Hexaquarks

find h-dibaryon

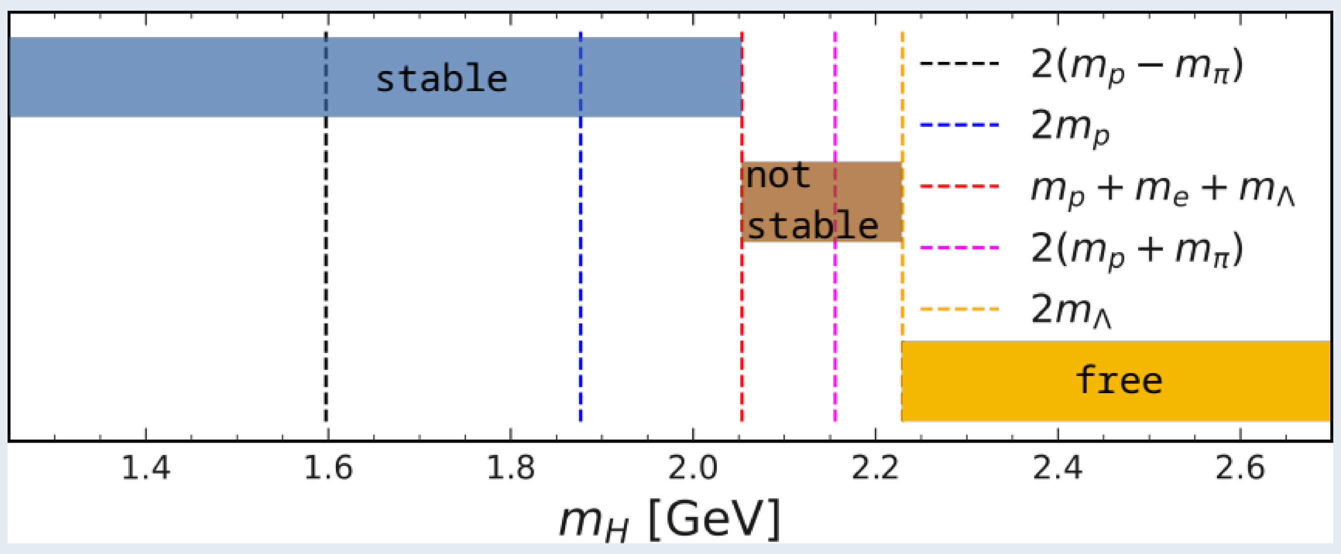
Search I'm feeling lucky

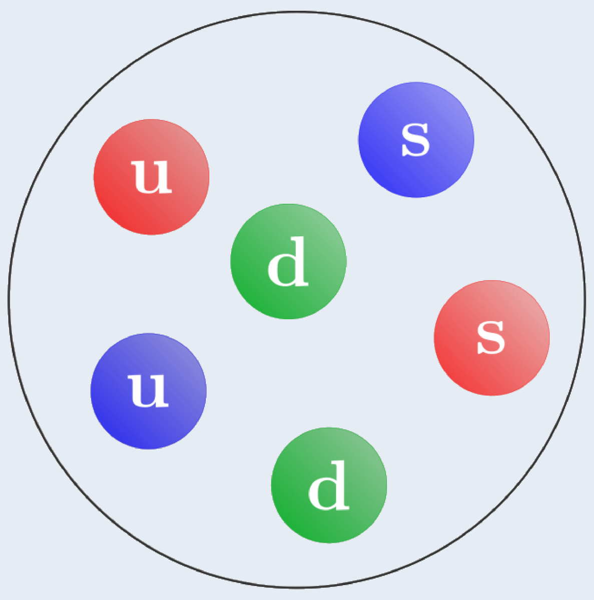
Felix M. Keil
Johannes Gutenberg-University Mainz, Germany

What About The H-Dibaryon?

- The H-Dibaryon (H) is a spin- and isospin scalar flavor-singlet six-quark state.
- The quark content is udsuds.

