Karlsruher Institut für Technologie

Steps towards a Track Quality Indicator from CDC Tracking

Roadmap from Curler Clone Rejection to a full Quality Indicator

Michael Eliachevitch | 10 February 2018

ETP - KIT

Reminder: Track Quality Indicator

- Goal: Track quality estimator module, which generates a single number: quality indicator (QI)
  - encodes probability that track is correctly matched
  - store QI in tracks and give it to analysists
    $\rightarrow$ do cuts on QI to get wanted efficiency vs. purity
- Quality Estimator trained using MVA-package on features from:
  - quality indicators provided by track finders: VXDTF2, **CDC**, CKF
  - fitted tracks: track parameters, hit patterns
  - merger information
  - ...
- Overview talk given by Felix Metzner at B2GM and talk on studies of the full track estimator by Sebastian Racs at tracking meeting last week
- **CDC Track Quality Indicator**: needed for full QI
  $\rightarrow$ not implemented yet, will be done by me

Current Status Quality Estimation in the CDC

- CDC track finding already has an MVA fake filter: `TrackRejecter`
- calculates weight, which encodes probability that track is not fake (currently means “less than 80% CDC hit purity”)
- cuts on filter weights below 0.1, otherwise store it in `CDCTrack`
- currently used at the end of `TFCDC_SegmentTrackCombiner`, could also be used seperatly with `TFCDC_TrackRejecter` module

Can the fake rejecter weight be used as the CDC quality indicator?

- Yes, just needs to be stored in the corresponding `RecoTrack` object.
- Keep the cut?
- Problem: Does not encode probability if track is clone. But is that needed at reconstruction level? Possible?
- Further inputs for final CDC QI?

Local CDC Track Finding

- two track finding algorithms for CDC standalone tracking available:
  - global track finding with legendre algorithm
  - local track finding with cellular automaton (CA)
- currently, only global track finding used (and CA in segment finder)
- full CA track finding implemented by Oliver Frost:
  - creates tracks by combining axial-stereo segment pairs
  - finds single-segment tracks in first superlayer with > 15 hits
    $\rightarrow$ no z-information, need CKF from SVD to CDC
  - to enable: `add_cdc_track_finding(with_ca=True, ...)`
  - CA tracking runs after `SegmentTrackCombiner`
    $\rightarrow$ currently no fake filter applied to CA tracks

Enabling CA track finding leads to (as shown in the following slides)

- $\oplus$ slightly higher efficiency
- $\ominus$ higher clone rate due to curlers, where multiple loop arms are found

Rejection of clones from curlers found by CA

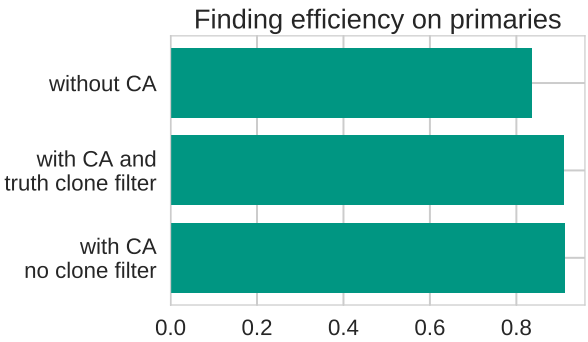
- Idea: An MVA filter can be used to detect clones from curlers
- The resulting filter weight can be used for rejecting clones or saved in track (eg. for CDC QI)
- Input variables:
  - features of the track itself: dangerous, might bias against secondaries
  - use variable that relates track to other tracks in event
- Goal: Reduce clone rate, but don't reduce finding efficiency on secondaries

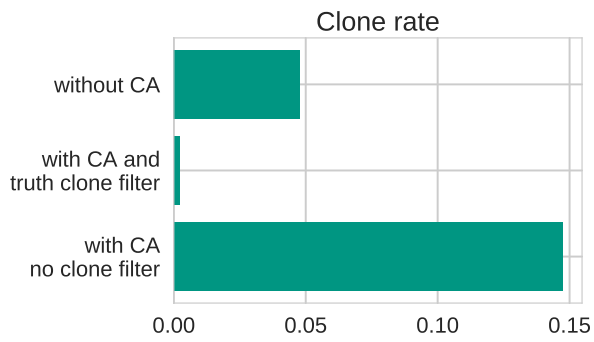
MC truth for clone filter: Clone or best match?

