Search for a Higgs Boson Decaying to Two Charm Quarks at the LHC

Bjorn Burkle
(Brown University)

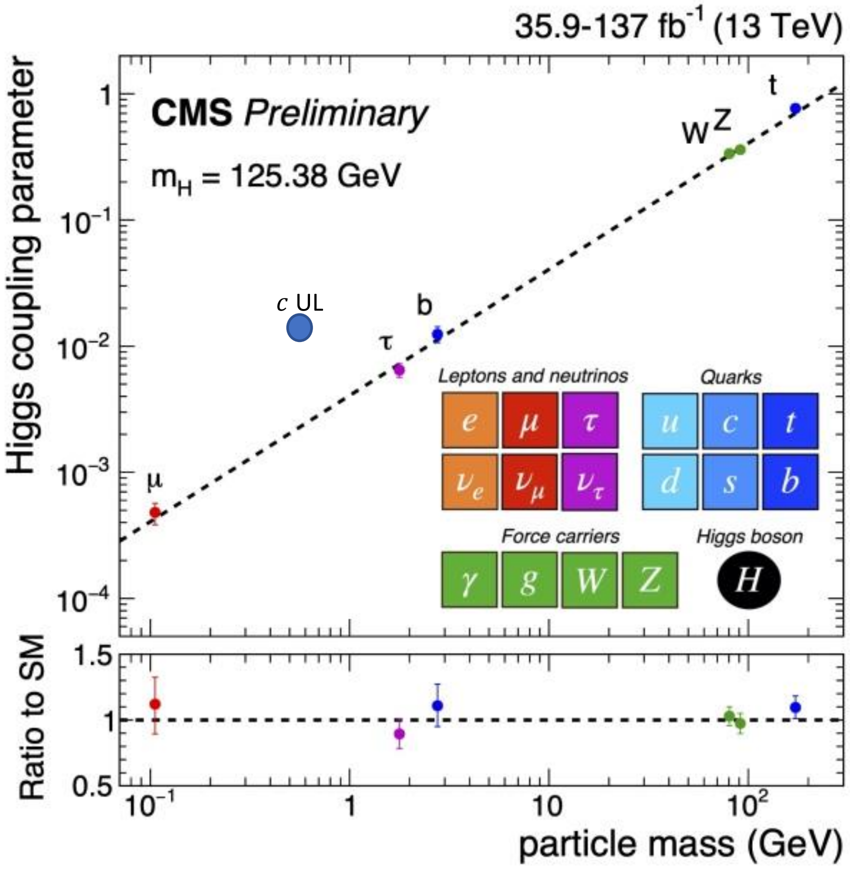
on behalf of the CMS and ATLAS Collaborations

Higgs 2020 Conference (Virtual)
October 26-29, 2020

arXiv:1503.00290

Introduction

- Since the observation of the Higgs, measuring its properties has been a primary goal of CMS and ATLAS
- Deviations from theory are signs of BSM
  - Need coupling to be on line for H to be responsible for particle mass as predicted by SM
- So far, have seen remarkable agreement with SM
- Based on indirect Higgs measurements using model dependent assumptions, expect that $\kappa_{c<6.2}$ [1]



Introduction

arXiv:2009.04363
arXiv:1610.07922
arXiv:2007.07830

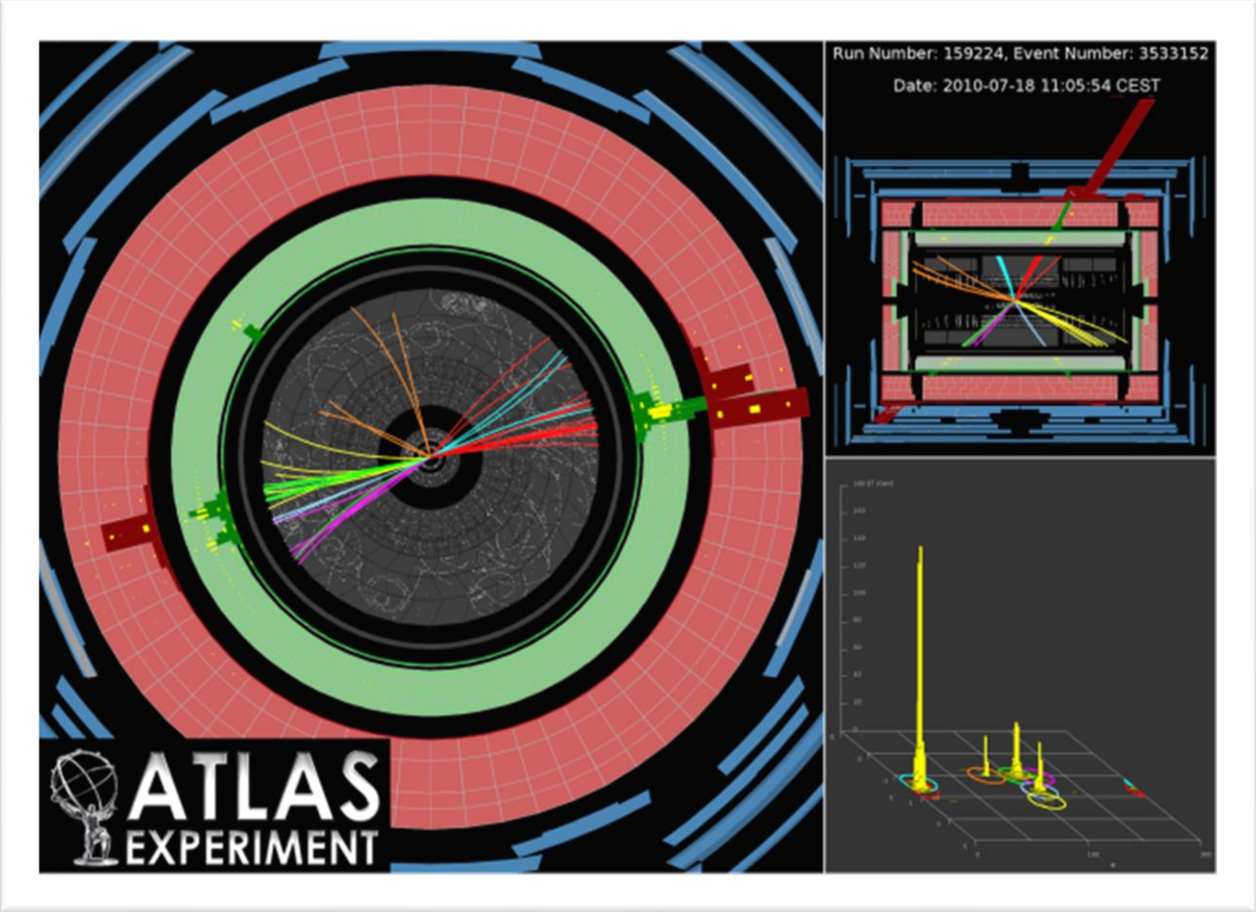
Decay Processes	BR (Theory*)
$H \to bb$	58%
$H \to WW$	22%
$H \to gg$	8.1%
$H \to \tau\tau$	6.2%
$\boldsymbol{H \to cc}$	**2.9%**
$H \to ZZ$	2.6%
$H \to \gamma\gamma$	0.23%
$H \to Z\gamma$	0.16%
$H \to \mu\mu$	0.02%

Orange entries have been observed to 5σ

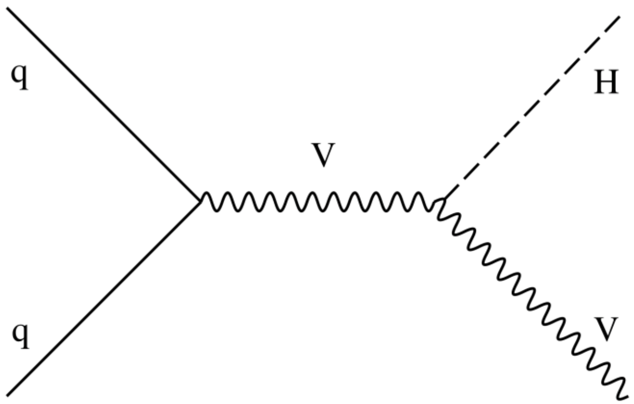
- Can measure couplings directly by measuring Higgs decay rates
- Evidence of $\mu\mu$ decay recently reported at 3.0σ by CMS and 2.0σ by ATLAS
- Dijet Higgs decays are difficult to detect at the LHC
  - Higgs to bb observed much later than ZZ or $\gamma\gamma$
  - BR$(H \to cc)$ > 10 × BR$(H \to \gamma\gamma)$ > 100 × BR$(H \to \mu\mu)$