- If it is stable or not depends on the mass / binding energy.
- Current Lattice-QCD calculations favour an H-Dibaryon with small binding energy of < 7 MeV.
- A small binding energy means a weak decay: H → Λ p π

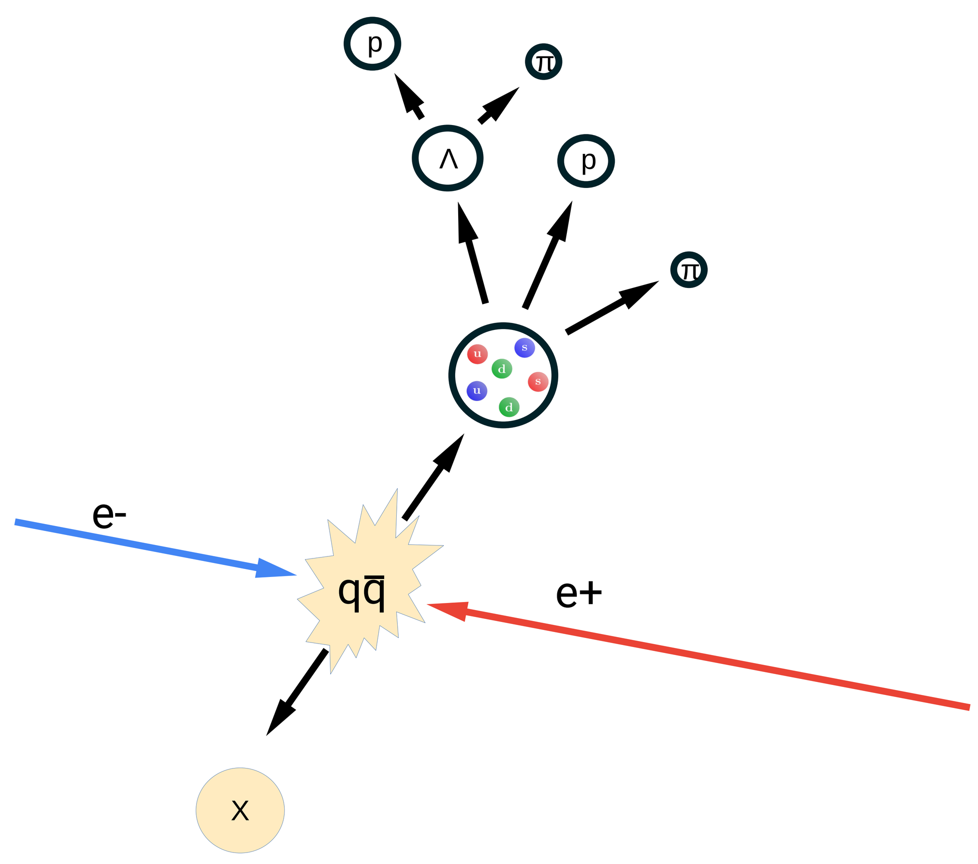




The Decay

We use continuum!

