New experimental techniques and new analysis ideas for probes of quark Yukawa interactions (excluding top)

António Jacques Costa,
on behalf of the ATLAS and CMS collaborations

LEVERHULME TRUST

MANCHESTER 1824
The University of Manchester

Standard Model at the LHC 2024

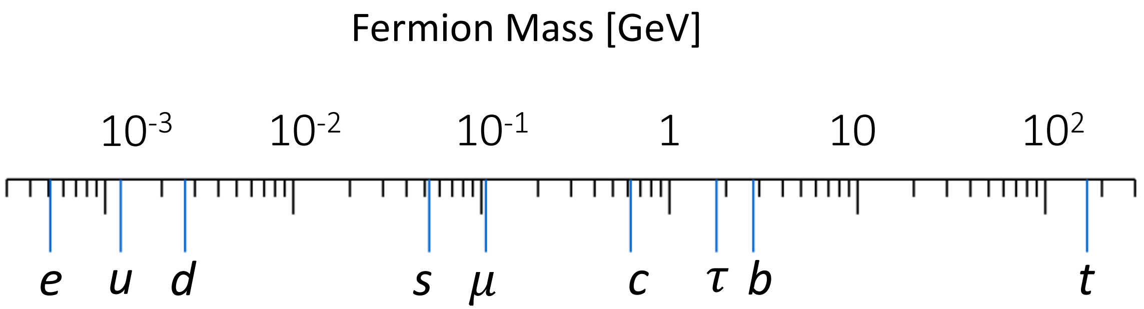
9th May 2024

CMS

ATLAS EXPERIMENT

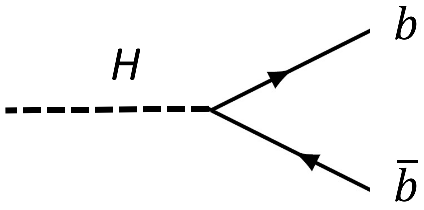
Introduction

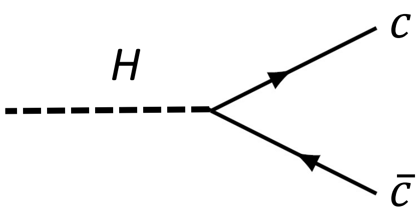
- Couplings of Higgs field to quarks and leptons – Yukawa couplings - are a potential source of the fermion masses
- Interaction so far only observed for 3rd generation of fermions (top, bottom and tau) and evidence found for coupling with muons
- Measurement of Yukawa couplings probe Standard Model expectation
- There is no guarantee all Higgs fermion couplings behave in a similar way
- Deviations could give insight into origin of the fermion mass hierarchy
- Of utmost importance to measure all Higgs couplings to fermions with best possible precision!



Heavy quark Yukawa couplings

- Probability of Higgs boson decays to bottom/charm quarks of 58.2%/3.9% in Standard Model
- **Standard Model Higgs** Yukawa **coupling** to charm quarks is rather **small** ($y_c = \sqrt{2}\, m_c(\mu{=}m_H)/\nu \simeq 0.2 \times y_b$)
- **One of largest contributions to Γ_H** (by SM expectations) **yet to be established experimentally**
- Both decay modes are susceptible to **significant modifications** in some **new physics** scenarios (e.g. two Higgs doublet models, EFTs)
- H → $b\bar{b}$ analyses in **precision measurement** phase of cross-section and Yukawa coupling
- H → $c\bar{c}$ analyses in search phase, **improving sensitivity** to decay and coupling





Probe options

- Different probes of the $Hb\bar{b}/c\bar{c}$ coupling have been proposed and investigated by the ATLAS and CMS experiments

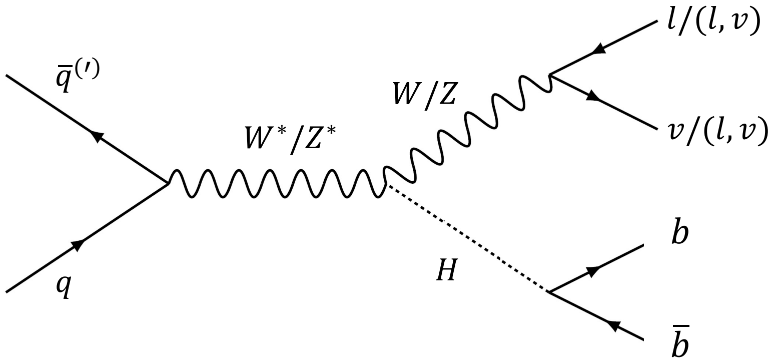
- **Inclusive** $H \to b\bar{b}/c\bar{c}$ **decays**
  - Direct access to coupling
  - Usually targeting **VH production (V = W, Z** boson**)**
    - **Enhanced Signal over Background ratio** w.r.t to inclusive Higgs production
    - Built around the use of **b/c-jet tagging algorithms**

    Leptonic mode

    - W/Z boson decays into leptons allow for a **convenient trigger strategy**
    - **Suppression of multi-jet backgrounds**

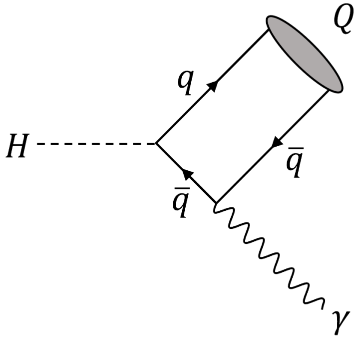
    All hadronic mode

    - W/Z boson decays into quarks, **increasing statistics**
    - **Dominant multi-jet background**

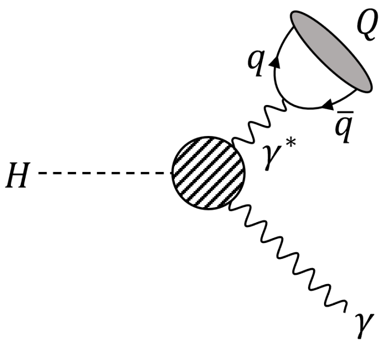


Probe options II

- **Exclusive** H→J/ψ γ **decays** (other rare decay modes can also probe lighter quark couplings)
  - Rare decay but **experimentally clean** probe (Phys. Rev. D 88 (2013) 053003)
  - Direct amplitude sensitive to the **magnitude and sign of the Higgs-charm coupling**
  - **Indirect amplitude** constitutes **dominant** contribution to the Higgs boson width
  - Production rate is dominated by the indirect contribution, hindering the sensitivity to the coupling



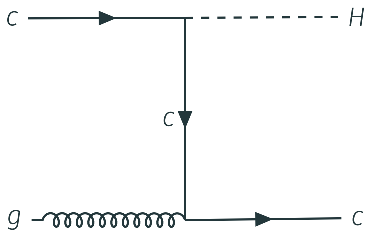
Direct contribution

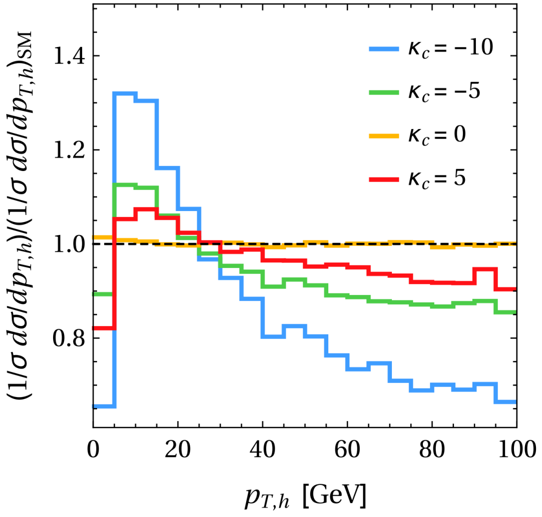


Indirect contribution

Probe options III

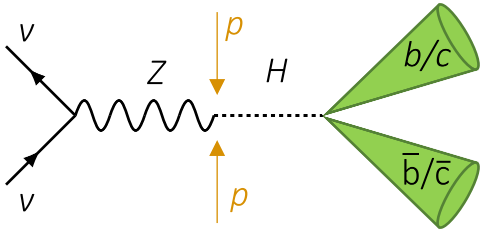
- **Loop induced gluon-gluon fusion and quark initiated Higgs boson production**
    - **Higgs production with quark contribution**, e.g., via gg→ Hg, gb/gc→ Hb/Hc, $b\bar{b}$/$c\bar{c}$→ Hg (Phys. Rev. Lett. 118 (2017) 121801)
    - Sensitivity to coupling via total production **differential cross-section measurements**
    - **p_T^H shape and normalisation** will be changed in case of Yukawa couplings different to SM prediction
    - Normalisation effects coupled to Higgs width changes
    - Obtained **constraints on $Hb\bar{b}$ ($Hc\bar{c}$)** coupling **not as stringent (comparable)** as the ones set by inclusive Higgs decays searches
    - Subject to **different assumptions and sources of uncertainty**
    - Complementary **sensitivity to the sign** of the coupling modifiers





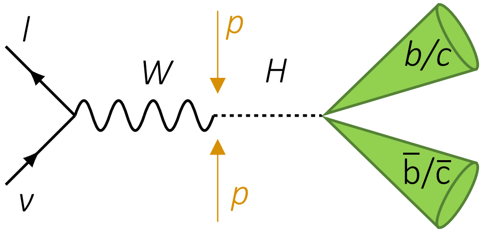
VH Production: “Resolved” regime

Note:
l = e, mu



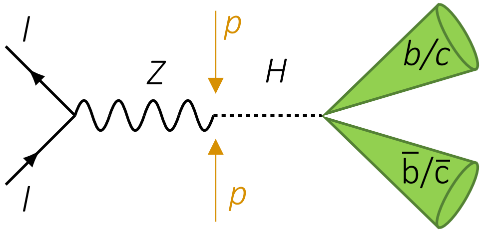
0 lepton channel

- Targets Z(*νν*)H(bb̅/cc̅) production
- Large missing transverse energy in events (due to neutrinos)



1 lepton channel

- Targets W(*lν*)H(bb̅/cc̅) production
- Large missing transverse energy
- One electron or muon



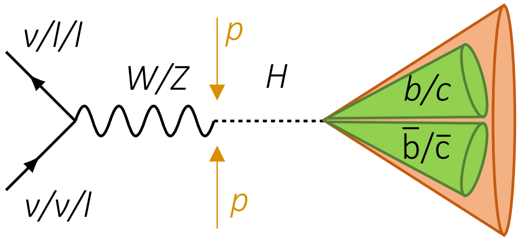
2 lepton channel

- Targets Z(*ll*)H(bb̅/cc̅) production
- e^+e^- or $\mu^+\mu^-$ pairs

Drives sensitivity up to p_T^V of 250-400 GeV (where p_T^V is p_T of associated vector boson produced)

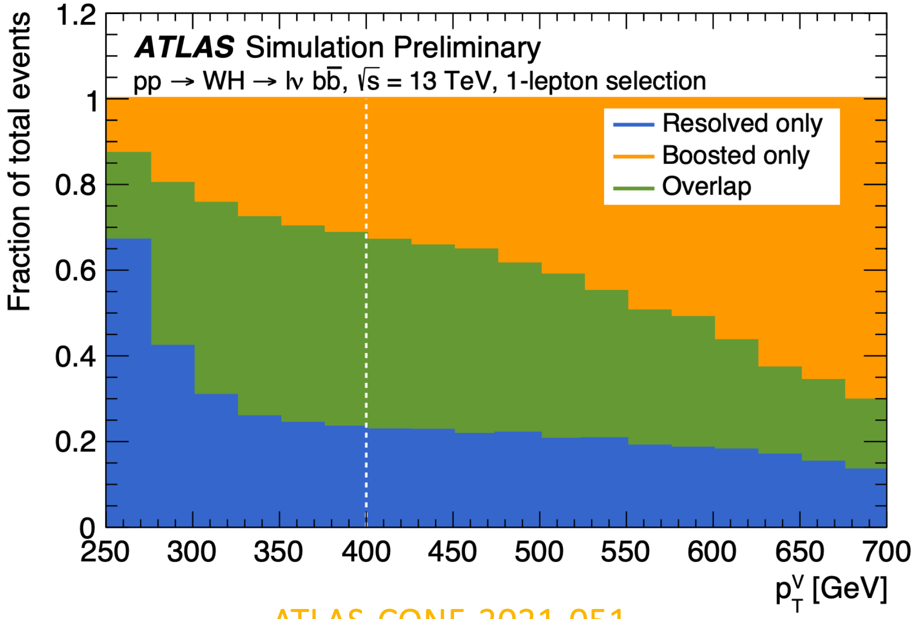
VH Production: “Boosted” leptonic regime

Note:
l = e, mu



0/1/2 lepton channel

- Targets Z(νν)H(bb̄/cc̄), W(lν)H(bb̄/cc̄), Z(ll)H(bb̄/cc̄) production
- Large radius jet, with substructure exploited



ATLAS-CONF-2021-051

Resolved/boosted overlap

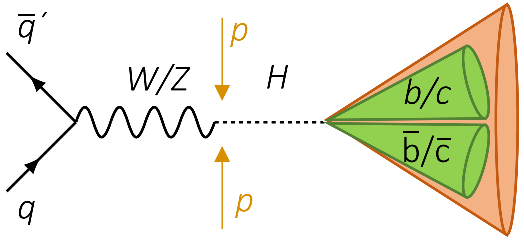
- Slightly different treatments between ATLAS and CMS
- Common factor – priority in event allocation given to resolved regime in overlap region as improves expected sensitivity and decreases uncertainties
- Events used only once

VH Production: "Boosted" hadronic regime

D/GNN: Deep/Graph Neural Network

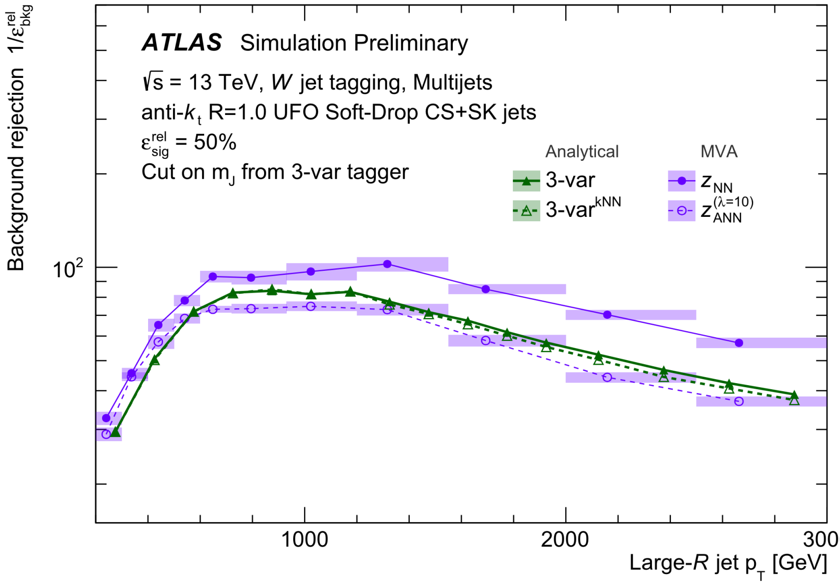
Hadronic channel

- Targets V(qq)H($b\bar{b}$/$c\bar{c}$) production
- Vector boson tagged via dedicated taggers