The Challenge

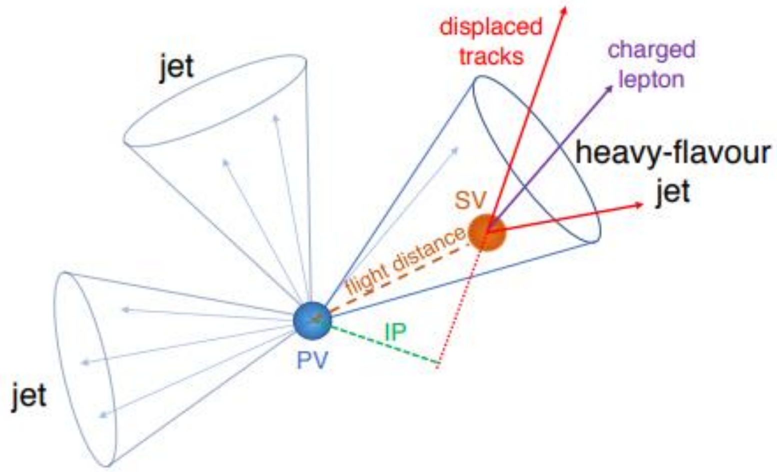
- Events with only jets are the most common events at the LHC
  - QCD dijet production have cross sections O(10^5 pb)
  - Higgs production have cross sections O(10 pb)
- c-jets have broader jet energy resolution
  - Larger systematic limits analysis sensitivity
- Charm tagging is a very difficult task
  - c-jets kind of look like b-jets and kind of look like l-jets
  - Similarities due to properties of the intermediate mesons created by b-, c-, and light quark hadronization
  - $H \rightarrow b\bar{b}$ is 20x as likely and $H \rightarrow l$-jets is 3x as likely



Higgs to $c\bar{c}$ Strategy



- Search carried out on targeting events where the Higgs boson is produced in association with vector boson (W or Z)
- Exploit W/Z leptonic final states to differentiate $W/Z + H(c\bar{c})$ process from common backgrounds
- Extensive use of advanced machine learning techniques to identify jets originating from c-quarks



arXiv:1802.04329

ATLAS Analysis

- 2015 and 2016 data with L = 36.1 fb^{-1}
- Perform search in the $Z(\ell^+\ell^-)H(c\bar{c})$ final state

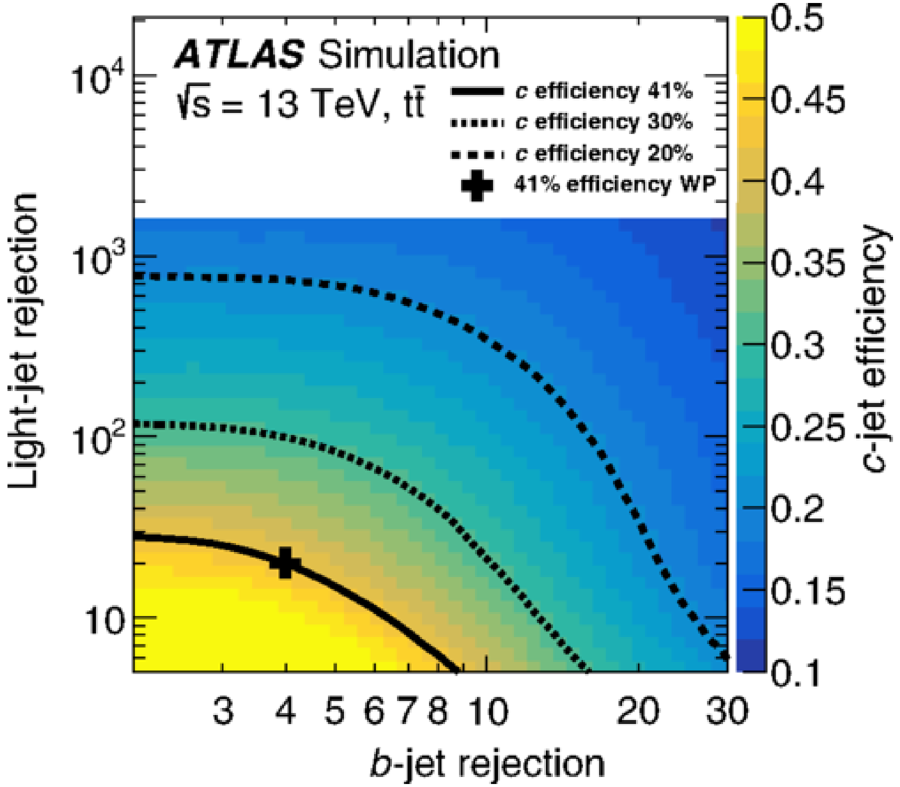
- Z reconstructed from same-flavor lepton pairs
  - OS requirement for muon pairs
  - $|\eta| < 2.5$ with lead (sub-leading) $p_T > 27$ (7) GeV
  - 81 GeV < $m_{\ell\ell}$ < 101 GeV

- Higgs reconstructed from 2 AK4 jets
  - $|\eta| < 2.5$ and $p_T > 20$ GeV
  - Constructed from 2 highest p_T jets
  - In MC, events are weighted based on jet flavor tagging efficiency

arXiv:1802.04329

ATLAS Analysis

- Two dedicated Boosted Decision Trees (BDTs) used for charm tagging – 41% c-tagging efficiency
  - Differentiate c- vs l-jets; 5% l-jet mistag
  - Differentiate c- vs b-jets; 25% b-jet mistag



- BDT trained on jet kinematics and properties of charged tracks contained in jet
  - Longitudinal and transverse impact parameters of tracks
  - Properties of displaced vertex (if present)
- Variables motivated by b- and c-hadron properties

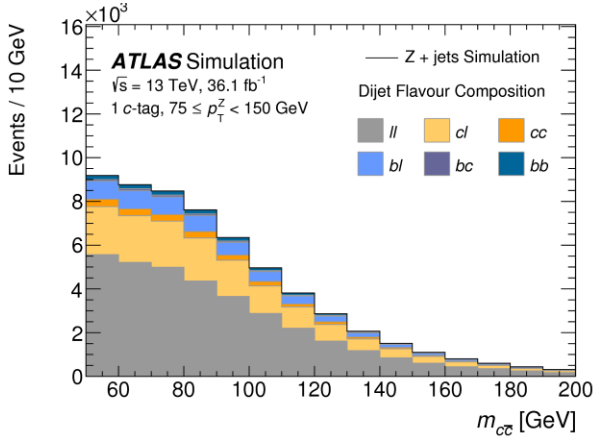
arXiv:1802.04329

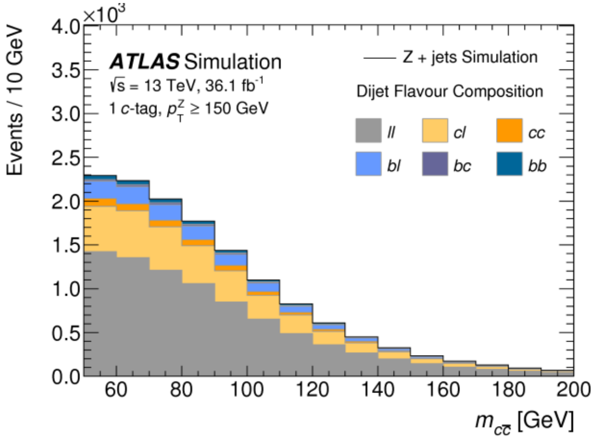
ATLAS Analysis

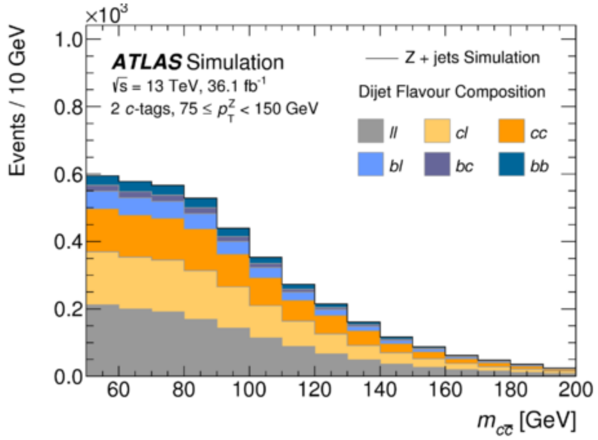
- Data split into 4 categories based on c-tags and p_T^Z
- Simultaneous profile-likelihood fit to $m_{c\bar{c}}$ in all categories with Z+Jets normalization floating

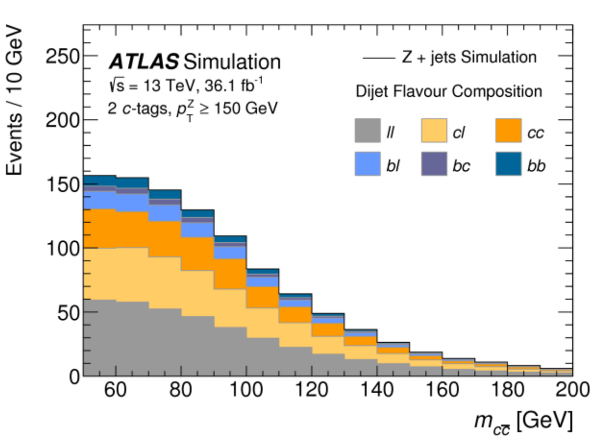
$m_{c\bar{c}}$	50 GeV $< m_{c\bar{c}} <$ 200 GeV			
c-tags	1 c-tag		2 c-tags	
p_T^Z	$[75\ \text{GeV}, 150\ \text{GeV})$	$[150\ \text{GeV}, \infty)$	$[75\ \text{GeV}, 150\ \text{GeV})$	$[150\ \text{GeV}, \infty)$

p_T^Z dependent $\Delta R_{c\bar{c}}$ cut also applied to exploit smaller jet separation observed in signal