$e^+ e^- \to q\bar{q} \to H [\to \Lambda p \pi] + X$

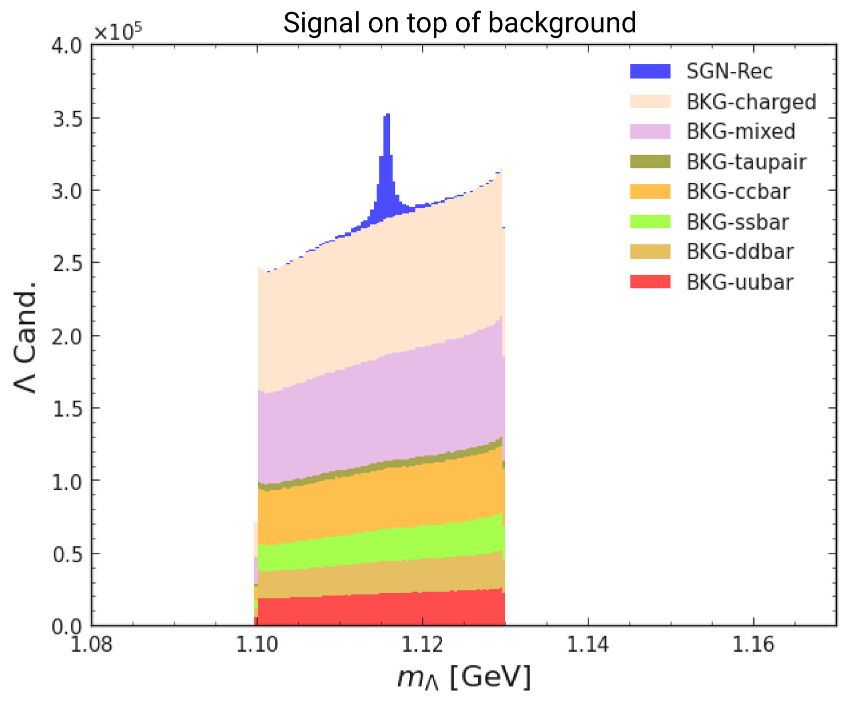


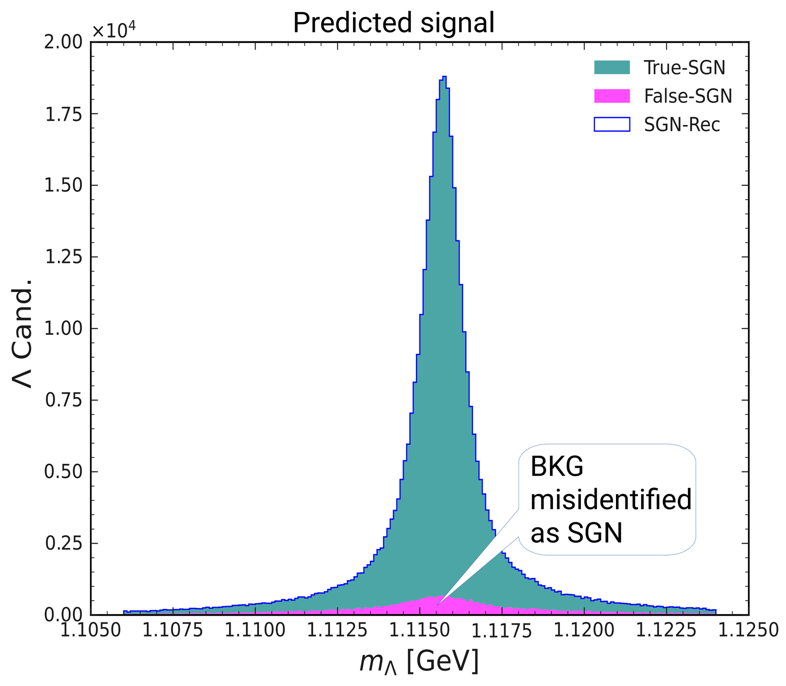
Optimizing Lambda Selection

Λ → p π⁻

- Involved in most of the decay chains of the H-Dibaryon
- Preprocessing to prepare data, eg. [1.10 GeV < m_Λ < 1.13 GeV]
- Serves as a perfomance test of the BDT-approach
- Detector-reconstruction ~ 34.55% efficiency and 4.09% purity

Started with high-eff. dataset!





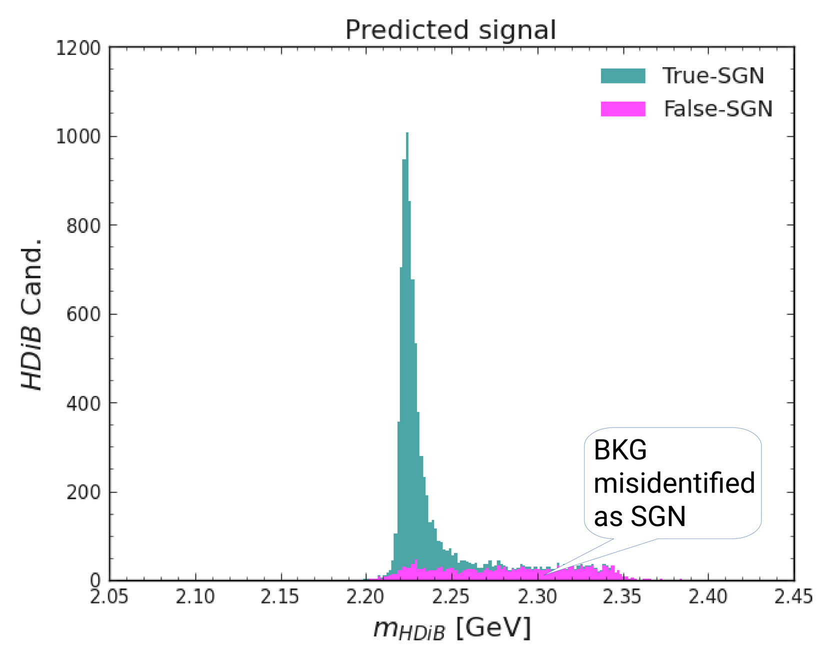
Reconstruction + BDT ~ 30.85% efficiency and **91.35%** purity

