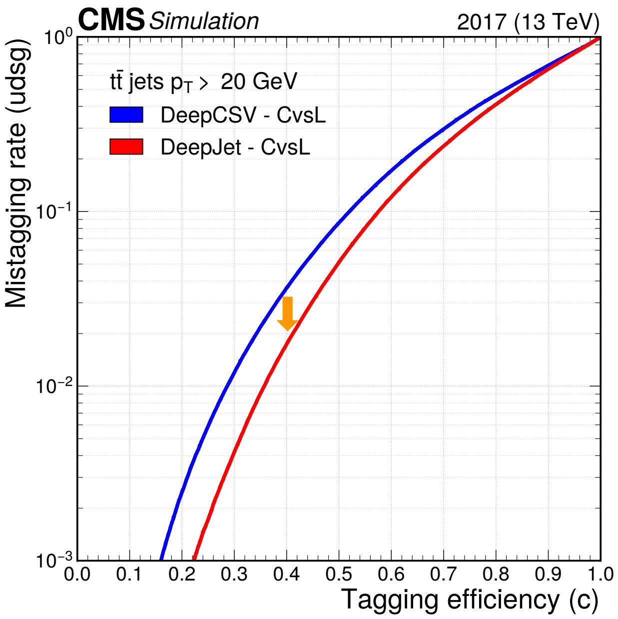
Resolved-jet topology

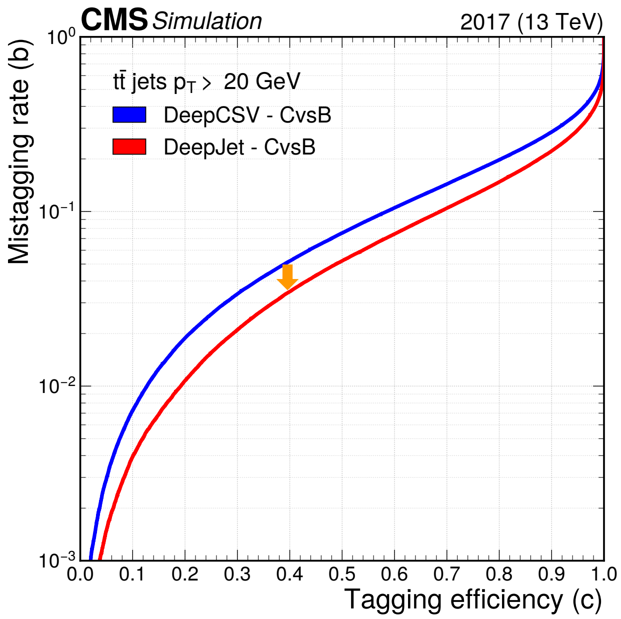


Charm quark identification

- 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

JINST 17 (2022) P03014

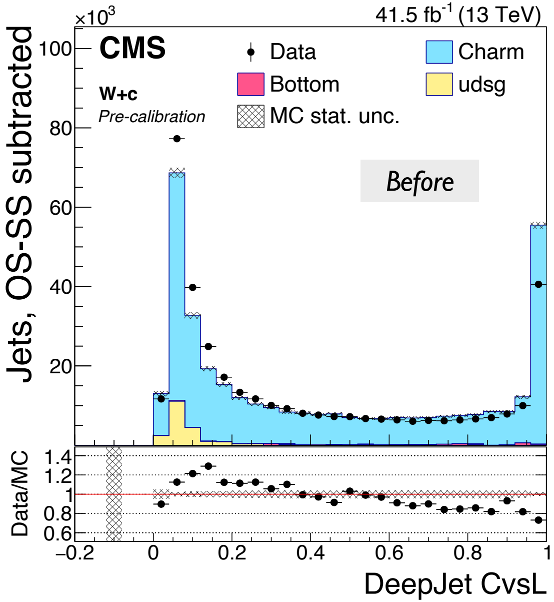


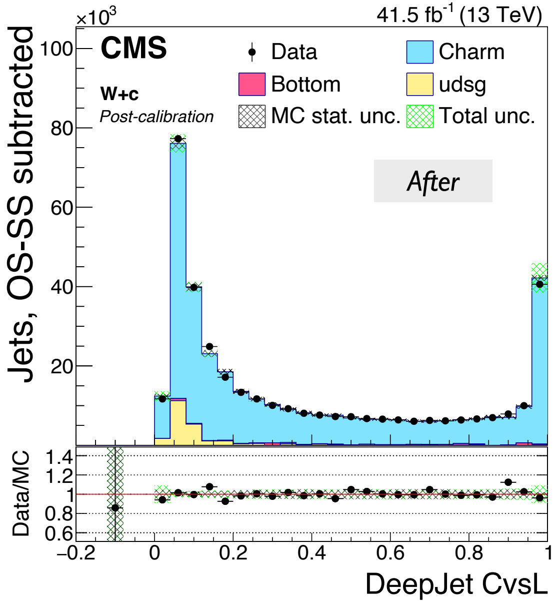


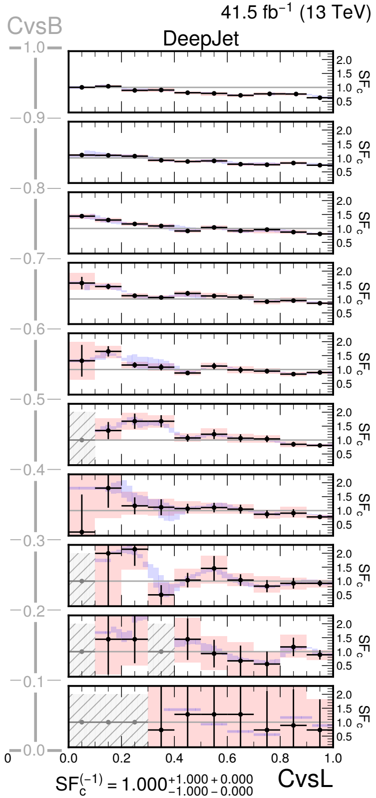
Charm tagging calibration

JINST 17 (2022) P03014

- Novel calibration method to correct the entire distributions of the c-tagging discriminants
  - per-jet SFs derived as a function of (CvsL, CvsB | truth flavor)
- SFs derived with an iterative approach using three samples
  - Z(ℓℓ) + jets (light jet enriched); W + c (c-jet enriched); tt (b-jet enriched)







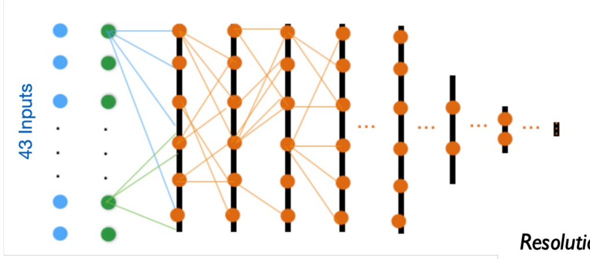
Charm jet energy regression

- 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→bb analyses
  - re-trained for c jets instead of b jets
    - c jets collected from W→cx 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

Inputs

43 high-level features:

- jet kinematics
- jet composition
- pileup information
- semi-leptonic decays
- secondary vertex



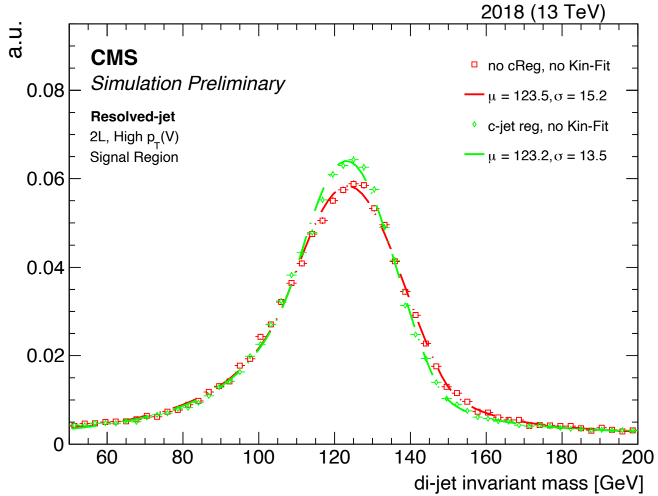
Outputs

- $\hat{y} = p_T(\text{gen})/p_T(\text{reco})$
- $\hat{y}_{0.25}$: 25% quantile
- $\hat{y}_{0.75}$: 75% quantile

Resolution estimator: $\hat{\sigma} \equiv \frac{1}{2}(\hat{y}_{75\%} - \hat{y}_{25\%})$

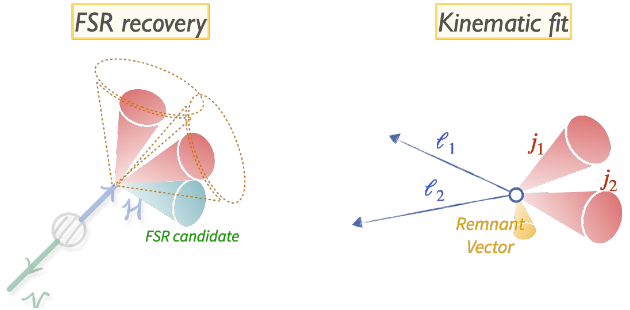
Joint loss function for correction (Huber) and resolution (quantiles) :

