# (re)-interpretation discussion

Andrew Whitbeck

October 15, 2017

# Materials from BSM searches

**Are we handling this for MVA-based analyses well enough?**

- Efficiencies, cut flow tables, standalone code,
  - are code snippets useful? Should we be doing more of this?
- Background predictions (+ correlation matrices)
  - can build simplified likelihood
    - these are not always an accurate representation, especially in the tails of distributions...
- Fake rates?
  - fake rates typically are relevant for backgrounds and covered by previous bullet
  - however, signals can also have significant “fake”

- What about the truth level information for BSM searches?

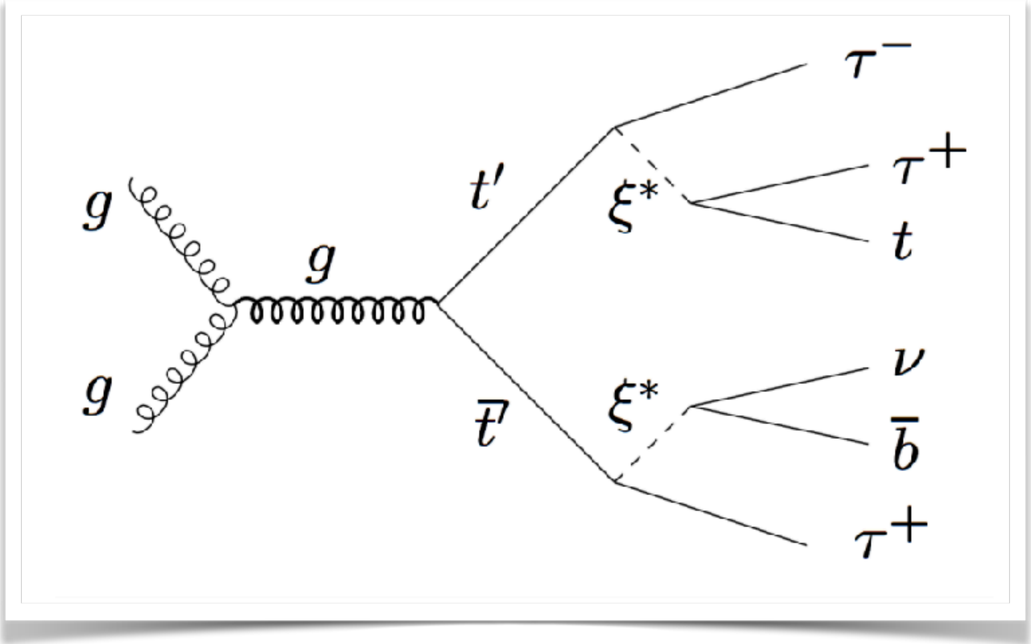
- In multidimensional interpretation, release significance plots

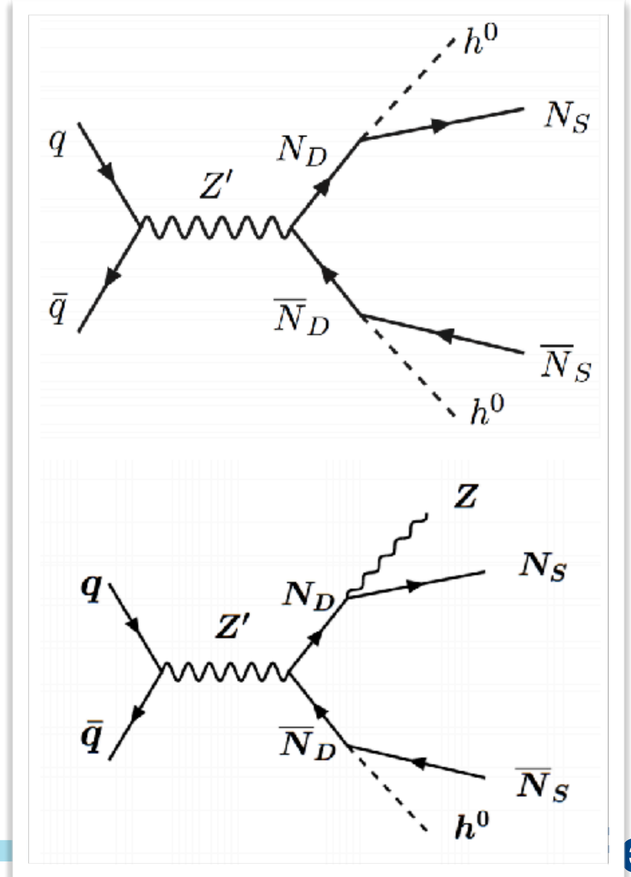
# Get creative …

- We had a lot of new ideas proposed by Bogdan
  - instead of have many resonance searches target multiboson searches check non-resonant searches, like SUSY searches

More LHC signatures: $t\bar{t} + 4\tau$, $t\bar{t}\tau^+\tau^-\nu\nu$, $tb\tau + 3\nu$ or $t\bar{t} + 4\nu$.

