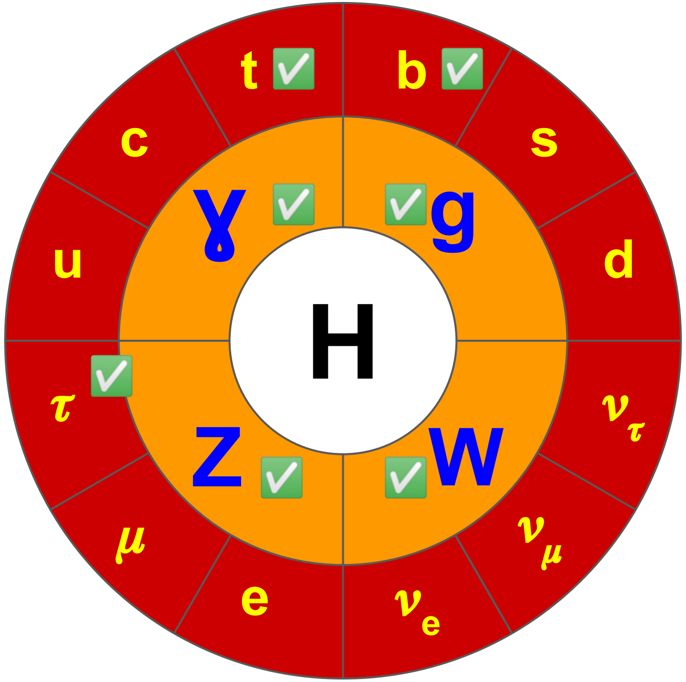
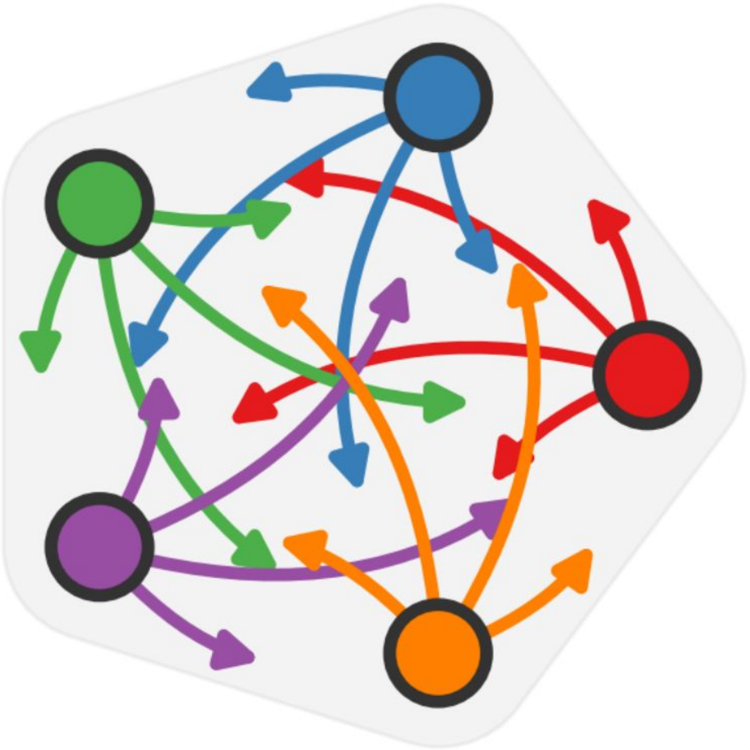
Measurements of Higgs bosons decaying into bb or cc pairs, and how to improve them with GNN-based flavour tagging techniques

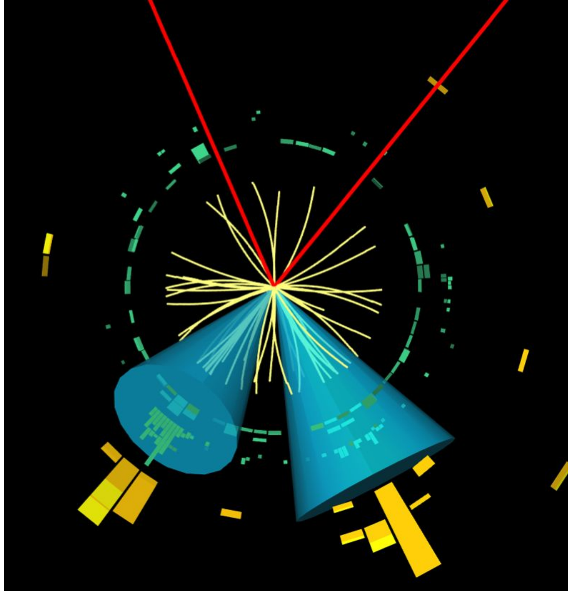
Martino Tanasini, on behalf of the ATLAS Collaboration

Rencontres de Moriond EW 2023
Young Scientists Forum


H
γ ✅
✅g
Z ✅
✅W
t ✅
b ✅
c
s
u
d
τ ✅
ν_τ
μ
ν_μ
e
ν_e



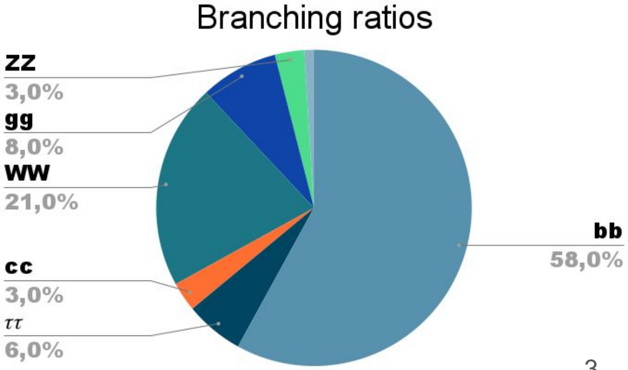
Search for VH(H→cc) decays



Motivation:
→ Higgs couplings to 3rd generation fermions (ttH, H→$\tau\tau$, H→bb) observed.
→ Probing couplings to lighter 2nd generation fermions could **open windows to new physics**.
→ H→cc: one of the **most common** yet unobserved decay modes.