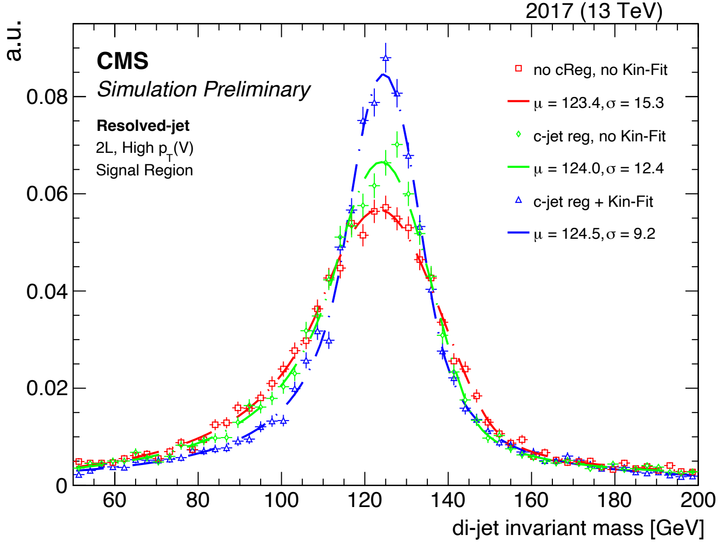
$$\boxed{Loss = Huber(y, F(x)) + \rho_{0.75}(y - F(x)) + \rho_{0.25}(y - F(x))}$$



Higgs boson candidate reconstruction

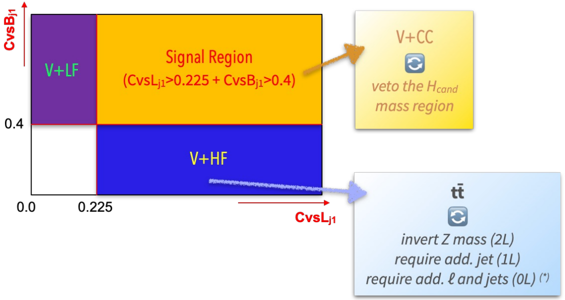
- Higgs boson candidate reconstructed using the two small-R jets with highest CvsL scores
- To improve the Higgs candidate mass resolution:
  - recovery of final state radiations (FSR)
    - additional jets within $\Delta R < 0.8$ from either of the two selected jets are included in the calculation of 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 $Z \to \ell\ell$ system
    - up to 30% improvement in Higgs candidate mass resolution



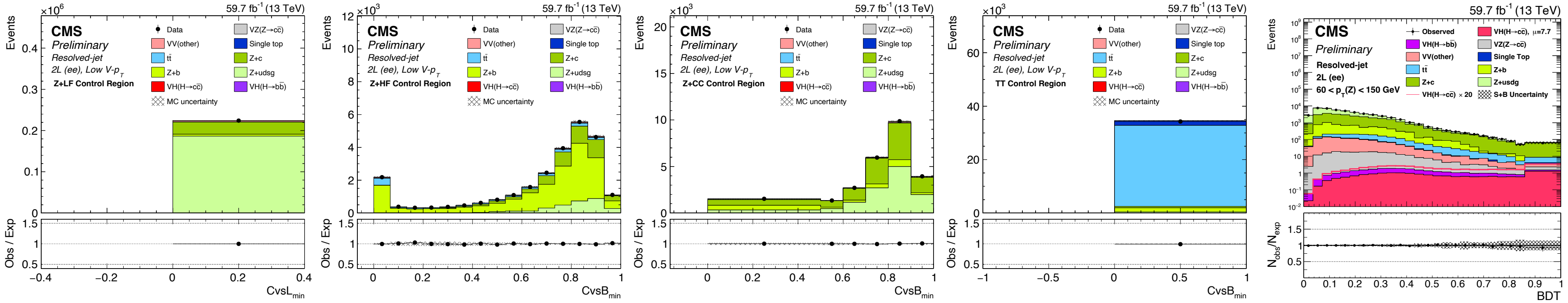


Signal extraction strategy

- 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 + jets, tt)
    - V + jets split based on flavor: V + b, V + c, V + udsg



- Simultaneous fit of SRs (BDT shapes) and CRs (c-tagging discriminants) for signal extraction

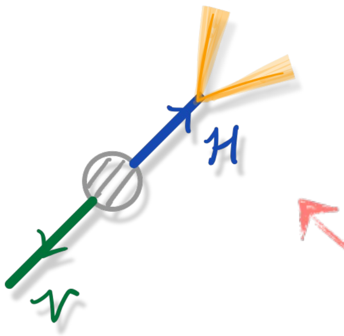


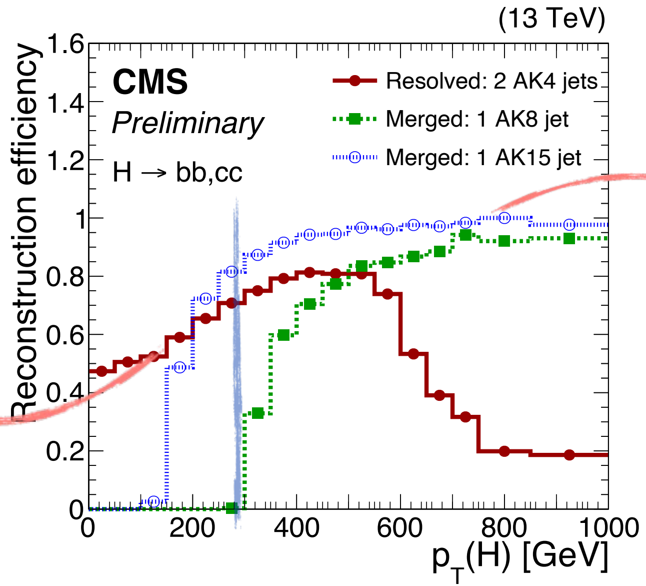
Results

Combination of the two topologies

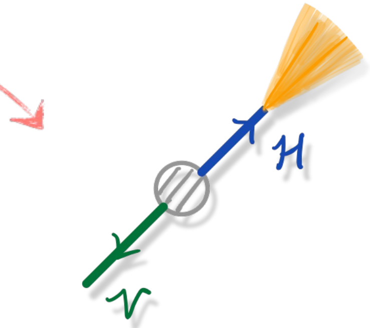
- The two topologies are made orthogonal via the presence of large-R jet with p_T > 300 GeV
  - p_T threshold chosen to maximize expected sensitivity

"Resolved-jet"





"Merged-jet"



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

Uncertainties

- Systematic uncertainties correlated between topologies, except:
  - background normalizations for V+jets and tt
  - charm quark identification efficiencies
- Main uncertainties
  - limited statistics of the data set
  - size of simulated samples (especially NLO V+jets)
  - charm quark identification efficiencies

Relative contributions to the total uncertainty on μ

Uncertainty source	$\Delta\mu/(\Delta\mu)_{tot}$
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%

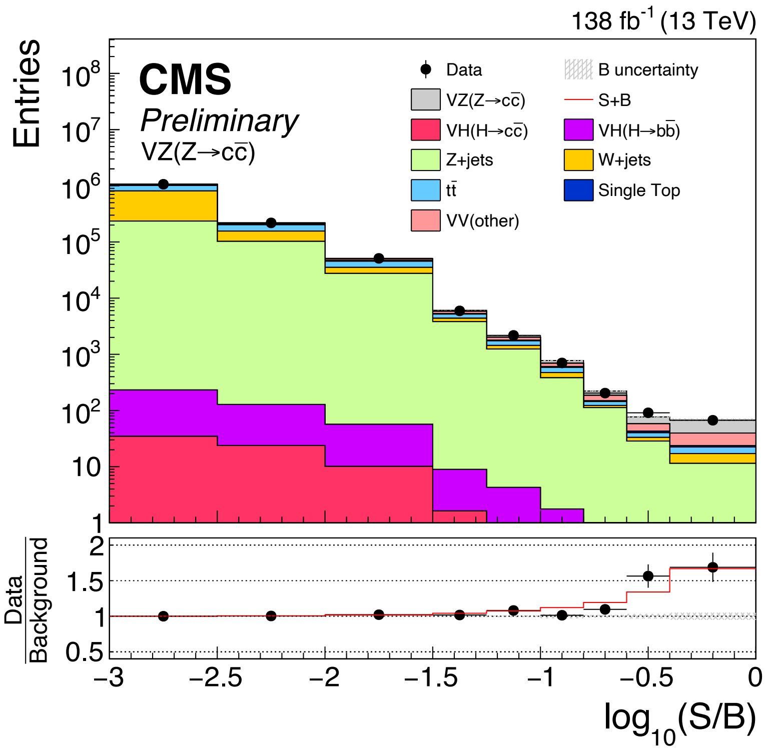
VZ(Z → cc) results

- The full analysis procedure is validated by measuring the VZ(Z → cc) process
  - resolved-jet topology:
    - BDT re-trained using VZ(Z → 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 → cc) results

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

arXiv:2205.05550



Observed significance for VZ(Z → cc): **5.7σ**

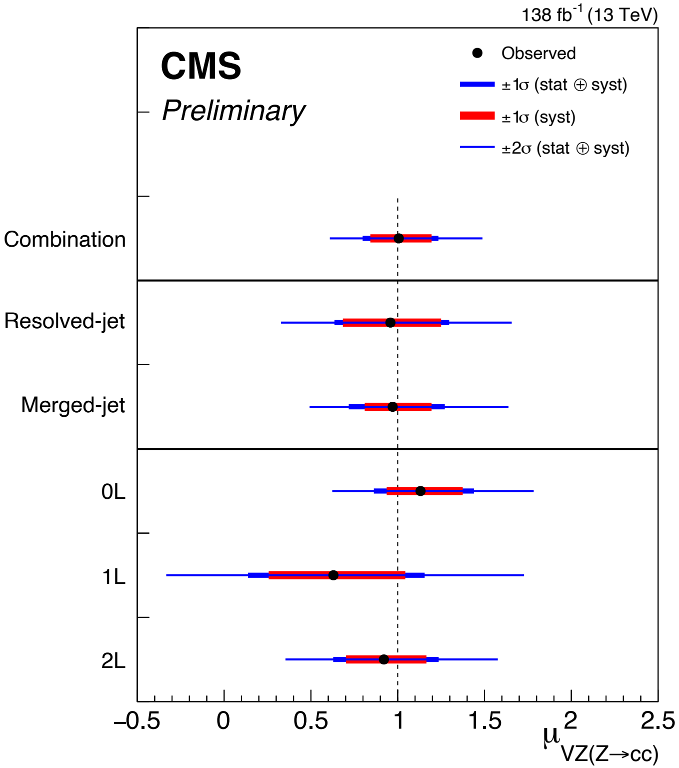
- expected significance: 5.9σ

First observation of Z→cc at a hadron collider!