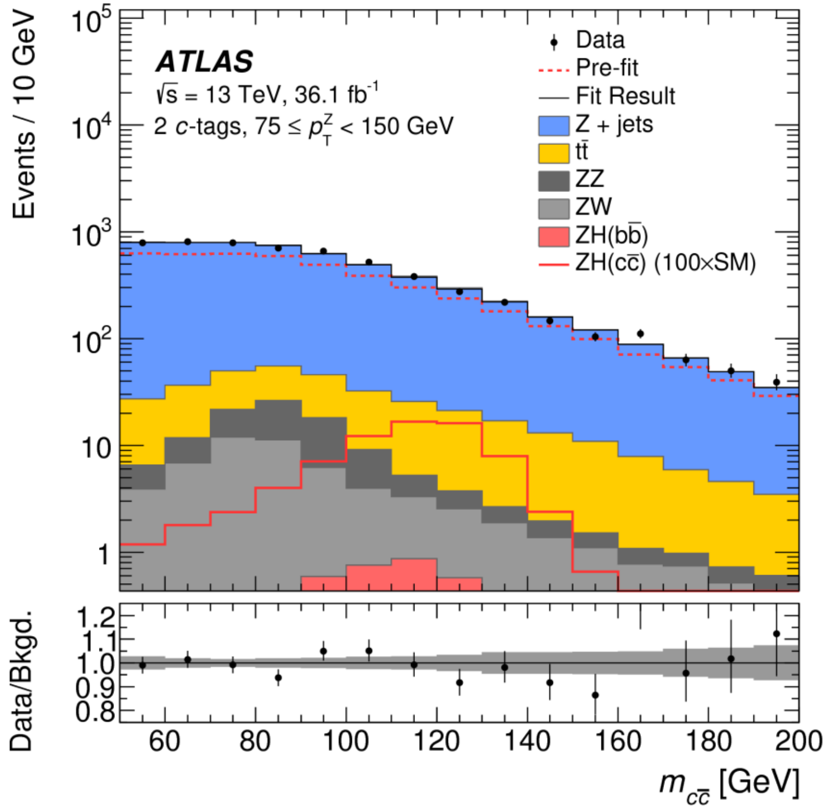


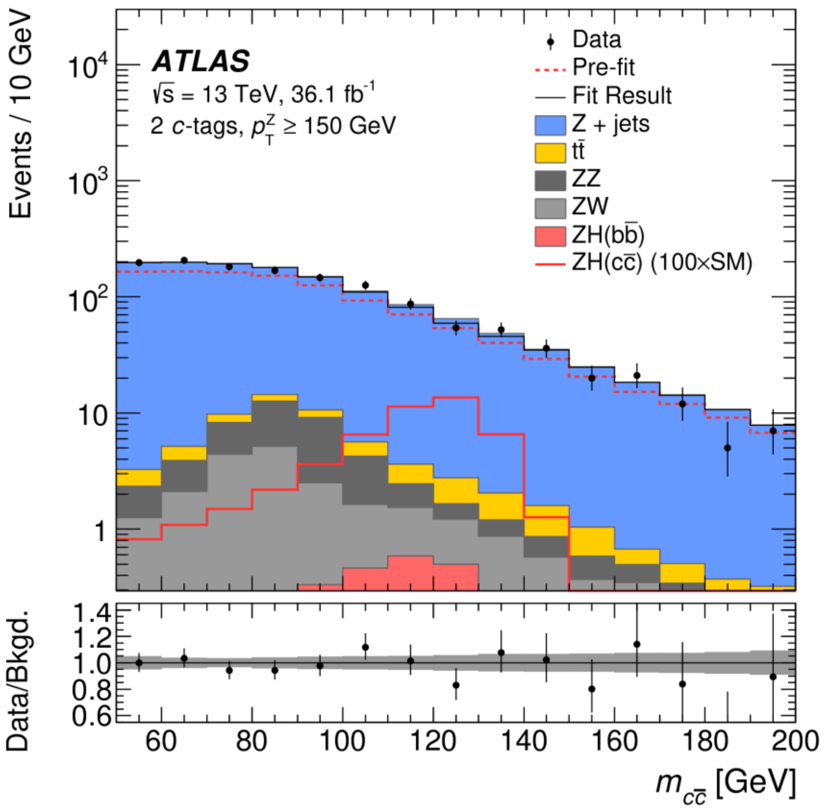
arXiv:1802.04329

ATLAS Analysis

Limits extracted from post-fit $m_{c\bar{c}}$ distribution

Obtained 95% CL observed (expected) upper limit on σ/σ_{SM} of $110\ (150^{+80}_{-42})$