- multi-lepton SUSY searches have many clean signal regions



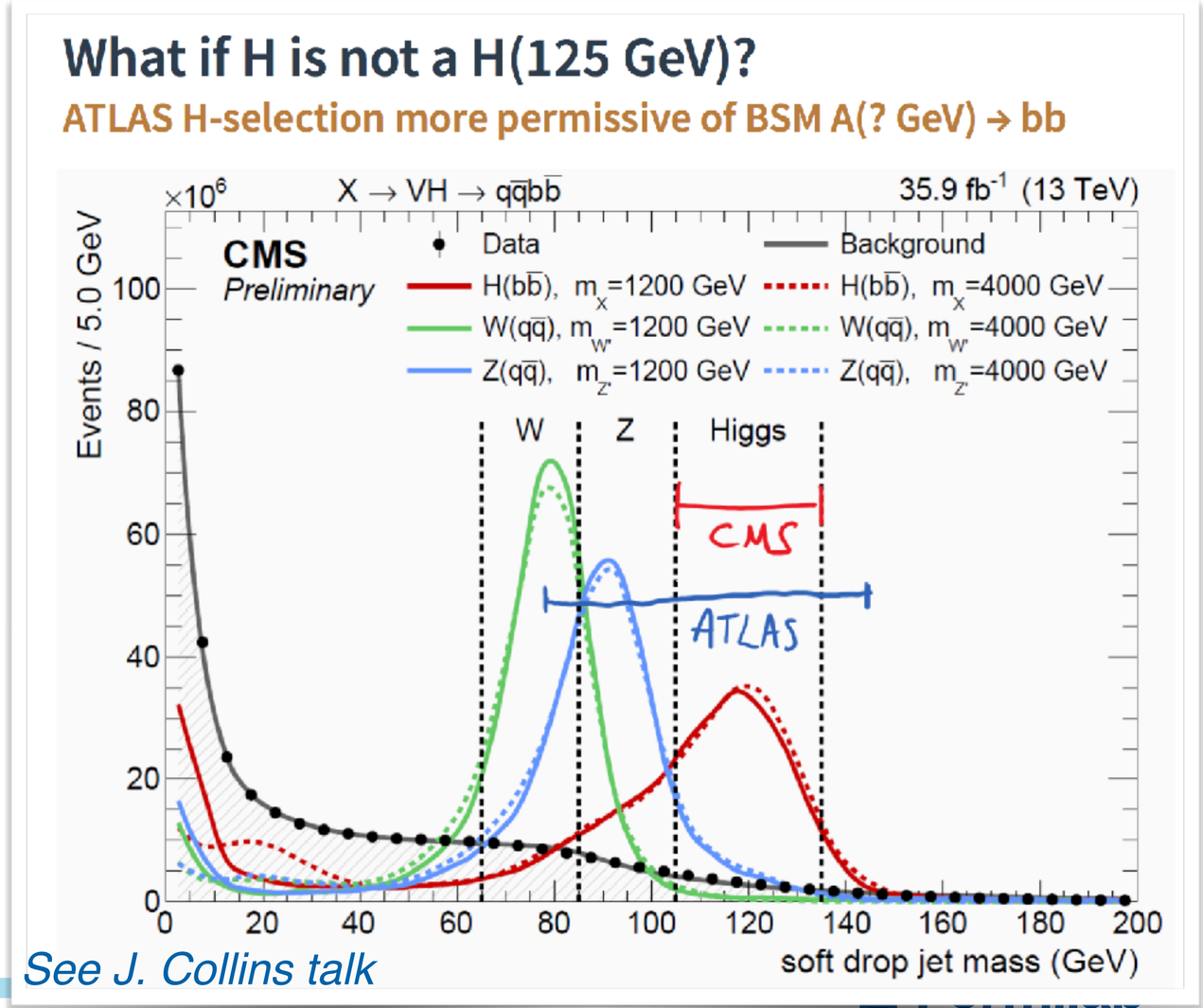



# Control regions

- In many searches, control regions don't significant signal contamination
- Sometimes this gets tricky:

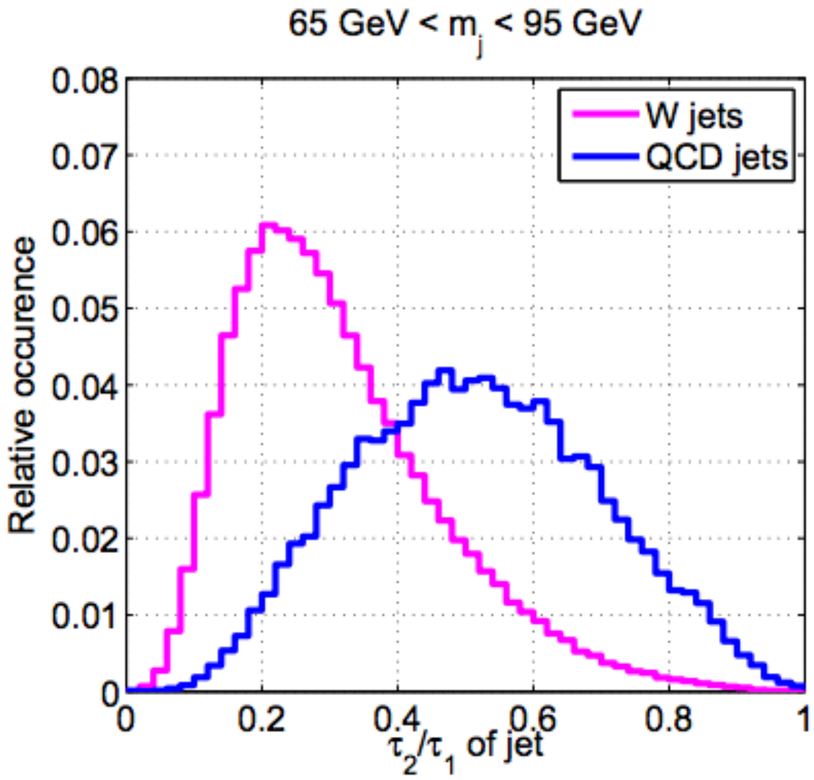
  "One man's onion is another man's plum"

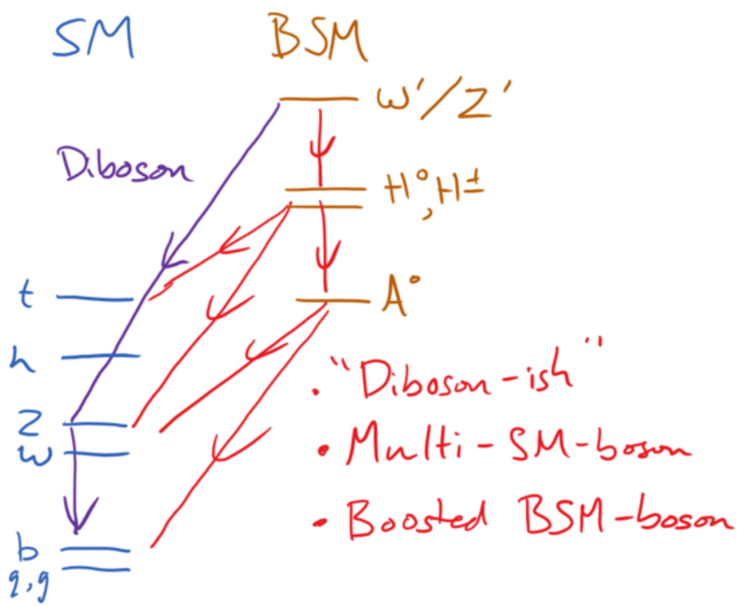
- Can we make control regions public without repercussions?



# Control regions

- Some taggers aren't very efficient — could mean there is a lot of signal in your control region
- Some mixed signals might populate control regions