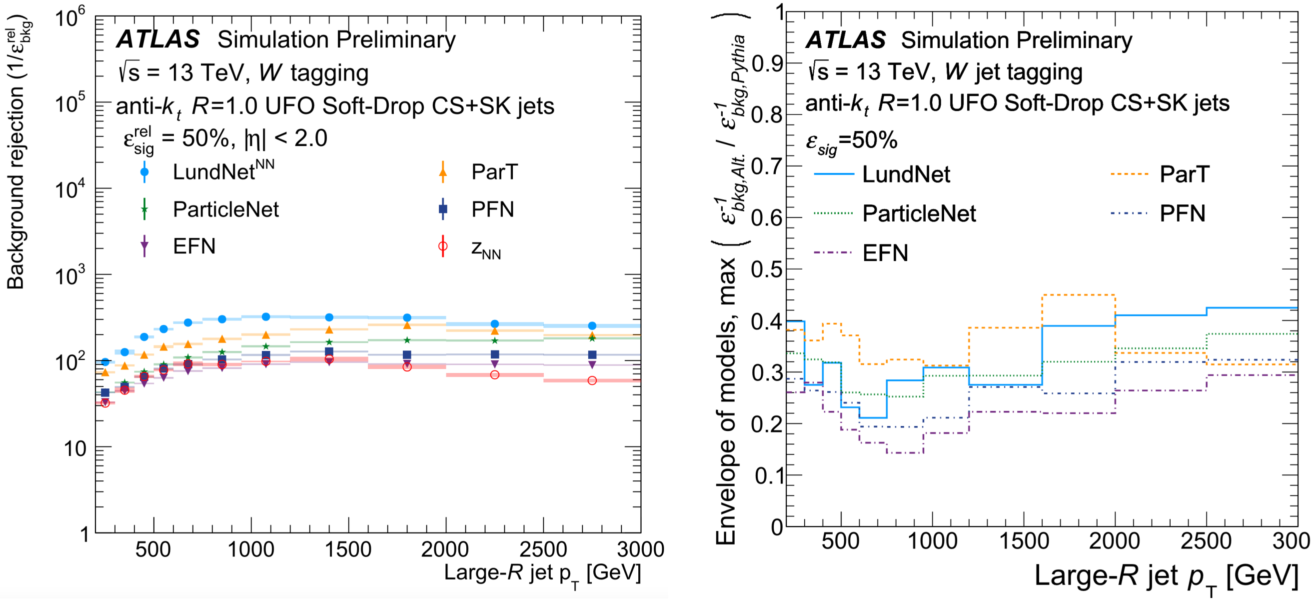


Hadronically decaying W/Z boson tagging

- First ATLAS **hadronic VH, H($b\bar{b}$)** analysis (PhysRevLett.132.131802) made use of a cut-based tagger, exploiting jet mass, number of tracks and substructure variables as energy correlation function ratios
- Performance improvements found in **DNN/GNN and transformers** architectures using jet constituents' energy and kinematics



JETM-2022-006

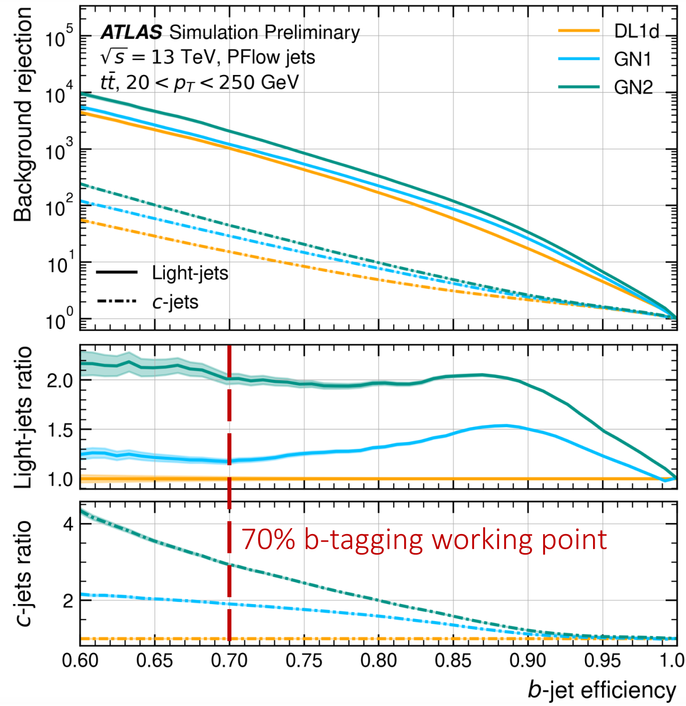


JETM-2023-003

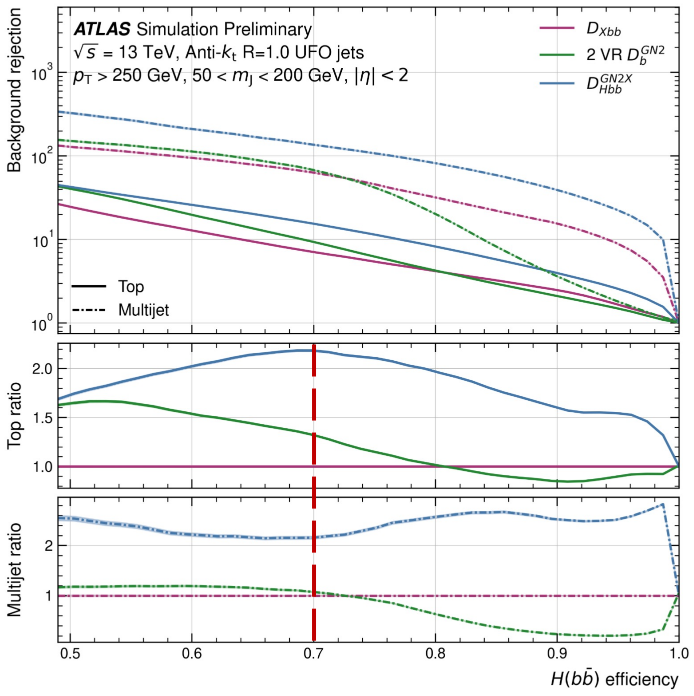
VH Production: Jet Heavy Flavour Tagging

CNN: Convolutional Neural Network

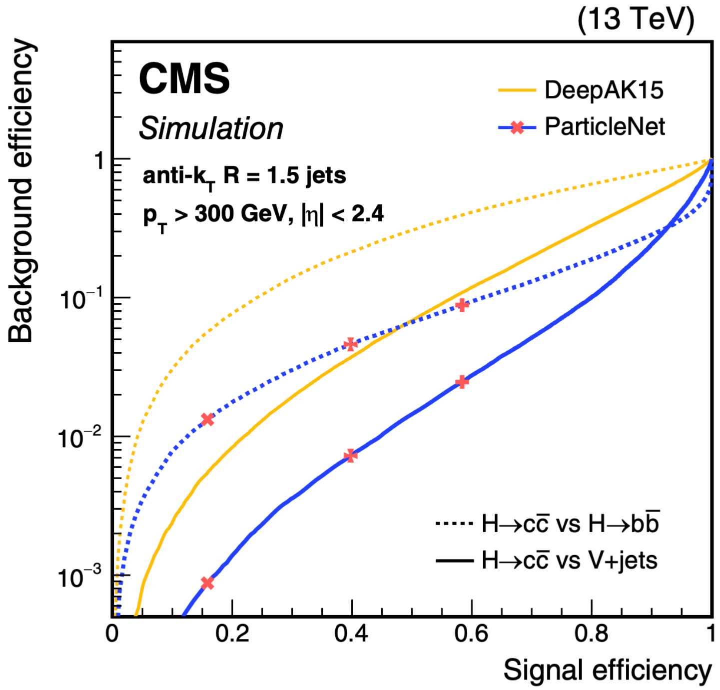
- **Bottom or charm Higgs candidate jet**s identified via **jet flavour tagging algorithms**
- Taggers optimised separately for **resolved or boosted regime**
- Targeting both **Higgs to bottom and charm decays**
- **DNN/CNN-based taggers** being used in most of latest analyses
- Further improvements found in **graph and transformer neural network** architectures, by **better exploiting jet internal structure**



ATLAS GNN Flavour Tagging



ATL-PHYS-PUB-2023-021



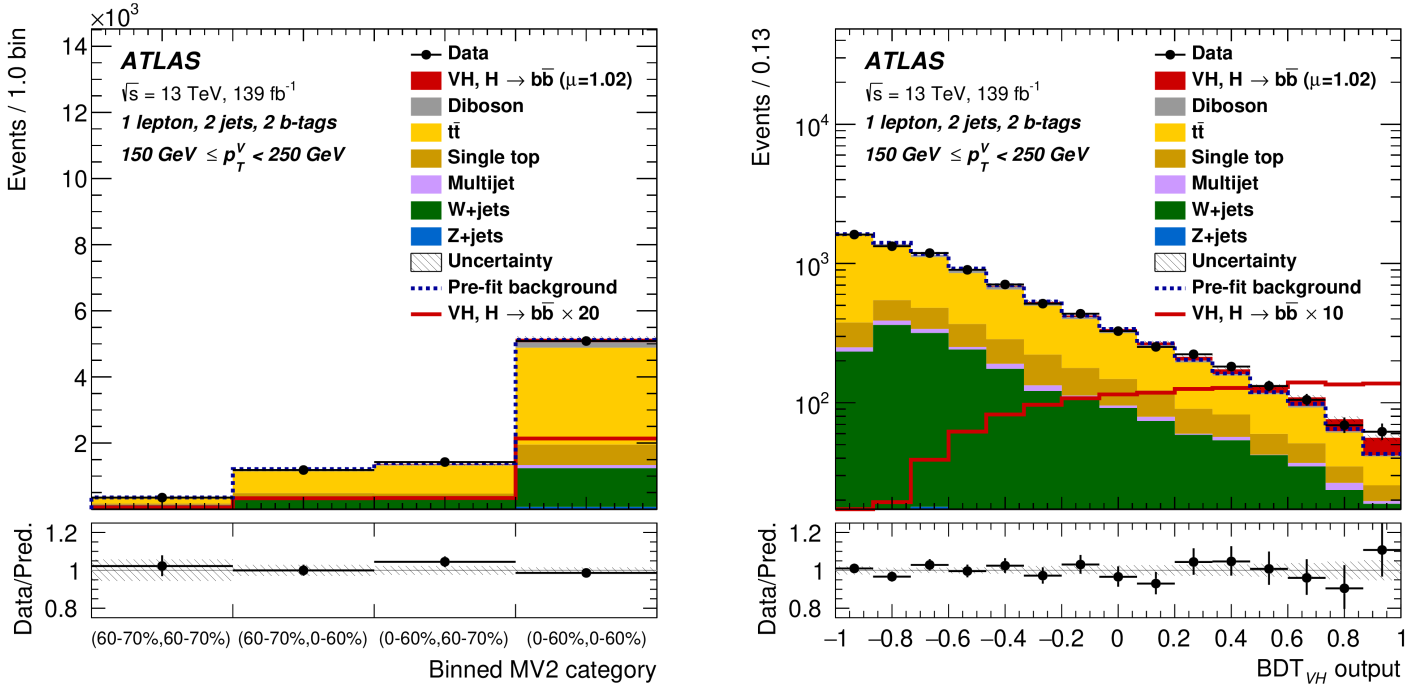
PhysRevLett.131.061801

VH Production: Event selection

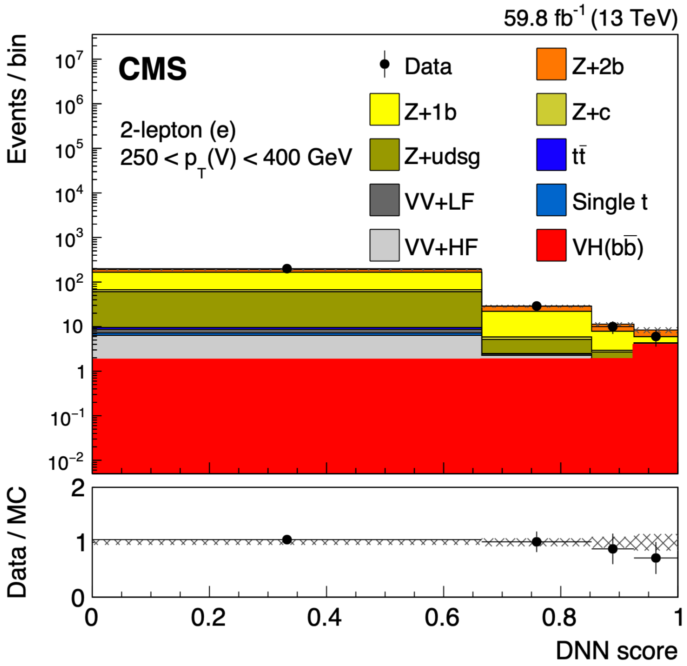
BDT: Boosted Decision Tree

Key aspects

- **Categorisation** of events for **optimised selections and signal sensitivity**
- **Data-driven methods/control regions** to estimate ttbar, W/Z+jets and multijet backgrounds
- Flavour tagging discriminants used along with jet mass and transverse momentum to **discriminate VH production from backgrounds** via BDTs/NNs



[Eur. Phys. J. C (2021) 81:178]



[arXiv:2312.07562]

VH Production: Main uncertainties

V(lep/had)H($b\bar{b}$)

- Starting to be dominated by **systematic uncertainties in resolved regime**
- **Statistical uncertainties** still very relevant **in boosted regime**

V(lep)H($c\bar{c}$)

- **Statistical and systematic uncertainties** of the **same magnitude**

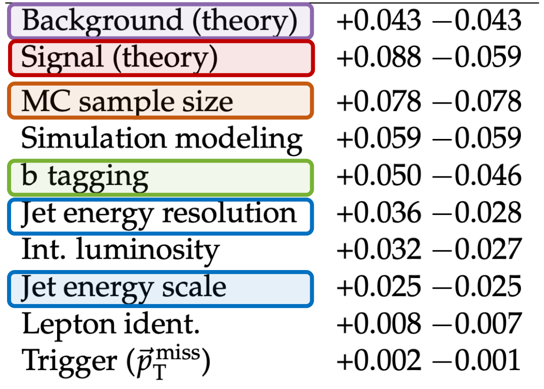
Leading systematic uncertainties

- Background modelling: ttbar, W/Z+jets, multijet
- Signal modelling
- Jet flavour tagging uncertainties
- Jet energy scale and resolution uncertainties
- Statistical uncertainty from limited size of MC samples

