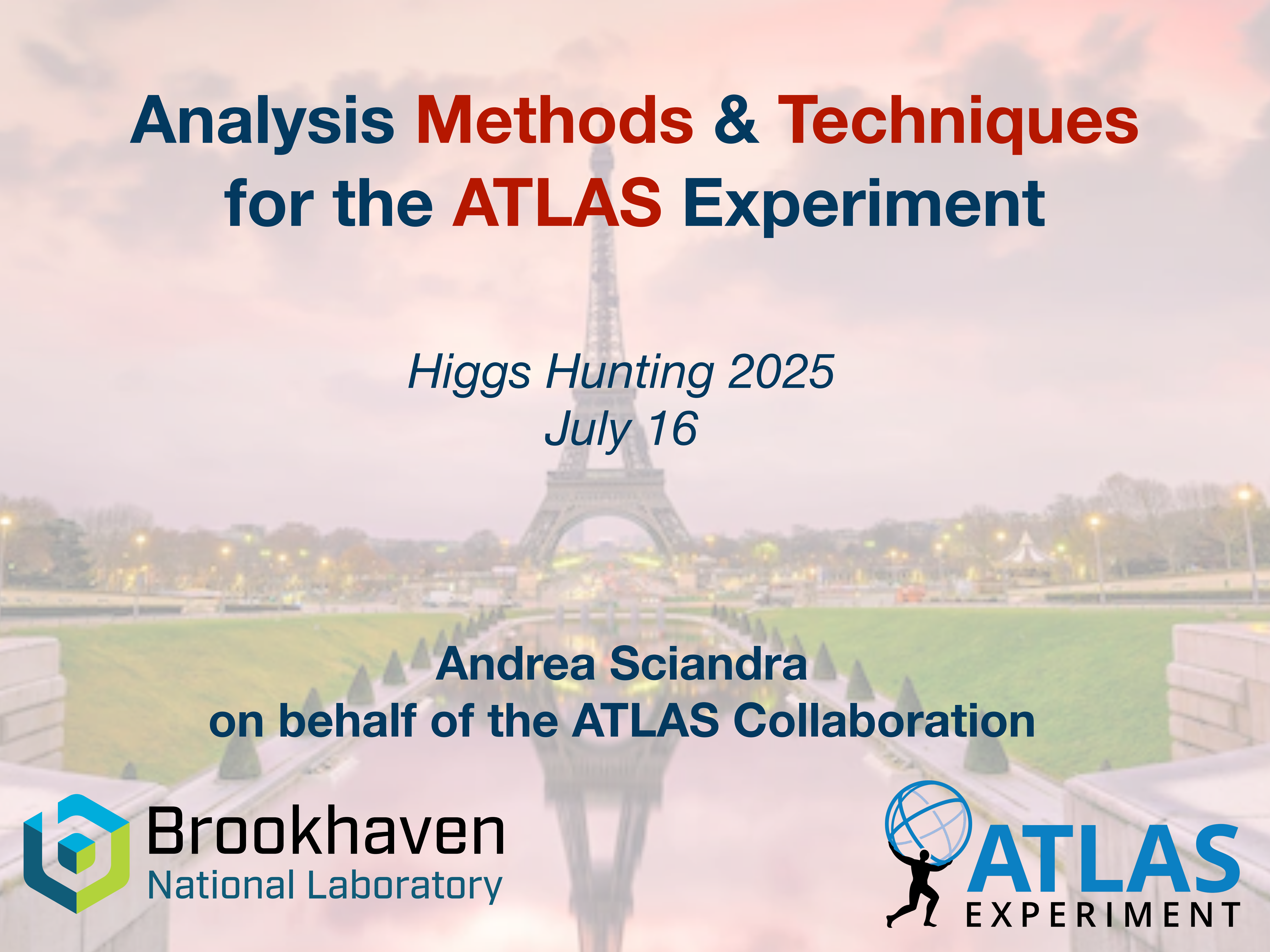


Analysis Methods & Techniques for the ATLAS Experiment

Higgs Hunting 2025
July 16

Andrea Sciandra
on behalf of the ATLAS Collaboration

Brookhaven National Laboratory

ATLAS EXPERIMENT

Introduction & Overview

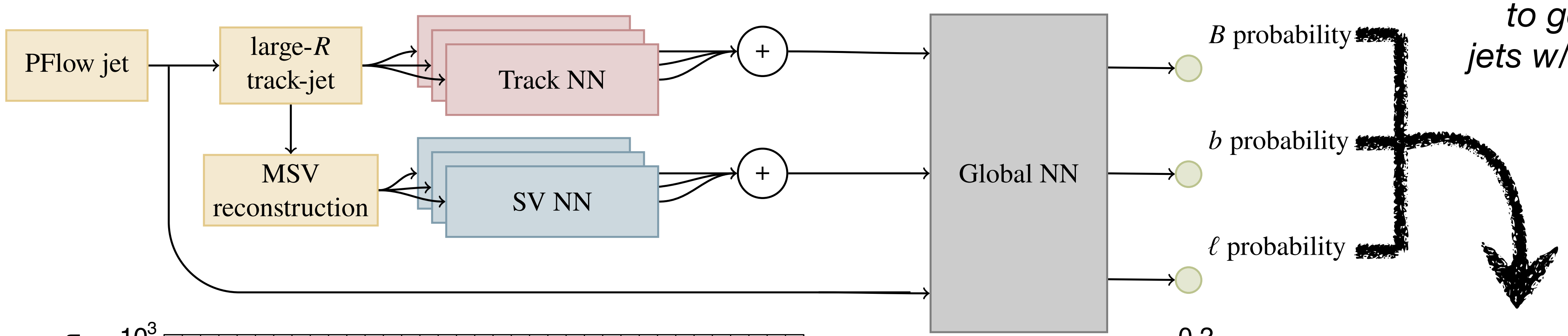
- Breakthrough innovations in methods, tools & techniques allow for striking improvements in LHC physics reach
  - Often come with proportionate challenges
- Focus on methods & techniques exploited by most recent (B)SM Higgs analyses published by ATLAS:
  - **Reconstruction & identification of *b, bb* & *ττ* topologies**
  - **Dedicated event-reconstruction algorithms**
  - **Weakly supervised anomaly detection**
  - **Neural Simulation-Based Inference (NSBI) for parameter estimation**

… and applications to real physics-cases!

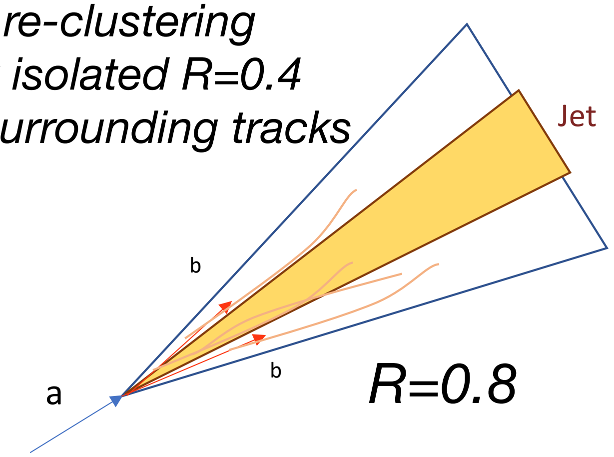
DeXTer: Specialised Flavour Tagging

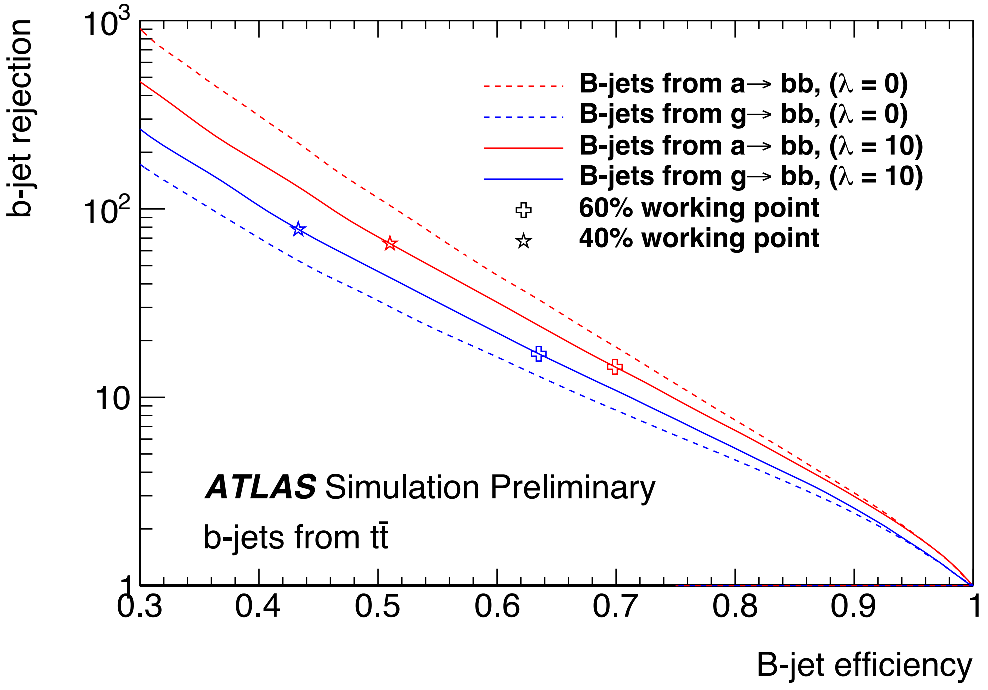
ATL-PHYS-PUB-2022-042

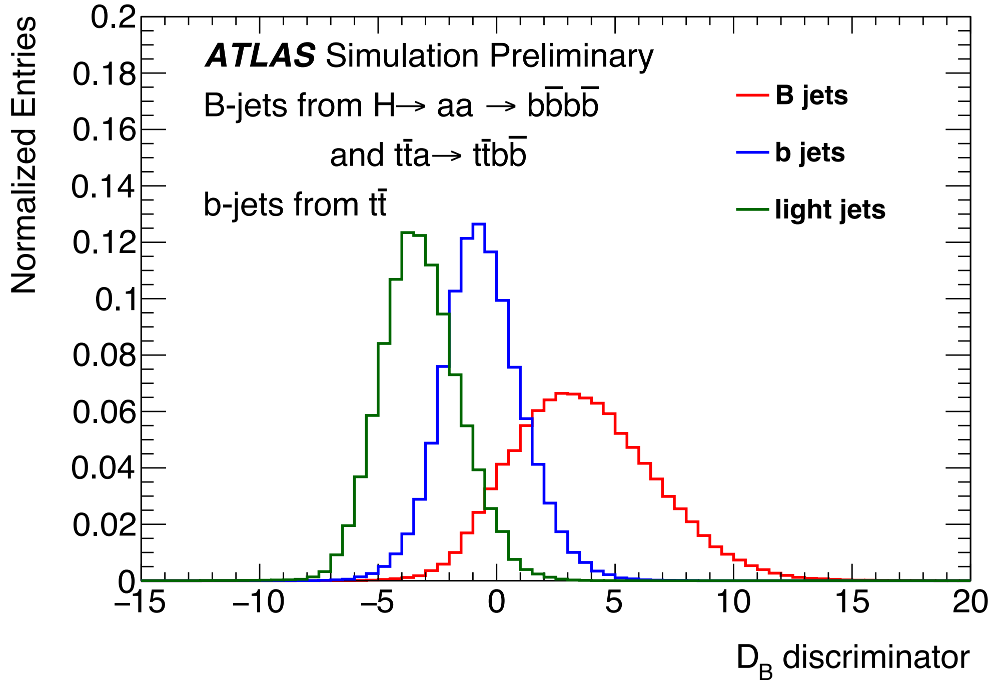
- **De**ep-set ***X***→*bb* **T**agg**er**: general-purpose **low-mass/p_T double-*b* tagging algorithm**
  - Specialised for jets 20<p_T<200 GeV: surrounding displaced tracks +multiple secondary vertices (MSV)
- **Domain-adversarial training**: remove different response on color singlets & octets
- Used in multiple new ATLAS BSM Higgs searches



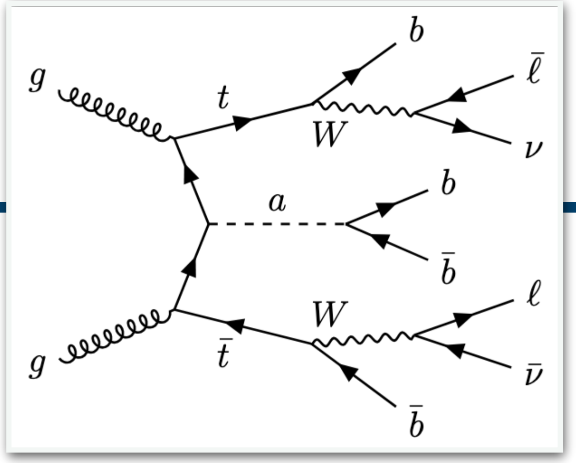
Jet re-clustering to get isolated R=0.4 jets w/ surrounding tracks



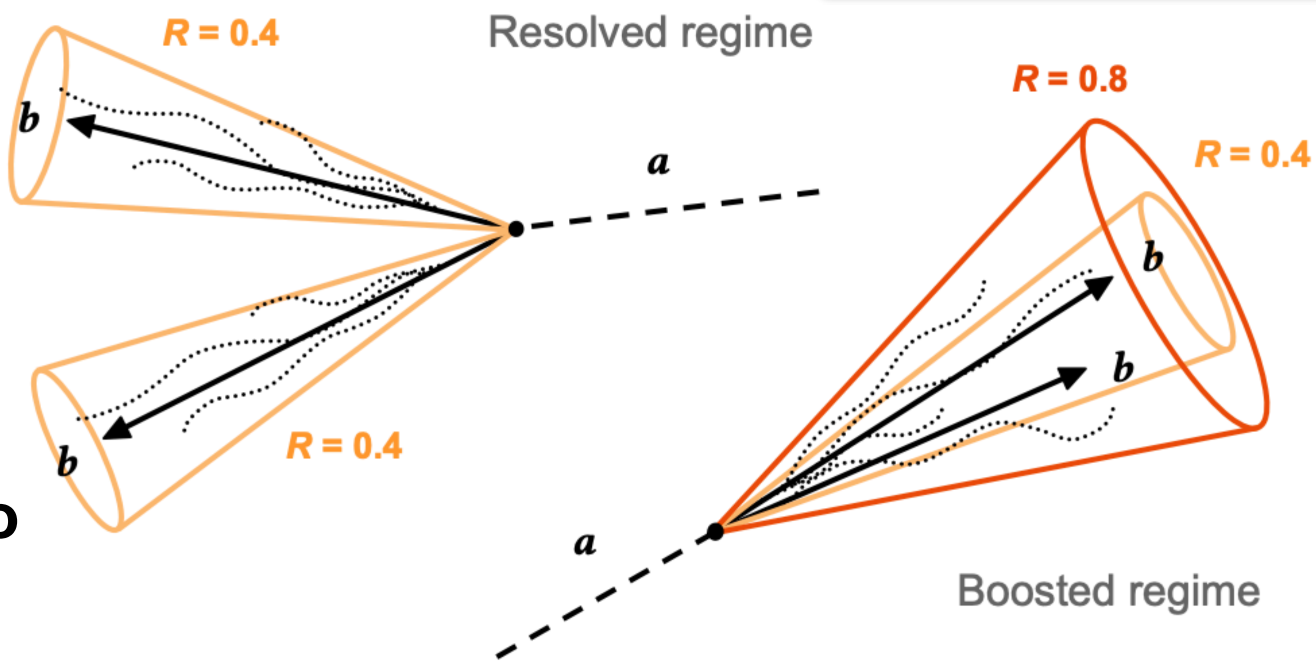


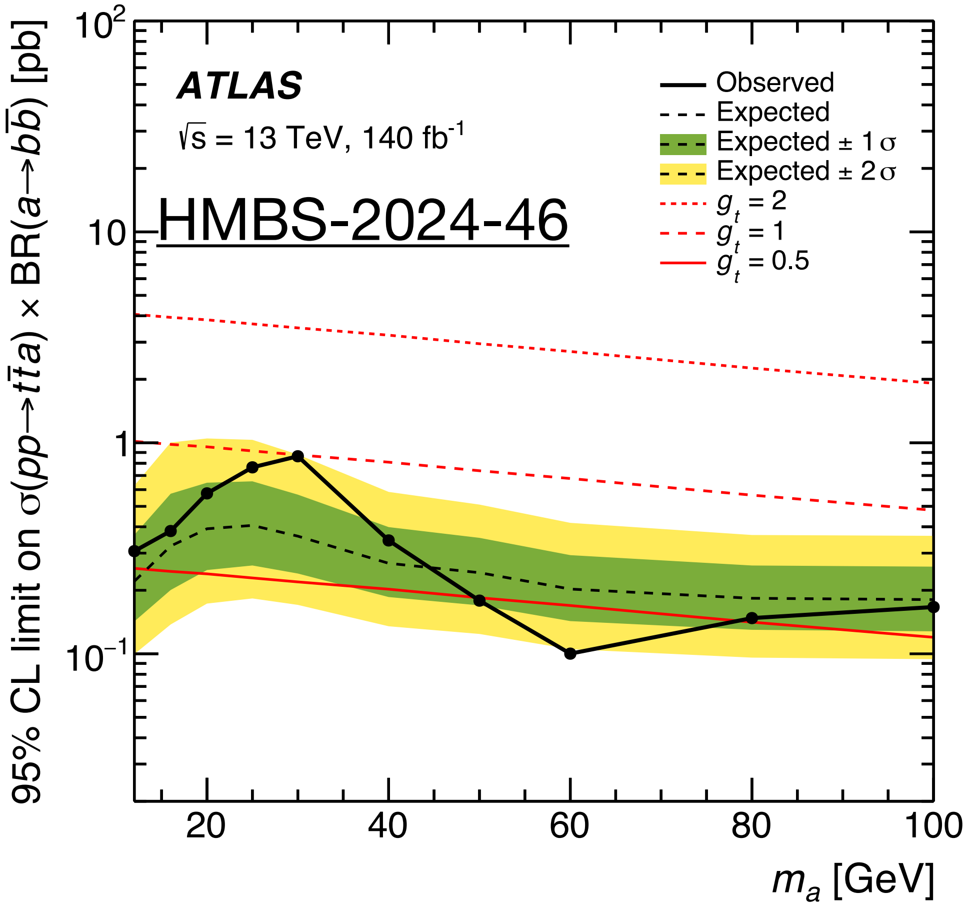


Search for New Pseudoscalar: *tt/tW+a(→bb)*



- Exploring both resolved & merged *bb* topologies in *tt/tW+a(→bb)* 2ℓ channel
  - **Std resolved** (DL1r) **& merged** (DeXTer) **flavour taggers combined**
- Extensive & combined usage of ML techniques
  - **Event reconstruction combines two BDTs, one for *t→jℓ* & one for *a→jj* reco**



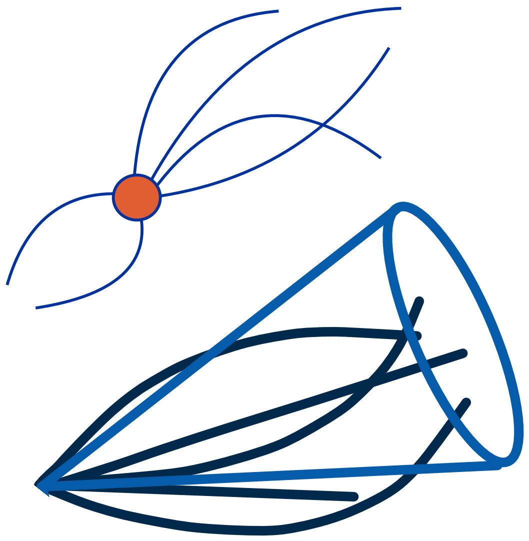


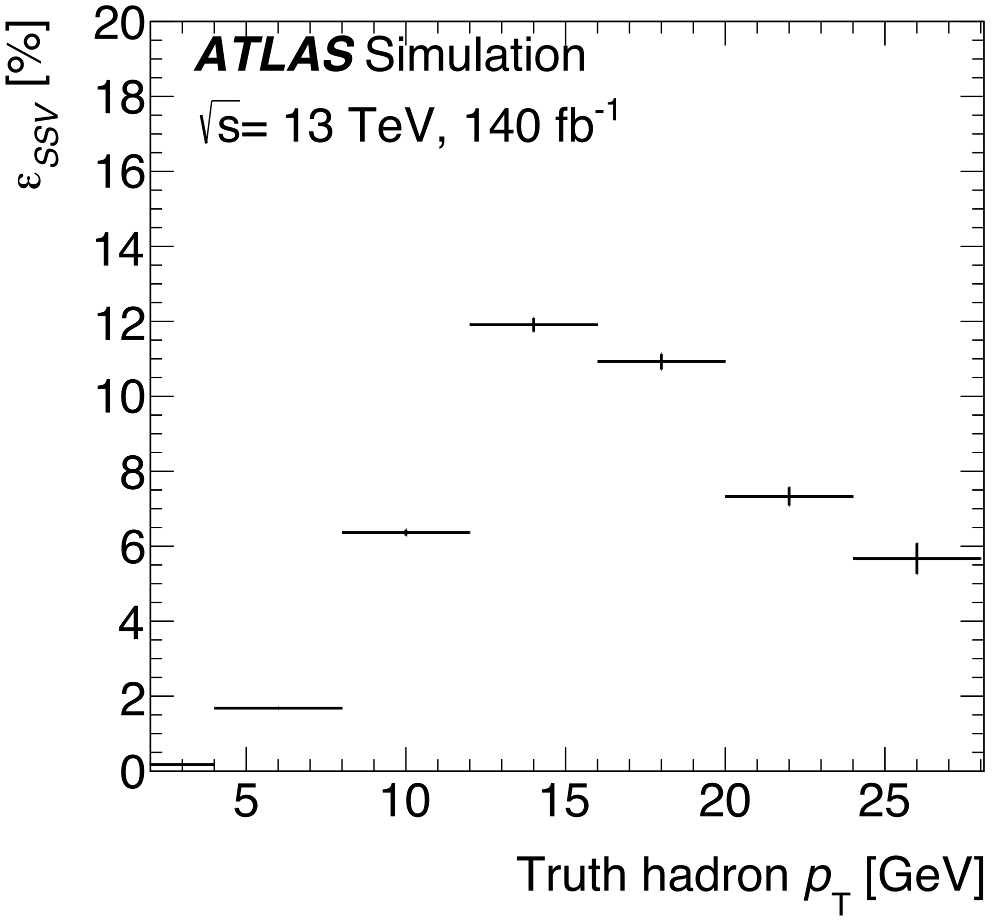
- Split into 4 SRs, based on small- (b) & large-R (B) tagged-jet multiplicity
- **One mass-parameterised deep NN per SR,** exploited as fit observable
- No excess found, competitive limits excluding 12<m_a<100 GeV for a coupling g_t=1