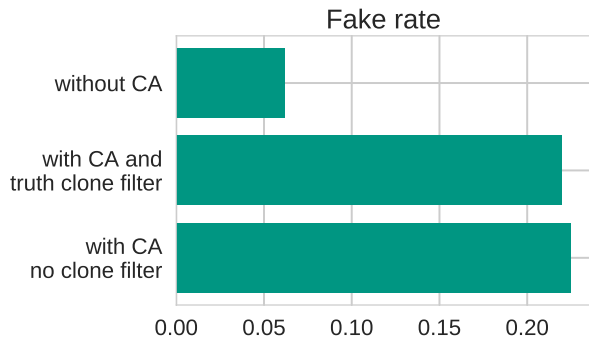
- So far, only MC truth filter implemented
- it will be used for MVA training
- current truth definition for “best match”:
  - matched track with lowest number of loops until first hit (usually 0)
  - if multiple tracks with same loop number, best match is track with highest number of matched hits
  - declare all other tracks “clone”
- other ideas for more sophisticated methods welcome

Figures of Merit

- based on 5000 MC events with phase 3 setup and background overlay from 16th campaign
- CDC-only reconstruction without CA, with CA and with CA and clone truth rejection
- **very preliminary**, don't trust exact numbers, mistakes possible . . .