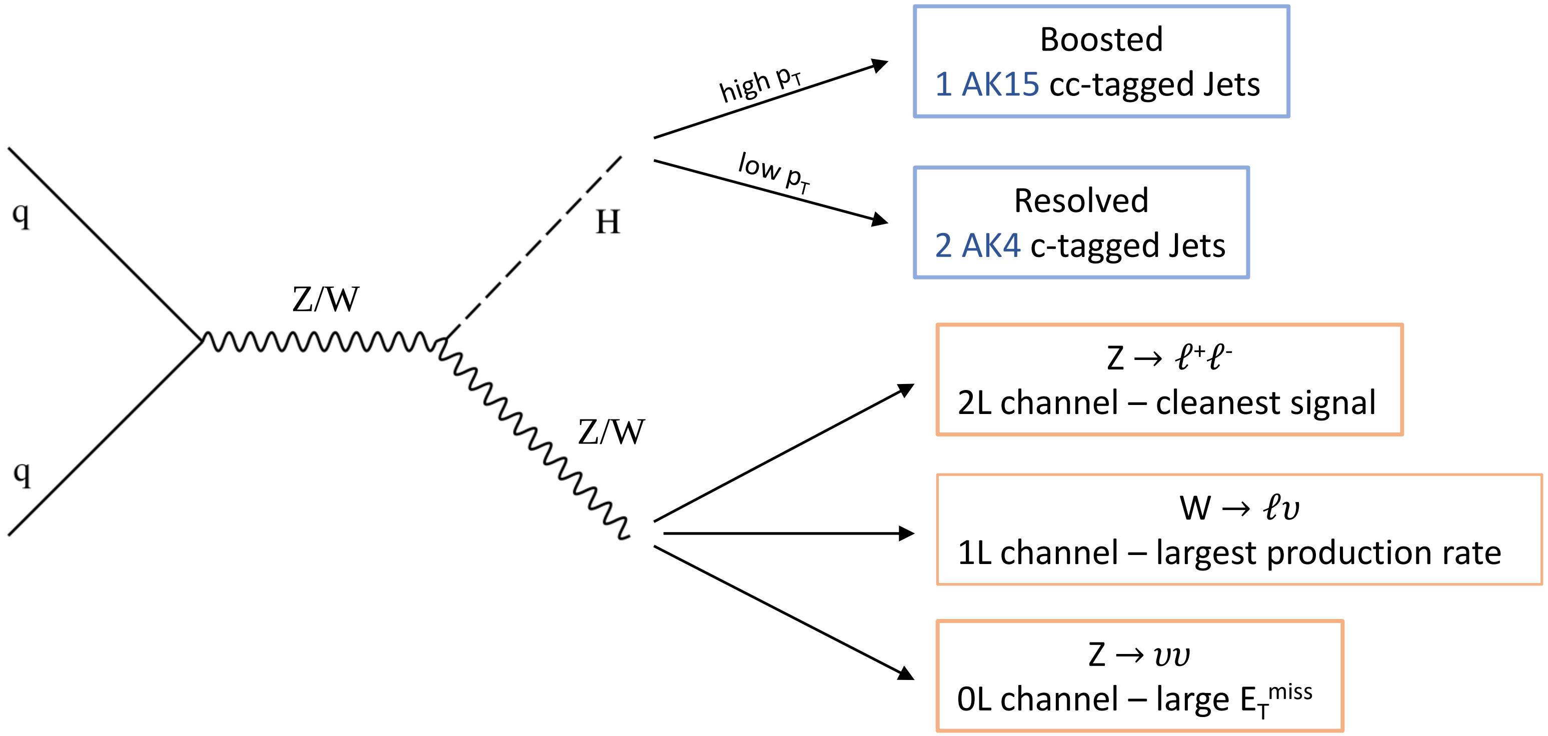




arXiv:1912.01662

CMS Analysis Strategy

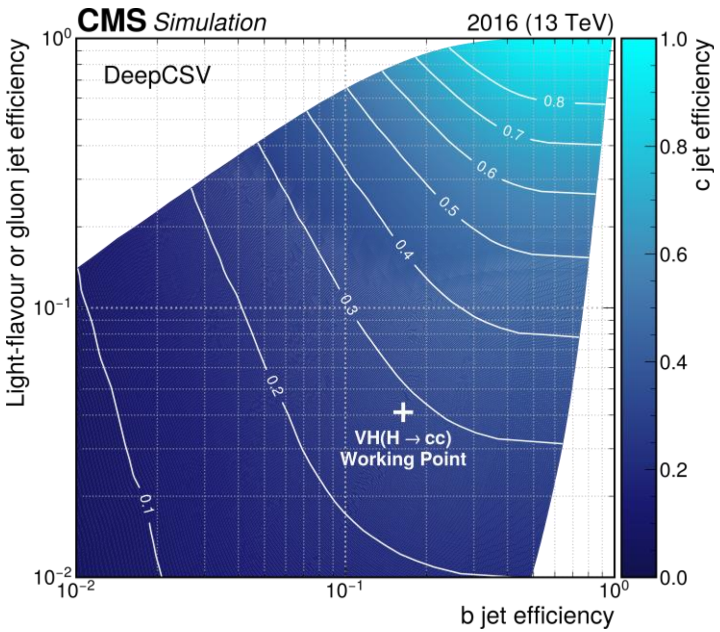
- 2016 data with L = 35.9 fb^{-1}
- Multi-channel analysis based on V type and decay
  - 2L and 1L further split by flavor of decay products (ee, $\mu\mu$, $e\nu$, $\mu\nu$)
- Split into two dedicated analyses based on $p_T(H)$
  - VH events have improved background separation at high p_T
  - 95% of signal in resolved regime

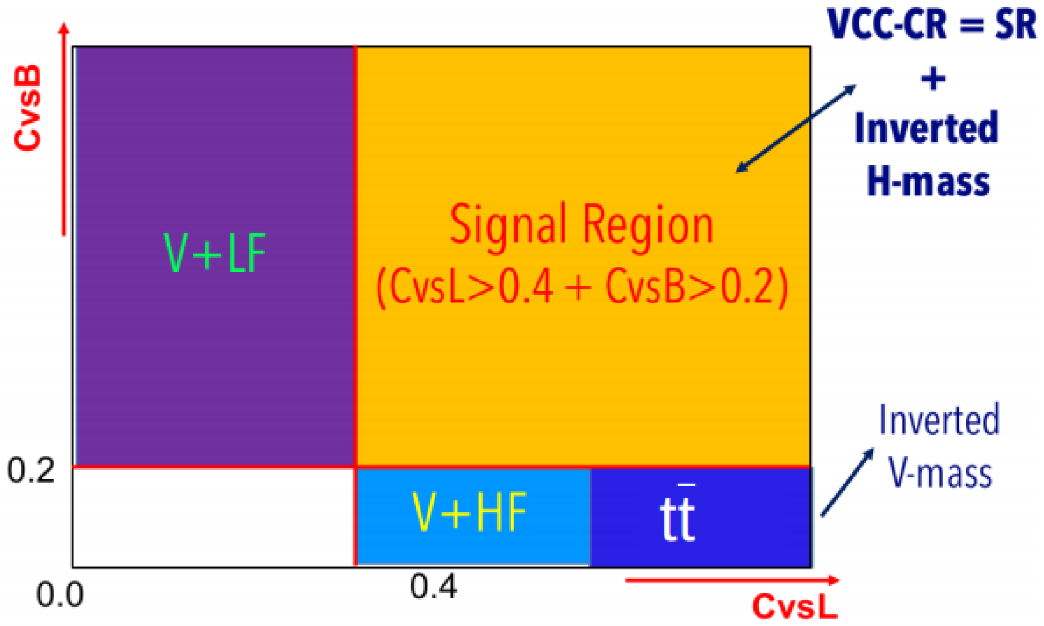


CMS – Resolved Analysis

arXiv:1912.01662
arXiv:1712.07158

- DeepCSV used to obtain c- vs l-jets and c- vs b-jets discriminants
  - DNN multi-classifier trained on track IP and secondary vertex properties
  - 28% c-jet efficiency at 15% b-jet and 4% l-jet mistag rate
  - Reshaping scale factors derived
- Higgs reconstructed from two AK4 jets with highest CvsL score





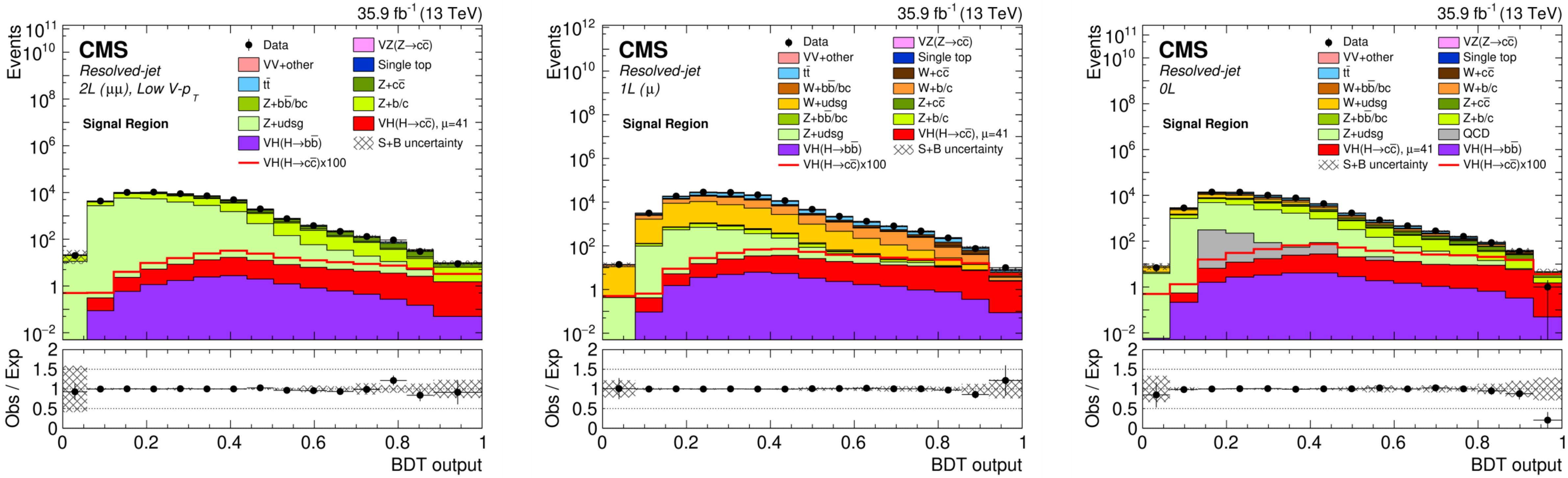
- CR used for background normalization produced using Tagger scores of H jets
  - Simultaneous normalization of 4 different backgrounds
- Background normalization derived from CvsB distribution of leading jet
  - Accurate reshaping scale factors vital for analysis

[arXiv:1912.01662](https://arxiv.org/abs/1912.01662)

CMS – Resolved Analysis

- Dedicated BDT trained on MC for signal extraction
- Simultaneous likelihood fit made to BDT in SR and CvsB in each CR

	0L	1L	2L	All Channels
Expected UL	84^{+35}_{-24}	79^{+34}_{-23}	59^{+25}_{-17}	38^{+16}_{-11}
Observed UL	66	120	116	75



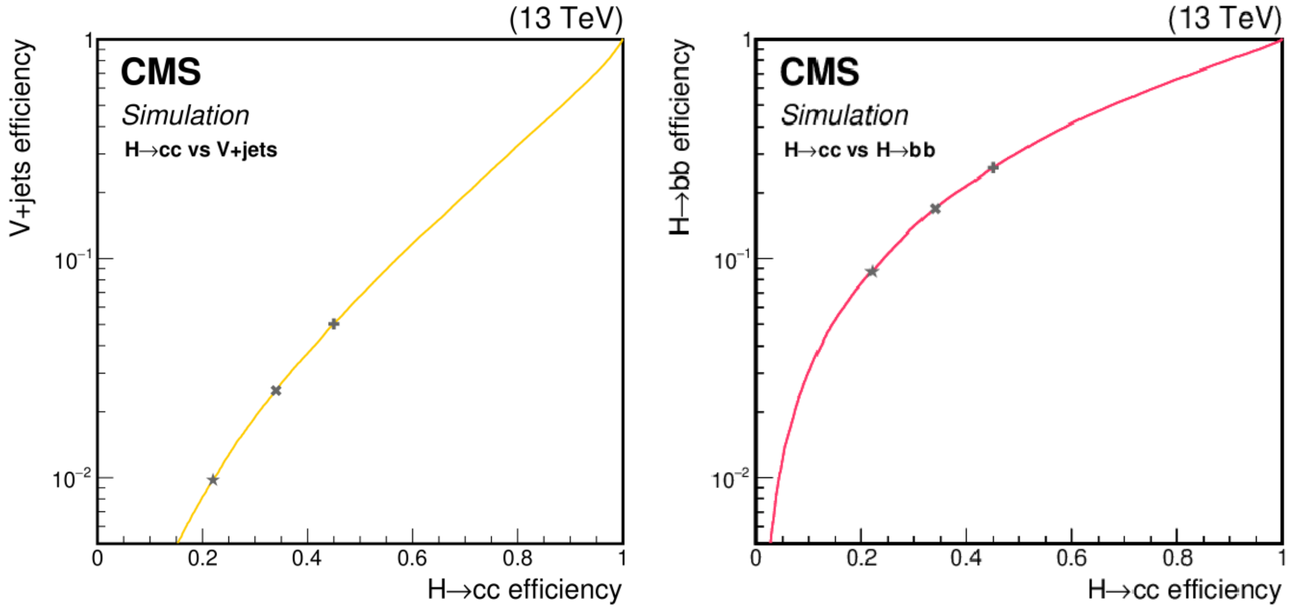
arXiv:1912.01662

CMS – Boosted Analysis

Dedicated object reconstruction and ML techniques developed to identify boosted Higgs jet

