

# First look at Cosmic CDC Tracks and Ideas for Estimating the Finding Efficiency

F2F Tracking Meeting

Michael Eliachevitch | 19 September 2018

ETP - KIT

- first cosmics data from Global Cosmic Run (GCR) 2017 available
- Warmup: Reproduction of kinematic distributions, Sim vs. MC
- Current tracking validation: Matching with MC Truth
  - ⇒ Is a data-only approach possible?
  - ⇒ confirmation of MC validation

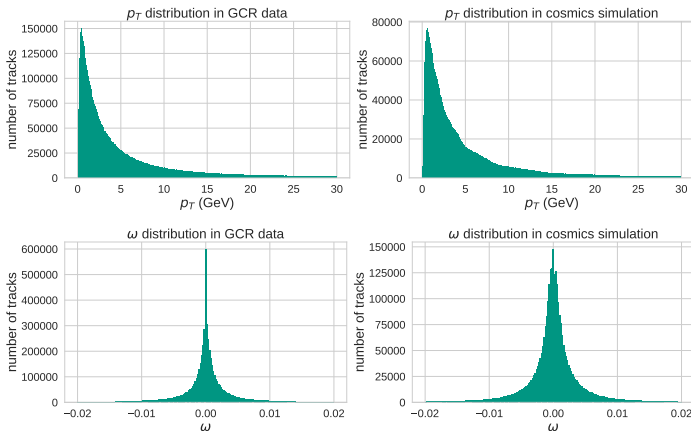
- use data from Global Cosmic Run (GCR) taken in July 2017
- use run numbers 3100–3370 (suggested by Dong Thang)  $\Rightarrow$  total 2.8 Million cosmic events with trigger selecting central tracks
- also produced 50 Million cosmic MC events with GCR setup
  - same as official MC group: large “accept box” of  $8\text{ m} \times 8\text{ m} \times 8\text{ m}$
  - no trigger in simulation, do kinematic cuts on central region ( $d_0, z_0$ )  
 $\Rightarrow \sim 10$  times less statistics than in data remain

## Links to information on data and MC production

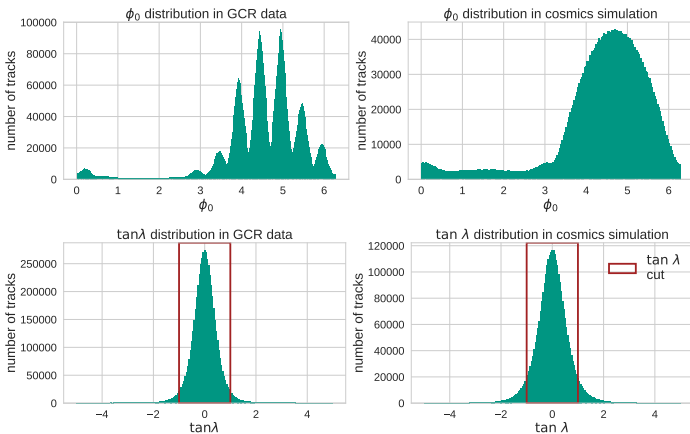
- Data:  
<https://confluence.desy.de/display/BI/Data+Production+Global+Cosmics+Run+Data#DataProductionGlobalCosmicsRunData-Runinfo>
- MC: <https://confluence.desy.de/display/BI/Data+Production+Global+Cosmics+Run+MC>

# Kinematic Distributions: Data and MC

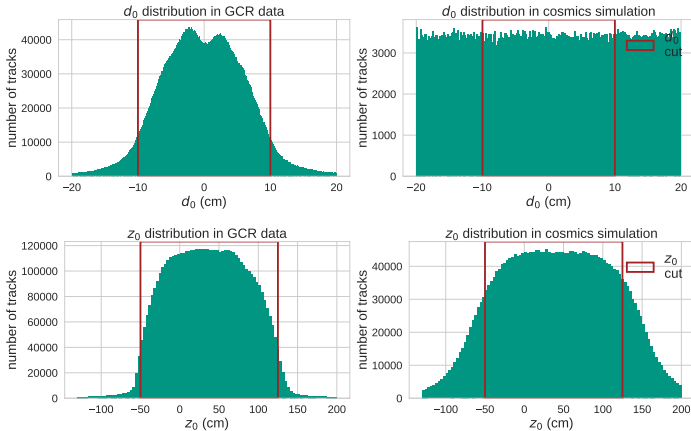
- left data (includes trigger), right MC (without trigger)
- use (preliminary) cuts for selection of central tracks (red lines)