Goal:
→ $\mu_{VH,Hcc}$ extracted with a fit on the invariant mass of the *c*-jet pair.

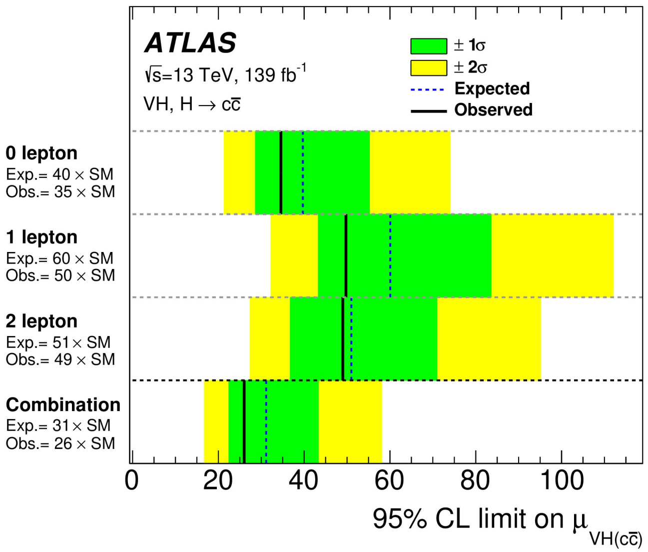
Focus of today:
→ **Deep learning based dedicated *c*-tagger** with 27% c-jet efficiency, 8.3% b-jet efficiency, and 1.7% light jet efficiency (on a ttbar sample).

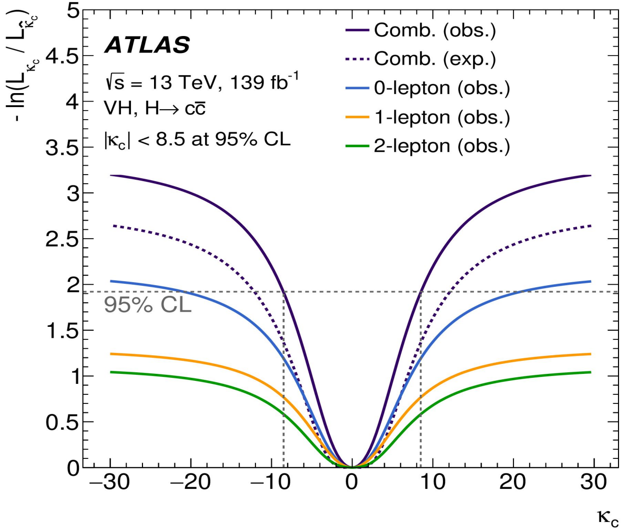


10.1140/epjc/s10052-022-10588-3

Results

10.1140/epjc/s10052-022-10588-3

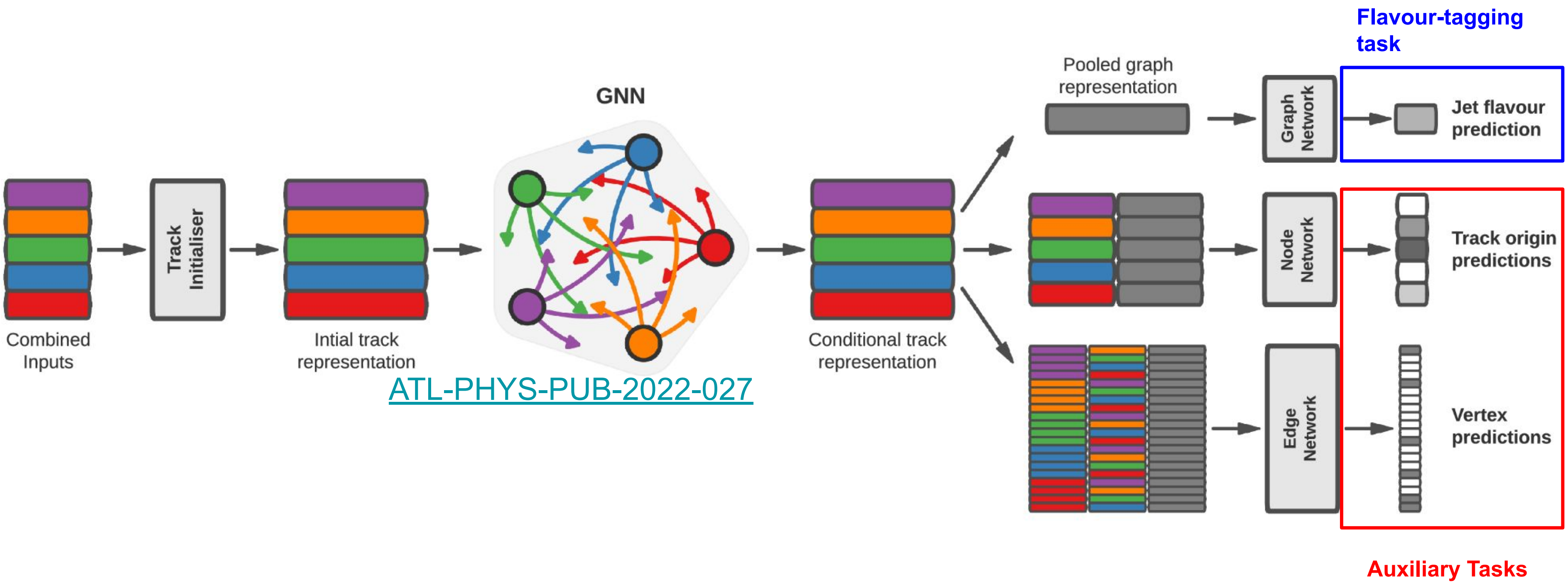




→ Upper limit is set on VH,H→cc process of **26 times the SM prediction**.