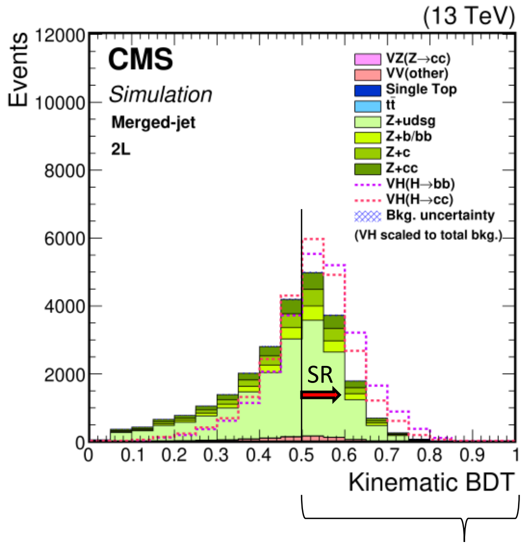
- Higgs reconstructed from the highest p_T large radius AK15 jet
  - Using a single large cone jet gives improved identification of correct jet pairs
- Dedicated DNN DeepAK15 trained to classify AK15 jets and perform $c\bar{c}$-tagging
- Events split into 3 regions based on $c\bar{c}$-tagging efficiency

	High Purity	Medium Purity	Low Purity
$c\bar{c}$-tag efficiency	23%	35%	46%
b-tag mis-Id	9%	17%	27%
Inclusive mis-Id	1%	2.5%	5%



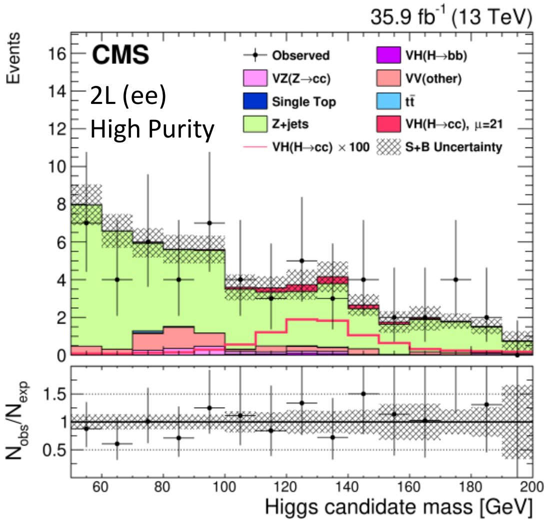
arXiv:1912.01662

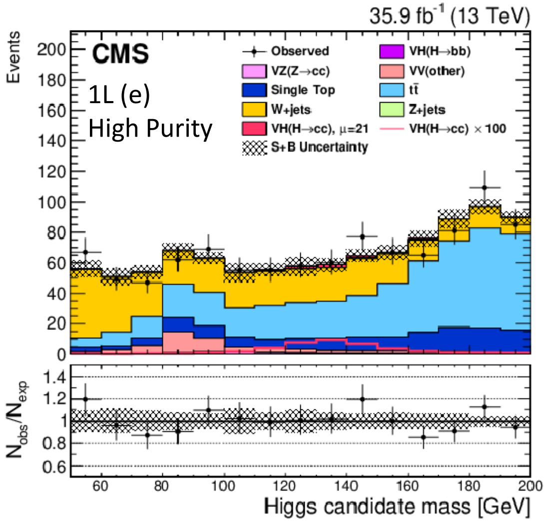
CMS – Boosted Analysis

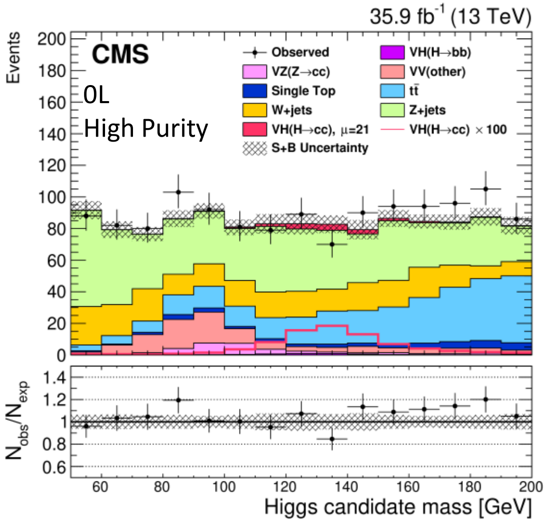


- Kinematic BDT used to produce Signal and Control regions
- Likelihood fit built from softdrop mass distribution of CRs and SRs for all purity categories

	0L	1L	2L	All Channels
Expected UL	81^{+39}_{-24}	88^{+43}_{-27}	90^{+48}_{-29}	49^{+24}_{-15}
Observed UL	74	120	76	71



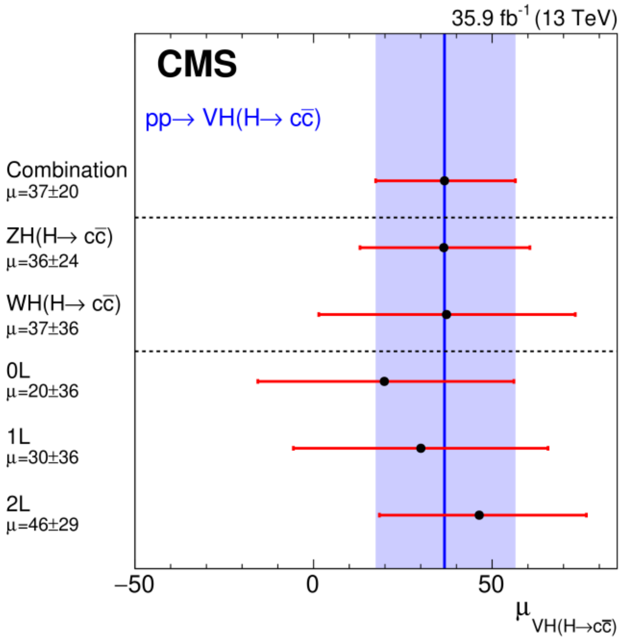


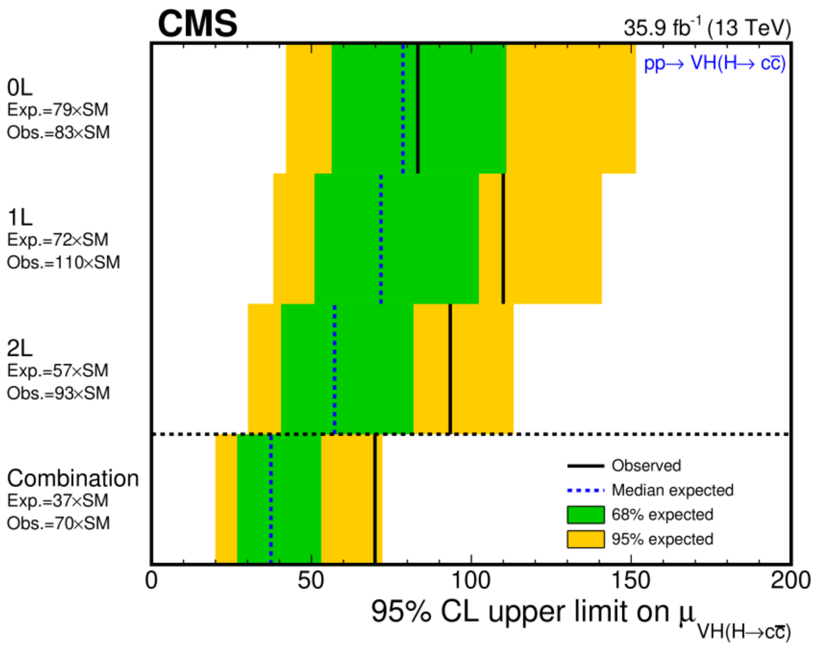


arXiv:1912.01662

CMS – Combined Analysis

- Boosted and resolved analyses combined to make final limits
- Additional $p_T(V)$ cut applied to ensure analyses are fully orthogonal
- Combined observed (expected) upper limit on μ of 70 (37^{+16}_{-11})



