

Investigative Steps Taken to Address DTrackTimeBased Result Inconsistencies Between JANA1 and JANA2

Event of focus:

15th event by event number is used from **hd_rawdata_121120_000.evio** data file for all this analysis.

hd_dump of event 15 with -DDTrackTimeBased shows missing cdc hits in JANA2 that are present in JANA1:

cdc hits are getting lost somewhere in JANA2 but are present in JANA1 for candidate 2. See the candidate 2 ndof in the DTrackTimeBased difference here <https://www.diffchecker.com/vlkvqcjL/>, the values of ndof should stay the same for a particular candidate. However, in case of JANA2, for candidate 2, it is 5 for 4 mass hypothesis and 17 for 5th mass hypothesis while for JANA1 it is 17 for all the 5 mass hypothesis of candidate 2. Moreover, the comparison of the hd_dump of event 15 along with listing associated objects for DTrackTimeBased also shows that for JANA1 there are 12 CDC Hits that are getting lost in JANA2. Compare the rows 1,3 and 4 of associated objects for event 15 both for JANA1 and JANA2 to see the difference. The link to the outputs and their difference <https://www.diffchecker.com/WD8TpWGN/> .

How to Reproduce?

To get this output one just need to run following command for both halld_recon with JANA1 and JANA2 and compare the results

```
hd_dump -PEVIO:NTHREADS=1 -a -s -q15 hd_rawdata_121120_000.evio  
-DDTrackTimeBased
```

Here is a short glimpse of the output too:

```
DCDCTrackHit 0x7fe46f7296c0  
    0x7fe46f7297a0  
    0x7fe46f729880  
    0x7fe46f729960  
    0x7fe46f72b100  
    0x7fe46f72b1e0  
    0x7fe46f72b2c0  
    0x7fe46f72b480  
    0x7fe46f72b560
```

```
0x7fe46f72b640
0x7fe46f72b720
0x7fe46f8a7950
DFDCPseudo 0x7fe47016dc70
0x7fe472f79df0
0x7fe472fe1a20
0x7fe472fe22c0
0x7fe4748fb950
```

Printing value of chi2_hit, cdc_chi2cut and k:

A print statement was added to `/src/libraries/TRACKING/DTrackFitterKalmanSIMD:4698` just after the initialization of `chi2_hit`:

```
std::cout<<std::endl<<"chi2_hit: "<<chi2_hit<<", cdc_chi2cut:
"<<cdc_chi2cut<<",k: "<<k<<:std::endl;
```

Now the code was rebuilt using the following instructions:

```
source /group/halld/Software/build_scripts/gluex_env_boot_jlab.csh
gxenv halld_recon_prereqs_version.xml
cd src
nice scons install -j32 //for normal building
nice scons install -j32 DEBUG=1 OPTIMIZATION=0 SHOWBUILD=1 // for
debugging efficient
```

Once building was done now following command was run:

```
JANA1:
hd_root -PEVENTS_TO_KEEP=20 -PPLUGINS=monitoring_hists
-PEVIO:NTHREADS=1 ../hd_rawdata_121120_000.evio
JANA2:
hd_root -Pjana:nevents=20 -PPLUGINS=monitoring_hists
-PEVIO:NTHREADS=1 ../hd_rawdata_121120_000.evio
```

The comparison of results generated by these led to following observations:

196 extra repetitions are happening for JANA1 that are not there for JANA2. The outputs difference can be seen here <https://www.diffchecker.com/QGQas6h4/>. Some important observations done about these differences are following:

k: 39 → very important number because all the refits (in the repetitions) that are

happening for JANA1 but not for JANA2 are starting from k:39 and then moving forward. But as for JANA2 k:39 is never showing again which shows that it is removed, thus jumping back to this hit is not possible, hence no refits are happening. To understand this further search k: 39 in the search bar on this link <https://www.diffchecker.com/QGQas6h4/>