VZ(Z → cc) results

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

arXiv:2205.05550



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

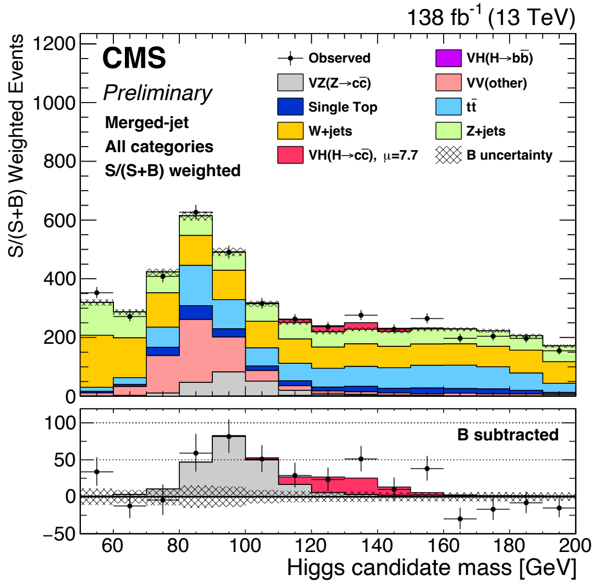
- very good agreement with SM expectation
- consistent results between topologies/channels

VH(H → cc) results

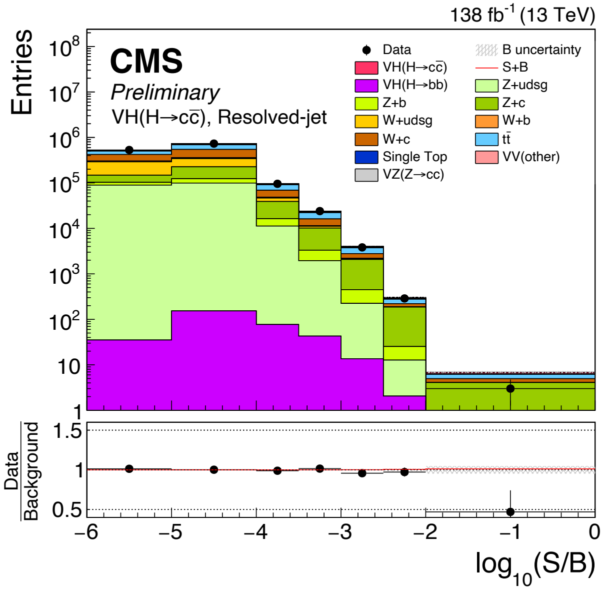
- Post-fit distributions in the two topologies and the combination

arXiv:2205.05550

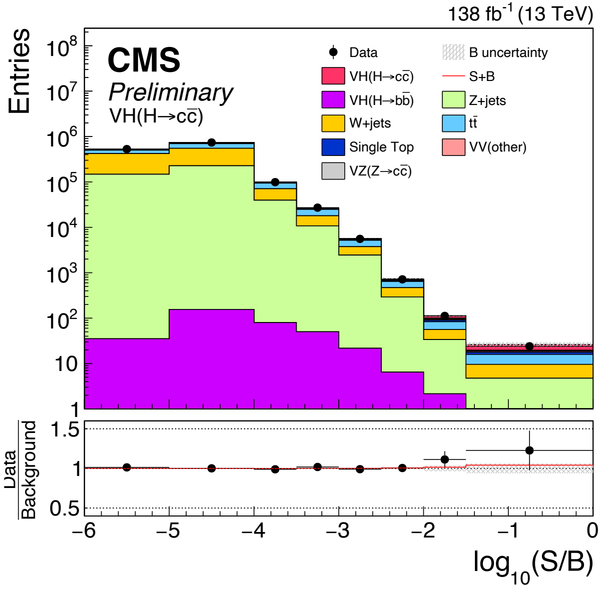
Merged-jet topology



Resolved-jet topology



Combination

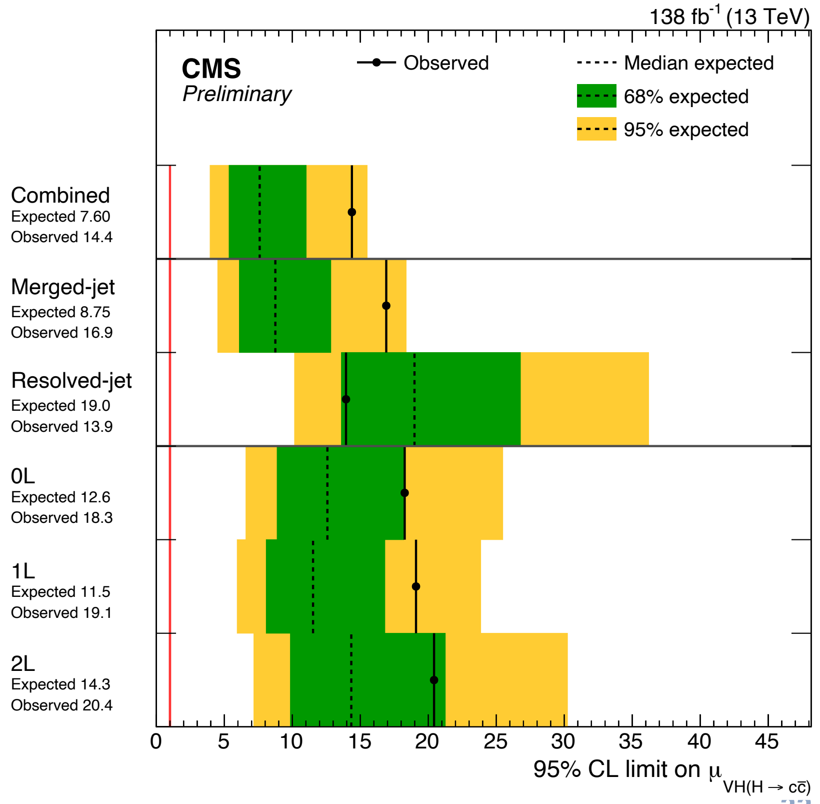


VH(H → cc) results

- Upper limits on the VH(H → cc) signal strength at 95% CL:
  - $\mu_{VH(H \to cc)}$ **< 14 (7.6) observed (expected)**
  - substantially stronger than ATLAS full Run 2 result
    - $\mu_{VH(H \to cc)}$ < 26 (31) obs. (exp.) [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

arXiv:2205.05550



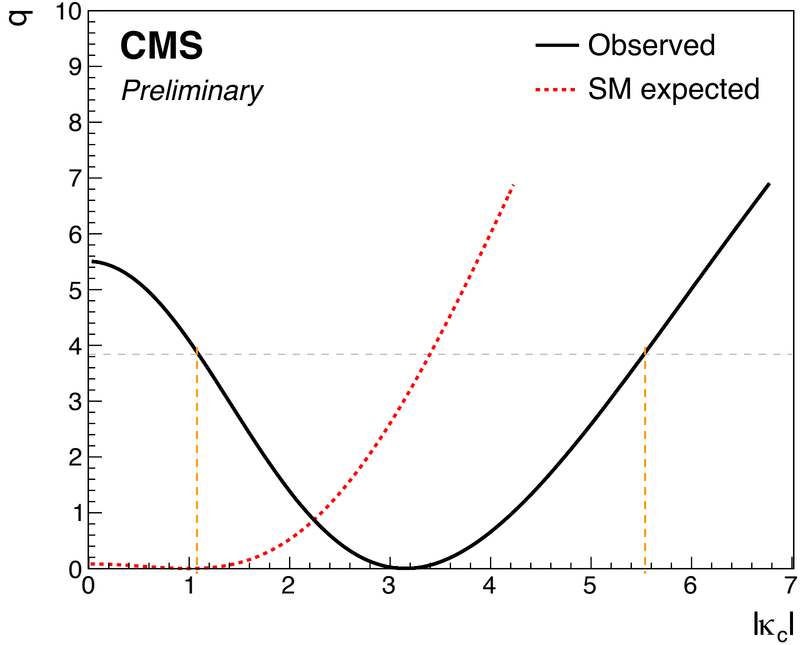
VH(H → cc) results

- Results used to set a constraint on the charm quark Yukawa coupling modifier $\kappa_c = y_c / y_c^{SM}$
  - for simplicity, only considering effects on B(H→cc) and fixing all other couplings at SM values:

$$\mu_{\mathrm{VH(H\to c\bar{c})}} = \frac{\kappa_c^2}{1 + \mathcal{B}_{\mathrm{SM}}\left(\mathrm{H} \to \mathrm{c\bar{c}}\right) \times \left(\kappa_c^2 - 1\right)}.$$

