



Contribution ID: 47

Type: **Poster**

Recent benchmarks in the Analysis Grand Challenge and integration with Combine (and HS3)

Monday 8 September 2025 11:00 (30 minutes)

The Analysis Grand Challenge (AGC) showcases an example of HEP analysis. Its reference implementation uses modern Python packages to realize the main steps, from data access to statistical model building and fitting. The packages used for data handling and processing (coffea, uproot, awkward-array) have recently undergone a series of performance optimizations.

While not being part of the HEP Python (PyHEP) ecosystem, the Combine tool is a pillar of CMS analyses, covering more than 90% of the analyses published in the last few years. As such, it is necessary to have Combine integrated in the PyHEP ecosystem, using the AGC as example.

This project also includes, in the long-term, providing support and integration for the High Energy Physics Statistics Serialization Standard (HS3), as a way to have a language-independent way of representing the likelihood and use different frameworks interchangeably.

In this talk we will cover part of the recent work performed on the AGC and Combine, including:

- performance benchmarks, covering benefits introduced by the recent improvements in the data processing packages;
- examples of how Combine can be integrated and run in a dedicated infrastructure (coffea-casa);
- examples and plans to integrate HS3 in Combine.

Significance

The topic is important as it shows how the recent improvements in some of the packages used in data analysis impact performance. Moreover, the goal of supporting HS3 in Combine is an important step towards interoperability between fitting frameworks.

References

Experiment context, if any

CMS

Author: GALLI, Massimiliano (Princeton University (US))

Co-authors: HELD, Alexander (University of Wisconsin Madison (US)); SHADURA, Oksana (University of Nebraska Lincoln (US))

Presenter: GALLI, Massimiliano (Princeton University (US))

Session Classification: Poster session with coffee break

Track Classification: Track 1: Computing Technology for Physics Research