Contribution ID: **666** Contribution code: **contribution ID 666**   Type: **Poster**

zfit: scalable pythonic fitting

Statistical modelling and likelihood inference is a key element in many sciences,
especially in High-Energy Physics (HEP) analyses. These require advanced features
such as handling large amounts of data, supporting binned, unbinned and mixed inference, using complicated
and often custom made model functions, and being highly performant.
In HEP, these features were covered in C++ frameworks such as ROOT/RooFit.
With the recent shift towards Python's scientific ecosystem, the lack of a fully featured Python fitting library
became a bottleneck, as existing Python libraries do not cover all the needs of typical HEP analyses; zfit was
created to fill this gap.

In this talk we will cover zfit, a likelihood fitting library in Python which covers most of the HEP needs
(composite models, uncertainty treatments, numerical integration and sampling methods, as well as straight-
forward ways to implement custom models, to name a few) in a simple and intuitive way.
A key point in zfit is to provide pythonic ease-of-use while having C++ like speeds and using modern ac-
celerators such as GPUs flexibly without modifications. This is achieved by making use of large scale ma-
chine learning libraries (mainly TensorFlow), since they provide high performance computing capabilities
with Numpy-like interfaces.
Therefor, we will also discuss zfits technical aspects, from the immense benefit that building on top of Ten-
sorFlow offers to its limitations. We will highlight how modern high performance computing libraries that
are built for big data analysis can be used to create computational scientific libraries with the ease-of-use of
Python and the same or better performance than their C++ counterparts on CPU and GPU.

References

CHEP 2019: https://doi.org/10.1051/epjconf/202024506025
Original paper: https://doi.org/10.1016/j.softx.2020.100508

Speaker time zone

Compatible with Europe

Significance

A preliminary version of zfit was already presented in the past introducing the
project and its capabilities to do unbinned fits.
These account for about
one third to half of the desired total scope.
Now, zfit extended its functionality significantly to include the two remaining main features: binned fits as
well as
combined binned and unbinned fits;
both are crucial for many HEP analysis.
Furthermore, zfit started with an early TensorFlow version which relied heavily on a graph based,
cumbersome computation model.
With significant changes in TensorFlow, zfit was largely rewritten to use this new and

more general computing model that compiles functions just-in-time and is used by other high performance libraries (numba, jax,...) as well. This allows for more general insights into the usage of this libraries and allows comparisons between them.

Primary authors: ESCHLE, Jonas (Universitaet Zuerich (CH));   SILVA COUTINHO, Rafael (Universitaet Zuerich (CH));   SERRA, Nicola (Universitaet Zuerich (CH));   MARINANGELI, Matthieu (EPFL - Ecole Polytechnique Federale Lausanne (CH));   PUIG NAVARRO, Albert (Universität Zürich (CH))

Presenter: ESCHLE, Jonas (Universitaet Zuerich (CH))

Session Classification: Posters: Orange

Track Classification: Track 2: Data Analysis - Algorithms and Tools