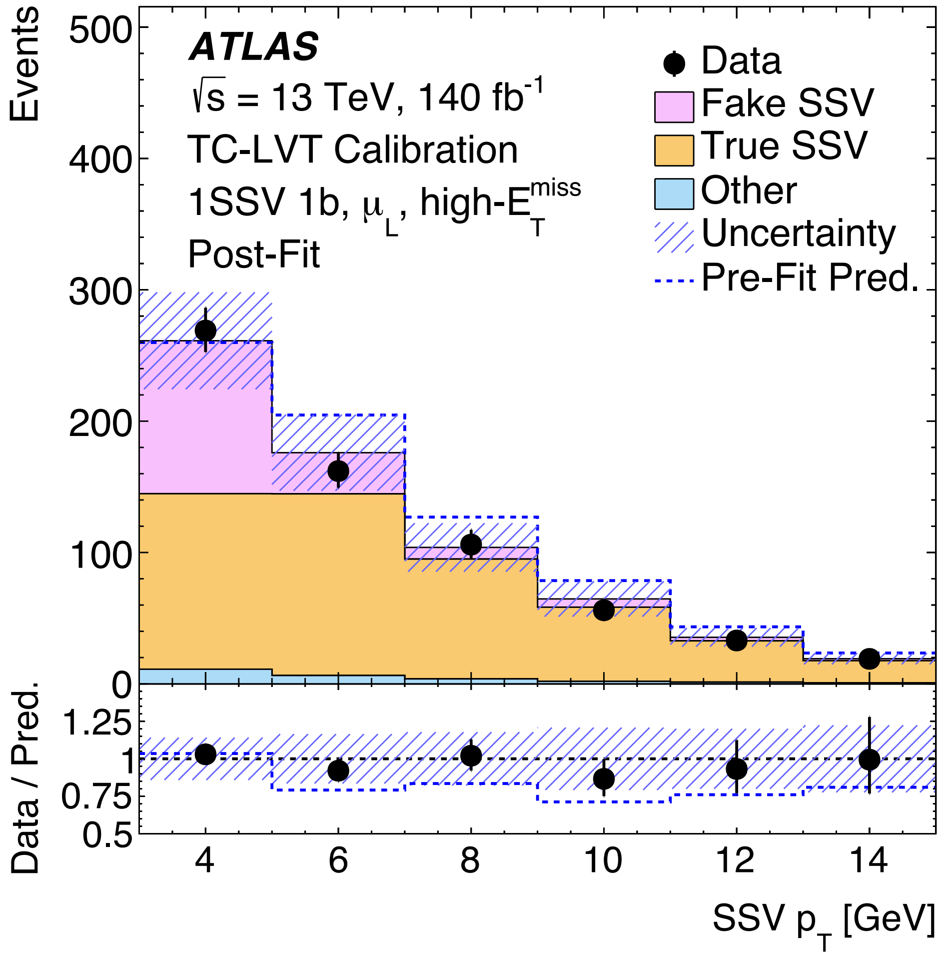
TC-LVT: “Jet-less” Flavour Tagging

FTAG-2023-02

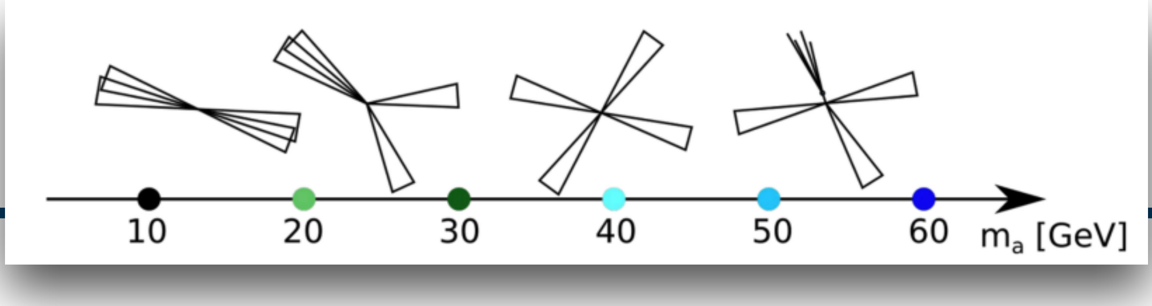
- Dedicated *b*-tagging algorithm to identify low-p_T (5-20 GeV) B-hadrons outside of jets, by reconstructing their **soft secondary vertices**
- **T**rack-**C**luster based **L**ow-p_T **V**ertex **T**agger
  - Based on std vertexing techniques, retuned & applied outside jets
  - Reconstruct displaced soft secondary vertices from seed tracks
  - Dedicated **calibration** on *tt* events



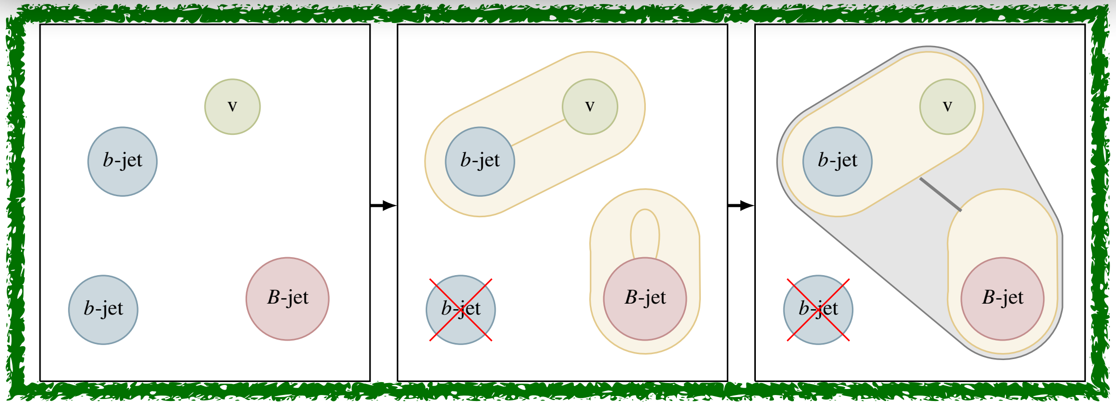




Search for $Z(\to\ell\ell/vv)H(\to 4b/6b)$

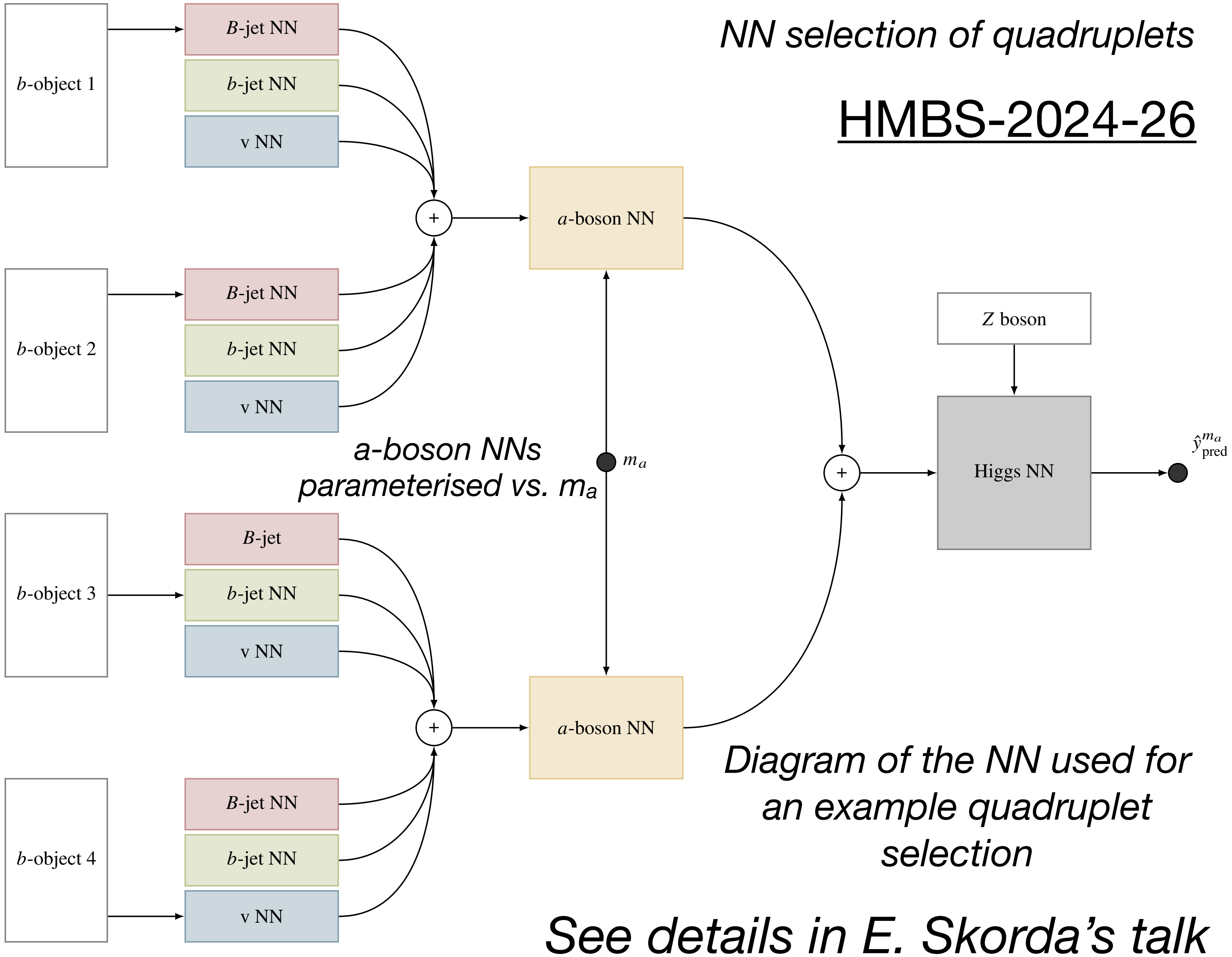


- Search for $H\to aa/a_1a_2\to 4b/6b$ for $12<m_a<60$ GeV with $Z\to\ell\ell/vv$
- Improves previous searches by means of **3 different tagging algs**:
  - Resolved jets (DL1r)
  - Low-mass merged jets (DeXTer)
  - **Soft secondary vertices** (TC-LVT)
- Strategies for different signatures:
  - 2ℓ 4b: **NN for jet-parton pairing** + BDT
  - 2ℓ 6b: BDTs for partially reco events
  - 0ℓ 4b: Cut-&-count
- No significant excess above SM bkg-only expectation



NN selection of quadruplets

HMBS-2024-26

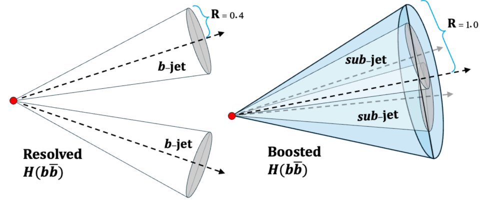


a-boson NNs parameterised vs. m_a

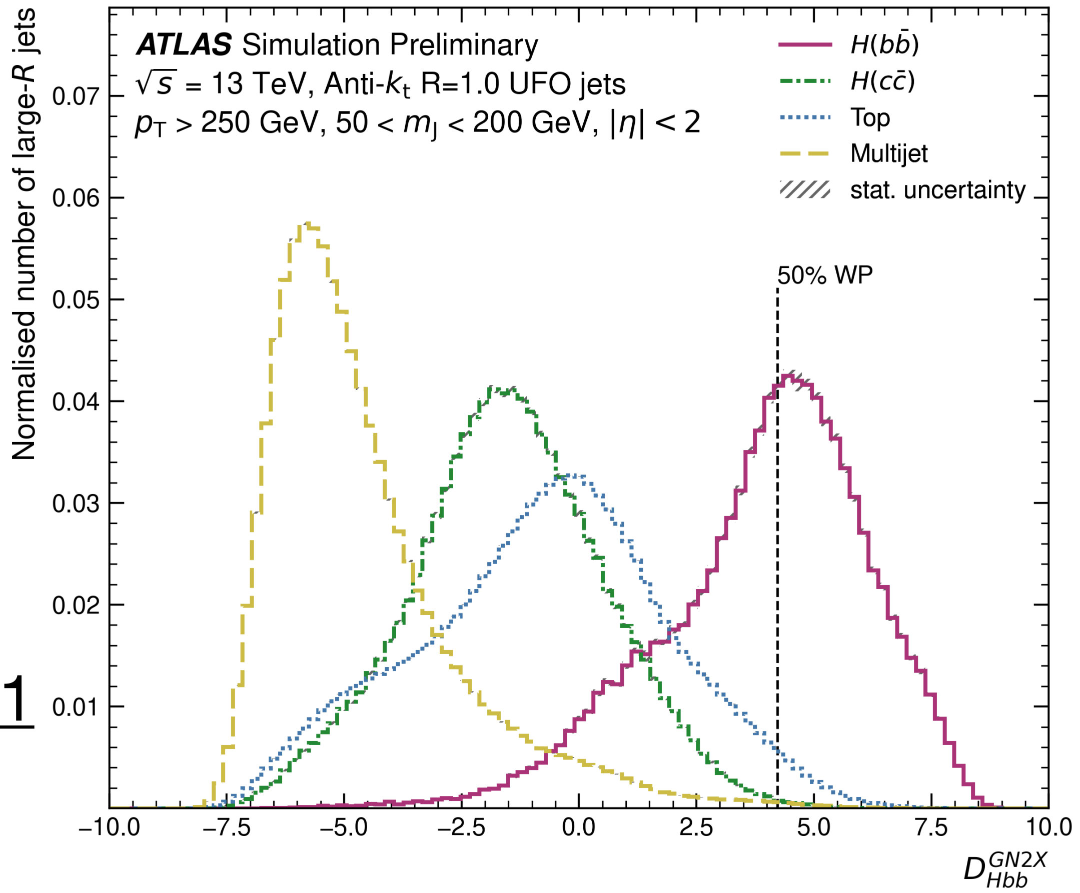
Diagram of the NN used for an example quadruplet selection

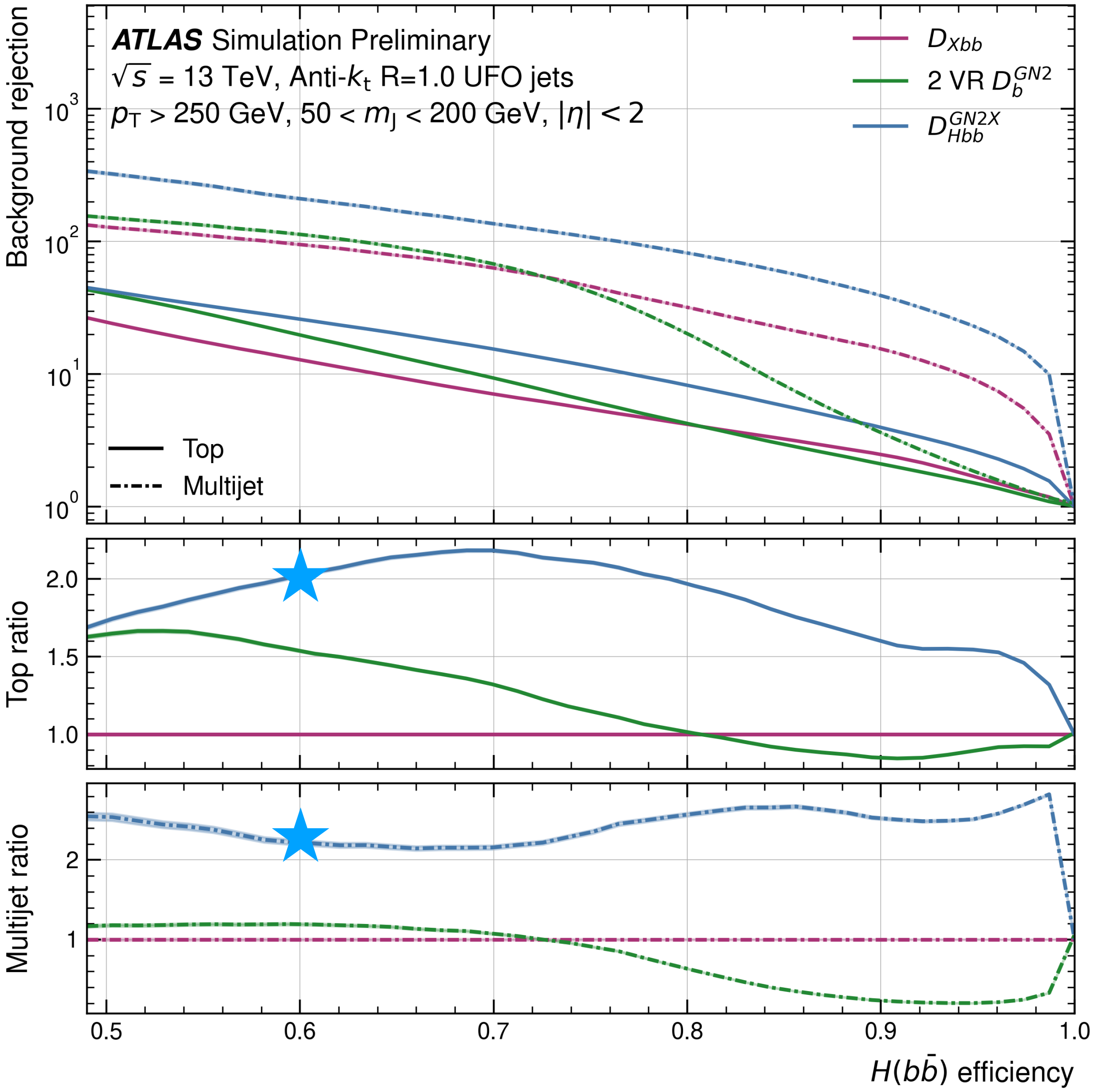
See details in E. Skorda's talk

GN2X *H(bb/cc)* at High Momentum



- **GN2X**: **transformer**-based ***X→bb* tagger** exploiting full info from tracks within large-R jet
- Trained on mass-decorrelated samples to discriminate **boosted *H→bb*, *H→cc*, had top & QCD jets**