- $p_T$  distributions seem similar, but more events with  $\omega = 0$



■ modulation in  $\phi_0$  due to varying trigger efficiencies



- distribution in MC due to lack of trigger much wider, use cuts on central region

# Idea: Cosmics-based Estimation of Finding Efficiency

- Typical event: Single muon track, usually no secondaries
- tracks passing through the central SVD volume are split
- reconstructed as two `NonMergedRecoTracks`, which are then merged to `RecoTracks`
- get estimate of finding efficiency from events where two tracks expected, but only one found (finding fails)

# Idea: Cosmics-based Estimation of Finding Efficiency

- Typical event: Single muon track, usually no secondaries
- tracks passing through the central SVD volume are split
- reconstructed as two NonMergedRecoTracks, which are then merged to RecoTracks
- get estimate of finding efficiency from events where two tracks expected, but only one found (finding fails)

$$\text{Finding efficiency} = \frac{N_{2 \text{ tracks found}}}{N_{2 \text{ tracks expected}}} = 1 - \frac{N_{1 \text{ track found}}}{N_{2 \text{ tracks expected}}}$$

where  $N_{1 \text{ track found}}, N_{2 \text{ tracks found}} \in N_{2 \text{ tracks expected}}$ , so that  
 $N_{1 \text{ track found}} + N_{2 \text{ tracks found}} = N_{2 \text{ tracks expected}}$ .



# Selection of Expected Events with two Tracks in MC and Data

- need to select events where 2 (findable) RecoTracks are expected
- this needs to be independent of actual number of tracks in event

# Selection of Expected Events with two Tracks in MC and Data

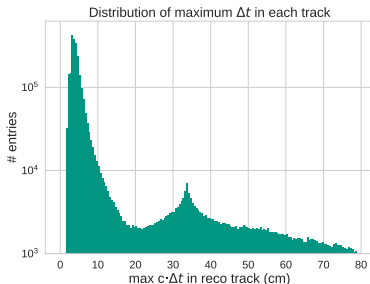
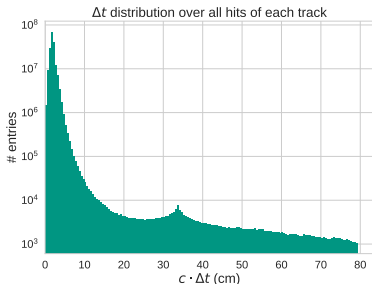
- need to select events where 2 (findable) RecoTracks are expected
- this needs to be independent of actual number of tracks in event
- use cuts:
  - kinematic cuts on  $d_0$ ,  $z_0$ , ... to select events where tracks went through SVD volume
  - cuts on hit content, e.g. minimum amount of CDC hits, hit positions?

# Selection of Expected Events with two Tracks in MC and Data

- need to select events where 2 (findable) RecoTracks are expected
- this needs to be independent of actual number of tracks in event
- use cuts:
  - kinematic cuts on  $d_0$ ,  $z_0$ , ... to select events where tracks went through SVD volume
  - cuts on hit content, e.g. minimum amount of CDC hits, hit positions?
- the choice of cuts and the assessment of the selection quality require MC truth
- **Problem:** Tracks not split in MC, only one MRecoTrack per particle

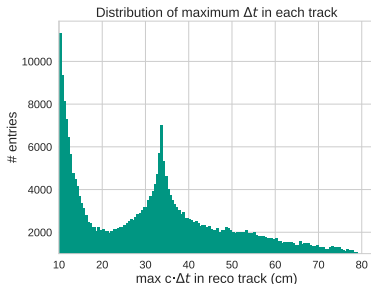
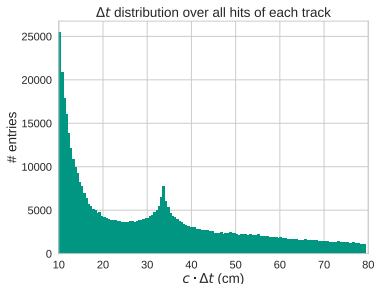
# Split MC Tracks on $\Delta t$ between Hits

- split MCRecoTrack when time between subsequent CDCSimHits is larger than chosen  $\Delta t$
- add new parameter SplitAfterDeltaT to TrackFinderMCTruthRecoTracksModule
- [pull request](#) has already been merged
- choose high enough  $\Delta t$  to only split on passing through SVD volume



