New analysis methods: discussion

Higgs Hunting 2025

Based on talks

by Maria Mazza and Andrea Sciandra

Julie Malclès

Remarks

Breakthrough innovations in methods bringing striking improvements in physics results

Shown today:

- Tagging/ID (specially b/c/s, tau, boosted topologies, but could be generalised to further objects/properties)
- Anomaly detection methods
- Simulation-based inference methods

Many others in use/development:

- Dedicated reconstruction of collimated objects with ML (photons, electrons,…)
- Dedicated pT/mass regressions, kinematic fits
- Data/MC correction methods (chained quantile-regression, normalising flows)
- Background modelling (generative networks, generator tuning,…)
- Systematic uncertainty aware deep learning methods

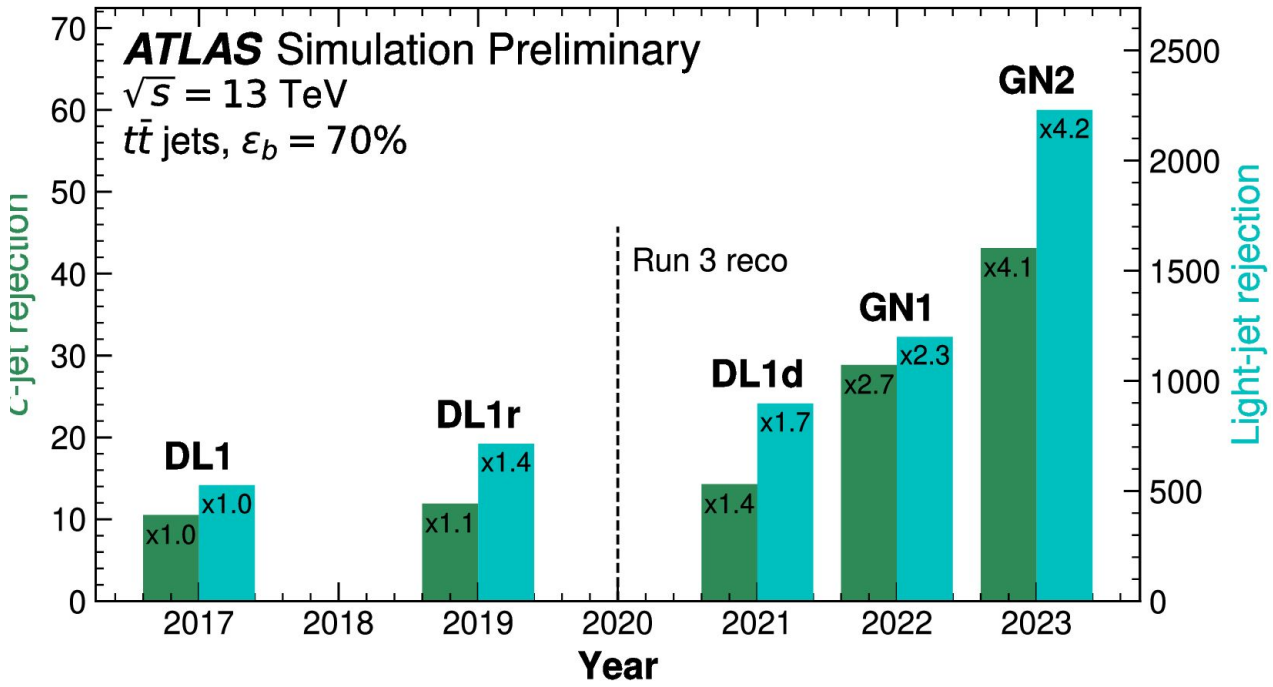
These new methods are often:

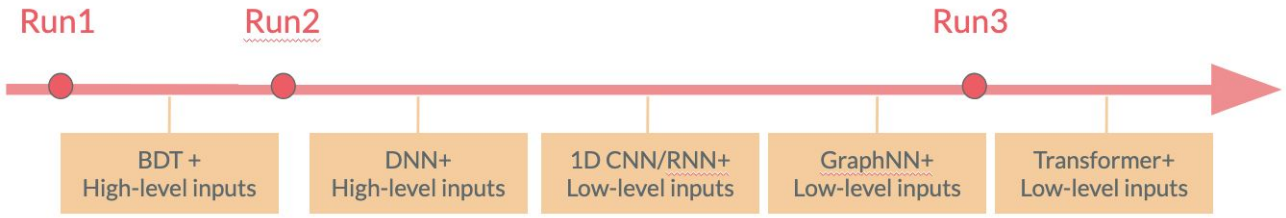
- Based on advanced machine learning, computationally expensive, needing high MC statistics
- Coming with big challenges and a long time of development
  - many of the methods being developed in collaborations are not yet in use in physics analyses
  - would call for early developments for phase II detectors
- Many are backward compatible: calling for reanalysis of past data when the improvement is sizable i.e. bringing more than X% of sensitivity. What would be your X?