→ Results interpreted in the kappa framework constraining the coupling modifier **$|k_c|$<8.5 @ 95%CL** (fixing the other couplings to the SM prediction, assuming no BSM decay modes, considering modifications to decays only).

GN1: flavour-tagging with Graph Neural Networks



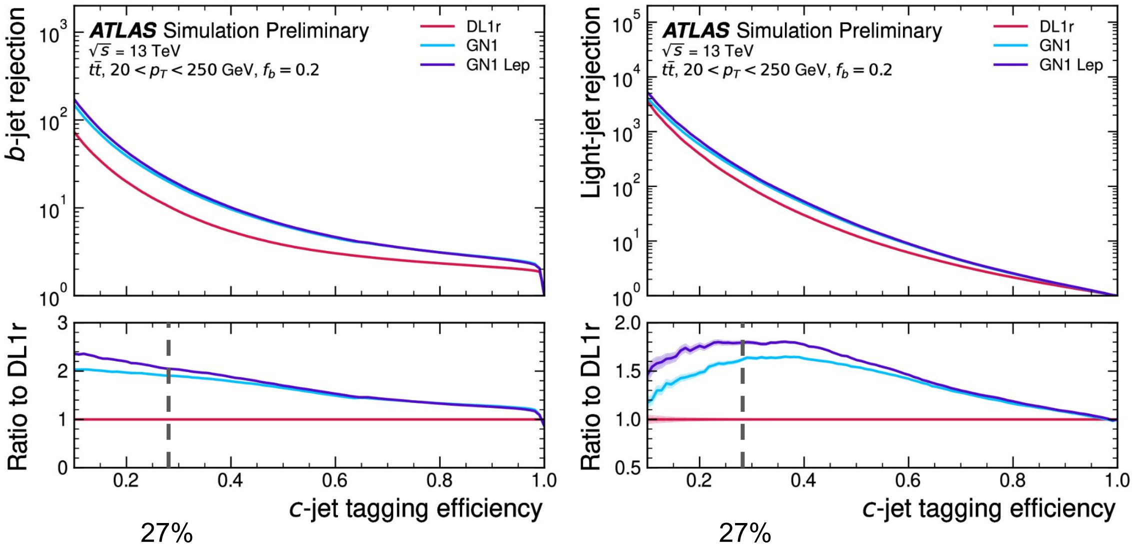
ATL-PHYS-PUB-2022-027

→ **GN1:** state-of-the-art jet flavour-tagging algorithm in ATLAS.

→ Represents jets as **graphs of tracks of charged particles.**

GN1: Improved *c*-tagging performance

ATL-PHYS-PUB-2022-027



→ Factor ~2 improvement for b and light-flavour jet rejection @ a 27% c-jet tagging efficiency.
→ It will directly impact the sensitivity of the ATLAS experiment to the VH,H→cc process.

Boosted VH→bb

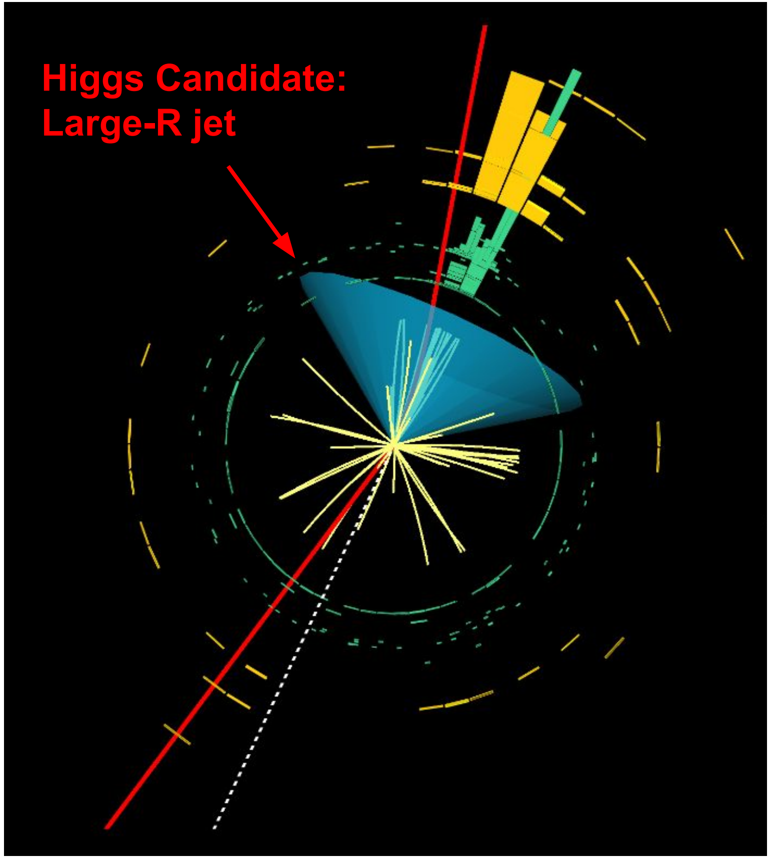
Motivation:

→ VH→bb process observed by ATLAS and CMS in 2018.