- The 95% CL interval on κ_c:
  - observed: **$1.1 < |\kappa_c| < 5.5$**
  - expected: **$|\kappa_c| < 3.4$**
  - **most stringent constraint on κ_c to date!**
    - comparable to the previous projection for HL-LHC w/ 3000 fb^{-1}: $|\kappa_c| < 3.0$ [ATL-PHYS-PUB-2021-039]

arXiv:2205.05550



Prospects: HL-LHC

Projection at HL-LHC: Setup

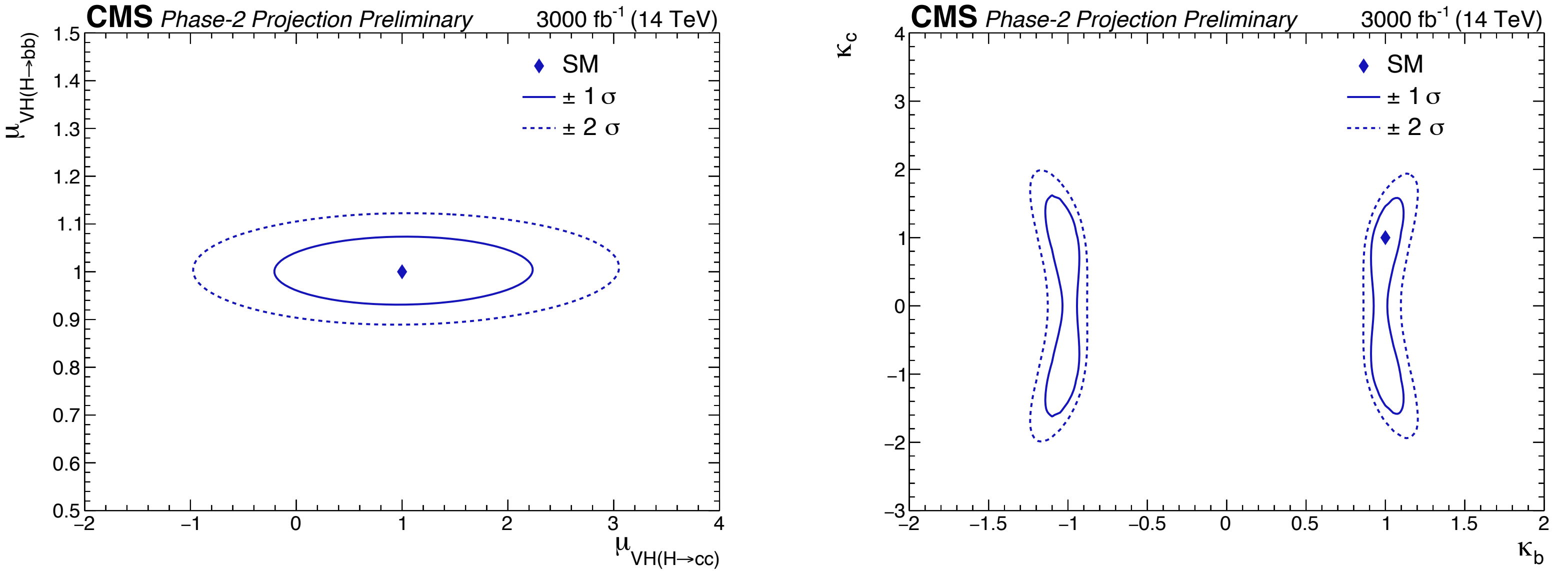
- Extrapolation of the merged-jet analysis to HL-LHC with 3000 fb^{-1} data
- Modifications to the Run 2 analysis to allow for a simultaneous constraint on $H \rightarrow bb$ and $H \rightarrow cc$
  - addition of 3 categories enriched in $H \rightarrow bb$ decays, selected with the ParticleNet bb-tagging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

CMS-PAS-HIG-21-008

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



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

Prospects: ML for jet tagging

Beyond ParticleNet

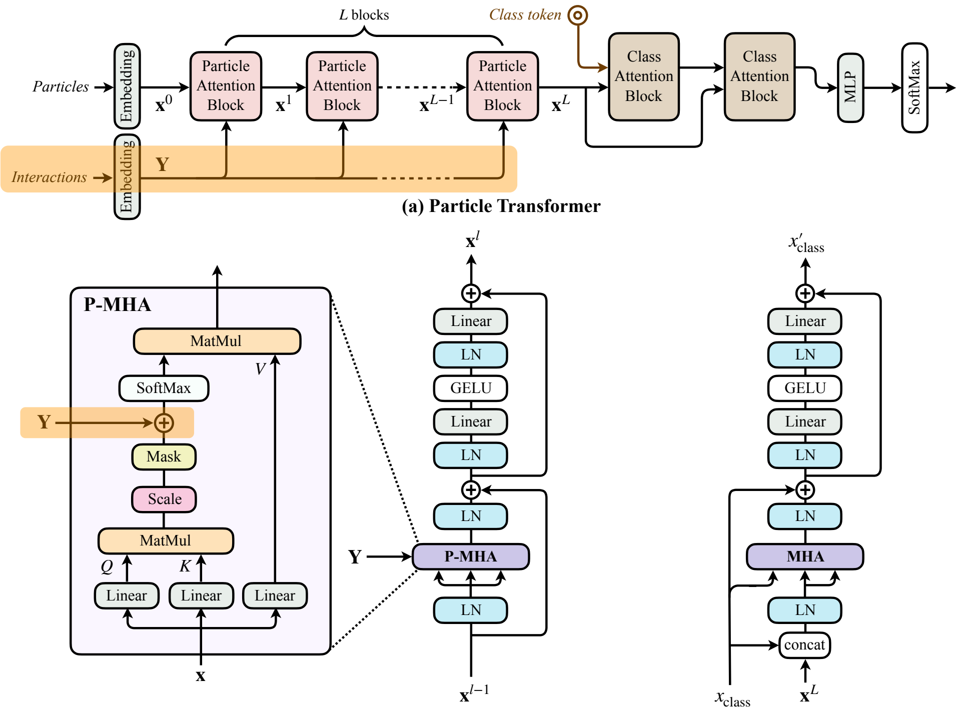
- **Particle Transformer (ParT)**: a new Transformer-based architecture for jet tagging
  - Transformer + *interaction*

Interaction features

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

Motivated by LundNet [F. Dreyer and H. Qu, JHEP 03 (2021) 052]

H. Qu, C. Li, S. Qian,
arXiv:2202.03772
GitHub

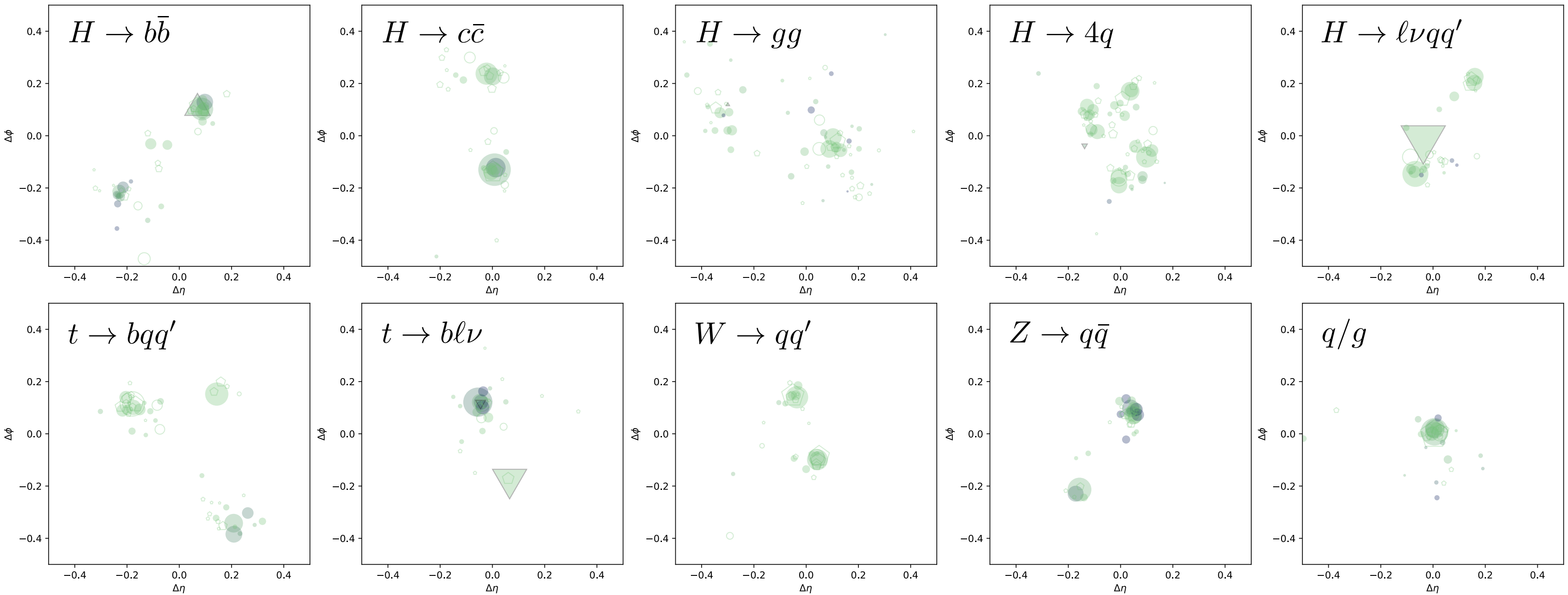


(a) Particle Transformer
(b) Particle Attention Block
(c) Class Attention Block

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
GitHub



Beyond ParticleNet: Particle Transformer

H. Qu, C. Li, S. Qian,
arXiv:2202.03772, GitHub

Performance comparison on the JetClass dataset

Rej$_{50\%}$ = 1/ε_{bkg} @ ε_{sig} = 50%

	All classes		$H \to b\bar{b}$	$H \to c\bar{c}$	$H \to gg$	$H \to 4q$	$H \to \ell\nu qq'$	$t \to bqq'$	$t \to b\ell\nu$	$W \to qq'$	$Z \to q\bar{q}$
	Accuracy	AUC	Rej$_{50\%}$	Rej$_{50\%}$	Rej$_{50\%}$	Rej$_{50\%}$	Rej$_{99\%}$	Rej$_{50\%}$	Rej$_{99.5\%}$	Rej$_{50\%}$	Rej$_{50\%}$
PFN	0.772	0.9714	2924	841	75	198	265	797	721	189	159
P-CNN	0.809	0.9789	4890	1276	88	474	947	2907	2304	241	204
ParticleNet	0.844	0.9849	7634	2475	104	954	3339	10526	11173	347	283
ParT	**0.861**	**0.9877**	**10638**	**4149**	**123**	**1864**	**5479**	**32787**	**15873**	**543**	**402**
ParT (plain)	0.849	0.9859	9569	2911	112	1185	3868	17699	12987	384	311

Fine-tuning result on Top-tagging Benchmark (~2M jets) [SciPost Phys. 7 (2019) 014]