Flavour-tagging

Courtesy Angela Maria Burger and Angela Zaza @ Higgs Pair

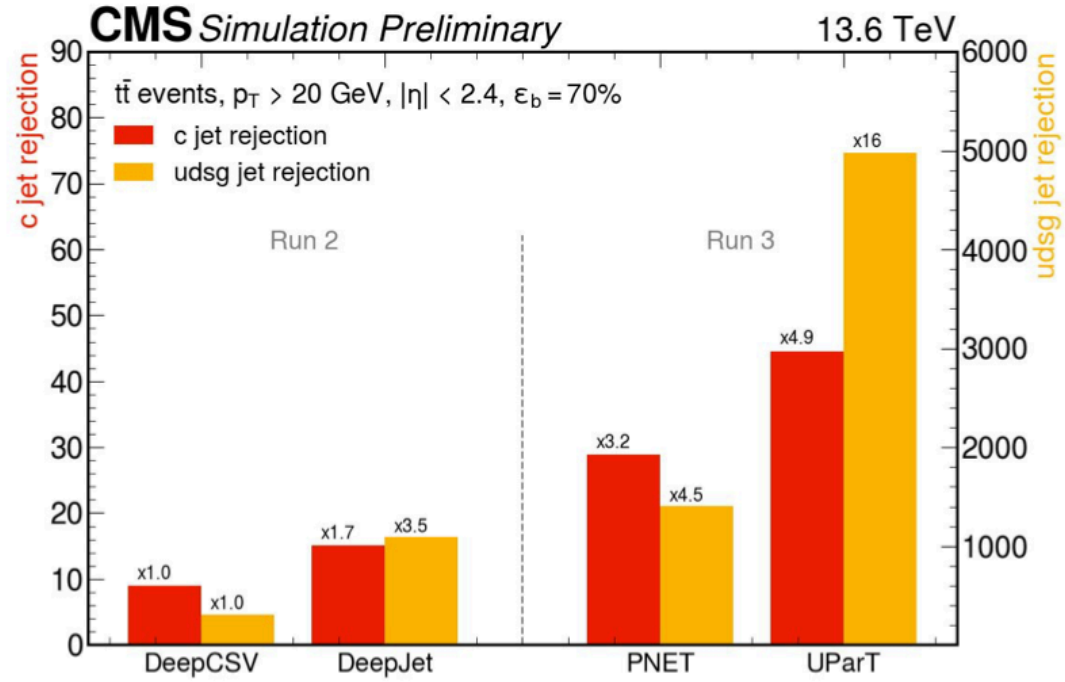
FTAG-2023-01





CMS (similar for ATLAS)

CMS-DP-2024-066

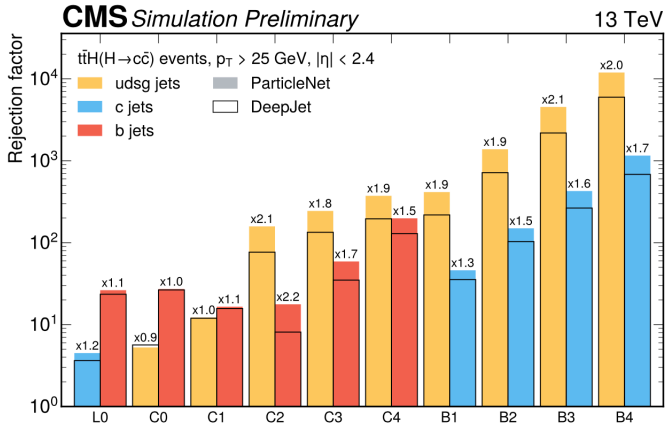


Heavy-flavour tagging performance greatly benefits from new developments in ML architectures

Flavour-tagging/tau-ID

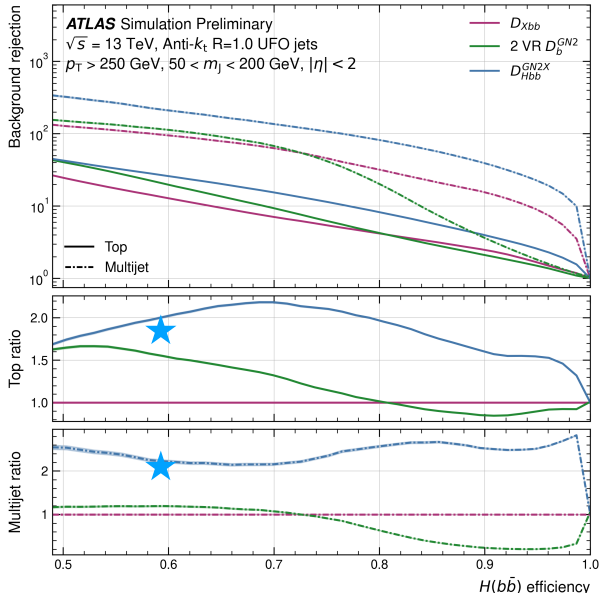
Besides improved b versus light-jet discrimination, many innovations:

- Multi-task algorithms:
  - mass/pT/resolution regressions
  - tau-h and c-quark tagging capabilities
  - Impacting importantly triggers as well



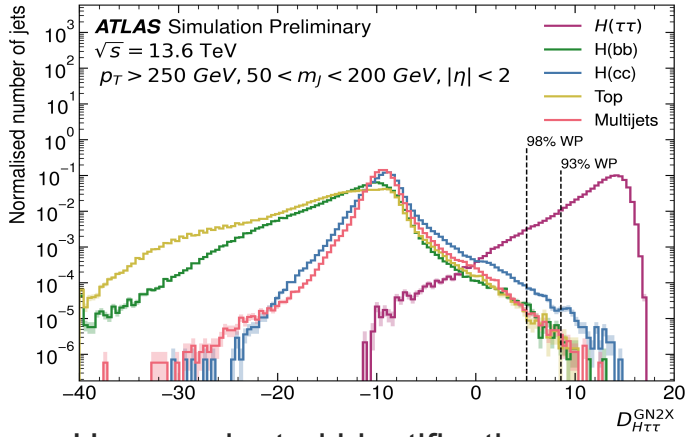
~ x 2 background rejection improvement over end Run 2 algo

- Boosted jet tagging/double b-tagging:
  - tagging for large radius jets,
  - specialised low pT double b-tagging
  - “Jet-less” Flavour Tagging: very low pT B-tagging outside jets



~ 2 x better bkg rejection for same H→bb efficiency

- Improved tau ID/reco techniques:
  - Improved TauID (deepTau, RNN)
  - mu-tau removal for merged” τ_{had}-τ_{μ}
  - double hadronic-tau tagging



Unprecedented identification of merged X->$\tau_h\tau_h$

Flavour-tagging/tau-ID

Besides improved b versus light-jet discrimination, many innovations:

- Multi-task algorithms:
  - mass/pT/resolution regressions
  - tau-h and c-quark tagging capabilities
  - Impacting importantly triggers as well

- Boosted jet tagging/double b-tagging:
  - tagging for large radius jets,
  - specialised low pT double b-tagging
  - “Jet-less” Flavour Tagging: very low pT B-tagging outside jets

- Improved tau ID/reco techniques:
  - Improved TauID (deepTau, RNN)
  - mu-tau removal for merged” τ_{had}-τ_{μ}
  - double hadronic-tau tagging

Major impact on very many analyses

- All Higgs analyses including b jets:
  - ttH in all channels, especially bb and cc

  Example: most stringent limit on K_c (CMS)
  - HH in bbyy, bbtautau, bbbb, bbWW.

  Example: most stringent single channel limit on HH production from bbyy including reanalysis of run2 (ATLAS) - 20% from b-tagging at same luminosity
  - VH/VBF in bb and cc
- Several BSM searches driving/benefiting from new techniques:

Examples: a->bb searches, a -> τ_{had}-τ_{μ}

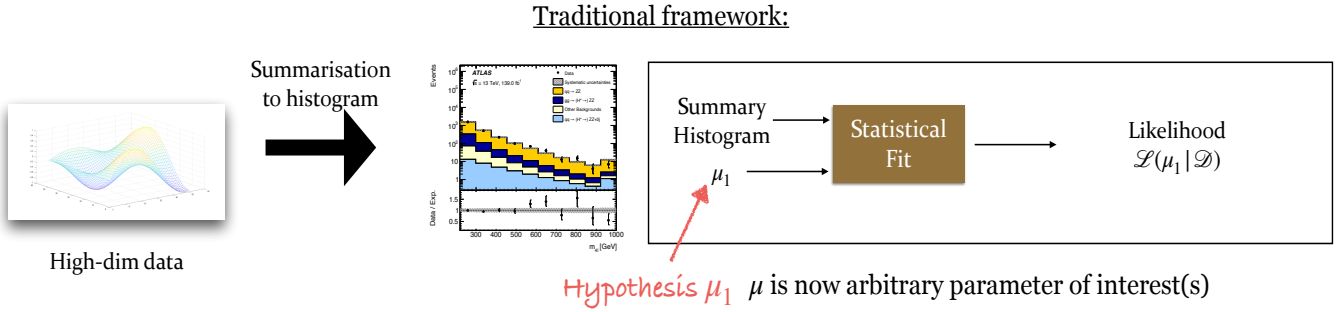
- Some improvements not yet is use in analyses
- Gains sometimes equivalent to a sizable lumi increase calling for Run 2 re-analyses

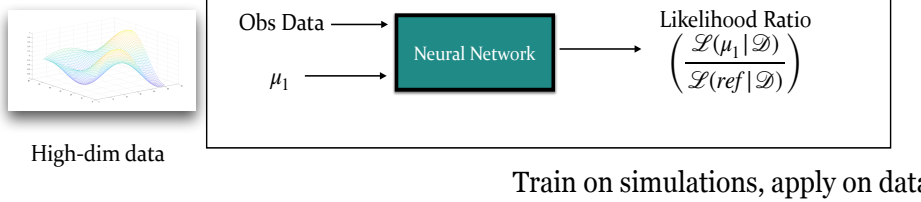
Simulation-based inference methods

Complete information exploited with simulation-based inference methods:

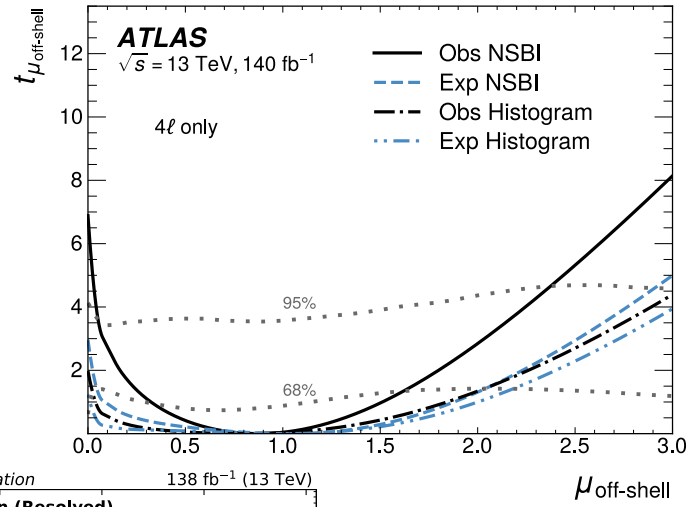
- Train a ML model to learn the likelihood ratio from simulation
- Handle many observables without loss of information
- Scalability to high-dimensional parameter space

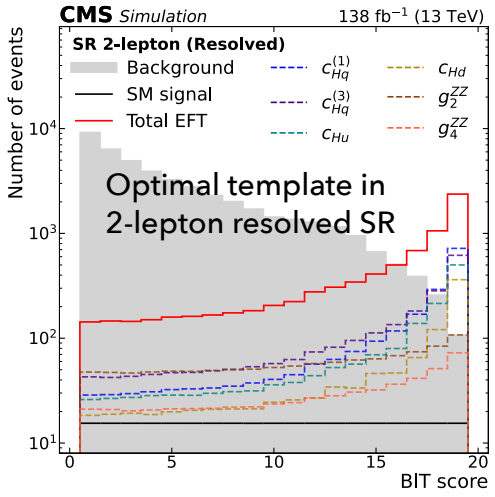
Courtesy Aishik Ghosh @ CERN Seminar





Off-Shell H->ZZ* (Γ_H) ~13% relative improvement in 95% CL upper limit on Γ_H, compared to standard histogram analysis





VHbb EFT interpretation: Build optimal observables with simulation-based inference methods

Complex in practice: systematics uncertainties to incorporate to the likelihood (large number of networks), several processes and parameters (ensemble of networks, parameterization)

Discussion time!

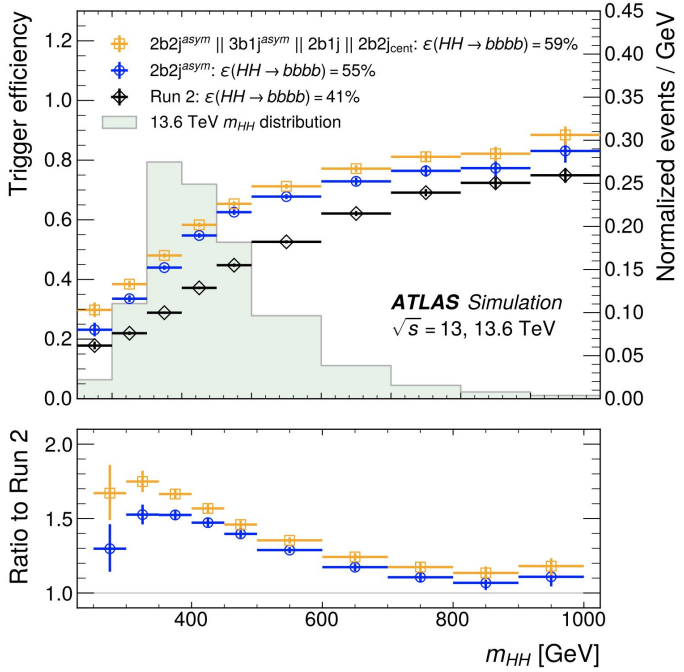
Trigger

Courtesy Liaoshan Shi and Silvio Donato @ HiggsPair

ATLAS

- Delayed stream since 2022.
- Merged back to main stream from 2024.
- 45% improvement w.r.t. Run 2.