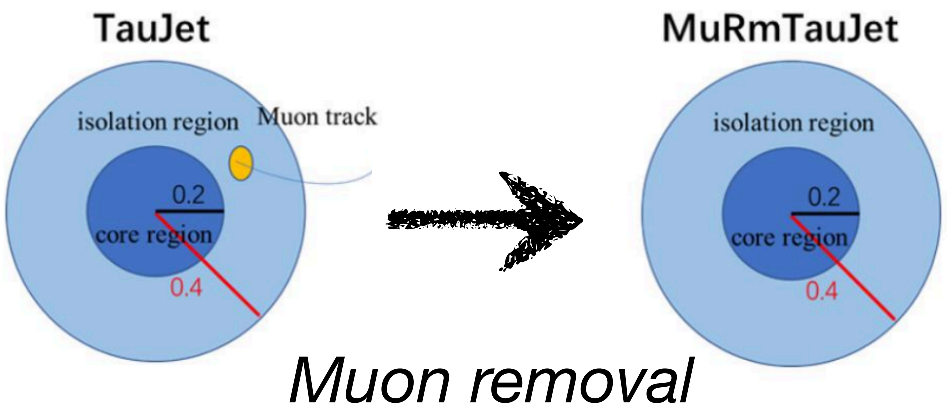
ATL-PHYS-PUB-2023-021

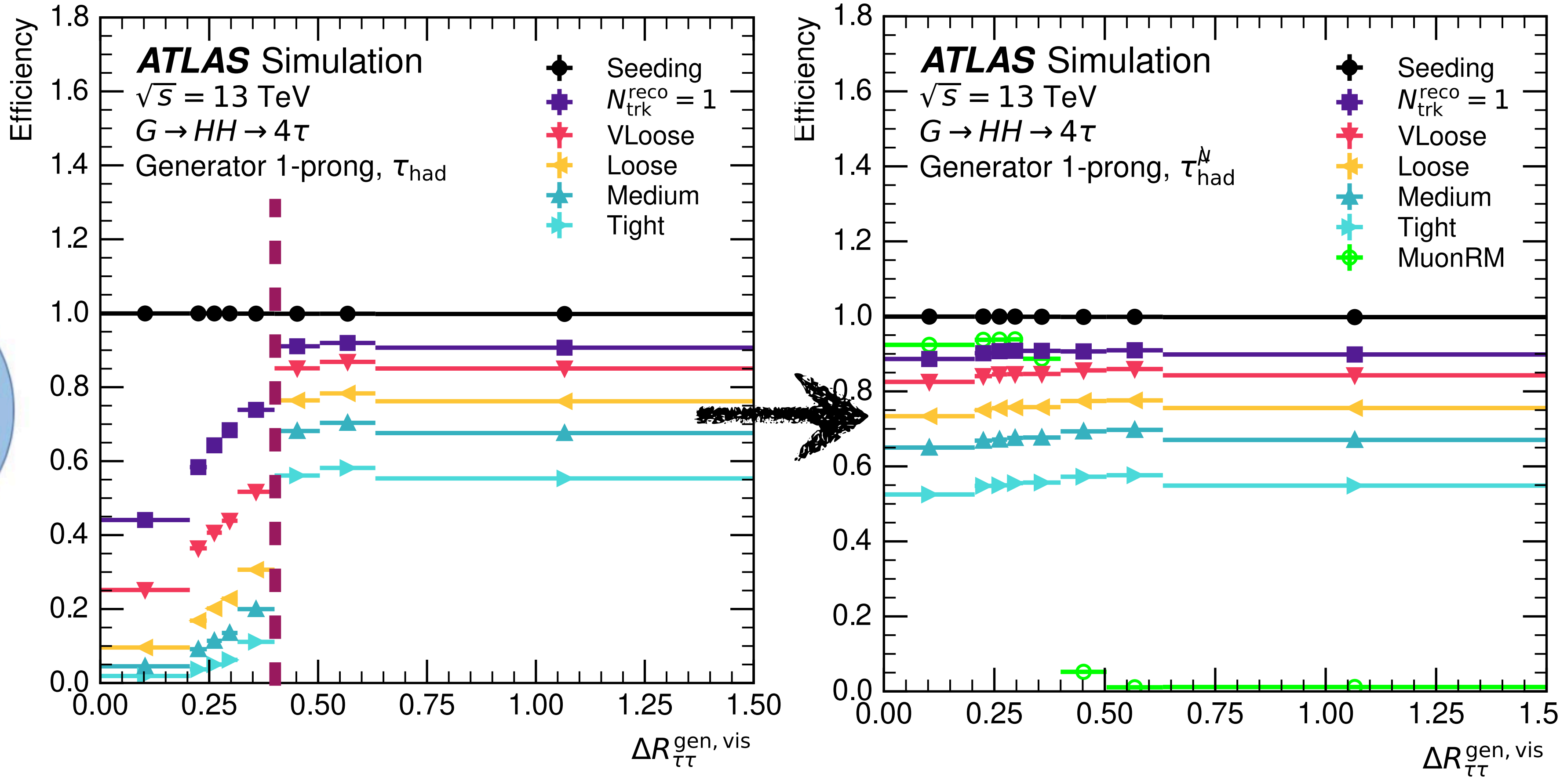
- At 60[50]% *H→bb*[*cc*] efficiency >2[3-5] better top, QCD [and *H→bb*] rejection wrt standard tagging
- **Boost high-p_T measurements/searches for *H→bb* & *HH→4b* (e.g. *VBF HH→4b* & κ_{2V})**

τ Reco with μ-τ Removal in $\tau\tau \to \mu\nu_\mu\nu_\tau$ had+ν_τ

TAUP-2023-02

- Standard τ_{had} reco based off anti-k_t R=0.4 topological calo-cluster jets
  - Identification through RNN exploiting info from tracks & calo energy clusters
- When decay products overlap, sub-optimal τ_{had} reco & identification
- μ leave tracks in ID & MS: identification independent of isolation
- Both ID tracks & calo clusters associated with **μ removed from τ_{had}**
- **τ identification RNN input variables recalculated after μ removal**
  - **Performance fully recovered**, stability observed & validation on data performed

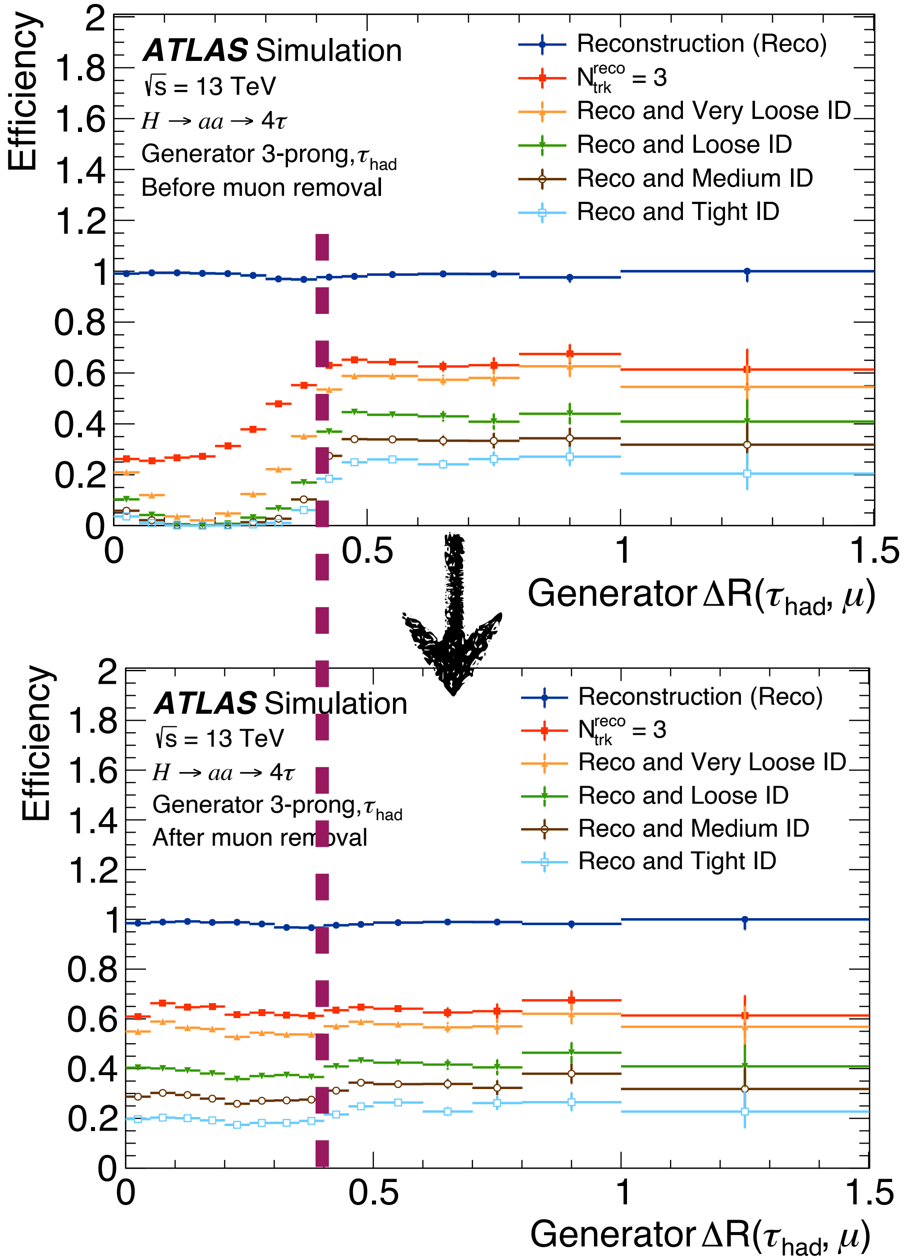


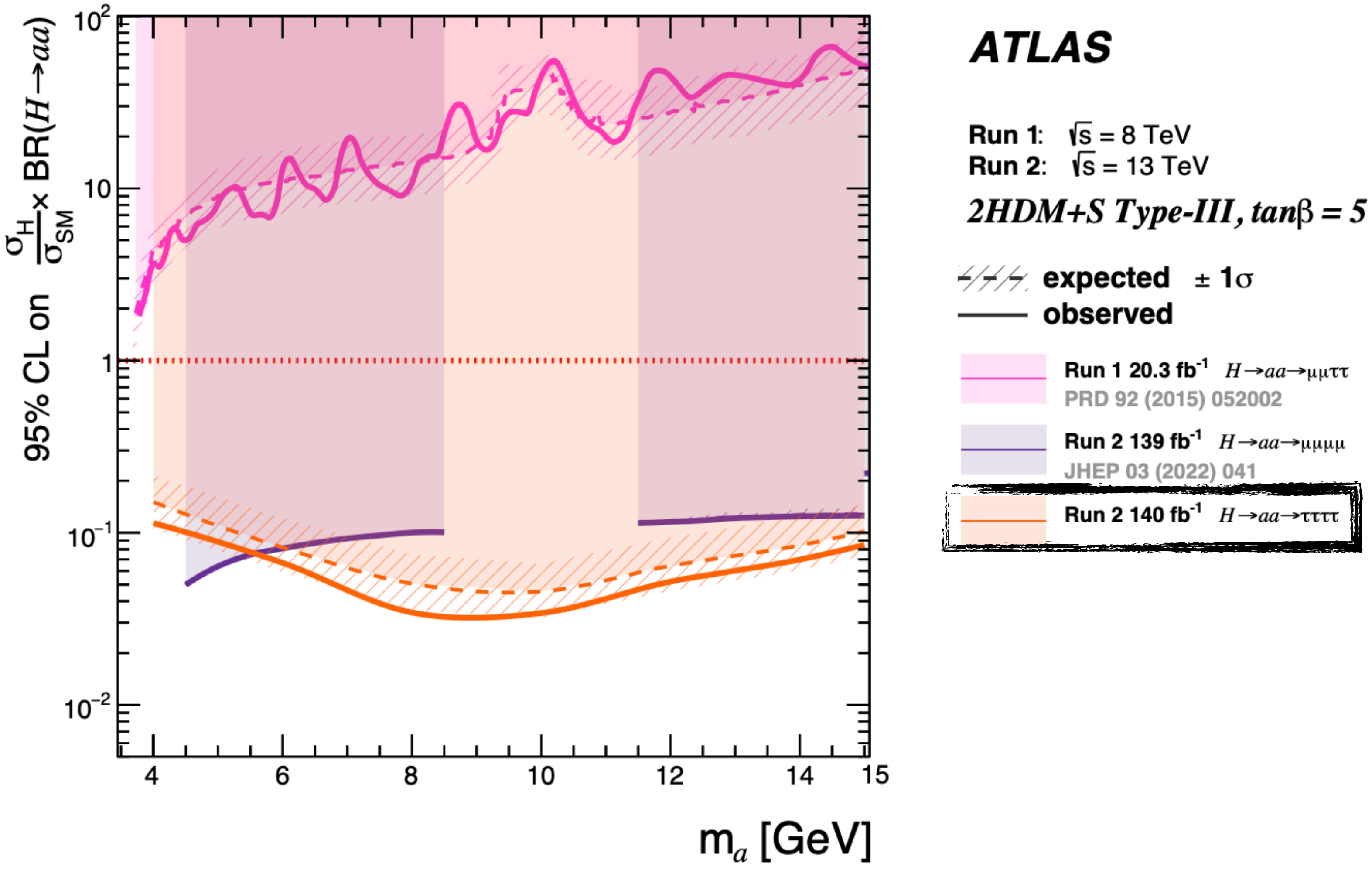


Search for Merged $H\to aa\to\tau_\mu\tau_{had}\ \tau_\mu\tau_{had}$

HMBS-2024-25

- Model-independent search for $\tau_\mu\tau_{had}$ (~23% of $\tau\tau$, low bkg) resonance
  - Focus on 4<m_a<15 GeV
  - **First boosted low-mass 4τ search at ATLAS**, both SS & OS $\mu\mu$ pairs considered
- ***μ-τ* removal technique implemented to resolve merged di-*τ* identification**
- No excess found: **stringent 95% CL upper limits** set

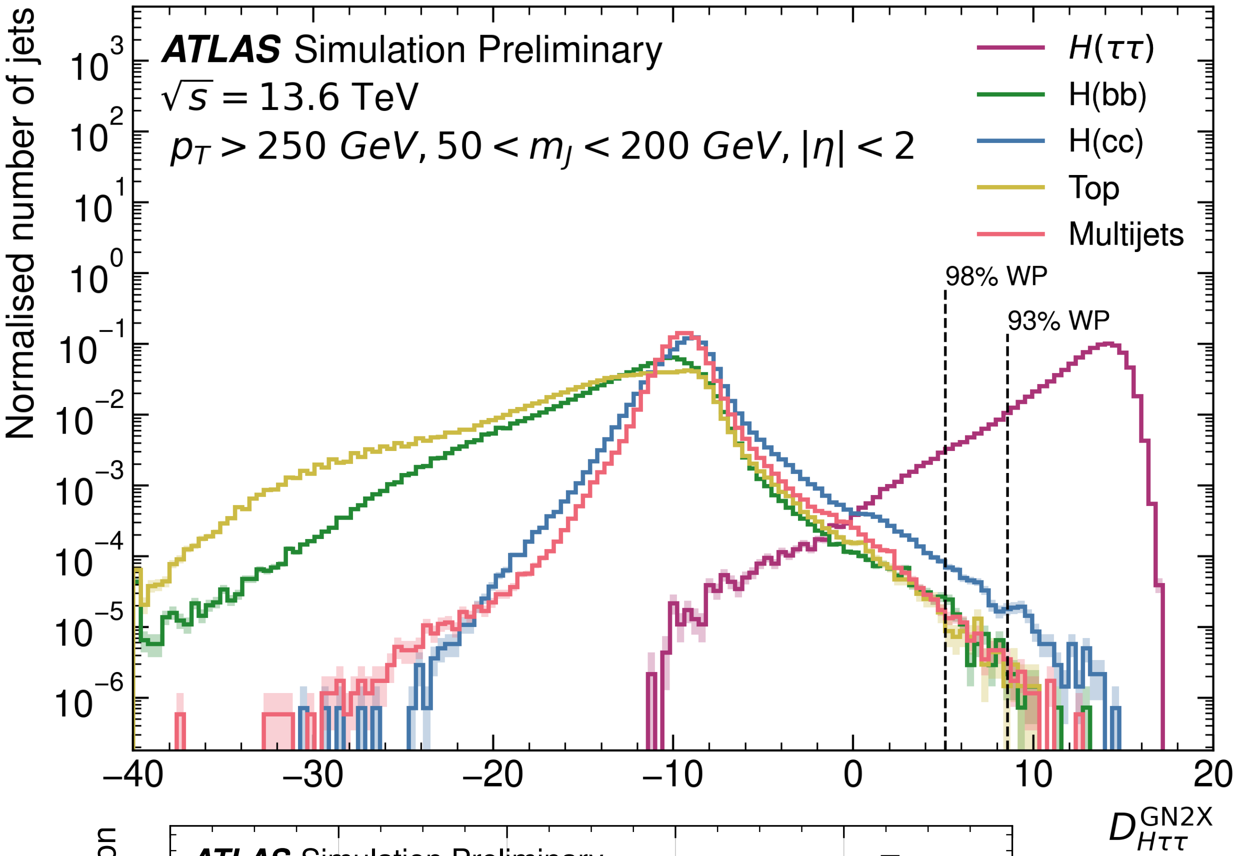


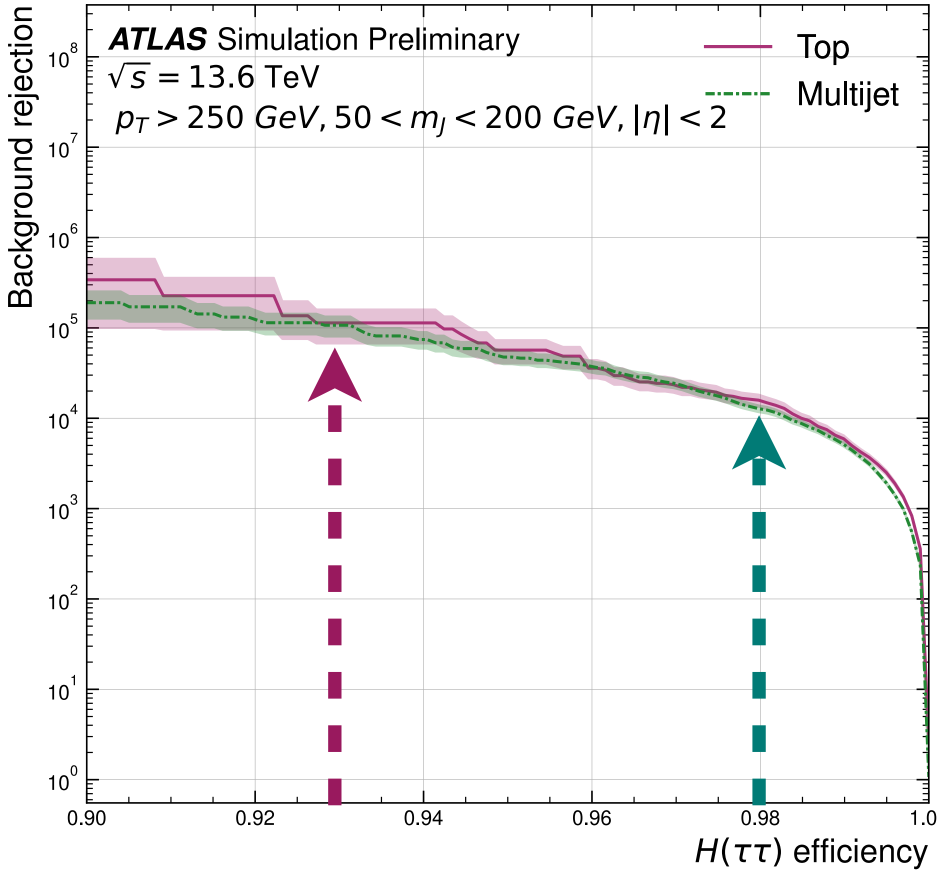


See details in E. Skorda's talk

GN2X$\tau\tau$ for Merged $\tau_{had}\tau_{had}$ at High Momentum

- Extension of GN2X to tag ***H(ττ)***
- Covers $\tau_{had}\tau_{had}$ use-case, with large BR
- Discriminant shows **significant top & MJ rejection** within 250 GeV<p_T<1.5 TeV & 50<m_J<200 GeV
  - rej~10^4 for ε=98%
  - rej~10^5 for ε=93%
- Provides us with **unprecedented identification of merged *X*→$\tau_{had}\tau_{had}$** topologies
  - Expect **significant improvements in sensitivity of $\tau_{had}\tau_{had}$ channels to high p_T^H measurements/new resonant mass searches**