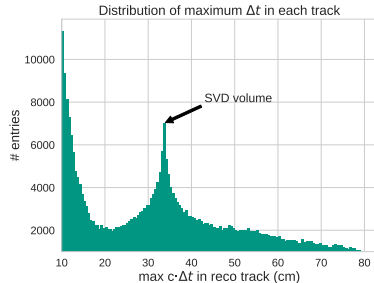
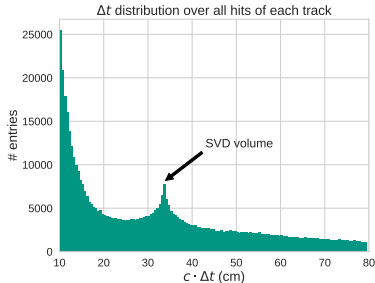
# Split MC Tracks on $\Delta t$ between Hits

- split MCRecoTrack when time between subsequent CDCSimHits is larger than chosen  $\Delta t$
- add new parameter SplitAfterDeltaT to TrackFinderMCTruthRecoTracksModule
- [pull request](#) has already been merged
- choose high enough  $\Delta t$  to only split on passing through SVD volume



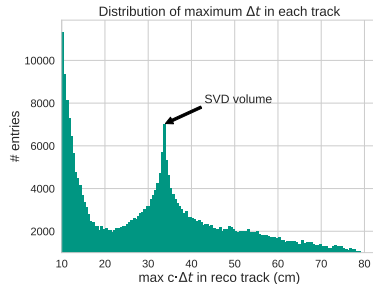
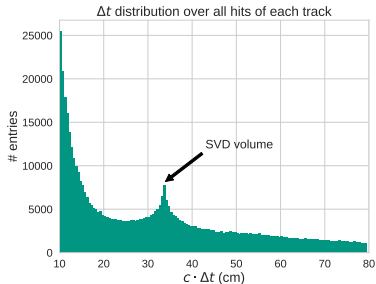
# Split MC Tracks on $\Delta t$ between Hits

- split MCRecoTrack when time between subsequent CDCSimHits is larger than chosen  $\Delta t$
- add new parameter SplitAfterDeltaT to TrackFinderMCTruthRecoTracksModule
- [pull request](#) has already been merged
- choose high enough  $\Delta t$  to only split on passing through SVD volume



# Split MC Tracks on $\Delta t$ between Hits

- split MCRecoTrack when time between subsequent CDCSimHits is larger than chosen  $\Delta t$
- add new parameter SplitAfterDeltaT to TrackFinderMCTruthRecoTracksModule
- [pull request](#) has already been merged
- choose high enough  $\Delta t$  to only split on passing through SVD volume
- method not yet used, still a **WIP**



# Naive, preliminary Cuts for testing (not from MC)

- MC truth for cut optimization and testing has just become available, not used yet
- before that, I still wanted to test the method for efficiency estimation
- previously shown kinematic cuts to select central events:  
 $|d_0| < 10 \text{ cm}, -50 \text{ cm} < z_0 < 125 \text{ cm}, |\tan \lambda| < 1$
- chose preliminary, “intuitive” cuts on hit content, based on distributions of hit content with 1/2 *found* tracks



# Naive, preliminary Cuts for testing (not from MC)

- MC truth for cut optimization and testing has just become available, not used yet
- before that, I still wanted to test the method for efficiency estimation
- previously shown kinematic cuts to select central events:  
 $|d_0| < 10 \text{ cm}, -50 \text{ cm} < z_0 < 125 \text{ cm}, |\tan \lambda| < 1$
- chose preliminary, “intuitive” cuts on hit content, based on distributions of hit content with 1/2 *found* tracks

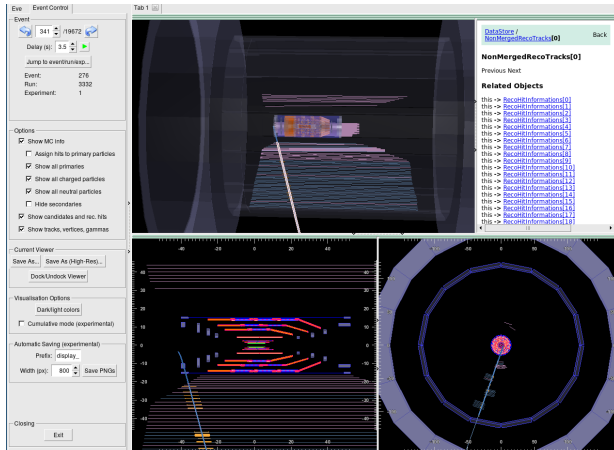
## Naive selection of candidate events for efficiency estimation

cuts on hit content for one track events only:

- # hits in events  $> 85$
- # missing CDC hits  $> 45\%$
- **Uncontrolled systematics, biased, just for testing!**