→ More data will improve results directly and indirectly (better modelling, reduced uncertainties)

arXiv:2312.07562

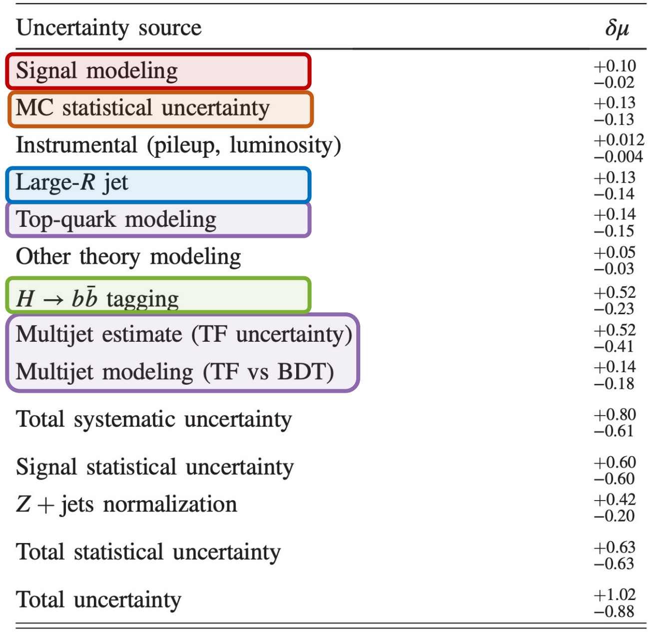
CMS V(lep)H(bb)



	$\Delta\mu$
Background (theory)	+0.043 −0.043
Signal (theory)	+0.088 −0.059
MC sample size	+0.078 −0.078
Simulation modeling	+0.059 −0.059
b tagging	+0.050 −0.046
Jet energy resolution	+0.036 −0.028
Int. luminosity	+0.032 −0.027
Jet energy scale	+0.025 −0.025
Lepton ident.	+0.008 −0.007
Trigger ($\vec{p}_{\mathrm{T}}^{\mathrm{miss}}$)	+0.002 −0.001

PhysRevLett.132.131802

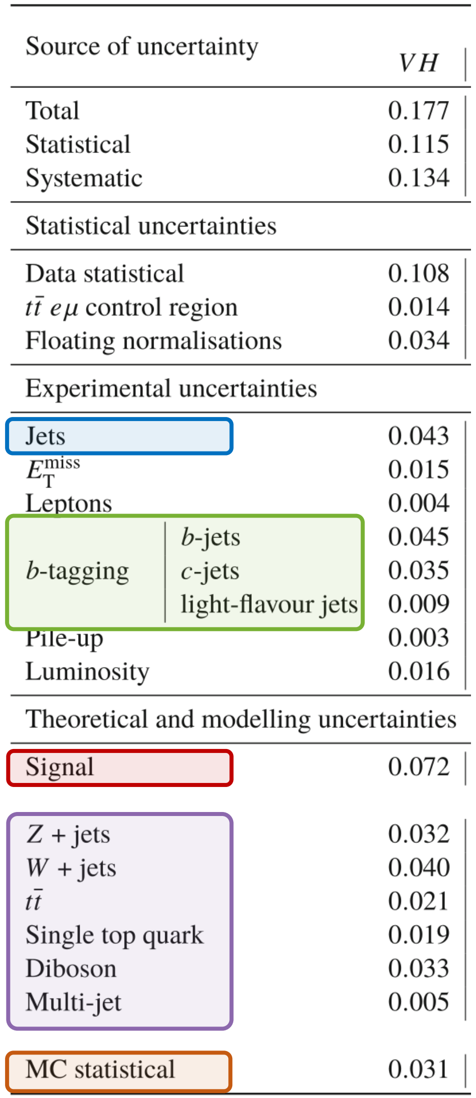
ATLAS boosted V(had)H(bb)



Uncertainty source	$\delta\mu$
Signal modeling	$^{+0.10}_{-0.02}$
MC statistical uncertainty	$^{+0.13}_{-0.13}$
Instrumental (pileup, luminosity)	$^{+0.012}_{-0.004}$
Large-R jet	$^{+0.13}_{-0.14}$
Top-quark modeling	$^{+0.14}_{-0.15}$
Other theory modeling	$^{+0.05}_{-0.03}$
$H \to b\bar{b}$ tagging	$^{+0.52}_{-0.23}$
Multijet estimate (TF uncertainty)	$^{+0.52}_{-0.41}$
Multijet modeling (TF vs BDT)	$^{+0.14}_{-0.18}$
Total systematic uncertainty	$^{+0.80}_{-0.61}$
Signal statistical uncertainty	$^{+0.60}_{-0.60}$
$Z +$ jets normalization	$^{+0.42}_{-0.20}$
Total statistical uncertainty	$^{+0.63}_{-0.63}$
Total uncertainty	$^{+1.02}_{-0.88}$

Eur. Phys. J. C (2021) 81:178

ATLAS resolved V(lep)H(bb)



Source of uncertainty		VH
Total		0.177
Statistical		0.115
Systematic		0.134
Statistical uncertainties		
Data statistical		0.108
$t\bar{t}$ $e\mu$ control region		0.014
Floating normalisations		0.034
Experimental uncertainties		
Jets		0.043
$E_{\mathrm{T}}^{\mathrm{miss}}$		0.015
Leptons		0.004
b-tagging	b-jets	0.045
	c-jets	0.035
	light-flavour jets	0.009
Pile-up		0.003
Luminosity		0.016
Theoretical and modelling uncertainties		
Signal		0.072
$Z +$ jets		0.032
$W +$ jets		0.040
$t\bar{t}$		0.021
Single top quark		0.019
Diboson		0.033
Multi-jet		0.005
MC statistical		0.031

H(b$\bar{\text{b}}$) Overview

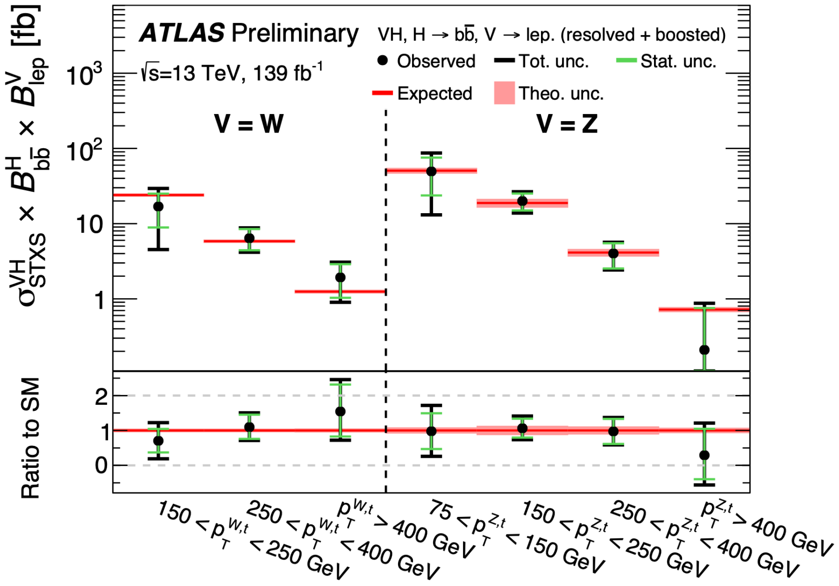
VH(b$\bar{\text{b}}$) analyses

Analysis	Fitted signal strength $\mu_{VH\ (H \to b\bar{b})}$
ATLAS resolved V(lep)H(b$\bar{\text{b}}$)	$1.02^{+0.18}_{-0.17} = 1.02^{+0.12}_{-0.11}\text{(stat.)}^{+0.14}_{-0.13}\text{(syst.)}$
ATLAS boosted V(lep)H(b$\bar{\text{b}}$)	$0.72^{+0.39}_{-0.36} = 0.72^{+0.29}_{-0.28}\text{(stat.)}^{+0.26}_{-0.22}\text{(syst.)}$
ATLAS resolved+boosted V(lep)H(b$\bar{\text{b}}$)	$1.00^{+0.18}_{-0.17} = 1.02^{+0.12}_{-0.11}\text{(stat.)}^{+0.14}_{-0.13}\text{(syst.)}$
CMS resolved+boosted V(lep)H(b$\bar{\text{b}}$)	$1.15^{+0.22}_{-0.20}$
ATLAS boosted V(had)H(b$\bar{\text{b}}$)	$1.4^{+1.0}_{-0.9} = 1.4^{+0.6}_{-0.6}\text{(stat.)}^{+0.8}_{-0.7}\text{(syst.)}$

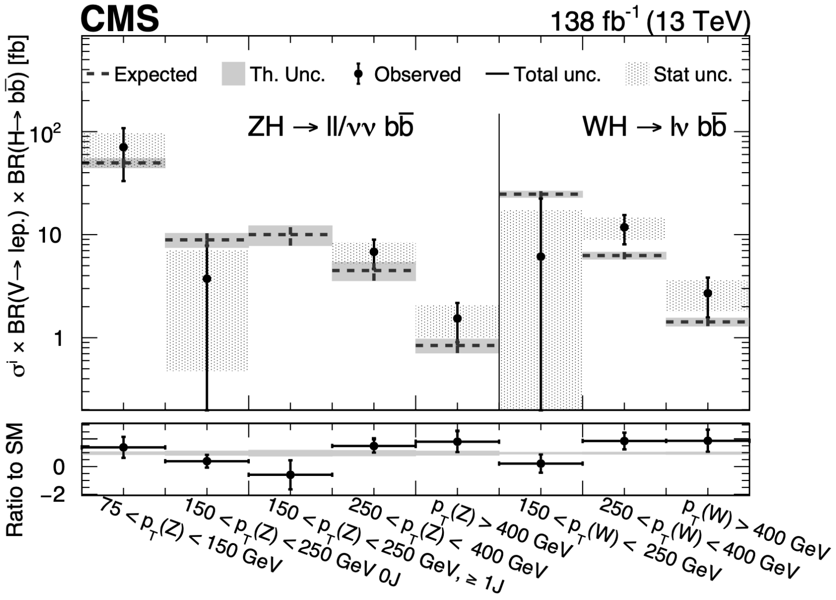
- Moving towards precision measurements
    - SXTS and differential cross-section measurements being undertaken

STXS Measurements

- Simplified Template Cross-Sections (STXS) involve categorisation of events in bins designed to separate different production modes



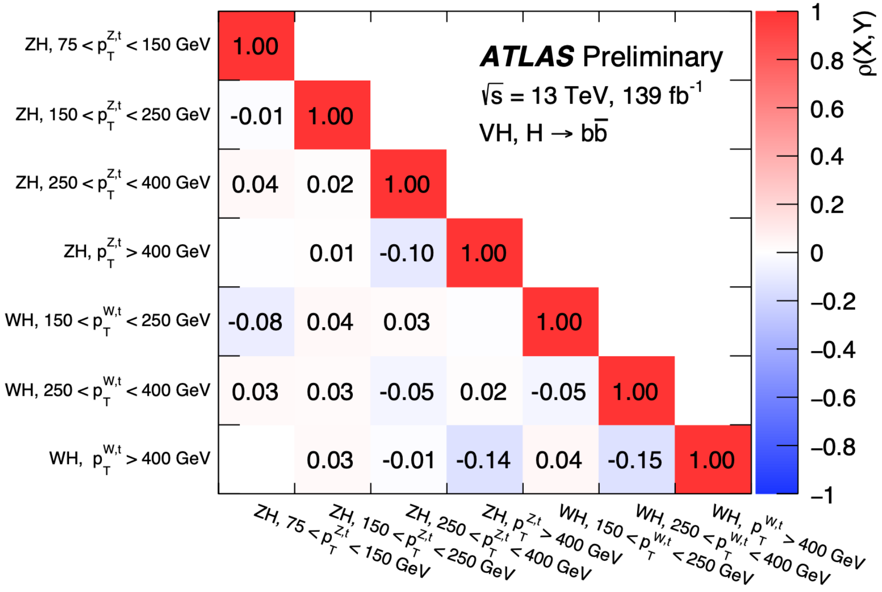
ATLAS-CONF-2021-051



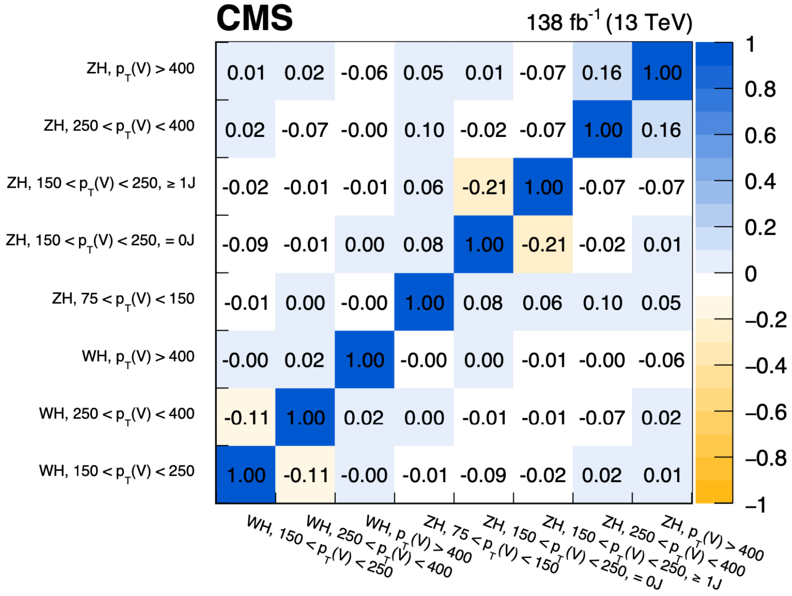
arXiv:2312.07562

STXS Measurements II

- Cross-contamination generally negligible between different STXS production bins
- Some WH contribution in ZH bins via 0 lepton channel



	ZH, $75 < p_T^{Z,t} < 150$ GeV	ZH, $150 < p_T^{Z,t} < 250$ GeV	ZH, $250 < p_T^{Z,t} < 400$ GeV	ZH, $p_T^{Z,t} > 400$ GeV	WH, $150 < p_T^{W,t} < 250$ GeV	WH, $250 < p_T^{W,t} < 400$ GeV	WH, $p_T^{W,t} > 400$ GeV
ZH, $75 < p_T^{Z,t} < 150$ GeV	1.00						
ZH, $150 < p_T^{Z,t} < 250$ GeV	-0.01	1.00					
ZH, $250 < p_T^{Z,t} < 400$ GeV	0.04	0.02	1.00				
ZH, $p_T^{Z,t} > 400$ GeV		0.01	-0.10	1.00			
WH, $150 < p_T^{W,t} < 250$ GeV	-0.08	0.04	0.03		1.00		
WH, $250 < p_T^{W,t} < 400$ GeV	0.03	0.03	-0.05	0.02	-0.05	1.00	
WH, $p_T^{W,t} > 400$ GeV		0.03	-0.01	-0.14	0.04	-0.15	1.00