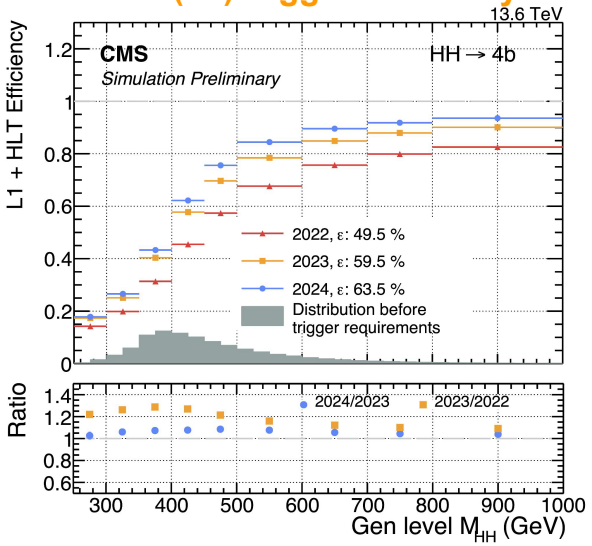
HH(4b) trigger efficiency



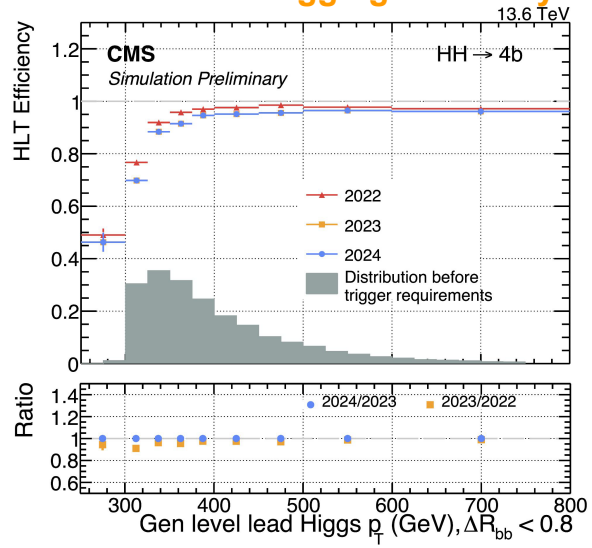
CMS:

- High-rate parking strategy since 2023.
  - 60% improvement w.r.t Run 2.
  - 20% improvement from 2022 prompt strategy.
- Loosened thresholds in 2024.
- Dedicated trigger for boosted H(bb)