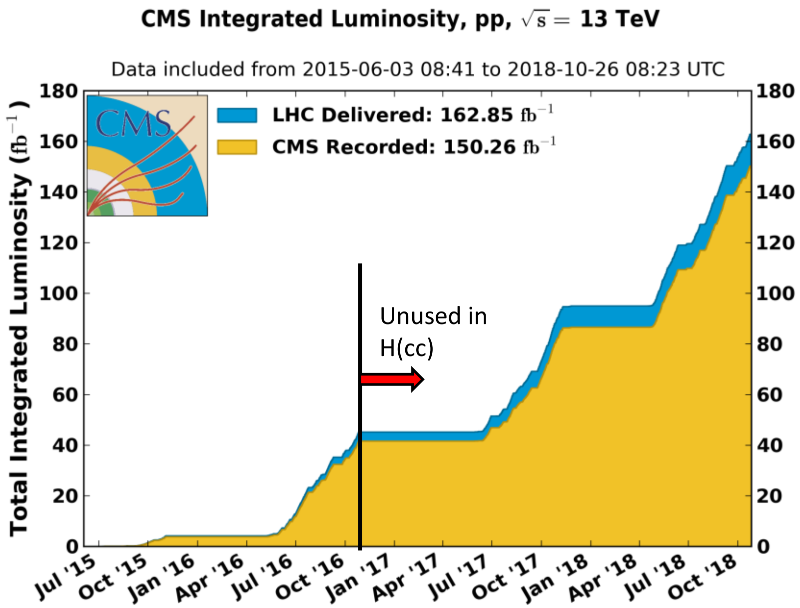


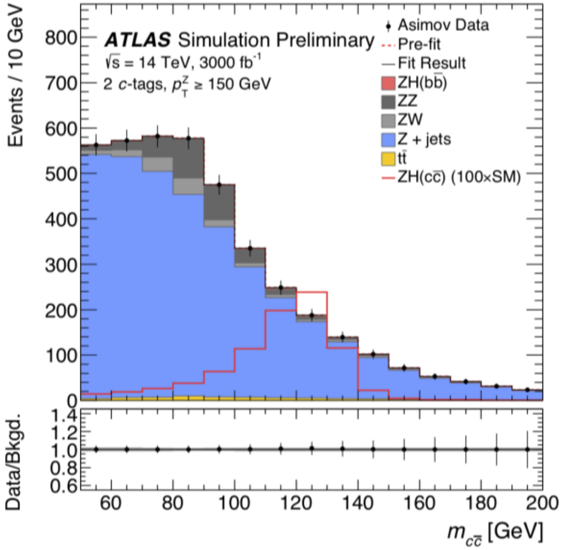
Future of $H(c\bar{c})$

arXiv:1610.07922
ATL-PHYS-PUB-2018-016

Still room for improvement

- Increase in statistics – SM prediction for $Z(\ell^+\ell^-)H(c\bar{c})$ is 26 fb which means few expected events in this search
    - Both CMS and ATLAS have used less than half of the now available run 2 data
    - Run 3 projected to provide another 150 fb^{-1}
    - HL-LHC projected to provide 3000 fb^{-1}
        - ATLAS projects sensitivity of $\mu_{ZH(c\bar{c})} < 6.3^{+2.5}_{-1.8}$
- Improvements in analysis techniques
- Improvements in c-jet tagging





Conclusion

arXiv:1802.04329
arXiv:1912.01662

- Both CMS and ATLAS have started to conduct searches for the $H \to c\bar{c}$ decay mode

- Make use of the VH production mode and analysis made possible by machine learning based c-taggers

- Observed (expected) 95% CL upper limits on signal strength at $110\ (150^{+81}_{-42})$ by ATLAS and $70\ (37^{+16}_{-11})$ by CMS

References

Analyses can be further studied in the following papers:

- Search for the Decay of the Higgs Boson to Charm Quarks with the ATLAS Experiment: 10.1103/PhysRevLett.120.211802
- A search for the standard model Higgs boson decaying to charm quarks: JHEP 03 (2020) 131

Other References:

- Constraining the Charm Yukawa and Higgs-quark Coupling Universality (arXiv:1503.00290)
- Evidence of Higgs boson decay to pair of muons (arXiv:2009.04363)
- A search for the dimuon decay of the Standard Model Higgs boson with the ATLAS Detector (arXiv:2007.07830)
- Handbook of LHC Higgs Cross Sections: 4. Deciphering the Nature of the Higgs Sector (arXiv:1610.07922)
- Identification of heavy-flavour jets with the CMS detector in pp collisions at 13 TeV (10.1088/1748-0221/13/05/P05011)
- The anit-k_t jet clustering algorithm (arXiv:0802.1189)
- Expected performance of the ATLAS b-tagging algorithms in Run-2 (ATL-PHYS-PUB-2016-012)
- Optimization of the ATLAS b-tagging performance for the 2016 LHC Run (ATL-PHYS-PUB-2016-012)

Questions?

Backup

Analysis Validation

arXiv:1802.04329
arXiv:1912.01662

Both CMS and ATLAS performed a validation procedure by measuring $Z \rightarrow c\bar{c}$ branching ratio with same analysis strategy

Reperformed analysis with $ZZ(c\bar{c})$ and $WZ(c\bar{c})$ as signal

	ATLAS	CMS (resolved)	CMS (merged)
Signal Strength $\mu_{VZ(Z\rightarrow c\bar{c})}$	$0.6^{+0.5}_{-0.4}$	$1.35^{+0.94}_{-0.95}$	$0.69^{+0.89}_{-0.75}$
Observed (expected) significance	$1.4\ (2.2)\sigma$	$1.5\ (1.2)\sigma$	$0.9\ (1.3)\sigma$

arXiv:0802.1189

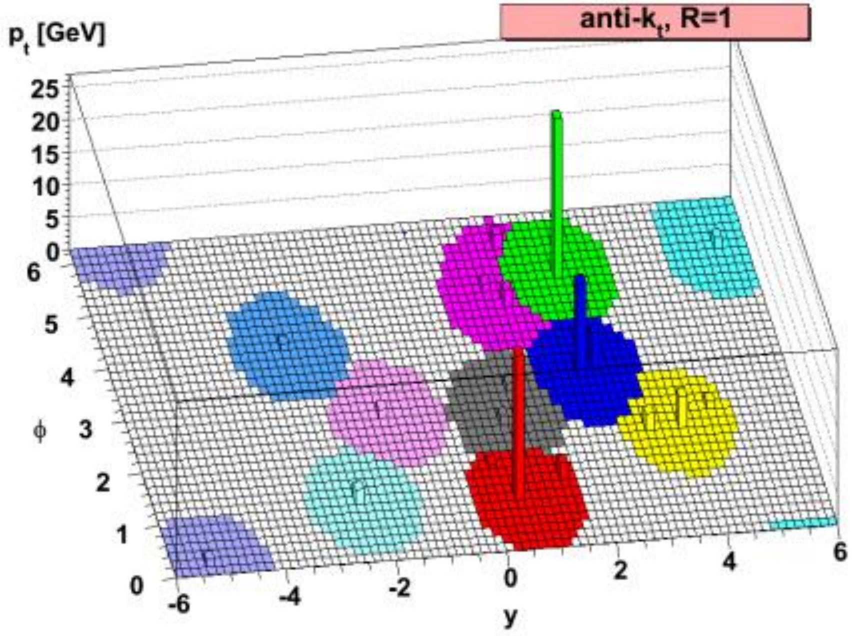
Anti k_T Algorithm

- Aims to form jets by clustering soft particles to the closest hard particle

- Does this by clustering particles based off a distance parameter

$$d_{ij} = \min(k_{t\ i}^{2p}, k_{t\ j}^{2p}) \frac{\Delta_{ij}^2}{R^2}$$

$p = -1$, Δ_{ij} is the angular separation, R is the max jet radius



ATLAS Triggers

arXiv:1802.04329

	2015	2016
Single Electron	$p_T(e) > 24$ GeV	$p_T(e) > 26$ GeV
Single Muon	$p_T(\mu) > 20$ GeV	$p_T(\mu) > 26$ GeV