→ High-p_T Higgs boson production sensitive to BSM scaling with $\sim (\frac{p_T^H}{\Lambda})^2$

Goals:

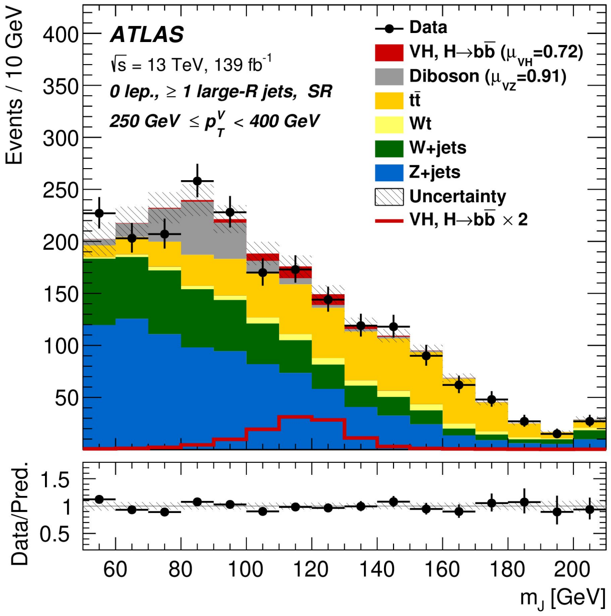
→ $\mu_{VH,Hbb}$ and cross section measurements in bins of p_T of the vector boson.

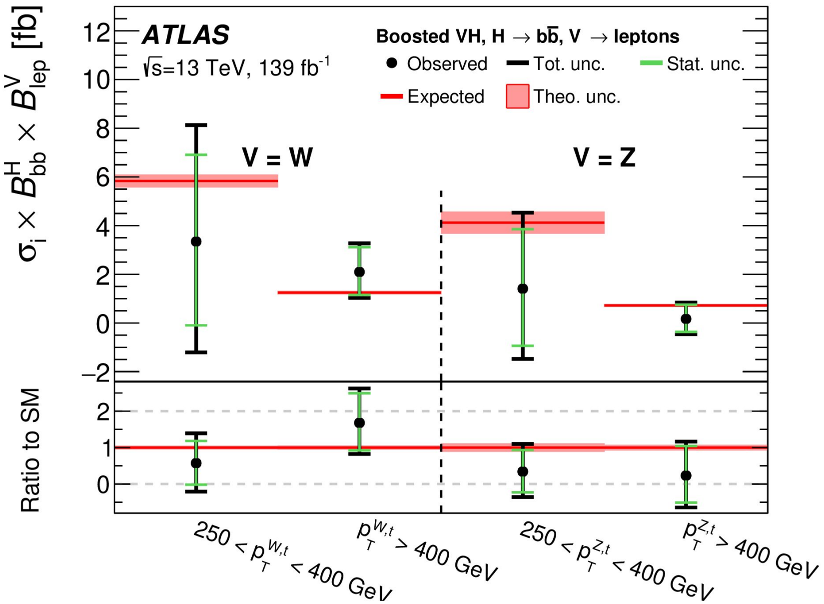


https://doi.org/10.1016/j.physletb.2021.136204

Results

https://doi.org/10.1016/j.physletb.2021.136204





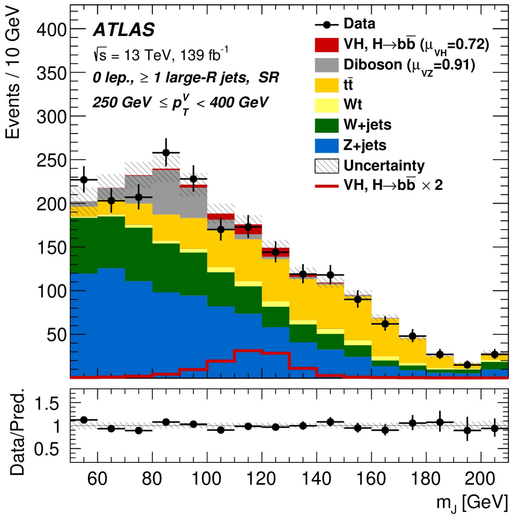
→ Fit of the invariant mass m_J of the Large-R jet:

- $\mu_{VHbb} = 0.72\,^{+0.39}_{-0.36} = 0.72^{+0.29}_{-0.28}$ (stat.) $^{+0.26}_{-0.22}$(syst).
- Cross section measurements in bins of $p_T^{W,Z}$ compatible with the SM (interpreted in an Effective Field Theory to constrain new-physics effects).

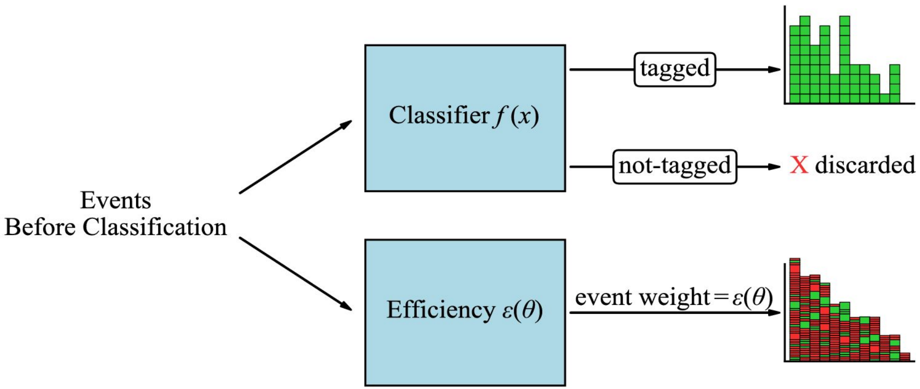
Truth-Tagging Technique

→ Main backgrounds simulated with MC generators.
→ **Truth-Tagging:** weighting events with their *probability* of passing the flavour-tagging cuts, instead of applying the cuts themselves.
→ Important for light-jets (cuts discard most of simulated events).