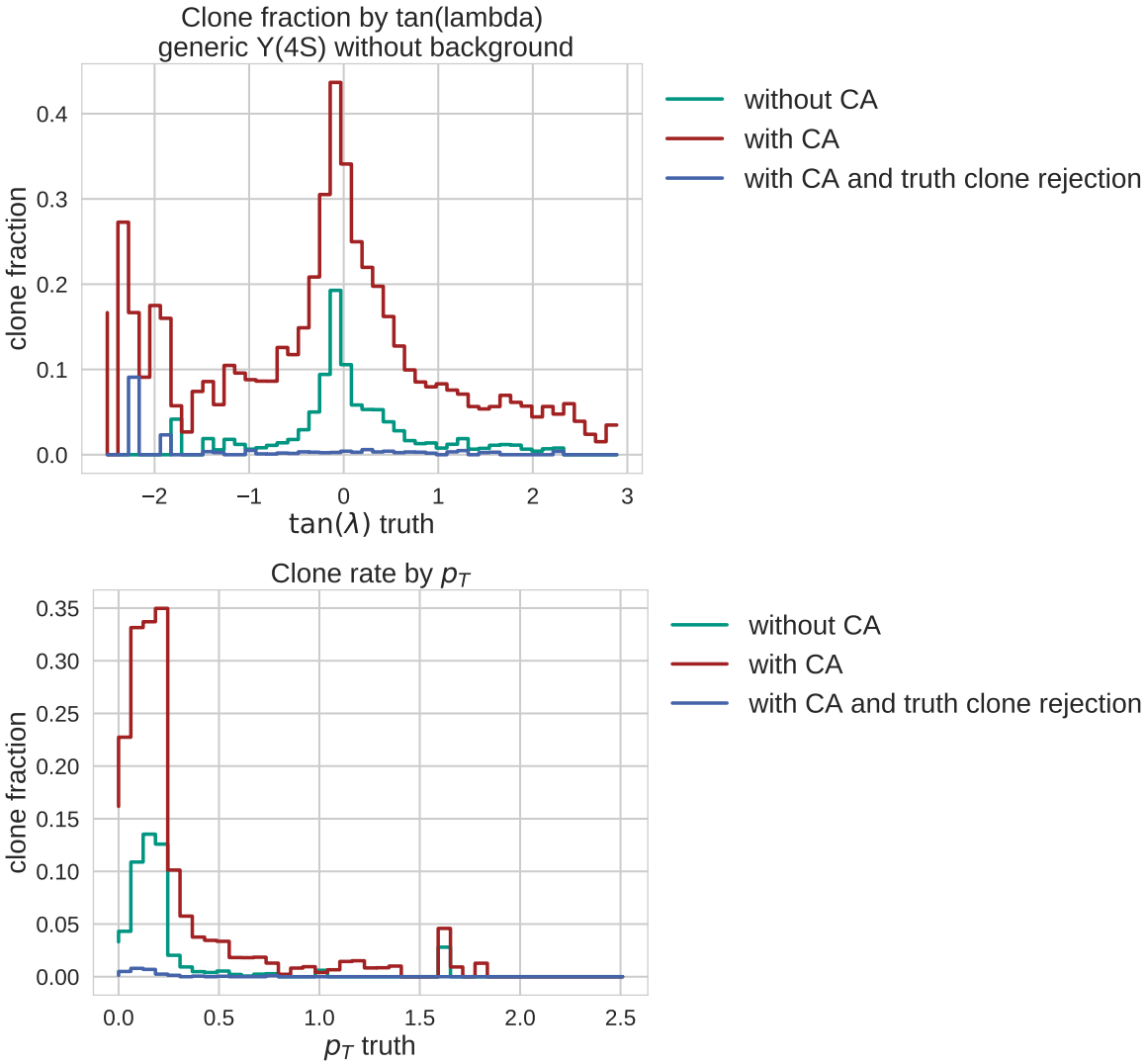






- changes to finding efficiency and clone rate as expected
- Why is fake rate so high? Maybe mistake in matching

Clone distributions



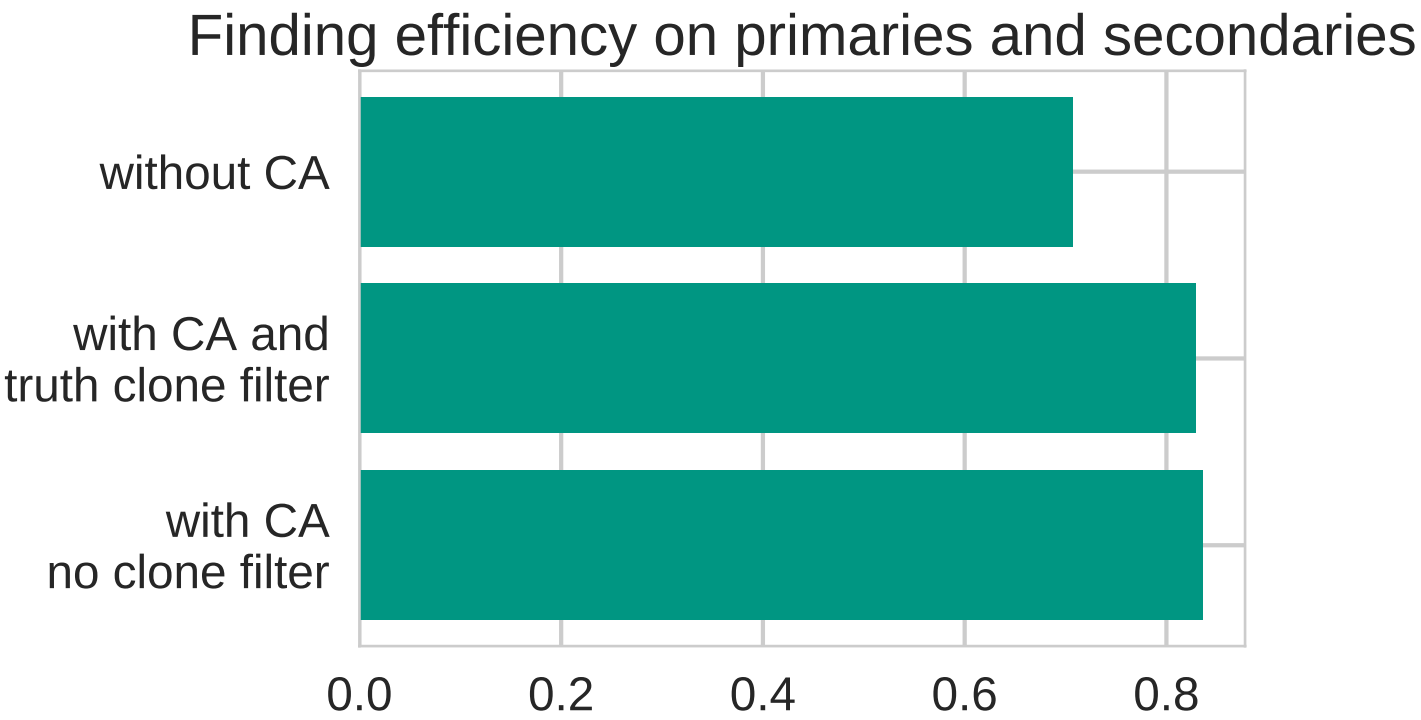
- clones come from region around $\tan(\lambda) = \cos(\theta) = 0$
- this indicates that they are indeed mostly from curlers