HH(4b) trigger efficiency



Online bb-tagging efficiency

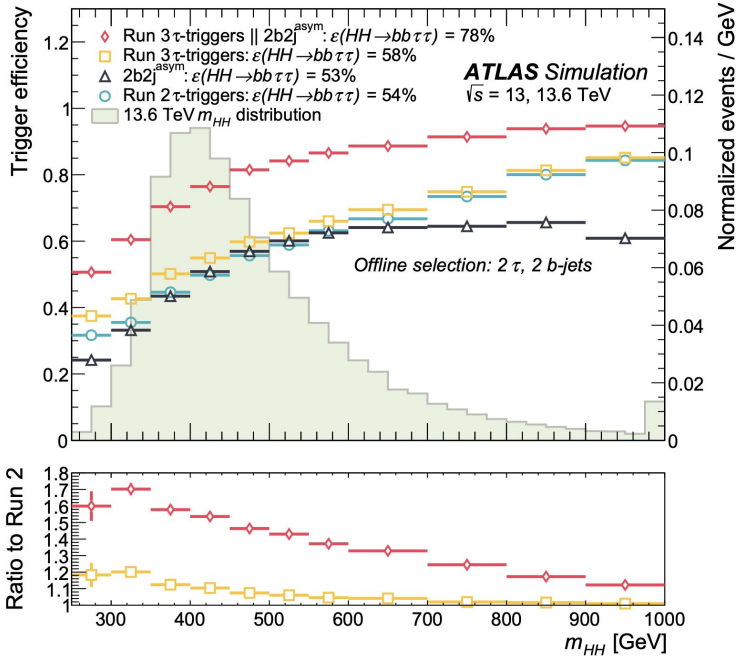


Trigger

Courtesy Liaoshan Shi and Silvio Donato @ HiggsPair

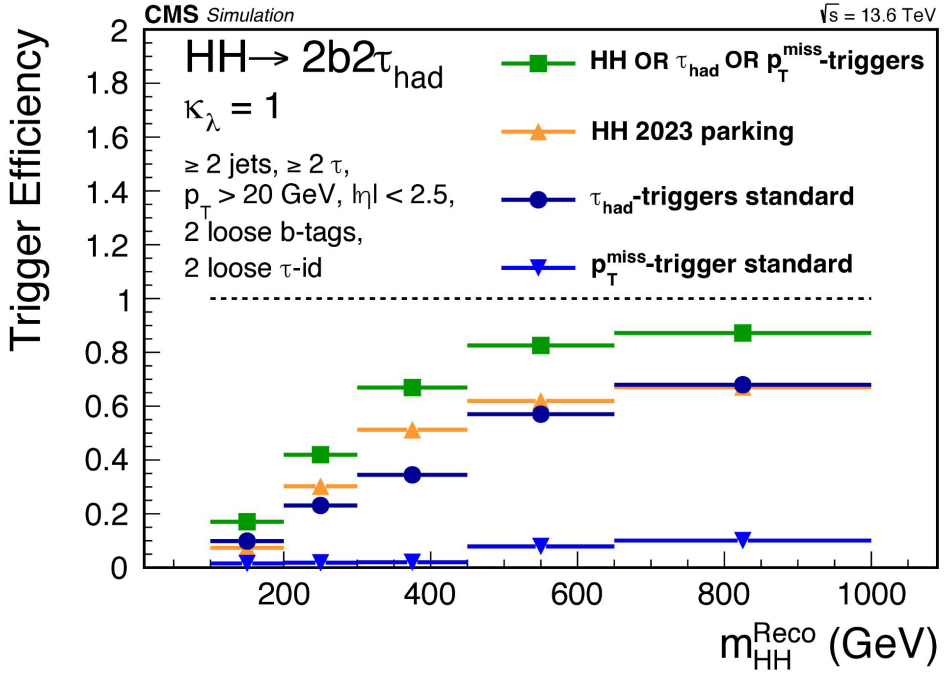
ATLAS

- Complementary phase space selected by bbjj and tau triggers.
- More than 40% unique efficiency gain when combining triggers.



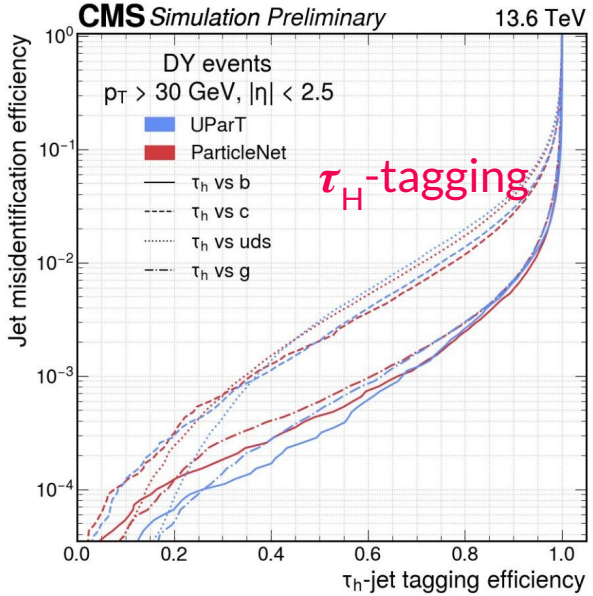
CMS

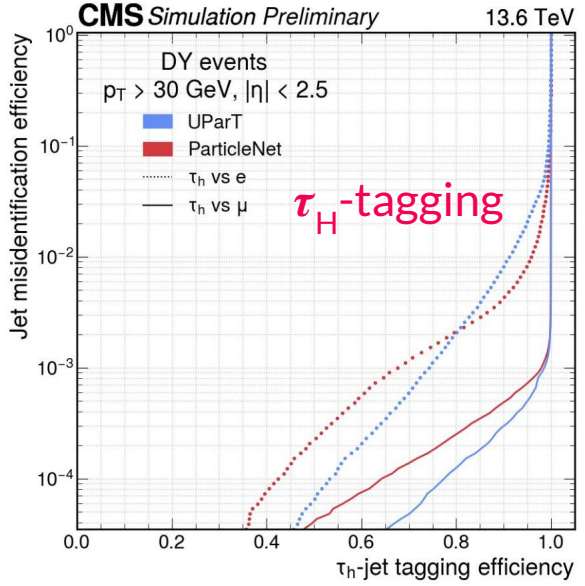
- 25% improvement w.r.t. tau-triggers
- 70% unique efficiency gain when combining triggers.