The Truth-tagging technique strongly relies on the knowledge of the tagging efficiency ϵ for each jet.



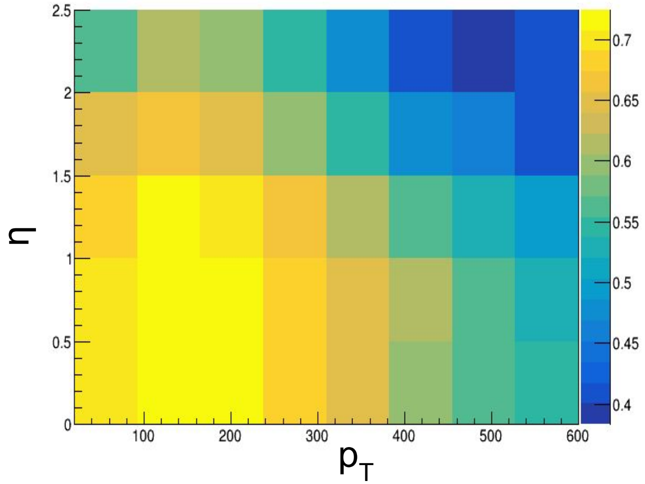
https://doi.org/10.1016/j.physletb.2021.136204



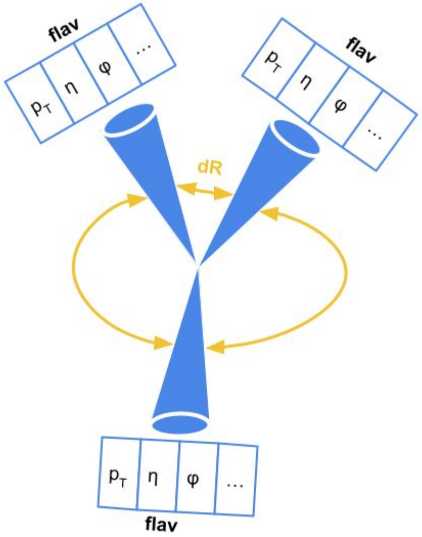
ATL-PHYS-PUB-2022-041

Truth-Tagging with GNNs

$\epsilon_{jet} = \epsilon_{jet}(\vartheta)$, with ϑ set of parameters (e.g. p_T and η) of each jet/event.



VS



Flavour-tagging Efficiency Maps:

→ Simple jet by jet approach.

→ Limited in number of parameters.

GNNs:
→ Events as graphs of fully-connected jets.
→ Allow to scale the problem to higher dimensionalities.
→ Can take into account dependence on event-level variables.

Performance

Good parametrization of ϵ_{jet} → Good closure on cut-based **Direct Tagging,** and less **statistical uncertainty.**

→ **GNN** better than **Maps** at low dR.

VH, H→cc: impact of uncertainties

Source of uncertainty		$\mu_{VH(c\bar{c})}$	$\mu_{VW(cq)}$	$\mu_{VZ(c\bar{c})}$
Total		15.3	0.24	0.48
Statistical		10.0	0.11	0.32
Systematic		11.5	0.21	0.36

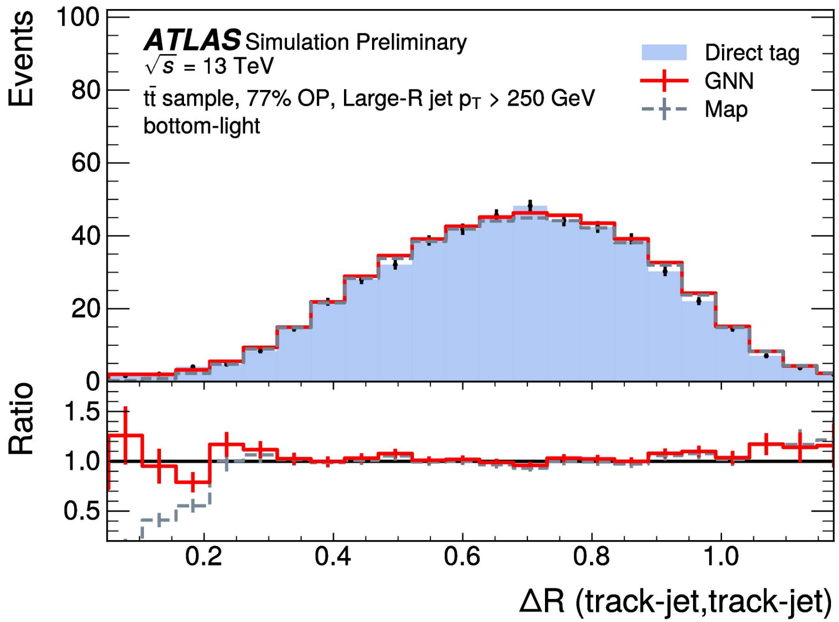
...

