Collider data analysis strategies:
Thoughts on object identification and event classification

Loukas Gouskos (CERN)

Snowmass community planning kick-off meeting
October 2020

Setting the scene

- It is still extremely exciting to be in Particle Physics [& in Collider Physics]
  - LHC: Discovery of a Higgs boson; exhaustively studying its properties
    - Many more to learn; Higgs becomes an exploration tool for BSM physics
  - Unique opportunity: direct impact on the choice & design of next generation particle collider(s)
    - Various options: e^+e^- linear vs. circular (even $\mu^+\mu^-$) and/or hadrons

- Main physics priorities:

e^+e^- (focus on low/med E_{cm})

- Measure EW/H/top properties <1% [necessary to probe M_{NP}>O(1TeV)]
- Flavor physics (e.g. FC violation)
- Very weakly coupled particles (e.g., RH neutrinos, dark photons)

hh (e.g., E_{cm}~O(100) TeV)

- Direct search of particles with M~30 TeV
- Usual suspects: DM, SUSY, ..
- EW/H/top physics and rare processes

- Success of the physics program requires coherent effort in all areas
  - detector design, triggers/DAQ, event reconstruction & ID, analysis techniques + TH
- This talk: focus more on thoughts/challenges related to future colliders
  - yet, many relevant for the upcoming LHC and HL-LHC runs

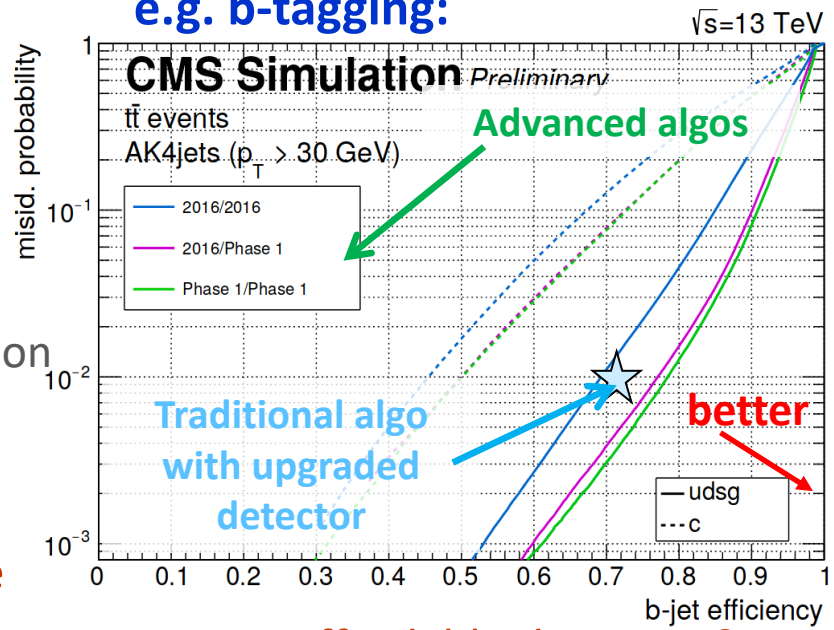
Some thoughts

- **Detector optimization:**
  - **e^+e^-:** Tracking: as low material as possible
  - **hh:** TRK/Calo granularity for ultra high-p_T objects and PU
- **Object reconstruction:**
  - e.g., e^+e^-: goal max efficiency and precision
    - jet clustering: inclusive, exclusive, full evt?
- **Object identification:**
  - Heavy flavor (b/c) ID: extend to s/ud/gluons
    Also: soft &/or displaced particles
- **Key player:** Advanced ML
  [e.g. lower-level info (particles →hits/clusters)]
  - But: understand what ML learns:
    - need uncertainties under control/calibration
    - EXP-TH x-talk: tune/develop MC GEN
- **Other considerations: Cost**
  - Absolute performance vs. cost-performance
  - Can improvements in algos/computing compensate more affordable detectors?

Successful LHC recipe

- Detector capabilities
- Theory/Pheno input
- Ingenuity in algo design
- Advanced tools (e.g., ML)
- Computing (GPU, HPC..)

e.g. b-tagging:

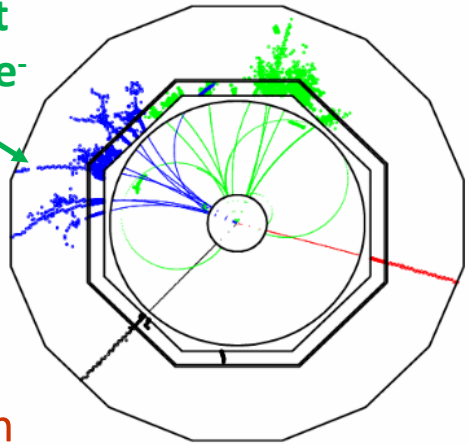


Some thoughts (II)

- Improving event categorization:
  - Novel techniques DL, multi-class evt categorization
  - Novel observables [e.g. track-based ones]
- Generally:
  - Increase effort on more complex topologies [LLP, dark sector]
  - Challenge [particularly in e^+e^-]:
    achieve systematic uncertainties **similar** to statistical precision

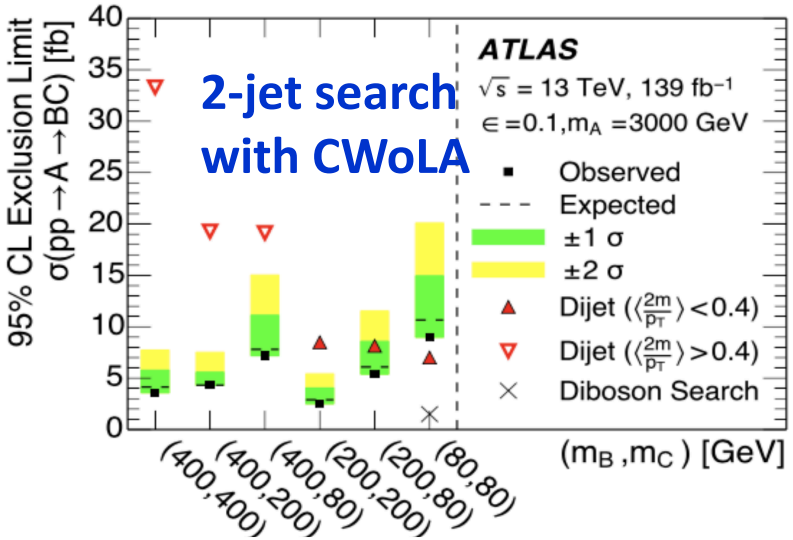
e^+e^-: Z(->μμ)H(->bb)

Full event reco in e^+e^-



- Current LHC results do not concretely point to any BSM scenario/ mass scale

  Explore the unknowns:
  - Alternative methods: VAE, CWoLA, …
  - e.g. Train on "known" jets types [QCD, top, H, Z, W] look for outliers: unconventional jets
    - Similarly in event classification [probably more relevant for hh colliders]



- All these strongly coupled to the advancements in computing
  - FPGAs for HPC, Quantum computing, ..