

ETP B-group meeting

## **MICHAEL'S CURRENT WIP**

2019-02-28

Michael Eliachevitch | ETP – KIT

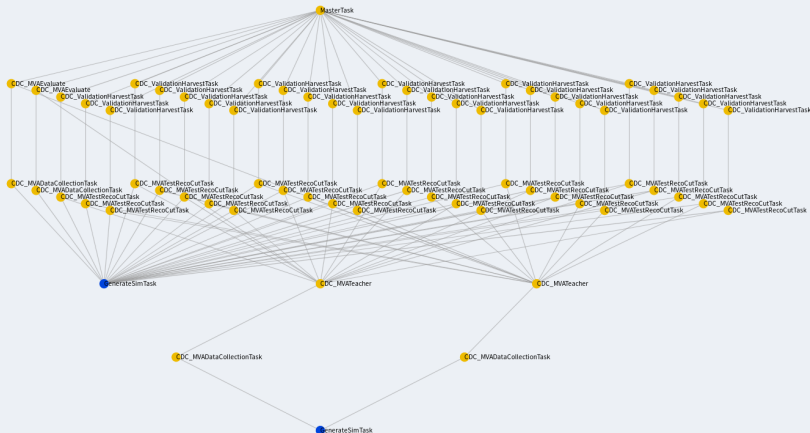
# RETRAINING AND REVALIDATING MASTER THESIS RESULTS VIA B2LUIGI

luigi dependency graph

MasterTask()

Dependency Graph

- Failed
- Running
- Batch Running
- Pending
- Done
- Disabled
- Unknown
- Truncated

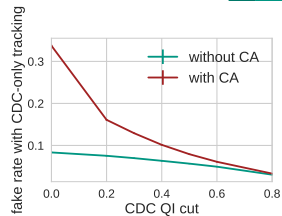
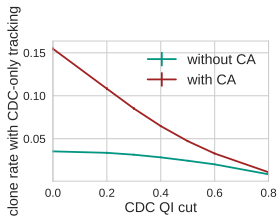
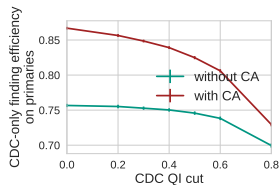


# RETRAINING AND REVALIDATING MASTER THESIS RESULTS VIA B2LUIGI

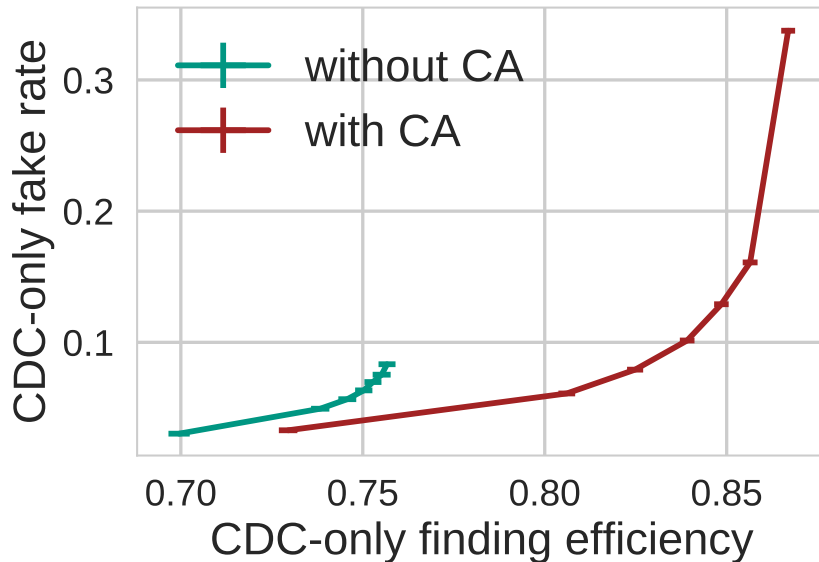
- I needed to retrain the old weightfiles for the CDC QI
- my old scripts for training and validation needed to be updated (they were a mess)
- with my experience from the semitaonic hackathon, I decided to refactor and port them to `b2luigi`
- it was easy to do, the reproducibility helps me to trust my own results and showed that my past results still hold
- will definitely use that more, e.g. for testing the full quality indicator and of course in analysis

# REDID THESIS EFFICIENCY PLOTS WITH EFFICIENCIES VIA UNCERTAIN\_PANDA

**Warning:** here the finding efficiency is with CDC-only tracking, but normalized to primary MC tracks which have hit anywhere in the detector (e.g. VXD only). The jobs with the correct MC matching for the CDC are still running.



# ROC CURVE



# GET GOING WITH THE MERGE REQUESTS

## ■ started with CDC QI (Link)

Feature/[BII-3423](#) qualityindicator mva for cdc tracks

Overview Diff Commits

### Details

 [Michael Eliachevitch](#) created a pull request 1 hour ago

I treat this as a WIP pull request, because I think that there is a lot that could be discussed, in particular because I had procrastinated that for forever, but here it finally is. Sorry for the messy commit history.

### New module, classes and other code changes

This PR introduces the `TFCDC_TrackQualityEstimator` module for an MVA estimation of the "quality" of a CDC track and for storing the MVA output in the RecoTracks exported from the CDC track finding. It is very similar and based on the `TFCDC_TrackRejecter` and uses the same variable set for training (improved in a [separate PR](#)).

However, it uses a different truth target, where the tracks which are clones are also treated as not signal. Therefore, I defined the best-matched track as the matched track with the highest number of matched hits. Therefore, I added the `CDCMCCloneLookup` singleton class, which in contrast to the other CDC MC lookup classes aggregates information of all tracks in an event. Therefore, in each event it has to be "filled" via the new `CDCMCCloneLookupFiller` class.

I already added the module to the default reconstruction in `add_cdc_track_finding`. It shouldn't change anything however, since it does not delete tracks in the default configuration, just sets their quality indicator. This is given further to the `"CDCTrackVector"` store array via the updated `TFCDC_TrackExporter` module.

Most other code changes are of cosmetic nature, when I worked on some classes and found things that might be improved (e.g. correct singleton constructor removal).

### Weight files

I have two weightfiles currently, in addition to the one trained for the current tracking chain one weightfile trained with the cellular automaton track finding enabled. The main reason is that I had tested enabling that track finding and using this quality estimator as a filter in the reconstruction can be used to bring the fake- and clone-rate below the level of the current CDC reconstruction chain while still keeping a net gain in finding efficiency.

The weight-files are currently part of the PR and in the `tracking/data` directory instead of their own payloads. I think this is okay since prior to the merge of the full track quality indicator (which should follow soon) the CDC track quality indicator is only for developers and of no physics-importance yet. But if people say that the git index should never see any data files, I can create a TUPPR.

### Possibilities for code quality improvement

- I feel that there is some probably avoidable code duplication here between the `TrackRejecter` and the `TrackQualityEstimator` for the CDC.

 1 build 

 [BII-3423](#)

 [Learn more](#)

(I shouldn't be so nervous about that)

# OTHER NEWS

- enjoyed some of the talks from the KSETA workshop
- flavor-physics was presented from theory-side by Marta and Danny van Dyk
- how do I delete a talk from the ekp talk page?