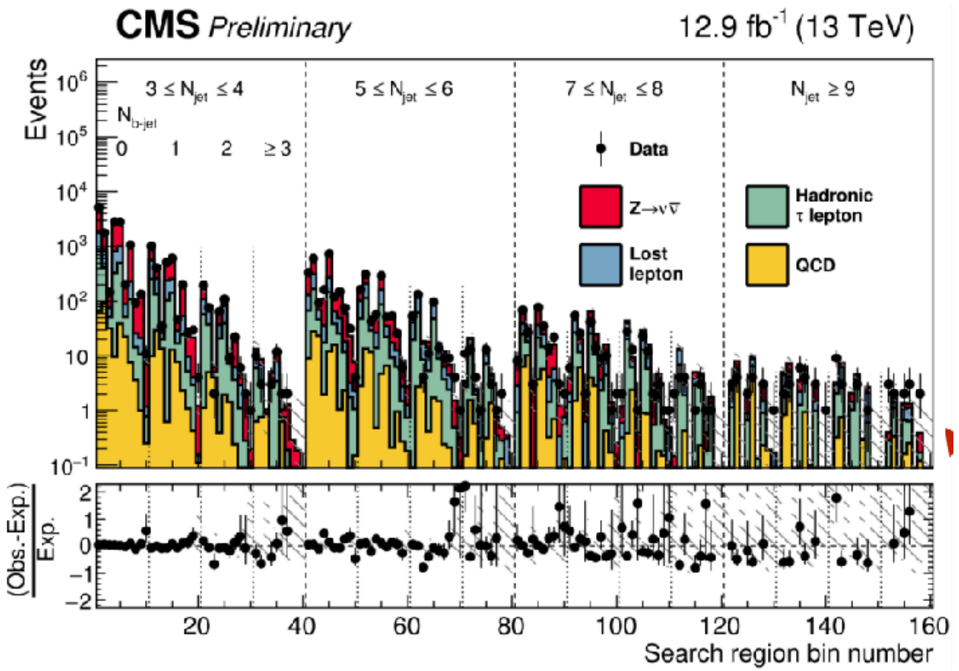



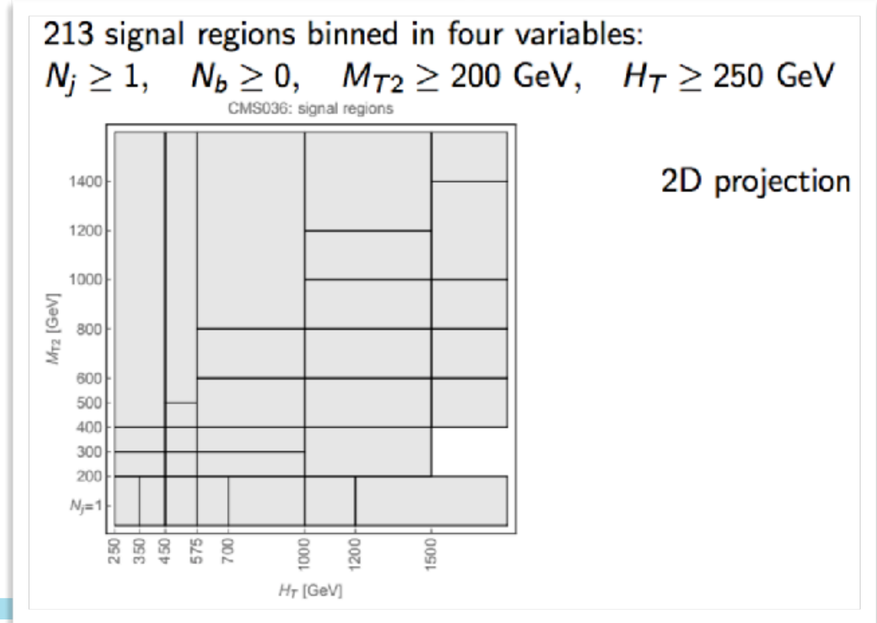
# Broadly applicable (sensitive)

- Especially in SUSY, there is a split in paradigm in CMS and ATLAS

- Many search bins are more optimal, but harder for reinterpretations
  - lots of progress over the last year to fix this. Is it enough?

- I think we have seen a success story in Angelo's talk … hopefully others will follow suite?
  - this focused mostly on the region where the simplified likelihood is expected to be well-behaved



# Recasting SUSY searches example…

CMS searches have so many signal regions (100+) that they are difficult to reinterpret. A proper approach would require sophisticated statistical methods combining multiple exclusive bins, using information (the correlation matrix of errors) that is not publicly available. In contrast, the ATLAS searches explicitly provide 95% CL limits on number of events due to new physics for each signal region, and generally have far fewer, coarser bins, allowing us to simply use the most sensitive SR to set a conservative (but reasonably accurate) limit.

One important case where CMS did include aggregate signal region information is the jets+MET search with $H_T^{\text{miss}}$ [52], which we find to be very powerful. The CMS jets+MET search has $b$-tagged categories, low-MET and high-MET categories, few jet and many jet categories. As a result, it is equivalent to the union of several different ATLAS jets+MET searches.[5]

Our simulation methodology is as follows. We generate hard events using MadGraph 5.2.3.3 [56], generating pairs of gluinos, squarks, and antisquarks in all possible combinations. Decay and showering is performed via Pythia 8.219 [57] (with which we implement the RPV and stealth decays of the higgsino), and we use Delphes 3.3.2 [58] with an ATLAS-approximating detector geometry for detector simulation. In order to speed up the event generation, we worked with unmatched samples in most cases.[6] The NLO cross sections for the superparticle production are obtained via Prospino [59–61]. The cutflows for the ATLAS and CMS analyses are recasted using Root 5.34 [62], and validated against the published experimental limits on supersymmetric simplified models.

Complaints 😕

Envious Statement 😕

Compliments 😀

Event Generator 🤔