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

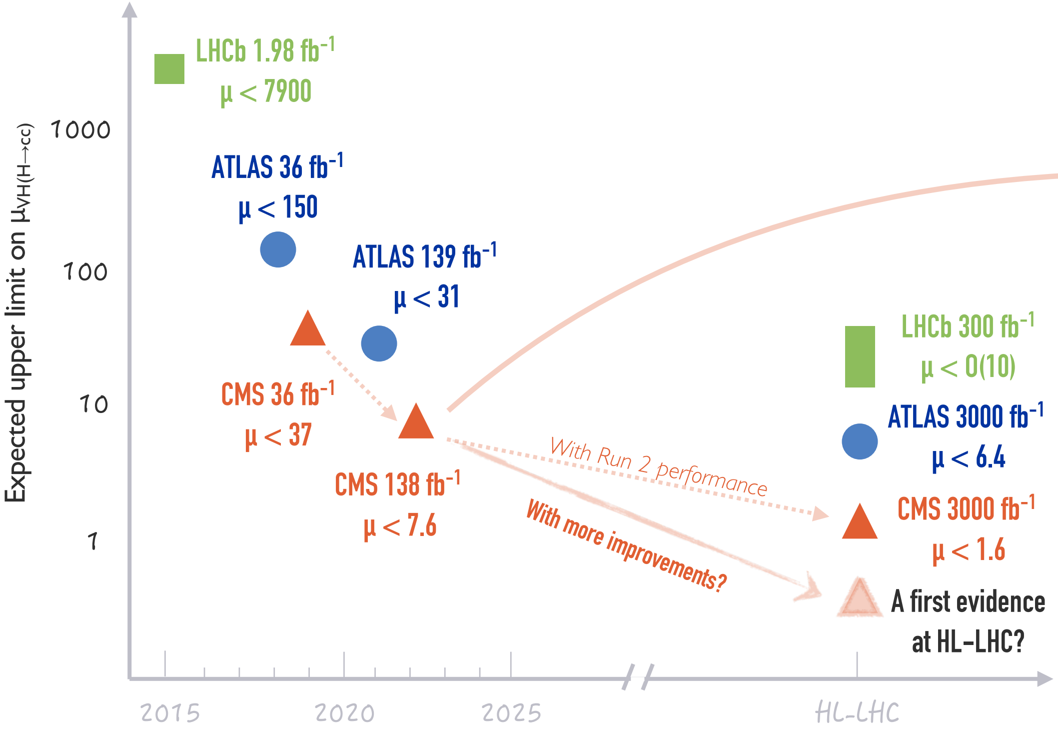
Model complexity

	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
ParT	**0.861**	2.14 M	340 M
ParT (plain)	0.849	2.13 M	260 M

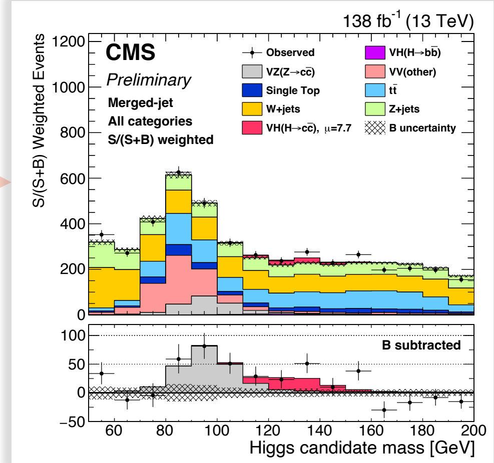
Significant performance improvement.
Similar computational cost.

Summary & Outlook

A charming journey



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!



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, …

A charming journey ahead!

Backups

H → 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})

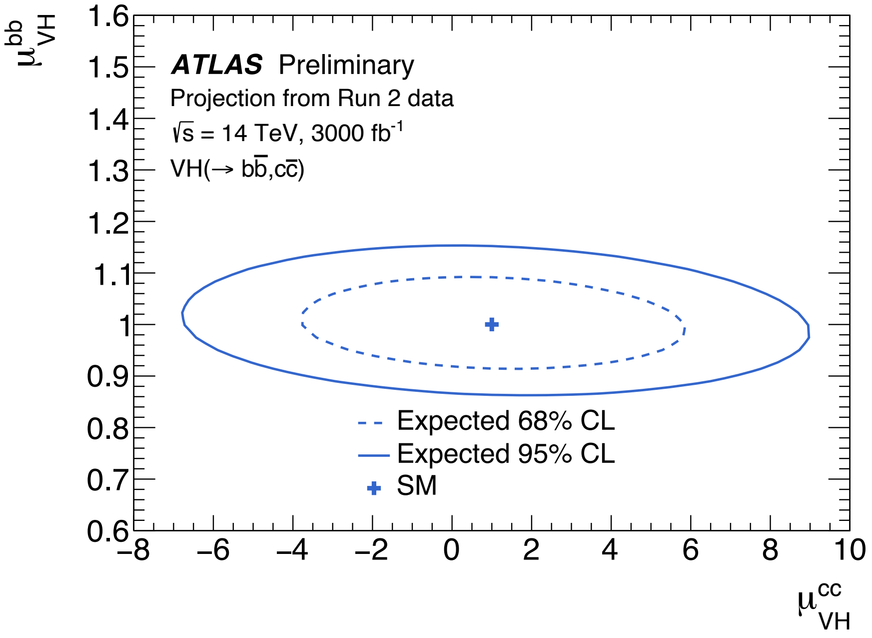
- LHCb:
  - [LHCb-CONF-2016-006] (1.98 fb^{-1})
  - [LHCb-PUB-2018-009] (HL-LHC projection, 300 fb^{-1})

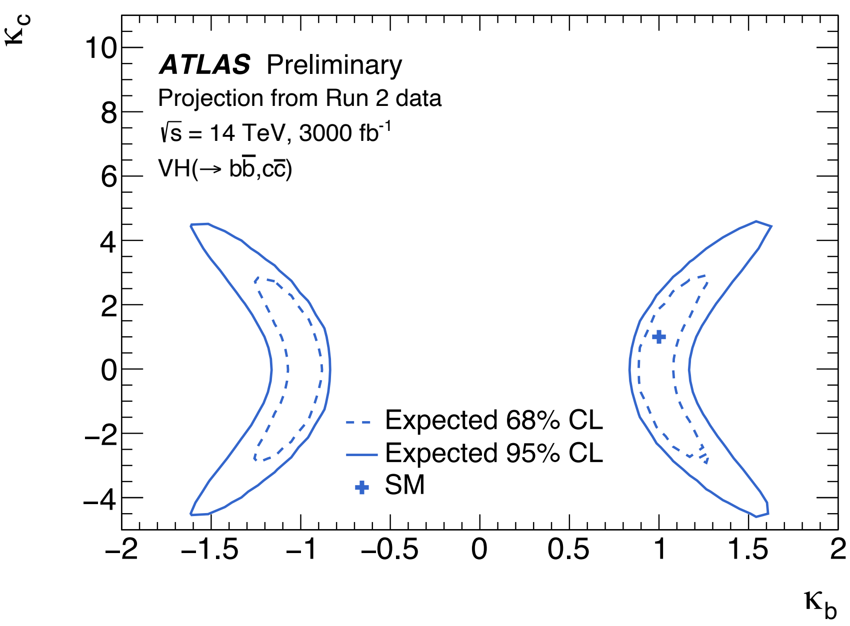
Higgs-charm coupling at HL-LHC

- Expected sensitivity at HL-LHC [CERN-2019-007]

$H \to cc$ decay	$\|\kappa_c\| <$ ~1-2
cH production	$\|\kappa_c\| <$ ~2-3
$H \to J/\Psi\gamma$	$\|\kappa_c\| <$ ~80
$p_T(H)$ distribution	$\|\kappa_c\| <$ ~10
WH charge assymetry	$\|\kappa_c\| <$ ~4-5

ATLAS HL-LHC projection for $H \to cc$





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
$N^{aj}_{\text{small-}R}$	<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
$N^{aj}_{\text{small-}R}$	—	<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 $\mu_{VH(H\to c\bar{c})}$ in the merged-jet analysis, with a best fit value $\mu_{VH(H\to c\bar{c})} = 8.7^{+4.6}_{-4.0}$.

Uncertainty source	$\Delta\mu / (\Delta\mu)_{tot}$
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 $\mu_{VH(H\to c\bar{c})}$ in the resolved-jet analysis, with a best fit value $\mu_{VH(H\to c\bar{c})} = -9.5 \pm 9.6$.

Uncertainty source	$\Delta\mu / (\Delta\mu)_{tot}$
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%