Flavour tagging	c-jets	1.6	0.05	0.16
	b-jets	1.1	0.01	0.03
	light-jets	0.4	0.01	0.06
	τ-jets	0.3	0.01	0.04
Truth-flavour tagging	ΔR correction	3.3	0.03	0.10
	Residual non-closure	1.7	0.03	0.10

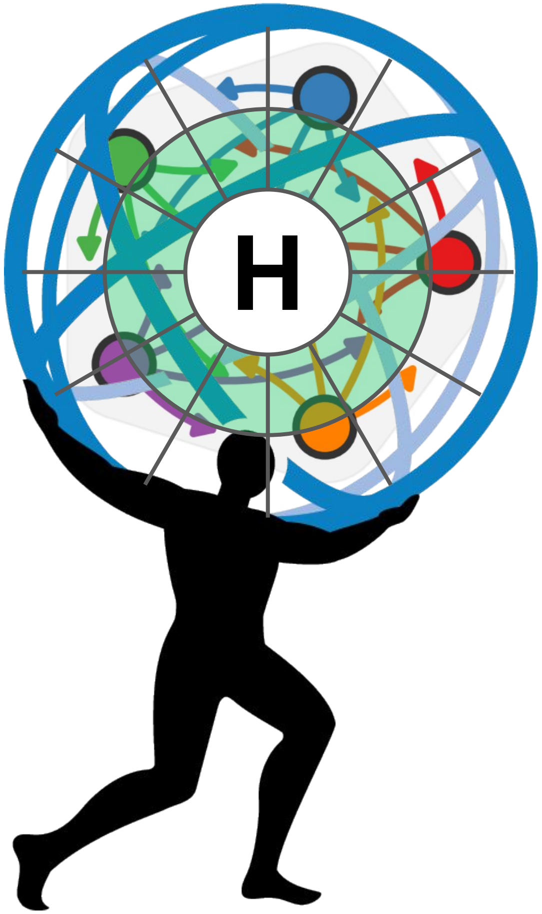
10.1140/epjc/s10052-022-10588-3

Non-closure of map-based truth-tagging

Boosted VH,H→bb: Closure

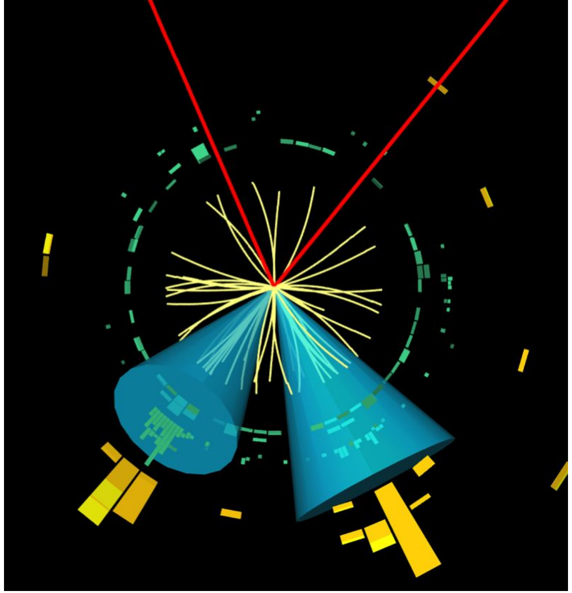


ATL-PHYS-PUB-2022-041


H

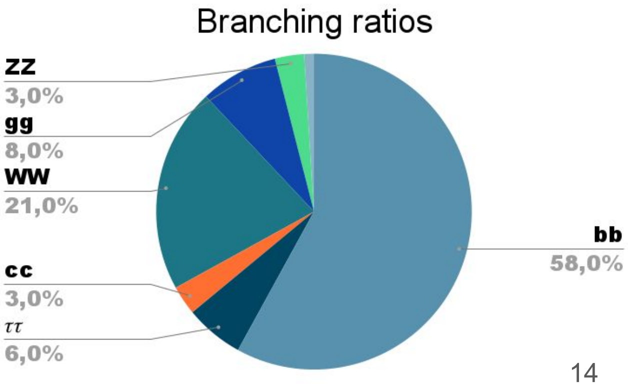
Backup

Search for VH(H→cc) decays

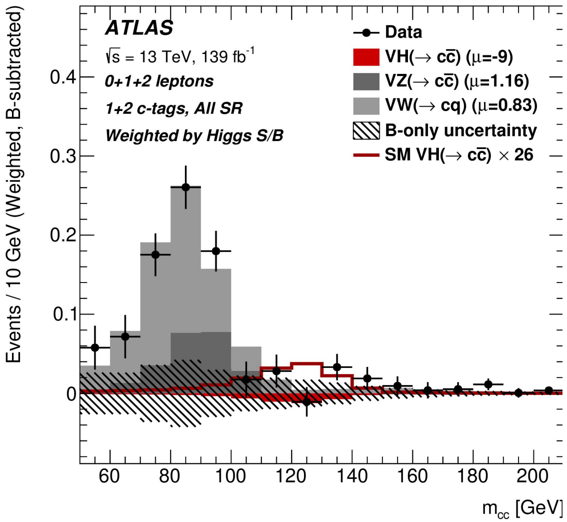


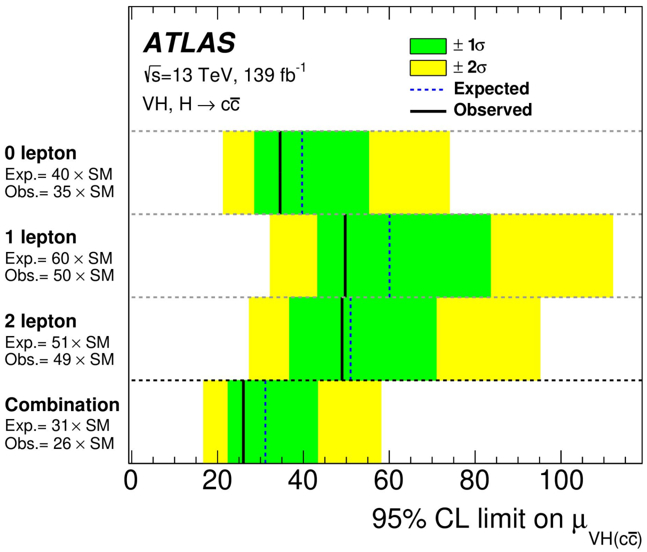
Motivation:
→ So far, we observed directly only Higgs couplings to 3rd generation fermions (ttH, H→$\tau\tau$, H→bb).
→ Probing couplings to lighter 2nd generation fermions could **open windows on new physics**.
→ H→cc is one of the **most common** unobserved decay modes of the Higgs boson.

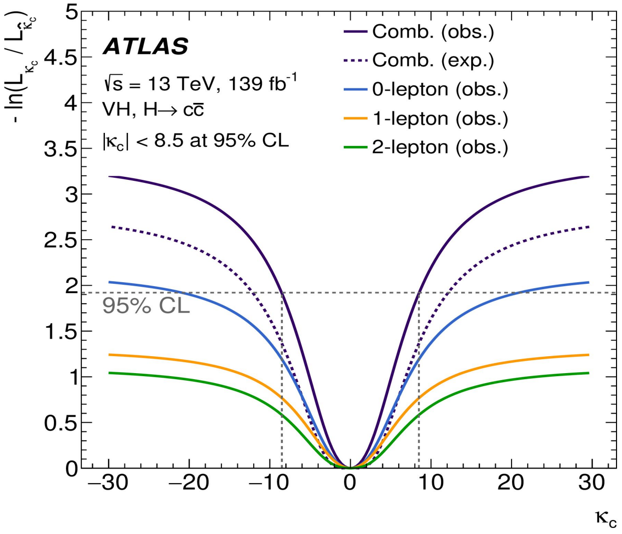
Strategy:
→ Overwhelming QCD background suppressed using leptonic decays of W and Z bosons.