	WH, $150 < p_T(V) < 250$	WH, $250 < p_T(V) < 400$	WH, $p_T(V) > 400$	ZH, $75 < p_T(V) < 150$	ZH, $150 < p_T(V) < 250$, = 0J	ZH, $150 < p_T(V) < 250$, ≥ 1J	ZH, $250 < p_T(V) < 400$	ZH, $p_T(V) > 400$
ZH, $p_T(V) > 400$	0.01	0.02	-0.06	0.05	0.01	-0.07	0.16	1.00
ZH, $250 < p_T(V) < 400$	0.02	-0.07	-0.00	0.10	-0.02	-0.07	1.00	0.16
ZH, $150 < p_T(V) < 250$, ≥ 1J	-0.02	-0.01	-0.01	0.06	-0.21	1.00	-0.07	-0.07
ZH, $150 < p_T(V) < 250$, = 0J	-0.09	-0.01	0.00	0.08	1.00	-0.21	-0.02	0.01
ZH, $75 < p_T(V) < 150$	-0.01	0.00	-0.00	1.00	0.08	0.06	0.10	0.05
WH, $p_T(V) > 400$	-0.00	0.02	1.00	-0.00	0.00	-0.01	-0.00	-0.06
WH, $250 < p_T(V) < 400$	-0.11	1.00	0.02	0.00	-0.01	-0.01	-0.07	0.02
WH, $150 < p_T(V) < 250$	1.00	-0.11	-0.00	-0.01	-0.09	-0.02	0.02	0.01

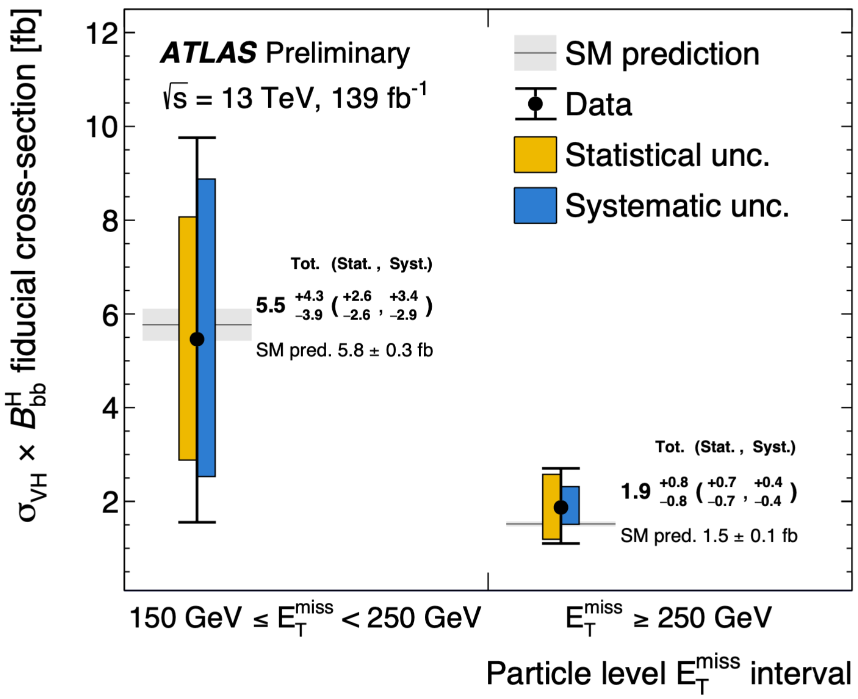
ATLAS-CONF-2021-051

arXiv:2312.07562

Differential Cross-section Measurements

- **Current statistics** and **understanding of systematic uncertainties** allow for differential cross-section measurements to be undertaken

- ATLAS performed such measurement in the 0 lepton channel (Z(*νν*)H($b\bar{b}$))
    - Remaining channels constrain the SM background via one-binned regions
- **Detector effects corrected** for via profile-likelihood unfolding procedure
- Impacted by **usual systematic uncertainties**, with some reduced contraints due to one-binned regions
- Results also affected by **migration of events from outside fiducial phase-space**



ATLAS-CONF-2022-015

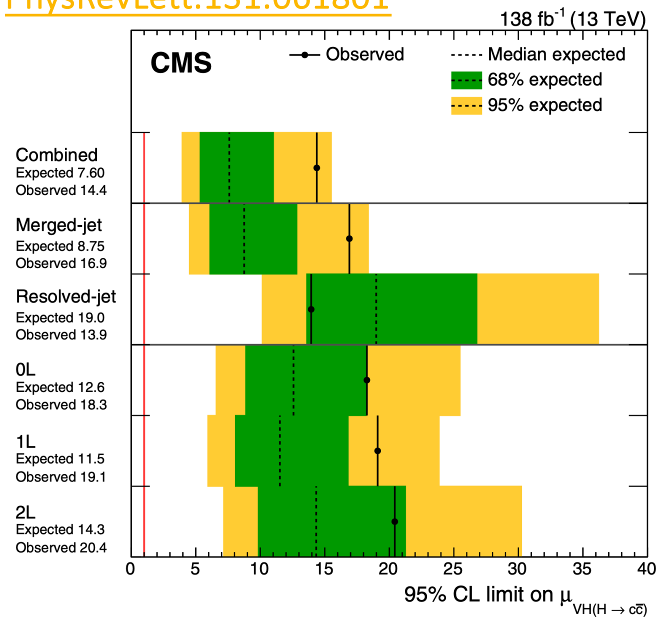
H($c\bar{c}$) Overview

H($c\bar{c}$) analyses

Analysis	Observed (Expected) 95% CL limit on $\mu_{VH\,(H\rightarrow c\bar{c})}$
ATLAS resolved V(lep)H($c\bar{c}$) H($b\bar{b}$)	$26.0 \times$ SM $(31^{+12}_{-8} \times$ SM$)$
CMS resolved+boosted V(lep)H($c\bar{c}$)	$14.4 \times$ SM $(7.6^{+3.4}_{-2.3} \times$ SM$)$
CMS boosted ggF H($c\bar{c}$)	$47.0 \times$ SM $(39 \times$ SM$)$

- Improving sensitivity
  - Use of MVA techniques in analysis design and jet flavour tagging methods bringing significant enhancements in sensitivity
  - Boosted regime can further exploit such tools to bring additional relevant contributions

PhysRevLett.131.061801



κ Framework

- Measure **deviations of Higgs couplings** from SM
- Inspired by **leading order diagrams** of Higgs couplings
- Assumes Higgs boson resonance at 125 GeV and narrow width approximation

$$\sigma \times \mathrm{BR}(i \to H \to f) = \frac{\sigma_i \times \Gamma_f}{\Gamma_H}$$

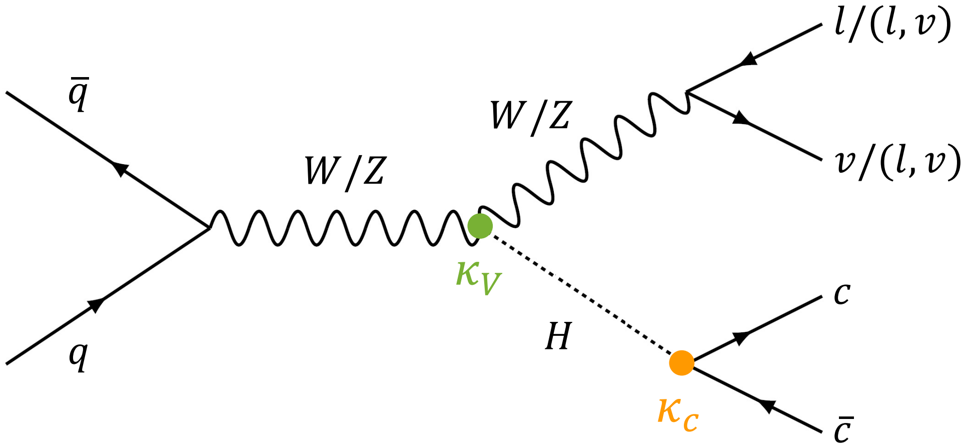
$$+$$

$$y_j \to \kappa_j y_j$$

$$\downarrow$$

$$\sigma \times \mathrm{BR}(i \to H \to f) = \frac{\sigma_i^{SM} \kappa_i^2 \times \Gamma_f^{SM} \kappa_f^2}{\Gamma_H^{SM} \kappa_H^2}$$

$$\downarrow$$

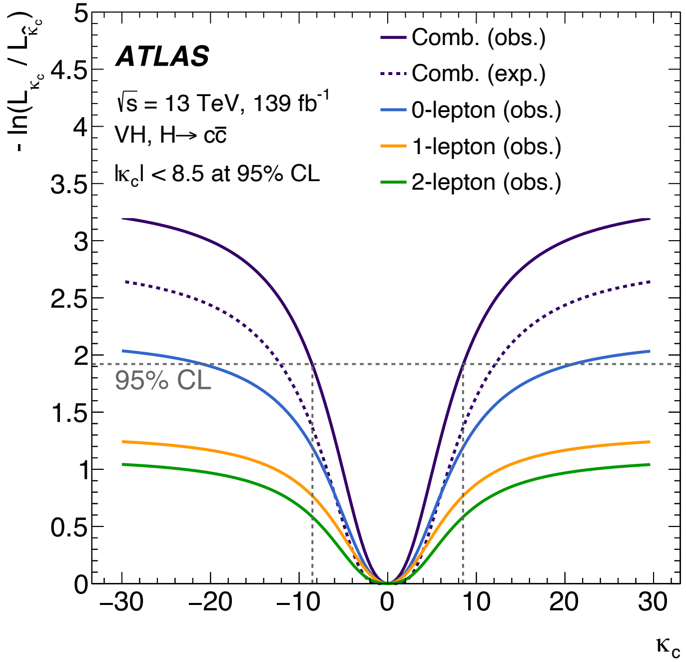


$$\sigma \times \mathrm{BR}(\mathrm{V} \to \mathrm{H} \to c\bar{c}) = \sigma_{SM} \times \mathrm{BR}(\mathrm{V} \to \mathrm{H} \to c\bar{c})_{SM} \times \frac{\kappa_V^2 \kappa_c^2}{\kappa_H^2}\ ,\ \kappa_H^2 = \frac{\Sigma_j \Gamma_j^{SM} \times \kappa_j^2}{\Gamma_H^{SM}}$$

κ_c Interpretation

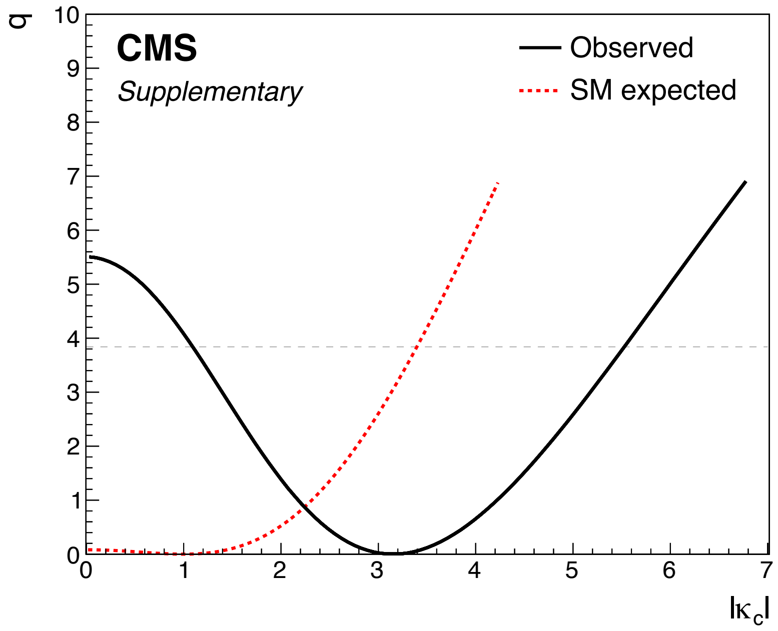
- V(lep)H(c$\bar{c}$) analyses sensitive to exclude some parts of κ_c parameter space (parametrisation below used)
- $\mu = \frac{\kappa_c^2}{[(1-BR_{Hcc})+BR_{Hcc}*\kappa_c^2]}$ (other coupling modifiers set to 1, no BSM contributions to Higgs width)

Eur. Phys. J. C (2022) 82:717



$|\kappa_c| \leq 8.5$ at 95% CL

PhysRevLett.131.061801



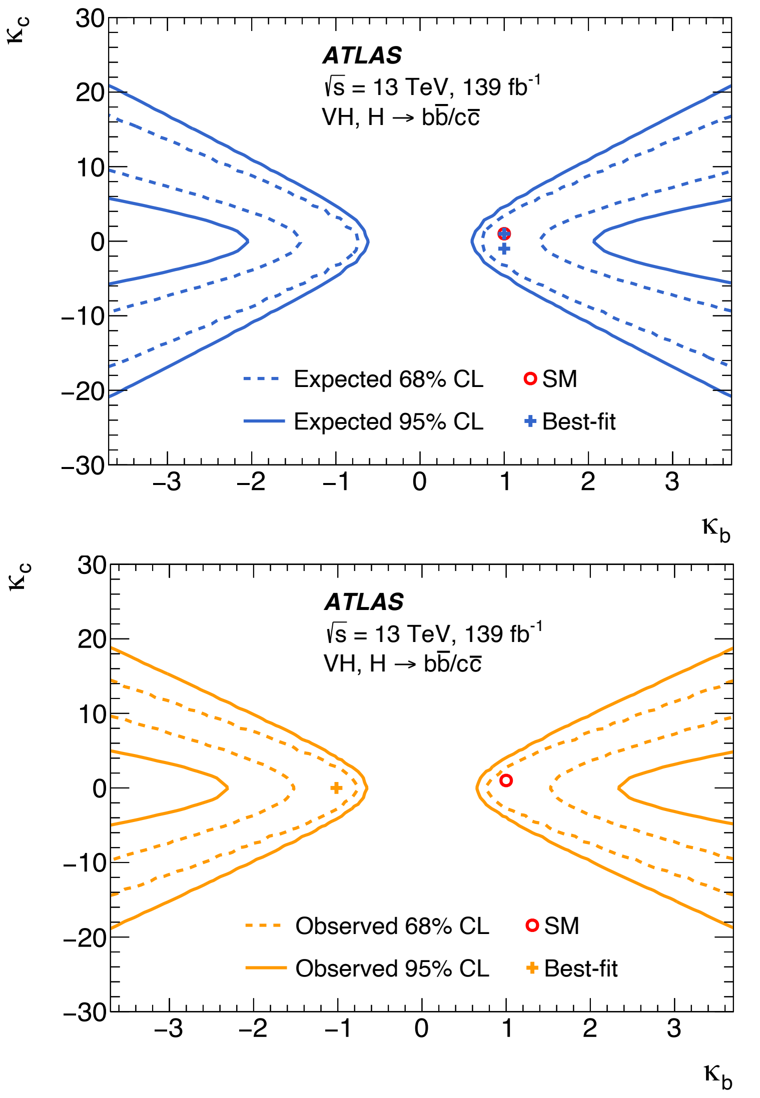
$|\kappa_c| \leq 3.4$ ($1.1 < \kappa_c < 5.5$) at 95% CL