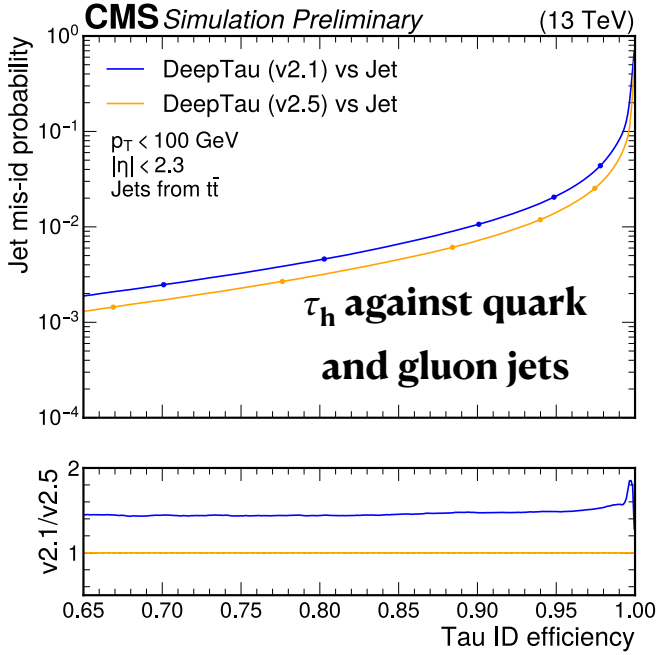


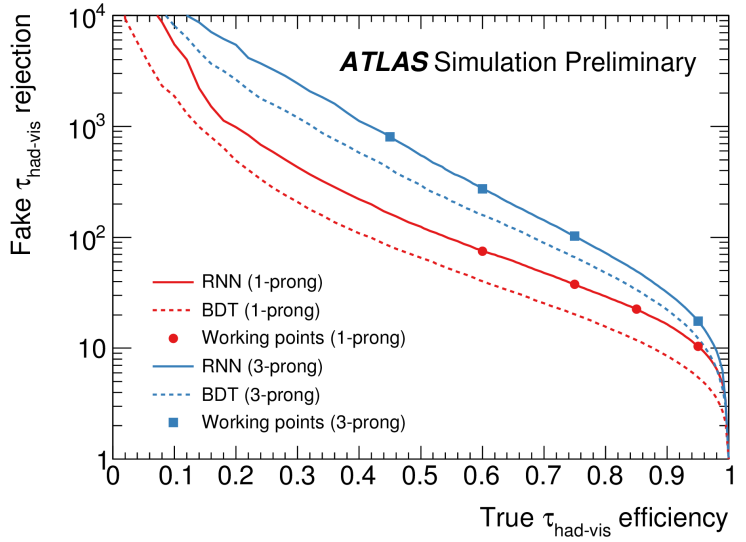
Tau-ID

Gains in tauID in both experiments with deep learning





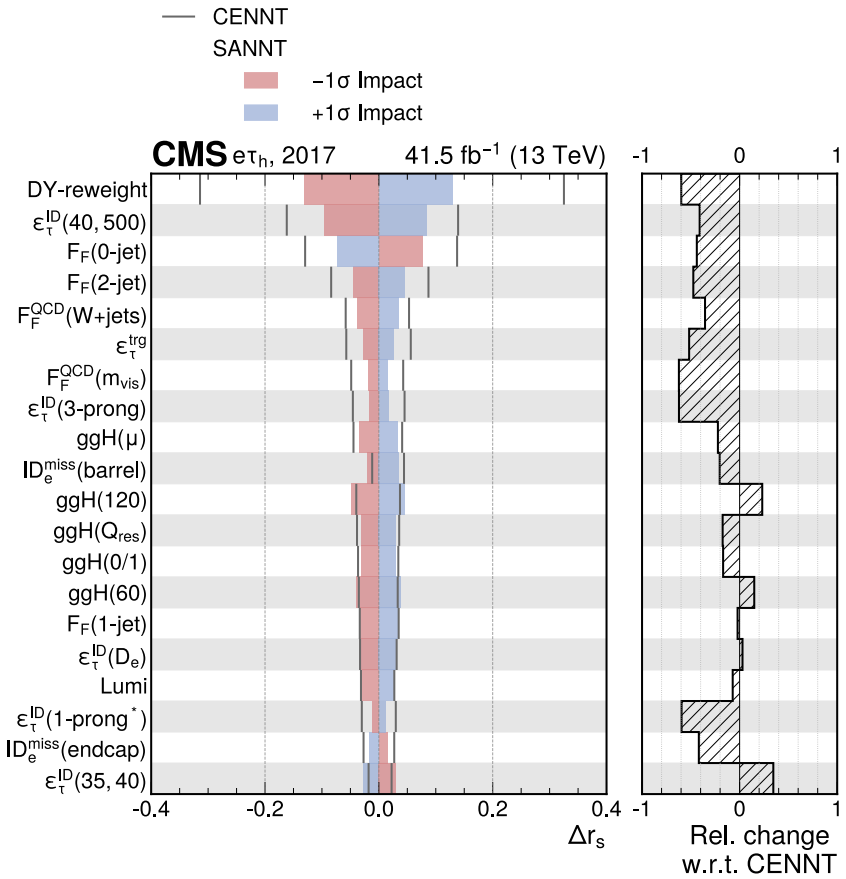


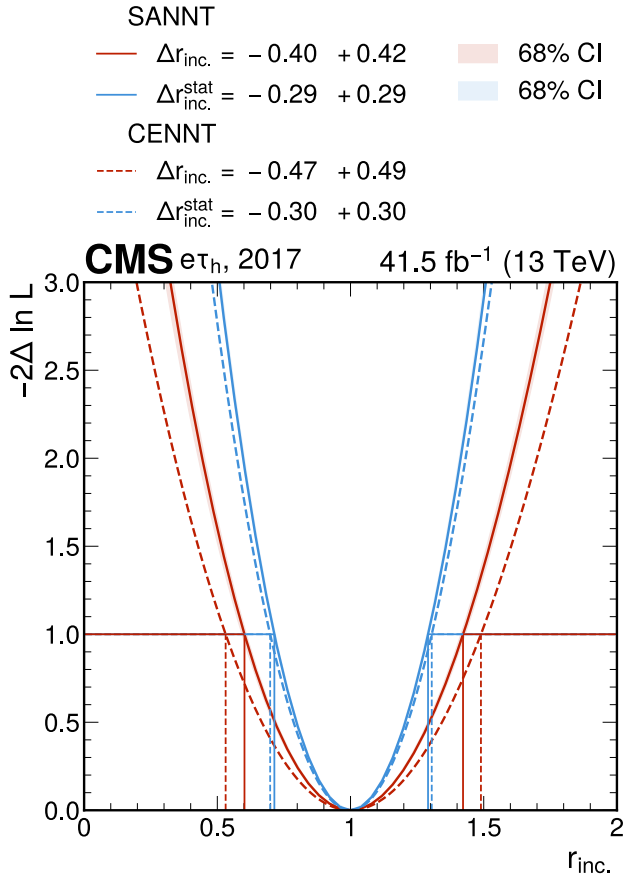


Systematics aware deep learning