BDT input variables: Merged-jet topology

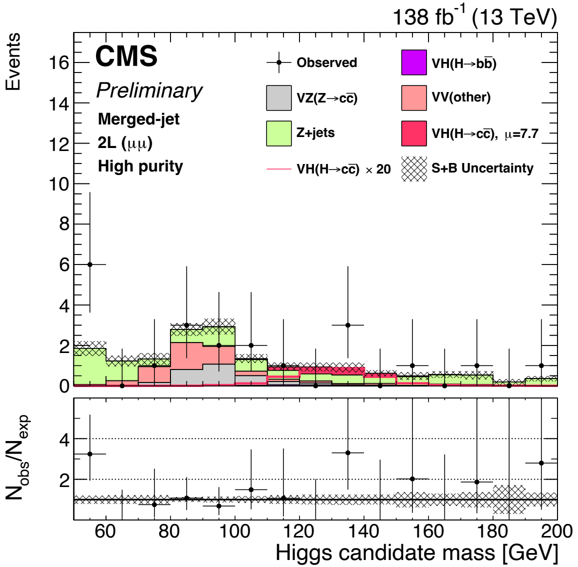
Variable	Description	0-lepton	1-lepton	2-lepton
$p_{\mathrm{T}}(\mathrm{V})$	vector boson transverse momentum	✓	✓	✓
$\Delta R(\ell,\ell)$	angular separation between the two leptons	—	—	✓
$p_{\mathrm{T}}\ (\mathrm{H}_{\mathrm{cand}})$	$\mathrm{H}_{\mathrm{cand}}$ transverse momentum	✓	✓	✓
$\lvert\eta(\mathrm{H}_{\mathrm{cand}})\rvert$	absolute value of the $\mathrm{H}_{\mathrm{cand}}$ pseudorapidity	✓	—	—
$\Delta\phi(\mathrm{V},\mathrm{H}_{\mathrm{cand}})$	azimuthal angle between vector boson and $\mathrm{H}_{\mathrm{cand}}$	✓	✓	✓
$p_{\mathrm{T}}^{\mathrm{miss}}$	missing transverse momentum	—	✓	—
$\Delta\eta(\mathrm{H}_{\mathrm{cand}},\ell)$	difference in pseudorapidity between $\mathrm{H}_{\mathrm{cand}}$ and the lepton	—	✓	—
$\Delta\eta(\mathrm{H}_{\mathrm{cand}},)$	difference in pseudorapidity between $\mathrm{H}_{\mathrm{cand}}$ and vector boson	—	—	✓
$\Delta\eta(\mathrm{H}_{\mathrm{cand}},\mathrm{j})$	min. difference in pseudorapidity between $\mathrm{H}_{\mathrm{cand}}$ and small-R jets	✓	✓	✓
$\Delta\eta(\ell,\mathrm{j})$	min. difference in pseudorapidity between the lepton and small-R jets	—	✓	—
$\Delta\eta(\mathrm{V},\mathrm{j})$	min. difference in pseudorapidity between vector boson and small-R jets	—	—	✓
$\Delta\phi(\vec{p}_{\mathrm{T}}^{\mathrm{miss}},\mathrm{j})$	azimuthal angle between $\vec{p}_{\mathrm{T}}^{\mathrm{miss}}$ and closest small-R jet	✓	—	—
$\Delta\phi(\vec{p}_{\mathrm{T}}^{\mathrm{miss}},\ell)$	azimuthal angle between $\vec{p}_{\mathrm{T}}^{\mathrm{miss}}$ and lepton	—	✓	—
m_{T}	transverse mass of lepton $\vec{p}_{\mathrm{T}} + \vec{p}_{\mathrm{T}}^{\mathrm{miss}}$	—	✓	—
$N_{\text{small-}R}^{\mathrm{aj}}$	number of additional small-R jets	✓	✓	✓

BDT input variables: Resolved-jet topology

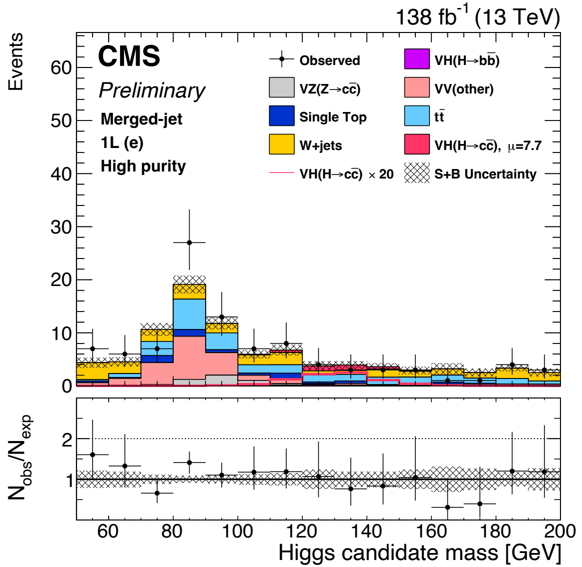
Variable	Description	0-lepton	1-lepton	2-lepton
$m\,(\mathrm{H_{cand}})$	$\mathrm{H_{cand}}$ mass	✓	✓	✓
$p_\mathrm{T}\,(\mathrm{H_{cand}})$	$\mathrm{H_{cand}}$ transverse momentum	—	✓	✓
$p_\mathrm{T}(\mathrm{V})$	vector boson transverse momentum	—	✓	✓
$m_\mathrm{T}(\mathrm{V})$	vector boson transverse mass	—	✓	—
$p_\mathrm{T}^\mathrm{miss}$	missing transverse momentum	✓	✓	—
$p_\mathrm{T}(\mathrm{V})/p_\mathrm{T}(\mathrm{H_{cand}})$	ratio between vector boson and H transverse momenta	✓	✓	✓
$CvsL_\mathrm{max}$	*CvsL* value of the leading *CvsL* jet	✓	✓	✓
$CvsB_\mathrm{max}$	*CvsB* value of the leading *CvsL* jet	✓	✓	✓
$CvsL_\mathrm{min}$	*CvsL* value of the subleading *CvsL* jet	✓	✓	✓
$CvsB_\mathrm{min}$	*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	✓	✓	✓
$\Delta\phi(\mathrm{V},\mathrm{H_{cand}})$	azimuthal angle between vector boson and H	✓	✓	✓
$\Delta R(\mathrm{j_1},\mathrm{j_2})$	ΔR between leading and subleading *CvsL* jets	—	✓	✓
$\Delta\phi(\mathrm{j_1},\mathrm{j_2})$	azimuthal angle between leading and subleading *CvsL* jets	✓	✓	—
$\Delta\eta(\mathrm{j_1},\mathrm{j_2})$	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	—	—	✓
$\Delta\phi(\ell_1,\mathrm{j_1})$	azimuthal angle between leading p_T lepton and leading *CvsL* jet	—	✓	—
$\Delta\phi(\ell_2,\mathrm{j_1})$	azimuthal angle between subleading p_T lepton and leading *CvsL* jet	—	—	✓
$\Delta\phi(\ell_2,\mathrm{j_2})$	azimuthal angle between subleading p_T lepton and subleading *CvsL* jet	—	—	✓
$\Delta\phi(\ell_1,p_\mathrm{T}^\mathrm{miss})$	azimuthal angle between leading p_T lepton and missing transverse momentum	—	✓	—
$\Delta\eta(\ell_1,\mathrm{b})$	difference in pseudorapidity between leading p_T lepton and b-tagged jet from top quark decay	—	✓	—
$\Delta\phi(\ell_1,\mathrm{b})$	azimuthal angle between leading p_T lepton and b-tagged jet from top quark decay	—	✓	—
$\Delta R(\ell_1,\mathrm{b})$	ΔR between leading p_T lepton and b-tagged jet from top quark decay	—	✓	—
$CvsL_\mathrm{b}$	*CvsL* value of the b-tagged jet from top quark decay	—	✓	—
$CvsB_\mathrm{b}$	*CvsB* value of the b-tagged jet from top quark decay	—	✓	—
$P(b{+}bb)_\mathrm{b}$	*DeepJet prob(b+bb)* value of the b-tagged jet from top quark decay	—	✓	—
$m(\mathrm{t})$	Reconstructed top quark mass	—	✓	—
$N^\mathrm{aj}_{\text{small-}R}$	Number of small-*R* additional jets after the FSR subtraction	—	✓	—
$\sigma_{cReg}(\mathrm{j_1})$	leading p_T jet resolution from c-jet energy regression	✓	✓	✓
$\sigma_{cReg}(\mathrm{j_2})$	subleading p_T jet resolution from c-jet energy regression	✓	✓	✓
$\Delta\eta(\mathrm{V},\mathrm{H_{cand}})\|_\mathrm{kinfit}$	difference in pseudorapidity between vector boson and $\mathrm{H_{cand}}$, after kinematic-fit	—	—	✓
$\Delta\phi(\mathrm{V},\mathrm{H_{cand}})\|_\mathrm{kinfit}$	azimuthal angle between vector boson and $\mathrm{H_{cand}}$, after kinematic-fit	—	—	✓
$m(\mathrm{H_{cand}})\|_\mathrm{kinfit}$	$\mathrm{H_{cand}}$ mass after kinematic-fit	—	—	✓
$p_\mathrm{T}(\mathrm{H_{cand}})\|_\mathrm{kinfit}$	$\mathrm{H_{cand}}$ transverse momentum after kinematic-fit	—	—	✓
$p_\mathrm{Tmax}\|_\mathrm{kinfit}$	p_T of the leading *CvsL* jet after kinematic-fit	—	—	✓
$p_\mathrm{Tmin}\|_\mathrm{kinfit}$	p_T of the subleading *CvsL* jet after kinematic-fit	—	—	✓
$p_\mathrm{T}(\mathrm{V})/p_\mathrm{T}(\mathrm{H_{cand}})\|_\mathrm{kinfit}$	ratio between vector boson and $\mathrm{H_{cand}}$ transverse momenta after kinematic-fit	—	—	✓
$\sigma(\mathrm{H_{cand}})\|_\mathrm{kinfit}$	$\mathrm{H_{cand}}$ invariant mass resolution from kinematic fit	—	—	✓