# False Finding Fail Candidate ("Background")

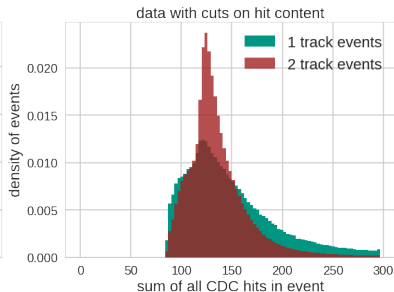
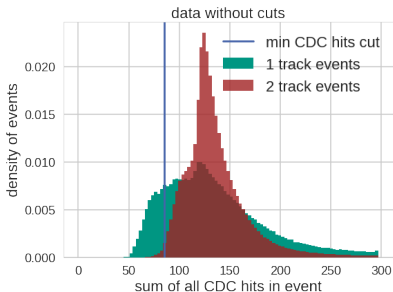
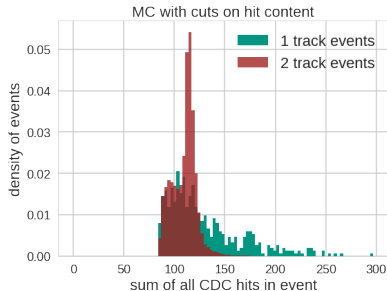
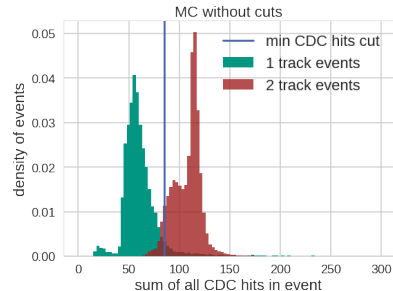
- many tracks leave acceptance after passing through SVD volume
- this event leaves enough hits in the second hough to pass through my naive cuts on hit content → need for more sophisticated methods



# WIP: Future Selection of Candidate Events

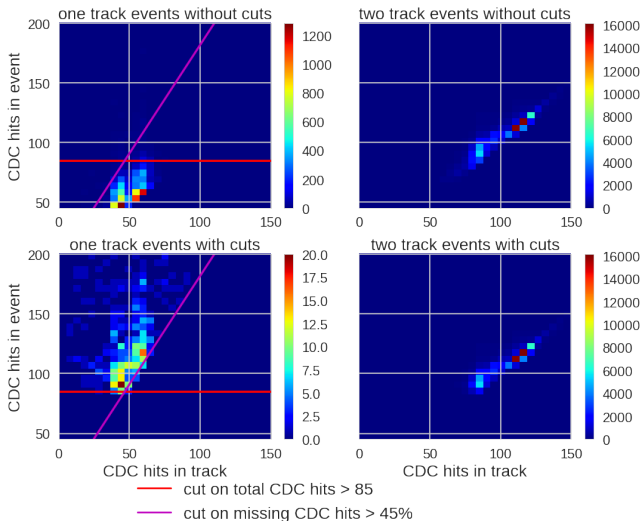
- choose cuts and analyze systematics based on MC Truth (from SplitAfterDeltaT and further constraints)
- candidate events with two expected tracks chosen without knowledge of how many tracks were found
- more sophisticated selection methods, use hit positions
  - e.g. select tracks where  $|\sum_i z_i| < \Delta z$

# Distributions of the number of CDC Hits



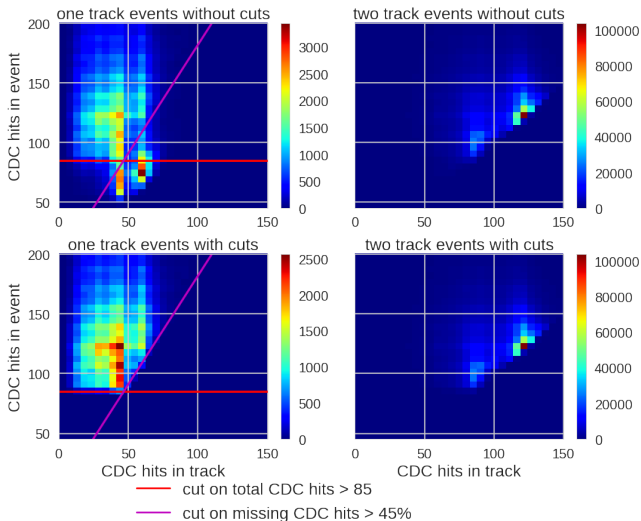
# Total CDC Hits vs. Matched CDC Hits

In MC:



# Total CDC Hits vs. Matched CDC Hits

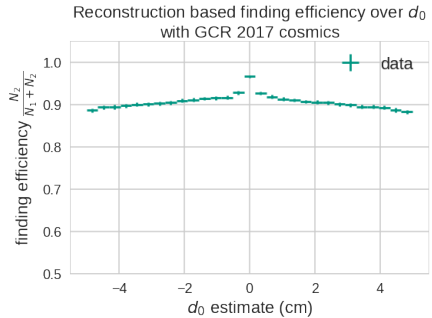
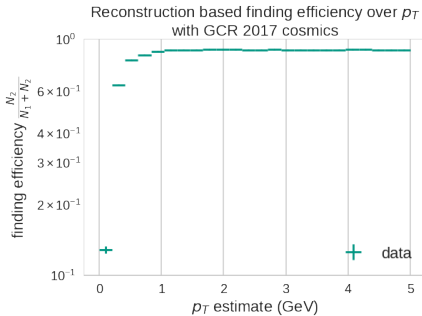
In Data:

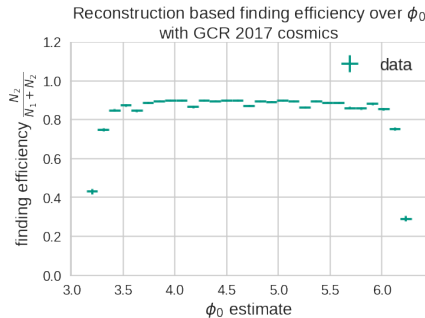
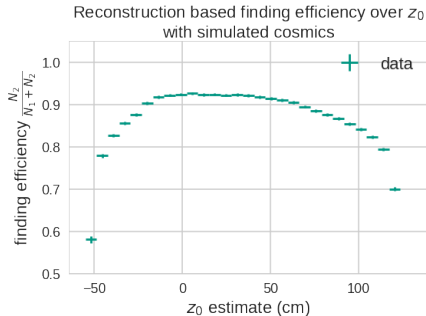
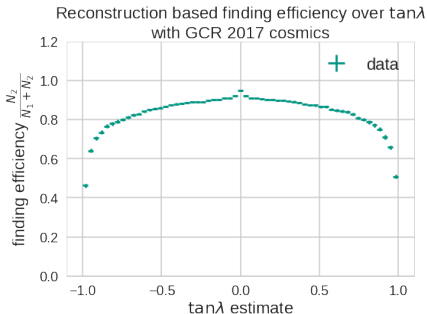


# “Efficiency” Profiles with Naive Cuts

## Reminder

Results for  $N_2/(N_1 + N_2)$  with naive cuts. Not necessarily related to actual efficiency yet!



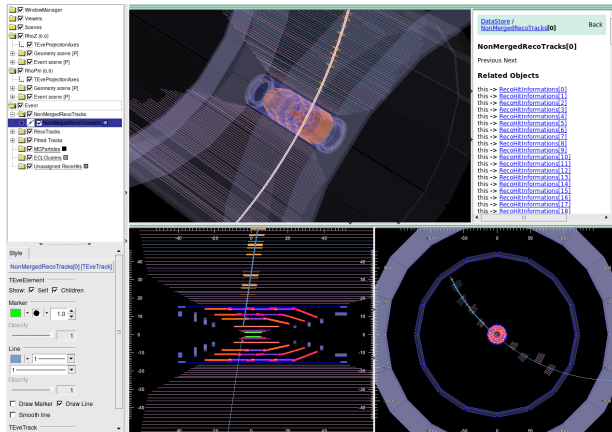




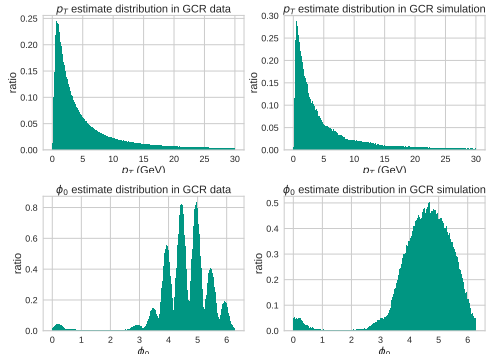
- get MC information on number of Reco Tracks with `SplitAfterDeltaT` and use it to test our methods
- with MC information, develop more sophisticated selection of events with two findable tracks
- finding efficiency profiles should be the same on MC and data
- see if efficiency profiles are similar to those of MC Matcher Truth