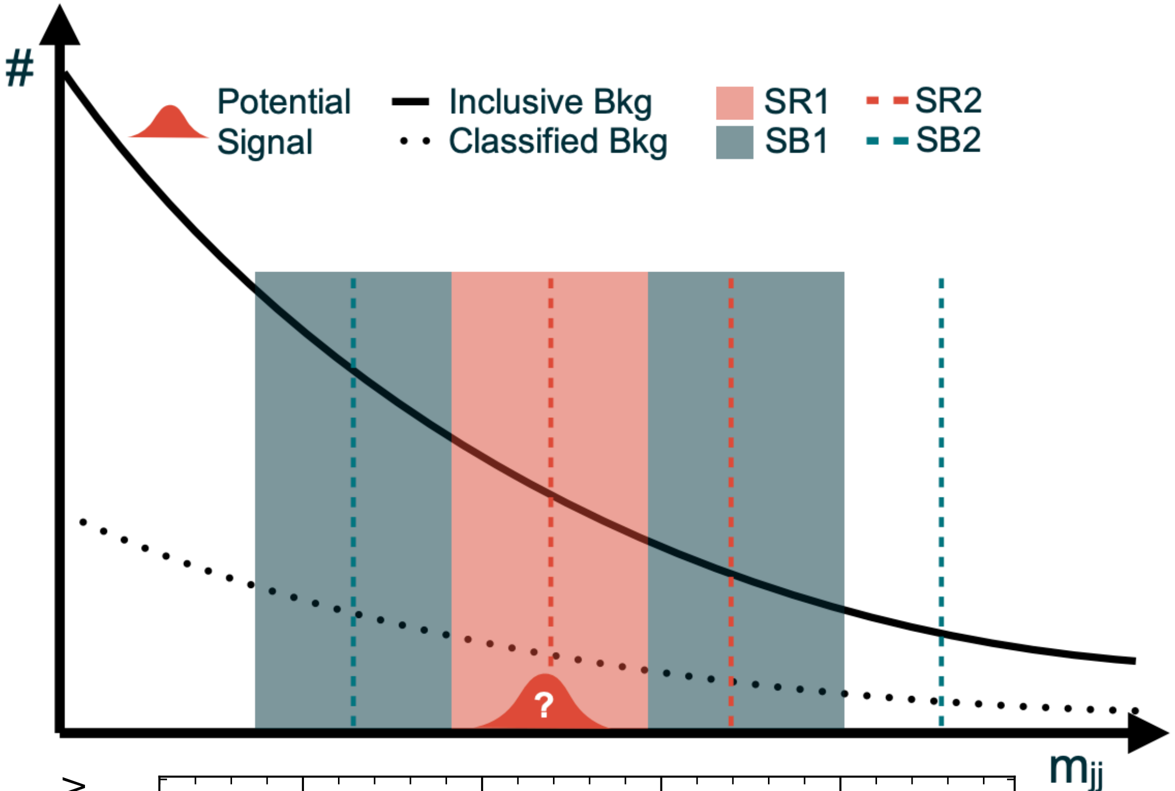


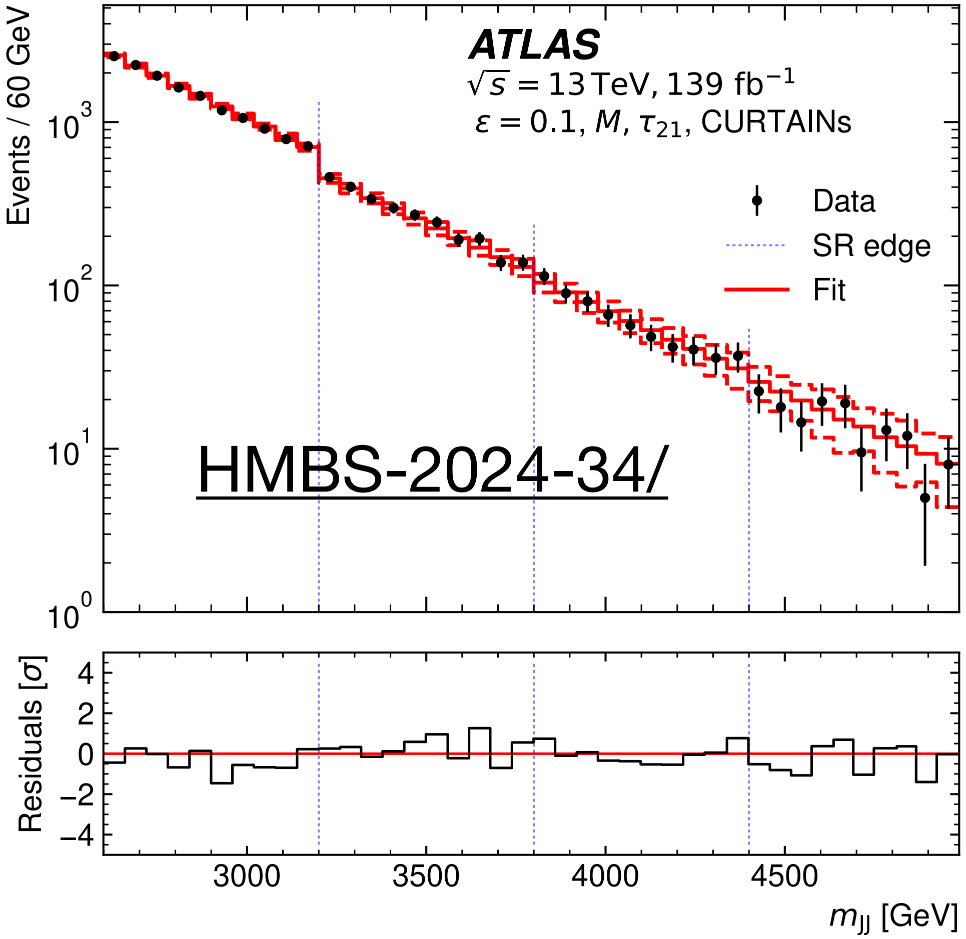


ATL-PHYS-PUB-2025-026

Weakly Supervised Anomaly Detection for Light Di-Jet Resonance

- # BSM models >> # dedicated searches
  - Agnostic searches may fill gap
- Classification Without Labels (**CWoLa**) paradigm & high-dimensional interpolation
- Compare **two NN methods** (SALAD & CURTAINs) **to estimate bkg** from SBs into SR
- **Generate reference Sample** from $p(x|m_{JJ})$ for $m_{JJ} \in$ SR
  - Features x={m_J,τ_{21},τ_{32}}, varying smoothly with m_{JJ}
- **Weakly supervised Anomaly Detection** to search for narrow resonance in di-jet events
  - Model-agnostic analysis *flagging* regions containing BSM signal for thorough studies

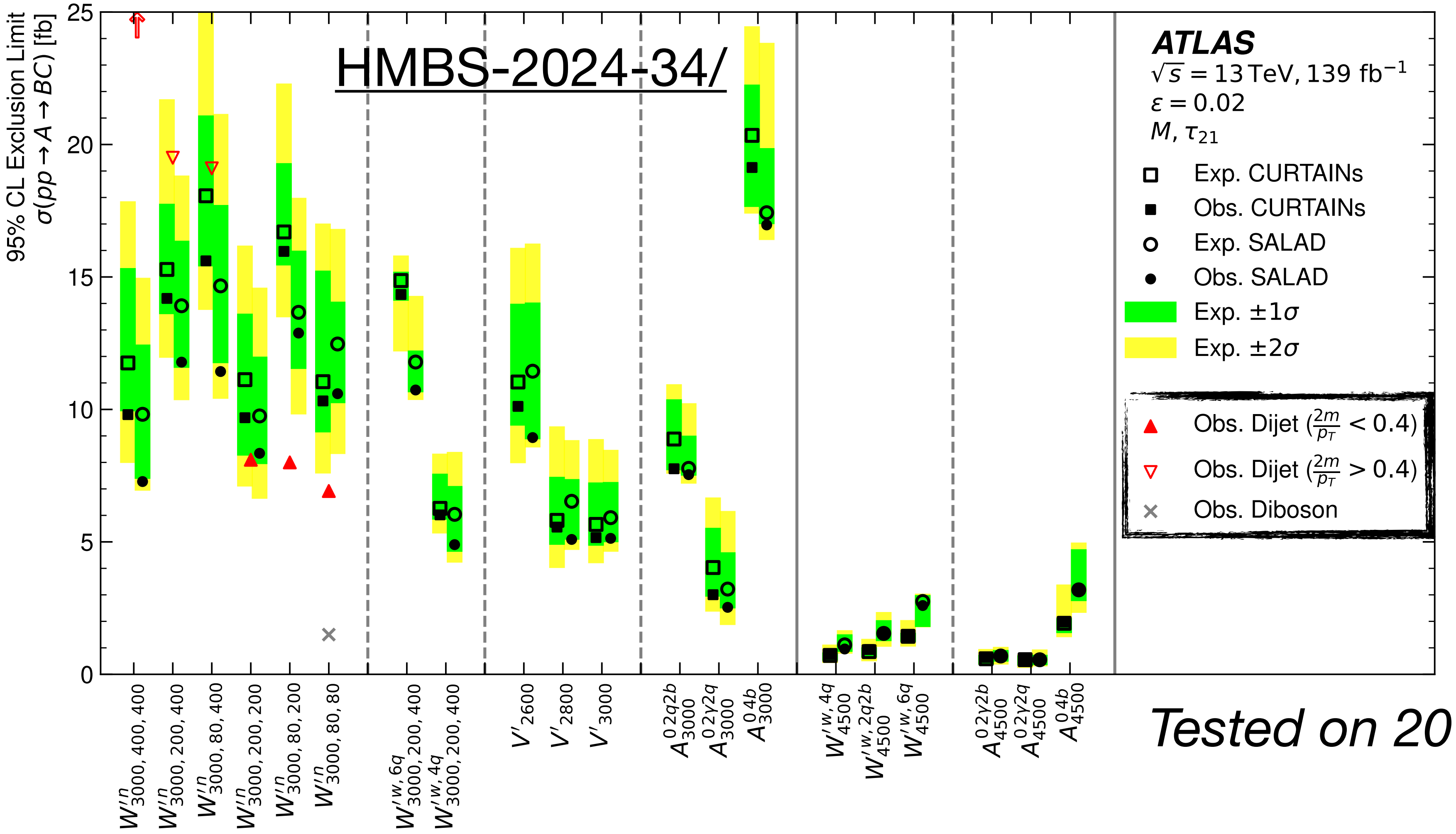




HMBS-2024-34/

Weakly Supervised Anomaly Detection for Light Di-Jet Resonance

- Analysis methodology:
  - **Three subsets of {m_J,τ_{21},τ_{32}}**: outputs of bkg estimate & inputs to classifier
  - Train Weakly Supervised classifier SR **data** vs **reference**: selection to increase S/B
  - **Extended bump hunt** with 6 (3x2 jets) features
  - **All analysis steps (Weakly Supervised class incl.) repeated for 7 diff SR/SBs**
- m_{JJ} spectrum between 2.6 & 5.0 TeV in steps of 300 GeV → no excess → **limits**



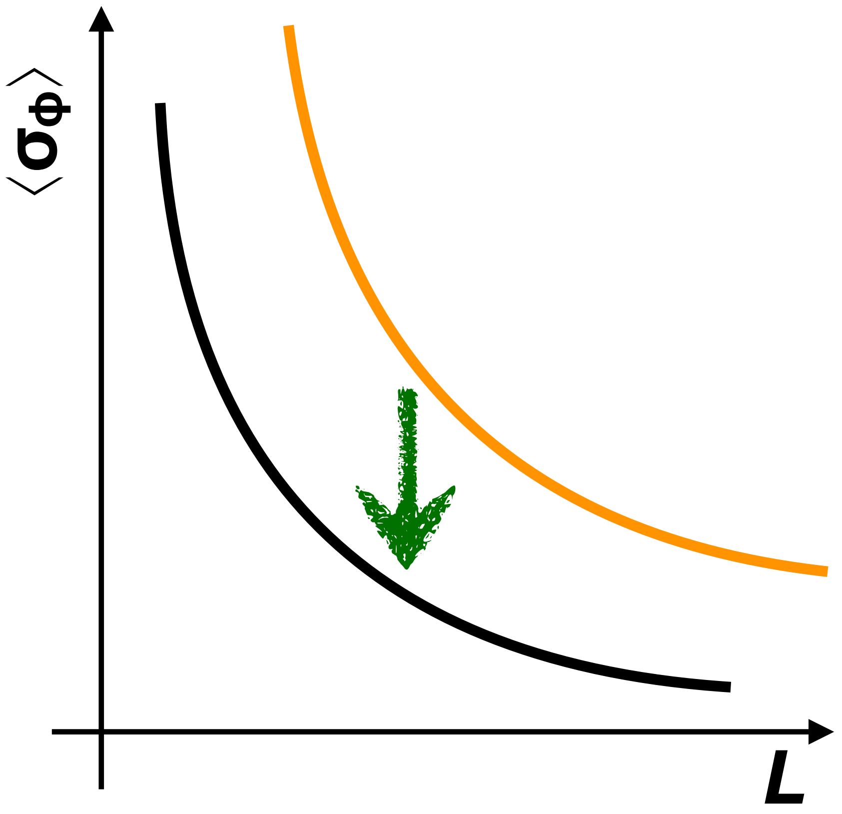
HMBS-2024-34/

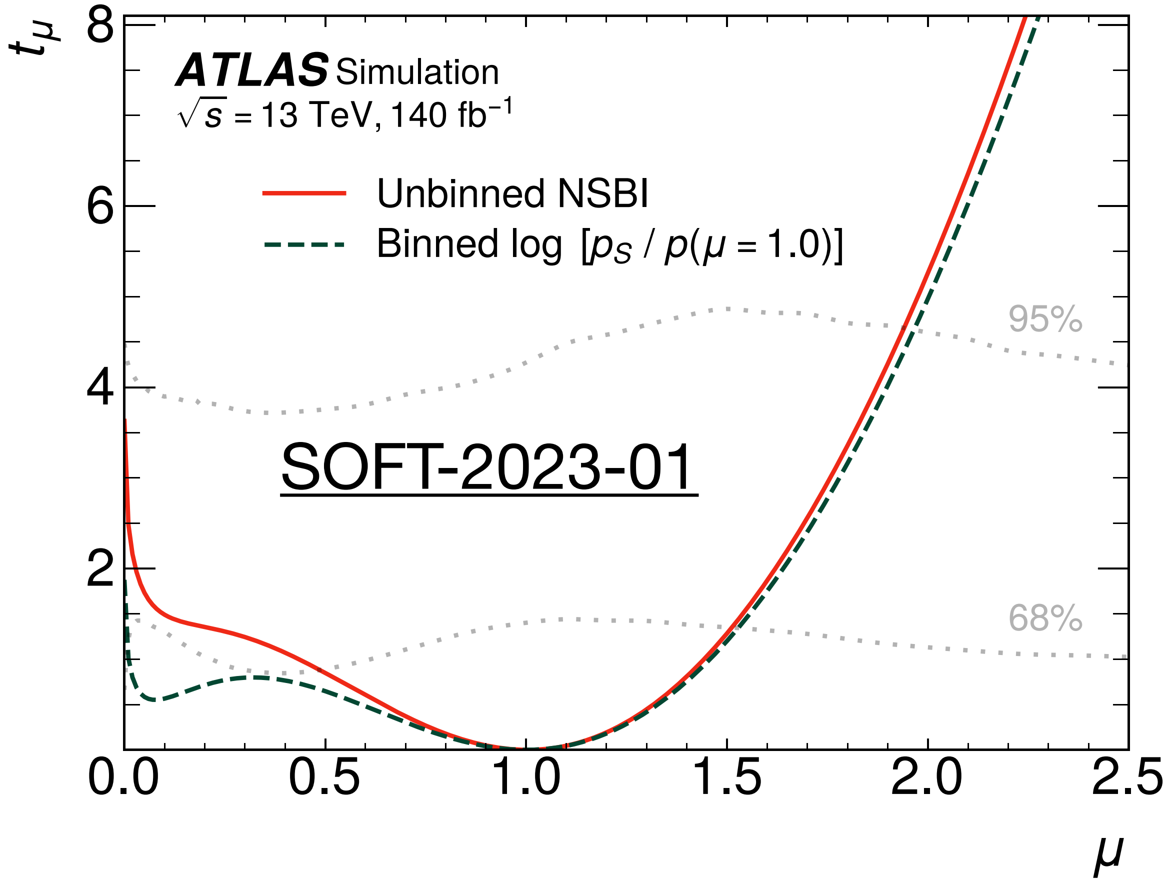
Previous di-jet/VV searches

Tested on 20 signal models

Neural Simulation-Based Inference (NSBI)

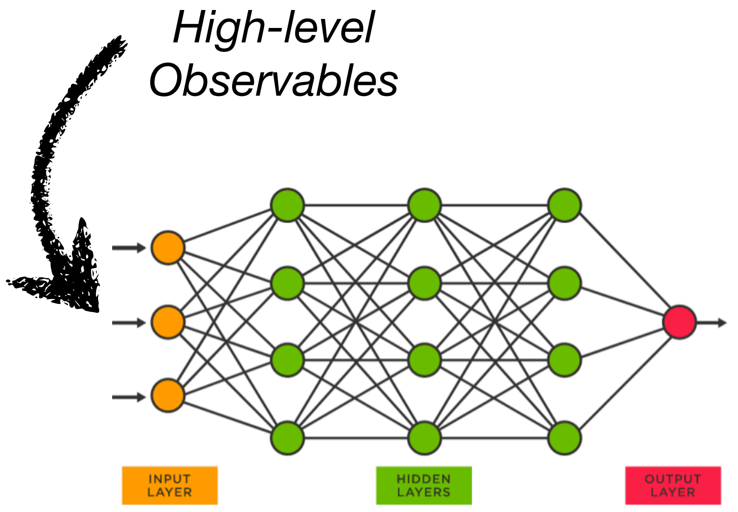
- Set of observations {x} sensitive to a parameter ϕ
- Given a simulator ϕ → p({x}) & a prior p(ϕ), posterior over ϕ: {x} → p(ϕ)
- **Variance of posterior scaling as ~1/|{x}| ~ 1/*L***
- A lossy function applied to {x} (e.g. binning) brings loss in constraining power
  - Same 1/*L* scaling, **constant offset generally worse**
- Better per-event use of x brings improved results
- **Neural Simulation-Based Inference (NSBI)** gives arbitrarily good approximation of true {x} → p(ϕ) → best sensitivity of analysis strategies based on {x}
- ATLAS building expertise for “real-life” applications





NSBI Application: Off-Shell H→ZZ* (Γ_H)

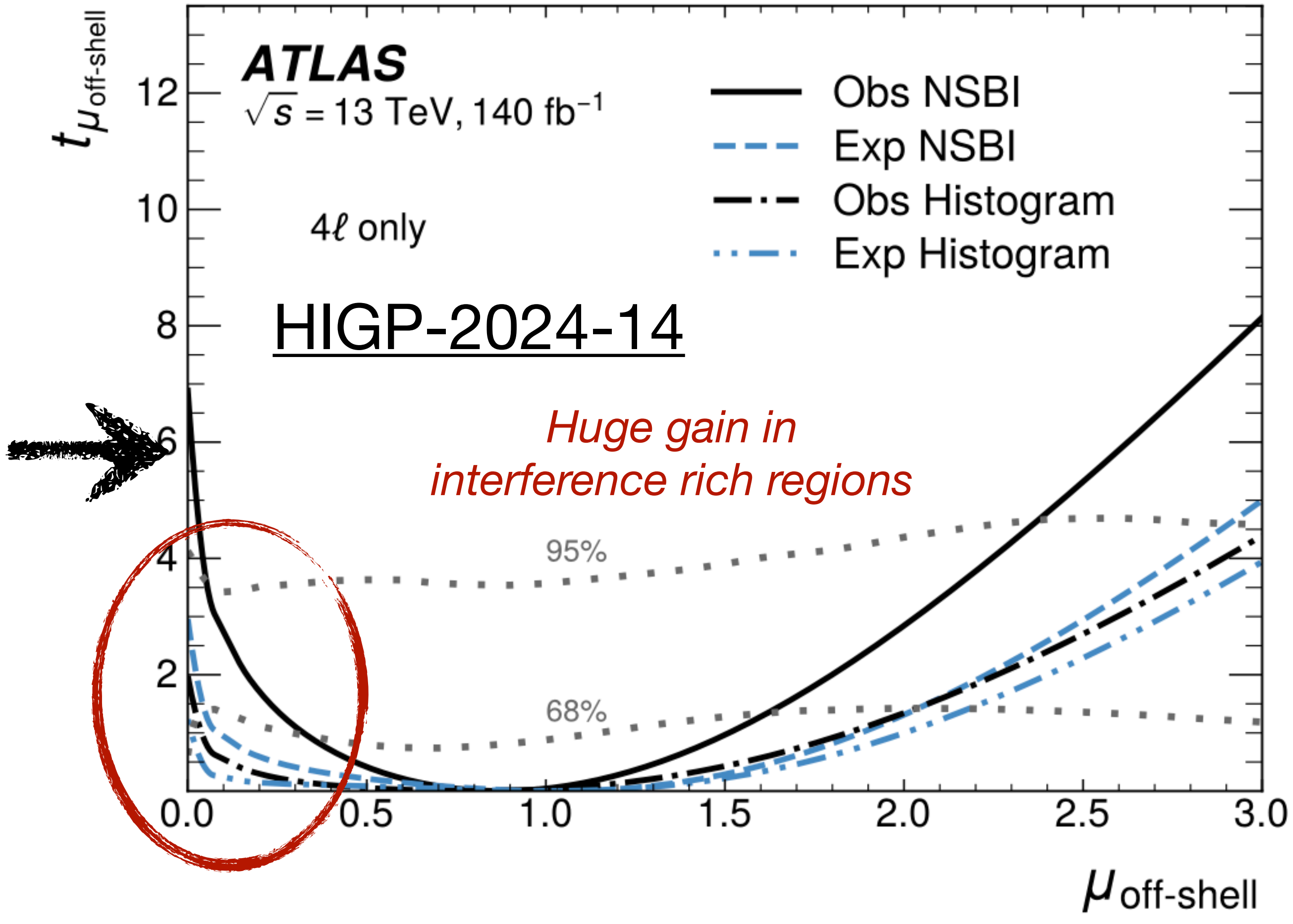
- Higgs boson production in the H→ZZ*→4ℓ decay channel on full Run-2 dataset (140 fb^{-1})
- Data analysed with NSBI strategy
  - **NNs used to estimate per-event contribution to likelihood ratio** (LLR) between different hypotheses: maximal sensitivity throughout parameter space
- Combined with H→ZZ*→2ℓ2ν decay channel
- **Evidence** for off-shell Higgs production
- **~13% relative improvement in 95% CL upper limit on Γ_H,** as compared to standard histogram analysis



Ensemble of NNs trained to minimise Binary Cross Entropy loss

$$-2\sum_{events}\log\left[\frac{p(x\,|\,\mu,\hat{\hat{\alpha}})}{p(x\,|\,\hat{\mu},\hat{\alpha})}\right]$$

Event-by-event parameterised LLR



Conclusion

- **Many new methods & techniques** developed within **ATLAS** Experiment
  - Some already helped us push our knowledge boundaries, some will soon!
- Discussed most recent developments & their applications to “real” physics cases
  - **Wide use of ML**: transformers, DNN, CWoLA, weakly supervised NN…
  - Dedicated object reconstruction & identification for **extension of analysis search/reach range**
  - Optimisation of analyses with **complex & mixed topologies**
  - ML-driven improvements in **better per-even use of observables** to extract physics results
- Stay tuned for more analysis methods to come!