Demonstration in the ττ decay channel

Example: arXiv:2502.13047v1 [hep-ex]





Improvement by 12% (16%) for ggH (VBF) signal strengths

Sensitivity

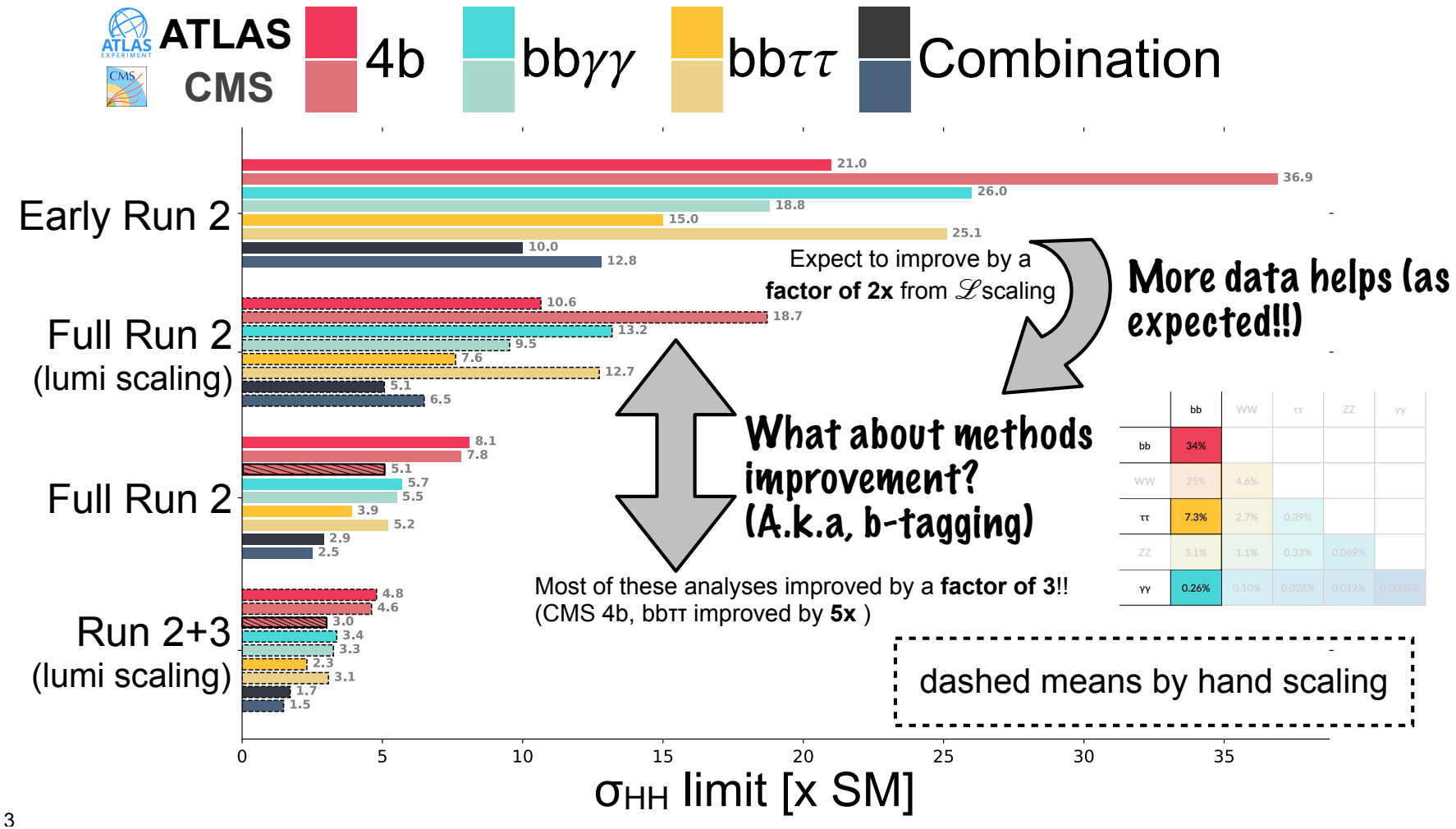
V. Cairo slide design

baseline

RNNs +Sets

Graphs

Trans-formers



13