A Python Package with Novel Raw Data Analysis Methods for Cryogenic Particle Detectors

Felix Wagner, D. Bartolot, J. Burkhart, D. Rizvanovic, D. Schmiedmayer
Institute of High Energy Physics, Austrian Academy of Sciences

cait

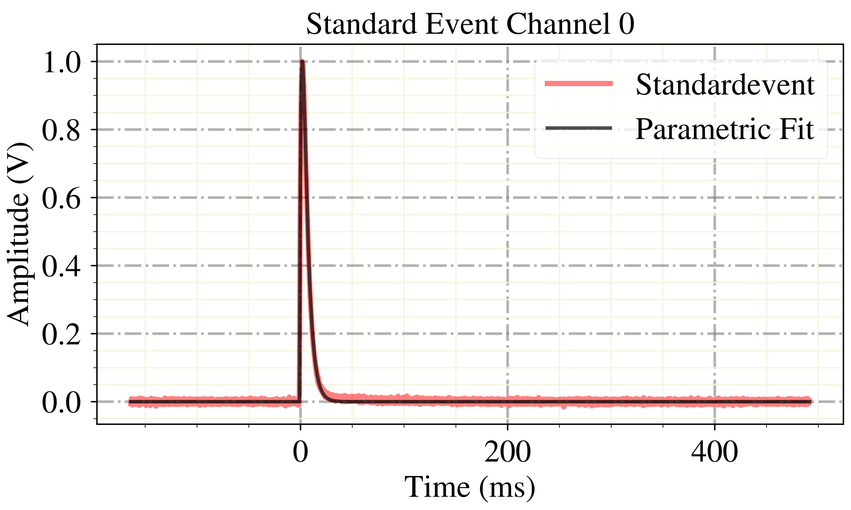
Getting Started

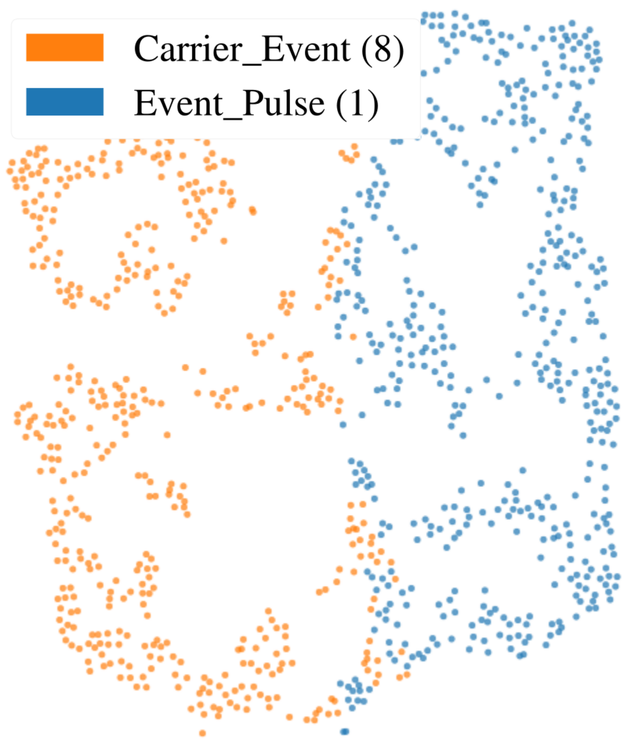
Novel cryogenic scintillating calorimeters achieve sub-keV recoil energy thresholds. Such low thresholds require a sensible raw data analysis of triggered events, to identify different types of particle recoils and artifacts despite the low signal-to-noise ratio, and reconstruct the corresponding recoil energy. We present for this purpose the Python package `cait` (Cryogenic Artificial Intelligence Tools, [1-4]), which utilizes new methods from data science and machine learning.

```
[1]: import cait as ai

dh = ai.DataHandler(nmbr_channels=1)
dh.set_filepath(path_h5='./',
                fname='dataset')

dh.calc_mp()  # main parameters
dh.calc_nps()  # noise spectrum
dh.calc_sev()  # standard event
dh.calc_of()  # optimum filter
dh.show_sev(channel=0)
```