BACKUP

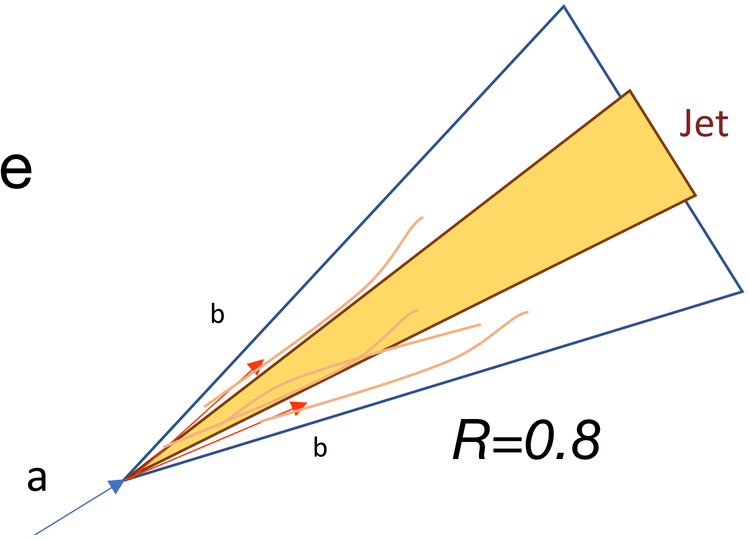
DeXTer: Specialised Flavour Tagging

DeXTer: Specialised Flavour Tagging

ATL-PHYS-PUB-2022-042

- DeXTer: capture particles from fragmentation or decay of heavy-flavoured hadrons from **multiple partons**
  - **Extended collection of tracks** to a reconstructed jet by clustering all PFlow jets & ID tracks matched to jets through ghost-association
  - **Re-clustering with anti-kt algorithm with R=0.8**
- **Combines TC-LVT**, to identify collection of tracks that may have >0 secondary vertices, & Multiple Secondary Vertex Finder (**MSVF**), to build all two-track proto-vertices consistent with (non-background) displaced tracks
- **Two track sub-jets** are reconstructed with **exclusive-kt algorithm**
  - Define fly direction of 2 sub-jets from *bb*
- **Jet, track & SV kinematics** fed as input to DeXTer architecture (see next slide)

Jet re-clustering to get isolated R=0.4 jets w/ surrounding tracks

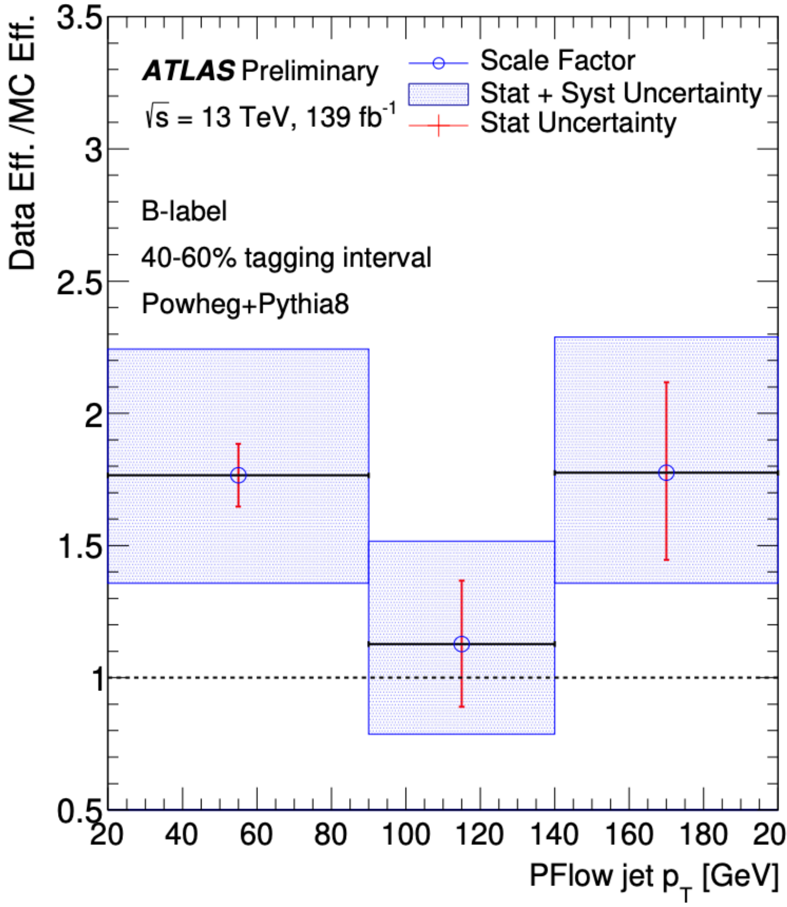


DeXTer: Specialised Flavour Tagging

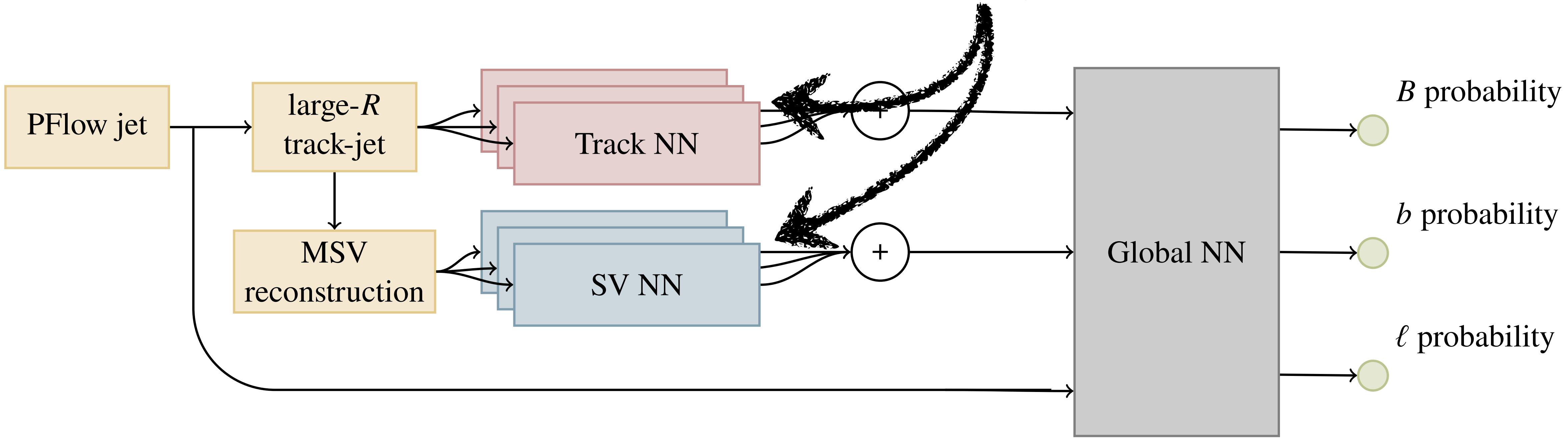
ATL-PHYS-PUB-2022-042

- Two **feature-extracting feed-forward NNs** used for track & SV observables
- **Final feed-forward global NN** used to interpret each output as a probability for each flavour
- Feature NNs: 2 hidden layers with 100 neurons & output layer with 128 features
- 256 features input to global NN with 3 hidden layers
- Only calibration sample available is $g \to bb$, B events mix of $g \to bb$ & $a \to bb$
    - **Colour-charge adversarial NN** in back-propagation gradients
- **Calibration on Z & tt events**

Data scale factors for B identification



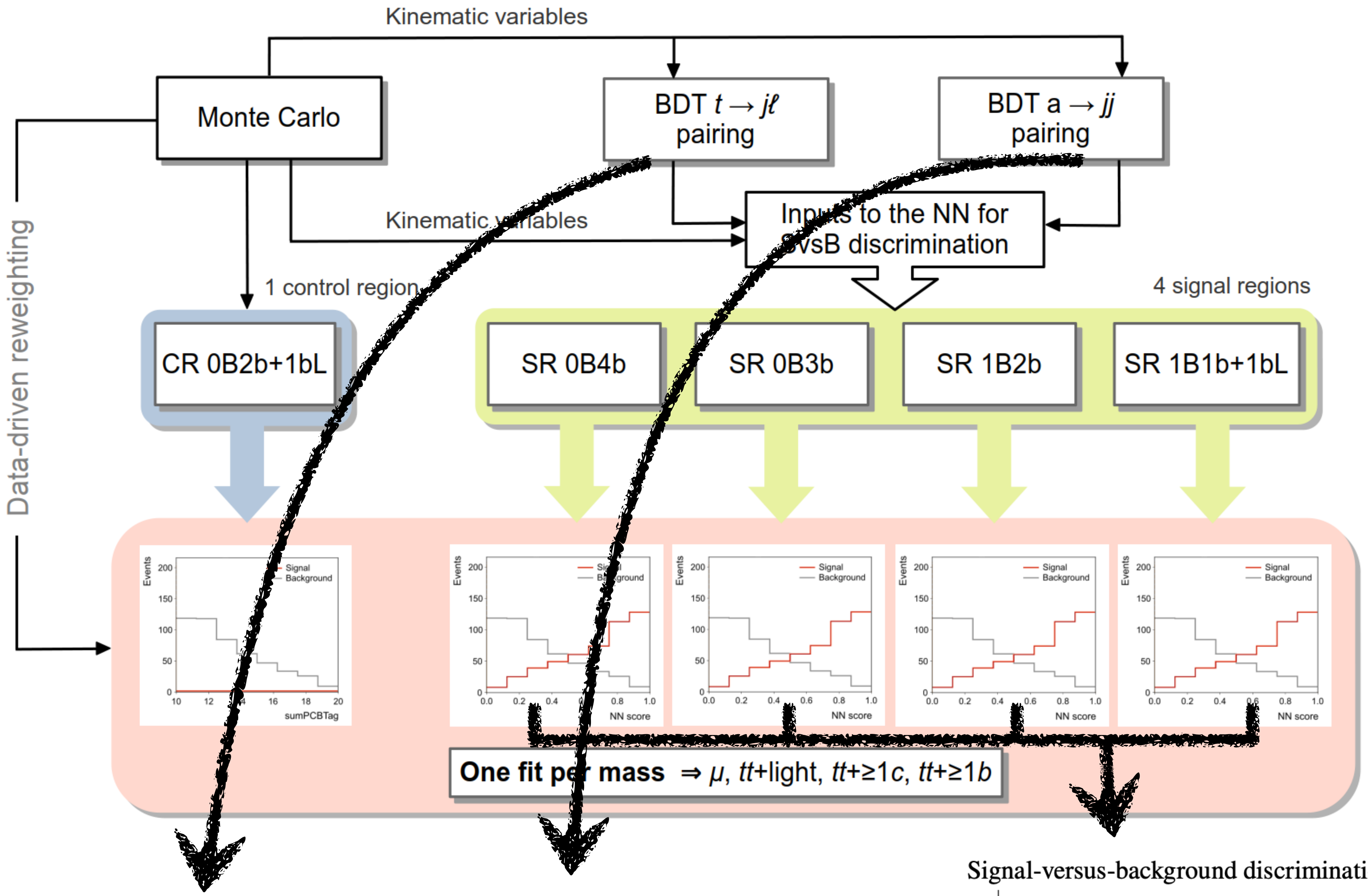
Adversarial backpropagation gradient minimises diff in singlets vs. octets



Search for New Pseudoscalar: $tt/tW+a(\rightarrow bb)$

Search for New Pseudoscalar: *tt/tW+a(→bb)*

HMBS-2024-46



Top quark/antiquark reconstruction BDT

Object	Variables
Full event	N_{jets}, $N_{b\text{-jets}}$
Lepton (tag, aux.)	p_{T}, η
Jet (tag, aux.)	p_{T}, η, PC b-tag, jet index
lj pair (tag, aux.)	m, p_{T}, η, ΔR
$t\bar{t}$ pair	m, p_{T}, η, ΔR, $\Delta\phi$
jj pair	ΔR

Pseudoscalar reconstruction BDT

Object	Variables
Full event	N_{jets}, $N_{b\text{-jets}}$, `sumPCBTag`
Jet (1st, 2nd)	p_{T}, η, PC b-tag, jet index
jj pair	m, p_{T}, η, E, ϕ, ΔR

Signal-versus-background discrimination NN

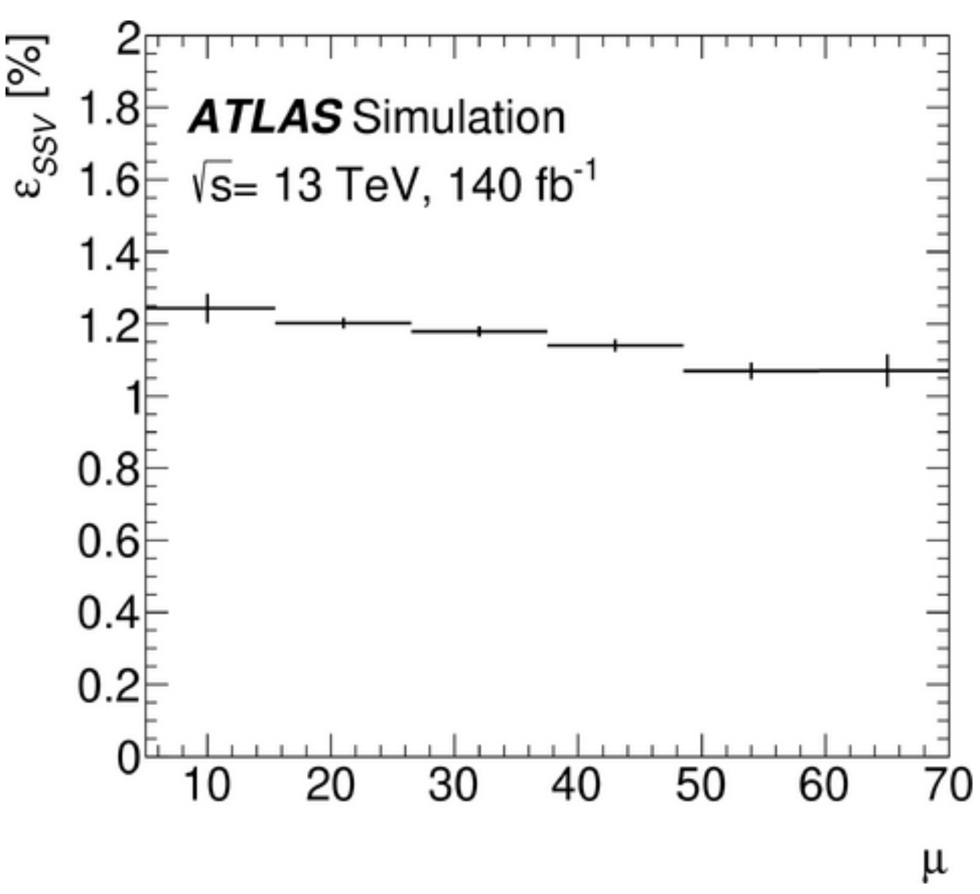
Object	Variables
Full event	N_{jets}, $H_{\text{T}}^{\text{jets}}$, $E_{\text{T}}^{\text{miss}}$
BDT $t \to lj$	Score, p_{T}^{lj}, ΔR_{lj}, $\Delta\eta_{lj}$, $\Delta\phi_{lj}$, jet index
BDT $a \to jj$	Score, p_{T}^{jj}, η_{jj}, m_{jj}, ΔR_{jj}, $\Delta\eta_{jj}$, $\Delta\phi_{jj}$, jet index
Leptons	ΔR_{ll}, $\Delta\eta_{ll}$, $\Delta\phi_{ll}$, $\Delta\phi_{E_{\text{T}}^{\text{miss}},l}$, $\Delta R_{ll,bb}$, $\Delta R_{ll,B}$, $\Delta R_{ll,b}$
Large-R jets	p_{T}, η, m, ΔR_{Bb}, $\Delta\phi_{E_{\text{T}}^{\text{miss}},B}$
Small-R jets	p_{T}^{bb}, m_{bb}, m_{bbb}, m_{bbbb}, ΔR_{bb}, $\Delta\eta_{bb}$, $\Delta\phi_{bb}$, $\Delta\phi_{E_{\text{T}}^{\text{miss}},b}$ p_{T}, η, PC b-tag

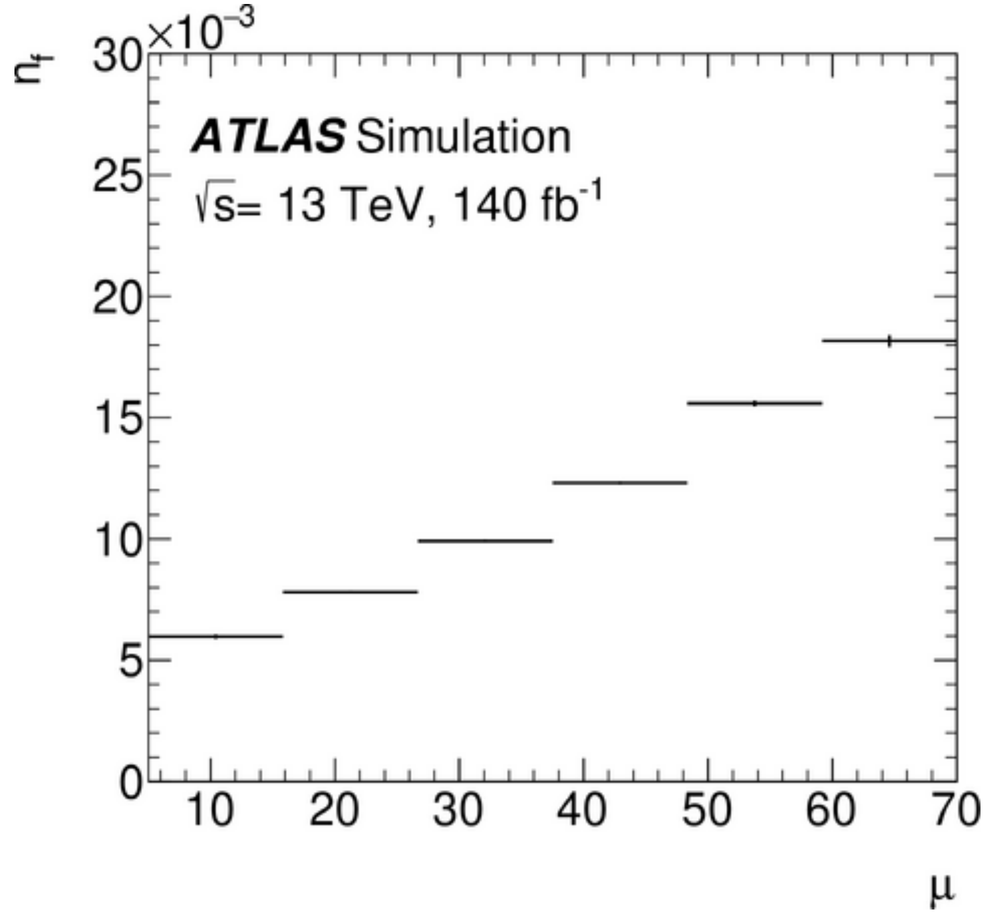
TC-LVT: “Jet-less” Flavour Tagging

TC-LVT: “Jet-less” Flavour Tagging

Seed track	Cluster tracks	Vertex
$\lvert d_0/\sigma(d_0)\rvert > 0.5$	$\lvert d_0/\sigma(d_0)\rvert > 1.5$	$600\,\text{MeV} < m_{\text{vtx}} < 6\,\text{GeV}$
$p_{\text{T}} > 1.5$ GeV	$\Delta R^{\text{track}}_{\text{seed}} < 0.75$	$p_{\text{T}}^{\text{vtx}} > 3$ GeV
	$d^{\text{track}}_{\text{seed}} < 0.25$ mm	

- TC-LVT inputs: seed, cluster tracks & vertices
  - Seed tracks -> cluster of tracks built around seed tracks by adding **additional high-displacement tracks**
  - For each identified cluster, SSVF executed on all tracks within ΔR=0.4 of vector sum of momenta of all tracks in cluster -> tracks inputs to **vertex fitter to get one SV**
- **Stable SSV efficiency** & non-negligible dependence of average number of fake SSVs per event as a function of average number of interactions per bunch crossing (μ)