Conclusion and ToDo's

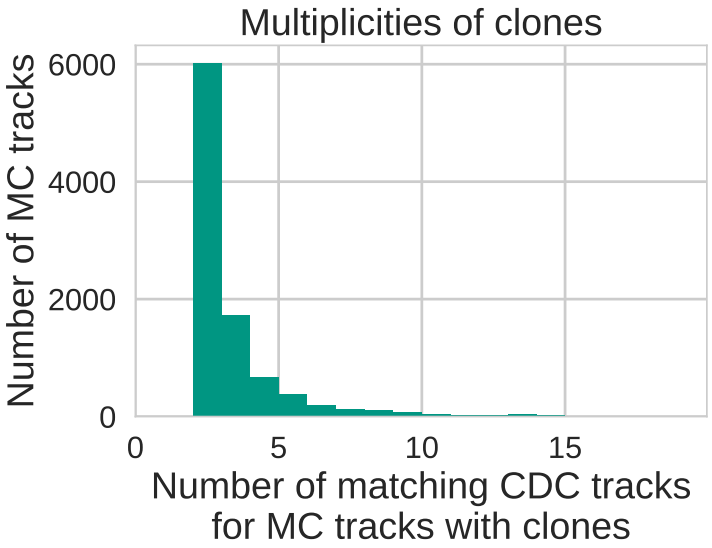
- CDC needs to provide its own quality indicator (stored in CDC `RecoTracks`)
- some functionality already in `TrackRejecter` (fake filter) module
- further studies CA track finding
  - train MVA clone filter (need to select features)
  - retrain fake filter with CA enables and apply it to CA tracks
  - check if filter work and if issues with fake rate can be solved
- combine individual filters weights into CDC quality estimator module or use them for cuts
- train full tracking quality indicator with CDC quality indicator included (first tests with fake filter weight ongoing)