# Backup

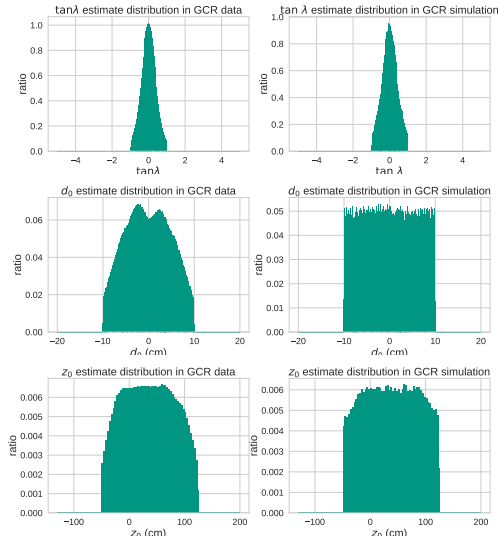
# Example Finding Fail Event



# Kinematic Distributions with Kinematic Cuts I



# Kinematic Distributions with Kinematic Cuts II

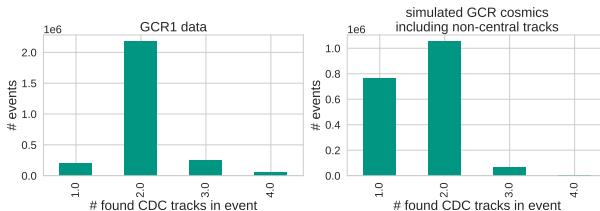


# Code for Splitting Tracks in MC Track Finder

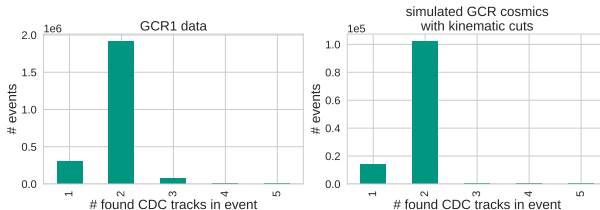
```
std::vector< std::vector<TimeHitIDDetector> > hitsWithTimeAndDetectorInformationVectors;  
  
if (m_splitAfterDeltaT < 0.0) { // no splitting, vector will only contain a single hitInformation  
    hitsWithTimeAndDetectorInformationVectors.push_back(hitsWithTimeAndDetectorInformation);  
} else { // split on delta t  
  
    std::vector<TimeHitIDDetector>::size_type splitFromIdx = 0; // whenever splitting subtrack,  
    for (std::vector<TimeHitIDDetector>::size_type i = 1; i != hitsWithTimeAndDetectorInformation.size(); i++)  
    {  
        double delta_t = (std::get<0>(hitsWithTimeAndDetectorInformation[i])  
                           - std::get<0>(hitsWithTimeAndDetectorInformation[i - 1]));  
  
        if (delta_t > m_splitAfterDeltaT) {  
            // push slice of 'hitsWithTimeAndDetectorInformation' between splitFromIdx and previous  
            hitsWithTimeAndDetectorInformationVectors  
                .emplace_back(hitsWithTimeAndDetectorInformation.begin() + splitFromIdx,  
                              hitsWithTimeAndDetectorInformation.begin() + i);  
            splitFromIdx = i;  
        }  
    }  
    // add subtrack after last splitting to list of tracks  
    hitsWithTimeAndDetectorInformationVectors  
        .emplace_back(hitsWithTimeAndDetectorInformation.begin() + splitFromIdx,  
                      hitsWithTimeAndDetectorInformation.end());  
}
```