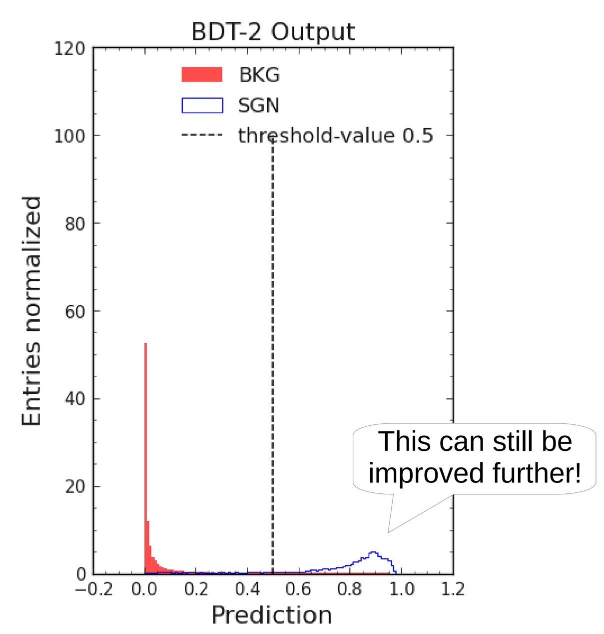
Optimizing H-Dibaryon Selection

H → [Λ → p π⁻] p π⁻

Can use selection from BDT-1!

- Mass set to 2.224 GeV
- Preprocessing to prepare data, eg. [2.15 GeV < m_{HDiB} < 2.35 GeV]
- Detector-reconstruction ~ 25.55% efficiency and 23.48% purity





- Trained on 70% of 15k generated MC – Validated on the remaining 30%
  → Using 15 'predictive' variables
- Tested on a separate 3k MC-sample
  → Reconstruction + BDT ~ 22.23% efficiency and **81.80%** purity