κ_c and κ_b Interpretation

Eur. Phys. J. C (2022) 82:717

- **V(lep)H(b$\bar{b}$)** and **V(lep)H(c$\bar{c}$) combined** likelihood parametrised in terms of both **κ_c and κ_b**
- Other coupling modifiers set to their SM predictions
- Gluon-initiated ZH production has no explicit parametrisation as function of κ_c
    - Only taken into account for κ_b
    - Only has an impact for large values of κ_b ($\kappa_b \gtrsim 10$)
- Non Drell-Yan-like qq → VH production also neglected given the low values of κ_c that the combination probes

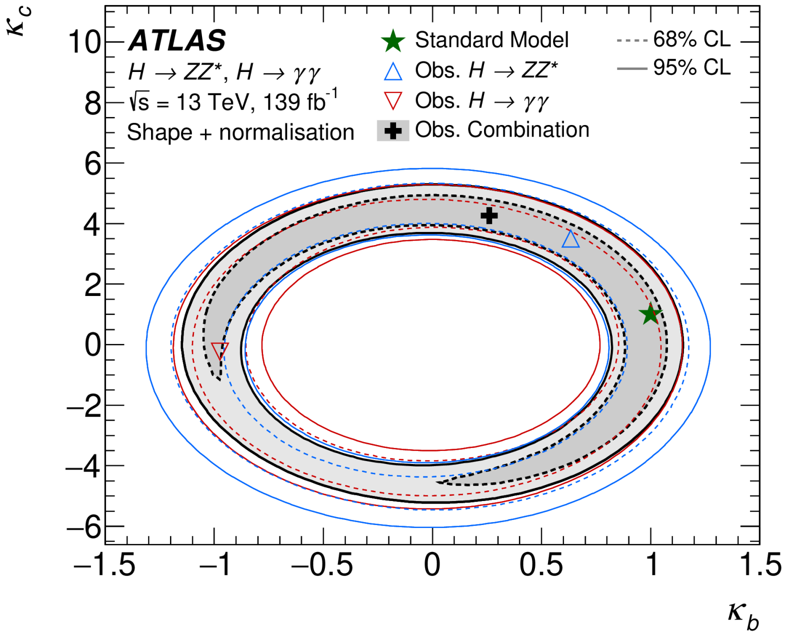
- Observed best fit value $(\kappa_b, \kappa_c) = (-1.02, 0)$
- Difference in value of the log-likelihood between best-fit value and $(\kappa_b, \kappa_c) = (+1.02, 0)$ is 0.02
    - Analysis has little power to constrain the sign of κ_b
    - Small likelihood asymmetry coming from b-quark loop contributions to gg → ZH production

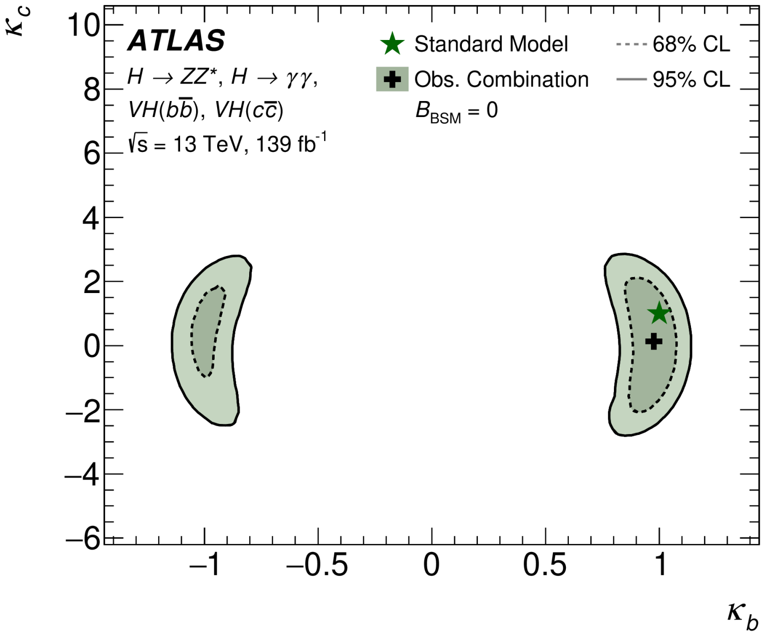


κ_c and κ_b Interpretation II

JHEP05(2023)028

- Precision measurements of **p_T^H cross-section in H → ZZ* and H → $\gamma\gamma$** channels provide **complimentary** information on couplings
- Individual analyses **dominated by statistical uncertainty**
- **Statistical combination between them** and in addition with **V(lep)H($b\bar{b}$)** and **V(lep)H($c\bar{c}$)** analyses allows for improved constraints
- Results for **shape and normalisation effects considered** shown here, different assumptions on backup

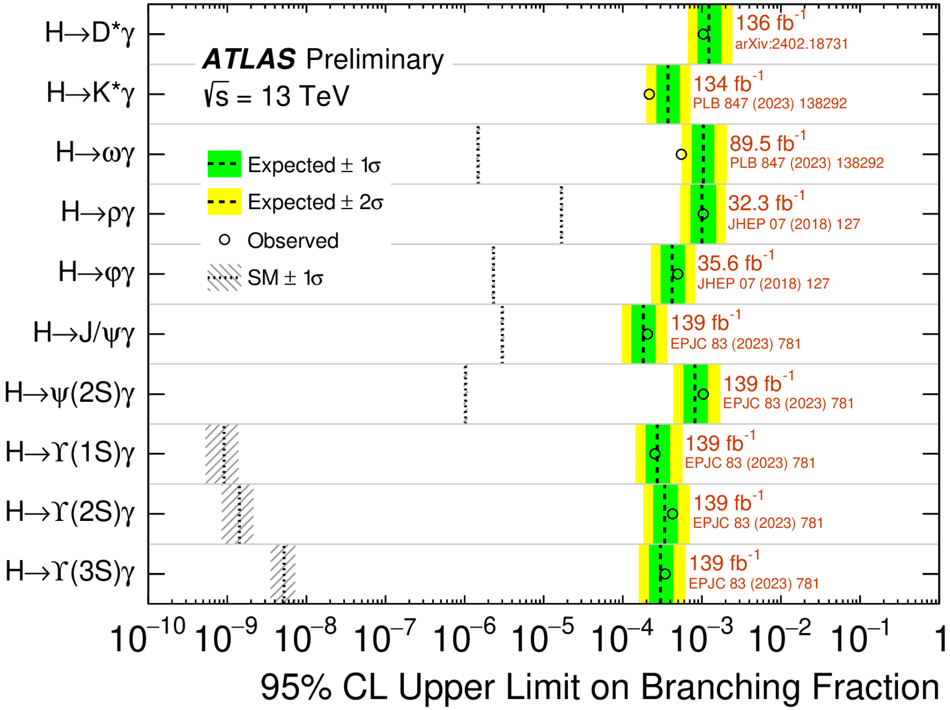




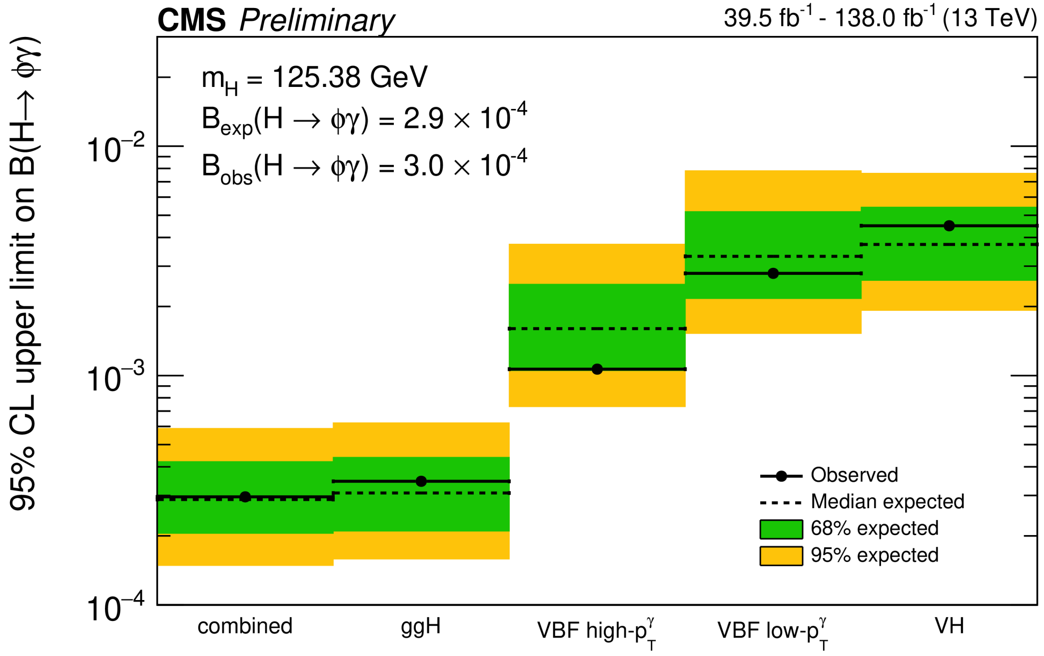
Probing lighter quark Yukawas

$q\bar{q} \rightarrow H(\rightarrow WW) + \gamma$ differential XS measurement as alternative probe: PhysRevLett.132.121901, in backup

- Rare **exclusive decays** of the Higgs boson into a meson and a photon allow **access to coupling to lighter flavour quarks** (up, down, strange) in addition to bottom and charm quarks
- Also used as probe of potential **flavour-violating Higgs boson interactions**
- **Dedicated triggers**, different **production modes**, **event selection with BDTs** and **background modelling enhancements improving sensitivity**



ATL-PHYS-PUB-2023-004



CMS-PAS-HIG-23-005, more CMS results on rare decays in backup

Summary

- Different probes sensitive to Higgs quark Yukawa couplings are available
  - Leading sensitivity to **heavy quark Yukawa** couplings currently via **Higgs production in association with W/Z bosons**
  - **Loop induced ggF and quark initiated Higgs production** and **exclusive Higgs boson decays into meson+photon** provide **complimentary information** on these couplings and give insight into **lighter quark couplings**

- Machine-learning improvements in analysis design and jet tagging techniques leading to continuous enhancements in sensitivity
  - Exploited both in **resolved** and **merged** jet topologies
  - Possible **tagging of Higgs candidates and/or W/Z bosons**
  - MVA discrimination of **Higgs signal** with respect to **backgrounds**

- Main uncertainties related to modelling, flavour tagging and jet contributions or statistics

- Different suite of measurements currently possible, from STXS bins to coupling modifiers

Backup

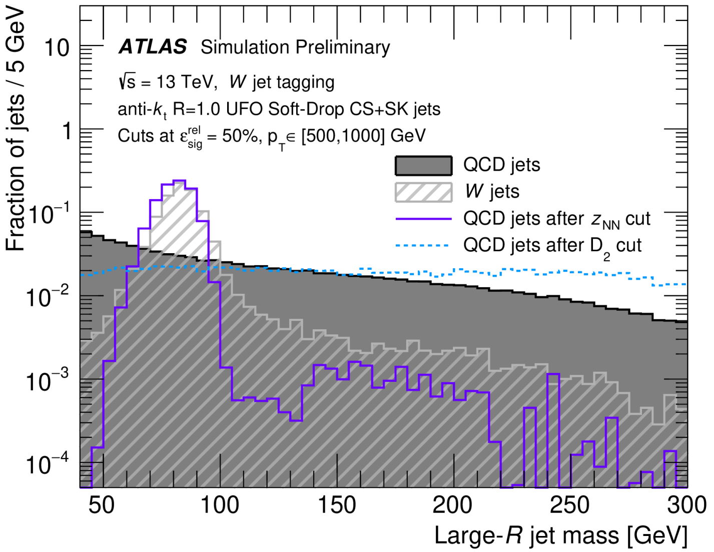
Vector boson DNN tagger

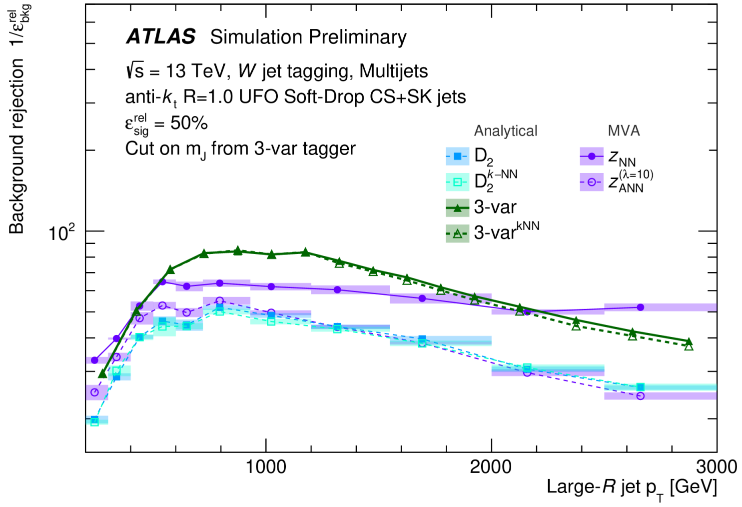
ATL-PHYS-PUB-2021-029

- DNN trained on high-level jet substructure quantities

Variable	Description
D_2, C_2	Energy correlation ratios
τ_{21}	N-subjettiness
R_2^{FW}	Fox-Wolfram moment
$\mathcal{P}$	Planar flow
a_3	Angularity
A	Aplanarity
Z_{cut}, $\sqrt{d_{12}}$	Splitting scales
$Kt\Delta R$	k_t-subjet ΔR

- Algorithm **decorrelated from jet mass** to prevent **bias in background jet-mass** distribution from being around the **W boson mass**
- Achieved via **sequential training of jet classification DNN and jet mass inferring adversarial NN** (ANN)
  - Classification DNN first optimised
  - DNN fixed, ANN optimised
  - Combined optimisation of both algorithms