# Track Numbers with different Cuts I

## ■ with no cuts:

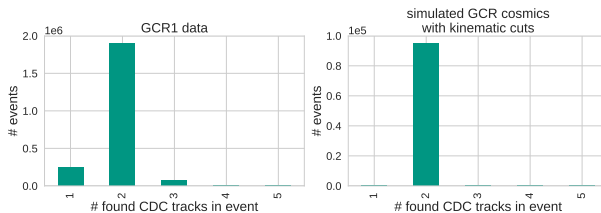


## ■ with kinematic cuts:



# Track Numbers with different Cuts II

## ■ with cuts on hit content:





# Runinfo for GCR data used by me I

Run	Date	Run time	Run Type	CDC	TOP	ECL	KLM	TRG	Trig type
3370	2017-07-13 00:07:25	1679.0	cosmic	ON	ON	ON	ON	ON	aux
3369	2017-07-12 22:58:35	4061.0	cosmic	ON	ON	ON	ON	ON	aux
3368	2017-07-12 21:58:58	3486.0	cosmic	ON	ON	ON	ON	ON	aux
3367	2017-07-12 20:52:52	3866.0	cosmic	ON	ON	ON	ON	ON	aux
3366	2017-07-12 20:28:56	1338.0	cosmic	ON	ON	ON	ON	ON	aux
3365	2017-07-12 20:15:13	697.0	cosmic	ON	ON	ON	ON	ON	aux
3364	2017-07-12 20:11:37	130.0	cosmic	ON	ON	ON	ON	ON	aux
3363	2017-07-12 19:34:08	2167.0	cosmic	ON	ON	ON	ON	ON	aux
3362	2017-07-12 19:23:05	564.0	cosmic	ON	ON	ON	ON	ON	aux
3361	2017-07-12 17:38:31	2632.0	cosmic	ON	ON	ON	ON	ON	aux
3360	2017-07-12 17:26:13	641.0	cosmic	ON	ON	ON	ON	ON	aux
3359	2017-07-12 17:23:03	65.0	cosmic	ON	ON	ON	ON	ON	aux
3332	2017-07-12 02:04:41	25316.0	cosmic	ON	ON	ON	ON	ON	aux
3330	2017-07-12 00:39:35	4918.0	cosmic	ON	ON	ON	ON	ON	aux
3329	2017-07-11 23:39:51	3343.0	cosmic	ON	ON	ON	ON	ON	aux
3328	2017-07-11 22:18:02	4808.0	cosmic	ON	ON	ON	ON	ON	aux
3327	2017-07-11 22:02:38	842.0	cosmic	ON	ON	ON	ON	ON	aux
3322	2017-07-11 19:53:51	2298.0	cosmic	ON	ON	OFF	ON	ON	aux
3321	2017-07-11 19:19:15	1991.0	cosmic	ON	ON	OFF	ON	ON	aux
3318	2017-07-11 18:50:44	1285.0	cosmic	ON	ON	OFF	ON	ON	aux
3317	2017-07-11 18:44:00	265.0	cosmic	ON	ON	OFF	ON	ON	aux
3315	2017-07-11 18:23:10	266.0	cosmic	ON	ON	OFF	ON	ON	aux
3314	2017-07-11 18:19:25	55.0	cosmic	ON	ON	ON	ON	ON	aux
3313	2017-07-11 18:09:17	518.0	cosmic	ON	ON	OFF	ON	ON	aux
3312	2017-07-11 18:06:47	56.0	cosmic	ON	ON	ON	ON	ON	aux
3311	2017-07-11 17:45:13	1232.0	cosmic	ON	ON	ON	ON	ON	aux

# Runinfo for GCR data used by me II

3310	2017-07-11 17:29:47	752.0	cosmic	ON	ON	ON	ON	ON	aux
3309	2017-07-11 17:26:44	99.0	cosmic	ON	ON	ON	ON	ON	aux
3307	2017-07-11 17:06:00	773.0	cosmic	ON	ON	ON	ON	ON	aux
3306	2017-07-11 17:02:35	63.0	cosmic	ON	ON	ON	ON	ON	aux
3295	2017-07-10 23:59:29	33091.0	cosmic	ON	ON	ON	ON	OFF	aux
3291	2017-07-10 23:40:31	524.0	cosmic	ON	ON	ON	ON	OFF	aux
3290	2017-07-10 23:23:45	824.0	cosmic	ON	ON	ON	ON	OFF	aux
3288	2017-07-10 23:17:08	123.0	cosmic	ON	ON	ON	ON	OFF	aux
3287	2017-07-10 22:30:22	2688.0	cosmic	ON	ON	ON	ON	OFF	aux
3286	2017-07-10 22:22:25	366.0	cosmic	ON	ON	ON	ON	OFF	aux
3280	2017-07-10 21:42:02	184.0	cosmic	ON	ON	ON	ON	OFF	aux
3279	2017-07-10 21:29:57	666.0	cosmic	ON	ON	ON	ON	OFF	aux
3278	2017-07-10 21:24:03	281.0	cosmic	ON	ON	ON	ON	OFF	aux
3273	2017-07-10 20:17:12	2319.0	cosmic	ON	ON	ON	ON	OFF	aux
3271	2017-07-10 19:58:58	963.0	cosmic	ON	OFF	ON	ON	OFF	aux
3270	2017-07-10 19:43:32	275.0	cosmic	ON	OFF	ON	ON	OFF	aux
3268	2017-07-10 19:10:50	562.0	cosmic	ON	OFF	ON	ON	OFF	aux
3267	2017-07-10 18:53:58	954.0	cosmic	ON	OFF	ON	ON	OFF	aux
3265	2017-07-10 18:45:50	182.0	cosmic	ON	OFF	ON	ON	OFF	aux
3264	2017-07-10 18:39:35	127.0	cosmic	ON	OFF	ON	ON	OFF	aux
3263	2017-07-10 18:36:54	50.0	cosmic	ON	OFF	ON	ON	OFF	aux
3262	2017-07-10 18:33:22	132.0	cosmic	ON	OFF	ON	ON	OFF	aux
3260	2017-07-10 18:15:01	121.0	cosmic	ON	OFF	ON	ON	OFF	aux
3218	2017-07-07 07:13:40	5602.0	cosmic	ON	ON	ON	ON	ON	aux
3217	2017-07-07 06:21:45	2972.0	cosmic	ON	ON	ON	ON	ON	aux
3216	2017-07-07 03:42:11	9420.0	cosmic	ON	ON	ON	ON	ON	aux
3215	2017-07-07 00:47:39	10347.0	cosmic	ON	ON	ON	ON	ON	aux
3214	2017-07-07 00:29:35	843.0	cosmic	ON	ON	ON	ON	ON	aux
3213	2017-07-06 23:21:17	3983.0	cosmic	ON	ON	ON	ON	ON	aux
3212	2017-07-06 23:14:30	77.0	cosmic	ON	ON	ON	ON	ON	aux