→ **Deep learning based dedicated *c*-tagger** with 27% c-jet efficiency, 8.3% b-jet efficiency, and 1.7% light jet efficiency.
→ Dedicated control regions to constrain main backgrounds (ttbar and V+jets).
→ Profile likelihood fit on the invariant mass of the *c*-jet pair to extract the three parameters of interest μ_{VHcc} , μ_{VWcc} and μ_{VZcc}. The VZ/VW → cc signal strengths are extracted to validate the strategy.



Results







→ The Diboson cross check measurement is in agreement with the SM.

→ Upper limit is set on VH→cc process of **26 times the SM prediction**.

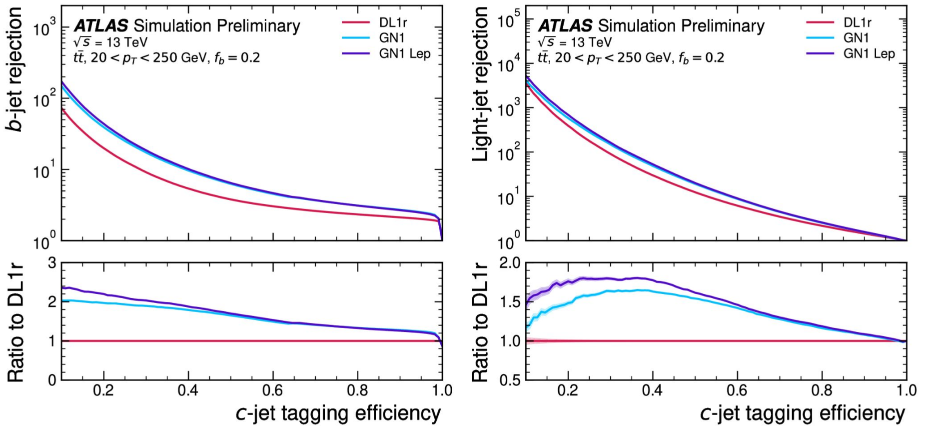
→ Results interpreted in the kappa framework as constraining the coupling modifier **$|k_c|$<8.5 @ 95%CL.**

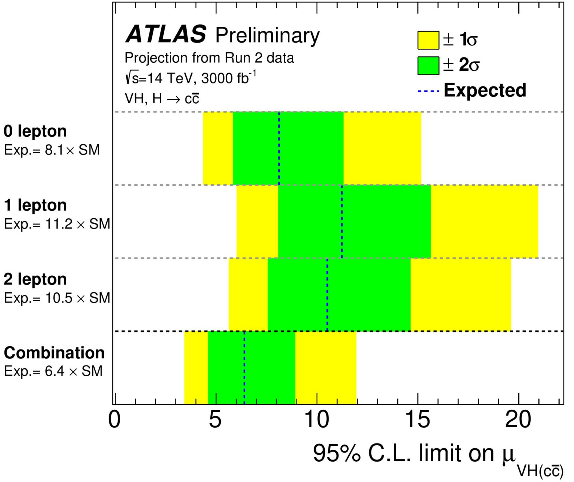
Improved *c*-tagging towards HL-LHC

→ GN1 outperforms its predecessors of factors up to 100% (80%) for b (light) flavour jet rejection @ a 27% c-jet tagging efficiency.

→ The impressive improvement will directly impact the sensitivity of the ATLAS experiment to the VH→cc process.