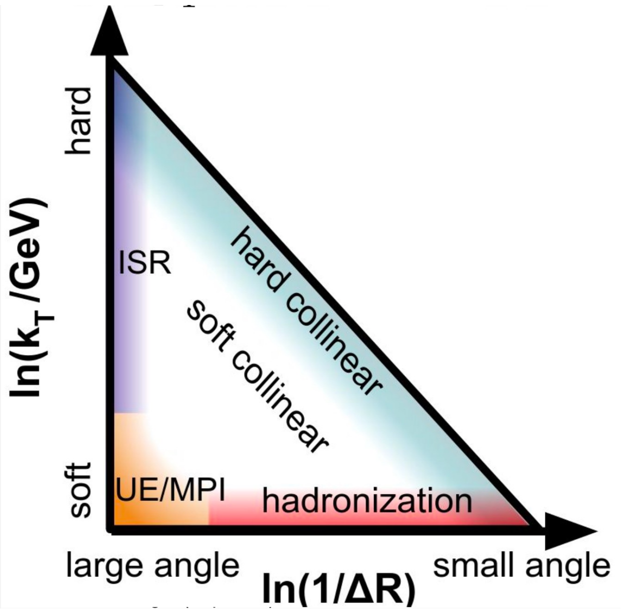


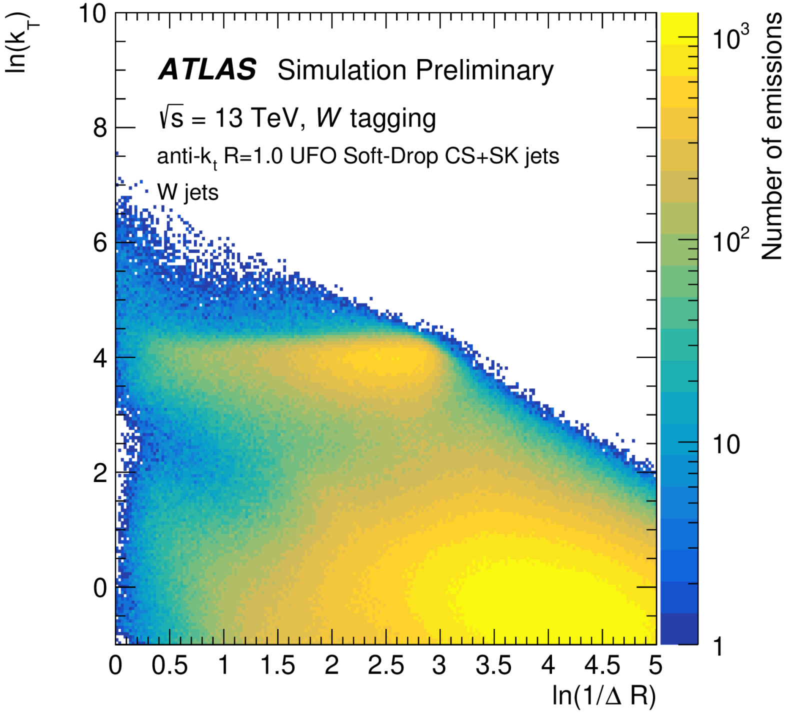


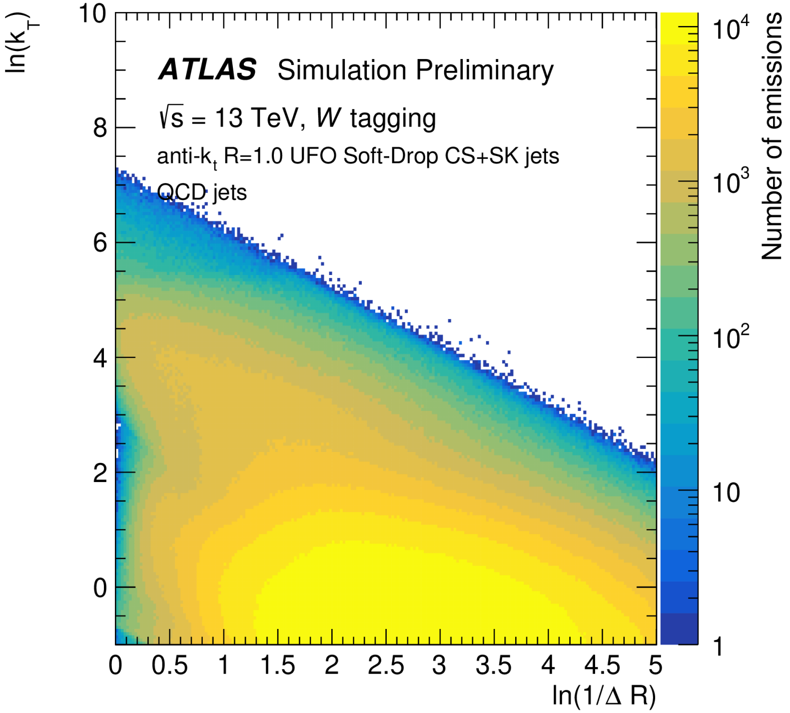
Lund jet planes

$$\Delta R_{ij} = \sqrt{\Delta y_{ij}^2 + \Delta \phi_{ij}^2}, \qquad z = \frac{p_T^j}{p_T^i + p_T^j}, \qquad k_t = p_T^j \Delta R_{ij}$$

- **Lund jet plane** represents **two-dimensional** space built from the **opening angle** and the **momentum fraction** of a given **gluon emission with respect to its emitter**







Jet substructure at high-pT

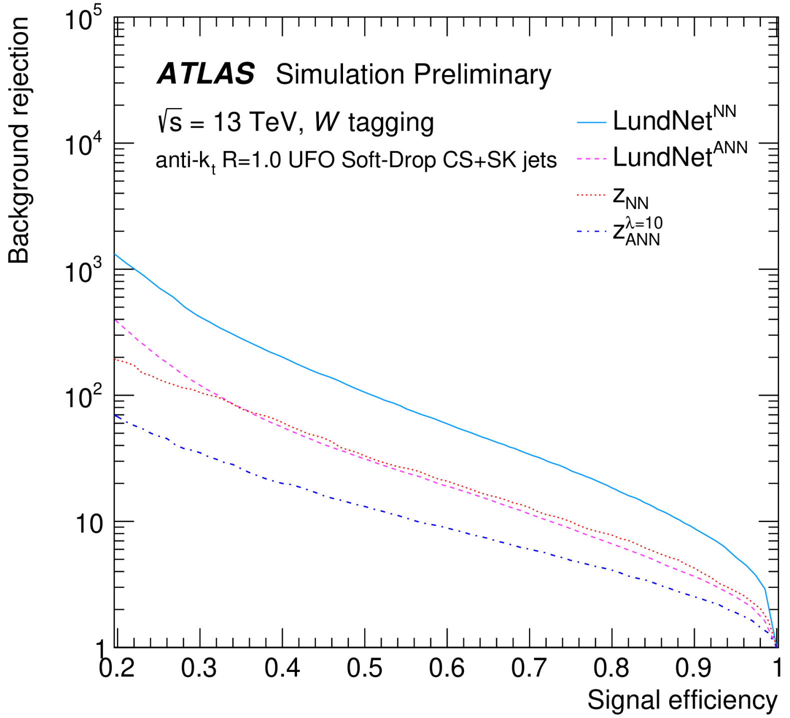
ATL-PHYS-PUB-2023-017

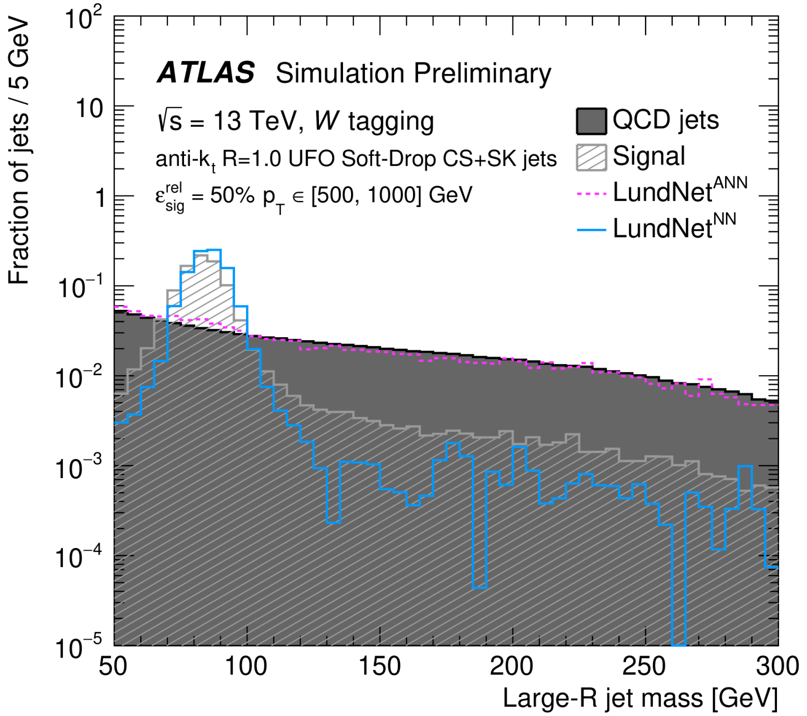
LundNet Performance

ATL-PHYS-PUB-2023-017

$$\Delta R_{ij} = \sqrt{\Delta y_{ij}^2 + \Delta\phi_{ij}^2}, \qquad z = \frac{p_{\mathrm{T}}^j}{p_{\mathrm{T}}^i + p_{\mathrm{T}}^j}, \qquad k_t = p_{\mathrm{T}}^j \Delta R_{ij}$$

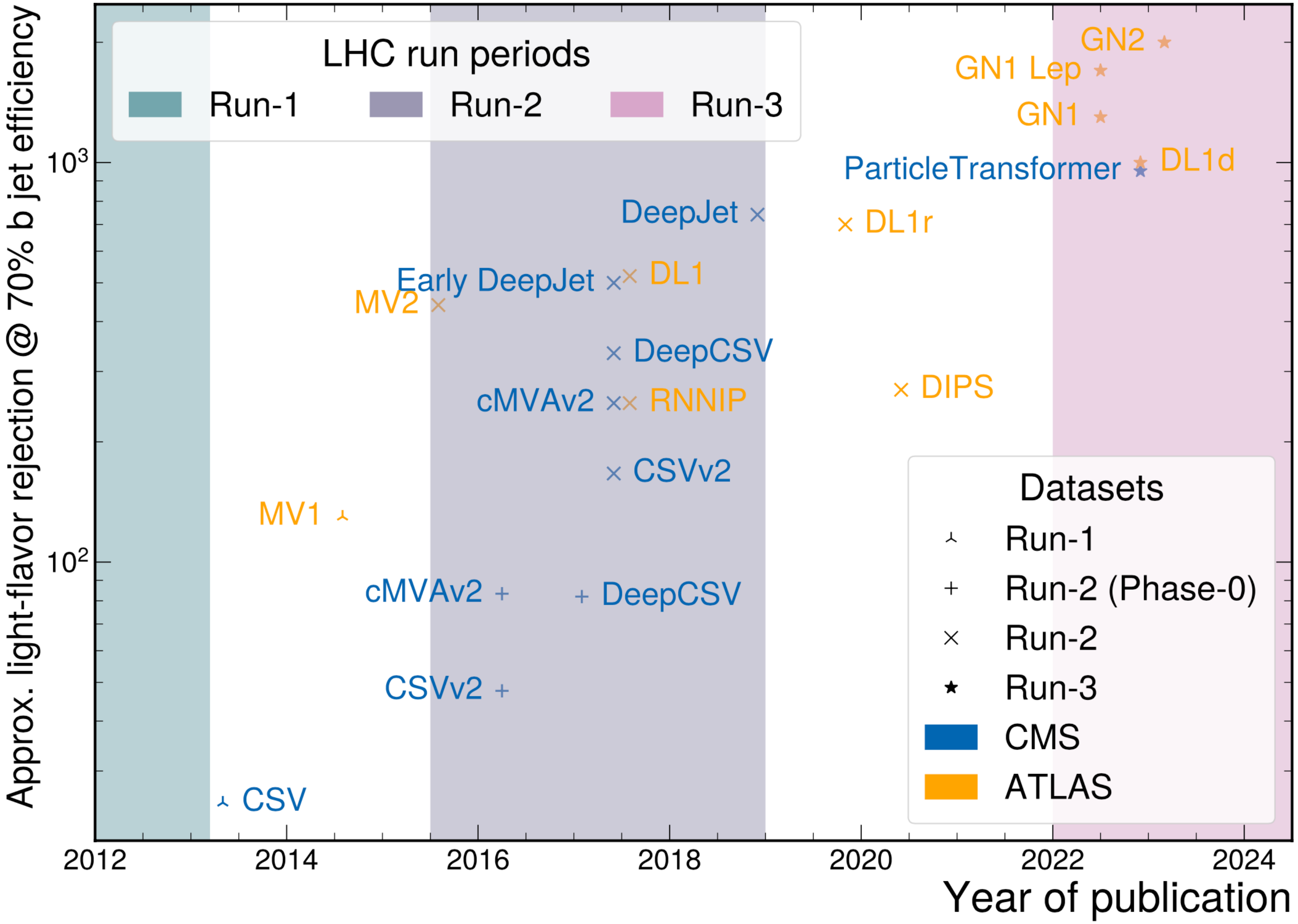
- **Sequence of jet emissions** coming from the full declustering of large radius jet can be **represented as a graph**
- **GNN inputs** include **ln(1/ΔR), ln(k), ln(1/*z*)** and **number of tracks**
- Algorithm **decorrelated from jet mass** to prevent **bias in background jet-mass** distribution from being around the **W boson mass**
- Achieved via **combined training of jet classification GNN and jet mass inferring adversarial NN** (ANN)