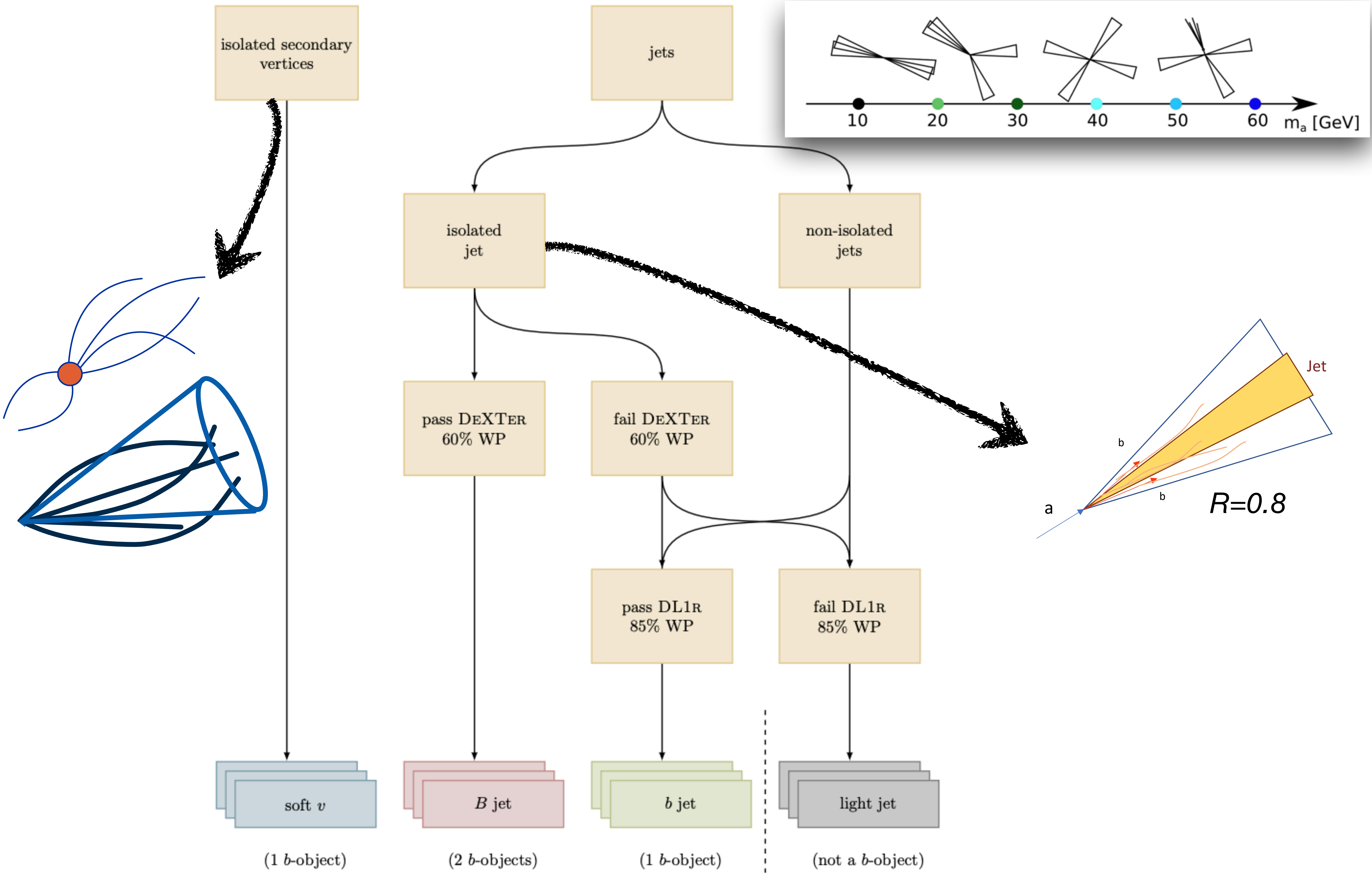


FTAG-2023-02

Search for $Z(\rightarrow \ell\ell/\nu\nu)H(\rightarrow 4b/6b)$

Search for *Z(→ℓℓ/νν)H(→4b/6b)*

HMBS-2024-26



isolated secondary vertices

jets

isolated jet

non-isolated jets

pass DeXTer 60% WP

fail DeXTer 60% WP

pass DL1r 85% WP

fail DL1r 85% WP

soft *v*

(1 *b*-object)

B jet

(2 *b*-objects)

b jet

(1 *b*-object)

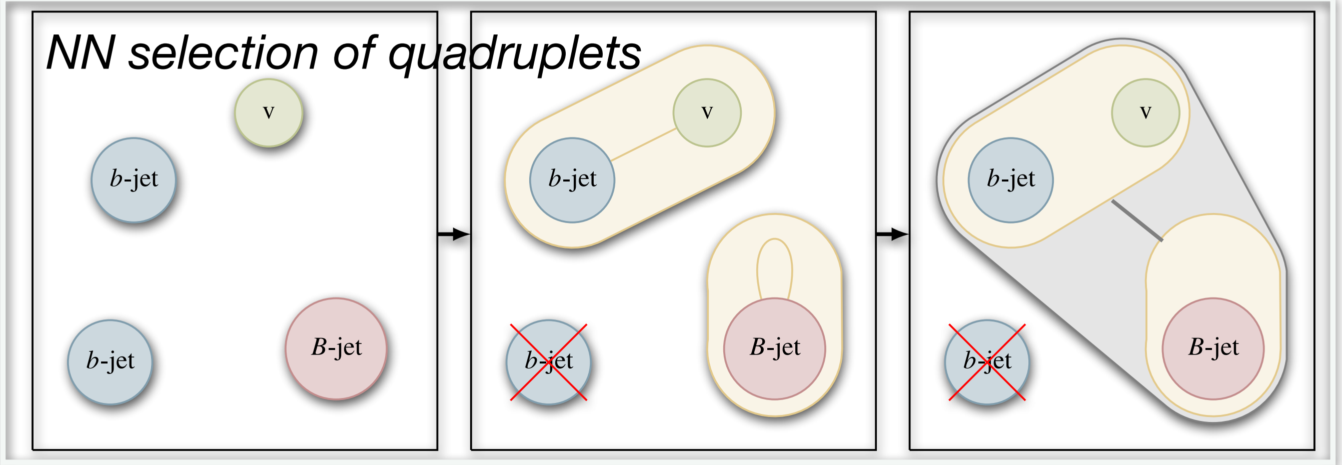
light jet

(not a *b*-object)

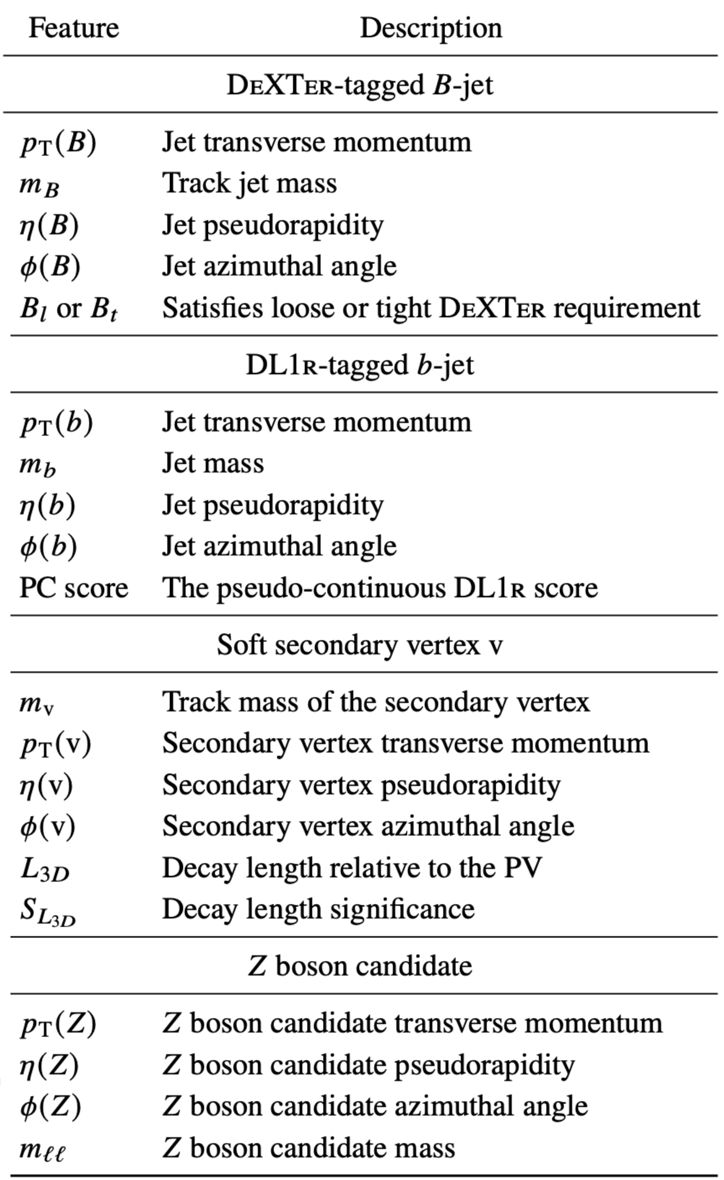
Search for Z(→ℓℓ/νν)H(→4b/6b)

HMBS-2024-26

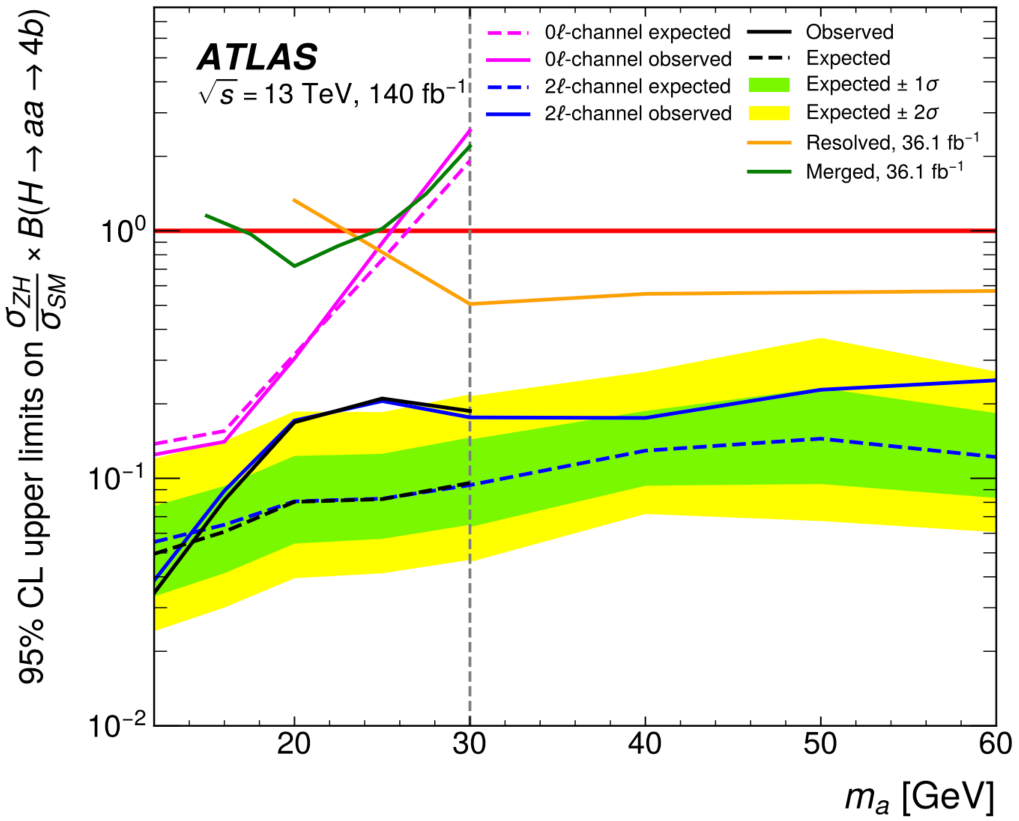
- Quadruplet selection NN in 2ℓ 4b
  - Select best combination of four b-objects to reconstruct decay chain $H \to 2a/a_1a_2 \to 4b$
  - Different encoder NNs for different b-objects
  - Fully connected encoders for b-jets, B-jets, & soft vertices
  - Encoded info used in pair of deep NNs, ensuring invariance under b & bb permutations
- NN built from up to 5 objects with ≤1 soft vertices
- Accuracy of associating b-partons to b-objects ranges from 60% to 98%



Quadruplets NN inputs



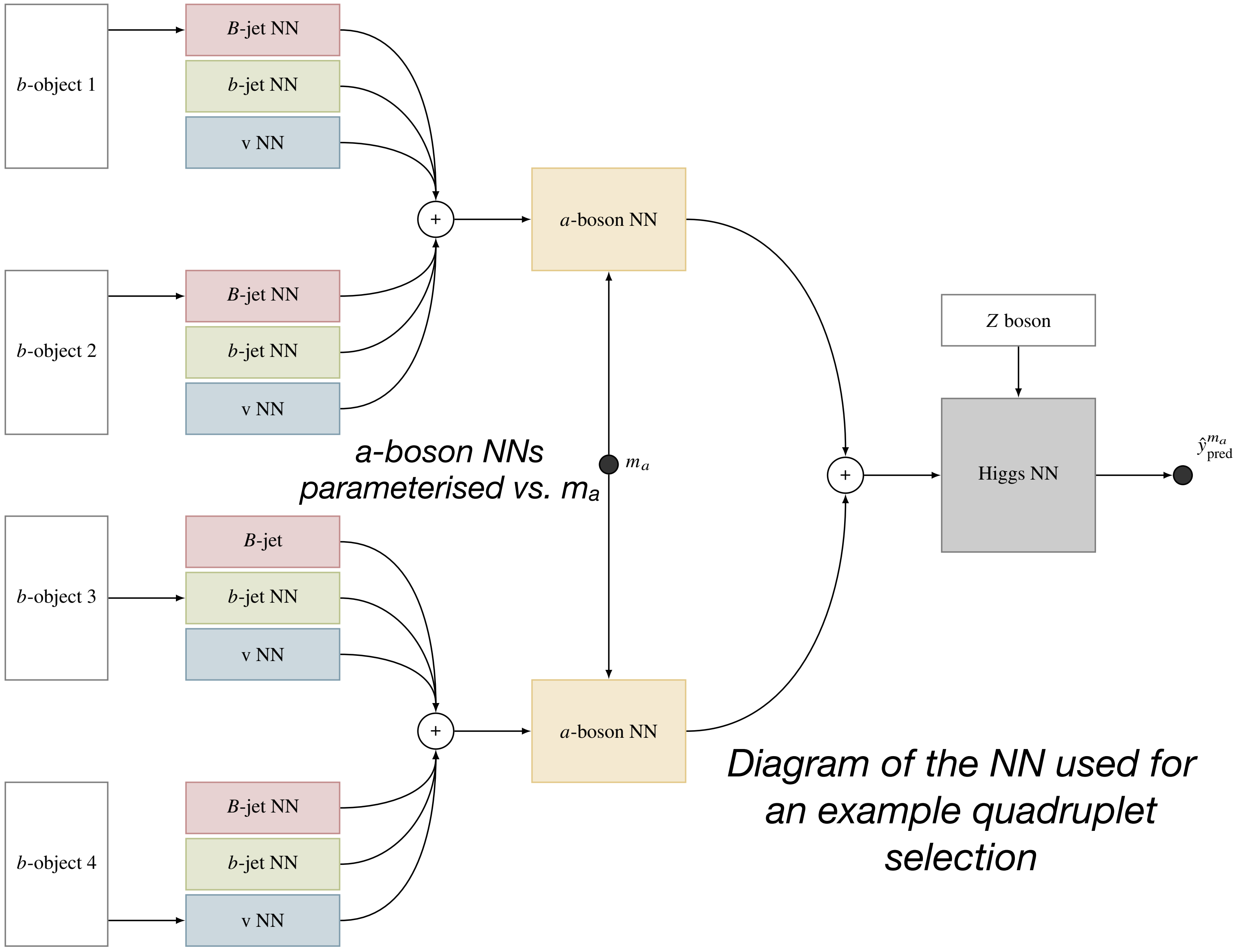
Feature	Description
DeXTer-tagged B-jet	
$p_T(B)$	Jet transverse momentum
m_B	Track jet mass
$\eta(B)$	Jet pseudorapidity
$\phi(B)$	Jet azimuthal angle
B_l or B_t	Satisfies loose or tight DeXTer requirement
DL1r-tagged b-jet	
$p_T(b)$	Jet transverse momentum
m_b	Jet mass
$\eta(b)$	Jet pseudorapidity
$\phi(b)$	Jet azimuthal angle
PC score	The pseudo-continuous DL1r score
Soft secondary vertex v	
m_v	Track mass of the secondary vertex
$p_T(v)$	Secondary vertex transverse momentum
$\eta(v)$	Secondary vertex pseudorapidity
$\phi(v)$	Secondary vertex azimuthal angle
L_{3D}	Decay length relative to the PV
$S_{L_{3D}}$	Decay length significance
Z boson candidate	
$p_T(Z)$	Z boson candidate transverse momentum
$\eta(Z)$	Z boson candidate pseudorapidity
$\phi(Z)$	Z boson candidate azimuthal angle
$m_{\ell\ell}$	Z boson candidate mass



Search for $Z(\to\ell\ell/\nu\nu)H(\to 4b/6b)$

HMBS-2024-26

- Quadruplet selection NN in 2ℓ 4b
  - Select best combination of four b-objects to reconstruct decay chain $H\to 2a/a_1a_2\to 4b$
  - Different encoder NNs for different b-objects
  - Fully connected encoders for b-jets, B-jets, & soft vertices
  - Encoded info used in pair of deep NNs, ensuring invariance under b & bb permutations
- NN built from up to 5 objects with ≤1 soft vertices
- Accuracy of associating b-partons to b-objects ranges from 60% to 98%



a-boson NNs parameterised vs. m_a

Diagram of the NN used for an example quadruplet selection

GN2X *H(bb/cc)* Tagging at High Momentum

GN2X *H(bb/cc)* at High Momentum

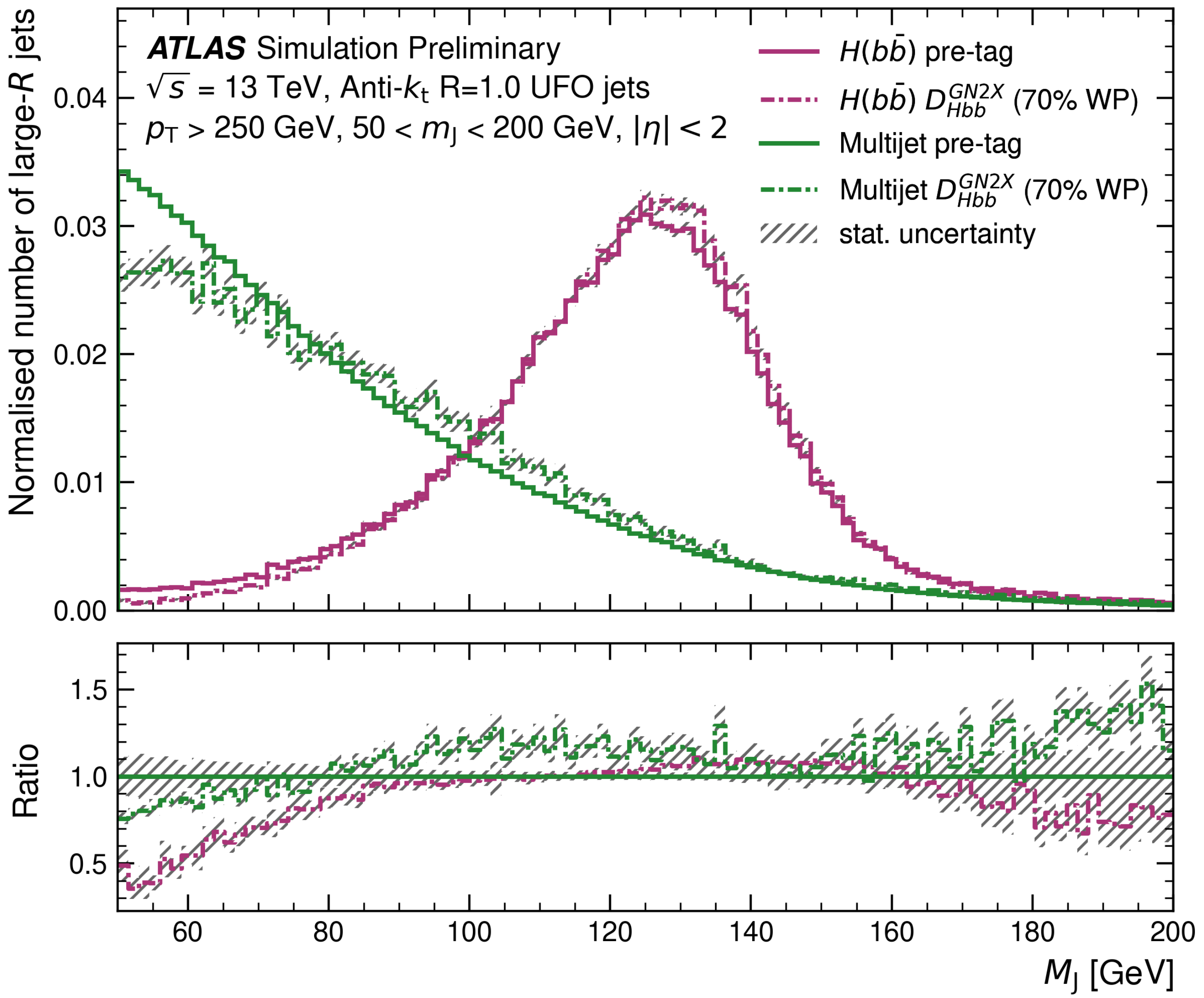
ATL-PHYS-PUB-2023-021

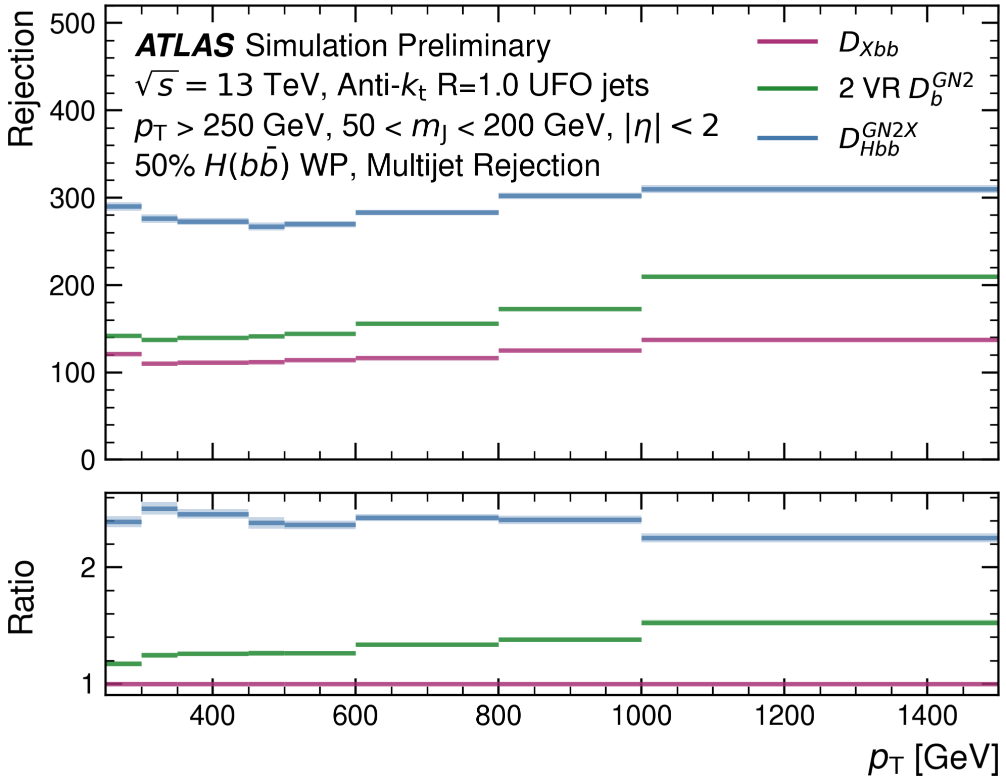
- Inputs: up to 100 tracks associated with jet
  - Two models with extra info: either kinematic and b-tagging info of variable-R track jets or Unified Flow Object (UFO) constituents
- Transformer network architecture
- Jet & track inputs concatenated: combined jet-track sequence vectors fed into per-track initialiser network
- Track representations fed into Transformer Encoder
  - 6 encoders with 4 attention heads
- Output representation of each track combined to form global representation of jet