ATLAS c-tag BDT

ATL-PHYS-PUB-2016-012
ATL-PHYS-PUB-2015-022

Input	Variable	Description
Kinematics	$p_T(jet)$	Jet transverse momentum
	$\eta(jet)$	Jet pseudo-rapidity
IP2D, IP3D	$\log(P_b/P_{\text{light}})$	Likelihood ratio between the b- and light jet hypotheses
	$\log(P_b/P_c)$	Likelihood ratio between the b- and c-jet hypotheses
	$\log(P_c/P_{\text{light}})$	Likelihood ratio between the c- and light jet hypotheses
SV	$m(\text{SV})$	Invariant mass of tracks at the secondary vertex assuming pion masses
	$f_E(\text{SV})$	Fraction of the charged jet energy in the secondary vertex
	$N_{\text{TrkAtVtx}}(\text{SV})$	Number of tracks used in the secondary vertex
	$N_{\text{2TrkVtx}}(\text{SV})$	Number of two track vertex candidates
	$L_{xy}(\text{SV})$	Transverse distance between the primary and secondary vertices
	$L_{xyz}(\text{SV})$	Distance between the primary and secondary vertices
	$S_{xyz}(\text{SV})$	Distance between the primary and secondary vertices divided by its uncertainty
	$\Delta R(\text{jet}, \text{SV})$	ΔR between the jet axis and the direction of the secondary vertex relative to the primary vertex
Jet Fitter	$N_{\text{2TrkVtx}}(\text{JF})$	Number of 2-track vertex candidates (prior to decay chain fit)
	$m(\text{JF})$	Invariant mass of tracks from displaced vertices assuming pion masses
	$S_{xyz}(\text{JF})$	Significance of the average distance between the primary and displaced vertices
	$f_E(\text{JF})$	Fraction of the charged jet energy in the secondary vertices
	$N_{\text{1-trk vertices}}(\text{JF})$	Number of displaced vertices with one track
	$N_{\geq\text{2-trk vertices}}(\text{JF})$	Number of displaced vertices with more than one track
	$N_{\text{TrkAtVtx}}(\text{JF})$	Number of tracks from displaced vertices with at least two tracks
	$\Delta R(\vec{p}_{\text{jet}}, \vec{p}_{\text{vtx}})$	ΔR between the jet axis and the vectorial sum of the momenta of all tracks attached to displaced vertices

arXiv:1802.04329

ATLAS Systematics and Yields

Source	$\sigma/\sigma_{\text{tot}}$
Statistical	**49%**
Floating $Z+$jets normalization	31%
Systematic	**87%**
Flavor tagging	73%
Background modeling	47%
Lepton, jet and luminosity	28%
Signal modeling	28%
MC statistical	6%

Sample	Yield, 50 $GeV < m_{c\bar{c}} <$ 200 GeV			
	1 c-tag		2 c-tags	
	$75 \leq p_{\text{T}}^{Z} < 150\,GeV$	$p_{\text{T}}^{Z} \geq 150\,GeV$	$75 \leq p_{\text{T}}^{Z} < 150\,GeV$	$p_{\text{T}}^{Z} \geq 150\,GeV$
$Z+$jets	69400 ± 500	15650 ± 180	5320 ± 100	1280 ± 40
ZW	750 ± 130	290 ± 50	53 ± 13	20 ± 5
ZZ	490 ± 70	180 ± 28	55 ± 18	26 ± 8
$t\bar{t}$	2020 ± 280	130 ± 50	240 ± 40	13 ± 6
$ZH(b\bar{b})$	32 ± 2	19.5 ± 1.5	4.1 ± 0.4	2.7 ± 0.2
$ZH(c\bar{c})$ (SM)	-143 ± 170 (2.4)	-84 ± 100 (1.4)	-30 ± 40 (0.7)	-20 ± 29 (0.5)
Total	72500 ± 320	16180 ± 140	5650 ± 80	1320 ± 40
Data	72504	16181	5648	1320

CMS Triggers

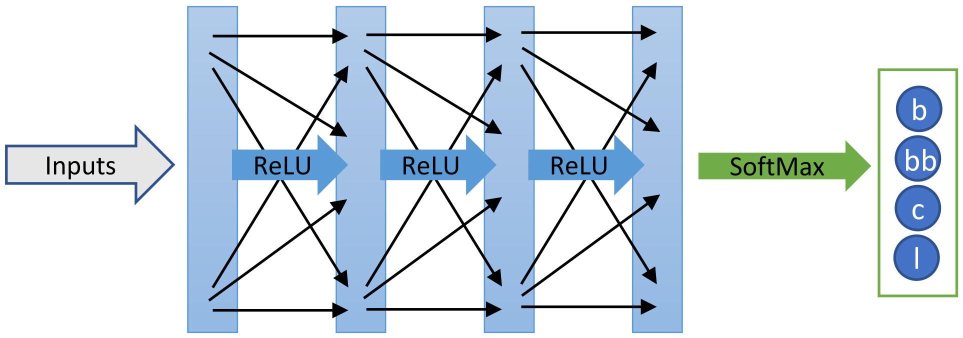
arXiv:1912.01662

0L	1L	2L
$p_T^{\text{miss}} > 170$ GeV	$p_T(e) > 27$ GeV	$p_T(e_1) > 23$ GeV && $p_T(e_2) > 12$ GeV
$p_T^{\text{miss}} > 110$ GeV && $\mathrm{H}_T^{\text{miss}} > 110$ GeV	$p_T(\mu) > 24$ GeV	$p_T(\mu_1) > 23$ GeV && $p_T(\mu_2) > 12$ GeV (resolved only)
		$p_T(\mu) > 24$ GeV (merged only)

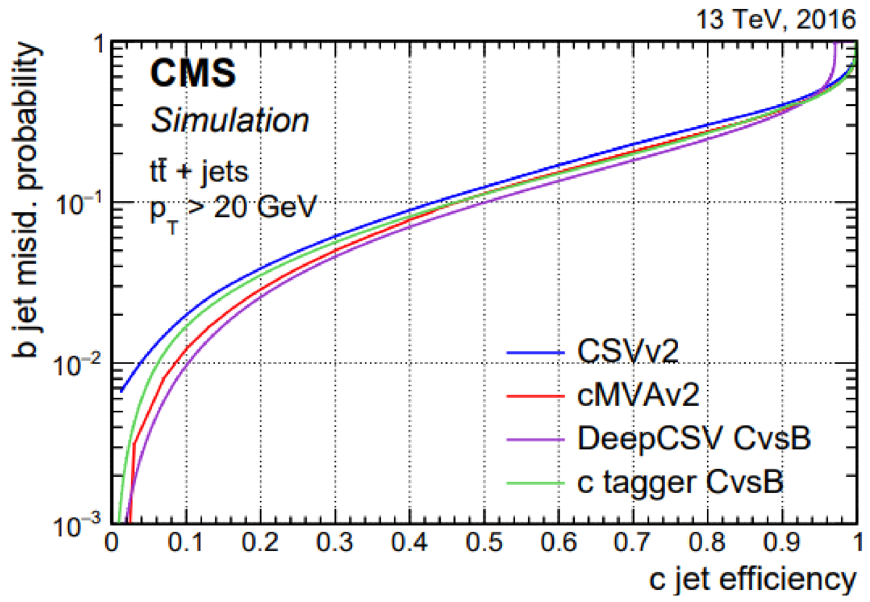
10.1088/1748-0221/13/05/P05011

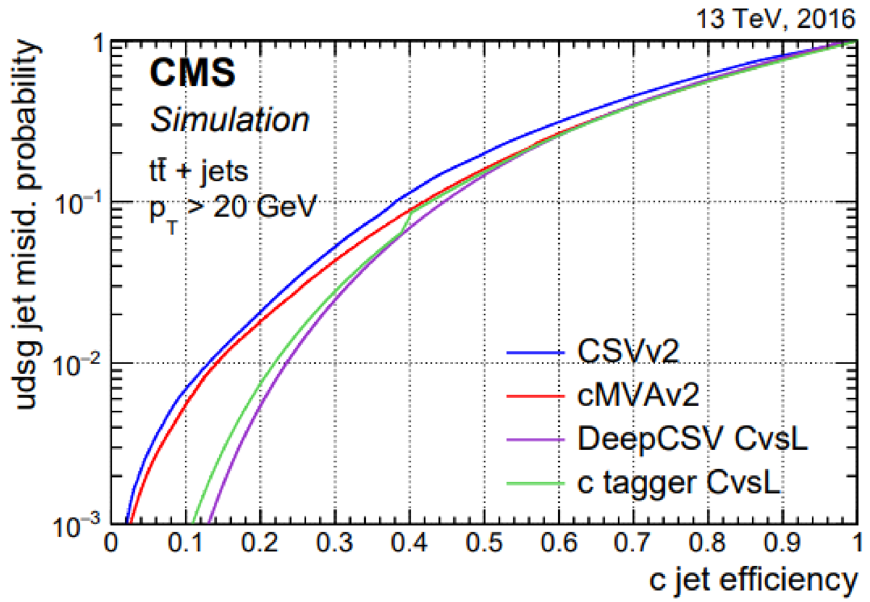
AK4 Jet Tagging – Deep CSV

- Deep network which classifies jets as coming from b, bb, c, or l
- 66 Input features with jet kinematic, track, and secondary vertex information