# Runinfo for GCR data used by me III

3210	2017-07-06 22:41:27	1584.0	cosmic	ON	ON	ON	ON	ON	aux
3209	2017-07-06 21:31:56	3860.0	cosmic	ON	ON	ON	ON	ON	aux
3208	2017-07-06 20:21:39	3878.0	cosmic	ON	ON	ON	ON	ON	aux
3207	2017-07-06 18:45:38	5678.0	cosmic	ON	ON	ON	ON	ON	aux
3206	2017-07-06 17:12:59	5473.0	cosmic	ON	ON	ON	ON	ON	aux
3175	2017-07-06 10:51:41	1346.0	cosmic	ON	OFF	ON	ON	ON	aux
3174	2017-07-06 09:48:58	3378.0	cosmic	ON	OFF	ON	ON	ON	aux
3173	2017-07-06 04:15:54	17627.0	cosmic	ON	ON	ON	ON	ON	trg1
3171	2017-07-06 00:40:30	1035.0	cosmic	ON	ON	ON	ON	ON	aux
3170	2017-07-05 22:42:40	6851.0	cosmic	ON	ON	ON	ON	ON	trg1
3169	2017-07-05 21:25:55	4498.0	cosmic	ON	ON	ON	ON	ON	trg1
3168	2017-07-05 20:08:21	4533.0	cosmic	ON	ON	ON	ON	ON	trg1
3167	2017-07-05 19:47:46	1111.0	cosmic	ON	ON	ON	ON	ON	aux
3165	2017-07-05 18:32:45	4234.0	cosmic	ON	ON	ON	ON	ON	trg1
3164	2017-07-05 18:21:11	618.0	cosmic	ON	ON	ON	ON	ON	aux
3163	2017-07-05 18:07:15	729.0	cosmic	ON	ON	ON	ON	ON	aux
3159	2017-07-05 16:53:46	531.0	cosmic	ON	OFF	ON	ON	OFF	aux
3158	—	0.0	cosmic	ON	OFF	ON	ON	OFF	aux
3157	2017-07-05 14:52:13	332.0	cosmic	ON	OFF	ON	ON	OFF	aux
3156	2017-07-05 14:46:51	101.0	cosmic	ON	OFF	ON	ON	OFF	aux
3120	2017-07-05 06:41:30	9146.0	cosmic	ON	ON	ON	ON	OFF	trg1
3119	2017-07-05 04:08:05	2977.0	cosmic	ON	ON	ON	ON	OFF	trg1
3118	2017-07-04 18:12	NaN	cosmic	ON	ON	ON	ON	NaN	trg1
3117	2017-07-04 17:43:11	1573.0	cosmic	ON	ON	ON	ON	OFF	aux
3116	2017-07-04 17:25:41	762.0	cosmic	ON	ON	ON	ON	OFF	aux
3114	2017-07-04 17:13:36	365.0	cosmic	ON	ON	ON	ON	OFF	aux
3112	2017-07-04 17:03:43	46.0	cosmic	ON	ON	ON	ON	OFF	aux
3100	2017-07-04 03:38:42	20012.0	cosmic	ON	OFF	ON	ON	OFF	aux

# “Efficiency” Profiles including MC

## Reminder

Approximation of trigger in MC is not well implemented yet, thus such different results. WIP.