→ Extrapolations of the expected sensitivities of the ATLAS detector to the VH→cc process at the HL-LHC, assuming a dataset of 3000 fb^{-1} of *pp* collisions, the presence of ITk **and a *b* (light) flavour jet rejections improved of a factor of 1.5 (3)**, predict an upper limit on μ_{VHcc} of 6.4 x SM, corresponding to $|k_c|<8.5$ @ 95%CL.





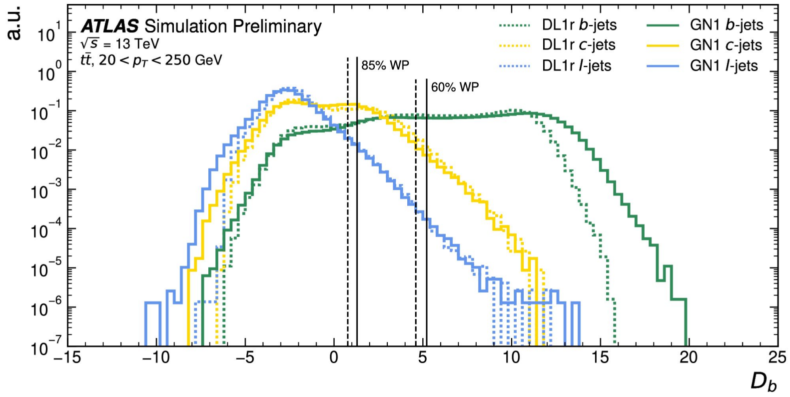
Truth-Tagging Technique

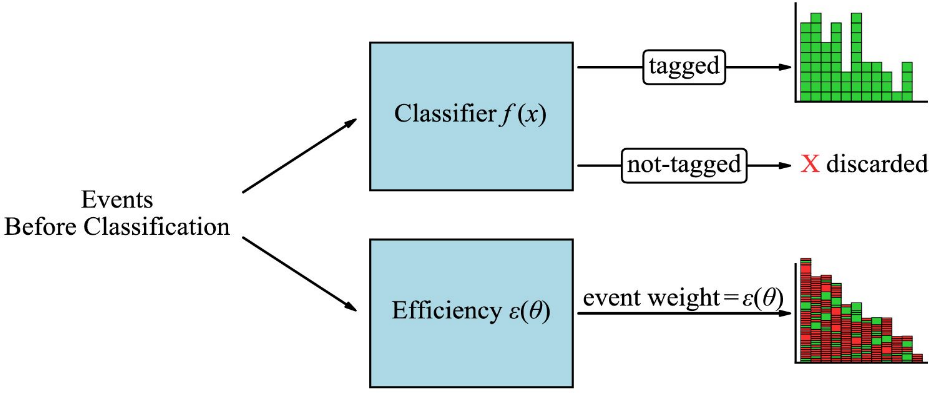
Direct Tagging (Pass or Fail): Define as "tagged" only jets passing a given cut on the flavour-tagging discriminant variable. Jets that don't pass the cut are discarded.

Truth Tagging: Weight jets with their *probability* of passing the cut and thus being tagged. No jet is discarded: **the statistical power of the simulated background samples is optimally exploited.**

The event weighting technique strongly relies on the *a priori* knowledge of the tagging efficiency ϵ for each jet.

ϵ is a function of a set of parameters (e.g. the phase space coordinates) ϑ of each jet: $\epsilon_{jet}=\epsilon_{jet}(\vartheta)$.





ATL-PHYS-PUB-2022-041

Source of uncertainty		$\mu_{VH(c\bar{c})}$	$\mu_{VW(cq)}$	$\mu_{VZ(c\bar{c})}$
Total		15.3	0.24	0.48
Statistical		10.0	0.11	0.32
Systematic		11.5	0.21	0.36
Statistical uncertainties				
Signal normalisation		7.8	0.05	0.23
Other normalisations		5.1	0.09	0.22
Theoretical and modelling uncertainties				
$VH(\rightarrow c\bar{c})$		2.1	< 0.01	0.01
$Z+$jets		7.0	0.05	0.17
Top quark		3.9	0.13	0.09
$W+$jets		3.0	0.05	0.11
Diboson		1.0	0.09	0.12
$VH(\rightarrow b\bar{b})$		0.8	< 0.01	0.01
Multi-jet		1.0	0.03	0.02
Simulation samples size		4.2	0.09	0.13
Experimental uncertainties				
Jets		2.8	0.06	0.13
Leptons		0.5	0.01	0.01
$E_{\mathrm{T}}^{\mathrm{miss}}$		0.2	0.01	0.01
Pile-up and luminosity		0.3	0.01	0.01
Flavour tagging	c-jets	1.6	0.05	0.16
	b-jets	1.1	0.01	0.03
	light-jets	0.4	0.01	0.06
	τ-jets	0.3	0.01	0.04
Truth-flavour tagging	ΔR correction	3.3	0.03	0.10
	Residual non-closure	1.7	0.03	0.10

Source of uncertainty		$\mu_{VH(c\bar{c})}$	$\mu_{VW(cq)}$	$\mu_{VZ(c\bar{c})}$
Total		15.3	0.24	0.48
Statistical		10.0	0.11	0.32
Systematic		11.5	0.21	0.36
Flavour tagging	c-jets	1.6	0.05	0.16
	b-jets	1.1	0.01	0.03
	light-jets	0.4	0.01	0.06
	τ-jets	0.3	0.01	0.04
Truth-flavour tagging	ΔR correction	3.3	0.03	0.10
	Residual non-closure	1.7	0.03	0.10