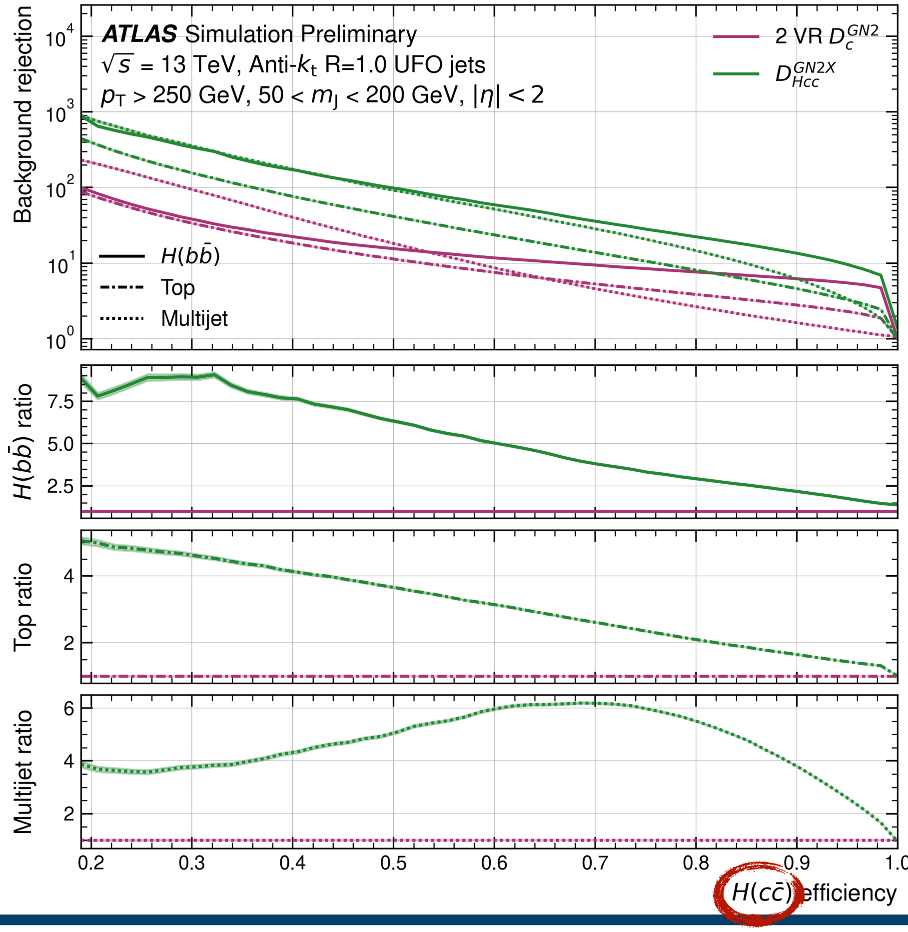
Jet Input	**Description**
p_T	Large-R jet transverse momentum
η	Signed large-R jet pseudorapidity
mass	Large-R jet mass
Track Input	**Description**
q/p	Track charge divided by momentum (measure of curvature)
dη	Pseudorapidity of track relative to the large-R jet η
dϕ	Azimuthal angle of the track, relative to the large-R jet ϕ
d_0	Closest distance from track to primary vertex (PV) in the transverse plane
$z_0 \sin\theta$	Closest distance from track to PV in the longitudinal plane
$\sigma(q/p)$	Uncertainty on q/p
$\sigma(\theta)$	Uncertainty on track polar angle θ
$\sigma(\phi)$	Uncertainty on track azimuthal angle ϕ
$s(d_0)$	Lifetime signed transverse IP significance
$s(z_0 \sin\theta)$	Lifetime signed longitudinal IP significance
nPixHits	Number of pixel hits
nSCTHits	Number of SCT hits
nIBLHits	Number of IBL hits
nBLHits	Number of B-layer hits
nIBLShared	Number of shared IBL hits
nIBLSplit	Number of split IBL hits
nPixShared	Number of shared pixel hits
nPixSplit	Number of split pixel hits
nSCTShared	Number of shared SCT hits
subjetIndex	Integer label of which subjet track is associated to (GN2X + Subjets only)
Subjet Input	**Description** (Used only in GN2X + Subjets)
p_T	Subjet transverse momentum
η	Subjet signed pseudorapidity
mass	Subjet mass
energy	Subjet energy
dη	Pseudorapidity of subjet relative to the large-R jet η
dϕ	Azimuthal angle of subjet relative to the large-R jet ϕ
GN2 p_b	b-jet probability of subjet tagged using GN2
GN2 p_c	c-jet probability of subjet tagged using GN2
GN2 p_u	light flavour jet probability of subjet tagged using GN2
Flow Input	**Description** (Used only in GN2X + Flow)
p_T	Transverse momentum of flow constituent
energy	Energy of flow constituent
dη	Pseudorapidity of flow constituent relative to the large-R jet η
dϕ	Azimuthal angle of flow constituent relative to the large-R jet ϕ

GN2X *H(bb/cc)* at High Momentum

ATL-PHYS-PUB-2023-021







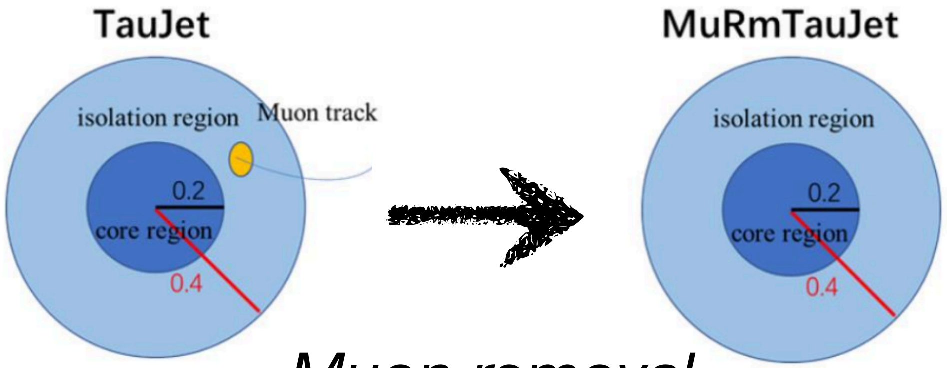
$$\underset{\textbf{(cc)}}{D_{\mathrm{Hbb}}^{\mathrm{GN2X}}} = \ln\left(\frac{p_{\mathrm{Hbb}}}{\underset{\textbf{(bb)}}{f_{\mathrm{Hcc}}} \cdot \underset{\textbf{(bb)}}{p_{\mathrm{Hcc}}} + f_{\mathrm{top}} \cdot p_{\mathrm{top}} + \left(1 - \underset{\textbf{(bb)}}{f_{\mathrm{Hcc}}} - f_{\mathrm{top}}\right) \cdot p_{\mathrm{QCD}}}\right)$$

τ Reco with μ-τ Removal in $\tau\tau \to \mu\nu_\mu\nu_\tau$ had+ν_τ

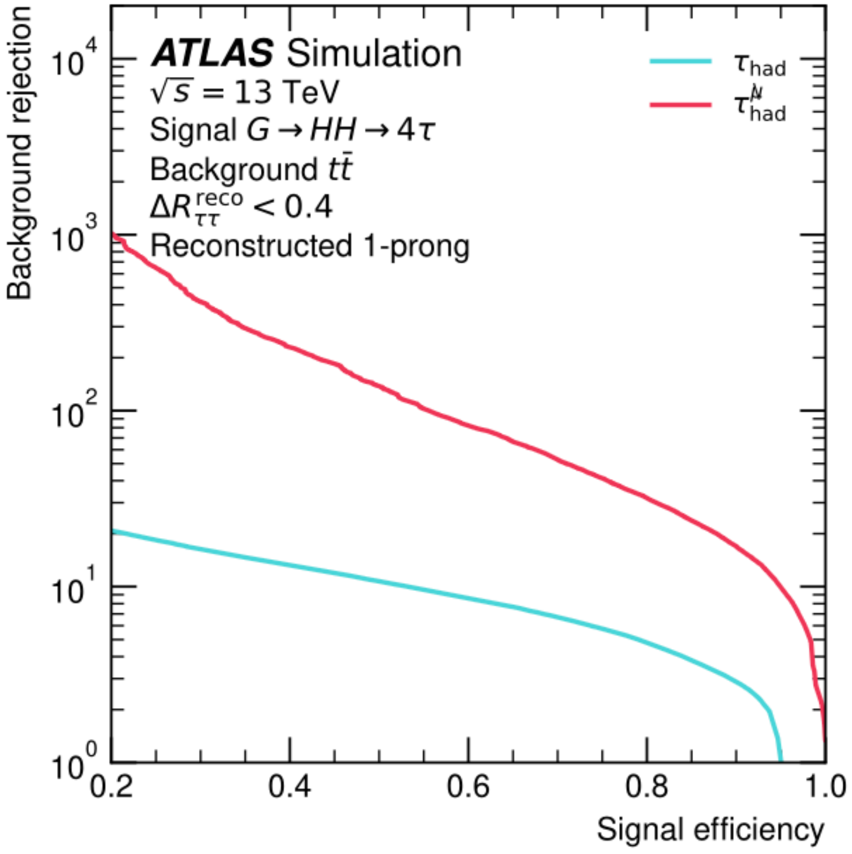
τ Reco with μ-τ Removal in $\tau\tau \rightarrow \mu\nu_{\mu}\nu_{\tau}$ had+ν_{τ}

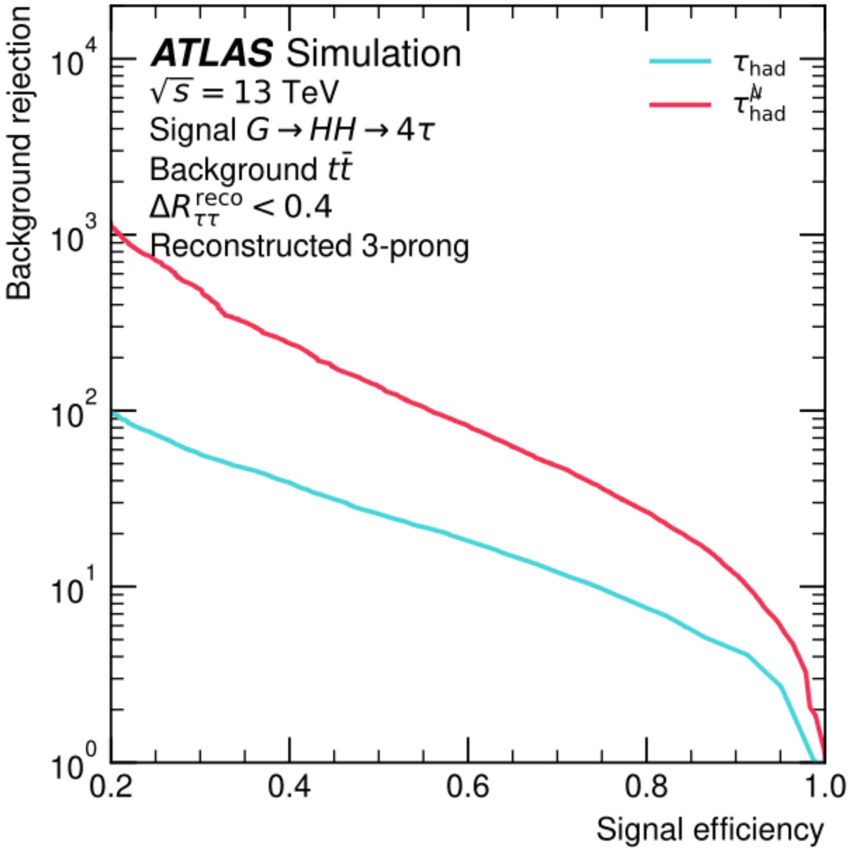
TAUP-2023-02

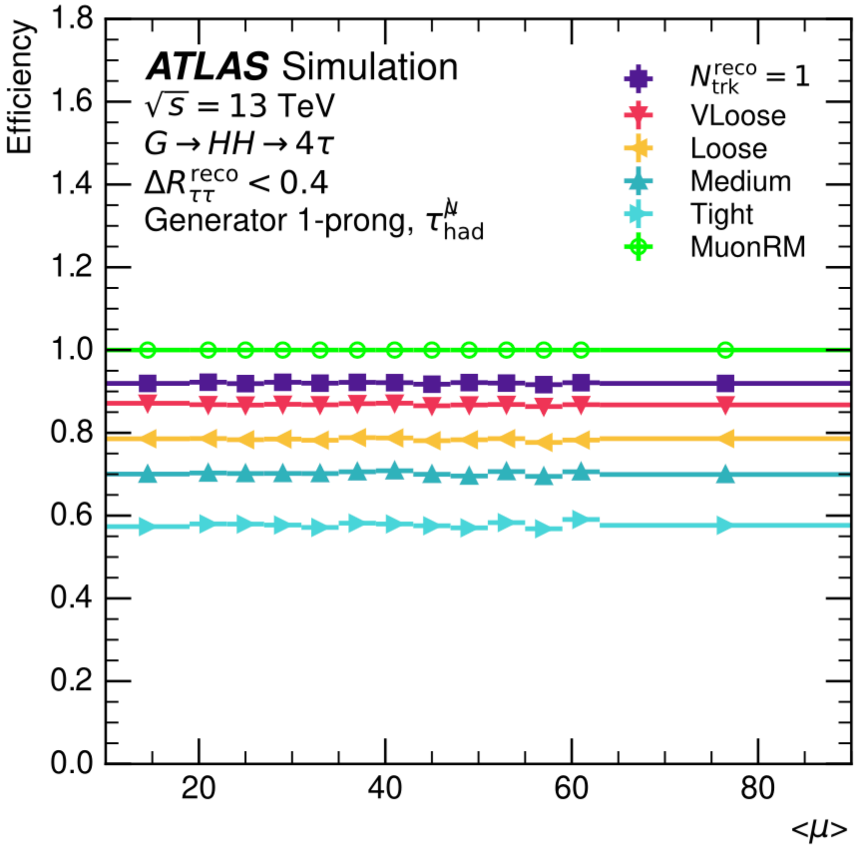
- Stability of performance vs. average pile-up
- An order of magnitude gain in rejection of “fake” τ