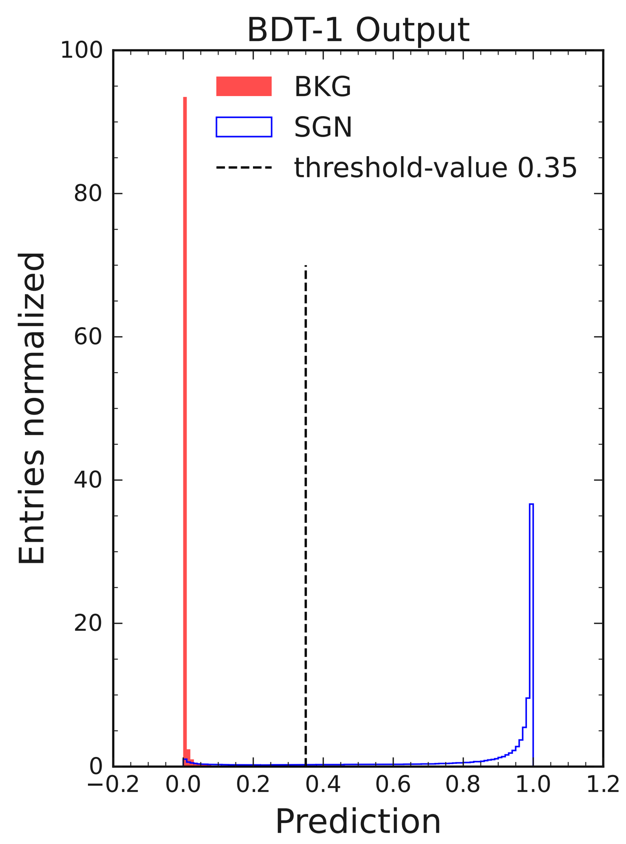
Machine Learning

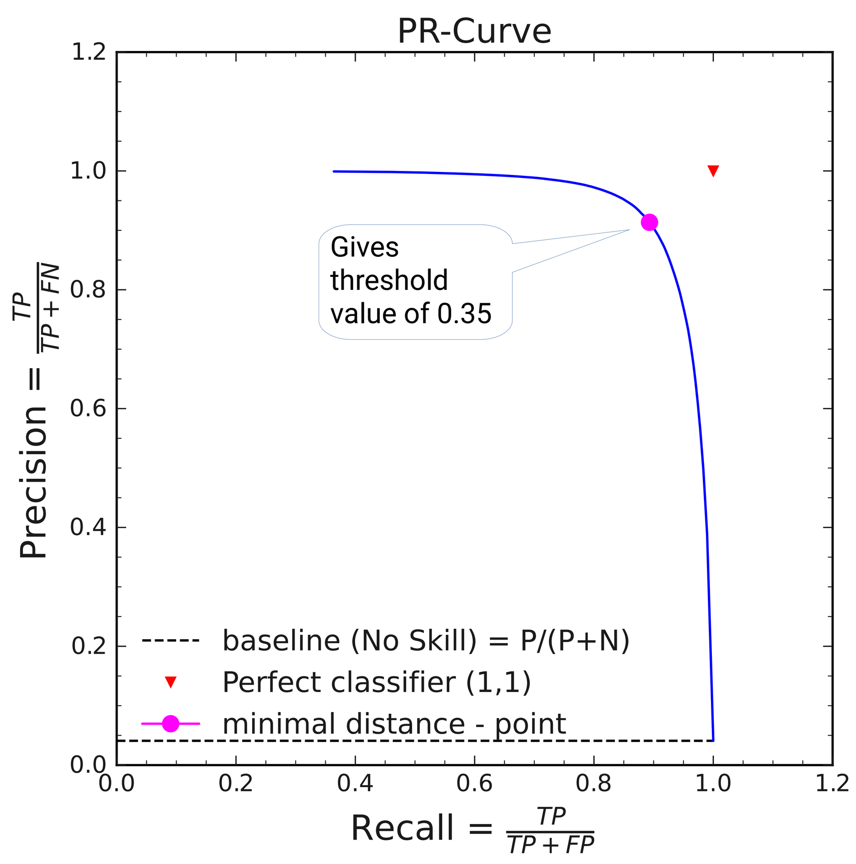
- Boosted-Decision-Tree (BDT) for signal/background separation

- Signal/background ratio is small, especially from continuum
  → Data is highly imbalanced
  → Precision-Recall (PR) metric prioritizes minority class (signal)
    → Perfect classifier at point (1,1)

- Predictive variables are determined using 'feature importance'
  → Make sure MC- and data-variables agree reasonably
- BDT with highest area under the PR-curve is considered best
- Threshold of the BDT-output = Point of minimal distance to (1,1)

- Hyperparameter optimization via optuna





Conclusion

- The purity was increased by more than a **factor of 22 and 3** for the Lambda and the H-Dibaryon respectively, using methods of machine learning.

- Due to the highly imbalanced (MC-)data at hand, the **PR-Curve** serves as a metric to prioritize the signal (minority class).

- To estimate the upper limits of the H → Λ p π decay channel, clean Lambda- and H-Dibaryon-lists are beneficial and look promising so far.

60th International Winter Meeting on Nuclear Physics
Bormio
Italy

Bundesministerium für Bildung und Forschung

JOHANNES GUTENBERG UNIVERSITÄT MAINZ