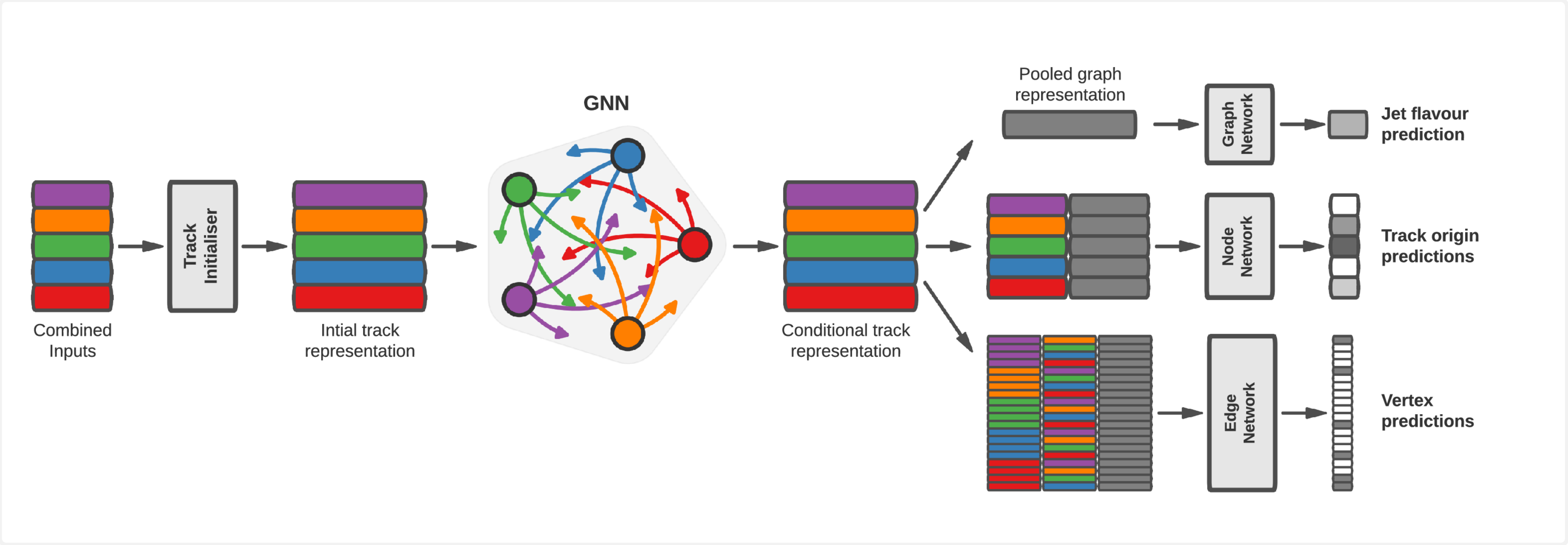
Comparison of different jet taggers

arXiv:2404.01071



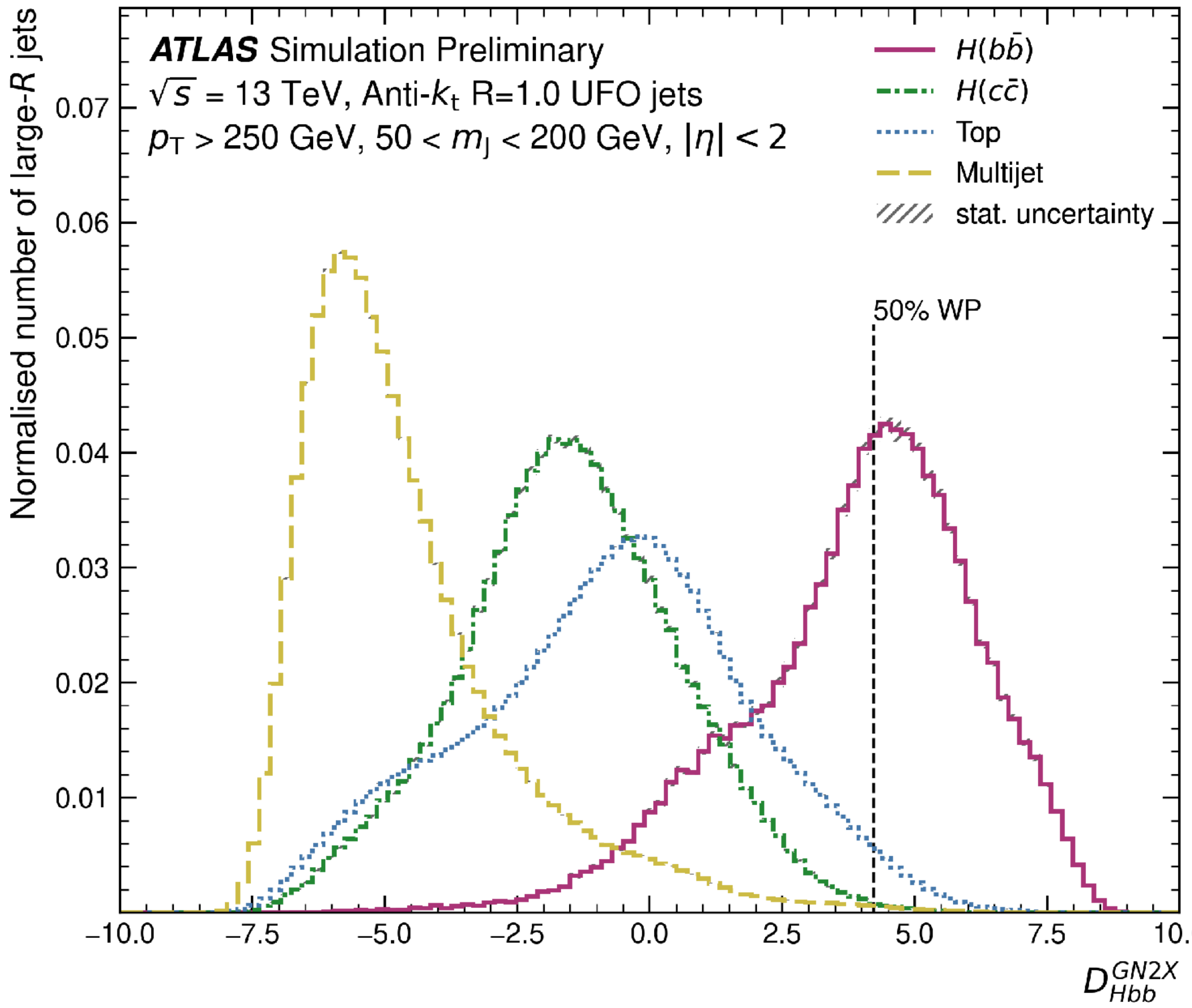
GN1 Architecture

ATL-PHYS-PUB-2022-027



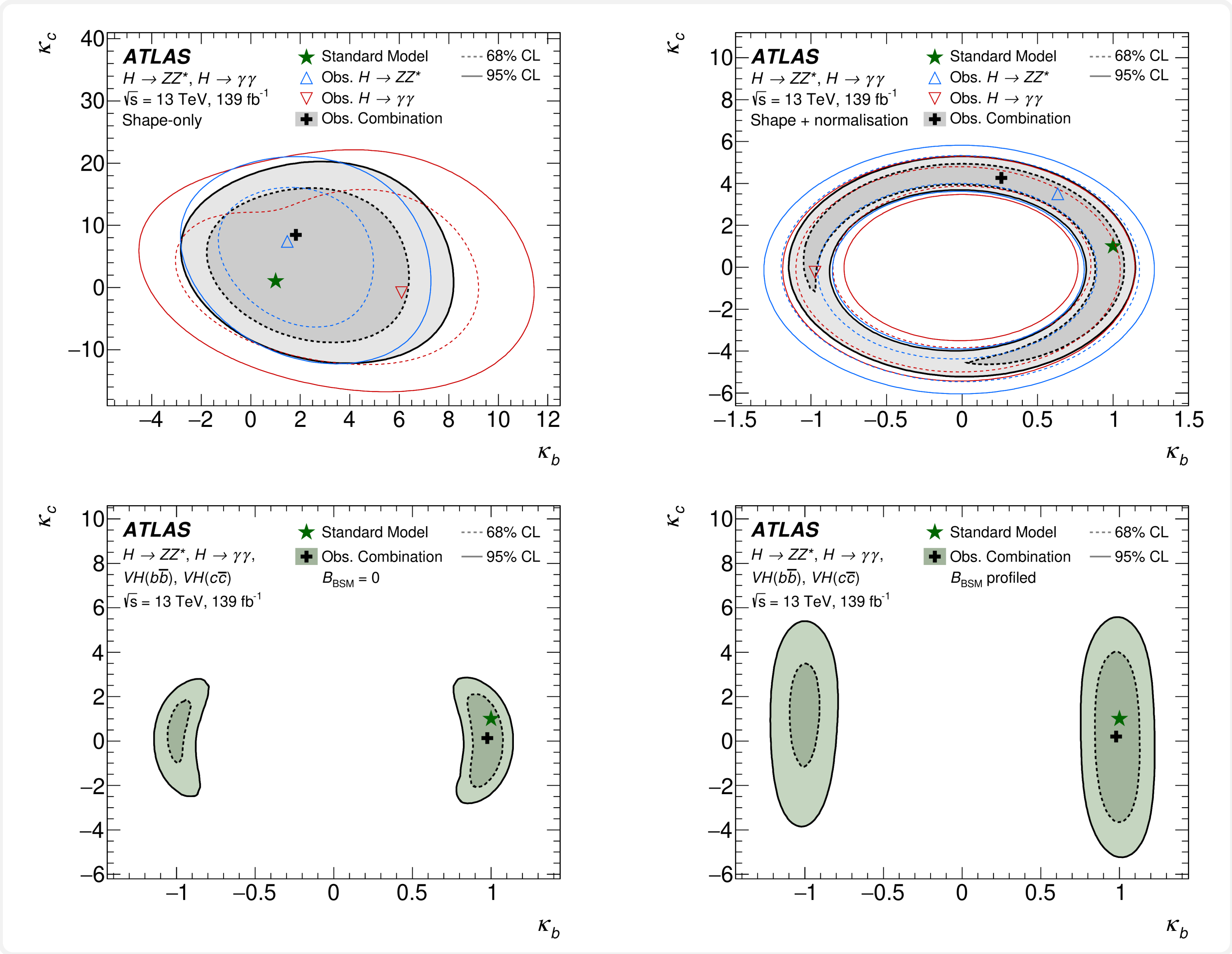
GN2X Performance

ATL-PHYS-PUB-2023-021

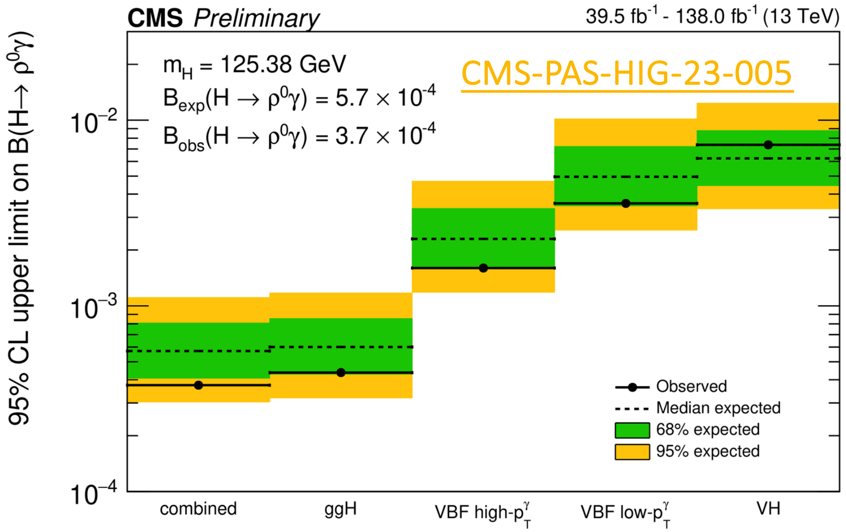


κ_c and κ_b Interpretation

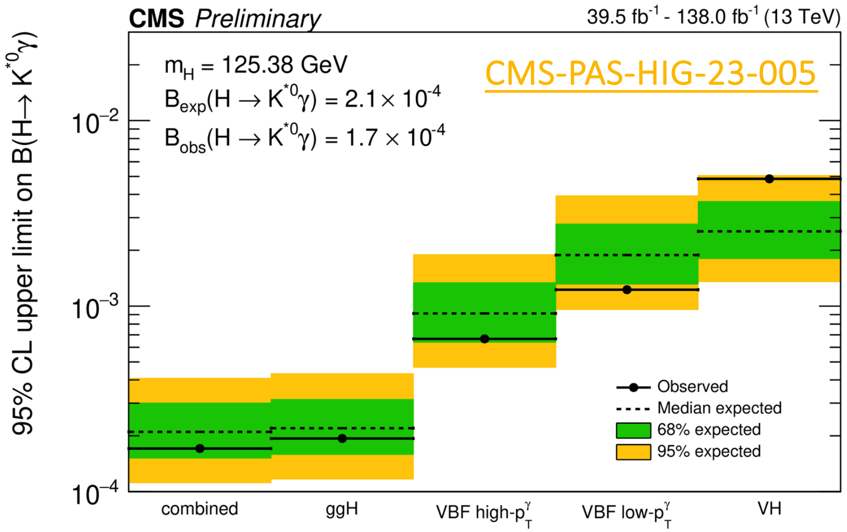
JHEP05(2023)028



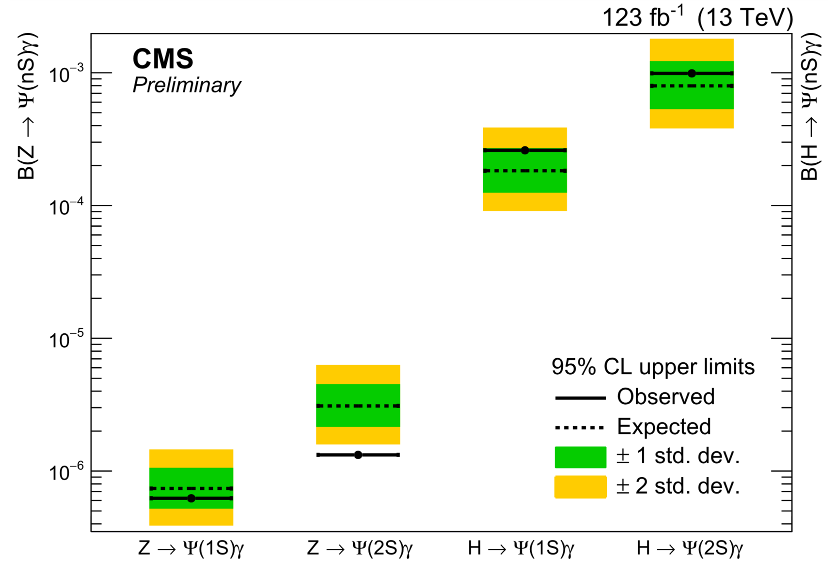
CMS Higgs to meson + photon searches



CMS-PAS-HIG-23-005



CMS-PAS-HIG-23-005

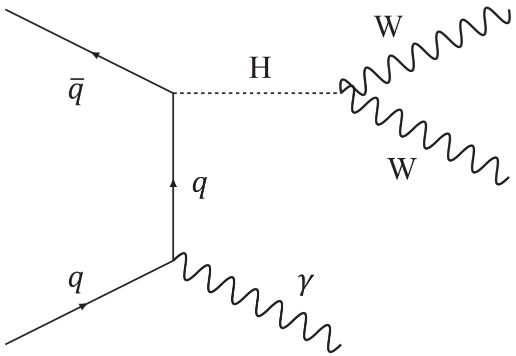


CMS-PAS-SMP-22-012

CMS Higgs + photon XS measurement

PhysRevLett.132.121901

- **Gluon initiated contribution disappears** according to Furry's theorem
- Sensitive to Higgs quark Yukawas
- Cut-based analysis, **statistical and systematic uncertainties** with similar magnitude
- Other coupling modifiers set to their SM predictions