JANA1: chi2_hit: 26.6706, cdc_chi2cut: 25,k: 39 → Appearing again in JANA1 although chi2_hit>cdc_chi2cut
JANA2: chi2_hit: 26.6706, cdc_chi2cut: 25,k: 39 → Not appearing again in JANA2 because chi2_hit>cdc_chi2cut

However, looking at other numbers we can see that JANA2 is not always removing the hit when chi2_hit>cdc_chi2cut. Both JANA1 and JANA2 are sometimes removing it and other times not while the right behavior would be to remove it and never show a k again once the condition for it is not satisfied. Examples below dig deeper into this:

k: 34 → Not appearing again both for JANA1 and JANA2 as chi2_hit>cdc_chi2cut

JANA1: chi2_hit: 62.6524, cdc_chi2cut: 25,k: 101
JANA2:chi2_hit: 62.6524, cdc_chi2cut: 25,k: 101

k: 108 → Appearing again both for JANA1 and JANA2 although chi2_hit>cdc_chi2cut

JANA1: chi2_hit: 110.456, cdc_chi2cut: 25,k: 108
JANA2: chi2_hit: 110.456, cdc_chi2cut: 25,k: 108

k: 99 → Appearing again in JANA1 event after chi2_hit>cdc_chi2cut for k=99
JANA1: chi2_hit: 84.9194, cdc_chi2cut: 25,k: 99
JANA2: chi2_hit: 84.9194, cdc_chi2cut: 25,k: 99

k: 95 → Not appearing again after chi2_hit>cdc_chi2cut for both JANA1 and JANA2
chi2_hit: 119.789, cdc_chi2cut: 25,k: 95
chi2_hit: 119.789, cdc_chi2cut: 25,k: 95

Debugging inside GDB:

1. Print statement added to the code to print current event number :

A print statement is added src/libraries/TRACKING/

[DTrackTimeBased_factory.cc:298](#) inside the event function right after myevt variable initialization:

```
jerror_t DTrackTimeBased_factory::evnt(JEventLoop *loop, uint64_t
eventnumber)
{
  // Save event number to help with debugging
  myevt=eventnumber;
  std::cout<<std::endl<<std::endl<<std::endl<<"I am event:"<<myevt;
```

This was done to confirm the event numbers that have already been processed even though we are not stopping at that. After adding this code line, we had to building again. For which following instructions were followed:

Building Halld_recon:

```
source /group/halld/Software/build_scripts/gluex_env_boot_jlab.csh
gxenv halld_recon_prereqs_version.xml
cd src
nice scon install -j32 //for normal building
nice scon install -j32 DEBUG=1 OPTIMIZATION=0 SHOWBUILD=1 // for
debugging efficient
```

2. Inside GDB:

To jump directly to event 15 and perform the deeper analysis at that event these instructions were followed in gdb:

JANA1

First set this break point

```
break halld_recon_jana1/halld_recon/halld_recon/src/libraries/TRACKING/
DTrackTimeBased\_factory.cc:290 if myevt == 15
```

Now run this command

```
run -PPLUGINS=monitoring_hists -PEVIO:NTHREADS=1 ../
hd_rawdata_121120_000.evio
```

Once above breakpoint is reached, add these two breakpoints and start your analysis

```
break src/libraries/TRACKING/DTrackFitterKalmanSIMD.cc:4697 → As from
hd_dump result it is now proven that problem is with cdc hits so the analysis was
focused on these hits only by setting the breakpoint that is supposed to set once
event 15 is reached to cdc_hit outlier setting only.
```

JANA2

First set this break point

```
break src/libraries/TRACKING/DTrackTimeBased\_factory.cc:298 if myevt == 15
```