Merged-jet topology: signal regions

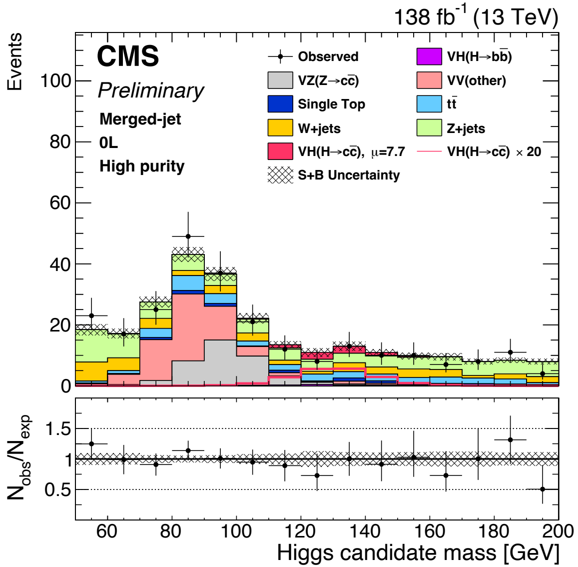
2L(μμ), high cc-purity



1L(e), high cc-purity

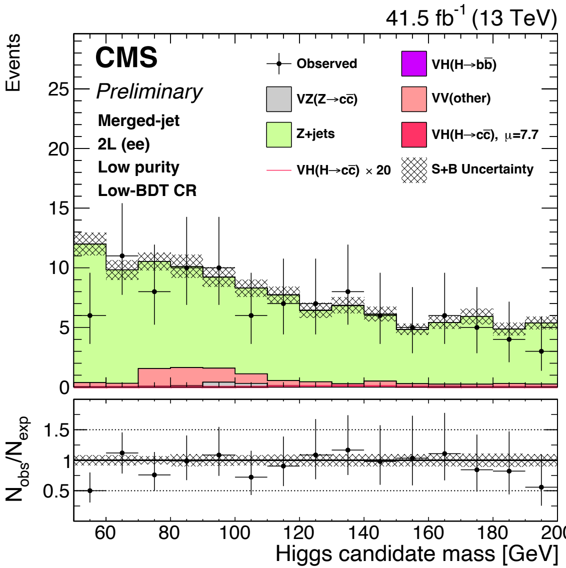


0L, high cc-purity

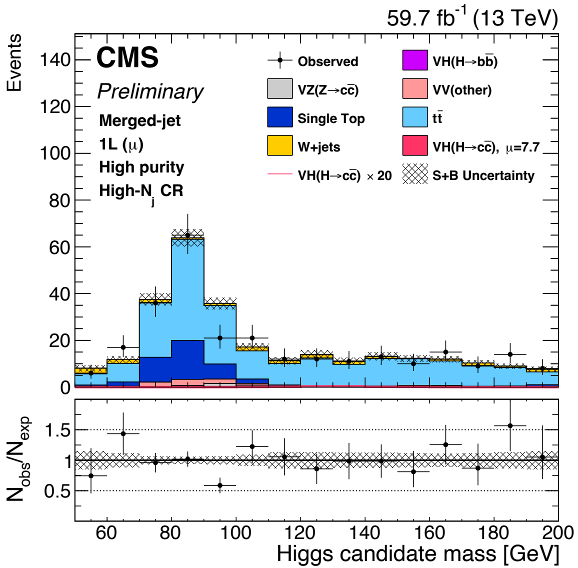


Merged-jet topology: control regions

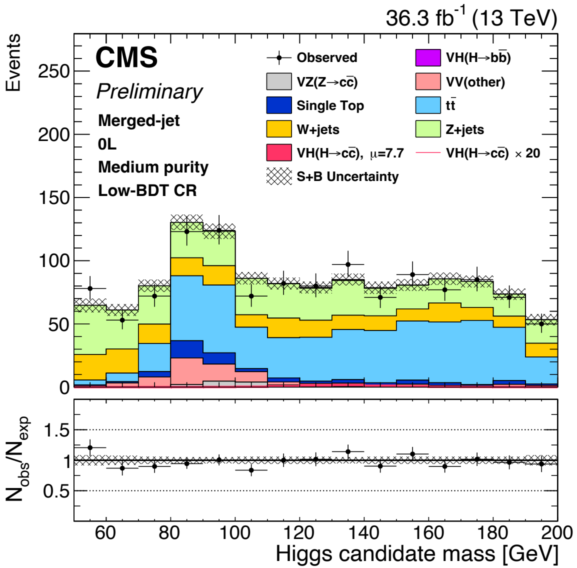
2L(ee), V+jets CR, low cc-purity



1L(μ), tt CR, high cc-purity

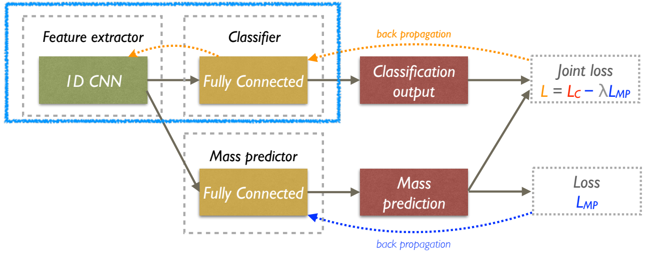


0L, V+jets CR, medium cc-purity

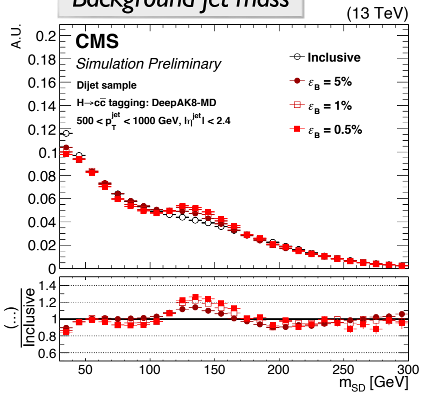


Comparison of mass decorrelation methods

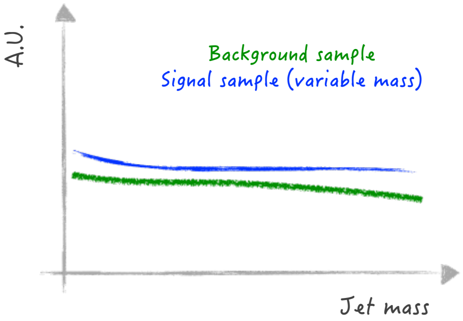
DeepAK8-MD



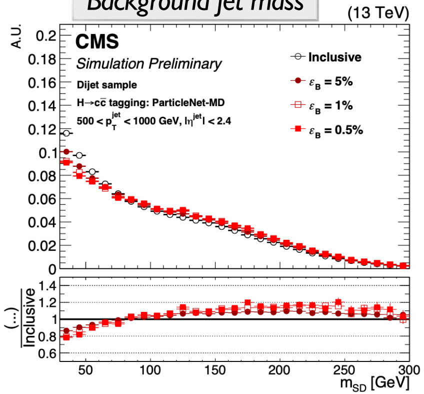
Background jet mass



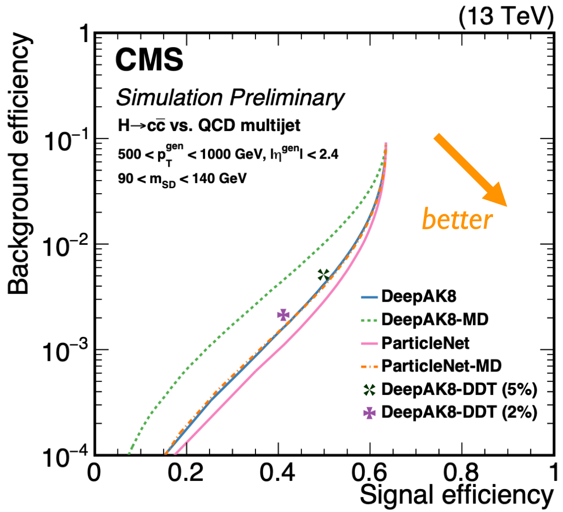
ParticleNet-MD



Background jet mass



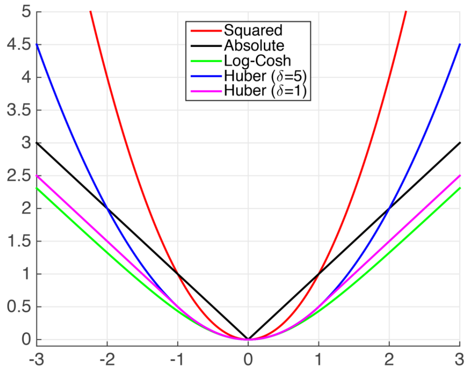
H→cc tagging performance



Large-R jet mass regression

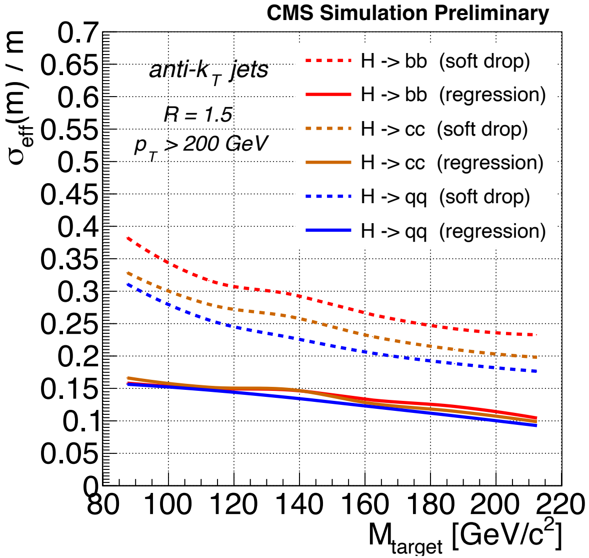
Loss function: LogCosh

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

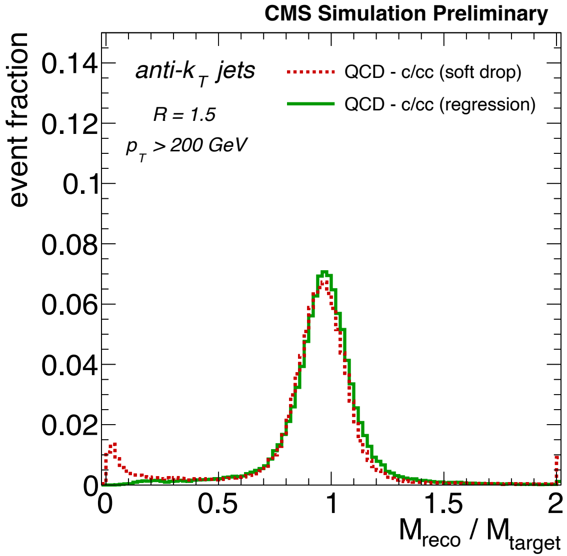


https://www.cs.cornell.edu/courses/cs4780/2015fa/web/lecturenotes/lecturenote10.html