4 hidden layers with 100 nodes



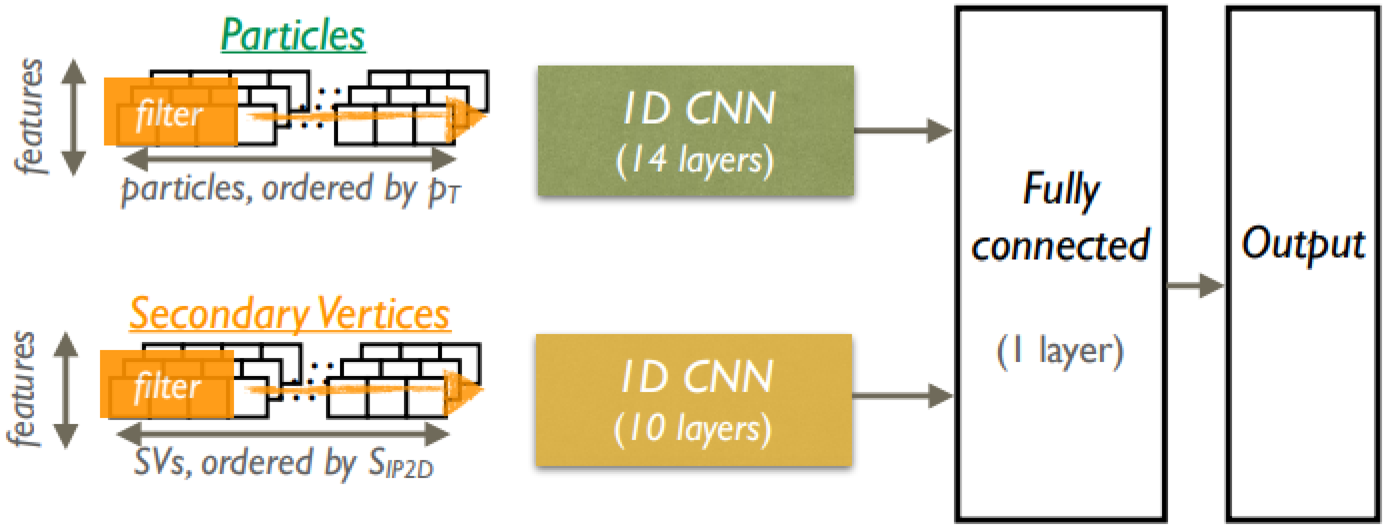


arXiv:1912.01662

CMS Resolved – BDT Variables

Variable	Description	0L	1L	2L
$m(\mathrm{H})$	H mass	✓	✓	✓
p_T (H)	H transverse momentum	✓	✓	✓
$p_\mathrm{T}(\mathrm{V})$	vector boson transverse momentum	✓	✓	✓
$m(\mathrm{V})$	vector boson mass	—	—	✓
$m_\mathrm{T}(\mathrm{V})$	vector boson transverse mass	—	✓	—
$p_\mathrm{T}^\mathrm{miss}$	missing transverse momentum	✓	✓	—
$p_\mathrm{T}(\mathrm{V})/p_\mathrm{T}(\mathrm{H})$	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	✓	✓	✓
$\Delta\phi(\mathrm{V},\mathrm{H})$	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	—	✓	—
$\mathrm{N}^\mathrm{aj}_{\text{small-}R}$	number of small-R jets minus the number of FSR jets	✓	✓	✓
N_5^{soft}	multiplicity of soft track-based jets with $p_\mathrm{T} > 5\,\mathrm{GeV}$	✓	✓	✓

AK15 Jet Tagging – DeepAK15



Particles

- Up to 100 PF candidates
- Sorted in descending p_T order
- Uses basic kinematic variables, PUPPI weights, and track properties (quality, covariance, displacement, etc.)

Secondary Vertices

- Up to 5 SVs (inside jet cone)
- Sorted in descending S_{IP2D} order
- Uses SV kinematics and properties (quality, displacement, etc)

arXiv:1912.01662

CMS Boosted – BDT Variables

Variable	Description	0L	1L	2L
$p_T(V)$	vector boson transverse momentum	✓	✓	✓
p_T (H)	H transverse momentum	✓	✓	✓
$\|\eta(H)\|$	absolute value of the H pseudorapidity	✓	—	—
$\Delta\phi(V, H)$	azimuthal angle between vector boson and H	✓	✓	✓
p_T^{miss}	missing transverse momentum	—	✓	—
$\Delta\eta(H, \ell)$	difference in pseudorapidity between H and the lepton	—	✓	—
$\Delta\eta(H, V)$	difference in pseudorapidity between H and vector boson	—	—	✓
$\Delta\eta(H, j)$	min. difference in pseudorapidity between H and small-R jets	✓	✓	✓
$\Delta\eta(\ell, j)$	min. difference in pseudorapidity between the lepton and small-R jets	—	✓	—
$\Delta\eta(V, j)$	min. difference in pseudorapidity between vector boson and small-R jets	—	—	✓
$\Delta\phi(\vec{p}_T^{miss}, j)$	azimuthal angle between $\vec{p}_T^{miss}$ and closest small-R jet	✓	—	—
$\Delta\phi(\vec{p}_T^{miss}, \ell)$	azimuthal angle between $\vec{p}_T^{miss}$ and lepton	—	✓	—
m_T	transverse mass of lepton $\vec{p}_T + \vec{p}_T^{miss}$	—	✓	—
$N_{small\text{-}R}^{aj}$	number of additional small-R jets	✓	✓	✓

arXiv:1912.01662

CMS – Uncertainties

Uncertainty source	$\Delta\mu \mid \mu = 37$	
Statistical	+17.3	−17.1
Background normalisations	+10.1	−10.2
Experimental	+7.6	−8.2
Charm tagging efficiencies	+5.6	−4.8
Simulation modeling	+4.2	−5.1
Jet energy scale and resolution	+2.4	−2.8
Lepton identification efficiencies	+0.4	−1.8
Luminosity	+1.6	−1.7
Statistics of the simulated samples	+0.5	−1.9
Theory	+6.5	−4.6
Signal	+5.0	−2.5
Backgrounds	+4.3	−3.9
Total	+20.0	−19.5

arXiv:1912.01662

CMS Upper Limits (w/ overlap)

	Resolved-jet (inclusive)				Merged-jet (inclusive)			
	0L	1L	2L	All channels	0L	1L	2L	All channels
Expected UL	84^{+35}_{-24}	79^{+34}_{-23}	59^{+25}_{-17}	38^{+16}_{-11}	81^{+39}_{-24}	88^{+43}_{-27}	90^{+48}_{-29}	49^{+24}_{-15}
Observed UL	66	120	116	75	74	120	76	71

arXiv:1902.00134

HL-LHC Prospects

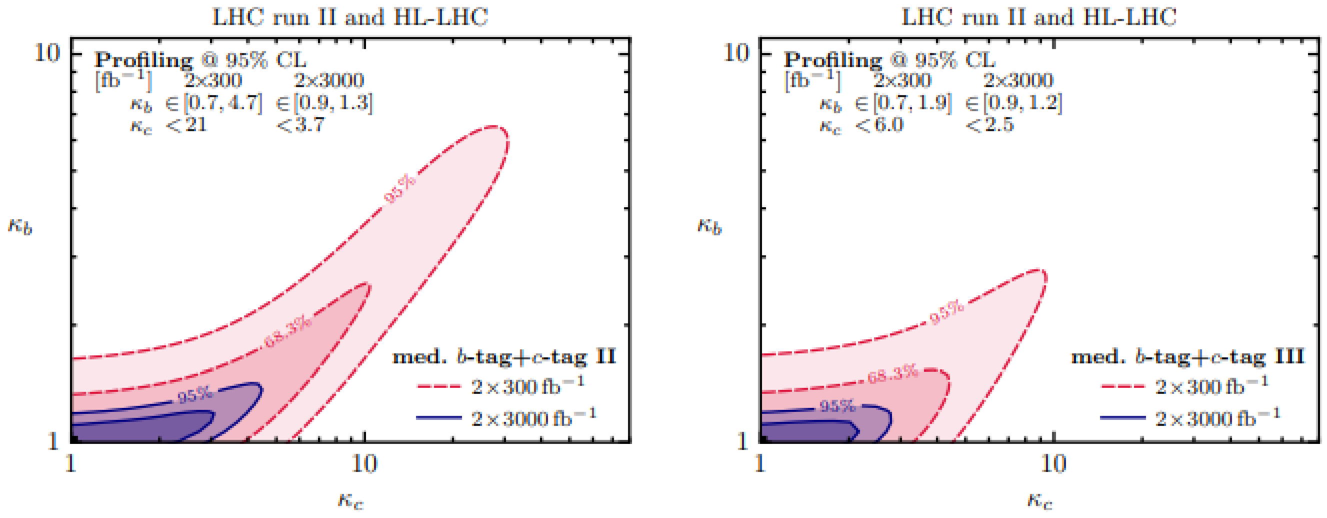


Fig. 120: Projections for measuring charm Yukawa modifications from an inclusive $H \to c\bar{c}$ search at $\sqrt{s} = 14\,\text{TeV}$ using two different c-taggers (left and right panel) [616]. In red the 95% CL region employing an integrated luminosity of $2 \times 300\ \text{fb}^{-1}$ and in blue the region employing $2 \times 3000\ \text{fb}^{-1}$.

Higgs Physics at the HL-LHC and HE-LHC, ch7.3.1