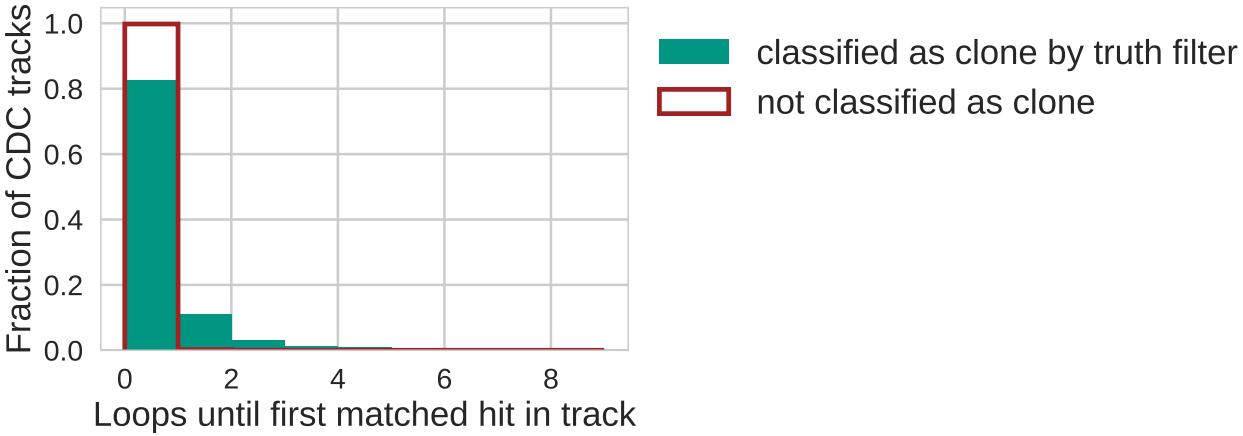
Backup Slides

Finding efficiency with secondaries

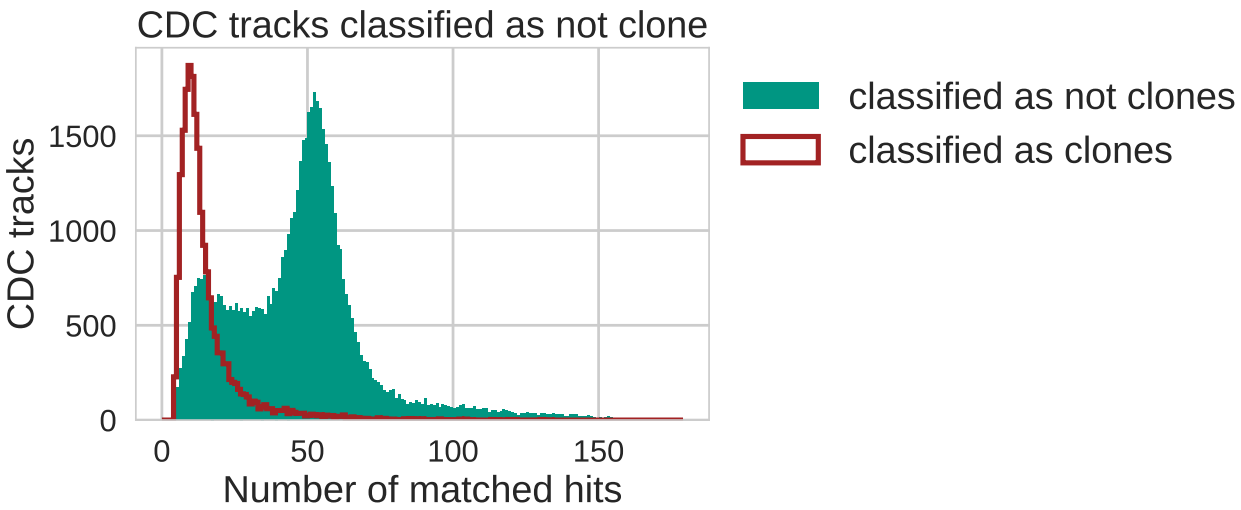


Plots from recording filter output I





Plots from recording filter output II



Features currently used for CDC Track Rejecter training

As defined in
`tracking/trackFindingCDC/filters/track/BasicTrackVarSet.h`

- size
- pt
- sz_slope
- drift_length_mean
- drift_length_variance
- drift_length_max
- drift_length_min
- drift_length_sum
- adc_mean
- adc_variance
- adc_max
- adc_min
- adc_sum
- empty_s_mean
- empty_s_variance
- empty_s_max
- empty_s_min
- empty_s_sum
- has_matching_segment
- s_range