Signal jet mass resolution



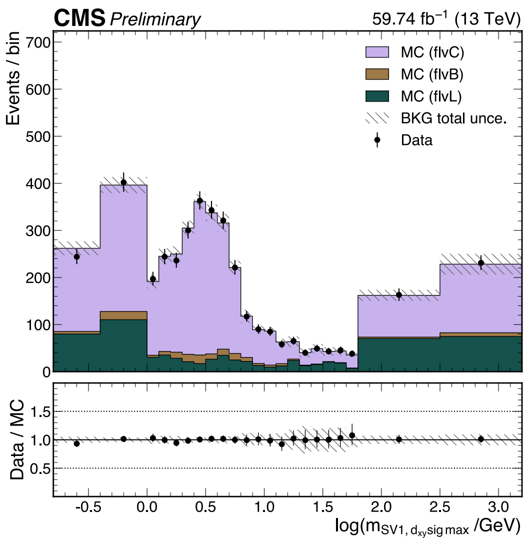
Background jet mass response



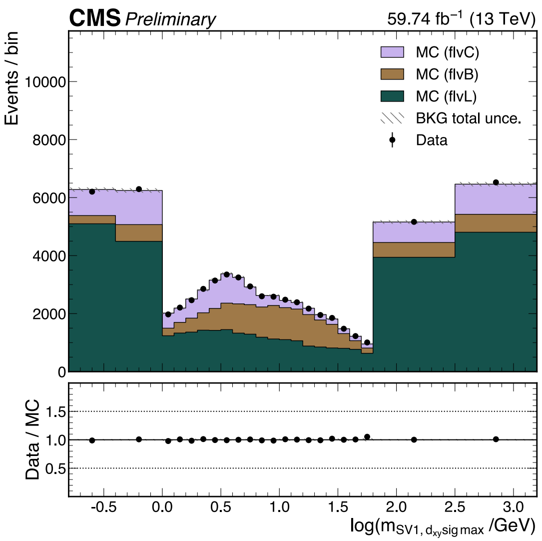
Calibration of the cc-tagger

CMS-DP-2022-005

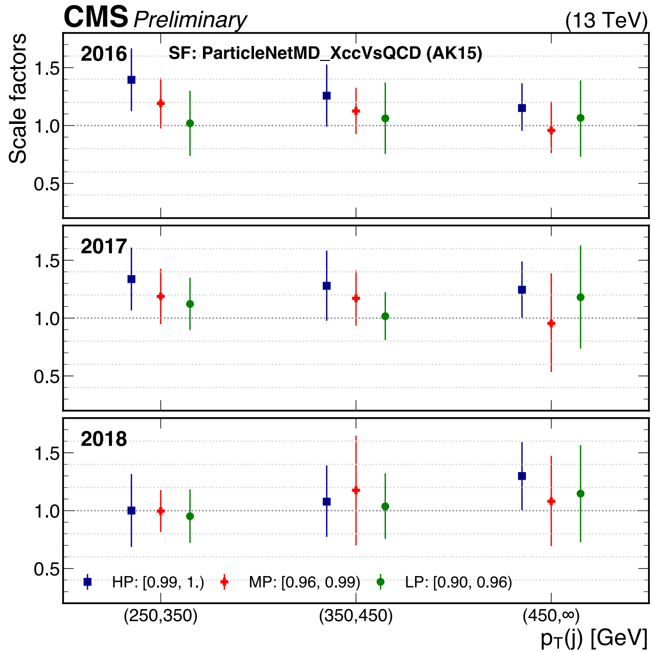
Passing category



Failing category

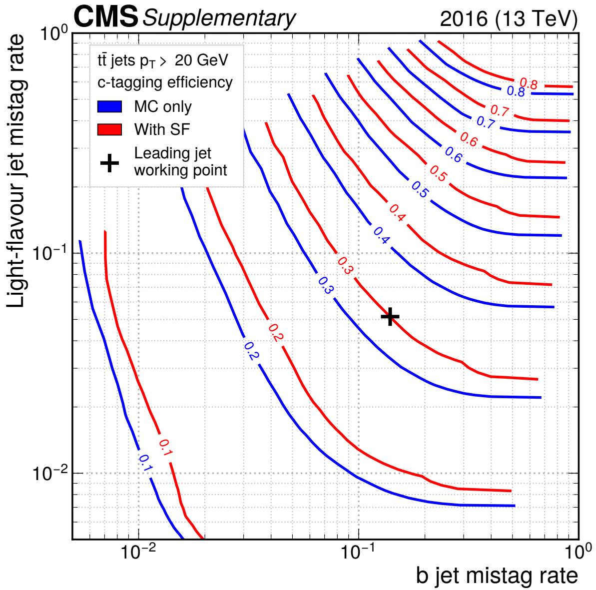


cc-tagging scale factors

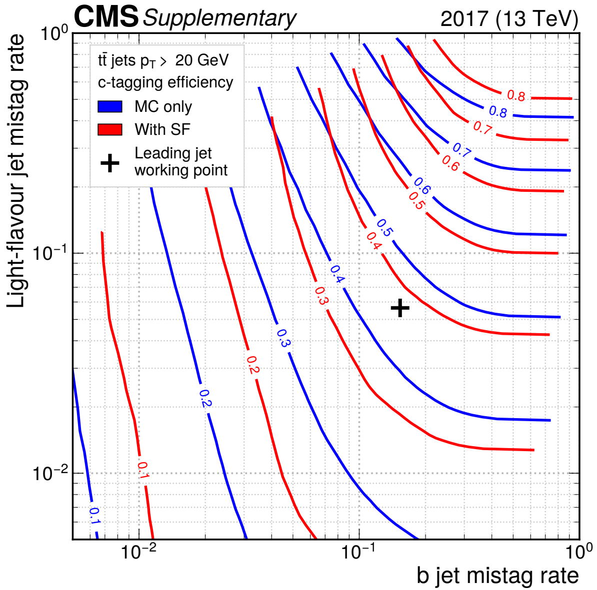


C-tagger ROC curves

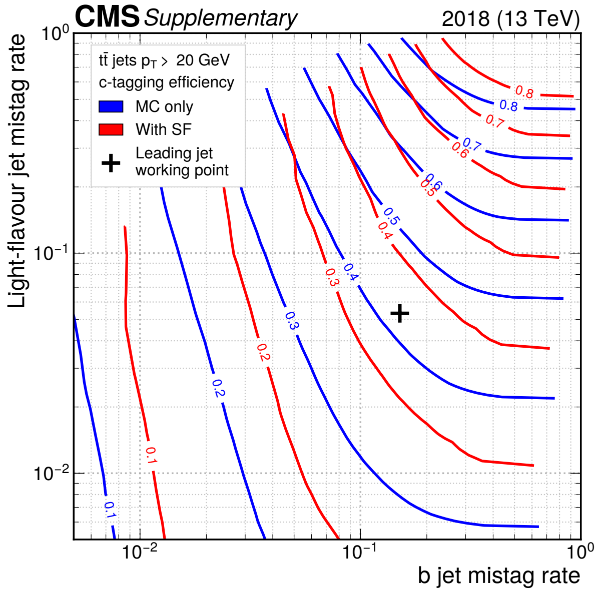
2016



2017



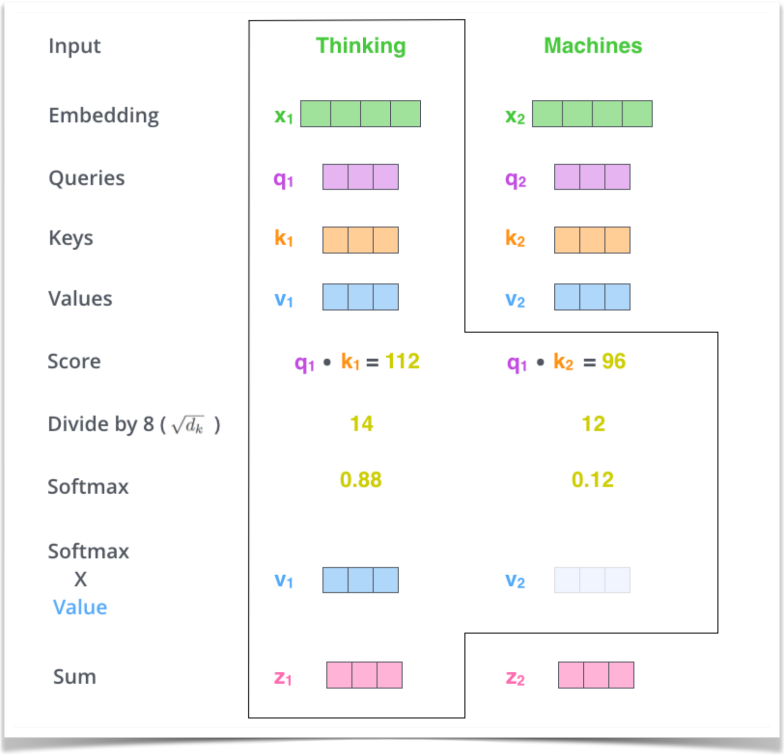
2018

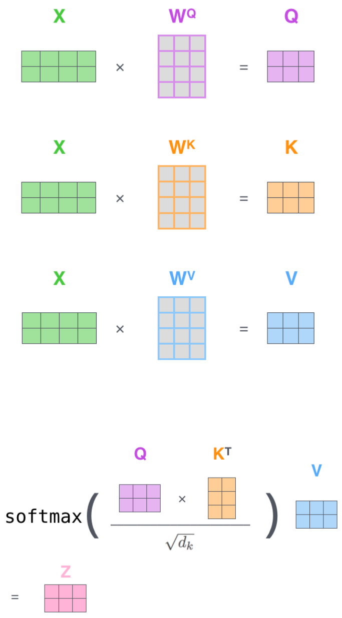


- CMS c-tagging WP: ~40% (c), ~16% (b), ~4% (light)
- ATLAS c-tagging WP [arXiv:2201.11428]: 27% (c), 8% (b), 1.6% (light)

Transformer 101

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





https://jalammar.github.io/illustrated-transformer/ X