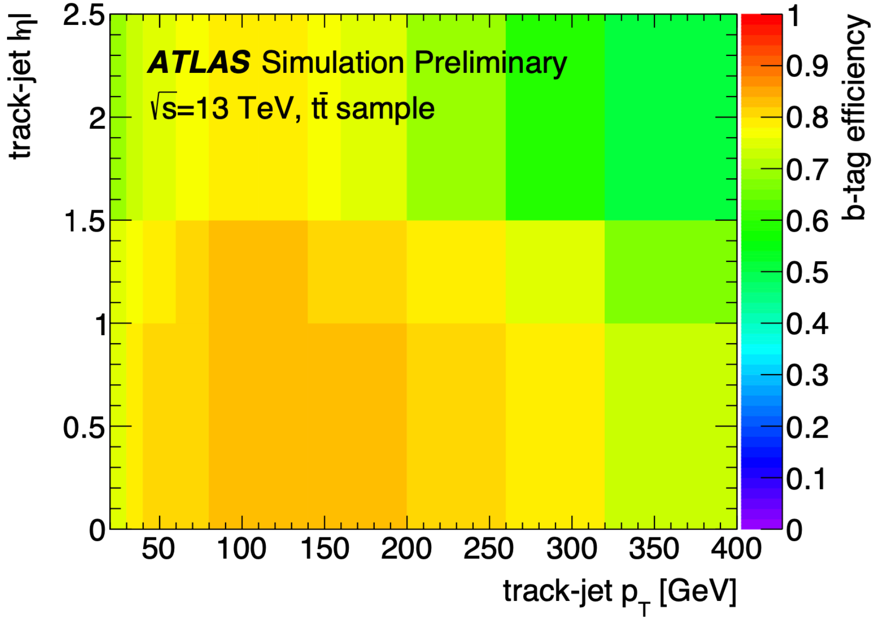


Process	σ upper limits obs. (exp.) [fb]	κ_q limits obs. (exp.) at 95% C.L.
$u\bar{u} \to H + \gamma \to e\mu\nu_e\nu_\mu\gamma$	86 (67)	$\lvert\kappa_u\rvert \leq 16000$ (13000)
$d\bar{d} \to H + \gamma \to e\mu\nu_e\nu_\mu\gamma$	72 (58)	$\lvert\kappa_d\rvert \leq 17000$ (14000)
$s\bar{s} \to H + \gamma \to e\mu\nu_e\nu_\mu\gamma$	66 (49)	$\lvert\kappa_s\rvert \leq 1700$ (1300)
$c\bar{c} \to H + \gamma \to e\mu\nu_e\nu_\mu\gamma$	88 (66)	$\lvert\kappa_c\rvert \leq 190$ (110)

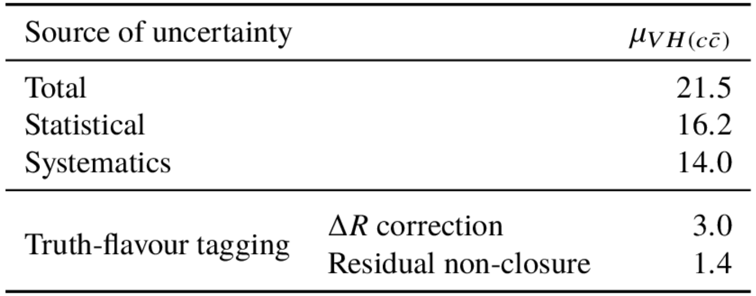
Truth flavour jet tagging

- Accepting/rejecting (“**direct tagging**”) events based on jet flavour tagging discriminants significantly **reduces available background** (e.g., V+light flavour jets processes) **statistics** due to **enhanced rejection** factors in latest algorithms
- ATLAS has implemented a **“truth-flavour tagging”** approach in recent **V(lep)H($b\bar{b}$)** and **V(lep)H($c\bar{c}$) analyses**
    - Consists in applying **tagging (in-)efficiency** of the Higgs candidate jets **as event weights**
    - **Efficiency maps** usually parametrised as **function of jet p_T and η**
- Leads to **reduction of background statistical uncertainties**, since all available events are used
- Some **systematic uncertainties can arise** from differences to direct tagging approach (e.g. due to close-by jets) – **still resulting in better sensitivity**

PhysRevLett.132.121901



Eur. Phys. J. C (2022) 82:717

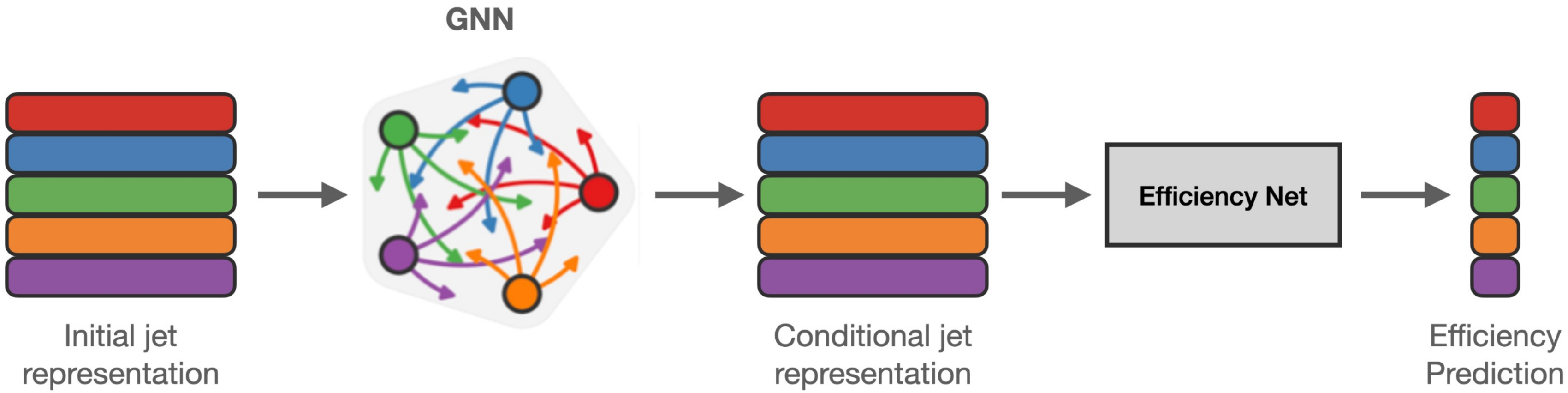


Source of uncertainty		$\mu_{VH(c\bar{c})}$
Total		21.5
Statistical		16.2
Systematics		14.0
Truth-flavour tagging	ΔR correction	3.0
	Residual non-closure	1.4

Truth flavour jet tagging with GNNs

PhysRevLett.132.121901

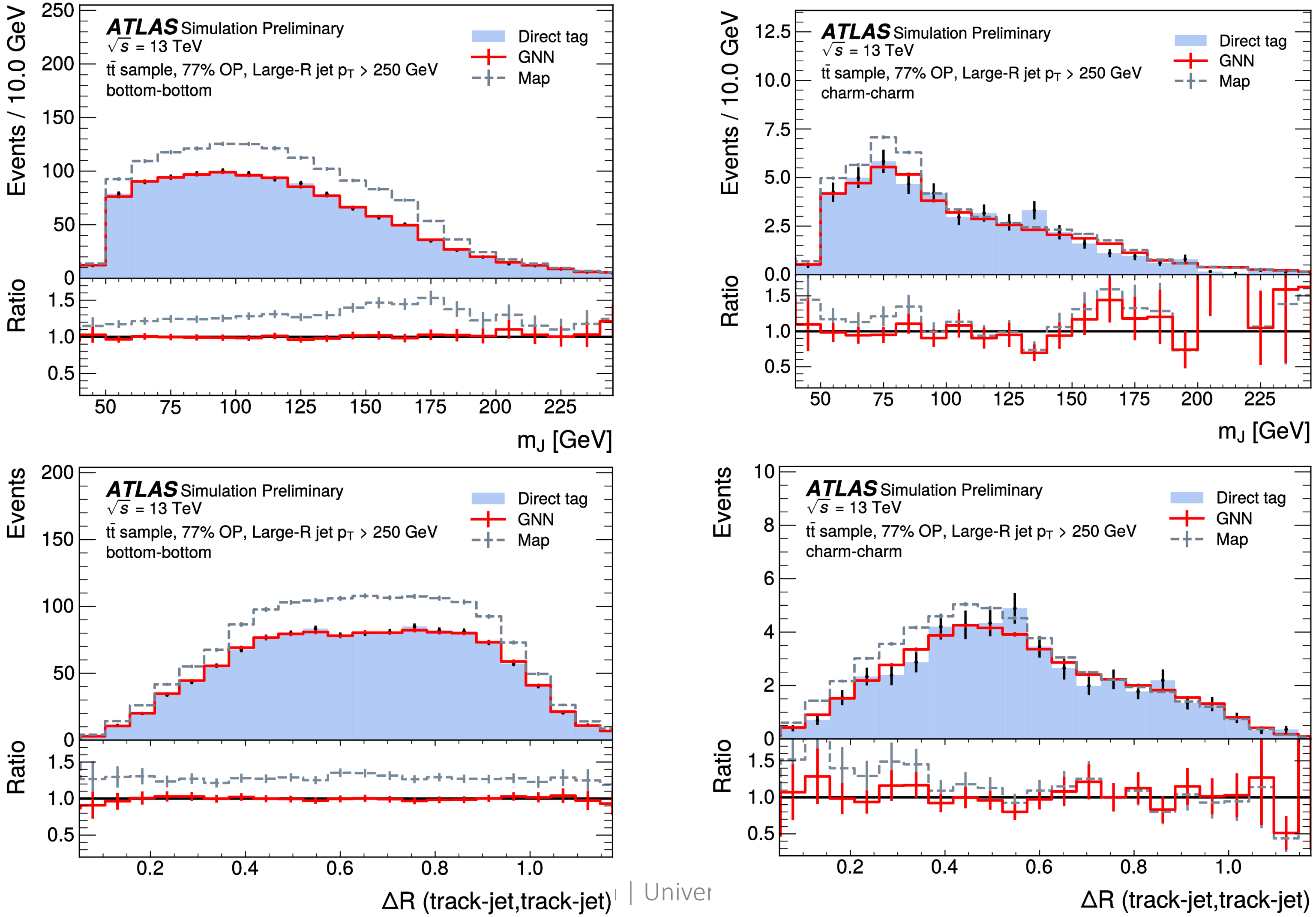
- **Efficiency maps fail to capture all kinematic correlations**, and can be **affected by binning choices**
- A **GNN** can potentially solve these problems, capturing **high-dimensional correlations between the jets** and leading to **smooth output efficiency distributions**
- **GNN inputs** include jet, event and jet-pair variables, trained with ttbar events and large radius jets



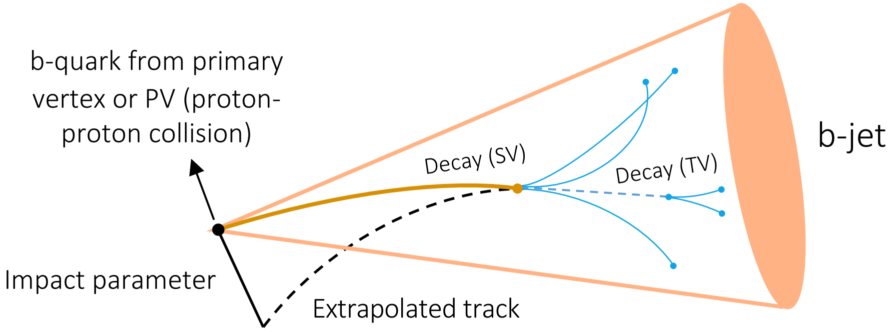
Truth flavour jet tagging with GNNs II

PhysRevLett.132.121901

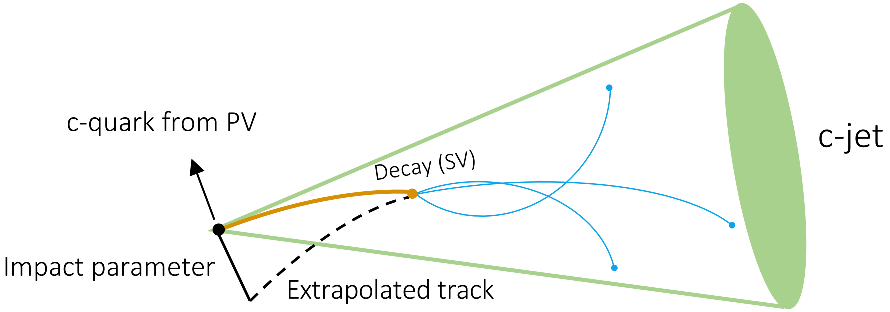
- **GNN-based approach** generally with **similar or better agreement with direct tagging** approach when **compared to efficiency maps** approach, particularly for very **close-by jets**



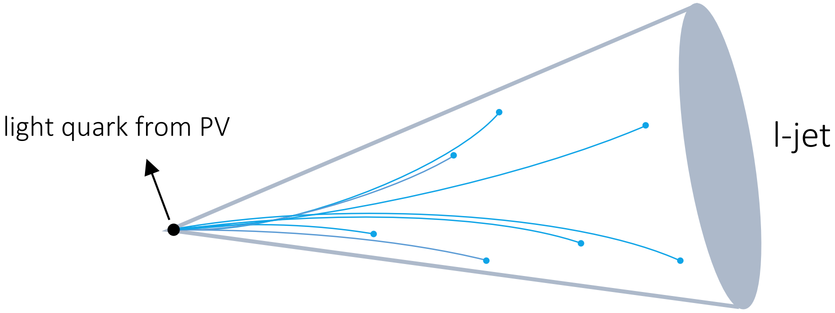
Flavoured Jets



- b-quark fragments into b-hadron carrying around **80% of the jet energy**
- **High** b-hadron **decay product multiplicities** (around 5 charged particles per decay)
- Most b-hadrons (≈ 90%) **decay into c-hadrons**
- b(c)-hadron **decay vertex often displaced** from the PV(SV) by a few mm
- Tracks from both of these vertices often have **large impact parameters**



- c-quark fragments into c-hadron carrying around **55% of the jet energy**
- 2 to 3 times **lower** c-hadron **decay product multiplicities** than for b-hadrons (around 2 charged particles/decay)
- c-hadron **decay vertex often displaced** from the PV by a few mm
- Tracks from this vertex can often have **large impact parameters**



- Light quark hadronises into **many light hadrons sharing the jet energy**
- Tracks from this vertex can often have **impact parameters consistent with zero**
- Long-lived light hadrons can be produced, but more likely to decay cms away from PV