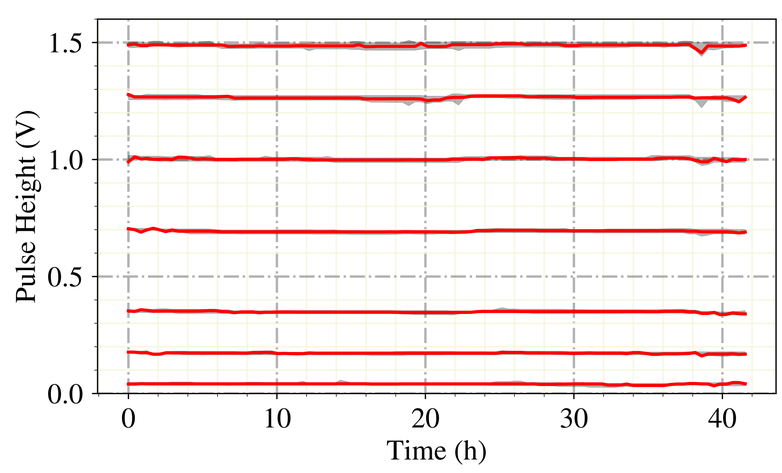


Open Source Release v1.0

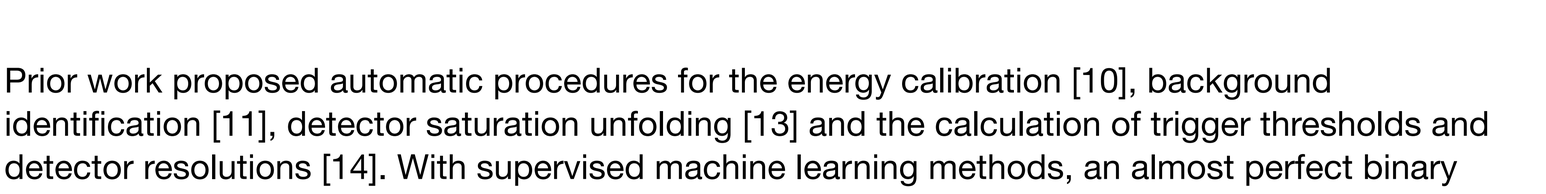
`cait` is developed open source and available on the Python Package Index [3]. Although initially meant for efficient prototyping of new analysis methods, its latest `v1.0` release renders it a fast and production-ready package for the whole analysis process. The package is currently tailored to the needs of the CRESST and COSINUS dark matter searches. Extensions for experiments with similar, time series-like, data formats are possible.

The implemented features include an interface for the user-friendly labeling of data for supervised models, the visualization of data distributions ([15], middle column) a range of event simulation tools for data augmentation, tailored data sets and data modules for Scikit-Learn, PyTorch and PyTorch lightning [5-7], as well as standard methods for processing of standard events (left column), fits, plots, triggering and the energy calibration (below).

The flexible HDF5-based data structure [8] enables integration with most Python scientific computing packages, which is convenient for fast prototyping of new analysis methods. The extensive documentation, tutorials and template files [4] provide a cornerstone for the onboarding of new users.



Towards A Fully Automated Analysis



Prior work proposed automatic procedures for the energy calibration [10], background identification [11], detector saturation unfolding [13] and the calculation of trigger thresholds and detector resolutions [14]. With supervised machine learning methods, an almost perfect binary discrimination accuracy between two types of simulated recoil events with energies slightly above detector resolution was achieved [9]. For an imbalanced multi-class classification problem of different recoil events and artifacts, accuracies above 90% were achieved [12]. Training on a large data set of augmented recoil events and artifacts might lead to a universal event selection model, which is the topic of ongoing research. Putting all these methods together reduces the human interaction, and human bias, to a minimum. `cait` holds implementations of all mentioned raw data analysis methods and is therefore a promising candidate to achieve full automation of the analysis chain.

References

[1] F. Wagner et. al. (2021), fewagner/cait: v1.0.0, http://doi.org/10.5281/zenodo.5091416
[2] Cait GitHub Repository, https://github.com/fewagner/cait
[3] cait 1.0.0 on PyPi, https://pypi.org/project/cait/
[4] Cait Documentation, https://cait.readthedocs.io/en/latest/
[5] Scikit-Learn Homepage, https://scikit-learn.org/stable/
[6] PyTorch Homepage, https://pytorch.org/
[7] PyTorch Lightning Homepage, https://www.pytorchlightning.ai/
[8] HDF5 for Python Homepage, https://www.h5py.org/
[9] F. Wagner (2020), Machine learning methods for the raw data analysis of cryogenic dark matter experiments, https://doi.org/10.34726/hss.2020.77322
[10] F. Reindl (2016), Exploring Light Dark Matter With CRESST-II Low-Threshold Detectors, http://mediatum.ub.tum.de/?id=1294132
[11] D. Schmiedmayer (2019), Calculation of dark-matter exclusions-limits using a maximum Likelihood approach, https://resolver.obvsg.at/urn:nbn:at:at-ubtuw:1-122743
[12] D. Bartolot, D. Rizvanovic (2021), Unpublished Projects
[13] M. Stahlberg (2020), Probing low-mass dark matter with CRESST-III : data analysis and first results, https://doi.org/10.34726/hss.2021.45935
[14] M. Mancuso, A method to define the energy threshold depending on noise level for rare event searches, https://doi.org/10.1016/j.nima.2019.06.030
[15] L. v. d. Maaten (2008), Visualizing Data using t-SNE, https://www.jmlr.org/papers/volume9/vandermaaten08a/vandermaaten08a.pdf

Institutional Partners

TU WIEN TECHNISCHE UNIVERSITÄT WIEN Vienna | Austria

HEPHY INSTITUT FÜR HOCHENERGIEPHYSIK

FFG Forschung wirkt.