Muon removal

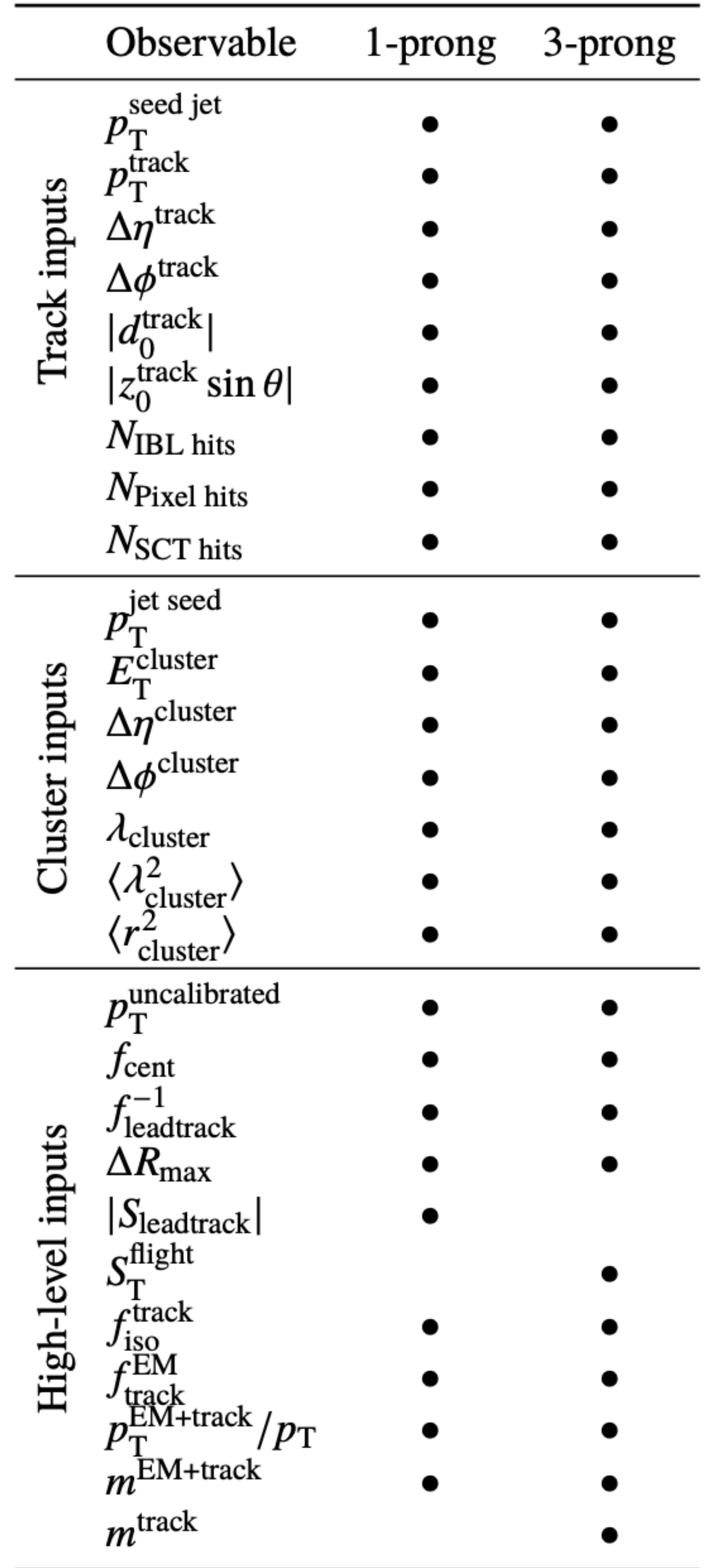






Inputs to τ Identification RNN

TAUP-2023-02

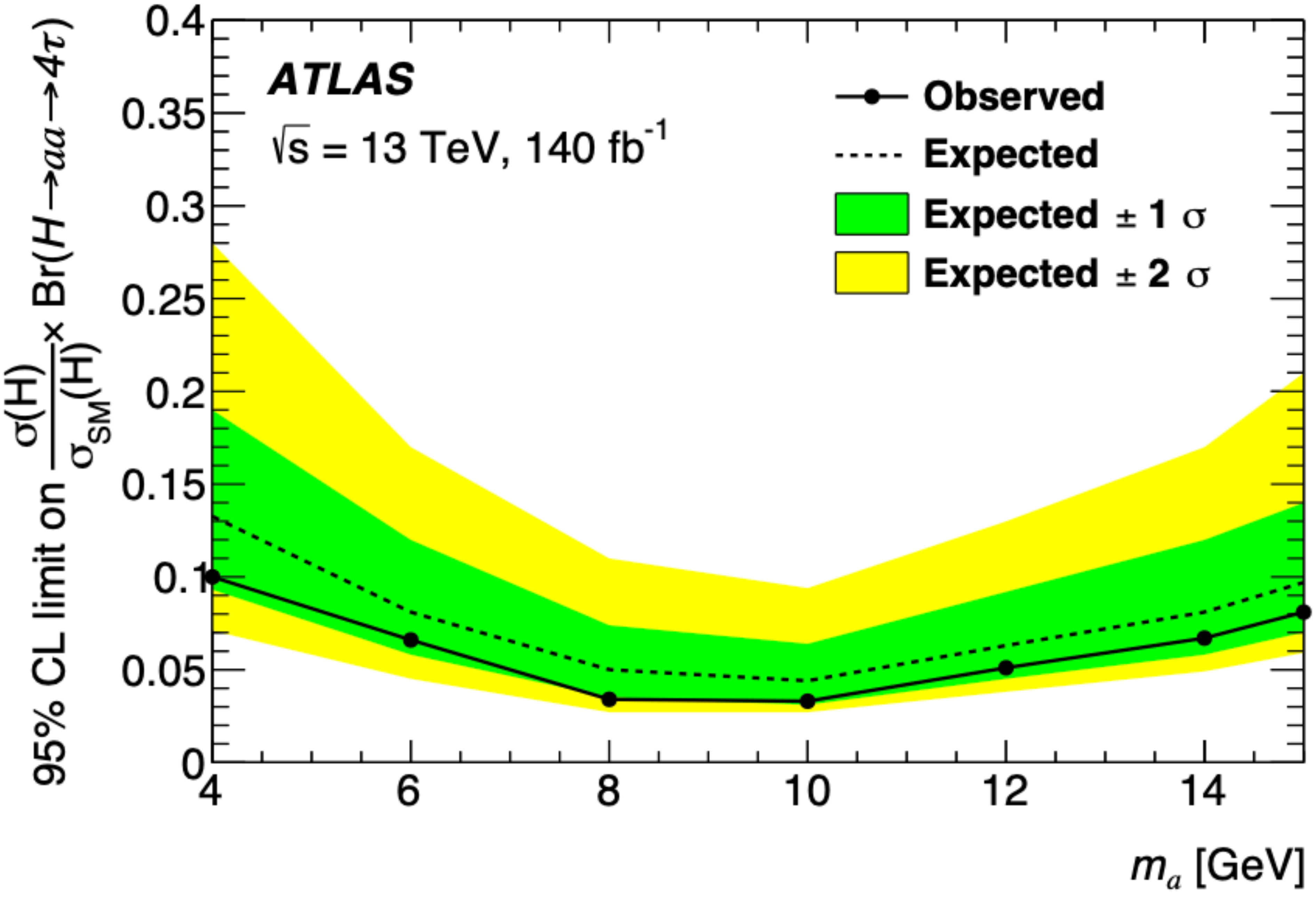


	Observable	1-prong	3-prong
Track inputs	$p_T^{\text{seed jet}}$	●	●
	p_T^{track}	●	●
	$\Delta\eta^{\text{track}}$	●	●
	$\Delta\phi^{\text{track}}$	●	●
	$\lvert d_0^{\text{track}} \rvert$	●	●
	$\lvert z_0^{\text{track}} \sin\theta \rvert$	●	●
	$N_{\text{IBL hits}}$	●	●
	$N_{\text{Pixel hits}}$	●	●
	$N_{\text{SCT hits}}$	●	●
Cluster inputs	$p_T^{\text{jet seed}}$	●	●
	E_T^{cluster}	●	●
	$\Delta\eta^{\text{cluster}}$	●	●
	$\Delta\phi^{\text{cluster}}$	●	●
	λ_{cluster}	●	●
	$\langle \lambda^2_{\text{cluster}} \rangle$	●	●
	$\langle r^2_{\text{cluster}} \rangle$	●	●
High-level inputs	$p_T^{\text{uncalibrated}}$	●	●
	f_{cent}	●	●
	$f^{-1}_{\text{leadtrack}}$	●	●
	ΔR_{max}	●	●
	$\lvert S_{\text{leadtrack}} \rvert$	●	
	S_T^{flight}		●
	$f_{\text{iso}}^{\text{track}}$	●	●
	$f_{\text{track}}^{\text{EM}}$	●	●
	$p_T^{\text{EM+track}}/p_T$	●	●
	$m^{\text{EM+track}}$	●	●
	m^{track}		●

Search for Merged
$H \rightarrow aa \rightarrow \tau_\mu\tau_{had}\ \tau_\mu\tau_{had}$

Search for Merged $H\rightarrow aa\rightarrow\tau_\mu\tau_\text{had}\ \tau_\mu\tau_\text{had}$

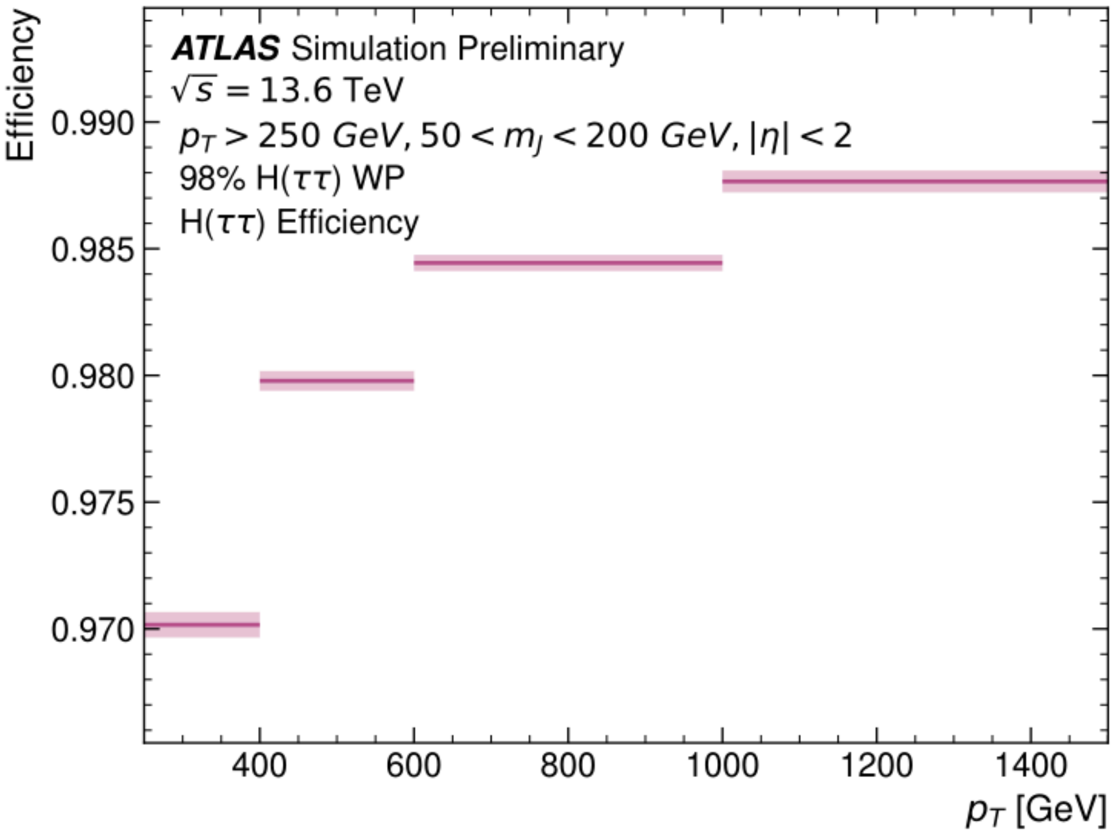
HMBS-2024-25

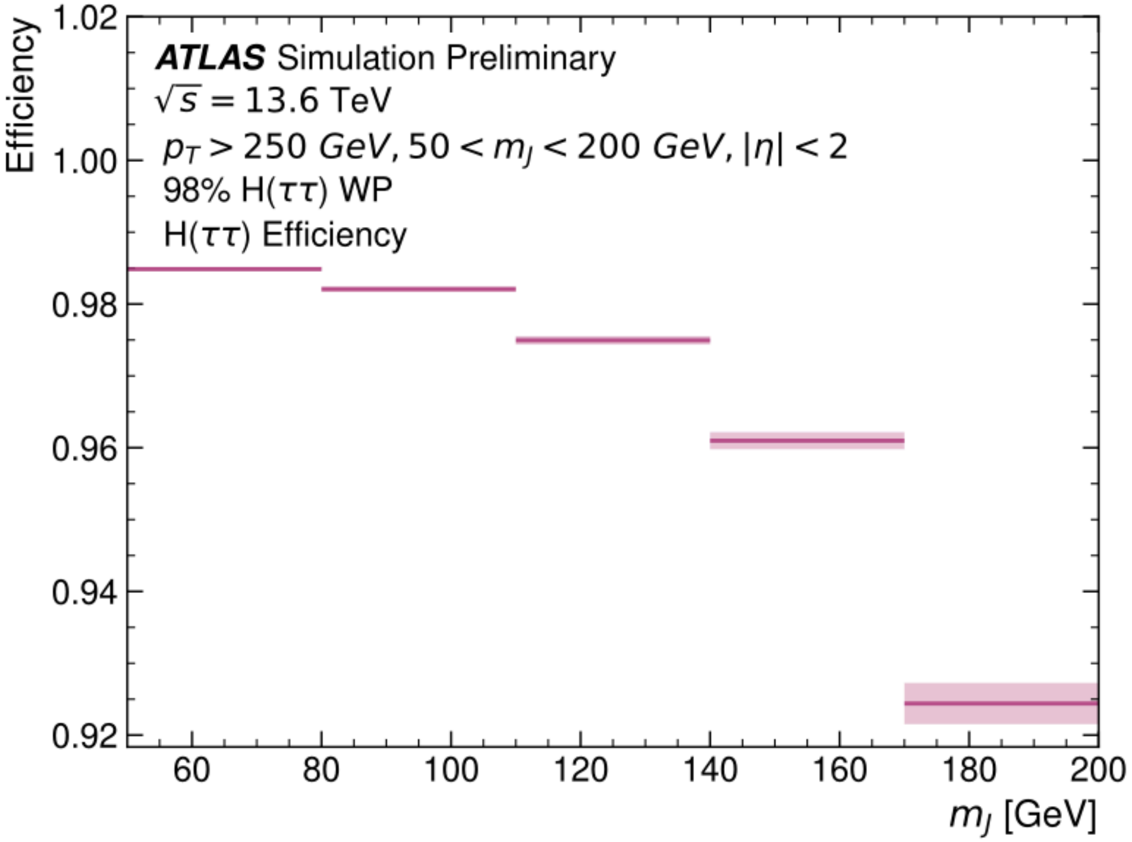


GN2X$\tau\tau$ for Merged $\tau_{had}\tau_{had}$ at High Momentum

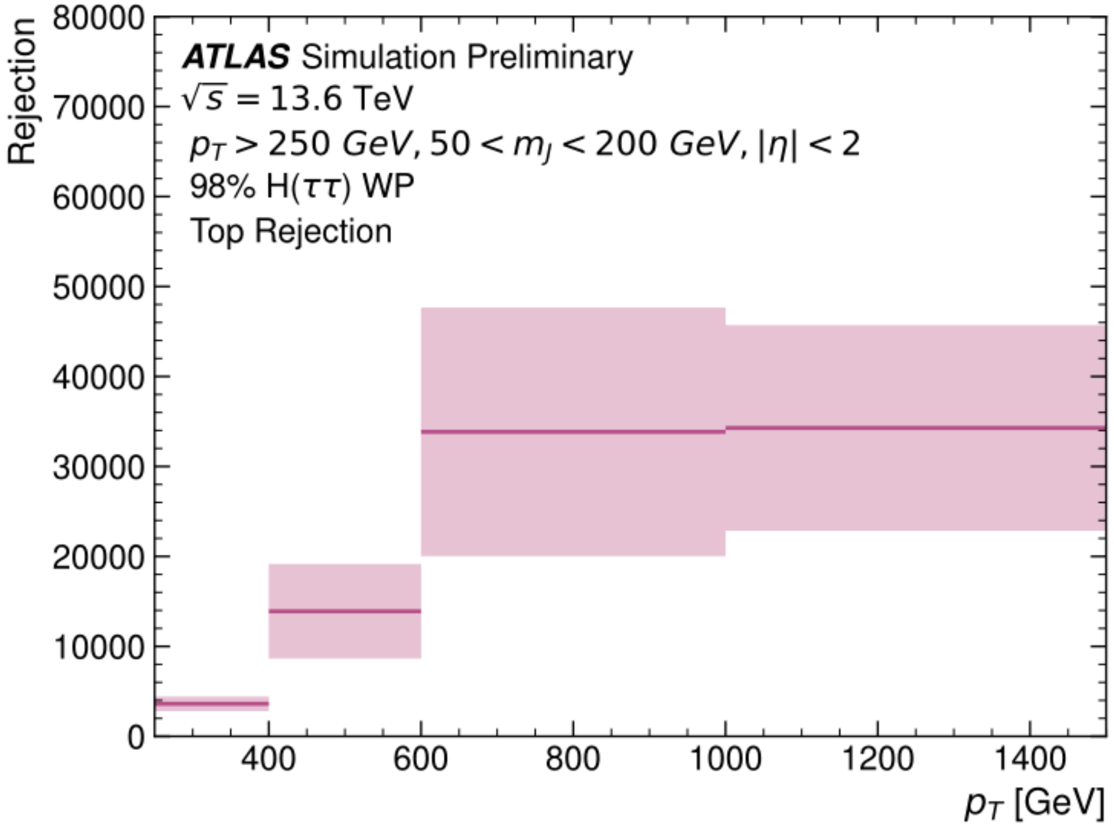
GN2X$\tau\tau$ for Merged $\tau_{had}\tau_{had}$ at High Momentum

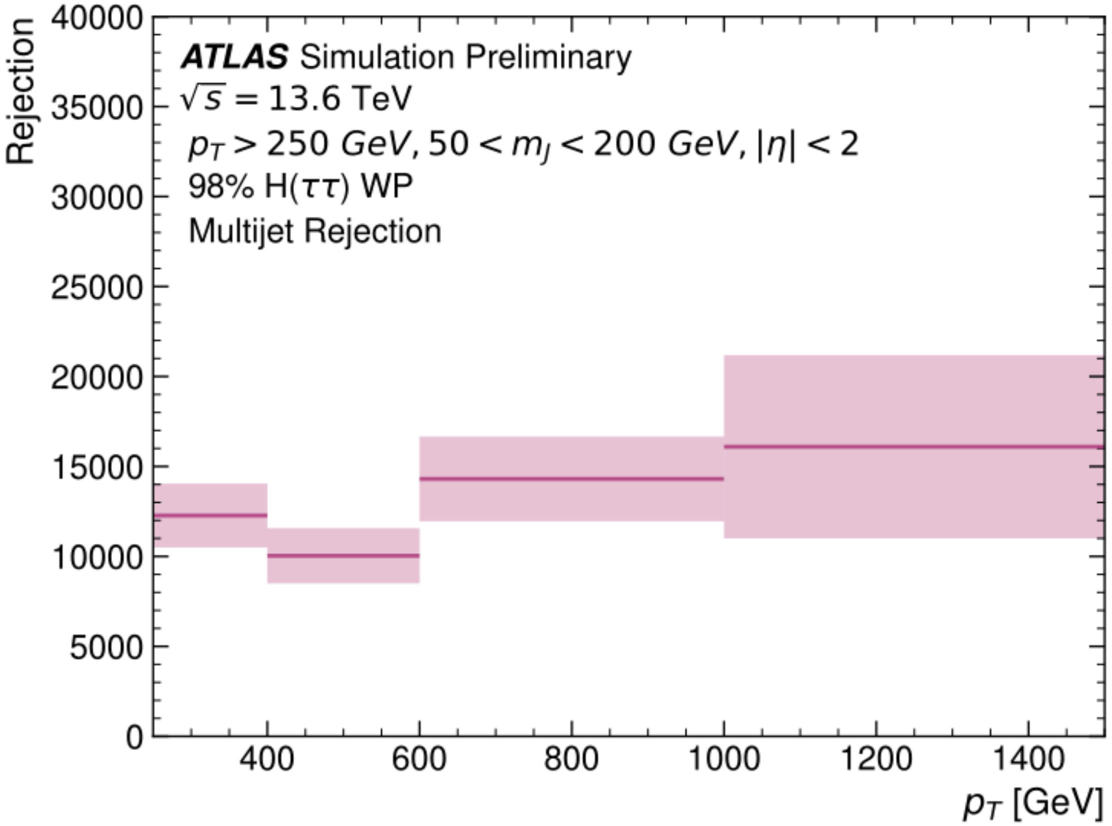
Efficiency vs. p_T/M





Rejection vs. p_T





ATL-PHYS-PUB-2025-026

Weakly Supervised Anomaly Detection for Light Di-Jet Resonance

Weakly Supervised Anomaly Detection for Light Di-Jet Resonance

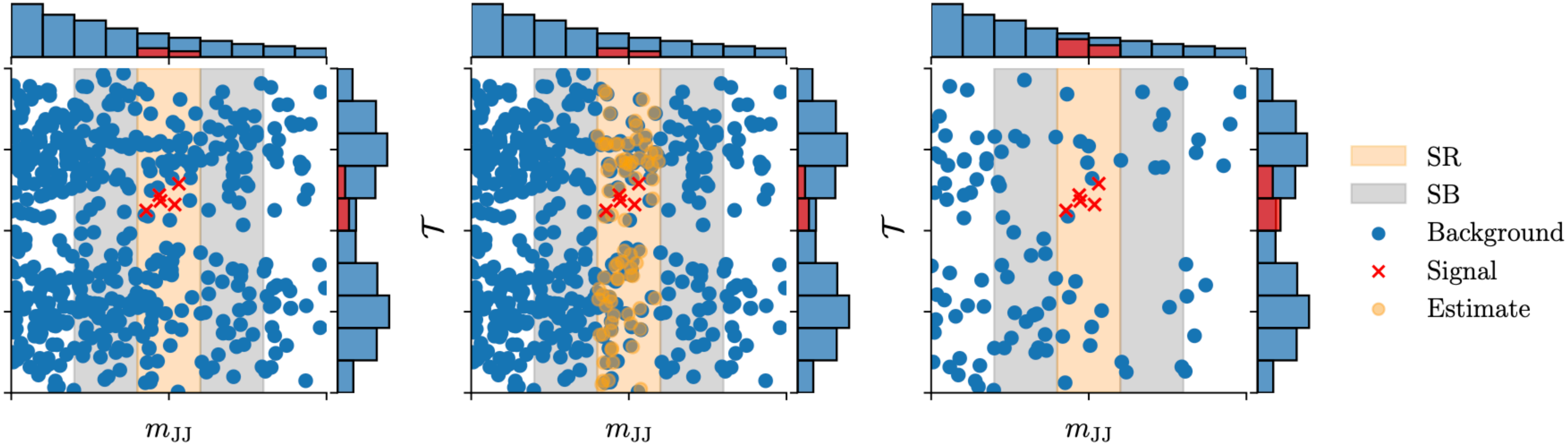
HMBS-2024-34/

- Di-jet mass sidebands (SBs) used for interpolation into SR when forming reference data sample: two methods explored
- Simulation Assisted Likelihood-free Anomaly Detection (SALAD)
    - Di-jet mass **conditional reweighting** function trained in SBs to corrects simulation in SR
- Constructing unobserved regions by transforming adjacent intervals (CURTAINs)
    - Di-jet mass **conditional morphing** function trained in SBs to correct SBs data to look like data in SR
    - Both approaches correct for correlations between M_{JJ} & classifier features

Weakly Supervised Anomaly Detection for Light Di-Jet Resonance

HMBS-2024-34/

- **Weakly supervised classifier** trained to distinguish between data & estimated background distribution in SR
    - Following **CWoLa** paradigm, NN to approximate optimal classifier between a potential signal & background (signal-depleted template)
    - **Five-fold cross validation** of training: each NN not applied to data used for its training
- **Cut on classifier** to enhance S/B
- Fit: smoothly falling function fit to SB bins until χ^2 p-value>5% first, post-hoc non-closure corrections & 95% CLs limits from intersection of limit cures on 20 different signal models



NSBI Application:

Off-Shell H→ZZ* (Γ_H)

NSBI & Application: Off-Shell H→ZZ* (Γ_H)

HIGP-2024-14

- NN classifiers can be used to discriminate between two hypotheses
- Train a classifier to estimate a decision function *s(x$_i$)* separating signal from background events using balanced class weights -> compute *r(x$_i$|S,B)* -> scaled to constructed LLR
- Application to off-shell H→ZZ* accounts for S-to-B interference, introducing a non-linearity in signal strength
- Architecture: feed-forward dense network with 5 hidden layers 1000 nodes each
  - Output layer with a single node and a sigmoid activation
- Improved Γ_H results

$$s(x_i) = \frac{p(x_i|\mathrm{S})}{p(x_i|\mathrm{B}) + p(x_i|\mathrm{S})}$$

Per-event probability density ratio

$$r(x_i|\mathrm{S},\mathrm{B}) = \frac{p(x_i|\mathrm{S})}{p(x_i|\mathrm{B})} = \frac{s(x_i)}{1 - s(x_i)}$$

LikeLihood Ratio

$$\frac{p(x_i|\mu)}{p(x_i|\mu=0)} = \frac{1}{(\mu\nu_\mathrm{S} + \nu_\mathrm{B})} \frac{\mu\nu_\mathrm{S}\, p(x_i|\mathrm{S}) + \nu_\mathrm{B}\, p(x_i|\mathrm{B})}{p(x_i|\mathrm{B})}$$
$$= \frac{1}{(\mu\nu_\mathrm{S} + \nu_\mathrm{B})} (\mu\nu_\mathrm{S}\, r(x_i|\mathrm{S},\mathrm{B}) + \nu_\mathrm{B}),$$