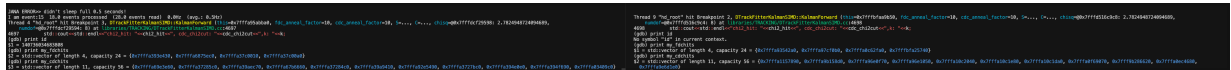
Now run this command

```
run -PPLUGINS=monitoring_hists -PEVIO:NTHREADS=1 ../  
hd_rawdata_121120_000.evio
```

Once above breakpoint is reached, add these two breakpoints and start your analysis

```
break src/libraries/TRACKING/DTrackFitterKalmanSIMD.cc:4698
```

After adding the last breakpoint, `continue` was entered as instruction in gdb terminal and first time this breakpoint was hit we got following values for id, my_fdchits, my_cdchits:



Now the breakpoint set in the file src/libraries/TRACKING/DTrackFitterKalmanSIMD.cc: at code line `if chi2_hit<cdc_chi2cut` inside KalmanForward was deleted and new breakpoints were set which were conditioned now to stop at if statement only if the condition is not met that is, chi2_hit>cdc_chi2cut:

JANA1:

```
break src/libraries/TRACKING/DTrackFitterKalmanSIMD.cc:4697 if  
chi2_hit>cdc_chi2cut
```

JANA2:

```
break src/libraries/TRACKING/DTrackFitterKalmanSIMD.cc:4698 if  
chi2_hit>cdc_chi2cut
```

Continuing this time and hitting the above breakpoint for the first time gave us following results:



OBSERVATION: The size of vector my_cdchits and my_fdchits is reduced when we are printing them out on first time hitting chi2_hit>cdc_chi2cut breakpoint.

After hitting the above breakpoint if I am deleting all the breakpoints and adding this new one only:

```
break src/libraries/TRACKING/DTrackFitterKalmanSIMD.cc:4458
```

And entering continue I am now getting the right value for id but the size of vectors is still reduced one because they were never supposed to get reduced in the size in the first place.

```
Thread 6 "hd_root" hit Breakpoint 3, DTrackFitterKalmanSIMD::KalmanForward (this=0x7fffa09310, fd_annoal_factor=18, cc_annoal_factor=18, S..., ...):
4459 bool skip_plane_my_fchits[id] = -1;
4460 print id;
4461 print my_fchits;
4462 m = stdevector of length 4, capacity 24 = (0x7fffa09310a8, 0x7fffa09310b8, 0x7fffa09310c8, 0x7fffa09310d8);
4463 print m;
4464 print forward_traj[k].h_id-1;
4465
```

Id variable value check:

At the breakpoints mentioned above we saw that id variable is uninitialized in case of JANA1 and undeclared in case of JANA2. This was concerning because id is a variable within the KalmanForward function, the same functions where previous breakpoints are set, and is initialized before the lines for which breakpoints were set above. To confirm if it was even initialized, a breakpoint was added immediately after the line `unsigned int id=forward_traj[k].h_id-1;`

break src/libraries/TRACKING/DTrackFitterKalmanSIMD.cc:4458

On hitting this breakpoint we observed that id is getting initialized inside the same function where the if statement for chi2_hit < cdc_chi2cut exist as it can be seen in the bt below when breakpoint is set at break src/libraries/TRACKING/DTrackFitterKalmanSIMD.cc:4458 which is a line next to id variable declaration.

We are able to see the value of this id at this breakpoint as shown below:

Hence some memory corruption is happening between the location where id is

initialized and if $\text{chi2_hit} < \text{chi2_cut}$ line.

Extra info:

`my_fdchits` has the actual hits and data related to them while `forward_traj` has the trajectory points for a particular trajectory that is generated so `forward_traj` has the predicted value and we will compare it with `my_fdchits` to see if the hit given for this `k` is resembling with the value predicted but if `chi2_hit` got really big then that `k` should be ignored forever.

```
4457: unsigned int id=forward_traj[k].h_id-1;
```

```
4478: Mdiff(1)=my_fdchits[id]->vstrip-vpred-doca*lorentz_factor;
```